Hindawi
Shock and Vibration
Volume 2018, Article ID 4168209, 13 pages
https://doi.org/10.1155/2018/4168209

Research Article

Theory and Numerical Analysis of Extreme Learning Machine and Its Application for Different Degrees of Defect Recognition of Hoisting Wire Rope

Zhike Zhao and Xiaoguang Zhang

School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China

Correspondence should be addressed to Xiaoguang Zhang; zhangxiaoguang2014@126.com

Received 29 December 2017; Revised 23 June 2018; Accepted 18 July 2018; Published 29 August 2018

Academic Editor: Mahmoud Bayat

An improved classification approach is proposed to solve the hot research problem of some complex multiclassification samples based on extreme learning machine (ELM). ELM was proposed based on the single-hidden layer feed-forward neural network (SLFNN). ELM is characterized by the easier parameter selection rules, the faster converge speed, the less human intervention, and so on. In order to further improve the classification precision of ELM, an improved generation method of the network structure of ELM is developed by dynamically adjusting the number of hidden nodes. The number change of the hidden nodes can serve as the computational updated step length of the ELM algorithm. In this paper, the improved algorithm can be called the variable step incremental extreme learning machine (VSI-ELM). In order to verify the effect of the hidden layer nodes on the performance of ELM, an open-source machine learning database (University of California, Irvine (UCI)) is provided by the performance test data sets. The regression and classification experiments are used to study the performance of the VSI-ELM model, respectively. The experimental results show that the VSI-ELM algorithm is valid. The classification of different degrees of broken wires is now still a problem in the nondestructive testing of hoisting wire rope. The magnetic flux leakage (MFL) method of wire rope is an efficient nondestructive method which plays an important role in safety evaluation. Identifying the proposed VSI-ELM model is effective and reliable for actually applying data, and it is used to identify the classification problem of different types of samples from MFL signals. The final experimental results show that the VSI-ELM algorithm is of faster classification speed and higher classification accuracy of different broken wires.

1. Introduction

Extreme learning machine (ELM) was proposed based on the single-hidden layer feed-forward neural network (SLFNN) [1]. Unlike the conventional network learning algorithm which must know the training samples before generating the parameters of the hidden node, ELM could generate randomly the parameters of the hidden node before understanding the training samples. ELM is characterized by the easier parameter selection rules, the faster converge speed, the less human intervention, and so on. However, due to the random generation mechanism of hidden nodes in ELM, there are still some urgent problems to be improved in ELM. The network structure is very crucial to the learning results and generalization ability of ELM. The network structure of ELM is determined by the number of hidden nodes. In recent years, the growth mechanism of hidden nodes has been extensively studied by many researchers. In order to obtain a better generalization of ELM, the performance of ELM should be optimized. Recently, there are many improved methods about ELM. The incremental extreme learning machine (I-ELM) was proposed by Huang et al. [1], which randomly adds hidden nodes one by one until it reaches the convergence requirement. But I-ELM does not recalculate the output weights of all existing nodes when a new node is added. To solve the disadvantages of I-ELM, the convex incremental extreme learning machine (CI-ELM) [2] and its improved method (ICI-ELM) [3] have been proposed. To decrease the calculation time of ELM, two different growth structures (increased structure and decreased structure) of hidden nodes were designed. The increased structure of hidden nodes includes the enhanced random search based on I-ELM (EI-ELM) [4], EM-ELM [5], and so on. The decreased structure of hidden nodes includes

P-ELM [6], OP-ELM [7], EM-ELM [8], and so on. The error-minimized extreme learning machine for single-hidden layer feed-forward neural networks was proposed for the problem of simultaneous learning. The optimum values of these parameters and the numbers of hidden neurons of ELM were obtained by using a genetic algorithm (GA), wavelet or particle swarm optimization (PSO). In addition, some new adaptive growth methods of hidden nodes were proposed, including AG-ELM [9] and D-ELM [10]. Apart from optimization constraints of ELM, ELM has a wide range of applications in data classification [11], nonlinear dynamic systems identification [12], pattern recognition [13–15], expert diagnosis [16], medical diagnosis [17], modelling permeability prediction [18], expert target recognition [19], human face recognition [20], and prediction interval estimation of electricity markets [21]. However, there are still some problems that need to be studied. All these have resulted in contradiction between the efficiency and the accuracy. This paper is based on deeply studying the improved ELM methods, and a new growth network structure of the ELM algorithm is proposed to gain better generalization. Due to the updating process being dynamically adjusted by the structure of hidden nodes by a variable step length, the method is referred to as the variable step incremental extreme learning machine (VSI-ELM). So VSI-ELM is characterized by the compact network structure, the fast running speed, and the better generalization ability.

Wire rope is widely used in coal mines, as the key component of a mine hoister, which is characterized by high intensity, lightweight, favorable flexibility, high reliability, better bending performances, and so on [22]. So wire rope is playing an increasingly important role in coal mining. Under the alternative load, the fatigue, wear, and corrosion of wire rope tend to happen and even result in the serious damage to broken rope [23]. Since some events may lead to wire rope with some risks to hosting persons, broken wire is not only the beginning of serious damage of broken rope but also difficult to be found previously, which cumulatively decreases the strength or even leads to fracture of wire rope [24, 25]. Therefore, it is important to study the nondestructive testing technique of wire rope.

The rest of this paper is organized as follows: Section 2 gives ELM algorithm theory and its improved VSI-ELM model. Section 3 gives data analysis and research of ELM, I-ELM, and VSI-ELM and the performance analysis of ELM by using the UCI data set. Section 4 introduces an automatic MFL detection system. In this section, VSI-ELM is applied to diagnosis of different broken wires. Section 5 concludes the paper indicating major achievements and future scope of this work.

2. ELM Theory

2.1. Traditional SLFNN Theory. Extreme learning machine (ELM) was proposed based on the single-hidden layer feed-forward neural network (SLFNN). ELM is characterized by the easier parameter selection rules, the faster converge speed, the less human intervention, and so on. The ELM algorithm has been widely used in many areas of image processing, machines vision, pattern recognition, decision and control, and so on. A typical SLFNN is mainly composed of the input layer, hidden layer, and output layer. ELM is a unified SLFNN with randomly generated input weights, bias, and hidden nodes. For any given N independent samples (x_i, t_i), $x_i = [x_{i1}, x_{i1}, \ldots, x_{\text{in}}]^T \in R^n$ and $t_i = [t_{i1}, t_{i1}, \ldots, t_{\text{im}}]^T \in R^m$.

Assume the input layer of SLFNN with n nodes, the hidden layer of SLFNN with L nodes, and the output layer of SLFNN with m nodes. A typical SLFNN model can be represented by

$$\sum_{i=1}^{L} \beta_i g\left(w_i \cdot x_j + b_i\right) = o_j, \quad j = 1, 2, \ldots, N, \tag{1}$$

where w_i is the connection weight between input layer nodes and hidden layer nodes and b_i is the bias. The two parameters w_i and b_i are independent not only of the training sample set but also of each other. β_i is the connecting weight between the ith hidden node and the output nodes. $g(x)$ is the activation function of hidden nodes. o_i is the output vector.

Unlike based on traditional gradient descent learning algorithms which only work for differentiable activation functions, ELM algorithm also can work for all bounded nonconstant piecewise continuous activation functions. The hidden node of ELM includes additive or RBF-type nodes, fully complex nodes, and wavelet nodes. The common activation functions of the hidden layer are shown in Table 1. For the traditional hidden layer activation function, the activation function parameters a and b are 1. And the different values will impact the performances of the ELM algorithm.

For a given standard set of training samples (x_i, t_i), if the outputs of the network are equal to the targets, we can get $\sum_{i=1}^{L} \|o_j - t_j\| = 0$:

$$\sum_{i=1}^{L} \beta_i g\left(w_i \cdot x_j + b_i\right) = t_j, \quad j = 1, 2, \ldots, N. \tag{2}$$

Equation (2) can be written compactly as

$$H\beta = T, \tag{3}$$

where

$$\begin{aligned} H &= H\left(w_1, w_2, \ldots, w_L, b_1, b_2, \ldots, b_L, x_1, x_2, \ldots, x_n\right), \\ &= \begin{bmatrix} g\left(w_1 \cdot x_1 + b_1\right) & \cdots & g\left(w_L \cdot x_1 + b_L\right) \\ \vdots & \cdots & \vdots \\ g\left(w_1 \cdot x_N + b_i\right) & \cdots & g\left(w_L \cdot x_N + b_L\right) \end{bmatrix}_{N \times L}, \\ \beta &= \begin{bmatrix} \beta_1^T \\ \beta_1^T \\ \vdots \\ \beta_1^T \end{bmatrix}_{L \times m}, \\ T &= \begin{bmatrix} t_1^T \\ t_1^T \\ \vdots \\ t_1^T \end{bmatrix}_{N \times m}, \end{aligned} \tag{4}$$

Table 1: The common hidden layer activation functions of ELM.

Types of activation functions	Formula of activation functions
Gaussian function	$G(a,b,x) = \exp(-b\|x-a^2\|)$
Hard-limit function	$G(a,b,x) = \begin{cases} 1, & \text{if } a\cdot x - b \geq 0, \\ 0, & \text{otherwise,} \end{cases}$
Sigmoid function	$G(a,b,x) = 1/(1+\exp(-(\mathrm{a}\cdot \mathrm{x} + \mathrm{b})))$
Multiquadric function	$G(a,b,x) = (\|x-a^2\| + b^2)^{1/2}$

where H is called the output matrix of the hidden layer in ELM, the ith column of H is the output vector of the ith hidden node with respect to the inputs, and β^T is the transpose of a vector β.

In practical applications, the number of training sample sets is greatly larger than the number of hidden nodes ($N \gg L$). In order to reduce calculation of ELM, the number of hidden nodes is generally selected less than the number of training samples N.

For a given minimum value $\varepsilon > 0$, ELM is of the universal approximation capability, as represented by the following equation:

$$\sum_{i=1}^{L} \left\| o_j - t_j \right\| < \varepsilon. \tag{5}$$

Under the constraint of the minimum norm least square, the weight between the hidden nodes and the output nodes can be calculated as

$$\min_{\beta} \|H\beta - T\|, \tag{6}$$

where H^+ is the Moore–Penrose generalized inverse of the output matrix of the hidden layer H.

The ELM has a three-step learning model and can be summarized below. Given a training sample set (x_i, t_i) and the activation function of the hidden node $G(a,b,x)$,

Step 1: assign randomly the input weight w_i, the bias b_i, and hidden layer nodes L

Step 2: calculate the output matrix of the hidden layer H

Step 3: calculate the output weight $\beta = H^+T$.

2.2. VSI-ELM Algorithm. Based on deeply studying the improved ELM methods, a new growth network structure of the ELM algorithm is proposed to gain better generalization. Due to the updating process being dynamically adjusted by the structure of hidden nodes by a variable step length, the method is referred to as the variable step incremental extreme learning machine (VSI-ELM). VSI-ELM is characterized by the compact network structure, the fast running speed, and the better generalization ability.

But due to lack of the selecting standard of hidden nodes, the initial value of hidden nodes L_0 is particularly important. If the number L_0 is far greater than the optimal value of hidden nodes, it can result in the increase of the training time and the decrease of the generalization ability. If the number of hidden nodes is too small, it can result in not only the lack of the fault tolerance ability but also the increase of the training error. According to the requirements between the number of hidden nodes and the resolution problem of ELM, together with the selecting experience of other neural networks, the initial value of hidden nodes L_0 in ELM is as follows:

$$L_0 = \lceil \sqrt{n+m} \rceil + q, \tag{7}$$

$$q = \begin{cases} q^{(k)} = \lceil -2^{k-1} \rceil, \\ q^{(k)} = \lceil +2^{k-1} \rceil, \end{cases} \tag{8}$$

where L_0 is the initial number of hidden nodes, n is the number of the input layer nodes, m is the number of the output layer nodes, q is the variable step length function, $q \in Z$, and k is the number of the iterations. When $q = 0$, L_0 is the initial number of hidden nodes.

Next, the update of hidden nodes L_0 is adjusted by (8). When the number of hidden nodes is close to the objective, ELM can adjust the smaller step to increase or decrease the number of hidden nodes. When the number of hidden nodes is far to reach the objective, ELM can adjust the larger step to increase or decrease the number of hidden nodes. VS-ELM reduces the computational complexity by only updating the output weights incrementally each time. The output weight β is calculated by the least-square criterion. And the computing process of VS-ELM is as follows.

Given a set of training samples (x_i, t_i), $x_i = [x_{i1}, x_{i1}, \ldots, x_{\mathrm{in}}]^T \in R^n$, $t_i = [t_{i1}, t_{i1}, \ldots, t_{\mathrm{im}}]^T \in R^m$, the expected learning accuracy $\varepsilon > 0$, and the maximum iteration r, the VS-ELM algorithm can be shown in three phases.

Phase 1: the initialization phase:

(1) Initialize the parameters of SLFNN with the mechanism of randomly generated w_i, b_i, and activation functions $g(x_j)$. There exists a positive integer k. The initial number of hidden nodes is $L_0 = \lceil \sqrt{n+m} \rceil$, and its error is E_{1k}, when $k = 0$.

(2) Calculate the output matrix of the hidden layer H_1:

$$H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_{L_0} \cdot x_1 + b_{L_0}) \\ \vdots & \cdots & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_{L_0} \cdot x_N + b_{L_0}) \end{bmatrix}_{N \times L_0}. \tag{9}$$

(3) Calculate the corresponding output error $E_{1k} = E(H_1) = \|H_1 H_1^+ T - T\|$.

Phase 2: the recursive learning phase, while $k < r$ and $E(H_k) > \varepsilon$:

(1) $k = k + 1$

There are two seeking directions of the growing mechanism of hidden nodes of ELM, including the increasing growth and the decreasing growth of the network structure as represented by formula (10) and formula (11).

The total number of hidden nodes can be added to the value L_k. It means adding the number of hidden nodes

(-2^{k-1}) and $(+2^{k-1})$ to the existing SLFNN, respectively. Calculate the corresponding output errors E_{2k} and E_{3k}.

(a) For the negative growth of the network structure, the number of hidden nodes is

$$\lceil \sqrt{n+m} \rceil + \left(-2^{k-1}\right), \quad k = 1. \tag{10}$$

(b) For the positive growth of the network structure, the number of hidden nodes is

$$\lceil \sqrt{n+m} \rceil + \left(+2^{k-1}\right), \quad k = 1. \tag{11}$$

Firstly, compare E_{11}, E_{21}, and E_{31}; the smallest of E_{11}, E_{21}, and E_{31} will be used as the number of hidden nodes. Suppose $k = 0$ and the SLFNN with the number of hidden nodes L_0, if the corresponding output error $E_{11} \le \varepsilon$, $E_{11} \le E_{21}$, and $E_{11} \le E_{31}$, the growing procedure gets finished. If E_{21} is the minimum value of E_{11}, E_{21}, and E_{31}, VS-ELM chooses the negative growth of hidden nodes. If E_{31} is the minimum value of E_{11}, E_{21}, and E_{31}, VS-ELM chooses the positive growth of hidden nodes. For example, if E_{31} is the minimum value of E_{11}, E_{21}, and E_{31}, the next update of the number of hidden nodes is $\lceil \sqrt{n+m} \rceil + (2^{2-1})$ and the corresponding output error is E_{32}.

Secondly, compare E_{32} and E_{31}; if $E_{32} < E_{31}$ and $\|E\| > \varepsilon$, the next update of the number of hidden nodes is $\lceil \sqrt{n+m} \rceil + (2^{3-1})$ and the corresponding output error is E_{33}. In addition, if $E_{36} > E_{35}$ and $\|E\| > \varepsilon$, the number of hidden nodes will stop positive growing, $\lceil \sqrt{n+m} \rceil + (2^{5-1})$ it is taken as the new initial of the number of hidden nodes. Next, the number of hidden nodes $\lceil \sqrt{n+m} \rceil + (2^{5-1}) + (2^{1-1})$ is updated. Using this method, we can find the best number of hidden nodes until $\|E\| < \varepsilon$.

Phase 3: if the corresponding output error $\|E\| < \varepsilon$ or $k > r$, the growing procedure gets finished.

End while.

3. Data Analysis and Research

3.1. The Performance Analysis of ELM. Through the above analysis of ELM theory and its improved methods, it is not difficult to find that the performance of ELM has a direct relationship with its algorithm structure. The input weight matrix of ELM is generated by a random pattern after the matrix of input neurons and the number of hidden layer neurons. The number of input neurons is determined by the size of the sample matrix (training or testing), while the number of hidden layer neurons is artificially set. Therefore, the size of the sample matrix also affects the performance of ELM. However, it is very important to adjust the hidden layer neuron nodes to improve the performance of the ELM without changing the sample size. Based on this, the accuracy of the ELM in regression or classification will be improved if the input weight matrix generated by a random pattern is the best match with the training sample. Therefore, it is of great significance to study the number of hidden layer neurons in ELM and optimize the parameters of the input weight matrix.

To investigate the effect of hidden layer neuron nodes on the performance of ELM, a performance test of the ELM algorithm was conducted using some sample sets provided by the University of California, Irvine (UCI). The regression and classification sets selected from the UCI data set are shown in Tables 2 and 3, respectively. Among them, the determination coefficient R and the root-mean-square error (RMSE) are selected as evaluation indexes. The smaller the root-mean-square error, the better the performance of the algorithm model. The determination coefficient R is within the range [0, 1], and the closer the coefficient to 1, the better the performance of the algorithm model. Conversely, the closer the coefficient to 0, the worse the performance of the algorithm model. The two indicators are calculated as follows:

$$\text{RMSE} = \sqrt{\text{MSE}\left(\widehat{y}_i - y_i\right)},$$

$$R = \frac{\left(n\sum_{i=1}^{n} \widehat{y}_i y_i - \sum_{i=1}^{n} \widehat{y}_i \sum_{i=1}^{n} y_i\right)^2}{\left(n\sum_{i=1}^{n} \left(\widehat{y}_i\right)^2 - \left(\sum_{i=1}^{n} \widehat{y}_i\right)^2\right)\left(n\sum_{i=1}^{n} \left(y_i\right)^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right)}, \tag{12}$$

where $\widehat{y}_i$ is the predicted value of the ith sample, y_i is the true value of the ith sample, and n is the number of samples.

In order to verify the impact of hidden layer nodes of ELM on the field of regression prediction, the selected four classes of regression data sets (spectra set, concrete set, fertility set, and sinc set) were tested from UCI provided in Table 2. Based on the performance indicators of the root-mean-square error (RMSE) and determination coefficient R, the performance index changes with the hidden layer node, as shown in Figure 1.

After analysis, it can be seen that the sample features can effectively describe the characteristics of the samples, which is the prerequisite for the regression prediction of the samples, which is the main reason for the unsatisfactory performance, as shown in Figure 1(b). In both the training samples and the testing samples, the closer the RMSE to 0, the closer the determination coefficient to 1. Similarly, the closer the determination coefficient to 1, the closer the RMSE to 0.

For the simple characterization of samples such as the sinc set, with the increase of hidden layer nodes of ELM, the closer the RMSE to 0 and the closer the decision coefficient R to 1, the higher the regression prediction accuracy of the testing samples than the training samples. At the same time, with the increase of hidden layer nodes of ELM, there is no fluctuation of their RMSE and decision coefficient anomaly. However, with the increase of hidden layer nodes of ELM, the abnormal fluctuation of RMSE and determination coefficient R appears on the regression prediction of the spectra set, concrete set, and fertility set, and these abnormal fluctuations occur in the hidden layer node numbered 30, 50, and 50, respectively. Therefore, it is not appropriate to improve the fitting accuracy of ELM regression only by adding hidden layer nodes without considering the ELM overlearning problem. In addition, for the spectra set, concrete set, and sinc set, there is no change when the hidden layer nodes reach 90, 100, and 10, respectively; since then, continuing to increase the hidden layer nodes will only increase the computing time of ELM, as shown in Figure 1.

Table 2: Data set description of the selected UCI regression problems.

Sample set	Number of characteristics	Total number of samples	Training samples	Test samples
Spectra	401	60	30	30
Fertility	9	100	50	50
Concrete	7	103	51	52
Sinc	2	10000	5000	5000

Table 3: Data set description of the selected UCI classification problems.

Data set	Sample category	Number of features	Total number of samples	Training samples	Testing samples
Abalone	3	8	4177	2000	2177
Statlog (Heart)	2	13	270	135	135
Diabetes	2	8	768	576	192
Parkinsons	2	22	192	95	97
Wdbc	2	30	569	284	285
Iris	3	4	150	75	75
Wine	3	13	178	90	88
Breast Tissue	6	9	106	44	62
Glass	6	9	214	110	104
Seeds	3	7	210	105	105
Waveform (version 2)	3	21	5000	2500	2500

In order to verify the effect of hidden layer nodes of ELM on the classification method, 11 classification data sets (abalone, statlog (heart), diabetes, parkinsons, wdbc, iris, wine, breast tissue, glass, seeds, and waveform (version 2)) are used to test the classification predictions of ELM. The classification accuracy of the classification results varies with the hidden layer nodes, as shown in Figure 2.

Through the analysis and comparison of the above data set, the conclusions are as follows:

(1) With the increase of hidden layer neuron nodes, the classification accuracy of the training samples has a sharp increase stage and then relatively slowly approaches the target value. If the neurons in the hidden layer continue to increase, the classification accuracy of the training samples can reach 100%.

(2) With the increase of hidden layer neurons, the classification accuracy of the testing samples also has a sharp increase phase, but it does not approach the target value relatively slowly afterwards as the training samples. Instead, the following possibilities exist: (1) When the classification accuracy increases sharply and reaches the maximum value, the classification accuracy decreases gradually, as shown in Figures 2(a), 2(e), and 2(k). (2) When the classification accuracy sharply increased to reach the maximum, the classification accuracy first decreased and then stabilized, as shown in Figures 2(b), 2(c), 2(d), and 2(i). (3) When the classification accuracy increased sharply and reached the maximum value, the classification accuracy first decreases and then rises and approaches a new stationary value, as shown in Figures 2(f), 2(h), and 2(j).

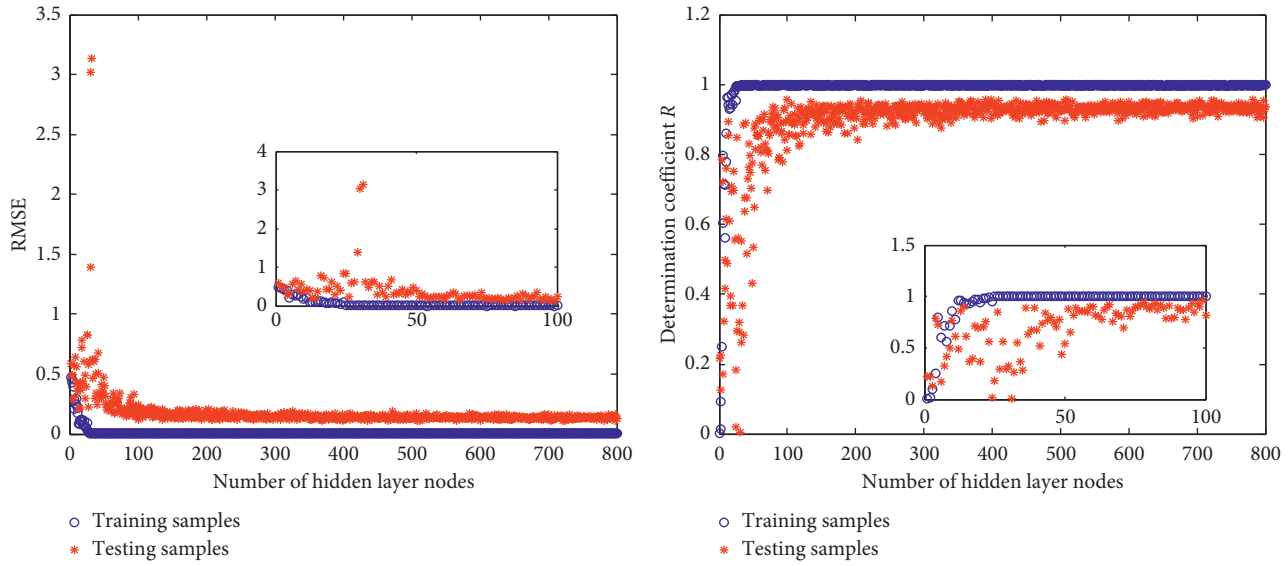
(3) There is a big difference in classification accuracy of ELM when the nodes of hidden layer neurons are not much different, and their classification accuracy even exceeds 20%. As shown in Figure 2(g), the classification accuracy of wine presents a banded distribution. The main reason for this situation is that ELM input weights are caused by the stochastic mode. However, the root cause is that every time the hidden layer node of ELM is updated, the input weight matrix is updated again, which leads ELM to lose self-optimizing ability and greatly increase the searching time of the ELM optimal structure. This is also the reason why the I-ELM algorithm uses hidden layer neuron nodes layer by layer.

3.2. The Different Growth Structure of Hidden Layer Nodes of ELM. There are two different ways of growth of hidden layer nodes in this article. I-ELM algorithm 1 needs to recalculate all the input weights according to the updated number of hidden layer neurons, and I-ELM algorithm 2 only needs to calculate the connection weights of the new added hidden layer neurons and original input and output neurons. I-ELM algorithm 2 makes full use of the previously calculated input weight matrix to reduce its calculation time. I-ELM algorithm 2 improves the algorithm structure only by adding hidden layer nodes, as shown in Figure 3.

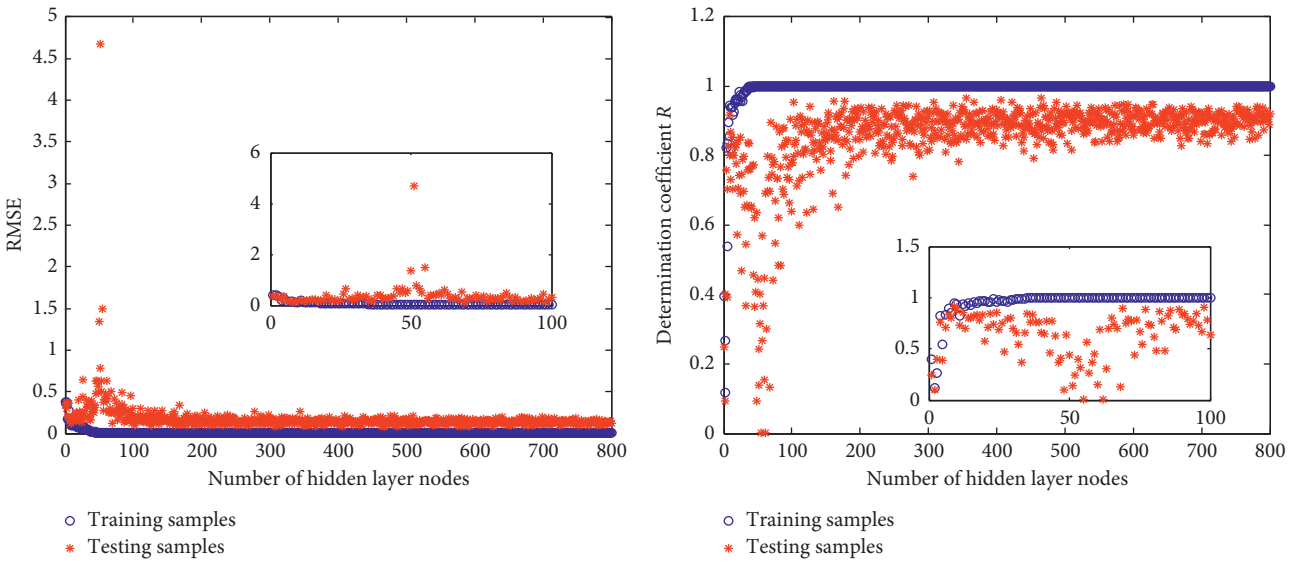
In order to verify the difference between the two methods, different updates of hidden layer neuron numbers and recalculation of input weights are done. In this paper, the two methods (I-ELM algorithm 1 and I-ELM algorithm 2) were tested by using the iris set, respectively. The results are shown in Figure 4. As can be seen in Figure 4, I-ELM algorithm 2 significantly converges faster than I-ELM algorithm 1, making the ELM structure more compact and avoiding unnecessary training time consumption.

3.3. The Performance Analysis of the VSI-ELM Algorithm. In order to compare the performance of the VSI-ELM algorithm and I-ELM algorithm, the UCI classification data set (statlog (heart), diabetes, parkinsons, and iris) provided in Table 3 was used to test the two algorithms. The update rate curves of the sample classification accuracy are shown in Figure 5. The detailed comparison results of the training time-consuming and hidden layer neuron node for the VSI-ELM algorithm and I-ELM algorithm are shown in Table 4. For iris, the preimprovement algorithm (I-ELM) takes twice the time-consuming training of the modified algorithm (VSI-ELM). So the VSI-ELM algorithm is of faster training speed of the multiclassification samples.

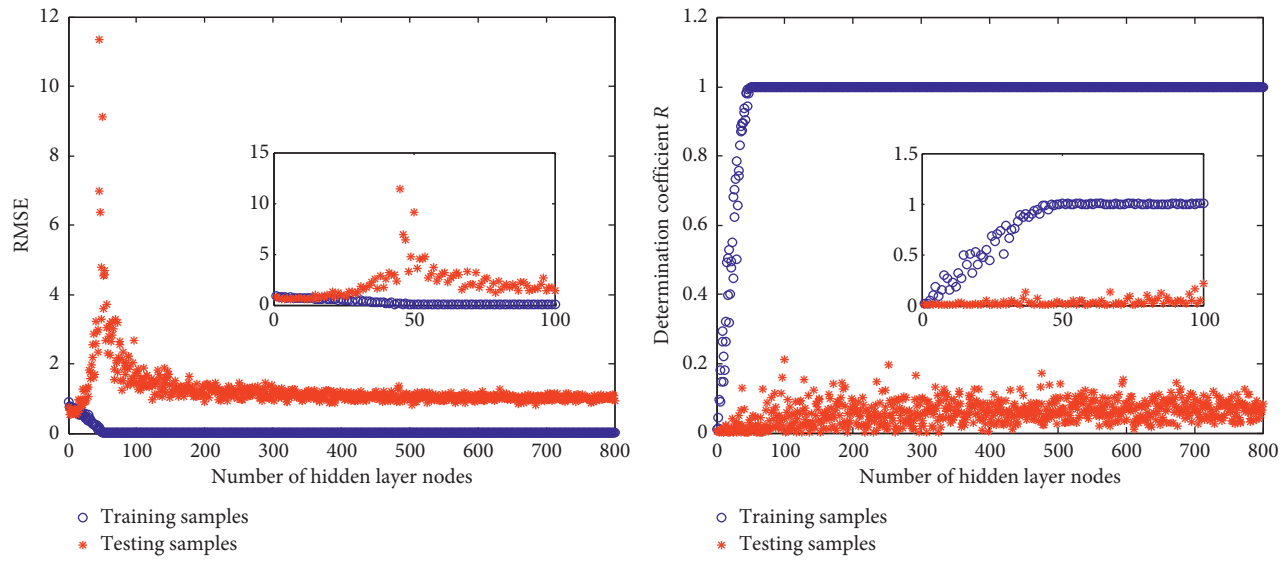
Numerical analysis shows that the VSI-ELM algorithm can guarantee the optimal number of neurons in the hidden layer and faster convergence than the I-ELM algorithm and



(a)

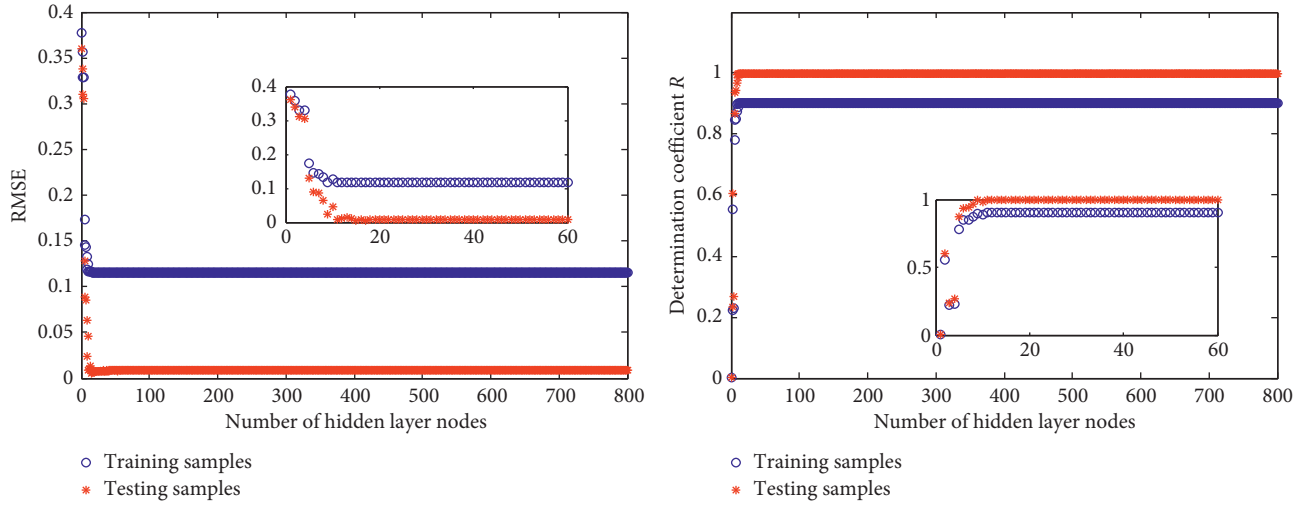


(b)



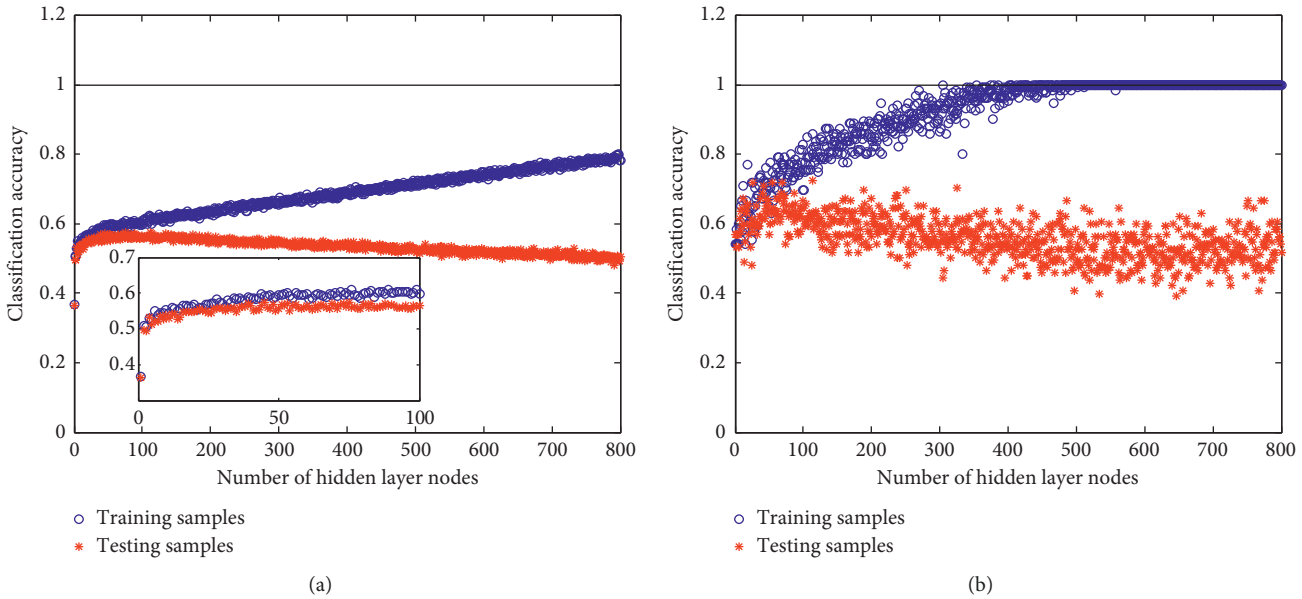
(c)

Figure 1: Continued.



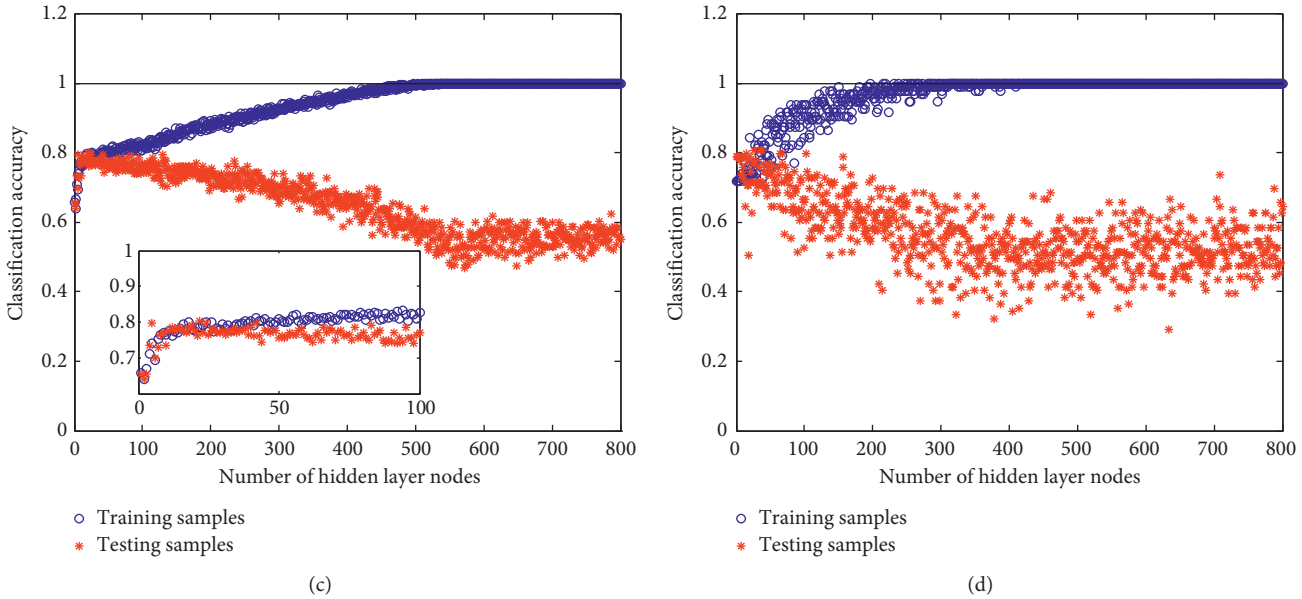
(d)

Figure 1: ELM training and testing results of the selected UCI regression data sets: (a) spectra; (b) concrete; (c) fertility; (d) sinc



(a)

(b)



(c)

(d)

Figure 2: Continued.

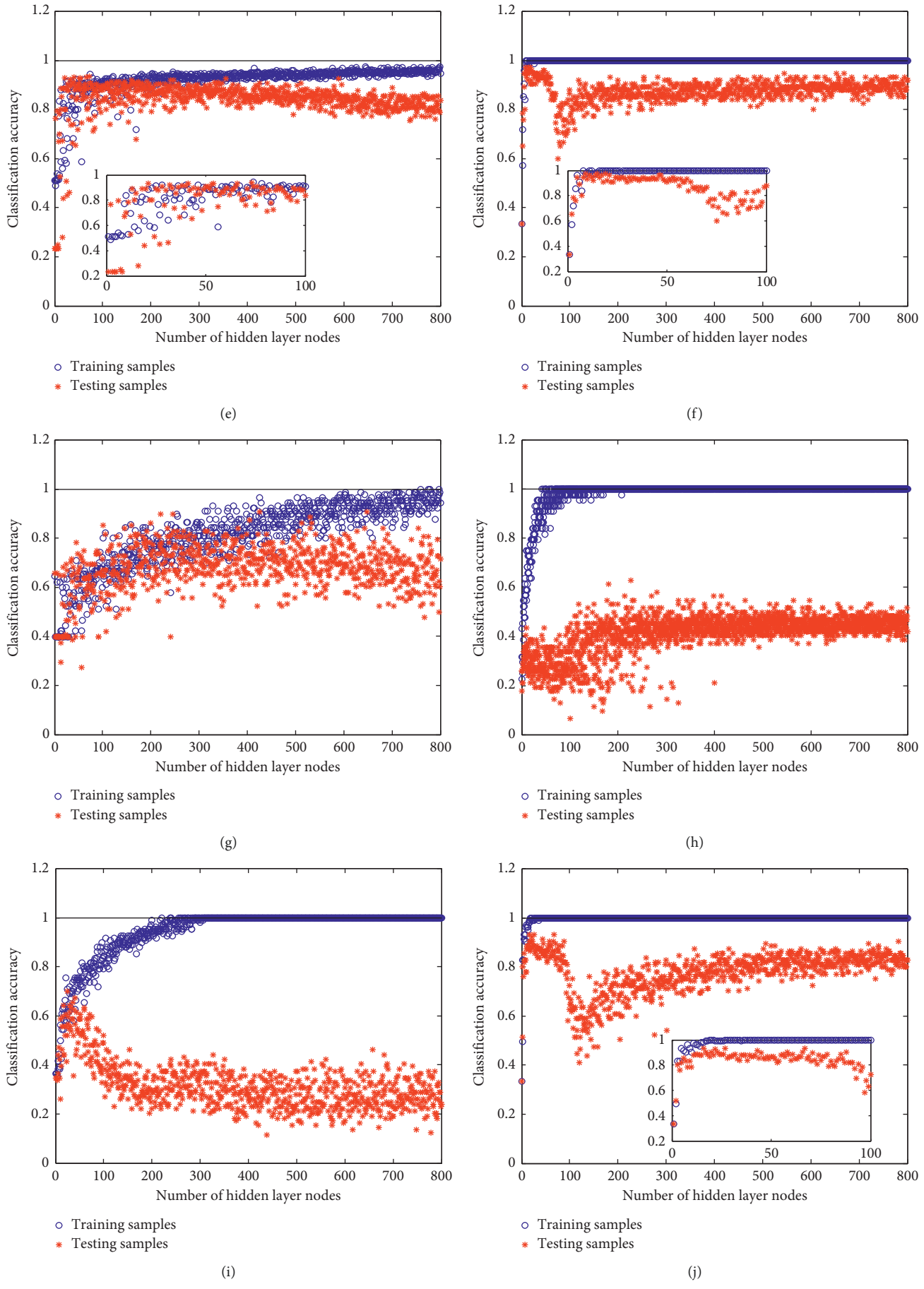
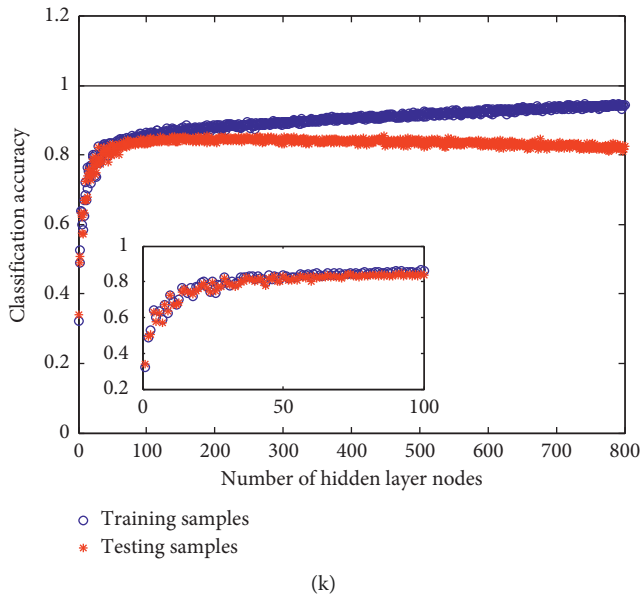


Figure 2: Continued.



(k)

Figure 2: ELM training and testing results of the selected UCI classification data sets: (a) abalone; (b) statlog (heart); (c) diabetes; (d) parkinsons; (e) wdbc; (f) iris; (g) wine; (h) breast tissue; (i) glass; (j) seeds; (k) waveform (version 2).

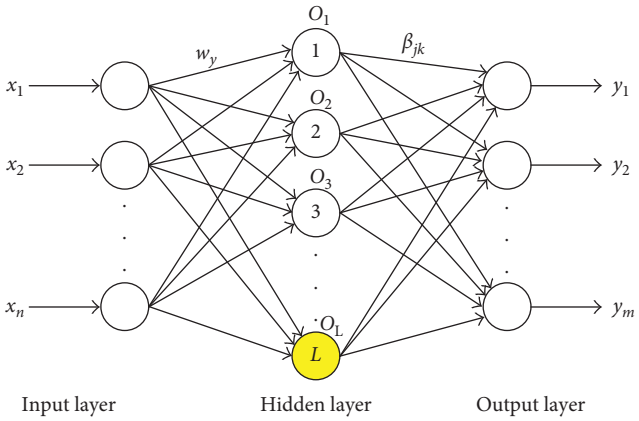


Figure 3: Network structure of I-ELM by adding hidden layer nodes.

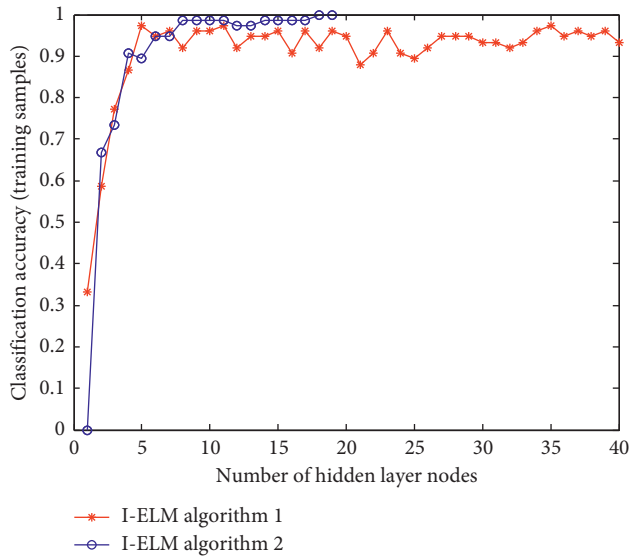


Figure 4: Comparison between I-ELM 1 and I-ELM 2 about classification accuracy updating curves of training samples for the iris data set.

make the ELM network structure more compact; generalization ability is also stronger.

4. Practical Application Based on the VSI-ELM Algorithm

4.1. Research Background of Broken Wire Detection. Mine lifting wire rope is one of the most critical components of the coal mine transportation system. It is responsible for the transportation of personnel, coal, and equipment, and its working condition is directly related to the safe and orderly production of coal mine. As the mine lifting wire rope is affected by the long-term friction, humidity, corrosion, and other harsh production conditions and bears the repeated tensile load and bending load, broken wire, abrasion, corrosion, and other structural damage will inevitably appear, which results in the strength reduction of wire rope and brings harm to the safe operation of the wire rope. With the increasing depth of mining, the requirements for wire rope that can withstand high-speed, long-term, and heavy-load conditions have become exigent. However, the complex structure of wire rope and uncertainty of damage type and location have brought a lot of technical problems to the wire rope nondestructive detection, especially the use of the magnetic flux leakage method. In that case, the relation between magnetic field change and structure, movement mode, and stress change of wire rope is becoming more complex, and it also brings troubles to the magnetic flux leakage signal detection. Magnetic flux leakage (MFL) is an

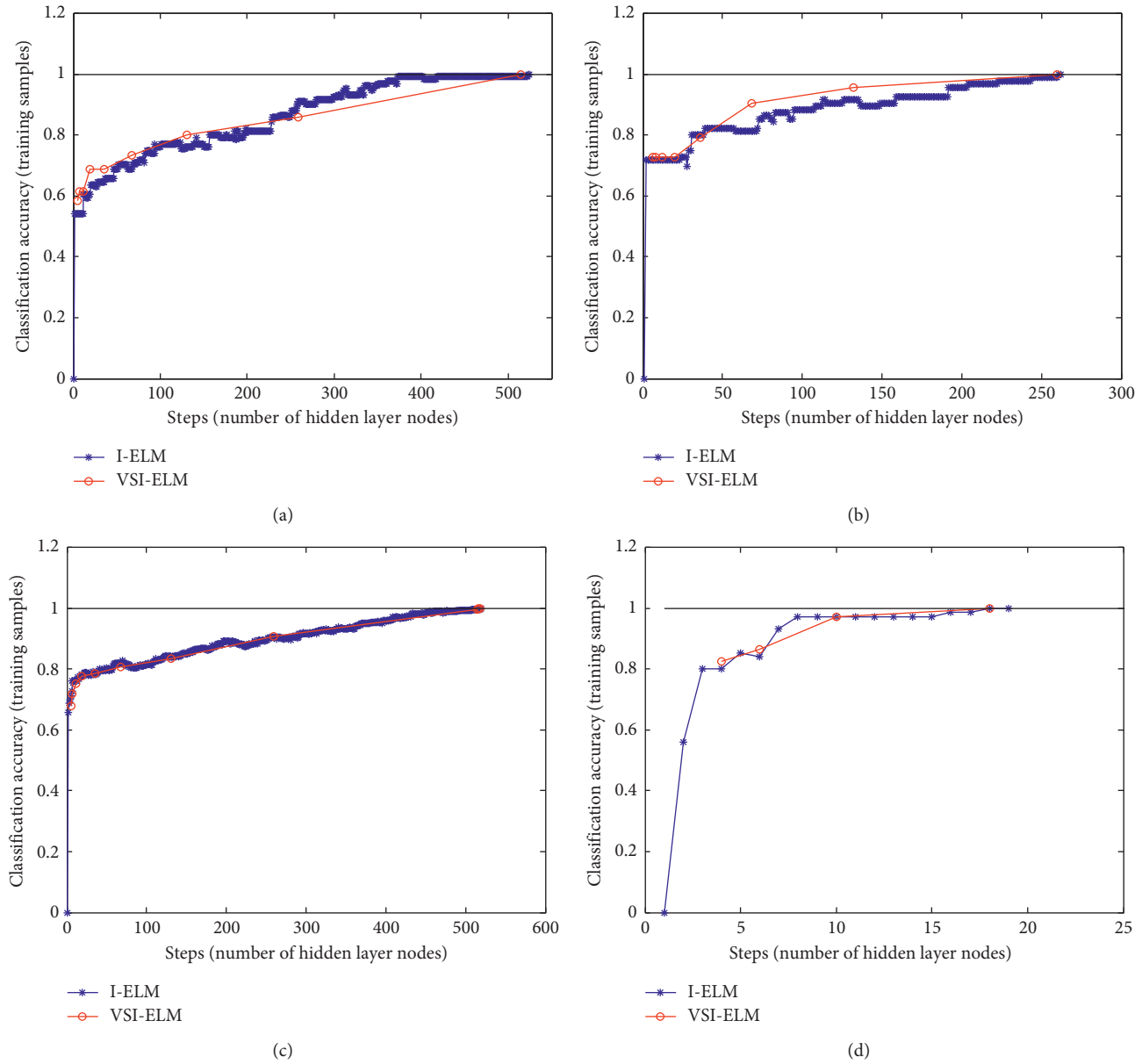


Figure 5: Classification accuracy updating curves of training samples of VSI-ELM and I-ELM: (a) statlog (heart); (b) parkinsons; (c) diabetes; (d) iris.

Table 4: Performance comparison of VSI-ELM and I-ELM in the UCI set classification problem.

Sample set	Algorithm	Time-consuming	Time-consuming ratio of I-ELM to VSI-ELM	Hidden layer neuron node
Statlog (Heart)	I-ELM	17.4606	50.64	537
	VSI-ELM	0.3448		515
Diabetes	I-ELM	69.796	18.08	520
	VSI-ELM	3.8594		518
Parkinsons	I-ELM	4.5163	28.89	262
	VSI-ELM	0.1563		260
Iris	I-ELM	0.3125	2.86	19
	VSI-ELM	0.1094		18

efficient nondestructive testing technique for the defected wire rope and plays an important role in the dynamic monitoring of wire rope [26–28]. Because of the intricate structure of wire rope, there is a complicated relation between the diverse damages and MFL signals. Permanent magnet is characterized by small volume, low cost, lightweight, high magnetic field, not requiring power, and easy to dispose and install. The MFL signals are gathered by some arrays of Hall effect sensors disposed at the circumference clinging to the outer surface of wire rope [29]. So the MLF signals are influenced by the lift-off distance, velocity effect, shaking, and various properties of the defects. The MFL signal of each channel is different from that of other channels in a multistage ring MFL detection device [30].

Certainly, all influencing factors are very important to study the design of the subsequent signal processing. In recent years, a large amount of the defect detection methods have gained great achievement in respect of monitoring of wire rope; meanwhile, there are also some issues that need to be resolved. Therefore, it is an important and urgent lesson in the research field that explains how to apply the simplest and fastest method for fault feature extraction of the broken wire of wire rope. The effect of variable tensile stress on the MFL signal response of defective wire ropes is analyzed and dealt with as needed [31]. The filtering system consisting of the Hilbert–Huang transform and compressed sensing is used to obtain the defect RMF image characteristics of wire rope, and the characteristics are extracted as the input of a radial basis function neural network to identify the defects of wire rope [32].

To remove the effects of channel-to-channel mismatch on the disposition, an adaptive method for MFL channel equalization is based on PCA and ELM [33]. For the classification of the MFL signal for different broken wires, the neural networks are very popular methods. The BP neural network was employed for the quantitative identification of broken wires [34]. The improved radial basis function neural network was applied for the quantitative identification of defected wire rope [35]. The wavelet neural network was used for the prediction and diagnosis of hoisting wire rope [36]. Therefore, the choice of the wire rope breakage identification method is to be solved.

4.2. Experimental Study of the Classification of Different Degrees of Broken Wires. In this paper, a new MFL detection device is used to obtain the MFL signal. The MFL detection device is shown in Figure 6. Twenty-four Hall sensors are distributed in space of the MFL detection device. Each of the three Hall sensors is a group. The acquisition board includes three diverse direction Hall sensor arrays. Each direction is composed of 8 channels of Hall sensors, which are uniformly arranged at the annular circuit board. There are 24 Hall sensors to measure the magnetic flux leakage of defected wire rope by using the necessary amplification and filter to record the MFL signal. The multichannel MFL signals are transmitted to the acquisition system. The time-domain and time-frequency domain characteristics of MFL signals of the diverse wire rope are analyzed. In order to train the VSI-ELM algorithm, some normal samples are needed in this experiment. The mixed-features vector can be used as the effective characteristic input of the quantitative identification when wire rope appears to be broken wires. To avoid the training sample set getting too large, the length of the sample set is set to a certain length (2048 data points). Table 5 shows the characteristic samples of MFL signals of broken wires, where n is the sequence number, P is the peak of the MFL wave, W is the width of the MFL wave, S is the area under the MFL wave, R is the diameter of wire rope, d is the lift-off distance, and k is the damage type. In this section, VSI-ELM was utilized to extract the characteristics of MFL signals of different broken wires.

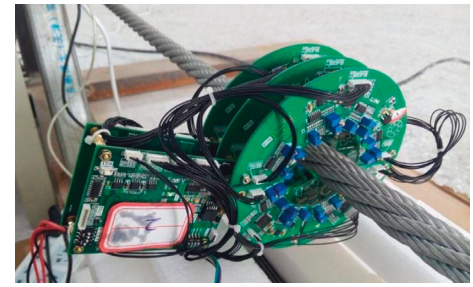


Figure 6: MFL detection device.

Table 5: The characteristic samples of MFL signals of broken wires.

n	P	W	S	k
1	0.274	0.572	0.321	1
2	0.267	0.592	0.311	1
3	0.258	0.502	0.325	1
4	0.345	0.632	0.442	2
5	0.324	0.616	0.431	2
6	0.317	0.634	0.421	2
7	0.495	0.772	0.599	3
8	0.494	0.796	0.560	3
9	0.501	0.762	0.551	3
10	0.591	0.857	0.709	4
11	0.599	0.856	0.706	4
12	0.603	0.861	0.713	4
13	0.751	0.887	0.819	5
14	0.769	0.886	0.836	5
15	0.743	0.891	0.813	5

For MFL signals, the characteristic samples include training samples and testing samples. The number of training samples is 80. The number of testing samples is 80. The classification accuracy of defected broken wires is up to the best value of 97.5% by using VSI-ELM. Compared to the I-ELM algorithm, VSI-ELM can not only gain the optimal number of hidden nodes but also the fast convergence rate. The experimental results show that the VSI-ELM algorithm is of faster classification speed and higher classification accuracy of different broken wires.

5. Conclusions

In this paper, the theory of ELM based on the single-hidden layer feed-forward neural network is reanalyzed. The classification model of ELM is theoretically deduced, and the existing improving methods of ELM are compared. The number of hidden layer nerves of ELM is emphatically analyzed. So the key is the hidden layer neuron growth strategy. This article focuses on the analysis of the influence of the number of hidden layer nodes on the performance of ELM.

The numerical simulation analysis of the UCI data set is used to test the effect of the number of hidden layer neuron nodes of ELM. Through comparative analysis, it is found that I-ELM algorithm 2 has better performance. It is verified that the I-ELM algorithm 2 is more conducive to finish of sample training by using stacking the hidden layer nerves. Based on

the above analysis, a novel adjustment strategy of hidden layer neuron nodes of the ELM (VSI-ELM) algorithm is proposed in this paper. The feasibility of VSI-ELM is verified by the UCI classification data set (statlog (heart), diabetes, parkinsons, and iris). The time-consuming ratio of I-ELM to VSI-ELM of statlog (heart), diabetes, parkinsons, and iris is 50.64, 18.08, 28.89, and 2.86, respectively. The experimental results show that the VSI-ELM algorithm can find the best number of hidden layer neuron nodes faster than the I-ELM algorithm. Finally, the VSI-ELM algorithm is applied to identify the characteristics of the MFL signal of different broken wires. The classification accuracy of defected broken wires is up to 97.5% by using VSI-ELM.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was financially supported by the Priority Academic Program Development (PAPD) of Jiangsu Higher Education Institutions, the University-Industry Cooperation Research Project in Jiangsu Province under Grant no. BY2016026-02 and the State Key Laboratory of Integrated Services Networks under Grant no. ISN10-10.

References

[1] G. B. Huang, L. Chen, and C. K. Siew, "Universal approximation using incremental constructive feed forward networks with random hidden nodes," *IEEE Transactions on Neural Networks*, vol. 17, no. 4, pp. 879–892, 2006.

[2] G. B. Huang and L. Chen, "Convex incremental extreme learning machine," *Neurocomputing*, vol. 70, no. 16–18, pp. 3056–3062, 2007.

[3] W. Wang and R. Zhang, "Improved convex incremental extreme learning machine based on enhanced random search," *Lecture Notes in Electrical Engineering*, vol. 238, pp. 2033–2040, 2012.

[4] G. B. Huang and L. Chen, "Enhanced random search based on incremental extreme learning machine," *Neurocomputing*, vol. 71, no. 16–18, pp. 3460–3468, 2008.

[5] G. R. Feng, G. B. Huang, and Q. P. Lin, "Error minimized extreme learning machine with growth of hidden nodes and incremental learning," *IEEE Transactions on Neural Networks*, vol. 20, no. 8, pp. 1352–1357, 2009.

[6] H. J. Rong, Y. S. Ong, and A. H. Tan, "A fast pruned-extreme learning machine for classification problem," *Neurocomputing*, vol. 72, no. 1–3, pp. 359–366, 2008.

[7] Y. Miche, A. Sorjamaa, and A. Lendasse, "OP-ELM: theory, experiments and a toolbox," *Lecture Notes in Computer Science*, vol. 5163, pp. 145–154, 2008.

[8] S. Balasundaram and Kapil, "Application of error minimized extreme learning machine for simultaneous learning of a function and its derivatives," *Neurocomputing*, vol. 74, no. 16, pp. 2511–2519, 2011.

[9] R. Zhang, Y. Lan, and G. B. Huang, "Universal approximation of extreme learning machine with adaptive growth of hidden nodes," *IEEE Transactions on Neural Networks*, vol. 23, no. 2, pp. 356–371, 2012.

[10] R. Zhang, Y. Lan, and G. B. Huang, "Dynamic extreme learning machine and its approximation capability," *IEEE Transactions on System*, vol. 99, pp. 1–12, 2012.

[11] E. G. Horta, C. L. DeCastro, and A. P. Braga, "Stream-based extreme learning machine approach for big data problems," *Mathematical Problems in Engineering*, vol. 2015, Article ID 126452, 17 pages, 2015.

[12] D. M. Salih, S. B. Mohd Noor, M. H. Merhaban, and R. Kamil, "Wavelet network: online sequential extreme learning machine for nonlinear dynamic systems identification," *Advances in Artificial Intelligence*, vol. 2015, Article ID 184318, 10 pages, 2015.

[13] D. M. Salih, S. B. Mohd Noor, M. H. Merhaban, and R. Kamil, "Vowel imagery decoding toward silent speech BCI using extreme learning machine with electroencephalogram," *BioMed Research International*, vol. 2016, Article ID 2618265, 11 pages, 2016.

[14] J. K. Deters, R. Zalakeviciute, M. Gonzalez, and Y. Rybarczyk, "Modeling $PM_{2.5}$ urban pollution using machine learning and selected meteorological parameters," *Journal of Electrical and Computer Engineering*, vol. 2017, Article ID 5106045, 14 pages, 2017.

[15] A. Khare, M. Jeon, I. K. Sethi, and B. Xu, "Machine learning theory and applications for healthcare," *Journal of Healthcare Engineering*, vol. 2017, Article ID 5263570, 2 pages, 2017.

[16] D. Avci1 and A. Dogantekin, "An expert diagnosis system for Parkinson disease based on genetic algorithm-wavelet kernel-extreme learning machine," *Parkinson's Disease*, vol. 2016, Article ID 5264743, 9 pages, 2016.

[17] K. Kaya and M. Uyar, "A hybrid decision support system based on rough set and extreme learning machine for diagnosis of hepatitis disease," *Applied Soft Computing Journal*, vol. 13, no. 8, pp. 3429–3438, 2013.

[18] S. O. Olatunji, A. Selamat, and A. Abdulraheem, "A hybrid model through the fusion of type-2 fuzzy logic systems and extreme learning machines for modelling permeability prediction," *Information Fusion*, vol. 16, no. 1, pp. 29–45, 2014.

[19] E. Avci, "A new method for expert target recognition system: genetic wavelet extreme learning machine (GAWELM)," *Expert Systems with Applications*, vol. 40, no. 10, pp. 3984–3993, 2013.

[20] A. A. Mohammed, R. Minhas, Q. M. Jonathan Wu, and M. A. Sid-Ahmed, "Human face recognition based on multidimensional PCA and extreme learning machine," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2588–2597, 2011.

[21] N. A. Shrivastava and B. K. Panigrahi, "Point and prediction interval estimation for electricity markets with machine learning techniques and wavelet transforms," *Neurocomputing*, vol. 118, no. 11, pp. 301–310, 2013.

[22] D. K. Zhang, K. Chen, and X. F. Jia, "Bending fatigue behaviour of bearing ropes working around pulleys of different materials," *Engineering Failure Analysis*, vol. 33, pp. 37–47, 2013.

[23] D. Basak, S. Pal, and D. C. Patranabis, "Inspection of 6×19 seale preformed haulage rope by nondestructive technique," *Russian Journal of Nondestructive Testing*, vol. 45, no. 2, pp. 143–147, 2009.

[24] W. X. Zhan, J. W. Tan, and Y. Wen, "Algorithm considering signal correlation for denoising of the magnetic flux leakage of wire rope," *Advanced Materials Research*, vol. 291–294, pp. 2652–2657, 2011.

[25] Y. H. Sun, S. W. Liu, and L. S. He, "A new detection sensor for wire rope based on open magnetization method," *Material Evaluation*, vol. 75, no. 4, pp. 501–509, 2017.
[26] G. Drummond, J. F. Watson, and P. P. Acarnley, "Acoustic emission from wire ropes during proof load and fatigue testing," *NDT & E International*, vol. 40, no. 1, pp. 94–101, 2007.
[27] X. L. Yan, D. L. Zhang, and S. M. Pan, "Online nondestructive testing for fine steel wire rope in electromagnetic interference," *NDT & E International*, vol. 92, pp. 75–81, 2017.
[28] B. Yu, J.-W. Kim, and S. Park, "Magnetic-flux-leakage-based pressure deformation detection of elevator wire rope," *Journal of the Korean Society for Nondestructive Testing*, vol. 37, no. 4, pp. 269–275, 2017.
[29] Q. S. Cao, D. Liu, and Y. H. He, "Nondestructive and quantitative evaluation of wire rope based on radial basis function neural network using eddy current inspection," *NDT & E International*, vol. 46, pp. 7–13, 2012.
[30] C. Jomdecha and A. Prateepasen, "Design of modified electromagnetic main-flux for steel wire rope inspection," *NDT & E International*, vol. 42, no. 1, pp. 77–83, 2009.
[31] G. H. Gao, M. J. Lian, and Y. G. Xu, "The effect of variable tensile stress on the MFL signal response of defective wire ropes," *Insight*, vol. 58, no. 3, pp. 135–141, 2016.
[32] J. W. Zhang, X. J. Tan, and P. B. Zheng, "Non-destructive detection of wire rope discontinuities from residual magnetic field images using the Hilbert-Huang transform and compressed sensing," *Sensor*, vol. 17, no. 3, pp. 1–19, 2017.
[33] Z. N. Wu, H. G. Zhang, and J. H. Liu, "A PCA and ELM based on adaptive method for channel equalization in MFL inspection," *Methematical Problems in Engineering*, vol. 2014, Article ID 124968, 8 pages, 2014.
[34] Z. Y. Tian, Y. Zhang, and J. W. Tan, "Quantitative test of broken wire for steel rope based on the back propagation artificial neural networks," *International Journal of Coal Science and Technology*, vol. 31, no. 2, pp. 245–249, 2006.
[35] H. Y. Wang, Z. Xu, and G. Hua, "Key technique of a detection sensor for coal mine wire ropes," *Mining Science and Technology*, vol. 19, pp. 170–175, 2009.
[36] X. J. Zhu, J. Y. Guo, and C. Y. Wei, "Prediction and diagnosis of mine hoist fault based on wavelet neural network," in *Proceeding of Chinese Control and Decision Conference*, pp. 598–601, Yantai, China, July 2008.