IADIS International Journal on Computer Science and Information Systems
Vol. 4, No. 1, pp. 42-53
ISSN: 1646-3692

A New Learning Algorithm for the Fuzzy Adaptive Resonance Theory: Multispectral Classification of the Algiers's Bay

Farid Alilat **Saliha Loumi** **Boualem Sansal**
Université des Sciences et de la Technologie Houari Boumediene, Faculté d'Electronique et Informatique, Laboratoire de Traitement d'Images et Rayonnement Atmosphérique,
BP 32 El Alia Bab-ezzouar 16111, Alger, Algérie

Abstract

This paper presents a new learning algorithm for the fuzzy adaptive resonance theory. The modification allows us to supervise the fuzzy ART and to simplify ARTMAP network. It consists to find network's parameters (comparison, training and vigilance) which gave the minimum quadratic distances between the output of the training base and those obtained by the network. A comparative study of these two parameterized network and an third modified fuzzy ARTMAP are done. In this last network, learning is done differently. We don't take account of the eight (08) values of network's parameters. As application we carried out a classification of the image of Algiers's bay taken by SPOT XS. The results of this study presented in the forms of curves, tables and images show that modified fuzzy ARTMAP presents the best compromise quality/computing time.

Keywords: Neural Networks, fuzzy ART, fuzzy ARTMAP, Remote sensing, Multispectral classification.

1 Introduction

Fuzzy ARTMAP Systems [1], [2], [3], [4] are neural networks based on knowledge (networks with supervised training), while the fuzzy ART systems [5] with not supervised training use data and operators of the logical fuzzy. These networks take a best place among the multitude of connectionist networks because of their aptitudes to solve problems which can be described by partially correct and/or incomplete data [6],

[7]. Their disadvantage is that they have too many parameters to be fixed correctly to make them converge towards the desired solution.

To supervise the fuzzy ART, we propose to vary its parameters with a step, to carry out for each case the (un supervised) training of the basis, to calculate the distance between the outputs obtained and wished, then to choose the parameters which gives the best distance.

The difficulty of the choice of the parameters of the neural fuzzy ARTMAP is solved, in a first technique, by the application of the process describes previously, and in a second technique by leaving this network in training as long as the objective is not achieved or that its architecture remains in evolution. The Fuzzy ART and fuzzy ARTMAP modified are presented in sections 4 and 6. The root mean square error, the training time and the rate of well classified points on a basis of control are used to evaluate performance.

The objective being the classification of the multi spectral image SPOT XS of Algiers's bay, result of the classifications and the experimental results of the study comparison are presented in section 7. Section 8 contain the conclusions of study.

2 Data and site of study

The site of study is the bay of Algiers which geographical co-ordinates are: (36° 39' 00 N, 36° 51' 00 N) and (3° 00' 30 E, 3° 16' 20 E). The data used represent a multi spectral image (XS1, XS2 and XS3) provided by HRV of SPOT sensor. Image was taken on April 1, 1997. This image represents part of the Mediterranean in north, the city and the wearing of Algiers along the coast, the Baïnem drill in the west of the city, and the naked ground mainly in the south.

The image size is 1500 pixels x 1000 pixels on 3 bands. From image, we have extracted 252 representative samples of four classes (87 for class 1, 38 for class 2, 63 for class 3 and 64 for the class 4) which will be useful as training bases, and 217 other samples (69 for class 1, 31 for class 2, 52 for class 3 and 65 for class 4) for the control of the studied neural classifiers.

3 The fuzzy ART network

The fuzzy ART network (Adaptive Resonance Theory) (Figure.1) is an unsupervised neural network. It propose a categorization with classes in hyper right-angled, each one represents a prototype (weight of the neuron). It is composed of three layers [5]:

1. A layer F0 (layer where the data are prepared) receiving the bodies of the vectors *a* (fuzzy input), it has a double number of nodes according to the size of the vector *a*, because of complement coding. Thus we generate the vector $I = (a, a^c)$.

2. A layer F1 for comparison, having the same number of nodes than F0. Each node of F1 is related to the same order of F0's node by a weight equal to one.

3. A layer F2 for competition entirely inter-connected with F1. Each node j of F2 is connected with all nodes of F1. The adaptive weight associated to the vector is noted W j. The vector T expresses the activation of F2.

The dynamics of the fuzzy ART network depends [6], [7] on the choice of the α parameter ($\alpha > 0$ used at the time of the competition between neurons in F2), the training parameter $\beta \in [0\ 1]$ fixing the speed of training, and the vigilance parameter $\rho \in [0\ 1]$ of defining the size of the right-angled hyper.

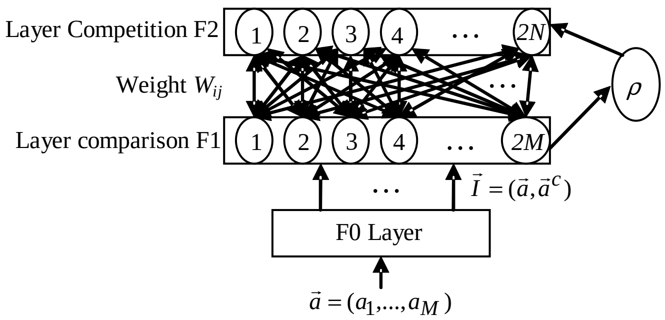


Figure 1: Neural network of fuzzy ART

The Fuzzy ART neural network algorithm used in this work is summarized below:

Step 1: Initializing the weights w_{ij} with one, ρ = [0 1], β=[0 1], and α>0.

Step 2: For each example, coding the input in complement, $I = (a, a^c)$

Step 3: calculate the Tj activity of each neuron of F2 by

$$T_j(I) = \frac{|I \wedge W_j|}{\alpha + |W_j|} \tag{1}$$

Where ^ the fuzzy intersection, is given by $(p \wedge q)_i = \min(p_i, q_i)$ and the norm |.| by : $|p| = \sum_i |p_i|$

Step 4: The neuron J having the highest activation T_j is selected like the winner neuron (Competition).

Step 5: Test of vigilance. It is carried out by checking (2).

$$\frac{|I \wedge W_J|}{|I|} \geq \rho \tag{2}$$

If the vigilance test is respected then the neuron J is updated (step 6). If not, this neuron is inhibited and another competition (step 4) takes place until a winning neuron respects the test of vigilance, or there is not active neurons (saturated network).

Step 6: The winner neuron is updated; these new weights are calculated by (3), and we reactive all neurons

$$W_J^{new} = \beta(I \wedge W_J^{old}) + (1-\beta)W_J^{old} \tag{3}$$

4 Modified Fuzzy ART

The fuzzy ART network is an unsupervised training network. The choice of vigilance parameters and training are strongly influences the result. To control the outputs so as to make them comparable with the desired outputs (to return it supervised), this paper proposes to find in the field of possible values of these parameters, those giving the best results.
The idea consist of varying α, ρ and β between 0 and 1 with a step λ, to carry out the training of fuzzy ART for each triplet, to calculate the root mean square error (MSE) between the outputs obtained by the network and desired output of the training base, and to take the triplet (α, ρ, β) giving the smallest MSE If this one is considered to be acceptable, if not we decrease the variation of step λ and remake the training. The modified Fuzzy ART algorithm is summarized below:

Step 1: We fixe the variation of step λ.

Step 2: For ρ going from 0 to 1 with a variation of step λ,

Step 2.1: For α going from 0 to 1 with a variation of λ,

Step 2.2: For β going from 0 to 1 with a variation of λ

To carry out the training of fuzzy ART, calculate MSE between the outputs obtained and the outputs of the training base. Take the best MSE and the associated parameters (α, ρ, β).

Step 3: If MSE obtained is not satisfactory, to decrease λ and go to step 2.

5 Fuzzy ARTMAP Network

The fuzzy ARTMAP network [8] is a supervised training neural network (the training is controlled by a base of examples, where each example is an association of a vector of input to a vector of desired output). Its architecture is evolutionary, and it is composed of two fuzzy ART networks [9], [10], [11], ARTa and ARTb. These two networks are bound by a network of a neural cells MAP (Figure.2).

ARTa receives the bodies of the vectors input of the examples, and ARTb receives the associated vector of desired output. Each fuzzy ART module has three layers:

- The coding layers F0 which generates the vector $A=(a, a^c)$ in ARTa and $B=(b, b^c)$ in ARTb. For reasons of simplification of the writing, we note I vector A or B according to the vector of input of ARTa or ARTb.

- The vector X (x^a for ARTa and x^b for ARTb) expresses the activation of F1.

- The vector of the adaptive weights binding F1 and F2 is noted W j (Wj^a for ARTa, and Wj^b for ARTb). The vector y (y^a for ARTa, and y^b for ARTb) expresses the activation of F2.

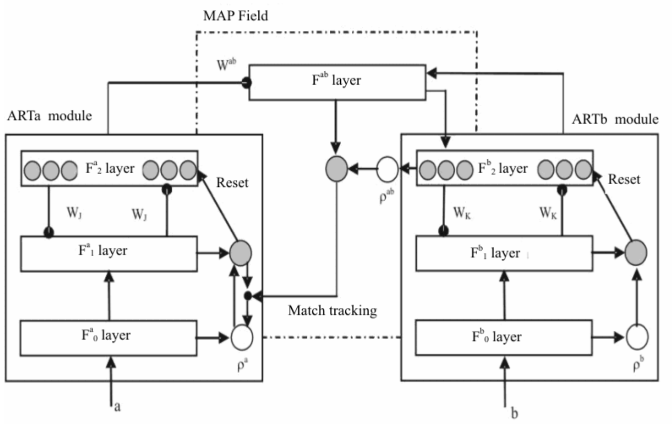


Figure 2: Block diagram of fuzzy network ARTMAP

The fuzzy ARTMAP [2], [3] has in addition to the three parameters of each fuzzy ART, three other parameters which are: The minimum value of the parameter of vigilance of ARTa noted $\overline{\rho}^a$, the vigilance parameter ρ^{ab} and the training parameter β^{ab} of layer MAP.

The phase of training of the fuzzy ARTMAP network consists of an adaptation of the architecture (numbers of cells of F2a, F2b and MAP) and of an update of the weights of various established connections [1], [3], [6]. This evolution of the architecture network rises owing to the fact that in this phase of training a winning neuron in each ART is required (competition), it is then compared with the vector of input (comparison). If this comparison is conclusive in each ART (the result of the comparison is higher than a threshold - criterion of vigilance -) an update of the weights is carried out. In the negative case a new research in corresponding ART is made. If no winning satisfies the criterion of vigilance, a neuron is created in the F2 layer of corresponding ART, and a layer or a neuron by layer (according to the evolution of architecture were done in ARTa or ARTb) is added in the MAP. Thus the algorithm of training of network ARTMAP [2] is presented as follows:

For each example (*a*, *b*) (*a* being the input vector of the example, and *b* its associated output) of the training base, we make the following step:

Step 1: Presentation of the inputs,

$$\rho^a = \overline{\rho}^a \quad , x^a = A \quad , x^b = B. ,$$

Step 2: Activation of F2

Step 2.1: Selection of a category.

For each module ART, we calculate the *Tj* activation (the degree with which the weight vector Wj is a subset of input *I*) for each node j of F2. Then we choose the node which has the highest value, it is considered as the winning neuron or the category (only one neuron can be faded for each input). T_j is defined by equation 1.

Step 2.2 For each node selected in step 2.1 (J in F2a and K in F2b), we calculate the function m_j (degree with which the input is a subset of the prototype W_j):

$$m_j(I) = \frac{\left|I \wedge W_j\right|}{\left|I\right|} \tag{4}$$

If this function for the node J and/or K is higher or equal to the criterion of vigilance ρ (ρ^a et/ou ρ^b) it will be supposed that there is resonance and that the respective F2 layer is activated: Thus for ARTa $y_J = 1$ and $y_j = 0$ for all $j \neq J$ (same for ARTb). In the contrary case, we return to step 2.1 and we select a new node in the respective module.
If no category could be chosen, one (several) new node(s) is (are) created dynamically. We note J and/or K this (these) node(s).

$$\begin{aligned} &W_j^a = 1,\ W_J^{ab} = 1,\ y_J^a = 1 \ \text{and} \ y_j^a = 0 \ \text{for} \ j \neq J \\ &\text{and / or} \\ &W_K^b = 1,\ W_K^{ab} = 1,\ y_K^b = 1 \ \text{and} \ y_k^b = 0 \ \text{for} \ k \neq K \end{aligned} \tag{5}$$

Step 3: Vigilance test in MAP layer. In this layer, we calculate

$$x^{ab} = W_J^{ab} \wedge y^b .$$

If $|x^{ab}| / |y^b| \geq \rho^{ab}$, we go to step 4. If not, go to step 5

Step 4: Training or update of the weights. w_j^a w_k^b Are updates following (3), and w_j^{ab} is update as follow:

$$W_j^{ab(new)} = \beta(y^b \wedge W_j^{ab(old)}) + (1-\beta)W_j^{ab(old)} \tag{6}$$

It is to be noticed that the fast training is obtained for the βvalue equal 1 (β= 1) in each layer.

Step 5: Change the criterion of vigilance of ARTa.
We done $\rho^a = m_J^a(A) + \varepsilon \ \text{and} \ T_J = 0$,
And go again at step 2.

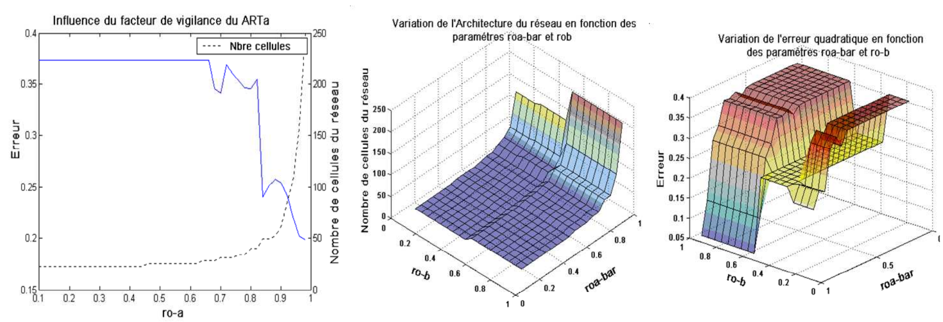
6 MODIFIED FUZZY ARTMAP

The fuzzy ARTMAP network has too many parameters to be fixed to reach a rate of reasonable training. These parameters are: vigilance coefficient and training coefficient of ARTa, ARTb and the MAP, and the comparison coefficient of ARTa and ARTb.

To highlight the difficulty of the choice of these parameters, we present the curves of error and the number of cells of the network according to the $\overline{\rho}^a$ parameters in (Figure 3a), and to ($\overline{\rho}^a, \rho^b$) parameters in (Figure 3b). The others parameters are fixed as follow:

$$\rho^b = 0.1, \rho^{ab} = 0.1, \beta^a = 1, \beta^b = 1, \beta^{ab} = 1, \alpha^a = 0, \alpha^b = 0$$

By the existence of the significant number of combinations, it appears clearly that it is very difficult to find the adequate values of these parameters to make the network convergent.



(a) Influence of $\overline{\rho}^a$ (b) influence of $(\overline{\rho}^a, \rho^b)$

Figure 3: Influence of the parameters on the MSE and the architecture of the ARTMAP.

A first idea consists, as for the fuzzy ART network, to vary the eight parameters of the fuzzy ARTMAP between 0 and 1 with a λ step, to carry out the training of the 8-uplet, to calculate the root mean square error (MSE) between the outputs obtained by the network and the desired outputs of the training base and to retain the 8-uplet offering the smallest MSE if this one is considered to be acceptable, if not we decrease the step λ and to remake the training. We will call this manner of process ARTMAP to research of the optimal parameters.

It is obvious that this manner of process will gave the best results according to the imposed distance, but the time of training is much more significant.

An improvement of training algorithm of the fuzzy ARTMAP was proposed [12]. It consists to do not pass the whole of the examples of the base of training only once as it is of use, but as many time as the network is in architectural evolution (i.e. until stability of the architecture network) or than the fixed error is not reached. This improvement is the consequence to the fact that the algorithm of training of fuzzy ARTMAP network as described in section 5, makes pass the examples one by one, and for each example an update of architecture and/or weights are carried out. Between the passage of an example being at the beginning of the base and the end of the training, the network will be strongly modified if the parameters are badly chosen, this modification influences negatively the degrees of training of the first examples. We will call this manner of process the modified ARTMAP.

7 RESULTS

In addition to the modifications of the fuzzy ART and fuzzy ARTMAP networks, the goal of this paper is to compare these networks and to evaluate the performances of each modification suggested. This comparison is carried out for a classification of the multi spectral image SPOT XS of Algiers's bay (section 2). The networks have learned on the training basis of and evaluated on the control basis (section 2).

The fuzzy ART network with fixed architecture has a F0 layer with six cells (three fuzzified entries and their complement) and a F2 layer of four cells (a cell by class). To be able to determine the values of the best network, we proceed as proposed in section 4. A study of the parameters according to the step value λ was carried out. Table 1 includes the minimal quadratic training error, the associated parameters as well as the control root mean square error It is to be noted that the root mean square errors at training and control are proportional to the λ step. The minimal error is obtained for a value of λ equal to 0.01.

These errors correspond to the values of the parameters α = 0.69, β = 0.11 and ρ = 0.93 of the fuzzy ART network offering the best classification. The mean square error of this network on the control basis is equal to 0.07.

Table 1: Evaluation of the Modified Fuzzy ART according to λ

λ	α- β- ρ Parameters	Learning MSE	Control MSE
0.01	0.69 - 0.11 - 0.93	0.1038	0.0699
0.05	0.7 - 0.15 - 0.9	0.0793	0.1038
0.1	0.7 - 0.9 - 0.8	0.1011	0.0992
0.2	0.8 - 0.8 - 0.8	0.1329	0.1313
0.5	0 - 1 - 0	0.3710	0.3456

ARTMAP being with evolutionary architecture during the training phase, only the numbers of cells of the F0 layers of ARTa and ARTb are fixed at six (three fuzzified entries and their complements) and eight (four classes and their complements) respectively. The research of the optimal parameters was carried out like previously. For each variation value λ of eight parameters, we carry out the training of the network and compute, for the various possibilities, the minimal mean square error. The graphs of Figure 4 illustrate the minimal error according to the λ step.

When the step λ is weak, the minimal root mean square error is small and better is the convergence of the network, but it is clear that the number of combinations in this case is more significant. A better root mean square error (RMSE) equal to 0.021 is obtained for a value λ equal to 0.05. Parameters giving this performance are: ρ^a=0.95, ρ^b=0.55, ρ^{ab}=0 .1, β^a=0 .8, β^b=0.1, β^{ab} = 0.1, α^a = 0.4, α^b=0.1.The number of combinations in this case are more significant, and the computing time is larger. Generalization to the control basis gave an RMSE equal to 0.036. Although this error is higher than that of the training, this result is satisfactory.

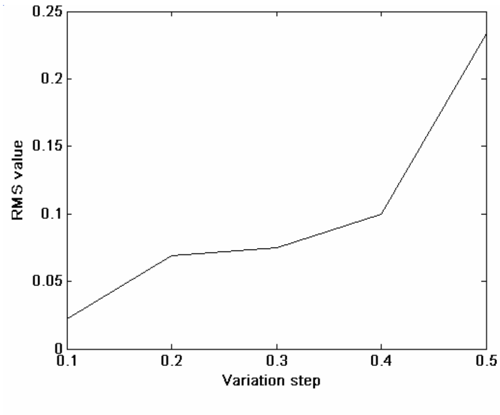


Figure 4: Minimal error MSE versus λ

The study undertaken the same basis but by making it spend several times (modified fuzzy ARTMAP) shows in figure 5 that the error decrease overall according to the passage (iteration) of the base.
We fixed the vigilance and training parameters to 0.75. The comparison parameter is selected weak equal to 0.5. The value 0.019 is obtained as final RMSE .Root mean square error on the control basis is equal to 0.035.

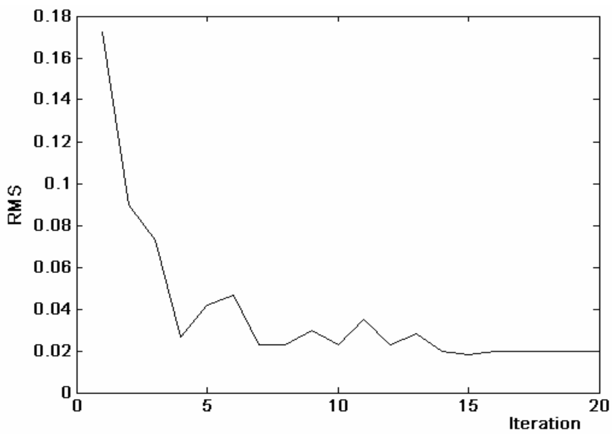


Figure 5: MSE according to the iteration of the modified fuzzy ARTMAP

The RMS error at control being a global criterion of evaluation, we suggest to estimate the quality of classification for the proposed network. For each class, we consider the rates of the well classified points of the control base. Figure 6 represents the result.

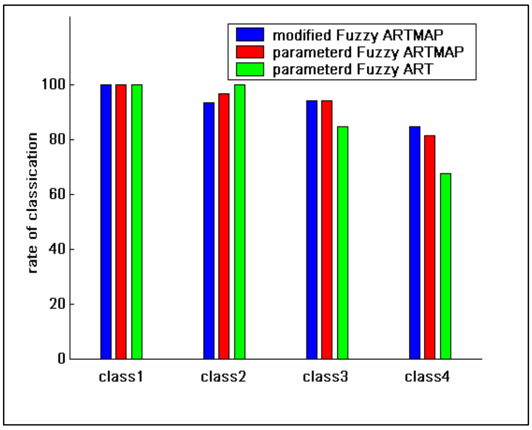


Figure 6: Rates of the well classified points per class of the control base

It is clear that with fuzzy ART and the fuzzy ARTMAP with research of the optimal parameters, the result is much better than the step λ is weak, but this with a significant computing time. The solution consisting to learn the fuzzy ARTMAP while making pass the base of training several times without worrying too much about parameters (modified fuzzy ARTMAP) offers the best compromise classification quality / computing time (Table 2).

Table 2: Performances of three techniques

Techniques	MSE Training	MSE Control	Accuracies classification
Parameterized Fuzzy ART	0.059	0.066	86.64%
Parameterized Fuzzy ARTMAP	0.021	0.036	92.63%
Modified Fuzzy ARTMAP	0.019	0.035	93.09%

For the supervised fuzzy ART, the training time is of $(1/\lambda)^3$ time of training time of the traditional network while for the fuzzy ARTMAP with search of optimal parameters this time is of $(1/\lambda)^8$ time of training time of a traditional ARTMAP for the same application. For the same application, modified ARTMAP produce a time proportional to the number of passage (much weaker than the firsts) which is multiply by the training time of a traditional ARTMAP. Figure 7 shows the generalization of the ARTMAP modified on the SPOT XS image of Algiers's bay.

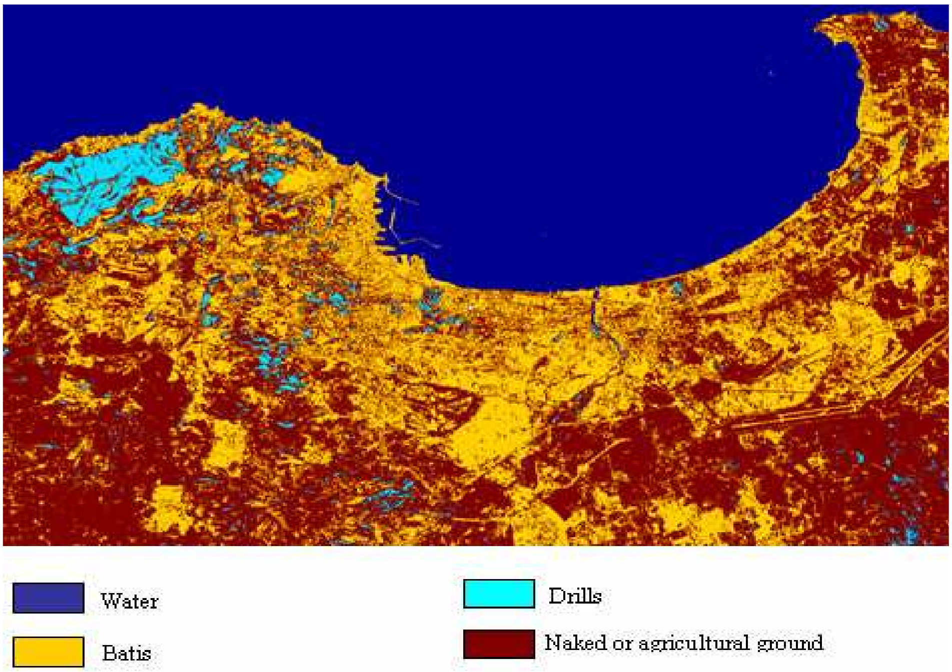


Figure 7: Classification of the Spot XS image of Algiers's bay by the modified fuzzy ARTMAP

8 Conclusion

The fuzzy ART network is an unsupervised network. To return it supervised, this study proposes to look for parameters which offer outputs closest to the training base in meaning of the root mean square distance. We propose parameters depending of a step λ; the result is much better than the step is weak and the time is proportional to $(1/\lambda)^3$.
The fuzzy ARTMAP network has too many parameters to be fixed to reach a rate of reasonable training. The difficulty of the choice of these parameters led us, in a first solution, to vary its eight parameters with a step λ, to make the training for each 8-uplet, and to choice the network with the parameters offering the best result. This solution is viable; it is much better than the step is weak, but very greedy in computing times (proportional to $(1/\lambda)^8$).
The second solution consisting in making as many pass the base of training time as the objective error is not reached or architecture remains in evolution, without worrying too much about the parameters, gave the best compromise quality of classification/computing time.

References

[1] A. H Tan, Cascade ARTMAP Integrating Neural Computation and Symbolic Knowledge Processing. *IEEE Transactions on Neural Networks*, Vol.8 No2, 1997.

[2] I. Dagher, M. Georgiopoulos, G. L. Heileman, G. Bebis,, An Ordering Algorithm for Pattern Presentation in Fuzzy ARTMAP That Tends to Improve Generalization Performance, *IEEE Transactions on Neural Networks*, Vol. 10 No. 4,1999

[3] D. Charalampidis, T. Kasparis, M. Georgiopoulos, Classification of Noisy Signals Using Fuzzy ARTMAP Neural Networks, *IEEE Transactions on Neural Networks*, Vol. 12, No. 5, 2001

[4] R. K. Aggarwal, Q. Y. Xuan, A. T. Johns, A Bennett, A Novel Approach to Fault Diagnosis in Multicircuit Transmission Lines Using Fuzzy ARTmap Neural Networks, *IEEE Transactions on Neural Networks*, Vol. 10, No. 5, 1999.

[5] G. A Carpenter, S. Grossberg, D. B. Rosen, Fuzzy ART : An Adaptive Resonance Algorithm for Rapid, stable classification of Analog Patterns, Proc. *International Joint Conference of Neural Networks*, IJCNN'91, Seattle, IEEE/INNS Inc, pp. II-411-416,1991.

[6] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, D. Rosen, Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps, *IEEE Transactions on Neural Networks*, Vol. 3 No. 5, 1992.

[7] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, A Fuzzy ARTMAP Nonparametric Probability Estimator for Nonstationary Pattern Recognition Problems, *IEEE Transactions on Neural Networks*, Vol. 6 No. 6, 1995.

[8] E. Gómez-Sánchez, Y. A. Dimitriadis, J. M. Cano-Izquierdo, J. López-Coronado, µARTMAP: Use of Mutual Information for Category Reduction in Fuzzy ARTMAP, *IEEE Transactions on Neural Networks*, Vol. 13, No1, 2002.

[9] T. Frank, K. F. Kraiss, T. Kuhlen, Comparative Analysis of Fuzzy ART and ART-2A Network Clustering Performance, *IEEE Transactions on Neural Networks*, Vol. 9, No. 3, 1998.

[10] P. Baldi, Gradient Descent Learning Algorithm Overview: A General Dynamical Systems Perspective, *IEEE Transactions on Neural Networks*, Vol. 6 No. 1, 1995.

[11] D. Charalampidis, G. C. Anagnostopoulos, M. Georgipoulos, T. Kasparis, Fuzzy ART and Fuzzy ARTMAP with adaptively weighted distances, *Proc. SPIE*, Vol 4739 pp 86-97, 2002.

[12] F. Alilat, S. Loumi, H. Merrad , B. Sansal, Nouvelle approche du réseau ARTMAP Flou Application à la classification multispectrale des images SPOT XS de la baie d'Alger *Revue Française de Photogramétrie et de Télédétection* SFPT, (2005-1),.no.177, pp 17-24, 2005.