International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.9, (2016), pp.279-290
http://dx.doi.org/10.14257/ijsip.2016.9.9.26

Class Incremental ELM and Application for Image Recognition

Wei Tao[1,2] and Ji Xin-Sheng[1]

[1]China National Digital Switching System Engineering＆Technological R＆D Center, Zhengzhou City, Henan Province 450002
[2]Computer College, Henan institute of engineering, Zhengzhou City, Henan Province, China, 451191
hnsqwt@163.com

Abstract

In image recognition field, the fact is that the trained image classifier can not recognize the images, whose class type is not the same as the training data. To resolve this problem, a new image classifier is proposed, which is based on the class incremental extreme learning machine. The new classifier can recognize the normal images well, label them with new labels, and update itself with the new labeled data. Tested on the real-world daily activity data set, the results show that our algorithm performs well.

Keywords: *machine learning; extreme learning machine; incremental learning; image recognition*

1. Problem Description

Some image recognition models may require incremental learning for a specific image class. In other contexts, it may be necessary to change the image classes so as to enhance the recognition ability for new image classes. For example, in image recognition of battlefield targets, there are only a few target classes to be recognized at the beginning. But as the military technologies are improving on both parties of the battle, new target classes will appear, including new weapons.

As a result, the overall distribution of image data to be recognized by the users will change, leading to a disparity from the distribution of the training data. In that case, the model will not be able to recognize the new images, and the new images will be mistakenly recognized as existing images. There will be an increase in false alarm rate.

We have drawn on a type of class-incremental extreme learning machine [1], which transfers the ELM model trained in the source domain to the target domain. The recognition model is constantly adjusted according to the new calibration images in the target domain, while the framework of ELM model trained on data in the source domain remains constant. Due to class-incremental learning, the output nodes will increase along with the extension of output weight matrix, but the input nodes, hidden nodes and their structure will not be affected. Thus the knowledge in the source domain can be preserved to the maximum extent. By virtue of class-incremental learning, the model cannot only learn new image classes, but also preserves the recognition capacity for existing image classes.

The contents of this article are organized as follows. Section 2 is a description of relevant work on incremental learning; Section 3 presents adaptive adjustment of the model structure and on-line incremental learning by using class-incremental learning; Section 4 shows the experimental results; Section 5 is the concluding section.

2. Relevant Work

The researches on incremental learning fall into the following categories:

ISSN: 2005-4254 IJSIP

(1) Incremental Learning based on Radial Basis Function Networks (RBF Networks) and Adaptive Resonance Theory.

Incremental learning method for feedforward neural network has been studied more extensively in recent years. Radial basis function network (RBF network) is an emerging network which can dynamically insert new RBF center based on the cumulative error of classification and determine the radius of the new center using Neural Gas algorithm. GAP-RBF algorithm is also a type of feedforward neural network with RBF nodes, which can automatically delete the unimportant nodes from the network. Although this method also aims to improve the learning speed by simplifying the incremental learning algorithm, the learning speed is still low for many applications because of the need for knowledge of input data distribution or sampling range. ARTMAP is an incremental learning network based on adaptive resonance theory, with fuzzy ARTMAP (FAM) being one variant. Fuzzy set theory is utilized to adjust the ARTMAP network dynamically. Probabilistic FAM is a mixture of FAM and probabilistic neural networks (PNNs) for the purpose of achieving on-line learning and prediction. This technique improves the classification precision of FAM by introducing PNNs, but the learning speed is low and vulnerable to the impact of noise data. This may lead to overfitting and poor generalization. Literature [2] proposes a type of incremental learning algorithm based on PNN and adjustable fuzzy clustering algorithm (AFC algorithm). This method can easily extract knowledge from the new training samples so as to improve the recognition capacity. Moreover, it can freely enhance the recognition for new classes while deleting the recognition for uninterested classes. There is no need to store the previously used training data and hence the storage space is saved. For about 1000 samples, incremental learning will take 264.3 sec with the use of 5-fold cross-validation. This is unacceptable for intelligent terminals with limited resources.

(2) Incremental Learning based on ELM

Huang *et al.,* from Nanyang Technological University proposed ELM [3-5]. As a single hidden layer feedforward neural network, ELM is different from conventional single hidden layer neural network in that it does not need iterations to adjust the input weights and offset by iterations. Instead the output weights are calculated by obtaining Moore-Penrose pseudoinverse solution and norm minimum-norm least-squares solution to the linear system of equations. Therefore, ELM has superior calculation speed, robustness and generalization, though it is confined to batch learning [6-7]. On this basis, literature [8] describes on-line sequential ELM (OSELM), which is capable of incremental data processing. Only the new individual data or blocks of data need to be learnt. Output matrix obtained by learning the new samples is used to incrementally update the existing output matrix. In this way, the new output matrix not only preserves the knowledge of the existing data, but also learns new knowledge from the new data. Moreover, the learnt data will be precluded from training, thus saving the storage space. But OSELM can only process samples of the existing classes, but not samples of the new classes.

We make some improvements of the OSELM algorithm, and come up with an incremental learning algorithm that is capable of learning samples of new classes while preserving the knowledge of the existing classes.

3. Image Classification Technique Based on Class-Incremental Learning

3.1. Framework of the Algorithm

We propose a non-class-specific recognition framework to preserve the recognition capacity for the existing image classes while incrementally learning the new image classes, as shown in Figure 1.

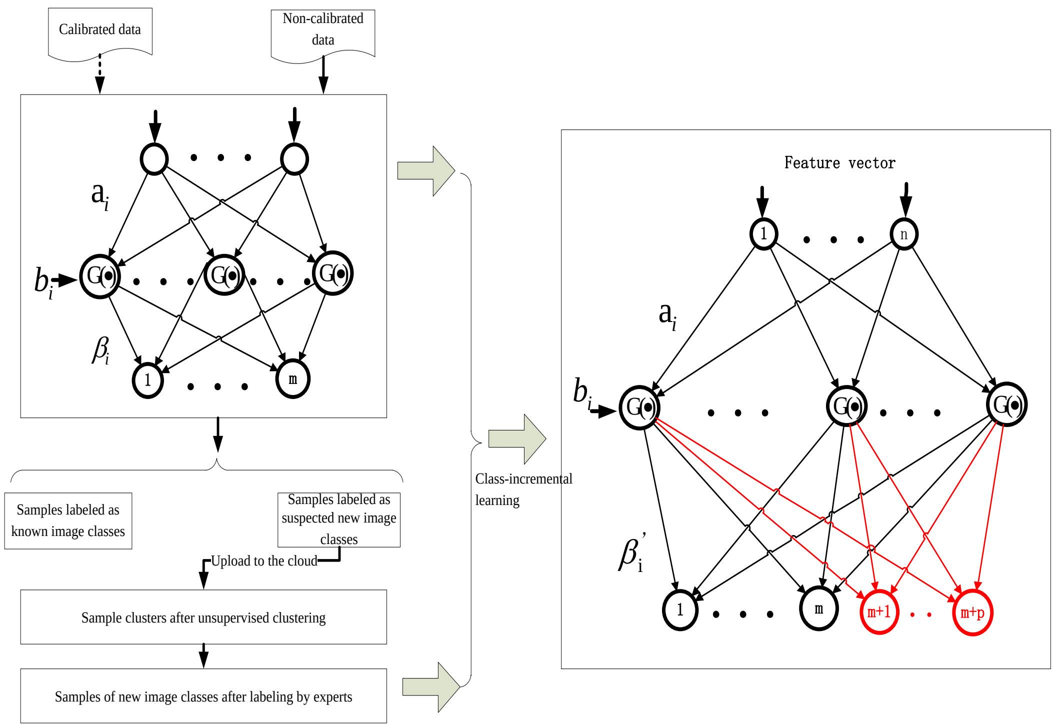

Figure 1. Framework of Image Recognition Method based on ELM

The off-line process is shown in the upper part of the diagram by dotted lines. Calibration is performed using some of the image data collected in the preliminary study or using the published datasets. Thus we build the image recognition model that has recognition capacity for suspected new images based on ELM. At this stage, the model can recognize only a few image classes. This model is deployed to the fighters where image recognition is needed and the images are acquired by the camera equipment. As shown by the solid lines in the diagram, from top to bottom, the image recognition model first makes preliminary judgment on the collected image data. For samples falling within the confidence interval, they are directly calibrated; but for those outside the confidence interval, they are labeled as suspected new image classes. When the suspected new images accumulate to a certain amount, they are uploaded to the cloud, where unsupervised clustering is carried out. Thus the suspected image samples are classified into different clusters. By expert labeling, new image classes are identified. Then the existing image recognition model is updated by the class-incremental approach.

3.2. ELM Model with Recognition Capacity for Suspected Abnormal Images

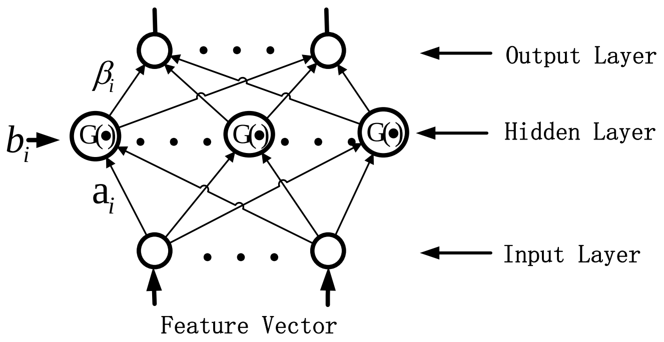

Figure 2. Framework of ELM Algorithm

Figure 2, shows the network structure of ELM algorithm, which is a neural network with a single hidden layer. Compared with other machine learning algorithms, it has higher recognition precision and less operation time. ELM does not require the adjustment of input weights and offset through training, but the output weights are calculated by obtaining Moore-Penrose pseudo-inverse solution and norm minimum-norm least-squares solution to the linear system of equations. The purpose of prediction is to determine the output by looking for nodes with maximum output. ELM algorithm can address the classification and regression problems simultaneously, but here we only consider the classification function of ELM algorithm. The training process of the single hidden layer ELM is shown below.

Algorithm 1: **ELM Algorithm**

Input: Let the training set be $\aleph = \{(x_i, t_i) | x_i \in R^n, t_i \in R^m, i = 1,2,\cdots,N\}$, excitation function G(a,b,x), and the number of hidden nodes $\tilde{N}$.

Output: Output the weight β .

Begin

1. Randomly generate input weight a_i and offset b_i, $i = 1,2,\cdots,\tilde{N}$.
2. Calculate the output matrix H of the hidden layer:

$$H = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_N) \end{bmatrix}_{N \times \tilde{N}}$$

3. Calculate the output weight matrix $\beta = H^{\dagger}T = (H^T H)^{-1} H^T T$, where $H^{\dagger}$ is Moore-Penrose pseudo-inverse of matrix H, $T=[t_1, t_2, \cdots, t_N]^T$.

End

When a new sample x is to be classified by ELM, the value of the output node is calculated as follows:

$$TY_{1\times m} = [G(a_1, b_1, x), \cdots, G(a_{\tilde{N}}, b_{\tilde{N}}, x)]_{1\times\tilde{N}} \cdot \beta_{\tilde{N}\times m}$$

Where m is the number of output nodes and also the number of classes. Correspondingly TY is a row vector containing m values.

ELM classifier then determines which component in vector TY is closest to 1, and the subscript j is taken as the class label of sample x.

$$j = \underset{1 \le j \le m}{\operatorname{argmin}} |1 - TY_j|$$

The confidence interval of sample x is calculated in vector TY as a measure of the subordination degree of sample x to class j. The confidence interval is calculated as follows:

1) Calculate the distance of each component in TY to 1: $D = |TY - 1|$

2) Calculate the inverse of each component in D and denote it as **D_Inverse**=$\frac{1}{D}$;

3) Calculate the proportion of the maximum value of **D_Inverse** in **D_Inverse** and take it as the confidence interval of sample x.

$$confidence = \frac{\operatorname{argmax} D\_Inverse}{\sum D\_Inverse}$$

Each sample collected and labeled off-line has a confidence interval. The set of confidence intervals for a specific class of samples is assumed to obey normal distribution. Thus according to the 3σ principle for normal distribution, the mean μ and standard deviation σ are related by the following expression: P (μ-σ<X≤μ+σ)=68.3%, P (μ-2σ<X≤μ+2σ)=95.4%, P (μ-3σ<X≤μ+3σ)=99.7%. The interval [μ-3σ, μ+3σ] is the normal interval. That is, when the sample falls within this confidence interval, the ELM prediction of the image sample is reliable. The interval (—∞, μ-3σ) or (μ+3σ, +∞) is the abnormal interval. That is, when the sample falls within this confidence interval, the ELM prediction of the image sample is inaccurate, and the sample is suspected as belonging to a new class or as abnormal compared with the existing samples (due to diseases or aging for reasons on the user's side).

Therefore, by differentiating between the normal and abnormal intervals, the ELM model can recognize abnormal or new images.

3.3. Identification and Labeling of Samples of New Image Classes

When the suspected samples have accumulated to a certain amount, they will be uploaded to the cloud, where unsupervised clustering is carried out. By expert labeling, new image classes are identified, and the existing model is updated by the class-incremental approach.

3.4. Transfer and Updating of ELM Model

OSELM is first analyzed with respect to its defects. Then, by modifying OSELM, we propose the ELM-based class incremental approach

3.4.1. OSELM: As described in Section 3.2, ELM model takes the entire training dataset as the input, based on which the training model is derived at one time. Huang establishes OSELM follows:

Step 1: Let the initial dataset be $\aleph_0 = \{(x_i, t_i)\}_{i=1}^{N_0}$. Randomly generate the input weight a_i and offset b_i. Let the excitation function be G(a,b,x), and the number of hidden nodes $\tilde{N}$. By implementing ELM algorithm, the output matrix H_0 of the hidden layer is expressed as follows:

$$H_0 = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_{N_0}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0}) \end{bmatrix}_{N_0 \times \tilde{N}} \quad (1)$$

Thus the output weight is

$$\beta_0 = K_0^{-1} H_0^T T_0 \tag{2}$$

Where:

$$K_0 = H_0^T H_0, \quad T_0 = [t_1, t_2, \cdots, t_{N_0}]^T \tag{3}$$

Step 2: Let the new block of data be $\aleph_1 = \{(x_i, t_i)\}_{i=N_0+1}^{N_0+N_1}$. $\aleph_1$ does not contain the data denoted with different class label compared with the preceding data $\aleph_0$. Then, by taking the same values of a_i, b_i, G and $\tilde{N}$ as in the first step, H_1 can be calculated as well:

$$H_1 = \begin{bmatrix} G(a_1, b_1, x_{N_0+1}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0+1}) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_{N_0+N_1}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0+N_1}) \end{bmatrix}_{N_1 \times \tilde{N}} \tag{4}$$

From this K_1 is calculated:

$$K_1 = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} = [H_0^T \quad H_1^T] \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} = K_0 + H_1^T H_1 \tag{5}$$

Therefore, the updated output weight is

$$\beta_1 = K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} = K_1^{-1} \left(K_1 \beta_0 - H_1^T H_1 \beta_0 + H_1^T T_1 \right) = \beta_0 + K_1^{-1} H_1^T (T_1 - H_1 \beta_0) \tag{6}$$

It can be seen from Formula (6) that β_1 is the function of β_0、 K_1、 H_1 and T_1 while it is unrelated to $\aleph_0$. The used data can be directly discarded, thus saving the storage space.

When there are no new image classes, the model can be updated by one incremental sample or by a batch of incremental samples. However, OSELM is incapable of this.

3.4.2. ELM-Based Class Incremental Approach for Structure Transfer: ELM-based class incremental approach not only aims to achieve incremental learning of the existing image classes, but also the learning of new image classes. For every new image class, the original network structure should be preserved (including input nodes, hidden nodes and connections between them). However, the existing network structure of ELM should be changed as well, by increasing the number of input nodes. Addressing this need, we propose the ELM-based class incremental approach that transfers the existing ELM network structure to the domain of new class data. The existing ELM network structure is updated with the new class data so that the model can deal with calibration samples of new image classes (Figure 3).

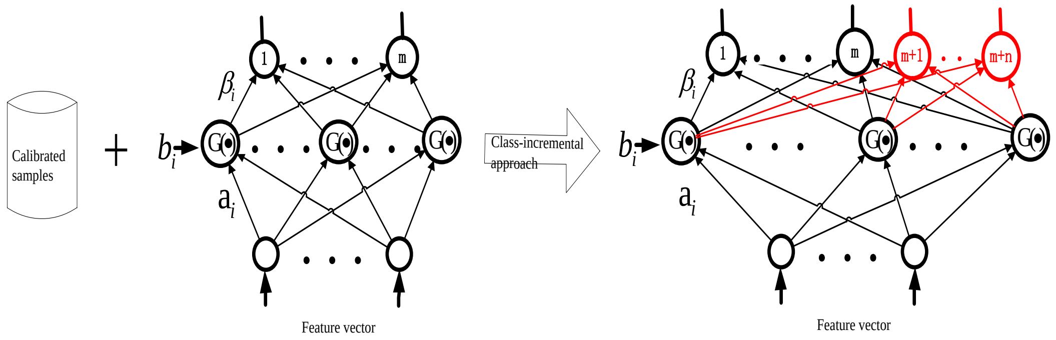

Figure 3. Network Structure of Class-Incremental Approach

The derivation of the new algorithm is presented below:

Suppose some of the calibrated data constitute the set $S = \left\{\left(x_s^{(i)}, y_s^{(i)}\right) \middle| i = 1,2,\dots,N_0\right\}$, and a user submits some of the new class data $D = \left\{\left(x_d^{(i)}, y_d^{(i)}\right) \middle| i = 1,2,\dots,N_1\right\}$. To simplify the description without losing generality, we suppose there is only 1 class label in data D and it is larger than the maximum class label in S by only 1.

The classification model is trained by S using ELM algorithm:

$$\beta_0 = K_0^{-1} H_0^T T_0 \tag{7}$$

Where

$$H_0 = \begin{bmatrix} G(a_1, b_1, x_s^{(1)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_s^{(1)}) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_s^{(N_0)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_s^{(N_0)}) \end{bmatrix}_{N_0 \times \tilde{N}} \tag{8}$$

$$T_0 = \begin{bmatrix} t_1^T \\ \vdots \\ t_{N_0}^T \end{bmatrix}_{N_0 \times m}, \qquad K_0 = H_0^T H_0 \tag{9}$$

Then for D, we calculate the value of H_1:

$$H_1 = \begin{bmatrix} G(a_1, b_1, x_d^{(1)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_d^{(1)}) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_d^{(N_1)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_d^{(N_1)}) \end{bmatrix}_{N_1 \times \tilde{N}} \tag{10}$$

Each class label in D is represented by a multi-dimensional vector, which has one more column compared with the class label in T_0, as expressed below:

$$T_1 = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix}_{N_1 \times (m+1)} \tag{11}$$

Taking S and the calibrated D as the training dataset, β_1 is obtained as follows:

$$\beta_1 = K_1^{-1}\begin{bmatrix}H_0\\H_1\end{bmatrix}_0^T\begin{bmatrix}T_0 \cdot M\\T_1\end{bmatrix} \tag{12}$$

Where

$$K_1 = \begin{bmatrix}H_0\\H_1\end{bmatrix}^T\begin{bmatrix}H_0\\H_1\end{bmatrix}, \qquad M = \begin{bmatrix}1 & \cdots & 00\\ \vdots & \ddots & \vdots 0\\ 0 & \cdots & 10\end{bmatrix}_{m\times(m+1)} \tag{13}$$

M is a transformation matrix, by which a column of 0 is added to the right of the matrix so make T_0 and T_1 have the same number of columns:
After derivation, β_1 is expressed as follows:

$$\beta_1 = K_1^{-1}\begin{bmatrix}H_0\\H_1\end{bmatrix}_0^T\begin{bmatrix}T_0 \cdot M\\T_1\end{bmatrix} = \beta_0 M + K_1^{-1}H_1^T(T_1 - H_1\beta_0 M) \tag{14}$$

It is easy to see that β_1 is no longer directly related to dataset S, which means the model can be updated by incremental learning to make it capable of recognizing new classes.

In practice, the data of both existing classes and new classes will appear as incremental data. The model can determine the classes of the data by reading the class label. If the data are of the existing classes, the network structure will be updated by the ELM-based class incremental approach. Otherwise, OSELM is implemented.

4. Experiment and Result Analysis

4.1. Data Preparation

Both two image datasets used in the experiment come from UCI machine learning repository (http://archive.ics.uci.edu/ml/), with one named image segments and the other satellite images.

Image segments, which are subimages totaling 2310, are extracted from 7 outdoor images, and each has a size of 3 pixels×3 pixels. From each subimage, 19 attributes are extracted, based on which the subimage is classified as 1 of the 7 original images.

Satellite images are scenes captured by landsat multispectral scanner with 4 channels, and therefore 4 images are obtained in each frame. One frame is selected, and the region of interest is delineated (82 pixels ×100 pixels) so that the subimages (3 pixels×3 pixels) are extracted. The purpose is to classify the central pixels into any of the 6 categories based on 36 attributes of the subimages (red soil, cotton crop, gray soil, damp gray soil, soil with vegetation stubble, and very damp gray soil).

Basic information of image segments database and satellite images database is shown in the table below.

Table 1. Information of UCI Datasets

Name of dataset	Number of attributes	Number of categories	Number of samples
Image segment	19	7	2310
Satellite image	36	6	6435

4.2. Algorithm Performance Evaluation

The proposed CIELM was verified on the above datasets. The performance was evaluated in three aspects: model parameter configuration, recognition capacity for abnormal images, and classifier performance under different number of new samples. An ordinary PC was used with CPU frequency of 2.6GHz and MATLAB2009a software. The source codes of ELM were downloaded from the Home Page of Huang, who invented CIELM1 (www.ntu.edu.sg/home/egbhuang).

4.2.1. Model Parameter Configuration: The network structure and the parameters of machine learning algorithm were first optimized, depending on the specific model used. For CIELM, only the number of nodes in the hidden layer needs to be optimized.

We performed training and verification to choose the optimal number of nodes in the hidden layer. The dataset was taken as the training dataset as well as the testing dataset. The dataset was traversed when the number of nodes in the hidden layer varied from 1 to 300. The results are shown in Figure 4.

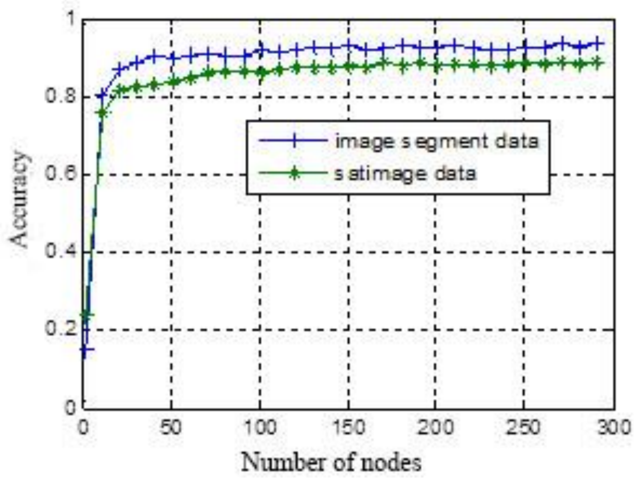

Figure 4. Performance of ELM Algorithm and Number of Nodes in the Hidden Layer

When the number of nodes in the hidden layer exceeds 100, the model achieves the best classification performance and stability. Thus, for subsequent experiment, the number of nodes in the hidden layer is set as 100.

4.2.2. Recognition Experiment for Abnormal Images: With one image class in the dataset designated as abnormal and the remaining classes as known classes, the ELM model was trained by the known image classes and tested by the abnormal image classes. This process was done for 6 times in the satellite image dataset [9-10] and for 7 times in the image segment dataset. Each time different images of the same abnormal class were used, and the average accuracy was calculated, as shown in Table 6.2.

Table 2. Result of Recognition Experiment for Abnormal Images

Name of dataset	Training data	Testing data	Average accuracy
Satellite image	Five image classes randomly	The remaining image class	98.02%
Image segment	Six image classes randomly	The remaining image class	97.05%

It can be seen from Table 2 that the model recognition accuracy for the abnormal image class reaches over 97%, which indicates excellent performance.

4.2.3. Classifier Performance VS Number of New Samples: We are concerned about the enhancement of classification capacity for new image classes with the increasing number of new samples. The number of samples needed for incremental learning is the key issue.

For the dataset containing m image classes, classification experiments were performed for *m* times. For the i-th experiment, the procedures are as follows:

(1) Select data of the i-th class as the incremental data, and the data of the remaining *m*-1 classes as the known data;

(2) Train the ELM model ELM_A using data of *m*-1 known classes, and the data of the i-th class are equally divided into n parts, each containing ChkSize data (depending on the properties of the dataset, 5 values of ChkSize are chosen);

(3) Update the ELM_A model using CIELM algorithm and ChkSize data each time. After updating, record the classification accuracy.

The experiment results are shown in Figure 5, for satellite image dataset.

It is easy to see that when the data block size is 60 for a new class in the satellite image dataset, the classification accuracy of 80% can be achieved after about 5 iterations. As more iterations are carried out, the classification accuracy is further increased. Since more knowledge of the new class is learnt with incremental learning of the new class data, the model classification accuracy increases correspondingly.

Similarly, for the image segment dataset, when the data block size is 6, the classification accuracy of 90% can be achieved after about 4 iterations. Classification accuracy will further increase with the number of iterations (Figure 6).

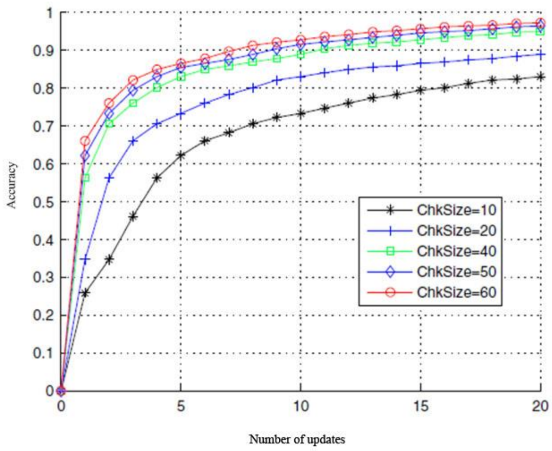

Figure 5. Performance Evaluation of CIELM based on Satellite Image Dataset

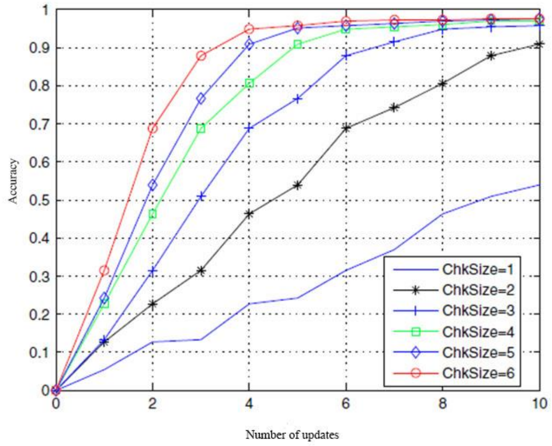

Figure 6. Performance Evaluation of CIELM based on Image Segment Dataset

5. Summary

We propose a non-class-specific image recognition method based on ELM, which is capable of incremental learning for new samples. Thus the model is automatically updated by incremental learning. The performance of the new algorithm is verified based on image datasets.

Acknowledgment

This work is supported by the Foundation research and advanced technology research project of Henan Province(No. 142300410374).

References

[1] Z. Zhao, Z. Chen, Y. Chen, S. Wang and H. Wang, “A Class Incremental Extreme Learning Machine for Activity Recognition”, Cognitive Computation, vol. 6, no. 3, **(2014)**, pp. 423-431.

[2] Z. Wang, M. Jiang, Y. Hu and H. Li, “An Incremental Learning Method Based on Probabilistic Neural Networks and Adjustable Fuzzy Clustering for Human Activity Recognition by Using Wearable Sensors”, Information Technology in Biomedicine, IEEE Transactions on , vol. 16, no. 4, **(2012)** July, pp. 691-699.

[3] 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, **(2011)**, pp. 107-122.

[4] Y. Chen, Z. Zhao, S. Wang and Z. Chen, “Extreme learning machine based device displacement free activity recognition model”, SoftComputing, vol. 16, no. 9, **(2012)** September, pp. 1617-1625.

[5] G. B. Huang, Z. Bai, L. L. C. Kasun and C. M. Vong, “Local Receptive Fields Based Extreme Learning Machine”, IEEE Computational Intelligence Magazine, vol. 10, no. 2, **(2015)**, pp. 18-29.

[6] G. Huang, G. B. Huang, S. Song and K. You, “Trends in Extreme Learning Machines: A Review”, Neural Netwokrs, vol. 61, no. 1, **(2015)**, pp. 32-48.

[7] G. B. Huang, “An Insight into Extreme Learning Machines: Random Neurons, Random Features and Kernels”, Cognitive Computation, vol. 6, **(2014)**, pp. 376-390.

[8] N. Y. Liang, G. B. Huang and P. Saratchandran, “Fast and accurate online sequential learning algorithm for feedforward networks”, IEEE Transactions on Neural Networks, vol. 17, no. 6, **(2006)**, pp. 1411-1423.

[9] J. Tang, C. Deng, G.-B. Huang and B. Zhao, “Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine”, IEEE Transactions on Geoscience and Remote Sensing, **(2014)**.

[10] Z. Zhao, Y. Chen, J. Liu, Z. Shen and Mi. Liu, “Cross-people Mobile-phone based Activity Recognition.IJCAI, **(2011)**, pp. 2545-2550.