

# Neural Network Combination by Fuzzy Integral for Robust Change Detection in Remotely Sensed Imagery

**Hassiba Nemmour**

*Signal Processing Laboratory, Faculty of Electronic and Computer Sciences, University of Sciences and Technology  
Houari Boumediene , 16111 Algiers, Algeria  
Email: hnemmour@lycos.com*

**Youcef Chibani**

*Signal Processing Laboratory, Faculty of Electronic and Computer Sciences, University of Sciences and Technology  
Houari Boumediene , 16111 Algiers, Algeria  
Email: ychibani@usthb.dz*

*Received 31 December 2003; Revised 5 December 2004*

Combining multiple neural networks has been used to improve the decision accuracy in many application fields including pattern recognition and classification. In this paper, we investigate the potential of this approach for land cover change detection. In a first step, we perform many experiments in order to find the optimal individual networks in terms of architecture and training rule. In the second step, different neural network change detectors are combined using a method based on the notion of fuzzy integral. This method combines objective evidences in the form of network outputs, with subjective measures of their performances. Various forms of the fuzzy integral, which are, namely, Choquet integral, Sugeno integral, and two extensions of Sugeno integral with ordered weighted averaging operators, are implemented. Experimental analysis using error matrices and Kappa analysis showed that the fuzzy integral outperforms individual networks and constitutes an appropriate strategy to increase the accuracy of change detection.

**Keywords and phrases:** remote sensing, change detection, neural network, fuzzy integral.

## 1. INTRODUCTION

Analysis of multitemporal images of remote sensing is used for multiple purposes like environment monitoring and wide-area surveillance. These applications involve the identification of changes in land cover and land use practices. Hence, even, a pair of spatially registered images acquired on the same ground area at different times is analyzed to identify areas that have changed. Commonly, the comparison of independently produced classifications of data is used since it provides complete knowledge upon the change [1]. However, there are major problems associated with this technique. On one hand, its accuracy is critically dependent upon the two individual classifications. On the other hand, it does not allow the detection of subtle changes within a land cover class [2]. Recently, an alternative approach based on simultaneous classification of multitemporal data begins to be used to overcome these drawbacks and allow an automatic extraction of different kinds of change [3, 4]. To develop such a change detector, one can adopt statistical classifiers that are widely used in remote sensing such as the max-

imum likelihood. However, these algorithms are based on hard and commonly untenable assumptions about the data. Therefore, nonparametric classifiers such as neural networks and fuzzy classifiers are increasingly being used. Presently, we focus our attention on artificial neural networks (ANNs), which have been successfully applied in a wide range of applications including classification and change detection in remotely sensed data [5, 6, 7]. Theoretically speaking, ANNs are able to achieve an accurate result with high generalization capacity. Nevertheless, in practice, their use poses several problems. In addition to the large variety of training algorithms, we are faced to a vast selection of possible network architectures and setup parameters. A bad choice of these factors affects subjectively the generalization capacity. Moreover, different neural networks may a priori perform differently from a land cover class to another. In this paper, we propose the combination (or fusion) of some neural networks to achieve the best possible performance of change detection. In fact, several researchers have attempted to use multiple neural networks with an appropriate collective decision strategy [8]. Among the available techniques, the fuzzy

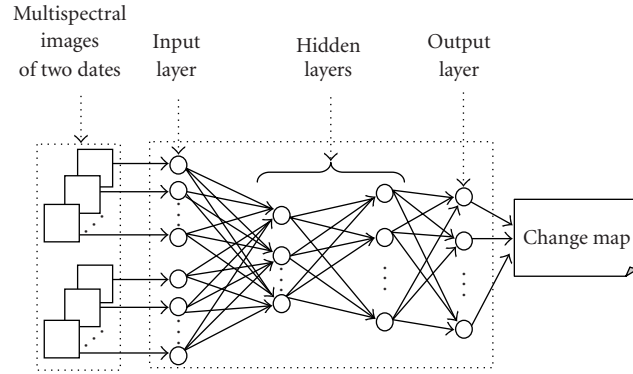


FIGURE 1: Neural network architecture for change detection.

integral has enjoyed a strong success in several applications of land cover classification, handwritten recognition, and image sequence analysis [8, 9, 10, 11, 12]. This method combines objective evidences in the form of network outputs according to expectations about their relevance. These expectations are estimated by using a fuzzy measure. In a first time, the experimental design includes architectural and training rule selection. Next, we carry out the combination of three neural networks using various forms of the fuzzy integral. Specifically, we use Sugeno integral, Choquet integral, and the extended Sugeno integral by two families of ordered weighted averaging operators (*OWA-OR* and *OWA-AND*). The rest of this paper is arranged as follows. Section 2 reviews the neural network change detector as well as its training rules. It presents also the combination via the fuzzy integral. Section 3 summarises experimental results including architectural and training algorithm selection for neural networks and the performance comparison evaluation of combination rules. Finally, Section 4 discusses and gives the main conclusions of the paper.

## 2. COMBINATION OF NEURAL-NETWORK-BASED CHANGE DETECTORS

### 2.1. Neural-network-based change detector

There are two approaches for change detection [2]. The first approach uses comparative analysis of independently produced classifications of data, in which pixel labels of two individual classifications are compared to detect changes. This approach gives complete information over the land cover change but errors in the two classifications appear in the final change map as missed or spurious changes. The second approach is based on simultaneous classification of multitemporal data. It overcomes some limitations of the first approach since the data are handled by the same classifier. Hence, land cover change classes are selected at the classifier input. For instance, if we are interested to assess changes from no urban to urban, we select pixels which were not in the urban class at  $t_1$  and are in this class at  $t_2$ . Similarly, if we want to extract a no-change class such as urban areas, we select pixels that belong to this class in both dates.

In this paper, we adopt the simultaneous analysis approach using artificial neural network classifiers. Recall that an artificial neural network is considered as a mapping device between an input set and an output set. Inside it is constructed from some processing units interconnected by weighted channels according to some architecture [5]. Thus, the neural network change detector has the following structure.

- (i) *Network input*: it receives input data extracted from two or more multitemporal images of the study area. Spectral channels of these images are spatially aligned and concatenated to form the input vector.
- (ii) *Network output*: the network output can be encoded by several ways. In our case, we use one output node per land cover category which is either change or no-change class.
- (iii) *Network architecture*: the number of hidden layers and their size are determined by the user. In general, for complex classification, like change detection, a network with two hidden layers achieves the best result in terms of the square error at convergence and the generalization ability [6].

For instance, Figure 1 shows the architecture of a two-hidden-layer network for change detection. Recall that supervised learning of this network aims to minimize the cost for all possible examples through the input-output relation by modifying iteratively synaptic weights (i.e., minimizing the mean squared error  $E$  between actual and desired outputs of the network which are, respectively,  $Y$  and  $Y_d$ ):

$$E = \sum_{j=1}^M (Y_j - Y_{d_j})^2, \quad (1)$$

$M$  is the number of the training patterns. Thereby, synaptic weights are updated at the iteration  $(t + 1)$  by

$$w(t + 1) = w(t) + \Delta w(t + 1), \quad (2)$$

$\Delta w(t + 1)$  corresponds to the weight change, which is

compued for the backpropagation algorithm by

$$\Delta w(t+1) = -\eta \frac{\partial E}{\partial w} + \alpha \Delta w(t), \quad (3)$$

$\eta$  is the step size and  $\alpha$  is the momentum. However, the main problem of the backpropagation is that many iterations are required to train a small network. Hence, an alternative algorithm based on the Kalman filtering has been proposed for accelerating the training stage of the neural network. For this training rule, the weight change is computed by means of the following equation:

$$\Delta w(t+1) = k_i e_i \mu_i, \quad (4)$$

$e_i$  is the error signal of the layer  $i$  computed at each node,  $\mu_i$  is the step size, and  $k_i$  is the Kalman gain (more details about this training rule are given in [13]).

### 2.2. Fuzzy measures and fuzzy integrals

In this section, we give the basic notions of the fuzzy measure and fuzzy integral.

#### (A) Fuzzy measure

Let  $Z$  be a finite set of elements, a set function  $g : 2^Z \rightarrow [0, 1]$  is called fuzzy measure if [7, 8, 9]

- (i)  $g(\phi) = 0, g(Z) = 1,$
- (ii)  $g(A) \leq g(B)$  if  $A \subset B.$

The fuzzy measure does not follow the addition rule, that is, if  $A, B \subset Z,$  then  $A \cup B = \phi:$

$$g(A \cup B) \neq g(A) + g(B). \quad (5)$$

However, while combining multiple sources, one must set the fuzzy measure of groups of sources. Therefore, Sugeno proposed the  $g_\lambda$  fuzzy measure so that

$$g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B). \quad (6)$$

#### (B) $g_\lambda$ fuzzy measure

The so-called  $g_\lambda$  fuzzy measure satisfies the following property [7, 8]: let  $Z = \{z_1, \dots, z_n\}$  be the set of available change detectors. For each change detector  $z_i$  to be combined, we associate a fuzzy measure  $g_k(z_i)$  indicating its performance in the class  $k$ . For a given pixel, let  $h_k(z_i)$  be the objective evidence of the change detector  $z_i$  for the class  $k$ . The set of change detectors is then rearranged such that the following relation holds:  $h_k(z_1) \geq \dots \geq h_k(z_n) \geq 0$ . We obtain an ascending sequence of change detectors  $A_i = \{z_1, \dots, z_i\}$ , whose fuzzy measures are constructed as

$$g_k(A_1) = g_k(z_1),$$

$$g_k(A_i) = g_k(A_{i-1} \cup z_i) = g_k(A_{i-1}) + g_k(z_i) + \lambda g_k(A_{i-1})g_k(z_i). \quad (7)$$

For each class  $k, \lambda$  is determined by solving an  $n-1$  degree equation:

$$\lambda + 1 = \prod_{i=1}^n (1 + \lambda g_k(z_i)). \quad (8)$$

Notice that  $\lambda \in ]-1, \dots, +\infty[$  with  $\lambda \neq 0$ . It is important to stress that (7) allows us to construct the fuzzy measures in order to provide both the weight of a single change detector as well as the weight of a subset of change detectors. However, there is no rule which would be followed to attribute  $g_k$  values. In fact, they can be subjectively assigned by an expert, or computed from the training data [8]. In this paper,  $g_k$  is expressed by the fuzzy accuracy per land cover class computed using a validation set.

#### (C) Sugeno integral

Sugeno integral  $I_S,$  of a function  $h : Z \rightarrow [0, 1]$  with respect to a fuzzy measure  $g$  over  $Z,$  is computed by

$$I_S(k) = \text{Max}_{i=1}^n [ \text{Min} (h_k(z_i), g_k(A_i)) ]. \quad (9)$$

This integral has been extended by using two special families of ordered weighted averaging operators OWA which are OWA-AND and OWA-OR operators [9].

#### (D) S-OWA-AND integral

With the OWA-AND operator, the objective evidences  $h_k(z_i)$  are transformed according to

$$\tilde{h}_k(z_i) = \frac{1-\alpha}{l+1} \sum_{i=1}^l h_k(z_i) + \alpha \text{Min}_{z_i \in Z} \{h_k(z_i)\}. \quad (10)$$

The new evidences are then utilized in (9) to compute the new form of Sugeno integral termed  $I_{\text{AND}}$ .

#### (E) S-OWA-OR integral

The OWA-OR performs on the values of the fuzzy integral computed for each class before evaluating the final aggregated decision:

$$I_{\text{OR}}(k) = \frac{1-\beta}{2^n} \sum_{k \in E} I_S(k) + \beta \text{Max}_{k \in E} \{I_S(k)\}. \quad (11)$$

Note that  $E$  is the set of classes. Parameters  $\alpha$  and  $\beta$  lie in the unit interval, and could provide somewhat different results when extending the fuzzy integral by OWA operators [9].

#### (F) Discrete Choquet integral

The discrete Choquet integral of a function  $h : Z \rightarrow R^+$  with respect to  $g$  is defined as [10]

$$I_C(k) = \sum_{i=1}^n \{h_k(z_i) - h_k(z_{i-1})\} g_k(A_i), \quad (12)$$

where indices  $i$  have been permuted so that  $0 \leq h_k(z_1) \leq \dots \leq h_k(z_n) \leq 1$  holds,  $A_i = \{z_i, \dots, z_n\},$  and  $h_k(z_0) = 0.$

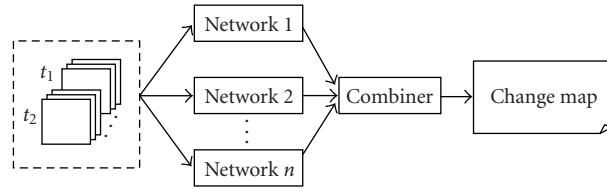


FIGURE 2: Neural network combination for change detection.

```

BEGIN
  For each class  $k$ 
    Calculate  $\lambda$ ; /* using the importance of each network */
  End For.
  For each class  $k$  do
    For each neural network  $z_i$  do
      Compute  $h_k(z_i)$ ;
      Compute  $g_k(z_i)$  using  $\lambda$ 
    End For.
    /* Compute the fuzzy integral for the class  $k$  */
    If (Choquet integral), then Compute  $I_C(k)$ ;
    Else
      If (Sugeno integral), then Compute  $I_S(k)$ ;
    Else
      If (S-OWA-AND integral), then Compute  $\tilde{h}(z_i)$ , Compute
       $I_{AND}(k)$ ;
    Else
      Compute  $I_S(k)$ , Compute  $I_{OR}(k)$ ;
    End For.
  Assign the pixel to the class corresponding to the maximal integral.
END.
  
```

ALGORITHM 1: Algorithmic implementation.

### 2.3. Algorithmic implementation

The basic idea of multiple neural networks is to develop  $n$  independently trained networks with relevant features, and combine their outputs to produce an average consensus decision [9] as shown in Figure 2. The fuzzy integral seeks the maximum grade of agreement between objective evidences  $h_k(z_i)$  according to their performances represented by the fuzzy measures  $g_k(z_i)$ . In Algorithm 1, we adapt the pseudocode of combination given in [8] to the work presented here. We define the performance of a neural network  $z_i$  in the class  $k$  by

$$g_k(z_i) = \frac{R_k}{R_k + R_c + R_o + R_r}, \quad (13)$$

where  $R_k$  denotes recognition rate,  $R_c$  commission rate,  $R_o$  omission rate, and  $R_r$  reject rate.

## 3. EXPERIMENTAL RESULTS

### 3.1. Description of the study area and evaluation criteria

The study area is located to the east of Algiers, Algeria. It is a coastal region which comprises the *Isser* River and watershed as well as several land cover types which are grouped as follows.

- (i) *Water bodies*: portion of Mediterranean sea, *Isser* river, lakes, and the *Isser* reservoir.
- (ii) *Vegetation*: areas covered by all kinds of vegetation and mixed forest.
- (iii) *Constructions*: areas covered by both dense and less dense settlements as well as single buildings.
- (iv) *Soil*: all kinds of bare soil.

Two SPOT high resolution visible (HRV) images acquired in May 1989 and June 1991, respectively, were selected for testing the validity of the proposed change detection method. During this period, the region has undergone rapid transitions from the classes' water, vegetation, and construction to soil. Thereby, we narrowed our attention to classify these land cover changes. However, satellite data depict other changes caused by the presence of clouds in the second image. To avoid all surprising behaviors of the change detection system towards these changes, an additional class " $X \Rightarrow$  clouds" was taken into account ( $X$  denotes whatever land cover class in the first image). The three spectral channels ( $XS_1$ ,  $XS_2$ , and  $XS_3$ ) of both images were spatially aligned by using ground control points with a second-order polynomial warp on the ENVI software. The registration was achieved with a residual error less than 0.13 pixels. As illustration, Figure 3 depicts only  $XS_1$  bands of both multitemporal images.

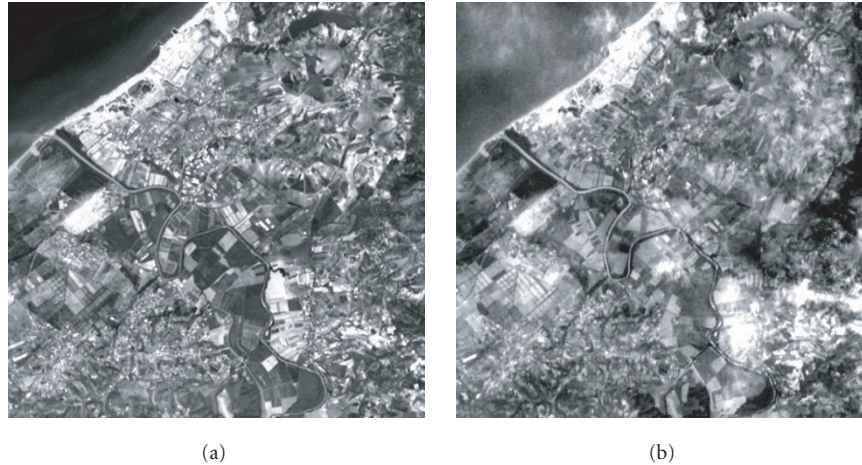


FIGURE 3: SPOT images of the eastern part of Algiers: (a)  $XS_1$  band of the image acquired in 1989, (b)  $XS_1$  band of the image acquired in 1991.

TABLE 1: Classes of interest.

Class labels	Class attribute (1989 $\Rightarrow$ 1991)	Training	Test	Validation
1	Water $\Rightarrow$ water	400	750	100
2	Vegetation $\Rightarrow$ vegetation	400	860	100
3	Construction $\Rightarrow$ construction	400	1020	100
4	Soil $\Rightarrow$ soil	400	950	100
5	Vegetation $\Rightarrow$ soil	400	780	100
6	Construction $\Rightarrow$ soil	400	660	100
7	Water $\Rightarrow$ soil	400	530	100
8	$X \Rightarrow$ clouds	400	420	100

Since the selection of the training data is crucial to the quality of the result, we used reference maps obtained from unsupervised classifications of available images. Unfortunately, these maps constitute the unique source of ground truth which can be used to select training data. In order to validate the results, the dataset for each land cover class was randomly split into three disjoint sets in order to be used, respectively, in the training stage, in the test stage, and for calculating fuzzy measures (Table 1). These data are linearly scaled between 0 and 1 by dividing radiometric values by 255.

Change detection performance is evaluated by using the usual error (or confusion) matrix which highlights the good class allocations or accuracy rates per land cover class. In addition, many accuracy measures can be derived from this matrix. We use then the overall recognition rate (ORR), which is computed by taking the ratio between the sum of the good allocations and the total number of test data, and the Kappa coefficient computed by the Khat. This latter is computed by using all elements of the error matrix. The more the *Khat* is closed to 1 (100%), the more the change detector is reliable [14].

### 3.2. Architectural and training rule selection

This experiment is conducted to seek the optimal neural network change detector in terms of architecture and training

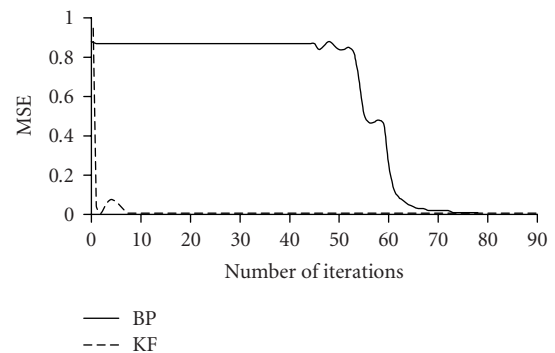


FIGURE 4: Mean squared error versus the number of iterations.

rule. First, the backpropagation (BP) and Kalman filtering (KF) algorithms were used to train the same network in order to compare their performances. Both algorithms were started from the same initial weights with random values ranging from  $-0.01$  through  $+0.01$  while setup parameters were selected to maximize the performance of each algorithm.

Figure 4 plots the evolution of the mean squared error (MSE) of the network across the time. As can be seen, KF converges more rapidly than the BP algorithm. This outcome

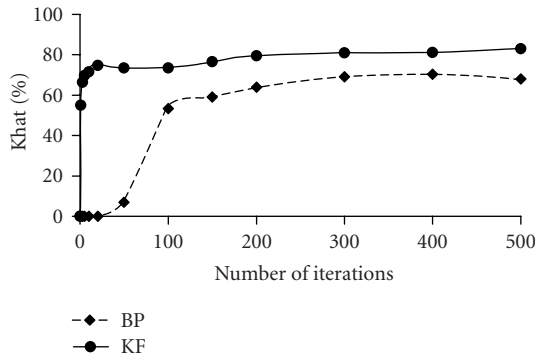


FIGURE 5: Khat according to the number of iterations.

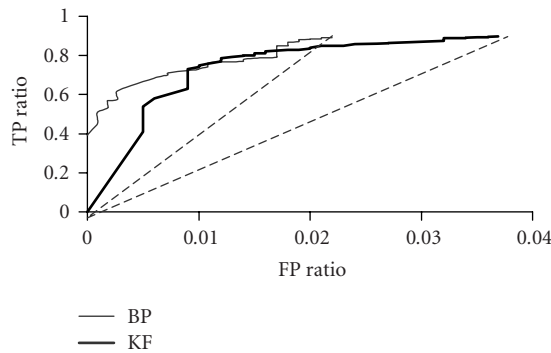


FIGURE 6: ROC curves.

is confirmed by the second test which calculates the *Khat* for various iterations. The result is depicted in Figure 5. We notice that KF rule reaches its maximal value of *Khat* after few iterations. In return, the *Khat* of the BP increases gradually across the time. In order to evaluate the ability to detect changes, we plot the ROC curves for both algorithms. These curves are formed by points whose coordinates are the false positive (FP) ratio (which expresses the amount of false alarms represented on the  $x$ -axis) and the true positive (TP) ratio (which expresses the amount of good detections represented on the  $y$ -axis). The change detection process is considered more effective if the area under the ROC curve and above the main diagonal is big. Figure 6 presents the obtained curves which show that KF algorithm is more accurate in terms of change detection. On the other hand, we perform another test to find the optimal architecture for change detection. Using the KF algorithm, we trained many neural networks comprising one and two hidden layers in which the number of nodes varies between 8 through 50. The result reported in Figure 7 indicates that in all cases the network with two hidden layers outperforms the one-hidden-layer net. Moreover, for both architectures, the increase of the number of hidden nodes does not necessarily lead to an improvement of performance. In fact, the best results are obtained with architectures in which the number of nodes per hidden layer varies between 10 and 20. Roughly speaking, the KF rule improves the performance of the neural network in both training and generalization stages. Nevertheless, different neural

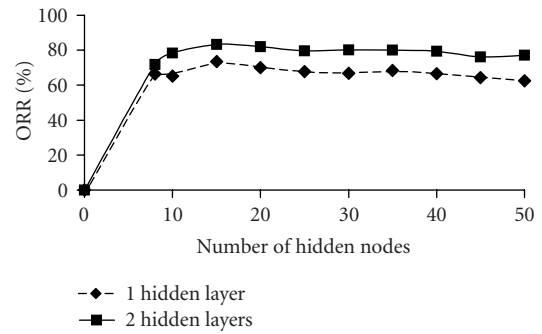


FIGURE 7: Influence of the number of hidden nodes over the change detection accuracy.

networks perform differently and make their errors in different regions of input space; we then conjecture that the combination of multiple neural networks may improve change detection.

### 3.3. Accuracy assessment of combination rules

The optimal number of experts to be combined depends on the task at hand, and should be found experimentally. This is not the purpose of the present work, we want just to evaluate the usefulness of the combination concept for land cover change detection. However, it is important to stress that when combining only two networks, if a network has a poor accuracy in a given class, the accuracy of the combination rule will be lower than that of the most precise network. Therefore, the number of combined networks must be superior to 2. Furthermore, it has been shown in [8] that when using the fuzzy integral as a combination rule, the recognition rate depends on the  $g$  values, if they change, the value of the fuzzy integral will change.

In the present work, three different networks having two hidden layers which contain, respectively, 10, 15, and 20 nodes were chosen to be combined according to the performances in terms of ORR given in Figure 7. The KF algorithm was used in the training stage which was stopped after 400 iterations. The values of  $g$  computed using (13) as well as the  $\lambda$  of each land cover class are reported in Table 2.  $g$  expresses the degree of importance of a given network in a particular class. Moreover, all  $\lambda$  values are closed to -1 because the sum of the different degrees of importance is greater than 1. In this case, the degree of importance may be interpreted as a plausibility value [8]. Banon [15] showed that  $\lambda \leq 0$  if  $g$  is a plausibility measure. On the other hand, parameters of *S-OWA-OR* and *S-OWA-AND* were experimentally fixed at 0.2 and 0.5, respectively.

Table 3 summarizes the error matrices obtained for the different change detectors, while Table 4 gives the corresponding ORR and *Khat* values. As expected, while the performance of individual neural networks varies from a land cover class to another, the combination rules give a significant improvement in the recognition rate of the majority of classes, especially the three change classes (classes whose labels are 5, 6, and 7). In consequence, they produce higher

TABLE 2: Values of  $g$  and  $\lambda$  for all land cover classes.

Classes	10	15	20	$\lambda$
1	0.975	0.962	0.975	-0.999
2	0.926	0.951	0.902	-0.999
3	0.741	0.857	0.612	-0.982
4	0.779	0.612	0.755	-0.969
5	0.678	0.761	0.735	-0.975
6	0.804	0.864	0.814	-0.994
7	0.901	0.865	0.867	-0.998
8	0.548	0.802	0.779	-0.881

TABLE 3: Accuracy rate per land cover class obtained for the different change detectors.

Classes	Individual networks %			$I_C$	Combined networks %		
	10 nodes	15 nodes	20 nodes		$I_S$	$I_{AND}$	$I_{OR}$
1	99.60	99.20	100.00	100.00	100.00	100.00	100.00
2	91.15	94.61	90.38	93.07	93.07	94.23	93.46
3	63.33	90.00	94.81	95.18	96.66	85.18	96.29
4	74.80	76.40	76.00	89.60	90.40	92.40	94.40
5	69.28	73.64	72.50	84.64	84.28	83.21	84.28
6	78.07	81.92	71.92	63.46	90.38	90.00	91.53
7	65.65	70.43	67.39	80.86	78.26	77.82	79.56
8	86.08	78.26	81.17	88.69	90.43	94.78	93.04

TABLE 4: ORR and Khat values obtained for the different change detectors.

	Individual networks %			$I_C$	Combined networks %		
	10 nodes	15 nodes	20 nodes		$I_S$	$I_{AND}$	$I_{OR}$
ORR	78.35	83.34	82.02	86.94	90.54	89.65	91.62
Khat	75.83	81.26	79.78	85.08	89.18	88.17	90.42

overall accuracy, with a gain more than 3% (with  $I_C$ ) and 8% (with  $I_{OR}$ ) over the best individual network. More specifically, the  $I_{OR}$  integral presents the most satisfactory results in terms of ORR and Khat. However, a surprising outcome in the class *Construction*  $\Rightarrow$  *soil* (class 6) which was the most critical one where the  $I_C$  exhibited still the bad accuracy (63.46%).

### 3.4. Visual inspection

As a result of implementing the different change detectors presented above, the maps depicted in Figure 8 were obtained. In this figure, only the change classes, which are namely, *construction*  $\Rightarrow$  *soil*, *water*  $\Rightarrow$  *soil*, and *vegetation*  $\Rightarrow$  *soil* are depicted since we are interested in change detection. It is easy to see that individual neural networks have limited generalization capacity. In fact, they cannot detect a large amount of areas of the class *vegetation*  $\Rightarrow$  *soil* (see circles in Figures 8a and 8b). Moreover, they produce an important number of spurious changes in the class *construction*  $\Rightarrow$  *soil* (see rectangles in Figures 8a and 8b). On the contrary, fuzzy integrals provide much cleaner change detection maps, with fewer numbers of missed and isolated spurious changes. Thus, we infer that the combination rules discriminate better

between classes, which demonstrates once again the effectiveness of change detector combination.

## 4. DISCUSSION AND CONCLUSIONS

In the last recent years, artificial neural networks have shown a particular relevance to land cover change detection because they provide complete information on the nature of change which is not allowed by all change detection schemes. However, with this method the user is faced to a large number of training rules and many possible architectures and setup parameters. The problem is that different networks produce different results since they make their errors in different regions of input space. In fact, there are various neural network optimization methods such as metropolis scheme and genetic algorithms [9] which derive the optimal value for only one parameter (the network architecture or the best sequence of the initialization of weights). Therefore, the combination of different neural networks seems to be more interesting since it aggregates the decision functions of different networks which can have different architectures, different training rules, and different setup parameters. In this paper, we investigated the potential of combining multiple neural

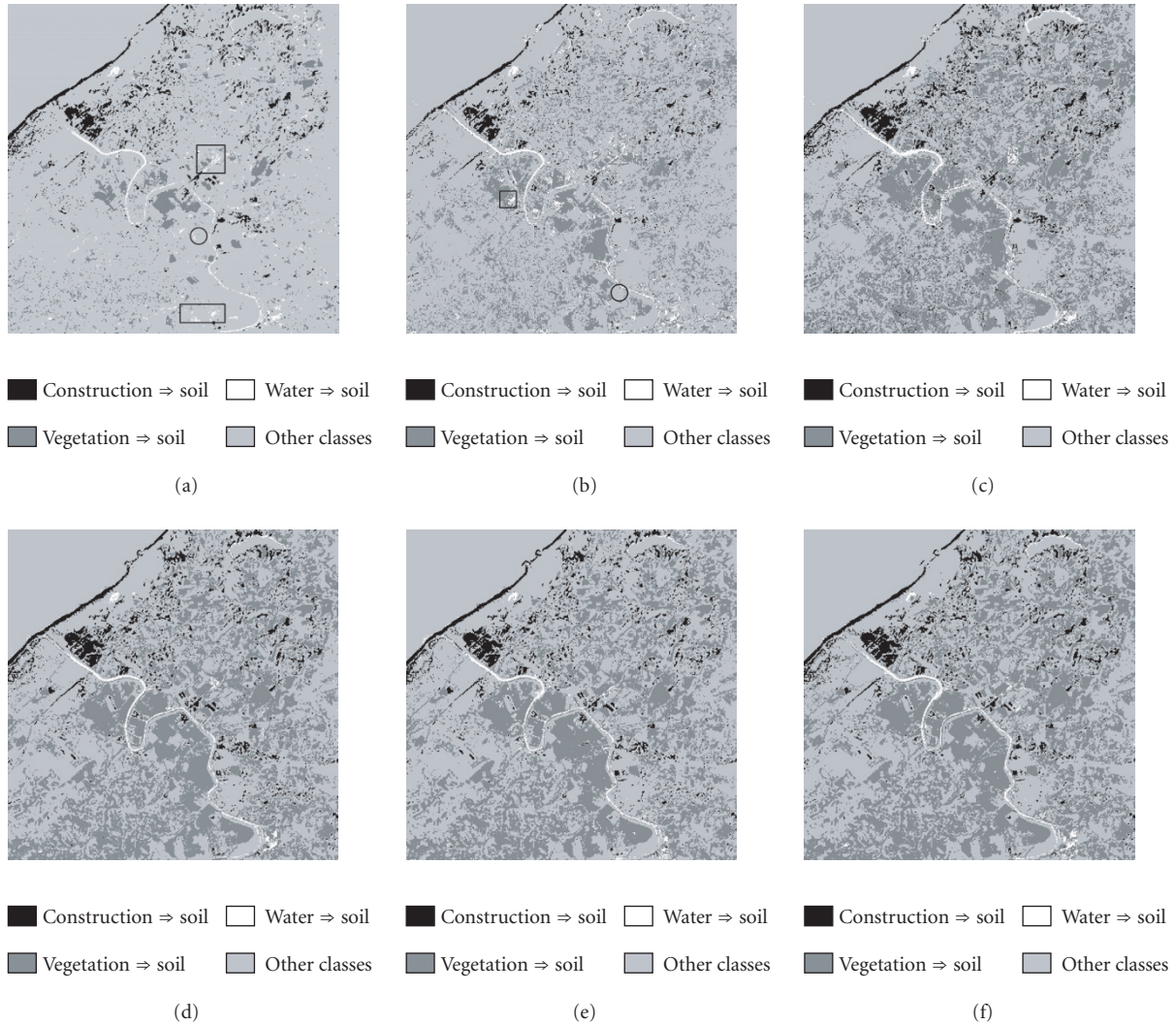


FIGURE 8: Change detection maps obtained for the different change detectors: (a) network with 15 nodes in each hidden layer, (b) network with 20 nodes in each hidden layer, (c) *Choquet* integral, (d) *Sugeno* integral, (e) *S-OWA-AND* integral, (f) *S-OWA-OR* integral.

networks for change detection in remotely sensed imagery. Our contribution was twofold. First, extensive experiments were carried out to obtain the optimal architectural selection for the neural network. In addition, the performance of the backpropagation algorithm which is the standard rule for training neural networks was compared to that of Kalman filtering algorithm. The results indicated that Kalman filtering algorithm is superior to the backpropagation in terms of convergence rapidity and change detection accuracy. Moreover, the use of artificial neural networks in remote sensing image processing was effectively highlighted for change detection. In a second step, we combined three different neural networks by using a strategy based on fuzzy integrals to increase change detection accuracy. The main advantage of this method is that it takes into account the reliability for individual networks as well as for a subset of networks. The conventional Sugeno and Choquet fuzzy integrals were used as combination operators. In addition, we implemented

two extensions of Sugeno integral based on *OWA-AND* and *OWA-OR* operators. Notice that the *OWA-AND* operator requires that all change detectors make the right decisions, while the *OWA-OR* requires that at least one change detector give the right decision. This is the reason for which  $I_{OR}$  and  $I_S$  outperformed  $I_{AND}$ . The results obtained by combining three different networks demonstrated the effectiveness of the combination concept. Specifically, it has been shown that each combiner provides an overall accuracy higher than those of individual networks and tends to improve the generalization performance by equalizing accuracy rates in individual classes. Furthermore, it has been shown that fuzzy integrals produced much closer results, which indicate that their different forms do not conceal the fact that they all stem from the same concept. Finally, throughout this study, the fuzzy integral appeared to be a straightforward computationally attractable approach which can enhance significantly the change detection in remotely sensed data.



## REFERENCES

- [1] P. Deer, *Digital change detection in remotely sensed imagery using fuzzy set theory*, Ph.D. thesis, Department of Computer Science, University of Adelaide, Adelaide, Australia, 1999, 185 pages, <http://www.kvocal.org/publications/1999.html>.
- [2] E. F. Lambin and A. H. Strahlers, "Change-vector analysis in multitemporal space: A tool to detect and categorize land-cover change processes using high temporal resolution satellite data," *Remote Sensing of Environment*, vol. 48, no. 2, pp. 231–244, 1994.
- [3] Y. Chibani and H. Nemmour, "Kalman filtering as a multi-layer perceptron training algorithm for detecting changes in remotely sensed imagery," in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS '03)*, vol. 6, pp. 4101–4103, Toulouse, France, July 2003.
- [4] H. Nemmour and Y. Chibani, "Comparison between object and pixel level approaches for change detection in multispectral images by using neural networks," in *Image and Signal Processing for Remote Sensing IX*, vol. 5238 of *Proceedings of SPIE*, pp. 551–559, Barcelona, Spain, September 2003.
- [5] G. M. Foody, M. B. McCulloch, and W. B. Yates, "Classification of remotely sensed data by an artificial neural network: issues related to training data characteristics," *Photogrammetric Engineering and Remote Sensing*, vol. 61, no. 4, pp. 391–401, 1995.
- [6] X. Dai and S. Khorram, "Development of a new automated land cover change detection system from remotely sensed imagery based on artificial neural networks," in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS '97)*, vol. 2, pp. 1029–1031, Singapore, Singapore, August 1997.
- [7] D. L. Civco, J. D. Hurd, E. H. Wilson, M. Song, and Z. Zhang, "A comparison of land use and land cover change detection methods," in *Proc. American Society for Photogrammetry and Remote Sensing (ASPRS '02)*, pp. 12–12, Washington, DC, USA, April 2002.
- [8] S.-B. Cho and J. H. Kim, "Combining multiple neural networks by fuzzy integrals for robust classification," *IEEE Trans. Syst., Man, Cybern. C*, vol. 25, no. 2, pp. 380–384, 1995.
- [9] S.-B. Cho, "Fuzzy aggregation of modular neural networks with ordered weighted averaging operators," *International journal of approximate reasoning*, vol. 13, no. 4, pp. 359–375, 1995.
- [10] A. Verikas, A. Lipnikas, K. Malmqvist, M. Bacauskiene, and A. Gelzinis, "Soft combination of neural classifiers: a comparative study," *Pattern recognition letters*, vol. 20, no. 4, pp. 429–444, 1999.
- [11] S.-B. Cho, "Fusion of neural networks with fuzzy logic and genetic algorithm," *Integrated Computer-Aided Engineering*, vol. 9, no. 4, pp. 363–372, 2002.
- [12] Z.-Q. L. Bruton, L. T. Bezdek, J. C. Keller, et al., "Dynamic image sequence analysis using fuzzy measures," *IEEE Trans. Syst., Man, Cybern. B*, vol. 31, no. 4, pp. 557–572, 2001.
- [13] R. S. Scalero and N. Tepedelenioglu, "A fast new algorithm for training feedforward neural networks," *IEEE Trans. Signal Processing*, vol. 40, no. 1, pp. 202–210, 1992.
- [14] R. G. Congalton, "Accuracy assessment and validation of remotely sensed and other spatial information," *International journal of Wildland Fire*, vol. 10, no. 4, pp. 321–328, 2001.
- [15] G. Banon, "Distinction between several subsets of fuzzy measures," *Fuzzy Sets and Systems*, vol. 5, no. 3, pp. 291–305, 1981.

**Hassiba Nemmour** received the M.S. degree from the Faculty of Electronic and Computer Sciences, USTHB University, Algiers, Algeria, in 2003. Actually, she is a Ph.D. student in the same faculty. Her research interests include machine learning, soft computing, and remotely sensed image processing.



**Youcef Chibani** was born in Algiers, Algeria. He received the Master's and State Doctoral degrees in electrical engineering from the University of Science and Technology Houari Boumediene, Algiers, Algeria. He has been teaching and researching as an Assistant Professor since 2002. His research interests include the use of the wavelet decomposition, neural networks, and support vector machines in many applications as multi-sensor image fusion, change detection, and multimedia signal processing. He coauthored many papers published in international peer-reviewed journals and conferences.

