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

Jiaoyan Chen [a], Huajun Chen [a] , Cong Fang [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 effiency and can't be expanded to parallel computing to deal with large scale 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 RS imagery classifier and the classifier is applied to time-series satellite imageries' analysis. In our experiment, we apply our method to the analyis of rapid land use change in Taihu Lake Area over the past decade due to accelerated urbanization.

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

1 Introduction

Nowdays, the available time-series RS imageris 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 method to be highly efficient and scalable. This paper proposes a novel elm-based land-cover change detection method with faster processing speed, high detection accuracy and less human intervene.

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

Submitted to ELM2012 *21 June 2012*

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 multi-spectral RS imageries are available for applications. By comparing two sets of RS imageries, 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 imagery, and provide an automatical change detection technology, which is available for large scale processing of RS imagery. In summary, time-series RS image-based land-cover change detection method is to identifiy the interesting land-cover changes between "before" time imagery and "after" time imagery through RS imagery digital processing and automatical comparison of land-cover classification mappings.

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 technology 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 directions 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 Landsat7 imagery collected in May of 1993 and 2000. Another typcial 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 applications of RS imagery, land-cover change detection has been a problem for some time. Among all the factors that cause land-cover change detectoin unavailable in real situations, change detection method or algorithm used plays a key role. Traditionally, the main challenge to land-cover change detection methods is the detection accuracy. However, with increasing 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. In the field of RS data processing, increasing data size also brings some new challenges to the traditional methods. To overcome the problems, our new land-cover change detection methods should be design with high efficiency and scapability. On one hand, the new method should improve its training and testing speed with new or improved machine learning algorithm. On the other hand, the new method should be designed to ba as automatic as possible, with least human intervene.

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. On one hand, our ELM-based application successfully analysises the rapid land use change in Taihu Lake Area over the past decade due to accelerated urbanization. Although ELM is a recently proposed and widely applied machine learning method, with the capability of multiclass classification and universal approximation[5][6][7][8], few applications of ELM have been done in the area of RS data processing. Our application firstly and sucessfully proves that ELM can be applied in time-series RS data processing with high accuracy and efficiency. On the other hand, the land-cover change detection results are evaluated and compared to the current methods, which use the classifier of SVM and BP network. Through the comparison, we prove ELM network's advantages of fast training speed and high generalisation performance in the area of RS imagery classification. What's more, through the analysis of our method and experiment results, we can find that our method detects the land-cover change totally automatically with least human intervene, which enables the method to be expanded to parallel computing. In conclusion, ELM is firestly applied to time-series RS data intelligent processing, and with its fast speed of training, high generalisation performance and least manual intervene, 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 be explored. Both traditional methods of post-classification and cross-correlation determine land-cover change through the comparison of "before" time and "after" time classifiction maps. Although both methods are more automatic than the others like image differencing, the unsupervised classifiers used result in low accuracy, which greatly limits their actual availability, while the traditional supervised classifiers used lead to more human intervene. 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 are 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 occured 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. But when applied in the field of RS data classification, the traditional supervised classifiers, like SVM and BP network, still have disadvantages of low training speed, low generalisation performance and two much huamn intervene. In comparison, unsupervised approaches require fewer manual intervenes[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 their application

2.2 ELM for Remote Sensing Application

Presently, little work of applying ELM to RS imagery processing has been done. Mahesh Pal[14] did the land cover supervised classification experiment with ELM using remote sensing imagery. However, only the classification accuracy and computational cost of ELM are tested simply and compared to BP network and no application of ELM is designed. Wu Jun[15] indicates that ELM can be naturally used for training the positive and negative fuzzy rule system quickly for image classification. Although ELM is applied to train the network for RS image classification, no time series images processing work are tried, not along large scale RS iamge processing method.

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 high processing efficiency and least human intervene, compared to the traditional neural network based methods.

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 with highest generalisation performance, through adjusting the hidden node number and activation function, using preprocessed training and testing samples. While the detection unit firstly calculates both "before" time and "after" time land-cover classification mappings with the trained network and input time-series RS imageries, and then works out the change detection statistics and change maps by means of comparing the land-cover classification mappings.

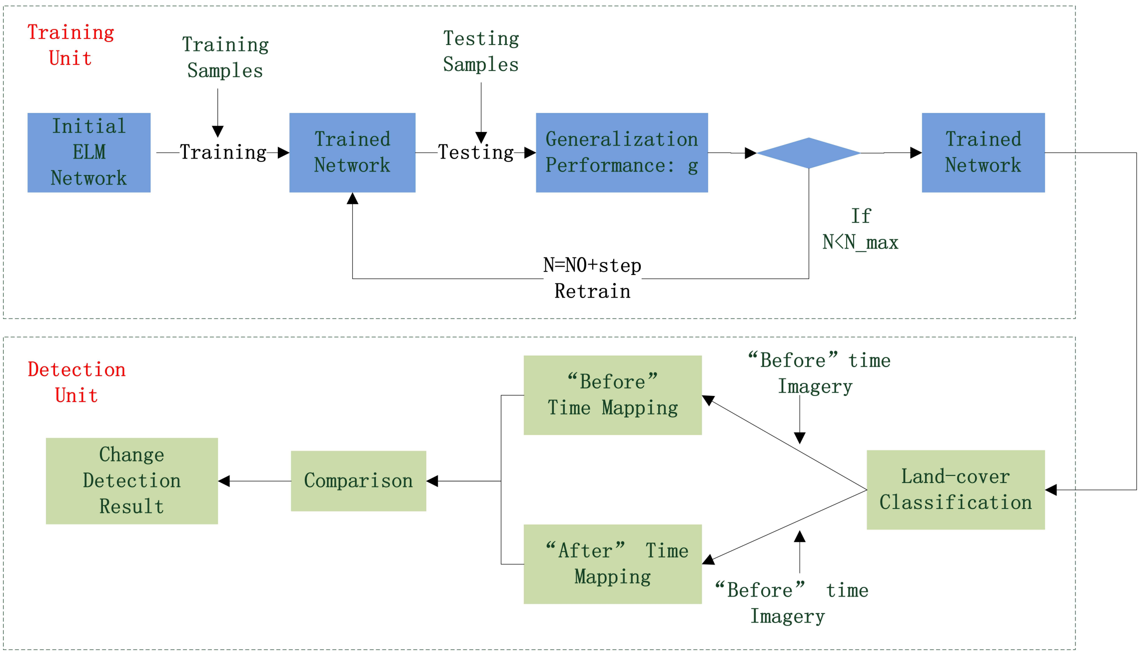


Figure 1. ELM-based Land-cover Change Detection Method

3.1.1 Land-cover Categories and Samples

According to different terrains area and resolution of RS imageries, the land-covery can be classified into specific categories. In fact, in order to assure the precision, statistical property and representative of the samples, training samples of each category should be selected through visual interpretation and field survey[16]. In our application, according to the investigation of Taihu Lake Area's terrains and the landsat imagery' resolution, we classify the land-cover into five types as described in Table 1.

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

To represent the categories of the land cover, every type of the result is assigned to an integer, ranging from 1 to 5. As we use multi-spectral RS images, every pixel is multi-dimensional, and thus every sample can be described:

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

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

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. In our application and experiments, sample is the pixel values of all the bands and is labeled manually in the RS imagery. In order to assure the network's generalisation performance for the application in detection unit, both "before" time and "after" sample imageries should contribute to the samples used in the training unit for network building.

3.1.2 Incremental Training Unit

ELM network with too few hidden nodes has no high classification capability, while too many hidden nodes will cause over-fitting, leading to low testing accuracy and high testing time. In method's training unit, an incremental algorithm with network continusous training and 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 $\chi = \{(\mathbf{x}_i, \mathbf{t}_i)|\mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in R^m, i = \{1, \cdots, N\}\}$, a testing samples set $\psi = \{(\mathbf{x}_i, \mathbf{t}_i)|\mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in R^m, i = \{1, \cdots, \widehat{N}\}\}$, 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 χ

c) calculate the average testing accuracy E of the *num* trained networks with the testing samples set ψ

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

d) $E_{highest} \leftarrow E$

e) $M_{highest} \leftarrow M$

end if

end while

3.1.3 Detection Unit with Land-cover Classification Mappings Comparison

In the detection unit of our method, the "before" time land-cover classification mapping and "after" time land-cover classification 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 imagery is classified into five categories, with computing every pixel's land-cover types. Known every pixel's land-cover type, we can draw the land-cover classification mapping, which visually display the coverage of the land surface. Moreover, land-cover classification pie charts are computed to accurately and visually compare the coverage of every land-cover category.

In the process of land-cover classification mappings comparsion, the trends of every land-covery category and every location's land-covery transition can be calcuated. In our application, as urbanization information is the main focus of decision makers, the expansion map of urban area is drawn.

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[17][18], and thus to achieve better generalization performance.

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

That means

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

where

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

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

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, β_i, through the following formula:

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

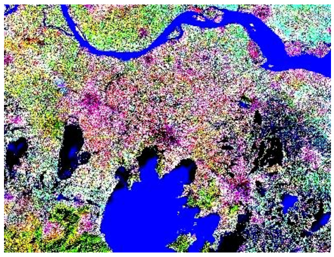
where $H^\dagger$ is the Moore-Penrose generalized inverse of matrix $\mathbf{H}$.

4 Evaluation

4.1 Incremental ELM Training

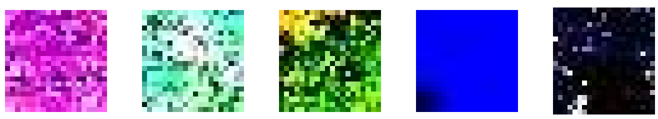
In our experiments, we firstly use our incremental ELM training algorithm to choose the proper hidden node number and activation function, and thus building the ELM network with high generalisation performance.

Firstly, we prepare our training and testing samples in the following way. The pair of time-series imageries, both token by LANDSAT-5 satellite with TM sensor mode, cover Taihu Lake area in east China. The "before" time imagery is token on Dec. 25, 2001, while the "after" time imagery is token on Feb. 28, 2008, as shown in Figure 2(a)(b). Through the investigation of Taihu Lake area's terrain, we manually label five 20×20 squares of different land-cover types on both imageries, as shown in Figure 2(c). From the labeled suquares, we collect a training samples set χ as well as a testing samples set ψ, both of which contain 3,000 pixels. In detail, every pixel can be represented as a four elements tuple, the first three elements of which are the RGB values, while the last element of which is the integer that represents the land-cover group.



(a) "Before" Time

(b) "After" Time



(c) Samples of Different Land-cover Types

Figure 2. Original Time-series RS Imageries

In our network building process, we incrementally train and test the ELM network with the hidden node number increasing from 10 to 400 three by three. Moreover, for specific hidden node number and activation function, the network is trained and tested 3 times and the average testing accuracy are calculated to assure the accuracy. Figure 3 shows the trends of testing accuracy with increasing hidden node number and different activative function. From the figure we can firstly conclude that all the activation functions, except for hardlim, performs similarly with different hidden node number. Secondly, the average testing accuracy of the network improves greatly before the hidden node number reaches 100, but are stabilized when the hidden node number exceeds about 150. Finally, the network with 350 to 400 hidden nodes and the activation function of tribas can achieve the testing accuracy of more than 98%, which will be applied in the detection unit.

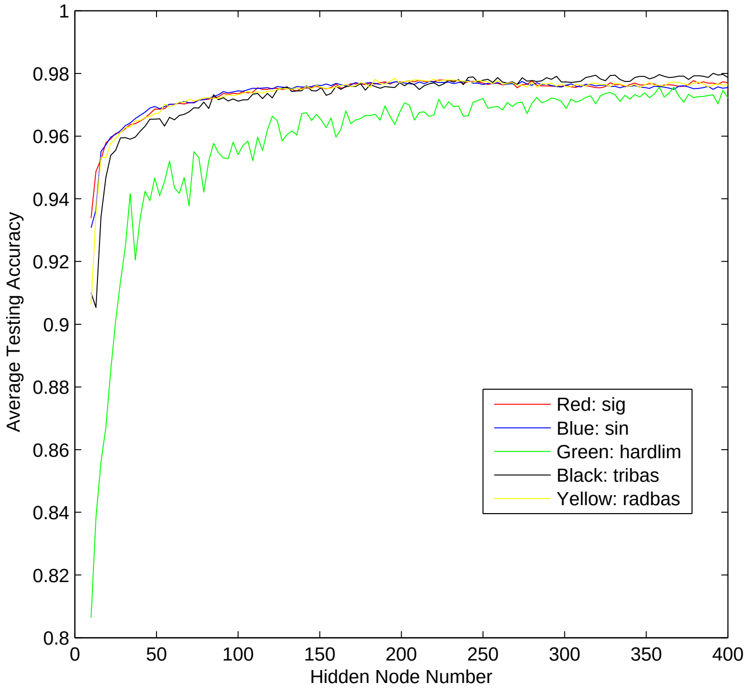


Figure 3. Incremental ELM Training Result.

4.2 Land-cover Classification Performance of ELM

The most outstanding advantages of ELM network that our land-cover change detection method based on are its fast learning spead and high generalisation performance. This experiment is to prove ELM network's advantages in RS imagery classification. Moreover, comparison between ELM and SVM, BP network is done to highlight ELM network's advantages.

One factor that we will compare is their training time, when specific generalisation performance is ensured. To start with, we set the hidden node number of ELM network to 100 and use sin as the activation function, which can assure testing accuracy to be more than 95% when 3,000 samples are used to train. Using the same training samples, we then adjust both the parameters of SVM and BP networks to assure their testing accuracy ranging between 95% and 96%. Finally, we compare the average training time of all the three methods using increasing training data size In this step, for specific size of training data size, training is performed several times and the average training time is calculated. The training time comparison result can be seen in Figure 4.

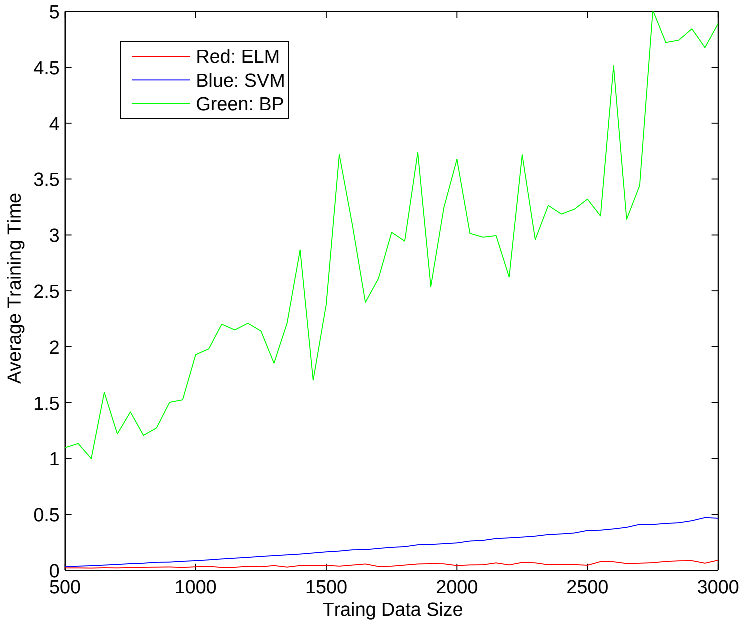


Figure 4. Average Training Time Comparision.

From the figure we can argue that ELM always costs the least training time. With the increasing training data size, the training time of BP network increases greatly and are unstable. From Figure 4 and Table 2, 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.

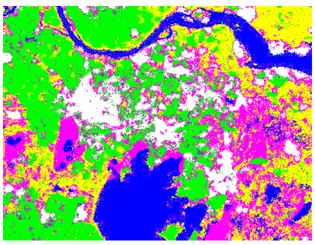
	ELM			BP			SVM		
Data Size	Training Time	Testing Time	Testing Acc	Training Time	Testing Time	Testing Acc	Training Time	Testing Time	Testing Acc
500	0.0180	0.0060	97.3200%	0.0312	0.0054	93.4000%	1.4743	0.0082	95.8000%
750	0.0270	0.0060	97.1333%	0.0582	0.0107	93.6000%	1.3774	0.0082	92.7333%
1000	0.0320	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.

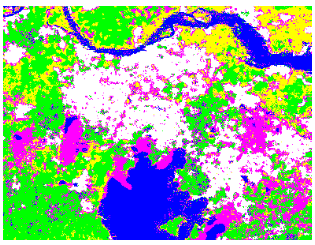
The other factor that we will compare is their generalisation performance under various scale of training and testing data sets. From Table 2, we can firstly conclude that the testing accuracy of ELM is always higher than that of BP and SVM networks. Second, ELM network can soon reaches high generalization performance, even with small training data set, while the other two networks need more training sampels to assure high generalizatio performance. Totally, we can argue that the ELM network has better generalization performance than BP and SVM in RS imagery classification, and it needs much less samples to train to get high generalization performance.

4.3 Change Detection Result Analysis

In this experiment, we firstly category the "before" time and "after" time imageries into five land-cover types and then the land-cover mappings are drawn as shown in Figure 5. Compared to the original imageries in Figure 2(a)(b), we can visually judge that both mappings truly reflect the actual ground surface of Taihu Lake area. Through comparison of the land-cover classification mappings, the change trend of every pixel can be drawn. The cyan area in Figure 6 shows the expansions of urban. From the change detection result, that fact that a lot of urban area has been expanded around the original urban area can be found intuitively.



(a) "Before" Time



(b) "After" Time

Figure 5. RS Imagery Land-cover Mapping. Whrite area represents urban, blue area represents water, yellow area represents arable land, green area represents vegetation, manenta area represents wetlands.

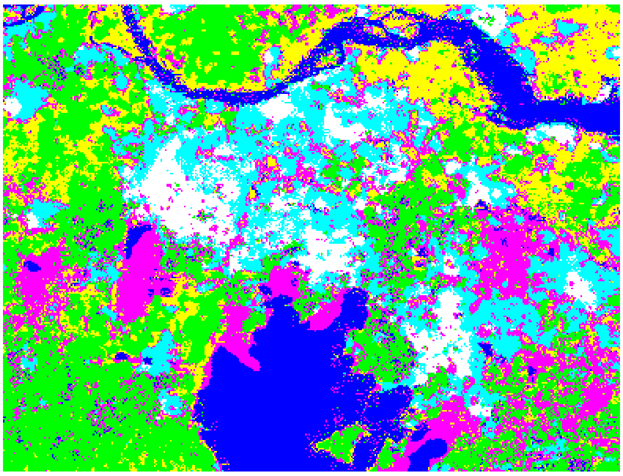
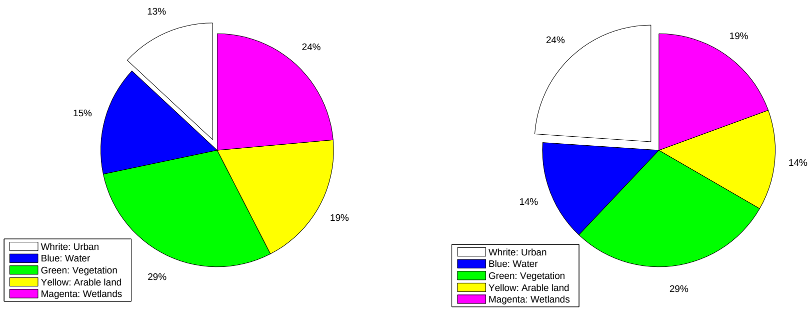


Figure 6. Urban Area Change Detection. The cyan are the new urban area that appears in year 2008, compared to the land-cover of year 2001.

Except for the change mapping of every land-cover type, such as urban expansions, we provide the proportion change of every type of land-cover in detail, as shown in Figure 7. From Figure 7, we can find that the urban area is rapidly expanding, from 30% in year 2001 to 24% in year 2008, as the fast urbanization of Taihu Lake area. In contrast, the area of water, arable land and wetlands is shrinking. What's more, Firgure 8 gives the comparison of every land-cover type between the year 2001 and the year 2008.



(a) "Before" Time (b) "After" Time

Figure 7. Land-cover Classification Result

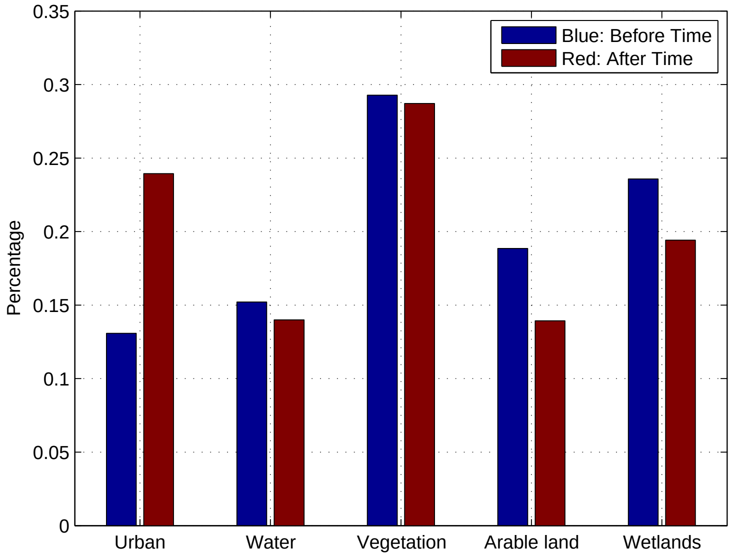


Figure 8. Change Detection Result of Five Different Land-cover Area .

5 Discussion

This paper contributes the first application of ELM to time-series RS imagery processing. We use ELM neural network to classify time-series RS imageries and accuratelly detect the land-cover change of Taihu Lake Area from the year 2001 to the year 2008. From the application result analysis, we can conclude that ELM can be sucessfully applied in the field of time-series RS imagery processing. Moreover, through our network training and testing experiments, we prove the ELM's advantages compared to traditional SVM and BP network in the field of RS imagery classification.

We design our ELM-based land-cover change detection method as shown in Figure 1. In the training unit, we use the simple incremental ELM training algorithm to find the proper hidden node number and suitable activation function, in order to achieve the highest generalization performance. Through the analysis of training result shown in Figure 3, we can find that the testing accuracy of our trained network exceeds 97% when more than 200 hidden nodes are used and any activation functions except hardlim is chosen. Thanks to the high generalization performance of ELM in RS imagery classification, once the neural network is trained, it can be widely applied to other RS imageries, which are token and generated by the same method and from the same original satellite data. And all the processes of the detection unit, from land-cover classification operation to land-cover mappings comparison, needs no human intervene. This characteristic makes it possible to be expanded to parallel computing or distributed computing on cloud infrastructure.

One application of our ELM-based land-cover change detection method is to analysis land use change of Taihu Lake Area from the year 2001 to the year 2008. The time-series land-cover mappings can be seen in Figure 5, while the urban area change trend mapping are shown in Figure 6. When compared to the original time-series RS imageries in Figure 2, we can visually judge that the classification mappings can accurately represent the distribution of five land-cover types. When we visually compare the before time and after time land-cover mappings, we can clearly find the expansion of urban area, which is totally the same result as shown in Figure 6. Moreover, our method can also provide the statistical results, as shown in Figure 7 and Figure 8, which provide more accurate decision making evidences. In summary, we can judge that the ELM-based method can accurately detect the change of land-cover.

In order to prove ELM's advantages of fast training speed and high generalisation performance, compared to traditional SVM and BP network, in the field of RS imagery classification, this paper constructs two experiments, whose results are shown in Figure 4 and Table 2. From Figure 4, we can directly conclude that the training time of ELM network, is much less than SVM and BP network. Moreover, average training time of BP network are unstable and increases greatly, while the average training time of ELM increases nearly linearly and the slope is very small. According to the fast speed of training, ELM can perform better than SVM and BP network when applied to large scale RS imagery samples training or online training with acceptable training time. According to Table 2, we can find that ELM's testing accuracy is always higher than SVM and BP network. We can see that when the training and testing data size increases to 3,000, the testing accuracy of ELM reaches more than 97%, while SVM and BP network only reach about 96% and 95%. The high generalisation performance makes ELM a better choice than SVM and BP network to RS imagery classification.

6 Future Work

Presently we have applied our ELM-based land-cover change detection method to the analysis of land use of Taihu Lake Area in the year 2001 and 2008. Our application and comparison experiments prove that ELM can be successfully applied to time-series RS data processing. However, more research topics of large scale time-series RS imagery processing with ELM have been proposed in our research group. In fact, the ELM-based land-cover change detection method is designed with high scability and are prepared to be expanded to parallel computing on cloud infrastructure. Our further work is to improve the processing efficiency of large scale RS imageries with the computing and storage capability of hdfs and mapreduce based cloud infrastructure. Nowdays, cloud based distributed computing has been one of the most popular solutions to deal with rapid growing data. To process fast increasing multi-spectral high resolution RS imagery, not only the efficient ELM-based method, but also the distributed computing, should be tried. Our furture work will firstly prove our ELM-based land-cover change detection method's high scability on cloud infrastructure. Moreover, based on ELM's capabilty of multi-class classification and regression, RS imagery applications, such as object identification, rainfall estimation, cloud cover evaluation and so on, will be developed. Finally, we hope to pay more attention to ELM's capability on online learning, such as OS-ELM, and we consider it a good resolution to the training of large scale RS imageries. Namely, with the help of cloud computing and storage, distributed training of ELM network based on large scale and online training samples is our following research focus.

7 Conclusions

Time-series RS imagery processing, including land-cover change detection, plays a more and more important role in various fields, but has been a problem for some time for various reasons, such as the rapid growth of satellite data size. This paper contributes the first application of ELM network to time-series RS imagery processing, and proposes an ELM-based land-cover change detection method with least human intervene and high processing efficincy. Thanks to ELM network's fast training speed, high generalisation performance and least human intervene, the land-cover change detection method is designed to deal with large-scale RS imageries. As a real application, the land-cover change of Taihu Lake Area, such as the expansion of urban area, from the year 2001 to the year 2008 is accurately detected using our method. What's more, our comparison experiments between ELM and SVM, BP network show that ELM do has better training and predicting performance in the classification of RS imagery.

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] 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.

[17] G.-B. Huang, "Extreme learning machine: a new learning scheme of feedforward neural networks," 2004.

[18] 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.