A Neural Network Approach for Land-Cover Change Detection in Multi-Temporal Multispectral Remote-Sensing Imagery

VICTOR-EMIL NEAGOE*, MIHAI NEGHINA*, and MIHAI DATCU**
Depart. Electronics, Telecommunications & Information Technology
*Polytechnic University of Bucharest
Splaiul Independentei No. 313, Sector 6, Bucharest, ROMANIA
ROMANIA
victoremil@gmail.com, mihai.neghina@gmail.com
**Remote Sensing Technology Institute (IMF)
German Aerospace Center, Oberpfaffenhofen,
GERMANY
mihai.datcu@dlr.de

Abstract: - This paper presents a neural network approach for land-cover change detection in remote-sensing imagery. One has considered the following supervised neural classifiers: Multilayer Perceptron (MLP), Radial Basis Function Neural Network (RBF), and Supervised Self Organizing Map (SOM). For comparison, we have chosen two well-known statistical classifiers (Bayes and Nearest Neighbour (NN)). The proposed model of change detection in multispectral satellite images has two main processing stages: (a) feature selection (using one of the three techniques: concatenation algorithm (CON), the algorithm based on absolute differences of pixels (ADIP), and the algorithm based on difference of reflectance ratios (DIRR)); (b) classification, using one of the above mentioned classifiers. The considered techniques are evaluated using a LANDSAT 7 ETM+ multi-temporal image, corresponding to a set of two images of the same zone (400 x 400 pixels) in the region Markaryd, Sweden taken in 2002 and 2006. One has the change reference map; we have used 2000 pixels for training and the rest of 158 000 pixels for test. The best experimental result leads to the change detection rate of 88.24 % for the test lot, proving the advantage of neural network models over the statistical ones.

Key-Words: - land-cover change detection, neural network classifier, multispectral multi-temporal images, Supervised Self Organizing Map (SOM), Multilayer Perceptron (MLP), Radial Basis Function (RBF) neural network.

1 Introduction

Automatic change detection is one of the most interesting problems of image processing, having a key function in many practical application areas [1], [3], [5], [8], [10], and [11]. The increasing interest in environmental protection and control has led this topic to have a great note in the remote sensing community. The European Environment Agency (EEA) has initiated computer-assisted image interpretation of earth observation satellite images to map the whole European territory into standard CORINE Land Cover categories [7]. Besides providing the status of the land cover at or around specific times, EEA also compiled vector databases for changes between those specific times. It has been the case with the CLC changes 2000-2006 database [12]. Land-cover change identification concerns the analysis of two registered remote sensed multispectral images acquired in the same geographical area at two different times. This is very useful in many applications, like land use change analysis, study on shifting cultivation, monitoring of pollution, assessment of burned areas, assessment of deforestation, and so on. Following the growing need and the increased data availability, numerous methods for the detection of changes have been developed over recent years [1], [3], [5], [8], [10], and [11]. Last years, Artificial Neural Networks (ANN) have emerged as an important tool for the classification of remote sensing images [2], [6], [9]. Change detection has emerged as one of the relatively new application areas of ANN [1], [8]. Several general advantages of applying neural networks for classification of satellite imagery are the following [8]: (i) ANN are data driven and self-adaptive since they can adjust themselves to the data without any explicit functional specification of the underlying physical model; (ii) ANN can provide universal functional approximations; (iii) the neural classifiers do not require initial hypotheses on the data distribution and they are able to learn non-linear and discontinuous input data.

In this paper we present and evaluate an approach of applying supervised neural network classifiers for change detection in multi-temporal and multispectral satellite imagery. We have evaluated the neural

ISBN: 978-1-61804-027-5

classifier change detection performances (Multilayer Perceptron (MLP), Radial Basis Function Neural network (RBF), and supervised Self-Organized Map (SOM)) versus statistical classifier performances (Bayes and_Nearest Neighbour (NN)). We have experimented the above mentioned algorithms using a multi-temporal LANSAT 7 ETM+ satellite image corresponding to a 400 x 400 pixel zone of the region Markaryd, Sweden.

2 Algorithm Description

The proposed processing cascade for change detection in multi-temporal and multispectral remote-sensing images consists of two main processing steps :

(a) feature selection using one of the following processing step:

(a1) concatenation of multispectral pixels (CON)

(a2) absolute differences of pixels (ADIP)

(a3) absolute difference of reflectance ratios (DIRR)

(b) supervised classification using an algorithm belonging to one of the following two categories:

(b1) neural network classifiers, consisting of one of the following three classifiers: Multilayer Perceptron (MLP); Radial Basis Function neural network (RBF); supervised Self Organizing Map (SOM)

(b2) statistical classifiers, consisting of one of the following two classifiers: Bayes and Nearest Neighbour (NN).

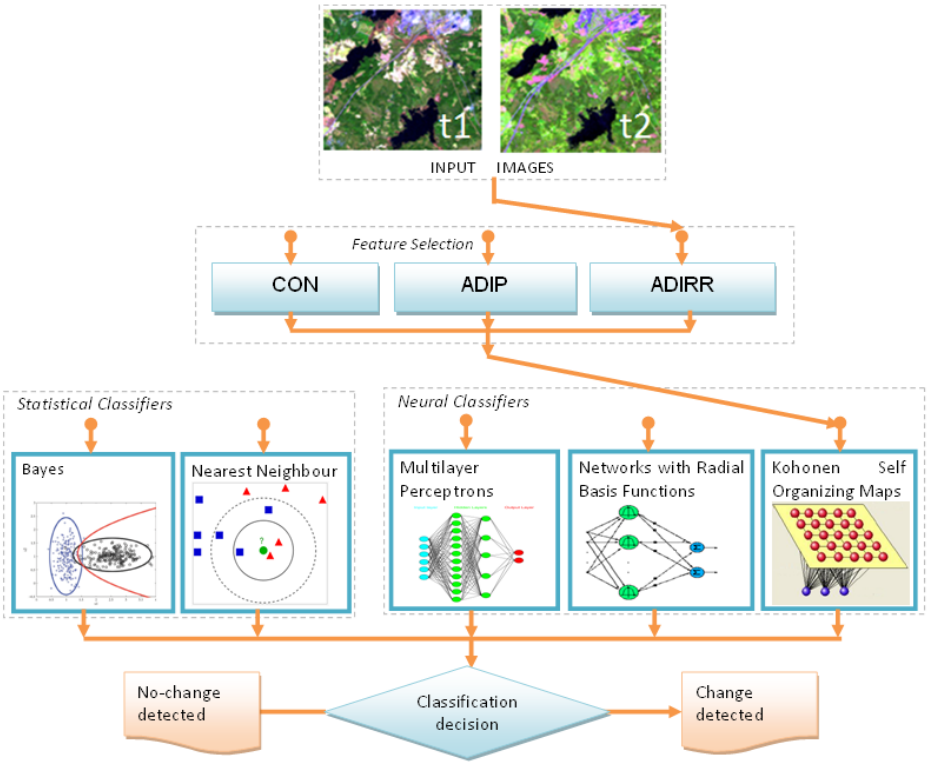


Fig. 1. Flowchart of the proposed land-cover change detection processing cascade.

2.1 Feature Selection

Performing change detection requires a robust feature selection technique. Let us consider two multispectral images X_1 and X_2, acquired in the same geographical area at two different times t_1 and t_2, co-registered and radiometric and calibrated. Each multispectral pixel is represented as an n-dimensional vector, where is the number of bands. One chooses one of the following three feature selection techniques further described in this chapter.

2.1.1 Concatenation of Corresponding Multispectral Pixels (CON)

For every pair of corresponding multispectral pixels belonging to the two multispectral images

$A^T = [a_1 \dots a_n]^T$ and $B^T = [b_1 \dots b_n]^T$,

belonging to the two corresponding multispectral images the concatenation result is

$V = [A^T, B^T]^T = [a_1 \dots a_n, b_1 \dots b_n]^T$.

This result is then transferred to the next module.

2.1.2 Absolute Difference between Corresponding Pixels (ADIP)

For every pair of corresponding pixels

$A^T = [a_1 \dots a_n]^T$ and $B^T = [b_1 \dots b_n]^T$,

belonging to the two multispectral images, the result

of the absolute difference of corresponding pixels is

$V = [\ |a_1-b_1|, |a_2-b_2|, \ldots |a_n-b_n|\ ]^T$

This vector is then transferred to the next module.

2.1.3 Absolute Difference of Reflectance Ratios (ADIRR)

We chose to work with reflectance ratios in order to reduce the effects of the different scene illuminations. For every pair of corresponding pixels A^T and B^T one computes the reflectance ratios [x]

$R_A = [a_1/a_2, a_1/a_3, .. a_1/a_n, a_2/a_3, \ldots, a_{n-1}/a_n\ ]^T$

$R_B = [b_1/b_2, b_1/b_3, ... b_1/b_n, b_2/b_3, \ldots, b_{n-1}/b_n]^T$

The number of elements of any of the vectors R_A, R_B *is* $n\ (n-1)/2$, where n is the number of bands (dimension of the multispectral pixel).

Then, for each pixel, one computes the absolute difference between the reflectance ratios R_A, R_B as one does in 2.1.2.

2.2 Neural versus Statistical Classification

One further considers change detection as a problem of binary classification: change/no change. We have evaluated the performances of neural classifiers versus statistical ones (Fig.1).

2.2.1 Neural Classifiers

A. Multilayer Perceptron (MLP).

MLP is the classical model of feedforward backpropagation neural network [2], [9]. For change detection in multispectral images, the input/output configuration for MLP is one input node for each feature of the input vector (corresponding to one of the three feature selection techniques) and one output node for each desired class label. Namely, our selected MLP configuration had *2n* input neurons for CON feature selection technique, *n* input neurons for ADIP variant, and *n (n-1)/2* neurons for ADIRR case, where n is the number of selected bands. One uses two output neurons, corresponding to *change/no change* labels. The number and sizes of hidden layers are not imposed.

B. Radial Basis Function Neural Network (RBF).

A RBF network [2] consists of an input layer of *m* virtual neurons that only distribute the information to the intermediate layer, an intermediate layer consisting of L neurons that implement the radial basis activation function (generally, a Gaussian function) as well as an output layer of N neurons (for our case, N=2), that performs a weighted sum of the outputs of the previous layer. Due to their non-linear approximation properties, RBF networks are able to model complex mappings, which MLP networks can only model by means of multiple intermediary layers. For the considered change detection application, the input/output configuration is the same as for MLP.

C. Supervised Self-Organizing Map (SOM).

SOM defines a mapping from the input n-dimensional input data space onto a regular one or two-dimensional (generally, m-dimensional, with $m < n$) array of M nodes [4]. With every node m, a weight n-dimensional vector is associated. An input vector $x \in R^n$ is compared with all the weight vectors and the best match is defined as "response": the input is thus mapped onto this location. During the training phase, the neurons (their corresponding weight vectors) become specifically tuned to various classes of patterns through a competitive, unsupervised or self-organizing learning. The spatial location of a neuron in the network (given by its co-ordinates) corresponds to a particular input vector pattern. We have used SOM as a *supervised* system. It requires that after *unsupervised* training to perform a stage of *calibration*. We have used the calibration SOM procedure described in [4].

2.2.2 Statistical Classifiers

D. Bayes Classifier

We have assumed that the conditional probability density functions $p(x|\omega_1)$ and $p(x|\omega_2)$ are normal of means μ_1 and μ_2, and covariance matrices Σ_1 and Σ_2, and the a priori class probabilities are $P(\omega_1)$, $P(\omega_2)$. Then, the Bayes decision rule becomes

$$(X-\mu_1)^T \times \sum_1^{-1} \times (X-\mu_1) - (X-\mu_2)^T \times \sum_2^{-1} \times (X-\mu_2) + \ln\frac{\det\Sigma_1}{\det\Sigma_2} \lessgtr 2\ln\frac{P(\omega_1)}{P(\omega_2)} \Rightarrow X \in \begin{cases}\omega_1\\ \omega_2\end{cases}$$

The parameters $P(\omega_1)$, $P(\omega_2)$, μ_1, μ_2, Σ_1 and Σ_2 are computed from the labeled training set.

E. Nearest Neighbour(NN)

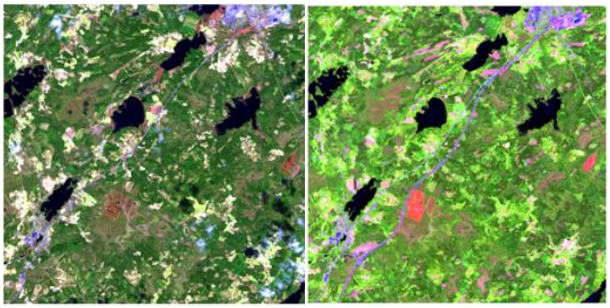
NN classification is one of the most fundamental and simple classification methods and should be one of the first choices for a classification task when there is little or no prior knowledge about the distribution of the data. The output class is given by the closest neighbour of the input vector belonging to the labeled training set.

3 Experimental Results

3.1 Satellite Image Database: Landsat 7 ETM+ Data Set over Markaryd, Sweden

The data used for the experiments are selections from the multi-temporal LANDSAT 7 ETM+ multi-temporal image consisting of the set :

LE71940212002095EDC00 (Acquisition date: 5 April 2002) and LE71940212006186ASN00 (Acquisition date: 5 Jul 2006), representing the same 400 x 400 pixel zone, from the region Markaryd, Sweden (Fig. 2). We have selected six bands: 1, 2, 3, 4, 5, 7, having all the same resolution of 30m.



(a) (b)

Fig. 2. LANDSAT ETM+ image sequence displayed with 3 bands (Red=Band 5, Green=Band 4, Blue=Band 3). (a) 5 Apr 2002. (b) 5 Jul 2006.

The selected region contains significant changes throughout the image, most notably a highway built along an already existing road, development of the urban area of the top-right corner city, as well as human-made buildings along the road and in some parts of the forest. According to the CORINE Land Cover Changes 2000–2006 [12], the European remote-sensing aim is to indicate land-cover changes that are larger than 5 ha, wider than 100m, and are detectable from satellite images.

For the considered multi-temporal image (Fig.2), there exist 5532 pixels of change (~3.45%) and 154468 pixels of non-change (~96.55%). All the 160000 pixels of the considered two-image sequence have a binary label *change/ (no change)*, according to CLC reference map database.

From the whole data set, we have selected 2000 pixels for training (1000 "change" + 1000 "no-change"), representing 1.25% din total, and the rest of 15800 pixels for test (98.75%).

3.2 Parameters for Performance Evaluation

In order to evaluate the performances of the proposed change detection algorithm, we have chosen the following parameters:

Correct Detection Rate [%]:

$$CDR = \frac{True\ Positives}{True\ Positives + False\ Negatives} \cdot 100 \quad [\%]$$

Correct Rejection Rate [%]:

$$CRR = \frac{True\ Negatives}{True\ Negatives + False\ Positives} \cdot 100 \quad [\%]$$

False Positive Rate:

$$FPR = \frac{False\ Positives}{True\ Negatives + False\ Positives} = 100 - CRR$$

Miss Rate:

$$MR = \frac{False\ Negatives}{True\ Positives + False\ Negatives}$$

Approximation of Total Success Rate [%]:

$$TSR = (CDR + CRR)/2,$$

where:

- *TP = true positives* = changes correctly detected
- *TN = true negatives* = no-changes detected correctly
- *FP = false positives* = no-changes detected as changes
- *FN = false negatives*=changes detected as non-changes

3.3 Experimental Results

The experimental results (recognition score for the test lot) obtained of applying the considered techniques of change detection for the above mentioned dataset are given in Tables 1-3 as well as in Figs. 3-8. An example of neural change detection map by comparison with CORINE Change Land-Cover reference map is given in Fig. 9.

Table 1. Change detection performances as a function of classifier type using multispectral pixel concatenation (CON) for feature selection.

Algorithm	Algorithm parameters							
	Network type	Dim.	Training	CDR [%]	CRR [%]	MR [%]	FPR [%]	(CDR + CRR)/2 [%]
Bayes	~~~	~~~	~~~	77.98	91.38	22.02	8.62	84.68
NN	~~~	~~~	~~~	84.07	85.81	15.93	14.19	84.94
SOM	rectang	50x50	99epochs	84.18	89.20	15.82	10.80	86.69
RBF	Spread = 49.5			85.44	81.77	14.56	18.23	83.60
MLP	1 hidden layer, 25 neurons			86.65	89.83	13.35	10.17	88.24

Table 2. Change detection performances as a function of classifier type using absolute difference of the corresponding pixels (ADIP) as a feature selection.

Algorithm	Algorithm parameters							
	Network type	Dim.	Training	CDR [%]	CRR [%]	MR [%]	FPR [%]	(CDR + CRR)/2 [%]
Bayes	~~~	~~~	~~~	74.47	90.10	25.53	9.90	82.29
NN	~~~	~~~	~~~	78.29	81.30	21.71	18.70	79.80
SOM	rectang.	50x50	99epochs	78.93	87.57	21.07	12.43	83.25
RBF	Spread = 473.0			77.85	88.16	22.15	11.84	83.00
MLP	1 hidden layer, 15 neurons			82.37	87.05	17.63	12.95	84.71

Table 3. Change detection performances as a function of classifier type using absolute difference of reflectance ratios (ADIRR) for feature selection.

Algorithm	Algorithm parameters							
	Network type	Dim.	Training	CDR [%]	CRR [%]	MR [%]	FPR [%]	(CDR + CRR)/2 [%]
Bayes	~~~	~~~	~~~	75.60	91.60	24.40	8.40	83.60
NN	~~~	~~~	~~~	77.91	80.60	22.09	19.40	79.26
SOM	hexag	5x10	30epochs	78.93	84.09	21.07	15.91	81.51
RBF	Spread = 34.5			82.30	89.56	17.70	10.44	85.93
MLP	2 hidden layers, 10 and 50 neurons			82.15	89.50	17.85	10.50	85.82

ISBN: 978-1-61804-027-5

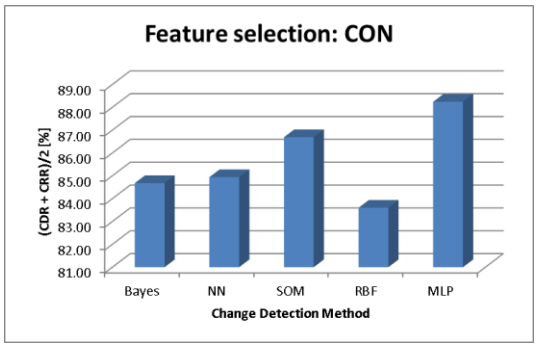


Fig. 3. Total Success Rate (TSR) as a function of classifier type using CON for feature selection.

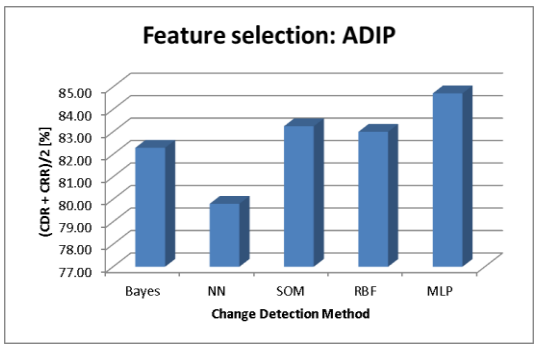


Fig. 4. Total Success Rate (TSR) as a function of classifier type using ADIP for feature selection.

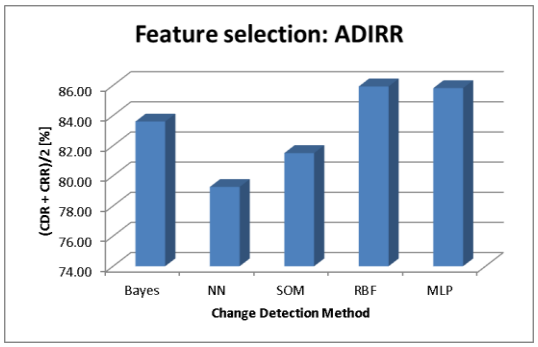


Fig. 5. Total Success Rate (TSR) as a function of classifier type using ADIRR for feature selection.

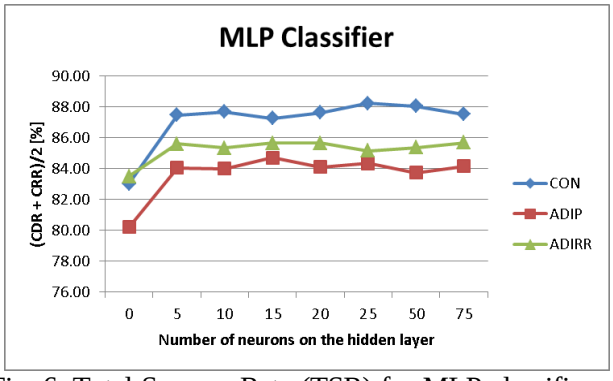


Fig. 6. Total Success Rate (TSR) for MLP classifier as a function of the number of neurons of the hidden layer.

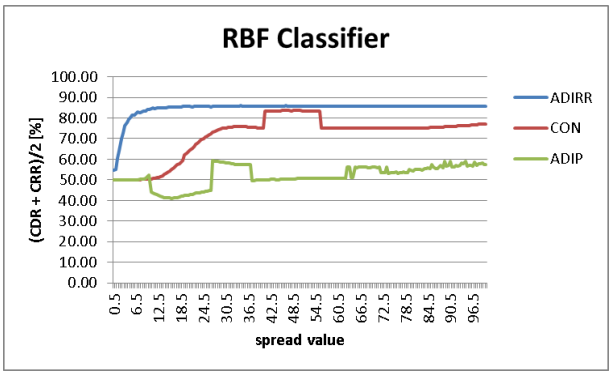


Fig. 7. Total Success Rate (TSR) for RBF classifier as a function of the spread parameter values.

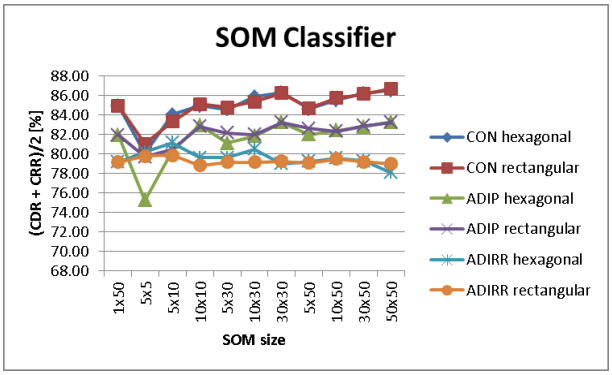


Fig. 8. Total Success Rate (TSR) for supervised SOM classifier as a function of SOM size and architecture (99 training epochs).

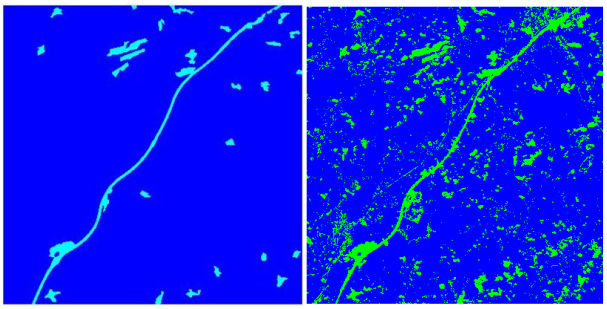


Fig. 9. Example of change detection map. (a) Reference change map (CORINE Change Land-Cover reference). (b) Change detection map using the cascade CON-MLP (1 hidden layer with 25 neurons; total success rate =88.24% on the test set).

4 Concluding Remarks

1) This paper presents a neural network approach for land-cover change detection in remote-sensing imagery. One has considered the following neural classifiers: Multilayer Perceptron (MLP), Radial Basis Function Neural Network (RBF), and Supervised Self

ISBN: 978-1-61804-027-5

Organizing Map (SOM). For comparison, we have also considered two well-known statistical classifiers (Bayes and Nearest Neighbour (NN)).

2) For feature selection we have chosen one of the three techniques: concatenation algorithm (CON), the algorithm based on absolute pixel differences (ADIP), and the algorithm based on difference of reflectance ratios (DIRR).
3) Globally, for all the feature selection techniques, one can deduce than the performances obtained by neural techniques are better than ones obtained by statistical ones.
4) Using feature selection by CON (Table 1, Fig. 3), the MLP classifier yields a 3.3% increase in performance (TSR) by comparison to NN, and even more compared to Bayes. The best SOM classifier variant has obtained also about 2% better performances over statistical ones.
5) Applying feature selection by ADIP (Table 2, Fig. 4), the MLP classifier yields a 2.4% increase with spread 34.5 in performance as compared to Bayes, and 5% more by comparison to NN. Also, SOM and RBF lead to about 1% better results than Bayes and more than 3% better by comparison to NN.
6) Using feature selection according to DIRR (Table 3, Fig. 5), the best results are obtained by RBF (for spread parameter 34.5). This means a 2.3% increase in performance as compared to Bayes, and about 6.6% more by comparison to NN. MLP has also obtained a 2.2% better performance than Bayes and 6.5% better one than NN.
7) For the MLP classifier experiments (Fig. 6), the influence of number of hidden layer neurons is rather small (maximum of 1% variation of the TSR for all the feature selection techniques).
8) The influence of the spread parameter over RBF classifier performance (Fig. 7) is a function of the feature selection technique. For DIRR, the TSR is almost flat for spread greater than 15.0, while for ADIP by increasing spread parameter we have obtained a steady improvement of TSR until the spread reaches values around 450.
9) In several cases, the SOM size has influence on the detection results (Fig. 8). For CON and ADIP, by increasing SOM size, one obtains better performance.
10) The best performance is obtained using the cascade CON-MLP (TSR=88.24%). The corresponding change detection map is shown in Fig. 9. The results may be further improved using a post-processing stage. One points out that we have used a training set of only 1.25% from the total data set (while the test set has been of 97.85%)!

References:

[1] A. D'Addabbo, G. Satalino, G. Pasquariello, P. Blonda, Three different unsupervised methods for change detection: An application, *Proc. IEEE Internat. Geosci. Remote Sens. Symposium, IGARSS 2004*, Anchorage (Alaska), Vol. 3, 2004, pp. 1980–1983.

[2] M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.

[3] L. Gueguen, C. Shiyong, G. Schwarz, M. Datcu, Multitemporal analysis of multisensor data: Information theoretical approaches, *Proc. Internat. Geosc. Remote Sens. Symposium (IGARSS), 2010*, pp. 2559–2562.

[4] T. Kohonen, *Self-Organizing Maps*, Springer Series in Information Sciences, Berlin, Heidelberg, Springer, 1995.

[5] D. Lu, P. Mausel, E. Brondízio and E. Moran, Change detection techniques, *Internat. J. Remote Sens.*, Vol. 25: No. 12, 2004, pp. 2365–2401.

[6] V. Neagoe and A. Ropot, A New Neural Approach for Pattern Recognition in Space Imagery, in *Harbour Protection Through Data Fusion Technologies, NATO Science for Peace and Security Series-C: Environmental Security*, Springer, 2009, pp. 283-289.

[7] M. Nunes de Lima (ed), *CORINE Land Cover Updating for the Year 2000. IMAGE2000 and CLC2000: Products and Methods*, European Environment Agency, Report EUR 21757, 2005, ISBN 92-894-9862-5.

[8] F. Pacifici, F. del Frate, C. Solimini, W. Emery, An Inovative Neural-Net Method to Detect Temporal Changes in High Resolution Optical Satellite Imagery, *IEEE Trans. Geosci. Remote Sens.*, Vol. 45, No. 9, 2007, pp. 2940-2952.

[9] J. D. Paola, R.A. Schowengerdt, A Detailed Comparison of Backpropagation Neural Network and Maximum-Likelihood Classifiers for Urban Land Use Classification, *IEEE Trans. Geosci. Remote Sens.*, Vol. 33, No. 4, July 1995, pp. 981-996.

[10] B.P. Salmon, J.C. Olivier, W. Kleynhans, K.J. Wessels, F. van den Bergh, The Quest for Automated Land Cover Change Detection using Satellite Time Series Data, in *Proc. of IGARSS (4)'2009*. pp. 244-247.

[11] A. Singh, Digital change detection techniques using remotely-sensed data, *Int. J. Remote Sensing*, Vol. 10, No. 6, 1989, pp. 989-1003.

[12] ***, *CLC2006 technical guidelines*, European Environment Agency, Technical Report No. 17/2007, ISSN 1725–2237, ISBN 978-92-9167-968-3.