

# Time-series Processing of Large Scale Remote Sensing Data with Extreme Learning Machine over Cloud Infrastructure

XXX<sup>a</sup>, xxx<sup>a,b</sup> and xxx<sup>a</sup>

<sup>a</sup>*Zhejiang University*

<sup>b</sup>*School*

---

## Abstract

Nowadays, land-cover change detection plays a more and more important role in environment protection and many other fields. However, the current land-cover change detection methods encounter the problem of low efficiency and can't be expanded to parallel computing to deal with fast-growing remote sensing(RS) data. To solve the above problems, we propose a novel ELM-based land-cover change detection method, in which the supervised classification capability, fast training speed and high generalisation performance of ELM is utilized to efficiently train the land-cover classifier and apply the classifier to time-series satellite imageries' analysis. As one experiment, we apply our method to the analysis of rapid land use change in Changjiang River Delta over the past two decades due to accelerated urbanization. Moreover, we prove our method's scalability by another experiment of large scale RS imageries processing over the cloud infrastructure.

*Key words:* Extreme learning machine, Remote Sensing, Classification, Change Detection, Time-series, MapReduce, Hadoop.

---

## 1 Introduction

Nowdays, the available time-series RS images provide a new way for land-cover change detection which is widely required in various fields. However, except for the traditional challenge of high accuracy, large scale RS data processing also requires the land-cover change detection to be efficient and highly scalable.

---

\* Corresponding author

*Email address:* xxx@zju.edu.cn (XXX).

This paper proposes a novel elm-based land-cover change detection method with faster processing speed and higher scalability, which is practiced on cloud infrastructure.

### *1.1 Land-cover Change Detection with Time-series RS Images*

Change detection is the process of identifying differences in the state of an object or phenomenon by observing it at different times[1]. With the development of satellite technology, massive time-series high resolution multispectral RS images are available for applications. By comparing two sets of RS images, taken of the same area at different time, we can manually handle the change detection job by view. In fact, it is possible for the computer to process the digital RS images, and provide an automatical change detection technology, which is available for large scale processing of RS images. In summary, time-series RS image-based land-cover change detection method is to identify the interesting land-cover changes between "before" time images and "after" time images through RS image digital processing and comparison of RS images in an automatical way.

In the recent decades, timely and accurate change detection of Earth's surface features plays a more and more important role in better decision making. Automatic change detection can be used in such diverse applications as land usage analysis, disaster monitoring, snow-melt measurements, forest coverage monitoring and other environment changes. Especially, land-cover change detection is one of the major direction of change detection application. Various papers[2][3][4] have presented their work of applying change detection technology to the analysis of land-use and land-cover. Ross S. Lunetta[4] performed the change detection experiment in the biologically complex landscape of the Neuse River Basin, North Carolina using Landsat5 and 7 imagery collected in May of 1993 and 2000. Another typical application is performed by Qihao Weng[2], to analysis the rapid land use change taken place in Zhujiang Delta over the past decades with the help of change detection technology.

### *1.2 Challenges of Land-cover Change Detection*

In the applications of RS images, land-cover change detection has been a problem for some time. Among all the factors that cause land-cover change detectoin unavailable to real situations, change detection methods or algorithms used play a key role. Traditionally, the main challenge to land-cover change detection methods is the detection accuracy. However, with increase scale of RS images and these images' high resolution, the method's processing speed and scability rise to be another two major challenges.

Presently, large data processing has become a hot research topic of computing science. Similarly, in the field of RS data processing, increasing data size also brings some new challenges to processing methods. On one hand, efficient processing method is one of the best solutions to reduce processing time, on the other hand, parallel computing or distributed computing can greatly improve the processing efficiency of large scale data. In order to support parallel computing, the scalability of the processing method become another challenge.

### *1.3 Major Contributions of Elm-based method*

This paper's major contribution is to propose a novel elm-based method for land-cover change detection. Using our method, the analysis of rapid land use change in Changjiang River Delta over the past two decades due to accelerated urbanization is done. Moreover, the land-cover change detection result is evaluated and compared to the current methods, which use the classifier of BP network and SVM. The novel approach supports high speed processing of single data set and can be extended to parallel computing on cloud infrastructure for the processing of large scale RS data. In our experiment, we expand the trained ELM network to hadoop based cloud infrastructure to prove the method's scalability. Through the evaluation of the performance improvement of parallel computing, we prove the method's capability to deal with increasing RS imageries.

ELM is a recently proposed and widely used machine learning method, with the capability of multiclass classification and universal approximation[5][6][7][8]. One of the major characteristics of ELM is its very fast training processing, which directly increases the processing efficiency of land-cover change detection method. The other advantage of ELM is the high generalisation performance. Generalisation is the key factor that will decide whether the trained network can be widely applied to large scale RS data. Moreover, our proposed change detection method also takes advantage of ELM's less human intervention for the process of network training. In conclusion, with ELM's high speed of training, high generalisation and less manual intervention, we design our novel land-cover change detection method to overcome the challenges of large scale RS data processing.

## 2 Related Work

### 2.1 *Methods for Land-cover Change Detection*

Land usage analysis and land cover mapping has long been an area of research focus, and a wide range of methods including cross-correlation analysis, post-classification, image differencing, image ratioing, principal components analysis and so on have been explored. Both the traditional methods post-classification and cross-correlation determine the land-cover change through the "before" time and "after" time classification map, and produce comparable result. Although both methods are more automatic than the others like image differencing, the unsupervised classifier used result in low accuracy, which is one of the most important challenges that change detection methods face. The methods of image differencing, image ratioing and principal components analysis involve transformations of the original spectral bands so as to enhance the land cover changes. These methods and their improvement do increase the accuracy of land-cover change detection, but it is limited to small scale data processing and are lack of automation.

With the advances of machine learning, several alternative land-cover change detection methods using neural network are proposed and applied in the analysis of landscape. Presently, both supervised and unsupervised neural network algorithms have made progress. Compared to unsupervised land-cover transitions detection methods, the supervised ones provide more information about the kinds of transitions that occurred on the ground and are less affected by the difference atmospheric conditions, sensor calibration, and ground condition[9][10]. The paper[10] presents a technique for using predictive modeling, which predicates "before" and "after" pixel value, to identify unusual changes in imageries. Begum Demir[9] present a novel iterative active learning(AL) technique aimed at defining effective multitemporal training sets to be used for the supervised detection of land-cover transitions in a pair of time-series RS imageries. In comparison, unsupervised approaches require fewer manual interventions[11][12][13], and are quite suitable in the situations that the ground-truth is always unavailable. However, the low accuracy of unsupervised classification and predication severely limit the application of unsupervised approaches.

Although the classification and predication capability of neural network brings much improvement to the traditional land-cover change detection methods, the common and widely applied neural networks, like BP network and SVM, exist several inadequacies. The most two important aspects that exist room for improvement may be the low speed of training and low generalization.

## 2.2 ELM for Remote Sensing Application

Presently, little work of applying ELM to the field of RS imagery processing has been done. Mahesh Pal[14] did the land cover supervised classification experiment with ELM using remote sensing images. Wu Jun[15] revealed the equivalence between ELM and the positive and negative fuzzy rule system. Moreover, Wu Jun indicate the claim that ELM can be naturally used for training the positive and negative fuzzy rule system quickly for image classification. However, only simple comparison of classification cabability between ELM and other classification network was done in both efforts, and no time series images processing work are tried, not along large scale RS iamge processing.

## 2.3 ELM Based Methods for Large-scale Data Set Processing

Extreme learning machine is a new learning algorithm for single-hidden layer feedforward neural networkd(SLFNs) which randomly chooses the input weights and analytically determines the output weights of SLFNs[16]. In fact, ELM is to minimize the training error as well as the norm of the output weights[16][17], and thus to achieve better generalization performance. As the hidden nodes, including the input weights and hidden nodes' threshold are randomly generated,  $\mathbf{H}$ , namely the hidden layer output matrix of the neural network[18][19], can be directly calcualted, thus greatly reducing the network training time.

As large-scale data set processing has been the trend of present applications, some researchers have paid their attentions to the scapability of ELM methods. Through improving the scapability, they extend the ELM variant methods to parallel computing infrastructures to deal with large-scale data set training. Qing He[20] proposed to implement an efficient paralalled ESVM based on the current and powerful parallel programming framework MapReduce. Moreover, an incremental learning algorithm for ESVM, which can meet the requirement of online learning to update the existing network, as well as its parallel version, is proposed in [20]. Mark van Heeswijk[21] compared three distinct ways of accelerating the evaluation of this ensemble of ELMs, and concluded that the GPU-accelerated and parallelized ELM ensemble achieved attractive speedups over using single CPU. Yongjiao Sun[21] propose an OS-ELM based ensemble classification framework for distributed classification in a hierarchical P2P network. In spite of the above contributions to the scability improvement, their proposed parallel computing methods are limited to the model building process. Meanwhile, no real applications are applied in and their efficiency on remote sensing imagery processing are known.

### 3 ELM-based Method for Land-cover Change Detection

In this paper, we propose our new land-cover change detection method based on extreme learning machine using time-series RS images. The change detection is calculated through the comparison between "before" time and "after" time land-cover mapping, which is the result of land-cover classification of RS images. With the high speed of network training and high generalisation performance of classification, the new method has advantage of fast processing and high scalability, compared to the traditional neural network based methods. Moreover, the new method is extended to cloud infrastructure to deal with large scale RS images processing.

#### 3.1 Land-cover Change Detection Method

The whole land-cover change detection method can be divided into two components: training unit and detection unit as shown in Figure 1. Training unit is to build an ELM networking using preprocessed training and testing samples, which has adjusted hidden node number and can be directly applied to real remote sensing image classification with the highest generalisation performance. While the detection unit firstly calculates both "before" time land-cover mapping and "after" time land-cover mapping through the classification of time-series remote sensing images using the trained network, and then works out the change detection statistics and change map by means of comparing the pair of time-series land-cover mapping.

##### 3.1.1 Land-cover Classification and Samples

According to different terrains area and resolution of remote sensing images, the land-cover can be classified into specific categories. In fact, in order to assure the precision, statistical property and representative of the sample, training sample of each category should be selected through visual interpretation and field survey[22]. In our application, according to the investigation of land-cover of Taihu Area and the landsat images' resolution, we classify the land-cover into four types as described in Table 1.

To represent the categories of the land cover, every type of the result is assigned to an integer, ranging from 1 to 6. As we use multi-spectral RS images, every pixel is multi-dimensional, and thus every sample can be described:

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n$$

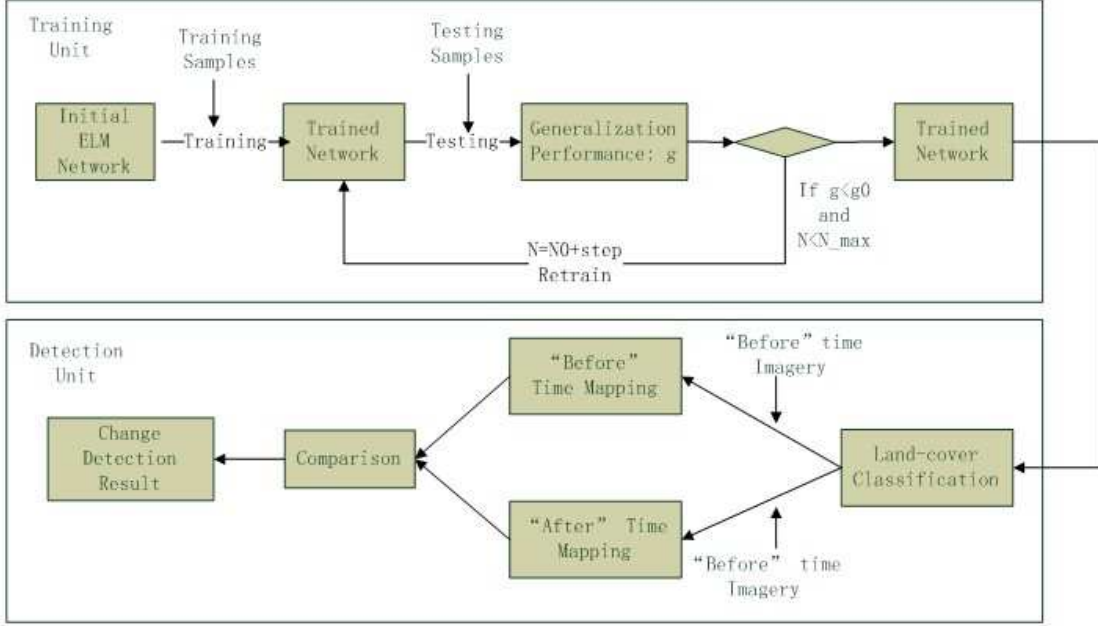


Figure 1. ELM-based Land-cover Change Detection Method

Ground object category	Image features	Ground object description
Urban	Purple, faint red	road or building
Water	Blue or deep blue	River or pond
Vegetation	Deep Green, deep Yellow	Forest,Hill
Arable land	Green, light green, cyan	land for agriculture
Wetlands	Black	Wetlands

Table 1  
Image Interpretation of Each Ground Object Category

$$t_i = f(x) = \begin{cases} 1, & \text{if type is Urban} \\ 2, & \text{if type is Water} \\ 3, & \text{if type is Vegetation} \\ 4, & \text{if type is Arable Land} \\ 5, & \text{if type is Wetlands} \end{cases}$$

where  $\mathbf{x}_i$  is the input,  $t_i$  is the output,  $n$  is the number of spectralis of the RS images,  $i$  ranges from 1 to  $N$ , and  $N$  is the number of samples. In order to obtain enough training data, we annotate six  $20 \times 20$  squares in the chosen RS image by visual, each of which represents one type of land cover. In result,  $20 \times 20 \times 5$  number of samples are provided for the ELM network training, namely,  $N = 4500$ . Enough samples for network testing can also be obtained in a similar way, but from different image individuals in the same image type, in order to assure that the network has enough generalization performance to fulfill large-scale change detection.

### 3.1.2 Incremental Training Unit

ELM network with too few hidden nodes has no capability to achieve enough training classification accuracy, while too many hidden nodes will cause over-fitting, causing low testing classification accuracy. In the training unit of our method, an incremental algorithm, with network testing is proposed to adjust the number of hidden nodes and activation function, finally to obtain an ELM network with highest generalization performance.

---

#### Algorithm 1 Incremental ELM Training Algorithm

---

**Require:** Given a training samples set  $\chi = \{(\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in R^m, i = \{1, \dots, N\}\}$ , a testing samples set  $\psi = \{(\mathbf{x}_i, \mathbf{t}_i) | \mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in R^m, i = \{1, \dots, \widehat{N}\}\}$ , activation function  $g(x)$ , maximum hidden node number  $M_{max}$ , minimum hidden node number  $M_{min}$ , increasing step  $step$ , training times of every step  $num$ :

Step 1) **Initialization:** Let  $M = M_{min}$ , testing accuracy  $E = 0$ , number of hidden nodes with highest testing accuracy  $M_{highest} = 0$ , highest testing accuracy  $E_{highest} = 0$

Step 2) **Learning Step:**

**while**  $M < M_{highest}$  **do**

a) increase by one the number of hidden nodes  $M : M = M + step$

b) training the new network  $num$  times with the training samples set  $\chi$

c) calculate the average testing accuracy  $E$  of the  $num$  trained networks with the testing samples set  $\psi$

**if**  $E > E_{highest}$  **then**

d)  $E_{highest} \leftarrow E$

e)  $M_{highest} \leftarrow M$

**end if**

**end while**

---

### 3.1.3 Detection Unit through Time-series Mappings Comparison

In the detection unit of our method, the classified "before" time land-cover mapping and "after" time land-cover mapping are compared to generated the land-cover transition map, namely the change detection result. Using the ELM network, builded in the training unit, the input time series RS imageries are classified into six categories of land-cover, with computing every pixel's classification. Known every pixel's land-cover type, we can draw land-cover mappings, which visually display the coverage of the land surface. Moreover, land-cover classification statistical tables are computed to accurately compare the coverage of every land-cover category.

In the process of land-cover mappings comparsion, both the trends of every land-cover category and every location's land-cover transition are calculated.



As one result of change detection, the whole area land-cover transition map is drawn, which visually display the change of land-cover of every pixel. As the other result of change detection, confusion matrix of land-cover detection, which represents every transition of coverage from one type to another type, is calculated.

### 3.2 ELM for Classification

ELM is a recently proposed and widely used machine learning method, with the capability of multiclass classification and universal approximation[5][6][7][8]. It is famous for its high speed of training and high generalisation performance for classification, as well as other characteristics like less human intervention. In short, ELM algorithm is to minimize the training error as well as the norm of the output weights[16][17], and thus to achieve better generalization performance.

$$\text{Minimize : } \left\| \mathbf{H}\beta - \mathbf{T} \right\|^2 \text{ and } \left\| \beta \right\|$$

That means

$$\mathbf{H}\beta = \mathbf{T},$$

where

$$\mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_M, b_1, \dots, b_M, \mathbf{x}_1, \dots, \mathbf{x}_N) = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_M \cdot \mathbf{x}_1 + b_M) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_N \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_M \cdot \mathbf{x}_N + b_N) \end{bmatrix}_{N \times M},$$

$$\beta = \begin{bmatrix} \beta_i^T \\ \vdots \\ \beta_M^T \end{bmatrix}_{M \times m}, \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N \times m} \quad \text{and}$$

As the  $M$  hidden nodes, namely the parameters  $\mathbf{w}_i$  and  $\mathbf{b}_i$ , are randomly generated, we can directly calculate the left parameters of the neural network,  $\beta_i$ , through the following formula:

$$\hat{\beta} = \mathbf{H}^\dagger \mathbf{T},$$

where  $H^\dagger$  is the Moore-Penrose generalized inverse of matrix  $\mathbf{H}$ .

### 3.3 Scalability with Cloud Infrastructure

With increasing number and increasing size of remote sensing images, our land-cover change detection faces the problem of large-scale data processing.

Moreover, the present parallel computing based cloud infrastructure has become an main resolution to deal with large-scale data in various fields. However, in order to expand to cloud infrastructure for distributed computing, our method should be designed to be scalable. Thanks to the high generalisation performance of ELM neural network and its less manual intervention in training and classification, the detection unit program of our algorithm can run individually on the slaves with mapped input data set. We describe the our method's expansion to mapreduce based parallel computing module in the following content.

Every pair of time-series RS images can be represented as  $d_i = (I_{before}, I_{after})$ ,  $i = 1, 2, \dots, l$ , where  $l$  is the number of pairs of time-series RS images,  $I_{before}$  and  $I_{after}$  represent the "before" time RS image and "after" time RS image individually. Thus we can represent the total RS images set as:

$$D = \{d_1, d_2, \dots, d_l\}$$

With the trained and tested ELM network, we can deticate the change of every sub-area that is coved by a pair of RS image  $d_i$  spearately. Similarly, we represent the detecated change statistics of every sub-area as

$$\mathbf{r}_i = [s_{i1}, \dots, s_{im}], i = 1, 2, \dots, l$$

where  $m$  is the number of land cover types, and  $s_{ij}$ ,  $j = 1, 2, \dots, m$  is the land cover statistic of jth land cover type difined. Moreover, the detected change statistics of the whole area, combined by the sub-area, is

$$\mathbf{R} = [s_1, \dots, s_m] = \mathbf{r}_1 + \mathbf{r}_2 + \dots + \mathbf{r}_l$$

Here is the algorithm for the processing of large scale RS image set  $D$  with MapReduce.

---

**Algorithm 2** Distributed Processing Algorithm for Large Scale RS Images with MapReduce

---

**Require:** number of hadoop job workers:  $v$

- a) **Map:** divide  $D$  into subsets:  $s_k = d_{(k-1) \times \frac{l}{v}}, \dots, d_{k \times \frac{l}{v}}$  and  $k = 1, \dots, v$
  - b) **Process:** Calculate  $r_i$  of every pair of time-series RS images on the slaver of hadoop cluster
  - c) **Merge:** Merge all the result,  $r_i$  of jth worker node into result set  $rs_j$
  - c) **Reduce:** Reduce the merged result  $rs_i$ ,  $i = 1, \dots, v$  into the final result set,  $R$
-

## 4 Evaluation

### 4.1 Incremental ELM Training

In our experiments, we will first use our incremental ELM training algorithm to train the ELM network with the highest testing accuracy. The key factors that affect the generalisation performance are the hidden node number and activation function.

Firstly, we prepare our training and testing samples in the following way. The time-series imageries, token by the satellite of LANDSAT-5 with the sensor mode of TM, cover the same area of Taihu Lake. The "before" time imagery is token on Dec 25 2001, while the "after" time imagery is token on Feb 28 2008, as show in figure 2.

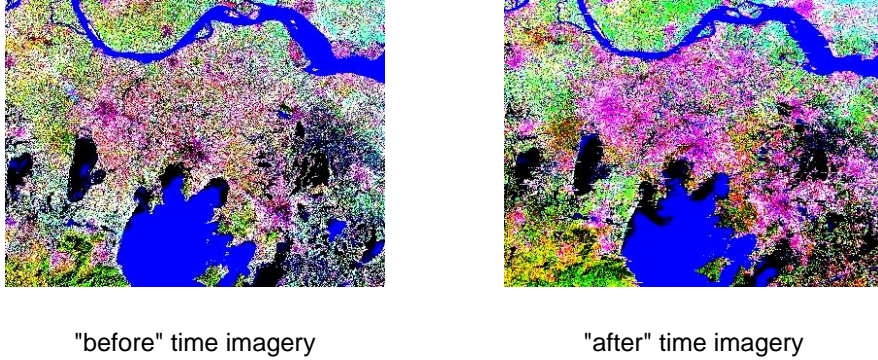


Figure 2. The Time-series Imageries.

Through the investigation of land-cover, we label the five different land-cover types on the pair of time-series RS imageries manually. To get the training and testing data sets, five  $20 \times 20$  labeled squares with different land-cover types are labeled on both the "before" and "after" time imageries as shown in figure 3.

In this experiment, we collect a training samples set  $\chi$  as well as a testing samples set  $\psi$  with the size of 3000. And every sample is a four elements tuple, in which the first three elements are the RGB values of every pixel, while the last element is the label of the pixel.

We incrementally train and test the ELM network with the hidden node number ranging from 10 to 400 and an increasing step of 3. Moreover, every network are trained and tested 3 times and the average testing accuracy are

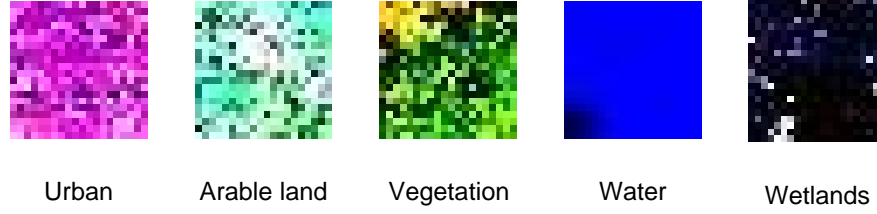


Figure 3. Five labeled different land-cover samples with the size of  $20 \times 20$  pixels.

calculated to assure the accuracy. The final result can be represented by figure 4, which shows the networks' testing accuracy trend with different hidden node number and activative function. From the figure we can firstly conclude that all the activate functions, except for hardlim, performs similarly the the hidden node number increasing. Secondly, the average testing accuracy of the networks with the activative functions of sig, sin, tribas and hardlim, improves greatly before the hidden node number reaches 100, but are stabilized when the hidden node number reaches about 150. Finally, we can see that choosing tribas as the activative function and 350 to 400 hidden node number will bring a little better testing accuracy, which is about 0.98, if the application can accept the increase of training time.

#### 4.2 Land-cover Classification Performance of ELM

The most valuable factors of ELM are the high learning speed and generalisation performance, which are also the two advantages that our land-cover change detection method's accuracy and scability are based on. In this experiment, we compare ELM to BP and SVM in the application of RS imagery classification.

One factor that we will compare is their training time, with their generalisation performance ensured. In this experiment, we set the hidden node number to 100 and use the active function of sin, which can generate an ELM network with testing accuracy of about 0.97 using 3000 training samples and 3000 testing samples. Similarly, we adjust both the parameters of SVM and BP networks to assure their testing accuracy ranging between 0.95 and 0.96 using the same training and testing samples. Moreover, we compare the training time of all the methods using increasing training data size, and to specific size of training data, multi training is performed and the average training time is calculated. The training time comparison result can be seen in figure 5.

From the figure we can argue that ELM always costs the least training time.

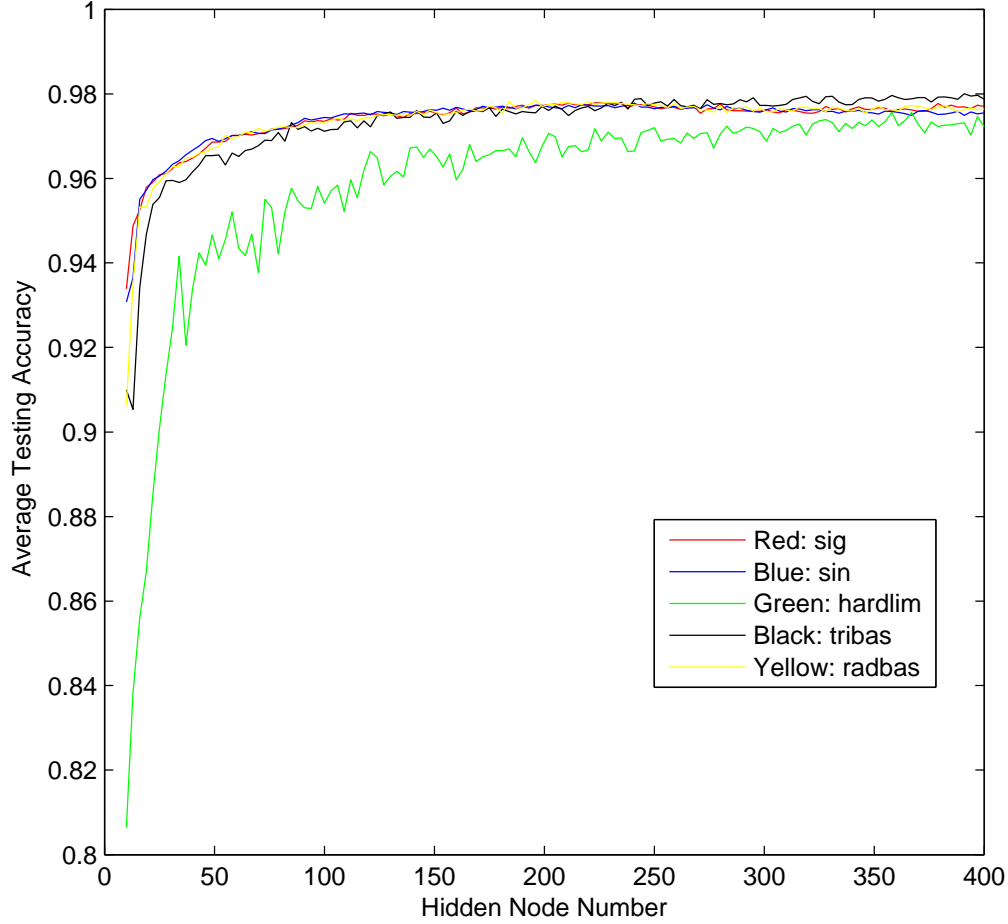


Figure 4. Incremental ELM Training Result.

With the increasing training data size, the training time of BP network increases greatly and are unstable. From the data of figure 6, we can see the BP network takes about 47 times more time than ELM to learn. Although SVM network training is much faster than BP network, it still takes nearly 7 times more time to learn than ELM.

The other factor that we will compare is their testing accuracy, with increasing training and testing samples.

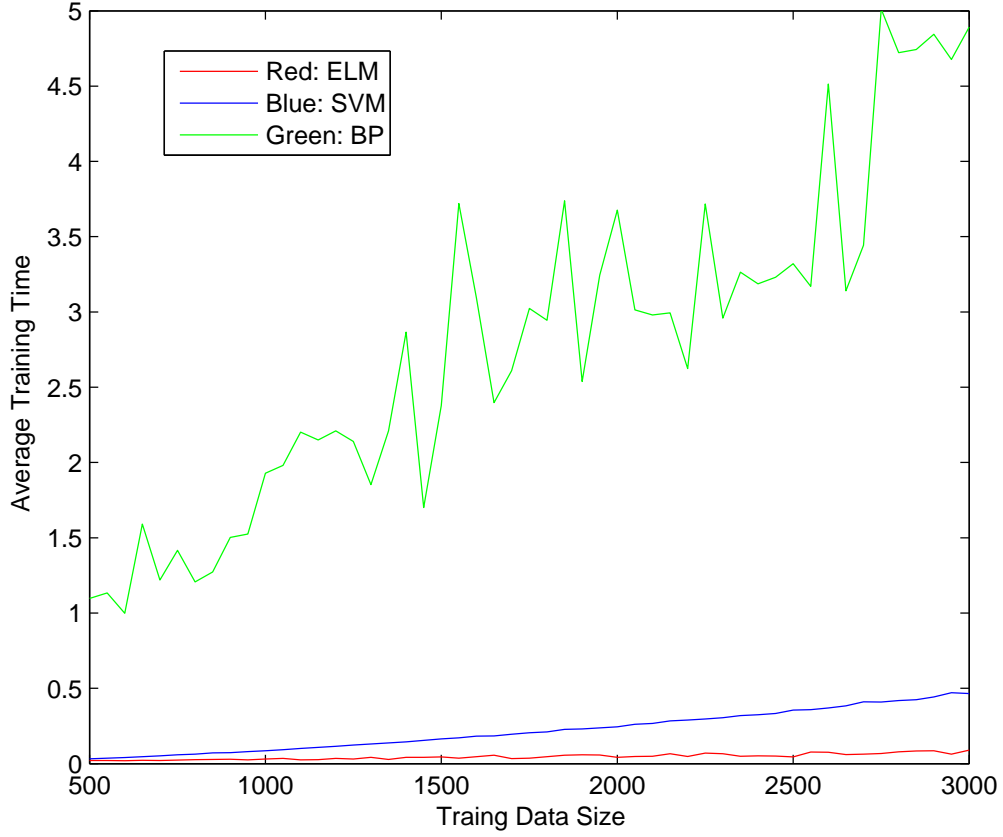


Figure 5. Average Training Time Comparison.

	ELM			BP			SVM		
Size	TrainTime	TestTime	TestAcc	TrainTime	TestTime	TestAcc	TrainTime	TestTime	TestAcc
500	0.018000	0.006000	0.973200	0.031253	0.005436	0.934000	1.474384	0.008222	0.958000
750	0.027000	0.006000	0.971333	0.058225	0.010785	0.936000	1.377492	0.008237	0.927333
1000	0.032000	0.007000	0.971000	0.084406	0.016624	0.943000	1.788723	0.008306	0.933300
1250	0.037000	0.006000	0.971360	0.117043	0.023910	0.945600	1.755750	0.008938	0.957760
1500	0.048000	0.012000	0.971733	0.159608	0.034197	0.951333	1.963340	0.008352	0.963467
1750	0.056000	0.011000	0.971429	0.217431	0.041878	0.951429	2.205566	0.008535	0.957943
2000	0.061000	0.012000	0.974150	0.247115	0.051873	0.955000	2.311722	0.008987	0.939900
2250	0.059000	0.014000	0.973422	0.322188	0.062474	0.951111	2.111166	0.009078	0.964089
2500	0.078000	0.014000	0.975560	0.360570	0.074148	0.953200	2.840775	0.008804	0.965120
2750	0.080000	0.015000	0.975382	0.416586	0.087604	0.952000	3.348820	0.009248	0.959673
3000	0.084000	0.015000	0.974500	0.455141	0.098914	0.952333	3.323896	0.009408	0.961267

Table 2

RS Imagery Classification Performance Comparison among ELM, BP and SVM

#### 4.3 Change Detection Result

#### 4.4 Scalability

## 5 Conclusions

## References

- [1] A. SINGH, “Review article digital change detection techniques using remotely-sensed data,” *International Journal of Remote Sensing*, vol. 10, no. 6, pp. 989–1003, 1989.
- [2] Q. Weng, “Land use change analysis in the zhujiang delta of china using satellite remote sensing, gis and stochastic modelling,” *Journal of Environmental Management*, vol. 64, no. 3, pp. 273 – 284, 2002.
- [3] J. M. Read and N. S.-N. Lam, “Spatial methods for characterising land cover and detecting land-cover changes for the tropics,” *International Journal of Remote Sensing*, vol. 23, no. 12, pp. 2457–2474, 2002.
- [4] R. S. Lunetta, J. Ediriwickrema, D. M. Johnson, J. G. Lyon, and A. McKerrow, “Impacts of vegetation dynamics on the identification of land-cover change in a biologically complex community in north carolina, usa,” *Remote Sensing of Environment*, vol. 82, no. 23, pp. 258 – 270, 2002.
- [5] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, pp. 513 –529, april 2012.
- [6] G.-B. Huang, Y.-Q. Chen, and H. A. Babri, “Classification ability of single hidden layer feedforward neural networks,” *IEEE Transactions on Neural Networks*, vol. 11, pp. 799–801, 2000.
- [7] R. Zhang, G.-B. Huang, N. Sundararajan, and P. Saratchandran, “Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis,” *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, vol. 4, pp. 485–495, July 2007.
- [8] G.-B. Huang, “Universal approximation using incremental constructive feedforward networks with random hidden nodes,” 2006.
- [9] B. Demir, F. Bovolo, and L. Bruzzone, “Detection of land-cover transitions in multitemporal remote sensing images with active-learning-based compound classification,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. PP, no. 99, pp. 1 –12, 2011.
- [10] C. Clifton, “Change detection in overhead imagery using neural networks,” *Applied Intelligence*, vol. 18, pp. 215–234, March 2003.
- [11] S. Patra, S. Ghosh, and A. Ghosh, “Unsupervised change detection in remote-sensing images using one-dimensional self-organizing feature map neural network,” in *Information Technology, 2006. ICIT ’06. 9th International Conference on*, pp. 141 –142, dec. 2006.
- [12] S. Patra, S. Ghosh, and A. Ghosh, “Unsupervised change detection in remote-sensing images using modified self-organizing feature map neural network,” in *Computing: Theory and Applications, 2007. ICCTA ’07. International Conference on*, pp. 716 –720, march 2007.

- [13] K. Chen, C. Huo, Z. Zhou, and H. Lu, "Unsupervised change detection in high spatial resolution optical imagery based on modified hopfield neural network," in *Natural Computation, 2008. ICNC '08. Fourth International Conference on*, vol. 4, pp. 281–285, oct. 2008.
- [14] M. Pal, "Extreme learning machine for land cover classification," *CoRR*, vol. abs/0802.1412, 2008.
- [15] W. Jun, W. Shitong, and F.-l. Chung, "Positive and negative fuzzy rule system, extreme learning machine and image classification," *International Journal of Machine Learning and Cybernetics*, vol. 2, pp. 261–271, 2011. 10.1007/s13042-011-0024-1.
- [16] G.-B. Huang, "Extreme learning machine: a new learning scheme of feedforward neural networks," 2004.
- [17] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 13, pp. 489 – 501, 2006. jce:title;Neural Networksj/ce:title; jce:subtitle;Selected Papers from the 7th Brazilian Symposium on Neural Networks (SBRN '04)i/ce:subtitle; jxocs:full-name;7th Brazilian Symposium on Neural Networksj/xocs:full-name;.
- [18] G.-B. Huang, "Learning capability and storage capacity of two-hidden-layer feedforward networks," 2003.
- [19] G.-B. Huang, "Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions," 1998.
- [20] Q. He, C. Du, Q. Wang, F. Zhuang, and Z. Shi, "A parallel incremental extreme svm classifier," *Neurocomputing*, vol. 74, no. 16, pp. 2532 – 2540, 2011. jce:title;Advances in Extreme Learning Machine: Theory and Applicationsj/ce:title; jce:title;Biological Inspired Systems. Computational and Ambient Intelligencej/ce:title; jce:subtitle;Selected papers of the 10th International Work-Conference on Artificial Neural Networks (IWANN2009)i/ce:subtitle;.
- [21] M. van Heeswijk, Y. Miche, E. Oja, and A. Lendasse, "Gpu-accelerated and parallelized elm ensembles for large-scale regression," *Neurocomputing*, vol. 74, no. 16, pp. 2430 – 2437, 2011. jce:title;Advances in Extreme Learning Machine: Theory and Applicationsj/ce:title; jce:title;Biological Inspired Systems. Computational and Ambient Intelligencej/ce:title; jce:subtitle;Selected papers of the 10th International Work-Conference on Artificial Neural Networks (IWANN2009)i/ce:subtitle;.
- [22] Z. J.-L. J. X.-G. Yu Xiu-Lan, Qian Guo-Hui, "Learning sample selection in multispectral remote sensing image classification using the bp neural network," 1999.