

Time-series Processing of Large Scale Remote Sensing Data with Extreme Learning Machine

Jiaoyan Chen ^a, Huajun Chen ^a , Cong Fang ^a, Ningyu Zhang ^a,
and Zhaohui Wu ^a

^a*College of Computer Science and Technology, Zhejiang University, Hang Zhou,
China, 310027*

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 not be expanded to parallel computing to deal with large scale remote sensing(RS) data. This paper presents a novel extreme learning machine(ELM) based land-cover change detection method with high accuracy, high efficiency and little human intervention. The evaluation results show that ELM outperforms the traditional methods, e.g., SVM and BP network, in terms of training speed and generalisation performance, when applied in remote sensing(RS) imagery classification. In our experiments, we apply our method to the analysis of rapid land use change in Taihu Lake region over the past decade.

Key words: Extreme Learning Machine, Remote Sensing, Classification, Change Detection, Time-series.

1 Introduction

In the field of land-cover change detection, detection accuracy is always one of the most important factors. Nowadays, with increasing scale of RS data, the new method is required to be highly efficient and scalable. This paper proposes a novel ELM-based land-cover change detection method with faster processing speed, higher detection accuracy and less human intervention.

* Corresponding author

Email address: huajunsir@zju.edu.cn (Huajun Chen).

1.1 Land-cover Change Detection with Time-series RS Imageries

Change detection is the process of identifying differences of the status of an object or phenomenon by observing it at different times[1]. The available time-series RS imageries, taken of the same area at different time, provide an efficient way for change detection. We can manually compare time-series imageries and handle the change detection job by view, but RS imagery digital processing can provide an automatic way. Briefly speaking, time-series RS imagery based land-cover change detection method is to identify the interesting land-cover changes between "before" time imagery and "after" time imagery.

Recently, land-cover change detection technology plays a more and more important role in various applications, such as land usage analysis, disaster monitoring, snow-melt measurements and so on. 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 Landsat7 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

Traditionally, the main challenge to land-cover change detection methods is the detection accuracy. However, with increasing scale of RS imageries and these imageries' high resolution, the method's processing speed and scalability rise to be another two major challenges.

In the field of RS data processing, increasing data size brings some new challenges to the traditional methods. Therefore, the new land-cover change detection method should be designed with high efficiency and high scalability. First, the new method should improve its training and testing speed with new or improved machine learning algorithm. Second, the new method should be designed to be as automatic as possible, with little human intervention.

1.3 Major Contributions

Briefly speaking, this paper's major contribution is to propose a novel ELM-based land-cover change detection method. In detail, it contains tree aspects:

- (i) First, our land-cover change detection method is the first work of applying ELM to time-series RS imagery processing. The application successfully analyses the rapid land use change in Taihu Lake region over the past decade. This proves that ELM can be applied in time-series RS data processing with high accuracy and high efficiency.
- (ii) Second, ELM-based RS imagery classifier is proved to have fast training speed and high generalisation performance. In this paper, classification performance of ELM is evaluated and compared to SVM and BP network.
- (iii) Third, our method detects land-cover change automatically with little human intervention. This enables the method to be expanded to parallel computing.

In conclusion, ELM is firstly applied to time-series RS data intelligent processing. And the evaluation proves that ELM outperforms the traditional classifiers, in terms of training speed and generalisation performance. With high accuracy, high efficiency and little manual intervention, our ELM-based land-cover change detection method is successfully applied to land usage analysis.

2 Related Work

2.1 *Methods for Land-cover Change Detection*

In the field of land usage analysis, a wide range of methods[1] have been explored. Here, we classify them into two categories: classification map based and non-classification map based.

- (i) First, post-classification and cross-correlation methods determine land-cover change through the comparison of "before" time and "after" time classification maps. Although both methods are more automatic, they face the problems of low detection accuracy and too much human intervention.
- (ii) Second, the non-classification map based methods, like 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 are lack of automation.

Among the classifiers used in change detection, both supervised and unsupervised algorithms have made progress. The supervised methods provide more information about the kinds of transitions that occurred on the ground. They are less affected by different atmospheric conditions, sensor calibration, and

ground conditions[5][6]. Begum Demir[5] presents a novel iterative active learning technique aimed at defining effective multitemporal training sets. It can be used for the supervised detection of land-cover transitions. In comparison, unsupervised approaches require fewer manual intervention[7][8][9], and are quite suitable in the situations that the ground-truth is always unavailable.

2.2 ELM for Remote Sensing Application

ELM is a recently proposed machine learning method with the capability of multiclass classification and universal approximation[10][11][12][13]. However, only a little work of applying ELM to RS imagery processing has been done until now.

Mahesh Pal[14] did the supervised land-cover classification experiment with ELM using remote sensing imagery. However, only the classification accuracy and computational cost of ELM are tested and compared to BP network. No application of ELM is designed in [14]. Wu Jun[15] indicates that ELM can be used for training the positive and negative fuzzy rule system quickly for specific imagery classification. Although ELM is applied to train the network for RS imagery classification, no time series imagery processing work are tried. Briefly, no application of ELM to land-cover change detection or time-series RS imagery processing has been done until now.

3 ELM-based Method for Land-cover Change Detection

In this section, we describe our ELM-based land-cover change detection method in detail. The following chapters will discuss the whole method, its two components: training unit and detection unit, as well as preparatory work.

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 network for RS imagery classification. Through adjusting the hidden node number and activation function, high generalisation performance of the classifier is achieved. Detection unit is responsible for classification mapping calculation and comparison.

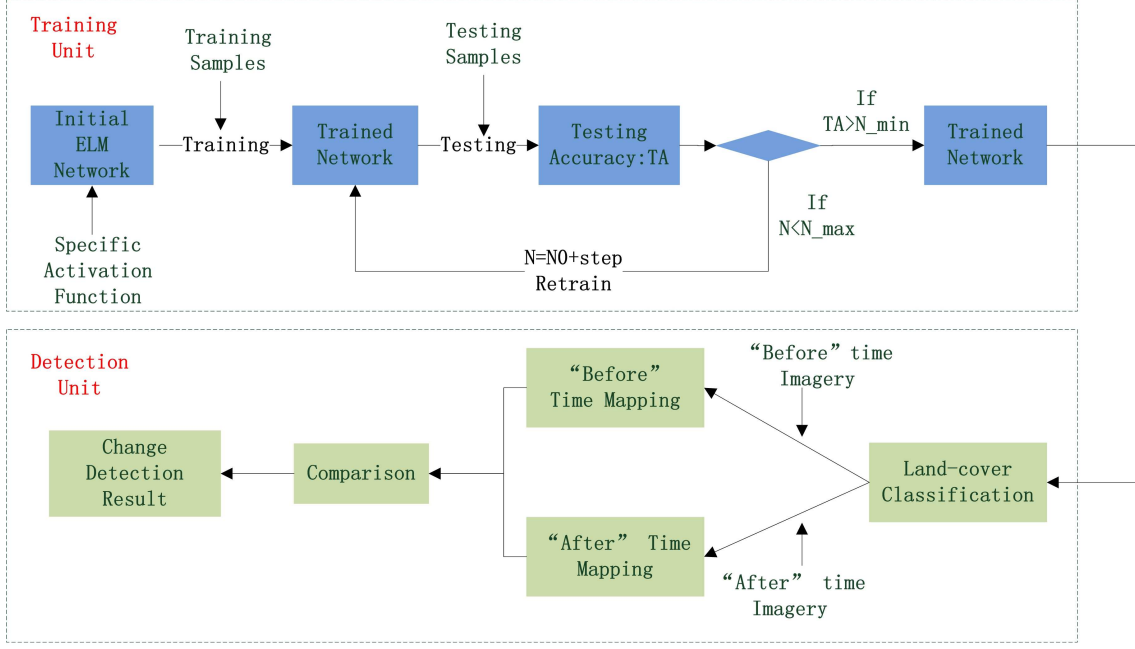


Figure 1. ELM-based Land-cover Change Detection Method

3.1.1 Preparatory Work: Land-cover Categories and Samples

According to different terrains area and RS imageries, the land-cover can be classified into specific categories. In our application, according to the investigation of Taihu Lake region's terrains and the chosen Landsat-7 imageries, we classify the land-cover into five types as described in Table 1. In order to ensure precision and representative, training and testing samples should be labeled through visual interpretation and field survey[16].

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

The land-cover category is represented by an integer, ranging from 1 to 5. Every pixel contains the values of all the bands of the multi-spectral RS imagery. So every sample can be described as:

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

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

, where \mathbf{x}_i is the input, \mathbf{t}_i is the output, n is the number of bands of the RS imagery, i ranges from 1 to N , and N is the number of samples.

As the preparatory work of our application and experiments, samples are manually labeled in the RS imagery. In order to ensure the network's actual availability, both "before" time and "after" sample imageries should contribute to the samples.

3.1.2 Training Unit

Training unit is to build the land-cover classifier for the detection unit. In the classifier, too few hidden nodes will lead to low classification capability, while too many hidden nodes will cause over-fitting. Moreover, the activation function will also greatly influence the classifier's performance. Therefore, our training unit is responsible for choosing proper hidden node number and activation function.

In training unit, we build the land-cover classification network in the following steps:

- (i) First, choose a specific activation function.
- (ii) Second, build the classifier using our land-cover classification network building procedure, as shown in Algorithm 1.
- (iii) Third, for every activation function, execute the procedure several times and calculate the average testing accuracy.
- (iv) Fourth, choose a proper activation function. It can build a classifier than can achieve the highest average testing accuracy.

In brief, we incrementally training and testing our network with different hidden node number for every activation function. The network with the highest average testing accuracy is the one that we build in training unit.

Algorithm 1 Land-cover Classification Network Building Procedurer

Require: Given training sample set χ , testing sample set ψ , activation function g :

Step 1) **Initialization:** Initialize the land-cover classifier with default hidden node number and g

Step 2) **Building Step:**

while no enough testing accuracy **do**

a) increase the hidden node number

b) training the new network with χ

c) calculate the network's testing accuracy with set ψ

if the hidden node number is too large **then**

d) record the network with the highest testing accuracy

e) break

end if

f) record the network with the highest testing accuracy

end while

Step3) return the network with the highest testing accuracy

3.1.3 Detection Unit

Detection unit is mainly responsible for two jobs: calculate the land-cover classification mappings and find the differences between time-series mappings.

Land-cover classification mapping represents the land-cover type of every pixel. It is drawn through RS imagery classification. In detail, the procedure contains the following steps.

- (i) Input the "before" time and "after" time imageries with the size of N .
 $N = W \times H$, where W is the width and H is the height.
- (ii) Preprocess the imagery and translate it into input set:

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

, where i ranges from 1 to N and n is the number of bands.

- (iii) Calculate the land-cover type matrix \mathbf{T} using the builded classifier. That is to calculate $\mathbf{T} = \mathbf{H}\beta$, where

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

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

\mathbf{w}_i , \mathbf{b}_i , M and β_i are the parameters of the builded classifier, m is the number of land-cover categories.

- (iv) Translate the vector \mathbf{t}_i^T into land-cover type of every pixel. The position of the largest element in \mathbf{t}_i^T is the integer that represents the land-cover type.
- (v) Draw the land-cover classification mapping. Different land-cover type of the pixel use different colour.

Detection unit finds the interesting changes through comparison of land-cover classification mappings. The comparison mainly contains two aspects.

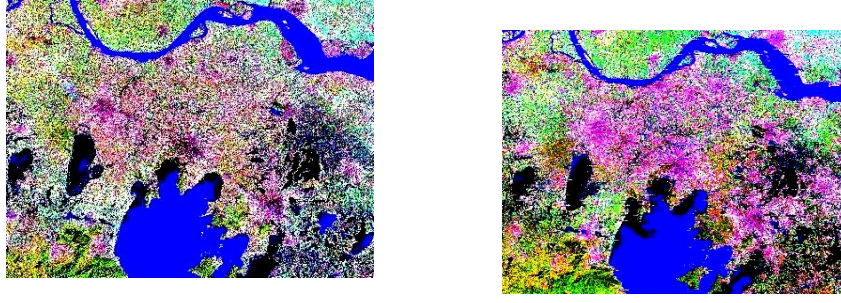
- (i) First, every location's land-cover transition can be calculated. This is done simply by checking the "before" time land-cover type and the "after" time land-cover type of every pixel. Moreover, the land-cover transition mapping, which represents the land-cover type change of every position, can be drawn.
- (ii) Second, the change of coverage of every land-cover type can be calculated. Through comparing the coverage area, we can find the expansion or shrinkage of every land-cover type. For example, the expansion map of urban area can be drawn. What's more, land-cover classification pie charts can be drawn to accurately and visually display the land-cover change of every type.

4 Evaluation

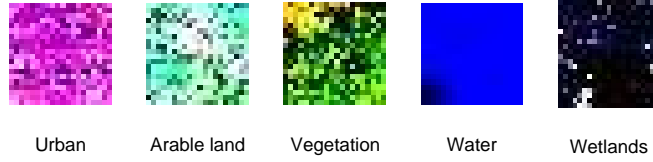
4.1 Training Unit Analysis

Training unit mainly builds the land-cover classifier. This chapter will analysis the classifier building procedure and its result.

Firstly, we show our original imageries and samples in Figure 2. The pair of time-series imageries are token by LANDSAT-5 satellite with TM sensor mode. They both cover Taihu Lake region in east China. Five labeled 20×20 squares of different land-cover types are shown in Figure 2(c).



(a) "Before" time imagery, token on Dec. 25, 2001 (b) "After" time imagery, token on Feb. 28, 2008



(c) Samples of Different Land-cover Types

Figure 2. Original Imageries and Samples

The land-cover classifier building procedure can be seen in Figure 3. In Figure 3, the hidden node number increases from 10 to 400 with a step of 3. Every-time, the network is trained and tested several times and the average testing accuracy is calculated.

From the figure, we can give three conclusions.

- (i) First, all the activation functions, except for hardlim, performs similarly with different hidden node number.
- (ii) Second, the average testing accuracy improves greatly before the hidden node number reaches 100, but is stabilized when the hidden node number exceeds about 150.
- (iii) Third, the network with 350 to 400 hidden nodes and the activation function of tribas can achieve the testing accuracy of more than 98%. This network can be chosen as the builded land-cover classifier of training unit.

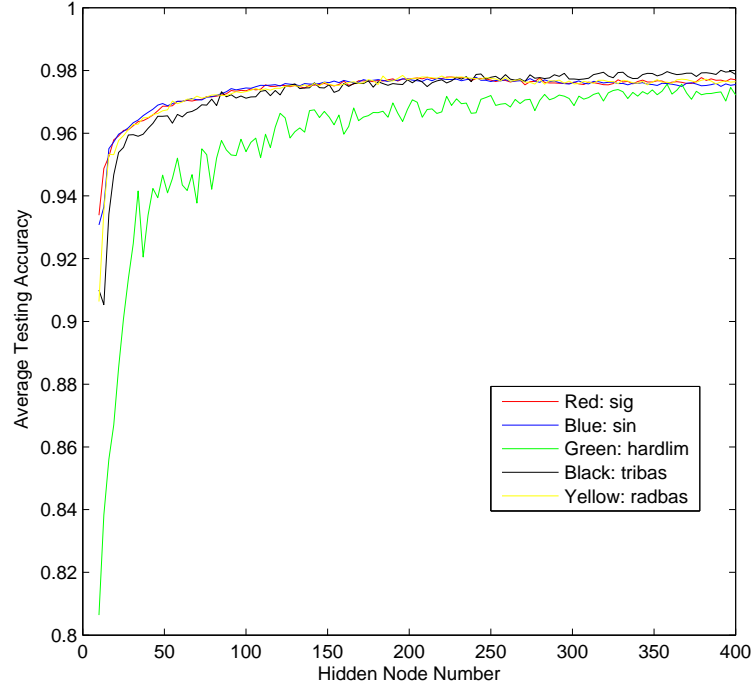


Figure 3. Land-cover Classifier Building Procedure and Result.

4.2 Land-cover Classification Performance of ELM

This chapter will analysis ELM network's fast training speed and high generalisation performance in RS imagery classification. Moreover, comparison between ELM and SVM, BP network will be evaluated.

One factor that we will compare is their training time. The hidden node number of ELM network is set to 100 and sin is chosen as the activation function. This ensures testing accuracy to be more than 95%. Similarly, both parameters of SVM and BP networks are adjusted to ensure testing accuracy to be more than 95%. The average training time of all the tree methods is shown in Figure 4.

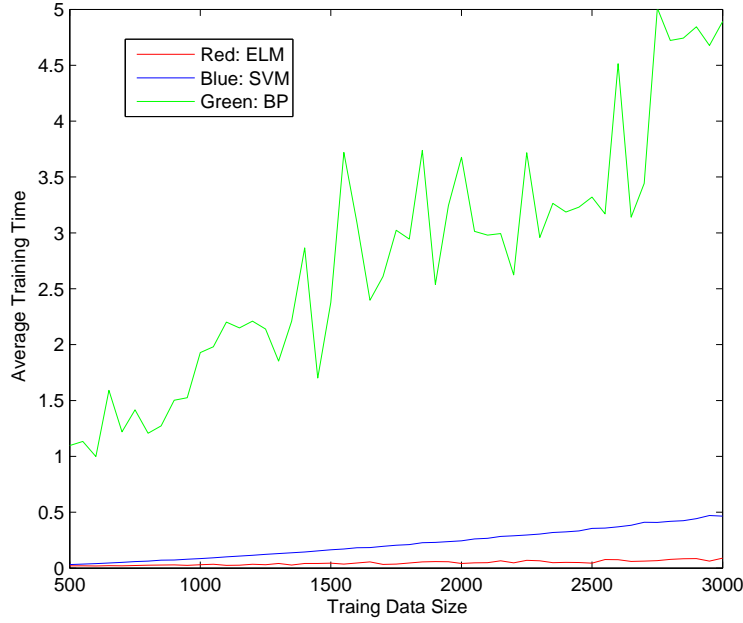


Figure 4. Average Training Time of ELM, BP and SVM.

From the figure we can argue that ELM always costs the least training time. When the data size increases, the training time of BP network increases greatly and are unstable. While the training time of ELM network increases linearly and the slope is small. From Figure 4 and Table 2, we can see the BP network takes about 47 times more time than ELM to learn. And SVM network takes nearly 7 times more time to learn than ELM.

Data Size	ELM			BP			SVM		
	Training Time	Testing Time	Testing Acc	Training Time	Testing Time	Testing Acc	Training Time	Testing Time	Testing Acc
500	0.0180	0.0060	97.3200%	0.0312	0.0054	93.4000%	1.4743	0.0082	95.8000%
750	0.0270	0.0060	97.1333%	0.0582	0.0107	93.6000%	1.3774	0.0082	92.7333%
1000	0.0320	0.0070	97.1000%	0.0844	0.0166	94.3000%	1.7887	0.0083	93.3300%
1250	0.0370	0.0060	97.1360%	0.1170	0.0239	94.5600%	1.7557	0.0089	95.7760%
1500	0.0480	0.0120	97.1733%	0.1596	0.0341	95.1333%	1.9633	0.0083	96.3467%
1750	0.0560	0.0110	97.1429%	0.2174	0.0418	95.1429%	2.2055	0.0085	95.7943%
2000	0.0610	0.0120	97.4150%	0.2471	0.0518	95.5000%	2.3117	0.0089	93.9900%
2250	0.0590	0.0140	97.3422%	0.3221	0.0624	95.1111%	2.1111	0.0090	96.4089%
2500	0.0780	0.0140	97.5560%	0.3605	0.0741	95.3200%	2.8407	0.0088	96.5120%
2750	0.0800	0.0150	97.5382%	0.4165	0.0876	95.2000%	3.3488	0.0092	95.9673%
3000	0.0840	0.0150	97.4500%	0.4551	0.0989	95.2333%	3.3238	0.0094	96.1267%

Table 2

RS Imagery Classification Performance Comparision among ELM, BP and SVM.

The other factor that we will compare is their generalisation performance. We

can evaluate it in the following two aspects.

- (i) First, the testing accuracy of ELM is always higher than that of BP and SVM networks.
- (ii) Second, ELM network can soon reaches high generalization performance, even with small training data set. while the other two networks need more training samples to assure high testing accuracy.

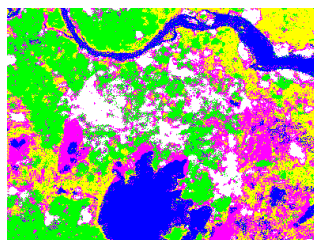
In summary, we can argue that ELM network cost much less time to train and can achieve higher generalisation performance than BP and SVM in RS imagery classification.

4.3 Change Detection Result Analysis

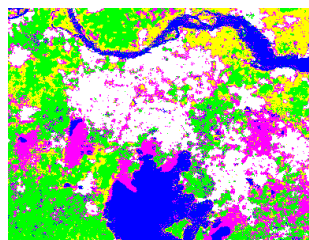
Two major contributions of detection unit are time-series land-cover classification mappings and land-cover change detection result. This chapter will evaluate the two contributions.

The time-series land-cover classification mappings are shown in Figure 5. Five different colours represent five different land-cover types. Compared to the original imageries in Figure 2(a)(b), we can visually judge that both mappings truly reflect the actual ground surface.

Coverage change mapping of every land-cover type is one of the formats to display land-cover change detection result. The cyan area in Figure 6 shows the expansions of urban. From the mapping, we can visually judge that a lot of urban area has been expanded around the original urban area.



(a) "Before" Time



(b) "After" Time

Figure 5. Time-series Land-cover Classification Mappings. Whrite area represents urban, blue area represents water, yellow area represents arable land, green area represents vegetation, manenta area represents wetlands.

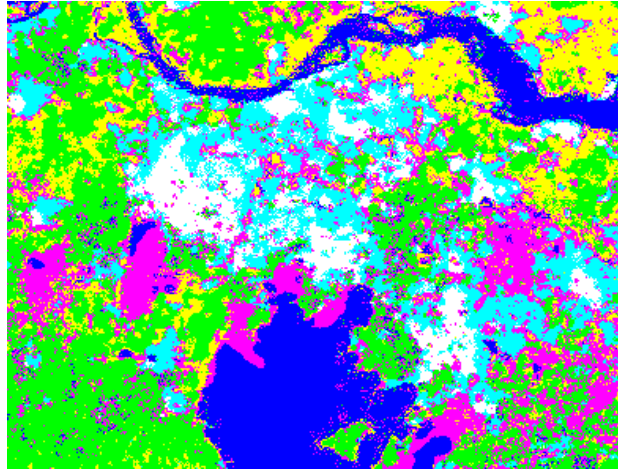


Figure 6. Urban Area Change Detection. The cyan are the new urban area that appears in year 2008, compared to the land-cover of year 2001.

Another format to display land-cover change is the proportion pie chart and histogram, as shown in Figure 7 and Figure 8. In both Figure 7 and Figure 8, the proportion of every land-cover type is displayed and compared. From the result, we can easily conclude that the urban area is rapidly expanding, from 13% to 24%. While the area of water, arable land and wetlands is shrinking.

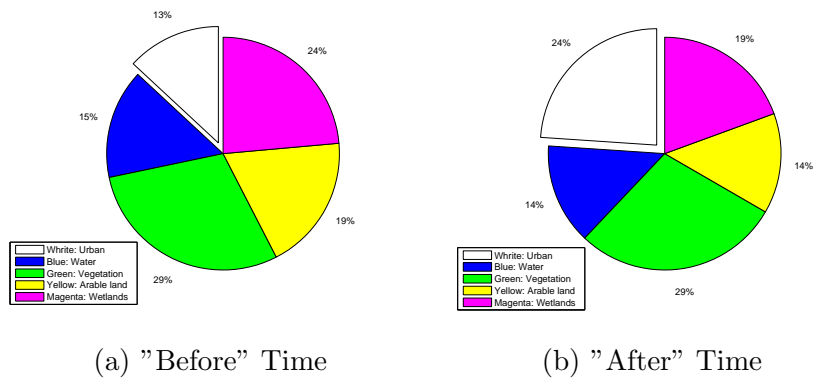


Figure 7. Land-cover Classification Result

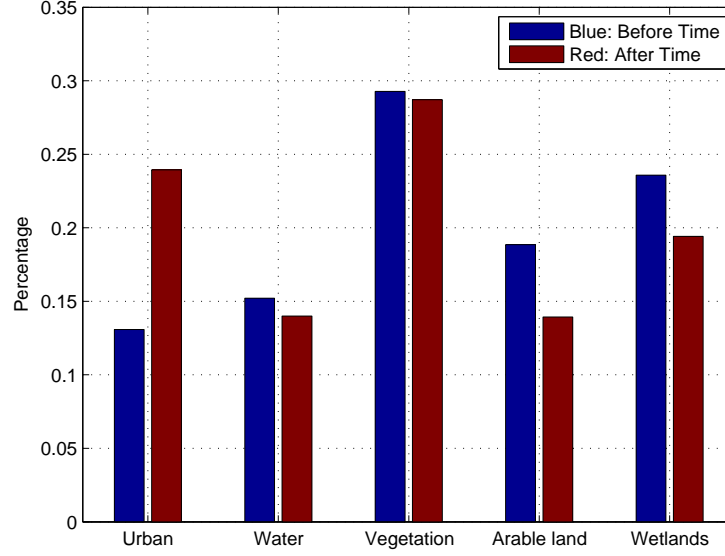


Figure 8. Change Detection Result of Five Different Land-cover Area .

5 Discussion

This paper proposes an ELM-based land-ocver change detection method with high accuracy, high efficiency and less human intervention. It is the first application of applying ELM to time-series RS imagery processing. The application result and comparison experiments prove that ELM can be sucessfully applied in RS imagery classification.

Our ELM-based land-cover change detection method is shown in Figure 1. It mainly contains the training unit and the detection unit. The training unit is to build the land-cover classifier. Through the analysis of Figure 3, we can find that the testing accuracy of our builded network can achieve more than 97%. This characteristic means that the classifier can be applied in the detection unit for the classification of other RS imageries. Detection unit will draw the land-cover classification mapping. And then calculate the land-cover change through comparison of time-series land-cover classification mappings.

Our method sucessfully analysis land usage change of Taihu Lake region from the year 2001 to the year 2008. The urban area expansion mapping is shown in Figure 6. Compared to the original time-series imageries in Figure 2(a)(b), Figure 6 accurately reflects the expansion of urban area. The proportion pie chart and histogram in Figure 7 and Figure 8 provide more statistic results of change detection. The result proves that ELM can be sucessfully applied in

time-series RS imagery processing.

In order to highlight ELM’s advantages in RS imageries classification, the paper evaluate the comparison experiments between ELM network and SVM,BP network. According to the training time comparison results in Figure 4 and Table 2, we can directly conclude that ELM network costs much less train time than SVM and BP network. Moreover, the train time of ELM network increases linearly and slowly when the training data size increases. From the average testing accuracy statistics in Table 2, we can find that ELM’s testing accuracy is always higher than SVM and BP network. For example, when the training and testing data size reaches 3,000, the testing accuracy of ELM exceeds 97%, while SVM and BP network only reach about 96% and 95%.

6 Future Work and Conclusion

More research topics of large scale time-series RS imagery processing with ELM have been proposed in our research group. Our future work is to expand the land-cover change detection method to parallel computing on hadoop-based cloud infrastructure. At the same time, ELM will be applied to other RS imagery applications, such as object identification, rainfall estimation, cloud cover evaluation and so on. What’s more, ELM’s capability of online learning, such as OS-ELM, will be explored. we consider it a good resolution to the training of large scale and increasing RS imageries.

Time-series RS imagery processing plays a more and more important role in various fields. But the new problems, such as increasing satellite data, are challenging the traditional methods. This paper contributes the first application of ELM network to time-series RS imagery processing. It proposes an ELM-based land-cover change detection method with less human intervention, higher processing efficiency and higher detection accuracy. Using this method, the land-cover change of Taihu Lake region from the year 2001 to the year 2008 is accurately detected. The result evaluation proves that ELM can be successfully applied in time-series RS imagery processing. Moreover, our comparison experiments show that ELM outperforms SVM and BP network in terms of training time and testing accuracy, when applied in RS imagery classification.

7 Acknowledgement

This work is funded in part by China 973 subprograms NO. 2003CB316906 and NO. 2003CB317006; China NSF programs NO. NSFC60503018, No. 60525202, and No. 60533040; and National Key Technology RD Program No. 2006BAH02A01.

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] 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.
- [6] C. Clifton, “Change detection in overhead imagery using neural networks,” *Applied Intelligence*, vol. 18, pp. 215–234, March 2003.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.

- [12] 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.
- [13] G.-B. Huang, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," 2006.
- [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] 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.