Modified Fuzzy ARTMAP and Supervised Fuzzy ART: Comparative Study with Multispectral Classification

F. Alilat, S. Loumi, H. Merrad and B. Sansal

Abstract—In this article a modification of the algorithm of the fuzzy ART network, aiming at returning it supervised is carried out. It consists of the search for the comparison, training and vigilance parameters giving the minimum quadratic distances between the output of the training base and those obtained by the network. The same process is applied for the determination of the parameters of the fuzzy ARTMAP giving the most powerful network. This modification consist in making learn the fuzzy ARTMAP a base of examples not only once as it is of use, but as many times as its architecture is in evolution or than the objective error is not reached. In this way, we don’t worry about the values to impose on the eight (08) parameters of the network. To evaluate each one of these three networks modified, a comparison of their performances is carried out. As application we carried out a classification of the image SPOT XS of Algiers’s bay taken by SPOT XS. We use as criterion of evaluation the training duration, the mean square error (MSE) in step control and the rate of good classification per class. The results of this study presented as curves, tables and images show that modified fuzzy ARTMAP presents the best compromise quality/computing time.

Keywords—Neural Networks, fuzzy ART, fuzzy ARTMAP, Multispectral Classification

I. 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 unsupervised 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 become the supervised fuzzy ART, we propose to vary its parameters with a step, to carry out for each case the (unsupervised) training of the basis, to calculate the distance between the outputs obtained and the 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 phase of 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 3 to section 6. To be able to evaluate the performances of these three techniques a comparison was carried out. The follows criterions of evaluations used are: The mean square error (MSE), the training duration and the rate of well classified points on a basis of control.

The objective being the classification of the multi spectral image SPOT XS of Algiers’s bay, results of the classifications carried out by the three networks put in competition, as well as the experimental results of the comparison of their performances are presented in section 7. Section 8 contains the conclusions of this study.

II. 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 port of Algiers along the coast, the Bâinem 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 this image, we have extracted 252 samples of four classes (87 for class 1, 38 for class 2, 63 for class 3 and 64 for 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.

III. THE FUZZY ART NETWORK

The fuzzy ART network (Adaptive Resonance Theory) (Fig. 1) is an unsupervised neural network. It proposes to a categorization with class in hyper right-angled. Each one represents a prototype (weight of the neuron). It is composed of three layers [5]:

Manuscript received March 28, 2006. This work was supported by the Sciences and Technology University of Houari Boumediene USTHB.

Farid Alilat, Saïlha Loumi, Hamoud Merrad, & Boualem Sansal are with Sciences and Technology University of Houari Boumediene. Faculté de Génie Electrique et d’Informatique, Laboratoire de Traitement d’Images et Rayonnement. BP 32, El-Alia Bah - Ezouzur, 16111, Alger Algérie. (Fax: +213 (21) 24-76-07 or +213 (21) 24-71-87. e-mail : falilat@yahoo.fr, salihaloumi@yahoo.fr, merradhamoud@yahoo.fr
• 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, and its complement. Thus we generate the vector I = (a, a').

• 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.

• 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 \( \alpha \) 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.

### Algorithm of Fuzzy ART

**Step 1:** To initialize the weights \( w_{ij} \) with one, \( \rho = "0-1", \beta = "0-1", \) and \( \alpha = 0.\)

**Step 2:** For each example, to generate the input I, I = (a, a')

**Step 3:** To calculate the \( T_j \) activity of each neuron of F2 by

\[
T_j(I) = \frac{|I \land W_j|}{\alpha + |W_j|}
\]

Where \( \land \) is the fuzzy intersection given by \( (p \land q)_i = \min (p_i, q_i) \) 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 is carried out by checking (2).

\[
\frac{|I \land W_j|}{|I|} \geq \rho
\]

If the test is respected then the neuron J is updated (step 6).

If not, this neuron is desactivated 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).

\[
W_j^{\text{new}} = \beta(I \land W_j^{\text{old}}) + (1 - \beta)W_j^{\text{old}}
\]

And activate again all neurons.

### IV. MODIFIED FUZZY ART

The fuzzy ART network is an unsupervised training network. The choice of vigilance parameters and training are strongly influence 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 network which gave the best results.

The idea consist to vary \( \alpha, \rho \) and \( \beta \) between 0 and 1 with a step \( \lambda \), to carry out the training of fuzzy ART for each triplet, to calculate the mean square error (MSE) between the outputs obtained by the network and desired output of the training base, and to retain the triplet \( (\alpha, \rho, \beta) \) giving the smallest MSE, if this one is considered to be acceptable, if not to decrease the variation of step \( \lambda \) and to remake the training.

**A. Algorithm of Modified Fuzzy ART**

**Step 1:** To fix the variation of step \( \lambda \).

**Step 2:** For \( \rho \) going from 0 to 1 with a variation of step \( \lambda \), carry out step 2.1

**Step 2.1:** For \( \alpha \) going from 0 to 1 with a variation of \( \lambda \), carry out step 2.2

**Step 2.2:** For \( \beta \) going from 0 to 1 with a variation of \( \lambda \), carry out step 2.2

**Step 3:** If MSE obtained is not satisfactory, to decrease \( \lambda \) and to remake starting from step 1.

### V. 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 an input vector to a desired output vector). 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 (Fig. 2).

ARTa receives the bodies of the vectors of 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 \( A = (a, a') \) in ARTa and \( B = (b, b') \) in ARTb. For reasons of simplification of the writings, let us note I vector A or B according to whether it is about 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 $W_j$ ($W_{ja}$ for ARTa, and $W_{jb}$ for ARTb). The vector $y$ ($y_a$ for ARTa, and $y_b$ for ARTb) expresses the activation of F2.

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 $\rho_a$, the vigilance parameter $\rho_{ab}$ and the training parameter $\beta_{ab}$ of layer MAP.

The training phase of the fuzzy ARTMAP network consists of an adaptation of its architecture (numbers of cells of F2a, F2b and MAP) and an updating of the weights of various established connections [1], [3], [6]. This evolution of the architecture network result to the fact that in this training phase a winner 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 contrary case, to return at step 2.1 and 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) (5)

$W_{ja} = 1$, $W_{jb} = 1$, $y_j = 1$ and $y_j = 0$ for $j \neq J$ and/or $K$

Step 2: Activation of F2

Step 2.1: Selection of a category.

For each module ART, we calculate the $T_j$ activation (the degree with which the weight vector $W_j$ is a subset of input $I_j$) for each node j of F2. Then we choose the node which has the greatest value. It is the winner neuron or the category (only one neuron can be faded for each input). $T_j$ is defined by (1)

$$ m_j = \left[ \frac{1}{|I \land W_j|} \right] $$

If this function for the node J and/or K is higher or equal to the criterion of vigilance $\rho$ ($\rho_a$ et/ou $\rho_b$) 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, to return at step 2.1 and select a new node in the respective module.

VI. 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 $\rho^a$ parameters in
(Fig. 3a), and to \((\beta^a, \rho^b)\) parameters in (Fig. 3b). The others parameters are fixed to:
\[
\rho^b = 0.1, \quad \rho^{ab} = 0.1, \quad \beta^a = 1, \quad \beta^{ab} = 1, \quad \alpha = 0, \quad \alpha^b = 0
\]

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 converge the network towards the desired result.

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 \(\lambda\) step, to carry out the training of the 8-uplet, to calculate the 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 which offer the smallest MSE, if this one is considered to be acceptable. If such is not the case, to decrease the step \(\lambda\) and to remake the training.

We will call this manner of process: “research of the optimal parameters of the ARTMAP” (parametrized ARTMAP).

It is obvious that this manner of process will done the network offering the best results within the meaning of the imposed distance, but it clear that 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 owing 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.

Table 1 includes the minimal mean square error (MSE) in training and control phase as well as the associated parameters.

It is obvious that the mean square errors at training and control are proportional to the \(\lambda\) step. The minimal error is obtained for a value of \(\lambda\) equal to 0.01. The corresponding parameters are \(\alpha = 0.69, \beta = 0.11\) and \(\rho = 0.93\) of the fuzzy ART network offering the best classification. In control phase, the mean square error is 0.07.

<table>
<thead>
<tr>
<th>(\lambda)</th>
<th>(\alpha), (\beta), (\rho) Parameters</th>
<th>Learning MSE</th>
<th>Control MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.69 - 0.11 - 0.93</td>
<td>0.1038</td>
<td>0.0699</td>
</tr>
<tr>
<td>0.05</td>
<td>0.7 - 0.15 - 0.9</td>
<td>0.0793</td>
<td>0.1038</td>
</tr>
<tr>
<td>0.1</td>
<td>0.7 - 0.9 - 0.8</td>
<td>0.1011</td>
<td>0.0992</td>
</tr>
<tr>
<td>0.2</td>
<td>0.8 - 0.8 - 0.8</td>
<td>0.1329</td>
<td>0.1313</td>
</tr>
<tr>
<td>0.5</td>
<td>0 - 1 - 0</td>
<td>0.3710</td>
<td>0.3456</td>
</tr>
</tbody>
</table>

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 to one) 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 \(\lambda\) was carried out.

VII. RESULTS
ARTa and ARTb are fixed at six (three fuzzified input and their complement to one) and eight (four classes and their complement) respectively. The research of the optimal parameters was carried out like previously.

For each variation of value \( \lambda \) of eight parameters, we carry out the training of the network and compute, for the various possibilities, the minimal mean square error. The Figure 4 illustrates the minimal error versus of the \( \lambda \) step.

When \( \lambda \) is weak, the minimal 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. Indeed a better mean square error (equal to 0.021) is obtained for a value \( \lambda \) equal to 0.05. Parameters giving this performance are: \( \rho_a=0.95, \rho_b=0.55, \rho_{ab}=0.1, \beta_a=0.8, \beta_b=0.1, \beta_{ab}=0.1 \). The number of combinations in this case are 218, and the computing time is important. We obtain 0.036 for mean square error in the control phase. Although this error is higher than that of the training, this result is satisfactory.

The study undertaken on the same basis but by making it spend several times (modified fuzzy ARTMAP) shows (Fig.5) that the error globally decrease according to the passage (iteration).

We fixed the vigilance and training parameters to 0.75. The comparison parameter is selected weak equal to 0.5. 0.019 and 0.035 are respectively mean square error obtained in training and control phases.

The criterion of the mean square error at control being a global criterion of evaluation, we suggest a better estimation by comparing the quality classification obtained by these three types of networks suggested. We represent (Fig.6) the rates of the well classified points per class in the control base.

It is clear that for fuzzy ART and fuzzy ARTMAP with optimal parameters, the result is much better than the step \( \lambda \) is weak. The computing time is more important. So the solution consisting in making learn the fuzzy ARTMAP while making pass the base of training several times without worrying to the parameters (modified fuzzy ARTMAP) offers the best compromise classification quality/computing time (Table II). For the supervised fuzzy ART, the training time is of \( (1+1/\lambda)^3 \) time of training time of the traditional network. For the fuzzy ARTMAP with search of optimal parameters this time is of \( (1+1/\lambda)^8 \) time of training time of a traditional ARTMAP for the same application. But at the same application, for the modified ARTMAP, this time is proportional to the number of passage (much weaker than the firsts) which is multiply by the training time of a traditional ARTMAP.

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Training MSE</th>
<th>Control MSE</th>
<th>Accuracies in classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parametrized Fuzzy ART</td>
<td>0.059</td>
<td>0.066</td>
<td>86.64%</td>
</tr>
<tr>
<td>Parametrized Fuzzy ARTMAP</td>
<td>0.021</td>
<td>0.036</td>
<td>92.63%</td>
</tr>
<tr>
<td>Modified Fuzzy ARTMAP</td>
<td>0.019</td>
<td>0.035</td>
<td>93.09%</td>
</tr>
</tbody>
</table>

The generalization of the ARTMAP modified on the SPOT XS image of bay of Algiers is illustrated by Fig.7.
VIII. CONCLUSION

The fuzzy ART network is an unsupervised network. To return it supervised, this study proposes to seek its parameters which offer the network giving the outputs closest to the training base within the meaning of the mean square distance. Parameter varying with a step $\lambda$, the result is of as much better than the step is weak, but time is proportional to $(1+1/\lambda)^3$.

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 $\lambda$, to make the training for each 8-uplet, and to keep the network with the parameters offering the best result. This solution is viable; it is of as much better than the step is weak, but very greedy in computing times (proportional to $(1+1/\lambda)^8$).

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 to the parameters, gave the best compromise quality of classification/computing time.

REFERENCES


Farid Alilat has received his diploma in electronics engineer at the University of Sciences and technology Houari Boumediene (USTHB) in 1988, and magister diploma in electronics in June 1992, Algeria. He is currently completing these of state doctorate. He is charged of research and lecturer at this university (USTHB) since 1993. He has done research in classification, pattern recognition, clustering, remote sensing, and data compression, artificial intelligence. His present research interests lie in applied mathematics, image processing and neural networks. Actually, F. Alilat is a reviewer for the International Journal of Remote Sensing (IJRS).

Saliha Loumi has received her diploma in electronics engineer from polytechnic school Algiers in 1983, and magister diploma in electronics in November 14, 1889. She is currently completing these of state doctorate at the University of Sciences and technology Houari Boumediene (USTHB). She is lecturer and charged of research 1992 à USTHB University. She has done research in pattern recognition, multi-image processing, remote sensing, texture analysis and data compression. Her present research interests lie in applied mathematical morphology, image processing.

Hamoud Merrad received his diploma in electronic engineer from the National Polytechnic school of Algeria in February 1986 and magister diploma in electronics in November 2005. He is currently a researcher and a lecturer at the University of Sciences and technology Houari Boumediene (USTHB). His research interests include remote sensing, methodologies for the determination of atmospheric and marine properties.

Sansal Boualem has received her professor diploma in electronics engineer from the university of England. In 1980, he has worked for the Nuclear Center of Research, in Algeria and in 1989 he joined the faculty of department of electronic and informatics of USTHB University. Prof. Sansal is responsible for development of GIP (Group Image Processing) and, a director of Image processing and radiation laboratory of Algeria University.