Neural Network Techniques for Automated Land-Cover Change Detection in Multispectral Satellite Time Series Imagery

Victor-Emil Neagoe, Mihai Neghina, and Mihai Datcu

***Abstract* —** This paper presents an advanced approach for land-cover change detection in remote-sensing imagery. Firstly, several supervised neural network change detection techniques have been considered and evaluated versus statistical supervised ones.; the chosen neural network models are Multilayer Perceptron (MLP), Radial Basis Function Neural Network (RBF), and Supervised Self Organizing Map (SOM), whereas the applied statistical classifiers are Bayes and Nearest Neighbor (NN). Secondly, we have investigated the following unsupervised change detection techniques: Self-Organizing Map (SOM) (neural clustering), versus K-means (statistical clustering), and Fuzzy C-means (FCM) (fuzzy clustering). The proposed model of change detection in multispectral satellite images has two main processing stages: (a) feature selection (using one of the three techniques: the concatenation of corresponding pixels (CON), the computation of absolute differences between corresponding pixels (ADIP), and the computation of absolute differences between reflectance ratios of corresponding pixels (ADIRR)); (b) classification, using one of the above mentioned supervised or unsupervised models (for the two-class case:"change", "no change"). The considered techniques are evaluated using a Landsat 7 ETM+ multi-temporal image, corresponding to a set of two images of the same area (400 x 400 pixels) in the region Markaryd, Sweden taken in 2002 and 2006. For model evaluation, a change map provided by the European Environmental Agency was taken as reference; we have used 2000 pixels for training and the rest of 158 000 pixels for test. The best experimental result using supervised techniques leads to the total success change detection rate of 88.24 % (CON-MLP) for the test lot, whereas among the unsupervised techniques, using all the pixels, the best result corresponds to a total success change detection rate of 78.22% (ADIP-SOM). The experimental results prove the advantage of the neural network change detection techniques over the statistical and fuzzy ones.

***Keywords*—**land-cover change detection, Multilayer Perceptron (MLP), multispectral multi-temporal images, neural clustering, neural network classifier, Radial Basis Function (RBF) neural network, Self Organizing Map (SOM).

Manuscript submitted September 25, 2011.

V.E. Neagoe is with the Department of Electronics, Telecommunications & Information Technology Dept., Polytechnic University of Bucharest, Romania
(email: victoremil@gmail.com).

M. Neghina is with the Department of Electronics, Telecommunications & Information Technology Dept., Polytechnic University of Bucharest, Romania (email: mihai.neghina@gmail.com).

M. Datcu is with the Remote Sensing Technology Institute (IMF), German Aerospace Center, Oberpfaffenhofen, Germany
(email: mihai.datcu@dlr.de).

I. 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], [6], [11], [13], and [15]. 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 [10], [14]. 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 [16].

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, such as land use change analysis, study on shifting cultivation, monitoring of pollution, assessment of burnt 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], [6], [9], [11], [13], and [15]. Artificial Neural Networks (ANNs) have emerged as an important tool for addressing many problems related to remote sensing images [2], [7], [8], [12], [14]. Change detection has emerged as one of the relatively new application areas of ANN [1], [11], regarding both supervised and also unsupervised approaches. Several general advantages of applying neural networks for classification of satellite imagery are the following [11]: (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 to both supervised and unsupervised pattern recognition for change detection in multi-temporal and multispectral satellite imagery. Firstly, we have experimented and evaluated the supervised neural classifier performances (Multilayer

Perceptron (MLP), Radial Basis Function Neural network (RBF), and supervised Self-Organized Map (SOM)) versus supervised statistical classifier performances (Bayes and Nearest Neighbor (NN)). Secondly, we have investigated the following unsupervised change detection techniques: Self-Organizing Map (SOM) (neural clustering), versus K-means (statistical clustering), and Fuzzy C-means (FCM) (fuzzy clustering). The considered techniques have been evaluated using a multi-temporal LANSAT 7 ETM+ satellite image of 400x400 pixels, corresponding roughly to 144 km2 of land in the vicinity of Markaryd – Sweden.

II. Supervised Change Detection Techniques

The proposed processing cascade for supervised change detection in multi-temporal and multispectral remote-sensing images consists of two main processing steps (Fig. 1):

(A) Feature selection:

(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 Neighbor.

A. 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 radiometrically calibrated. Each multispectral pixel is represented as an n-dimensional vector, where n is the number of bands. One can choose one of the three feature selection techniques described in the following sub-chapters.

A1. Concatenation of Corresponding Multispectral Pixels (CON)

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

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

the concatenation is computed according to

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

This result is then transferred to the next module.

A2. Absolute Difference between Corresponding Pixels (ADIP)

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

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

the absolute difference is computed according to

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

This vector is then transferred to the next module.

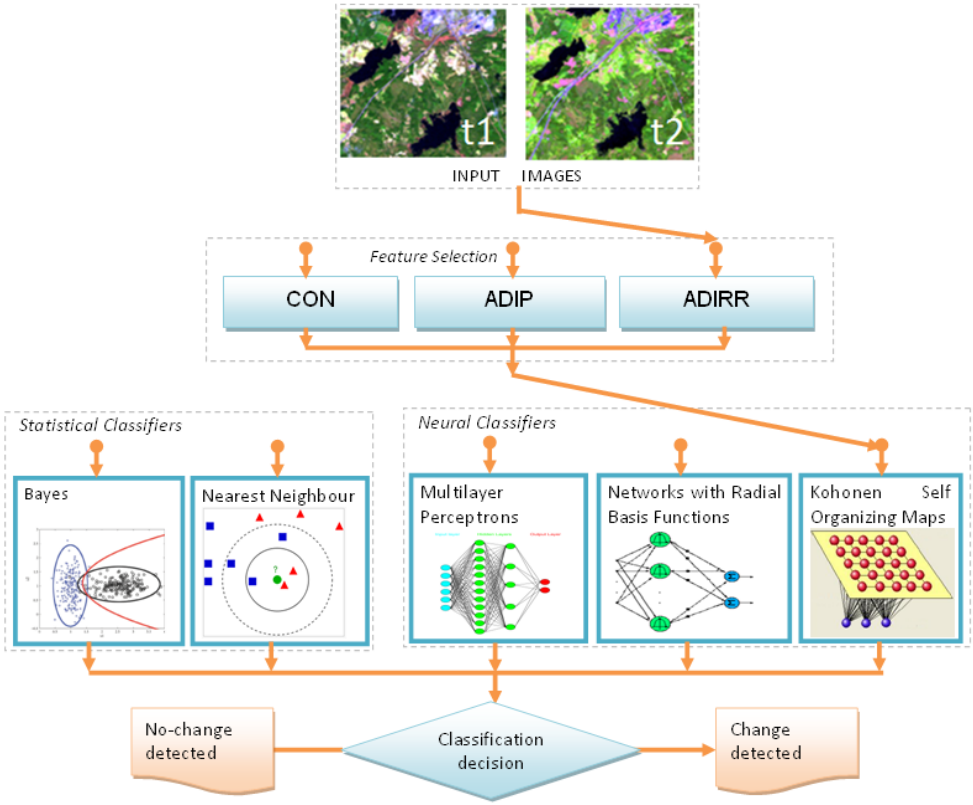


Fig. 1. Flowchart of the proposed land-cover change detection processing cascade using neural/statistical supervised classifiers

A3. 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 [11]

$$R_A = [a_1/a_2, a_1/a_3, .. a_1/a_n, a_2/a_3, ..., a_{n-1}/a_n]^T$$
$$R_B = [b_1/b_2, b_1/b_3, ... b_1/b_n, b_2/b_3, ..., b_{n-1}/b_n]^T$$

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). The absolute difference between R_A and R_B is then computed similarly as for *ADIP*, and the resulting vector $V = |R_A - R_B|$ is transferred to the classifiers.

B. Supervised Neural / Statistical Classification

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

B1. Neural Classifiers

Multilayer Perceptron (MLP)

MLP is the classical model of feed-forward back-propagation neural network [2], [12]. 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.

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.

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 [5]. With every node m, a *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 [5].

B2. Statistical Classifiers

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, covariance matrices Σ_1 and Σ_2, and that 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.

Nearest Neighbor (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 neighbor of the input vector belonging to the labeled training set.

C. Morphological post-processing

The morphological post-processing method is derived from the CORINE Land Cover Changes (CLCC) specifications on how the reference map was created. The particularities of the Landsat7 images, especially the 30m resolution, have also been taken into account. The following steps describe the post-processing algorithm:

Step 1. All regions of 8-connected pixels (of either class) are extracted

Step 2. For each region, the area (i.e. the number of pixels that form the region) is computed. If the region has fewer than 56 pixels (corresponding to 5ha), it changes the class label to the label of the other (surrounding) class.

Step 3. For each region, the slimness (i.e. the width of the bounding rectangle that covers the region, where the rectangle does not necessarily have horizontal and vertical sides) is computed. If the region is slimmer than 3.5 pixels (corresponding to 100m), it changes the class label to the label of the other (surrounding) class. ne further considers change detection as a

III. Unsupervised Change Detection Techniques

As any pattern recognition problem [2], the considered model of unsupervised change detection in multispectral satellite images has two main processing stages: (a) feature selection and (b) clustering (Fig. 2).

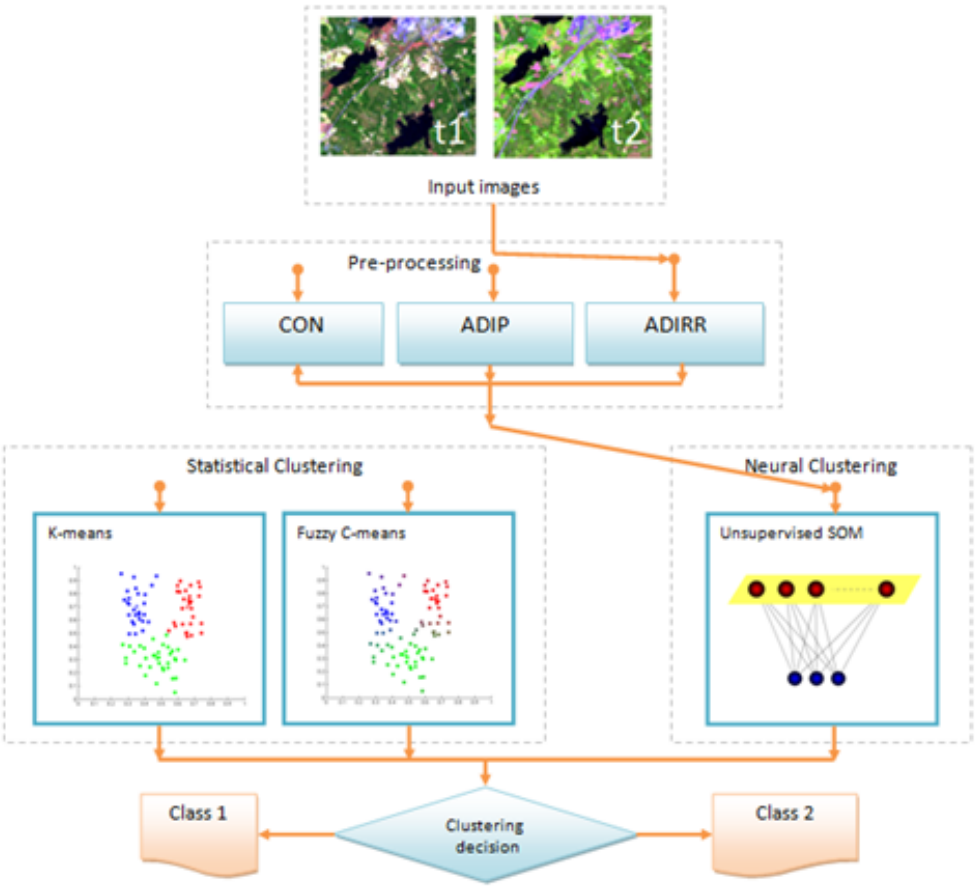


Fig. 2. Flowchart of the proposed land-cover change detection processing cascade using neural/statistical/fuzzy unsupervised classifiers

A. Feature Selection

One chooses one of the three feature selection techniques for land-cover change detection described in the previous chapter:

- Concatenation of Corresponding Multispectral Pixels (CON)
- Absolute Difference between Corresponding Pixels (ADIP)
- Absolute Difference of Reflectance Ratios (ADIRR).

B. Unsupervised Neural / Statistical /Fuzzy Classification

Neural clustering: Self-Organizing Map (SOM)

The SOM is built on the same underlying principles as described by T. Kohonen [5]. The training is a competitive, unsupervised and self-organizing learning process, mapping n-dimensional input data space onto an m-dimensional array of nodes. Each node has an n-dimensional weight vector associated. In the training phase, the input vector set is comprised of all feature vectors and the weight vectors are adjusted according to the values of the input vectors, the positions of the neurons in the m-dimensional array relative to the best-matching node and the learning rate. After the training phase is complete, the weights of the network are frozen and the test phase for all vectors begins. The left half of the output neurons represents one class and the right half represents the other class.

Statistical clustering: K-means

The method called K-means (also known as Basic ISODATA) is a simple and well known iterative clustering algorithm. It attempts to minimize the sum of point-to-centroid distances according to the cost function J:

$$J = \sum_{j=1}^{K} J_j = \sum_{j=1}^{K} \sum_{X_i \in \omega_j} \| X_i - \mu_j \|^2 \qquad (2)$$

where X_i represents the i-th vector
μ_j represents the centroid of class j
K is the number of classes ($K = 2$)

Fuzzy clustering: Fuzzy C-means (FCM)

In non-fuzzy or hard clustering, data is divided into crisp clusters, where each data point belongs to exactly one cluster. In fuzzy clustering, the data points can belong to more than one cluster. Membership grades are associated with each of the points. These grades indicate the degree to which the data points belong to the different clusters. The FCM algorithm, also known as Fuzzy ISODATA, is based on the minimization of the objective function $J(\mathbf{U},\mathbf{V})$ to achieve a good classification.

$$J(\mathbf{U},\mathbf{V}) = \sum_{i=1}^{N}\sum_{j=1}^{C}(u_{ij})^m \left\| X_i - V_j \right\|^2 \qquad (3)$$

where X_i represents the i-th vector

V_j represents the centroid of class j

u_{ij} is the membership degree of vector X_i to cluster j

m is the fuzziness index.

There is no theoretical basis for the optimal selection of m, but the value $m = 2.0$ is usually chosen.

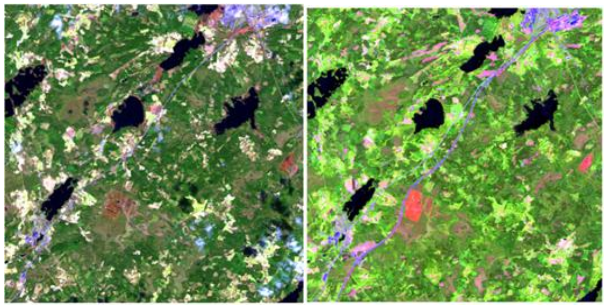
IV. Experimental Results

A. Satellite Image Database: Landsat 7 ETM+ Dataset 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 Apr 2002)
- LE71940212006186ASN00 (Acquisition date: 5 Jul 2006)

representing the same 400 x 400 pixel area, corresponding roughly to 144 km^2 of land from the region Markaryd, Sweden (Fig. 3). The selected image sequence is suitable for change detection since it 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. We have selected six bands for experiments: 1, 2, 3, 4, 5, 7, having the same resolution (30m).



(a) 5 Apr 2002 (b) 5 Jul 2006

Fig. 3. Landsat7 ETM+ image sequence displayed with 3 bands (Red=Band 5, Green=Band 4, Blue=Band 3)

For algorithm training and evaluation, we have used the reference change detection map built as a result of European Environmental Agency's effort of creating a standardized classification and indexation of the European territory, under the CORINE Land Cover project. The CORINE maps are generated based on the visual interpretation of experts, sometimes helped by aerial photographs, topographic maps and other additional information. In the context of this project, EEA also compiled vector databases for changes between some specific times. It has been the case with the CORINE Land Cover (CLC) Changes 2000–2006 database, part of which is used for the experiments described in this work [16]. 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.

According to the reference map, for the considered multi-temporal image (Fig. 3), there are 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 Changes 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%).

B. Parameters for Performance Evaluation

In order to evaluate the performances of the experimented change detection algorithms, 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

C. Experimental Results for Supervised Change Detection

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 I-III and Figs. 4-9. Examples of neural change detection maps by comparison with CORINE Land Cover (CLC) Changes reference map are given in Fig. 10 to Fig. 22.

Table I. Change detection performances as a function of classifier type using (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 II. Change detection performances as a function of classifier type using (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 III. Change detection performances as a function of classifier type using (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

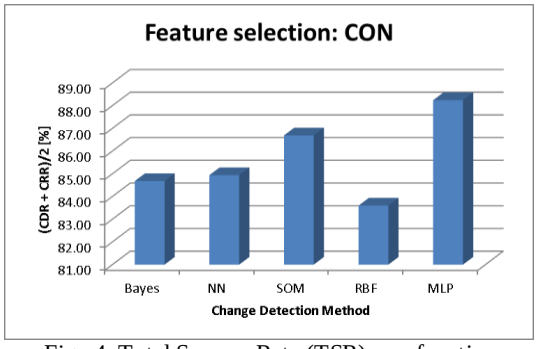


Fig. 4. Total Success Rate (TSR) as a function of classifier type using CON for feature selection

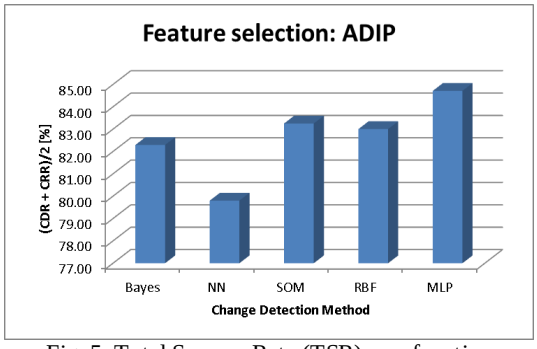


Fig. 5. Total Success Rate (TSR) as a function of classifier type using ADIP for feature selection

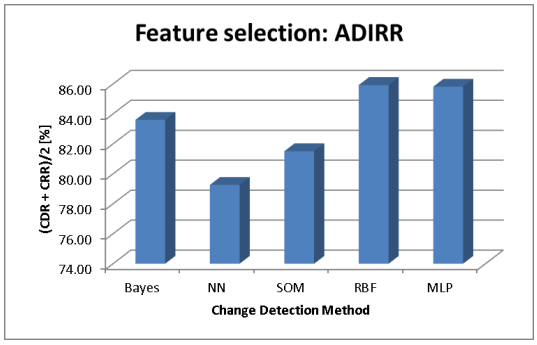


Fig. 6. Total Success Rate (TSR) as a function of classifier type using ADIRR for feature selection

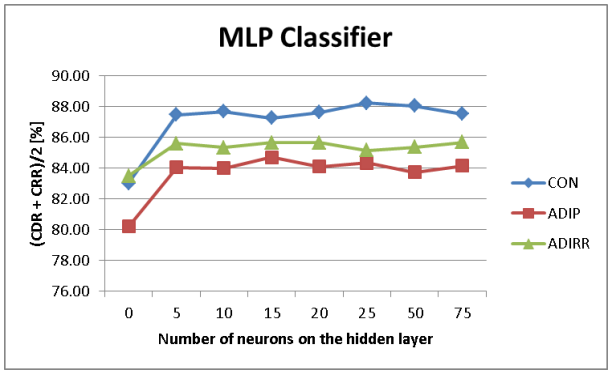


Fig. 7. Total Success Rate (TSR) for MLP classifier as a function of the number of neurons of the hidden layer

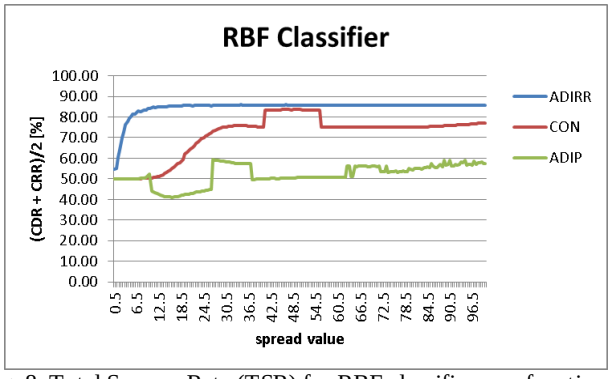


Fig. 8. Total Success Rate (TSR) for RBF classifier as a function of the spread parameter values

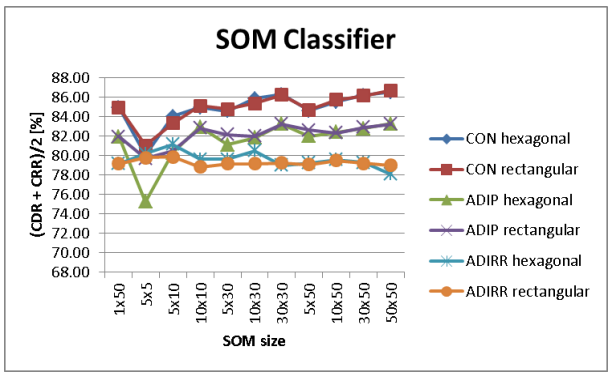


Fig. 9. Total Success Rate (TSR) for supervised SOM classifier as a function of SOM size and architecture (99 training epochs)

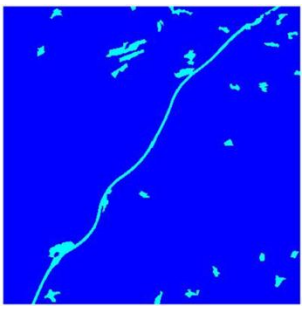
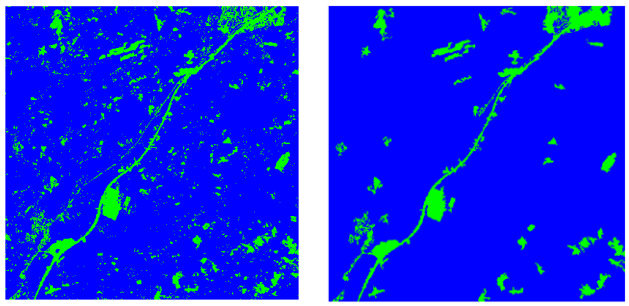
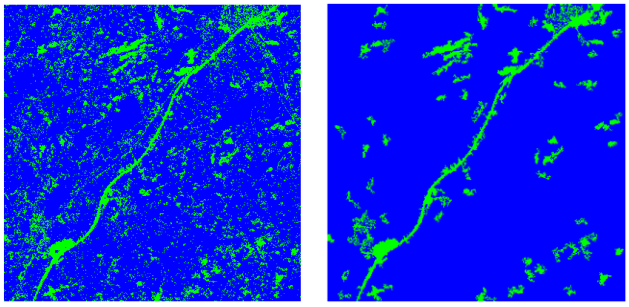


Fig. 10. CLCC reference change map



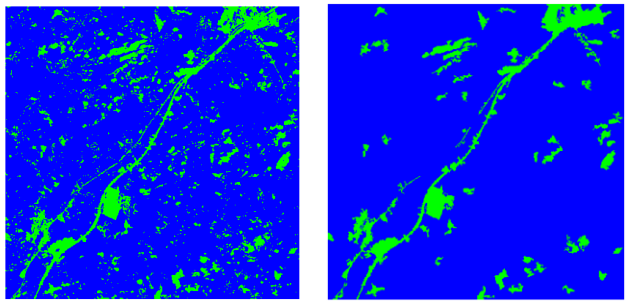
(a) (b)

Fig. 11. Change maps for the cascade CON-Bayes
(a) without post-processing, (b) with post-processing



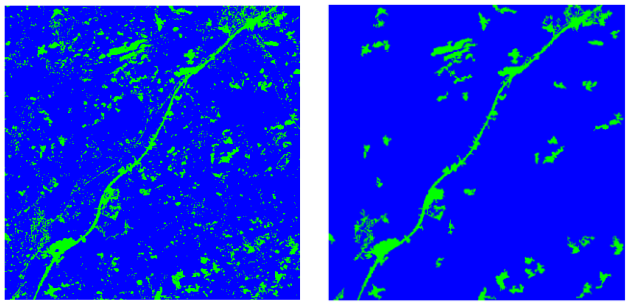
(a) (b)

Fig. 12. Change maps for the cascade CON-NN
(a) without post-processing, (b) with post-processing



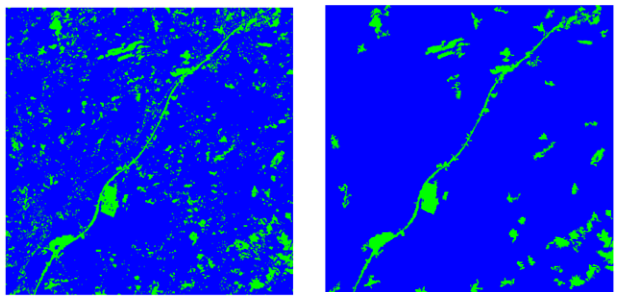
(a) (b)

Fig. 13. Change maps for the cascade CON-SOM
(a) without post-processing, (b) with post-processing



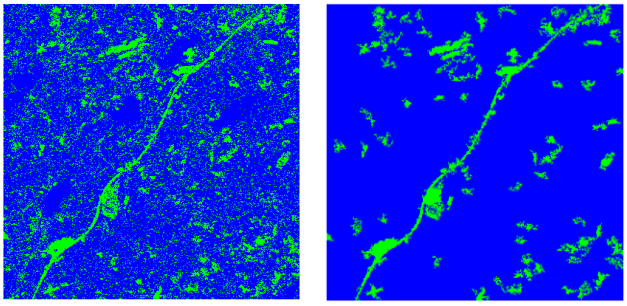
(a) (b)

Fig. 14. Change maps for the cascade CON-MLP
(a) without post-processing, (b) with post-processing



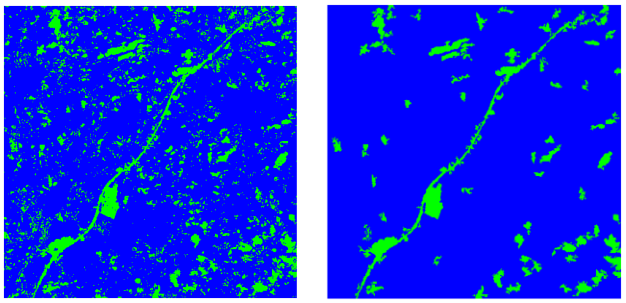
(a) (b)

Fig. 15. Change maps for the cascade ADIP-Bayes
(a) without post-processing, (b) with post-processing



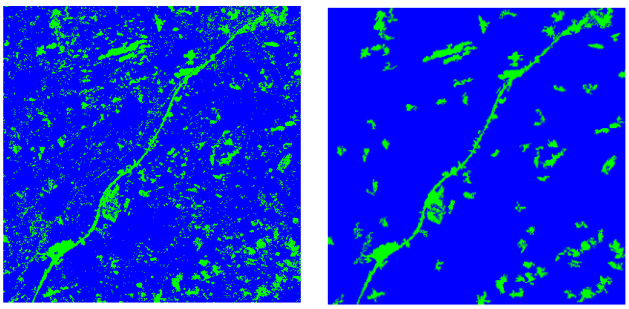
(a) (b)

Fig. 16. Change maps for the cascade ADIP-NN
(a) without post-processing, (b) with post-processing



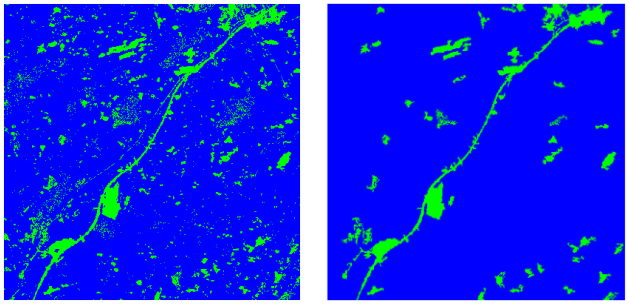
(a) (b)

Fig. 17. Change maps for the cascade ADIP-SOM
(a) without post-processing, (b) with post-processing



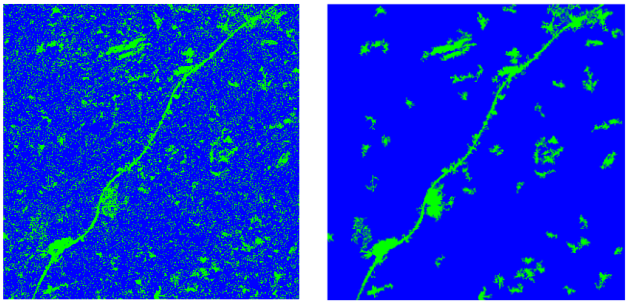
(a) (b)

Fig. 18. Change maps for the cascade ADIP-MLP
(a) without post-processing, (b) with post-processing



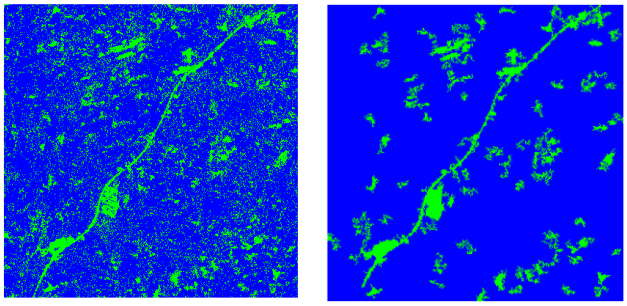
(a) (b)

Fig. 19. Change maps for the cascade ADIRR-Bayes
(a) without post-processing, (b) with post-processing



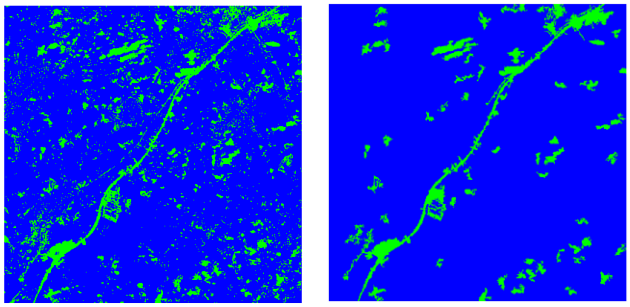
(a) (b)

Fig. 20. Change maps for the cascade ADIRR-NN
(a) without post-processing, (b) with post-processing



(a) (b)

Fig. 21. Change maps for the cascade ADIRR-SOM
(a) without post-processing, (b) with post-processing



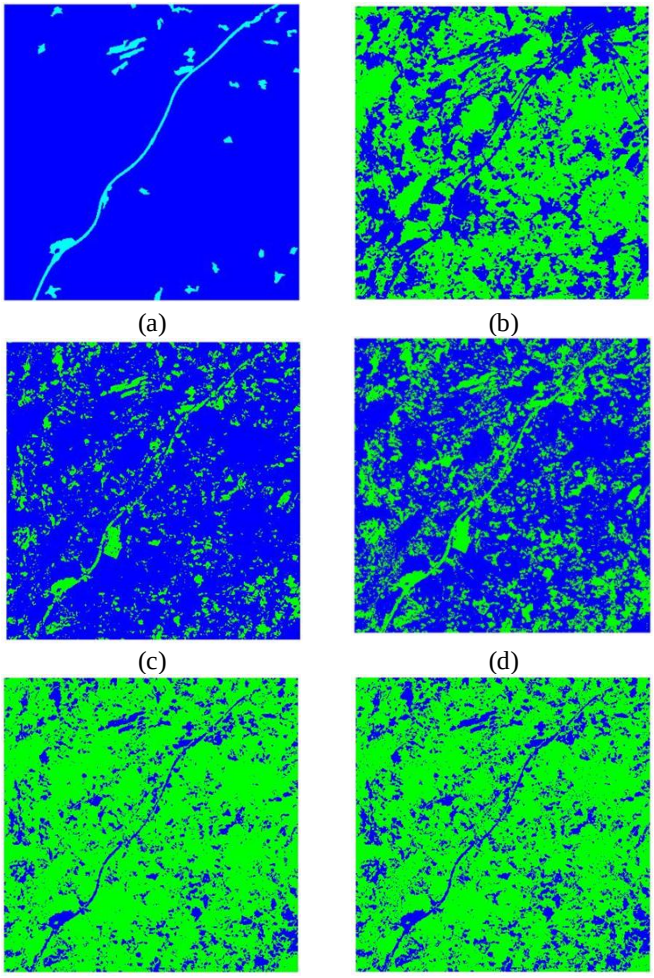
(a) (b)

Fig. 22. Change maps for the cascade ADIRR-SOM
(a) without post-processing, (b) with post-processing

D. *Experimental Results for Unsupervised Change Detection*

The experimental results obtained by applying the considered *unsupervised* change detection techniques are given in Tables IV-VI. Examples of change detection maps (for the best clustering result) are given in Fig. 23.

The lack of supervision inherent for these methods implies a lack of control over which algorithm label is associated with each class. As it can be clearly seen in Fig. 23, some methods have associated green with change and blue with non-change, whereas others have done the opposite. For a correct comparison of results, this correspondence mismatch has been corrected at the performance evaluation stage.



(a) (b)

(c) (d)

(e) (f)

Fig. 23. Example of unsupervised change detection maps:
(a) CLCC reference change map.
(b) Change detection map using the cascade CON – SOM
(c) Change detection map using the cascade ADIP – K-means
(d) Change detection map using the cascade ADIP – SOM
(e) Change detection map using the cascade ADIRR – FCM
(f) Change detection map using the cascade ADIRR – SOM

Table IV. Change detection performances as a function of clustering type using (CON) for feature selection

Algorithm	Parameters	CDR [%]	CRR [%]	MR [%]	FPR [%]	(CDR + CRR)/2 [%]
K-means	---	43.29	74.58	56.71	25.42	58.94
FCM	---	48.73	73.17	51.27	26.83	60.95
SOM	14 neurons	90.82	62.32	9.18	37.68	76.57

Table V. Change detection performances as a function of clustering type using (ADIP) for feature selection

Algorithm	Parameters	CDR [%]	CRR [%]	MR [%]	FPR [%]	(CDR + CRR)/2 [%]
K-means	---	67.55	87.16	32.45	12.84	77.35
FCM	---	70.21	85.59	29.79	14.41	77.90
SOM	2 neurons	84.38	72.06	15.62	27.94	78.22

Table VI. Change detection performances as a function of clustering type using (ADIRR) for feature selection

Algorithm	Parameters	CDR [%]	CRR [%]	MR [%]	FPR [%]	(CDR + CRR)/2 [%]
K-means	---	64.82	87.50	35.18	12.50	76.16
FCM	---	70.23	83.44	29.77	16.56	76.83
SOM	2 neurons	74.60	78.79	25.40	21.21	76.70

V. Concluding Remarks

A. Supervised Change Detection

1) First part of this paper presents a supervised 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 Organizing Map (SOM). For comparison, we have also considered two well-known statistical classifiers (Bayes and Nearest Neighbor (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 I, Fig. 4), 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 II, Fig. 18), 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 III, Fig. 6), 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. 7), 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. 8) 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. 9). 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. 14. The results may be further improved using a post-processing stage up to a TSR of 91.14%, also for cascade CON-MLP. 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%)!

B. Unsupervised Change Detection

1) Second part of this paper presents and compares several *unsupervised* land-cover change detection techniques in multi-temporal and multispectral satellite imagery, by using neural clustering (Self-Organizing Map (SOM)), versus statistical clustering (K-means), and versus fuzzy clustering (Fuzzy C-means) .
2) The three feature selection techniques that have been considered in experiments are: (CON) the concatenation of corresponding pixels, (ADIP) the computation of absolute differences of corresponding pixels, (ADIRR) the computation of absolute differences of reflectance ratios of corresponding pixels.
3) Concatenation does not give direct information about the amount of change present between corresponding pixels. While the SOM was flexible enough to overcome this disadvantage, leading to the total success rate (TSR) of 76.57%, the statistical methods were not. This explains the huge gap between the neural versus statistical and fuzzy techniques in Table IV (an advantage of 17.63% of SOM versus K-means as well as an advantage of 15.62% of SOM versus FCM).
4) Any pre-processing involving differences (as ADIP and ADIRR) re-distributes the information available so that all clustering methods achieve comparable global results. The differences of TSR are within 1%.
5) The statistical and fuzzy methods emphasize the correct detection of non-changes, whereas the SOM is emphasizing the correct detection of changes. For instance, in the ADIP-FCM processing chain, CDR = 70.21% and CRR = 85.59%. By contrast, in the ADIP-SOM processing chain, CDR = 84.38% and CRR = 72.06%. The two global results are balanced, within 0.35%, but the emphasis is different.
6) The best overall performance among clustering methods is obtained using the cascade ADIP-SOM (TSR=78.22%). The corresponding change detection map is shown next to the reference map in Fig. 23d.
7) The results may be further improved using a post-processing stage.

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] D. Klimešová, E.Ocelíková, GIS and Image Processing, *Internat. Journal of Mathematical Models and Methods in Applied Science*, issue 5, Vol. 5, 2011, pp.915-922.

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

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

[7] V. E. Neagoe and G. Strugaru, "Concurrent Neural Classifiers for Pattern Recognition in Multispectral Satellite Imagery", *Proc. 12th WSEAS Computer Conference*, Heraklion, Crete Island, Greece, July 22-25, 2008, pp. 893-898.

[8] V. E. 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.

[9] V.E. Neagoe, M. Neghina and M. Datcu, "A Neural Network Approach for Land Cover-Change Detection in Multi-Temporal Multispectral Remote Sensing Imagery", *Proc. 11th WSEAS Internat. Conf. Signal Processing, Computational Geometry and Artificial Vision (ISCGAV)*, Aug. 23-25, 2011, pp. 221-226.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] D. Samek, D. Mana, Artificial neural networks in artificial time series prediction benchmark, *Internat.Journal of Mathematical Models and Methods in Applied Science*, issue 6, Vol. 5, 2011, pp.1085-1093.

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

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

Victor-Emil I. Neagoe was born in Pitesti (Arges county, Romania) on May 31, 1947. From 1965 till 1970 he attended the courses of the Faculty of Electronics and Telecommunications, Polytechnic Institute of Bucharest, Romania. In 1970 he received the M.S. degree of diplomat engineer in electronics and telecommunications as a head of his series (with Honor Diploma, average of marks 9.97 out of 10). He also obtained the Ph.D. degree in the same field from the same institution in 1976 as well as the Postgraduate Master degree in Applied Mathematics and Informatics from the Faculty of Mathematics, University of Bucharest in 1981 (average of marks 10).

From 1970 till 1976 he has been an Assistant Professor at the Faculty of Electronics and Telecommunications, Polytechnic Institute of Bucharest, branches: Information Transmission Theory, Television, and Applied Electronics. From 1978 till 1991 he has been a Lecturer at the same Institute and Faculty, courses: Information Transmission Theory and Applied Electronics. Since 1991 he has been a Professor of the Polytechnic University of Bucharest, Romania, where he teaches the following courses: pattern recognition and artificial intelligence; detection and estimation for information processing; digital signal processing; computational intelligence; data mining. He has been a Ph.D. supervisor since 1990; he co-ordinates now ten Ph.D. candidates. He has published more than 120 papers; his research interest includes pattern recognition, nature inspired intelligent techniques (computational intelligence), multispectral and hyperspectral satellite/aerial image analysis, image compression, recognition, biometrics, sampling theory.

Prof. Victor-Emil Neagoe has been a Member of IEEE (Institute of Electrical and Electronics Eng., New York) since 1978 and a Senior Member IEEE since 1984. He has been included in Who's Who in the World and Europe 500; particularly, he has been recently included in *Who's Who in the World 2011* (28th Edition) as well as in *Who's Who in Science and Engineering 2011-2012* (11th Edition). Website: http://www.victorneagoe.com.

Mihai Neghina received the B.S. degree in mobile and satellite communications in 2006 and the M.S. degree in images, patterns, and artificial intelligence in 2008 from the University "Politechnica" of Bucharest UPB, Romania. He is currently preparing the Ph.D. thesis titled Computational Intelligence Techniques for Pattern Recognition in Multispectral Images at the same University.

His current research interests focus on multispectral imagery, artificial intelligence and pattern recognition.

Mihai Datcu received the M.S. and Ph.D. degrees in Electronics and Telecommunications from the University "Politechnica" of Bucharest UPB, Romania, in 1978 and 1986. In 1999 he received the title "Habilitation à diriger des recherches" from Université Louis Pasteur, Strasbourg, France. He holds a professorship in electronics and telecommunications with UPB since 1981. Since 1993 he is scientist with the German Aerospace Center (DLR), Oberpfaffenhofen. He is developing algorithms for model based information retrieval from high complexity signals and methods for scene understanding from synthetic aperture radar (SAR) and interferometric SAR data. He is engaged in research related to information theoretical aspects and semantic representations in advanced communication systems. Currently he is Senior Scientist and Image Analysis research group leader with the Remote Sensing Technology Institute (IMF) of DLR, Oberpfaffenhofen, coordinator of the CNES-DLR-ENST Competence Centre on Information Extraction and Image Understanding for Earth Observation, and professor at Paris Institute of Technology/GET Telecom Paris. His interests are in Bayesian inference, information and complexity theory, stochastic processes, model-based scene understanding, image information mining, for applications in information retrieval and understanding of high resolution SAR and optical observations. He has held visiting professor appointments from 1991 to 1992 with the Department of Mathematics of the University of Oviedo, Spain, from 2000 to 2002 with the Université Louis Pasteur, and the International Space University, both in Strasbourg, France. In 1994 he was guest scientist with the Swiss Center for Scientific Computing (CSCS), Manno, Switzerland and in 2003 he was visiting professor with the University of Siegen, Germany. From 1992 to 2002 he had a longer invited professor assignment with the Swiss Federal Institute of Technology ETH Zürich. He is involved in advanced research programs for information extraction, data mining and knowledge discovery and data understanding with the European Space Agency (ESA), Centre National d'Etudes Spatiales (CNES), NASA, and in a variety of European projects. He is a member of the European Image Information Mining Coordination Group (IIMCG).