

Nouvelle Approche du Réseau Artmap Flou : Application a la Classification Multi Spectrale des Images Spot XS de la Baie d'Alger *

Farid ALILAT, Saliha LOUMI, Hamoud MERRAD et Boualem SANSAL

Université des Sciences et de la Technologie Houari Boumediene,
Faculté de Génie Electrique et d'Informatique, Institut d'Electronique,
Laboratoire de Traitement d'Images et Rayonnement Atmosphérique,
BP 32, El-Alia Bab - Ezzouar, 16111, Alger Algérie
Télécopieur : +213 (21) 24-76-07 et +213 (21) 24-71-87

Email: falilat@usthb.dz ou faralilat@yahoo.fr, salihaloumi@yahoo.fr, lunerrad@usthb.dz

Résumé: Dans cet article nous proposons une modification de l'algorithme d'apprentissage du réseau de neurones ARTMAP (Adaptive Resonance Theory) flou. Cette modification consiste à faire apprendre au réseau une base d'exemples non pas une seule fois comme c'est d'usage, mais autant de fois que son architecture est en évolution ou que l'objectif erreur n'est pas atteint, et ceci sans trop se soucier des valeurs à imposer aux huit (08) paramètres du réseau ARTMAP flou. Jusqu'au jour d'aujourd'hui, il n'existe pas de méthodes universelles pour le choix de ces paramètres, leurs valeurs étant fixées suite à une série d'expériences. Pour évaluer les performances du réseau ainsi modifié, une comparaison entre ce dernier et un réseau PMC à rétro propagation du gradient optimisé par l'algorithme de Levenberg-Marquardt (LM) est réalisée. Comme application, nous avons effectué une classification de l'image multi spectrale SPOT XS de la baie d'Alger. Nous utilisons comme critère d'évaluation les temps d'apprentissage et de généralisation d'une part et l'erreur quadratique moyenne d'autre part. Les résultats de cette étude présentés sous forme de courbes, de tableaux et d'images montrent que la modification proposée contribue à une réduction importante du temps et de l'erreur d'apprentissage.

Mots-clés: Réseaux de neurones, PMC, Levenberg-Marquardt, ARTMAP flou, Télédétection, Classification.

Abstract: In this paper we suggest a modification of the algorithm of training of the fuzzy neural network ARTMAP. This modification consists in passing the base not only once time as usual but as many times as necessary.

At this time, the architecture of the network is in evolution to converge towards an imposed error. During this operation, we don't take into account of the choice of the eight parameters of the fuzzy neural network. Until now, there are not universal methods for the choice of these parameters, their values being fixed following a series of experiments. To evaluate the performances of the modified neural network, a comparison between this last and a neural network MLP with back propagation of the optimized gradient by the algorithm of Levenberg-Marquardt (LM) are realized. As application we carried out a classification of the image of the bay of Algiers taken by SPOT XS. We use as criterion of evaluation the time of training and generalization on the one hand and the average quadratic error on the other hand. The results of this study presented in the forms of curves, tables and images show that the suggested modification contributes to a significant improvement of time and error of training.

Keywords : Neural Networks, MLP, Levenberg - Marquardt, Fuzzy - ARTMAP, Remote Sensing, Classification.

1. Introduction

Les systèmes ARTMAP [Tan97] flous [Dagher99, Charalampidis99, Aggarwal 99] sont des réseaux de neurones basés sur la connaissance (réseaux à apprentissage supervisé), utilisant des données et des opérateurs de la logique floue. Ces réseaux ont trouvé une place de choix parmi la multitude des réseaux connexionnistes grâce à leurs aptitudes à résoudre des problèmes qui peuvent être décrits par des données partiellement correctes et/ou incomplètes [Carpenter92, 95].

* Article publié dans la revue Française de Photogrammétrie et de Télédétection n° 177 (2005-1), et repris dans le BSG Avec l'amiable autorisation de l'auteur.

Leur inconvénient est qu'ils possèdent beaucoup trop de paramètres à fixer correctement pour les faire converger vers la solution désirée. Pour palier cette difficulté, nous laissons, par la technique proposée dans cette étude, le réseau en phase d'apprentissage tant que l'objectif n'est pas atteint ou que son architecture reste en évolution. La présentation du ARTMAP flou et son amélioration fait l'objet de la section 3.

Pour pouvoir évaluer les performances du réseau ainsi modifié, nous l'avons comparé au perceptron multi couches (PMC) à apprentissage rapide à rétro propagation du gradient utilisant l'algorithme de Levenberg-Marquardt [Hagan94, Wilamowski01] que nous présentons dans la section 4. Les critères d'évaluation utilisés sont: l'erreur quadratique moyenne et les temps d'apprentissage et de généralisation.

L'objectif étant la classification de l'image multi spectrale SPOT XS de la baie d'Alger, les résultats des classifications réalisées par les deux réseaux mis en compétition, ainsi que les résultats expérimentaux de la comparaison de leurs performances sont présentés dans la section 5. La section 6 présente la conclusion de cette étude.

2. Données et site d'étude

Le site d'étude est la baie d'Alger de coordonnées géographiques 36° 39' 00 N à 36° 51' 00 N et 3° 00' 30 E à 3° 16' 20 E. Les données utilisées représentent une image multi spectrale (XS1, XS2 et XS3) fournie par le capteur HRV de SPOT (image prise le 1 avril 1997). Cette image représente une partie de la mer Méditerranée au nord, la ville et le port d'Alger le long de la côte, le bois de Baïnem à l'ouest de la ville, et de la terre nue ou agricole principalement au sud. La taille de l'image est de 1500 x 1000 pixels sur 3 bandes. De cette image, nous avons extrait 252 échantillons représentatifs des quatre classes (87 pour la classe 1, 38 pour la classe 2, 63 pour la classe 3 et 64 pour la classe 4) qui vont servir comme base d'apprentissage, et 217 autres échantillons (69 pour la classe 1, 31 pour la classe 2, 52 pour la classe 3 et 65 pour la classe 4) pour le contrôle des classifieurs neuronaux étudiés.

3. Le réseau ARTMAP flou

Le réseau ARTMAP flou est un réseau de neurones à apprentissage supervisé (l'apprentissage est contrôlé par une base d'exemples, où chaque exemple est une association d'un vecteur d'entrée à un vecteur de sortie désirée).

Son architecture est évolutive, il est composé de deux réseaux ART (Adaptive Resonance Theory) flou [Frank98, Baraldi02, Charalampidis02], ARTa et ARTb, ces deux réseaux sont liés par l'intermédiaire d'un réseau de cellules neuronales MAP (figure 1).

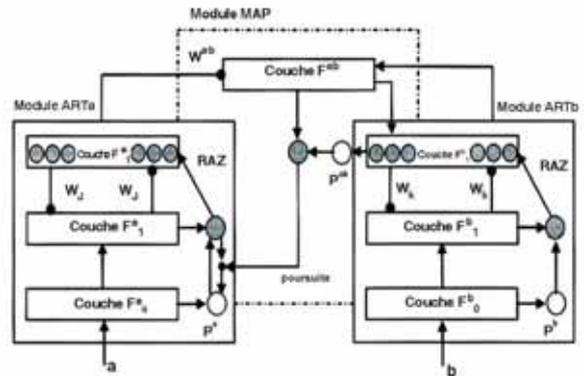


Fig. 1 Schéma bloc du réseau ARTMAP flou.

Chaque module d'ART flou est un réseau de neurones non supervisé composé de trois couches [Carpenter91]. Une couche F0 (couche de préparation des données), une couche F1 de comparaison et une troisième couche F2 de compétition.

L'ARTa reçoit les corps des vecteurs d'entrée des exemples, et l'ARTb reçoit le vecteur de sortie désirée associé.

Chaque module ART flou a trois couches:

- La couche de codage F0 possède un nombre de nœuds double de la taille du vecteur de l'entrée du module du fait d'un codage en complément, et génère le vecteur $A=(a,a^c)$ dans ARTa et $B=(b,b^c)$ dans ARTb, a et b étant des vecteurs flous (leurs éléments sont compris entre 0 et 1).

Pour des raisons de simplification des écritures, notons I le vecteur A ou B selon qu'il s'agisse du vecteur d'entrée du ARTa ou du ARTb.

- La couche de comparaison F1 possède le même nombre de nœuds que F0. Chaque nœud de F1 est lié au nœud de même ordre de F0 par un poids égal à un. Le vecteur X (x^a pour l'ARTa et x^b pour l'ARTb) exprime l'activation de F1.

- La couche de compétition F2 est entièrement interconnectée à F1. Un nœud j de F2 par exemple est connecté avec tous les nœuds de F1, le vecteur des poids adaptatifs associé est noté W_j (W_j^a pour l'ARTa, et W_j^b pour l'ARTb). Le vecteur Y (y^a pour l'ARTa, et y^b pour l'ARTb) exprime l'activation de F2.

La dynamique du réseau ART flou dépend [Carpenter92,95] du choix du paramètre $\alpha > 0$ (utilisé lors de la compétition entre neurones dans F2), du paramètre d'apprentissage $0 \leq \beta \leq 1$, et du paramètre de vigilance ρ ($0 \leq \rho \leq 1$).

L'ARTMAP flou possède en plus trois autres paramètres qui sont: La valeur minimum du paramètre de vigilance de l'ARTa noté $\bar{\rho}^a$, le paramètre de vigilance ρ^{ab} et le paramètre d'apprentissage β^{ab} de la couche MAP.

La phase d'apprentissage du réseau ARTMAP flou consiste en une adaptation de l'architecture (nombres de cellules de F2a, F2b et du MAP) et en une mise à jour des poids des différentes connexions établies [Tan97, Charalampidis01, Aggarwal99]. Cette évolution de l'architecture du réseau découle du fait que dans cette phase d'apprentissage un neurone gagnant dans chaque ART est recherché (compétition), il est ensuite comparé au vecteur d'entrée (comparaison). Si cette comparaison est concluante dans chaque ART ; le résultat de la comparaison est supérieur à un seuil (critère de vigilance) une mise à jour des poids est réalisée. Dans le cas contraire, une nouvelle recherche dans le ART correspondant est faite. Si aucun gagnant ne satisfait le critère de vigilance, un neurone est créé dans la couche F2 du ART correspondant, et une couche ou un neurone par couche (selon que l'évolution de l'architecture s'est faite dans l'ARTa ou l'ARTb) est rajouté dans le MAP.

Ainsi l'algorithme d'apprentissage du réseau ARTMAP flou [Carpenter92] se présente comme suit:

3.1 Algorithme

Pour chaque exemple (a,b) (a étant le vecteur d'entrée de l'exemple, et b sa sortie associée) de la base d'apprentissage, on réalise les étapes suivantes:
 Etape 1 : Présentation de l'entrée

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

Etape 2 : Activation de F2

Etape 2.1 : Sélection d'une catégorie.

Pour chaque module ART, calculer l'activation T_j (le degré avec lequel le vecteur poids W_j est un sousensemble de l'entrée I) pour chaque nœud j de F2. Choisir ensuite le nœud J qui a la plus grande valeur, il est considéré comme le neurone gagnant ou la catégorie (un seul neurone peut être altéré pour chaque entrée).

T_j est définie par:

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

Où \wedge est l'intersection floue donnée $(p \wedge q)_i = \min(p_i, q_i)$ et la norme $|\cdot|$ par:

$$|p| = \sum_i |p_i| \tag{2}$$

Etape 2.2: Pour chaque nœud pré sélectionné dans l'étape 2.1 (J dans F2a et K dans F2b), calculer la fonction m_j (degré avec lequel l'entrée est un sous-ensemble du prototype W_j):

$$m_j(I) = \frac{|I \wedge W_j|}{|I|} \tag{3}$$

Si cette fonction pour le nœud J et/ou K est supérieure ou égal au critère de vigilance ρ (ρ^a et/ou ρ^b) on supposera qu'il y a résonance et que la couche F2 associée est activée: ainsi pour l'ARTa $Y_j = 1$ et $Y_i = 0$ pour tout $j \neq J$ (pareillement pour ARTb). Dans le cas contraire, désactiver ce nœud en le mettant à zéro (RAZ), sélectionner un nouveau nœud et refaire l'étape 2.2 dans le module respectif.

Si aucune catégorie n'a pu être choisie, un (des) nouveau(x) nœud(s) est (sont) créé(s) dynamiquement, soit J et/ou K ce(s) nœud(s). Avec:

$$W_J^a = 1, W_J^{ab} = 1, y_J^a = 1 \text{ et } y_J^b = 0 \text{ pour } j \neq J$$

$$\text{et/ou } W_K^b = 1, W_K^{ab} = 1, y_K^b = 1 \text{ et } y_K^a = 0 \text{ pour } k \neq K$$

Etape 3 : Test de vigilance dans la couche MAP. Dans cette couche on calcule $x^{ab} = W_J^{ab} \wedge y^b$.

Si $|x^{ab}| |I| |y^b| \geq \rho^{ab}$ on exécute l'étape 4. Sinon il y a mauvaise catégorisation dans le ARTa on exécutera, alors l'étape 5.

Etape 4 : Apprentissage ou mise à jour des poids.

W_J^a, W_K^b : sont mis à jour comme suit:

$$W_j^{(\text{nouveau})} = \beta(I \wedge W_j^{(\text{ancien})}) + (1 - \beta)W_j^{(\text{ancien})} \tag{4}$$

Et W_J^{ab} est mis à jour comme suit:

$$W_j^{ab(\text{nouveau})} = \beta^{ab}(y^b \wedge W_j^{ab(\text{ancien})}) + (1 - \beta^{ab})W_j^{ab(\text{ancien})} \tag{5}$$

Il est à remarquer que l'apprentissage rapide correspond au choix $\beta=1$ dans chacune des couches. Etape 5 : Changement du critère de vigilance du ARTa.

Mettre $\rho^a = m_j^a(A) + \varepsilon$, et $T_j = 0$ et aller de nouveau à l'étape 2.

3.2 Modification du réseau ARTMAP flou

Une amélioration de cet algorithme est proposée dans ce papier. Elle consiste à faire passer l'ensemble des exemples de la base d'apprentissage non pas une seule fois comme c'est d'usage, mais autant de fois que le réseau est en évolution architecturale (c'est à dire jusqu'à stabilité de l'architecture du réseau) ou que l'erreur fixée n'est pas atteinte. Cette amélioration est la conséquence de deux principaux constats:

1°/ Le réseaux ARTMAP flou possède beaucoup trop de paramètres à fixer pour atteindre un taux d'apprentissage raisonnable.

Ces paramètres sont: les coefficients de vigilance et les coefficients d'apprentissage des ARTa, ARTb et du MAP, et les coefficients de comparaison des ARTa et ARTb.

Pour mettre en évidence cette difficulté de choix des paramètres, une étude consistant à faire varier un paramètre et à fixer les autres pour amener le réseau à classifier une base d'apprentissage de 252 exemples en 4 classes, a été réalisée. Cette étude consiste à déterminer l'erreur d'apprentissage et le nombre de cellules du réseau en fonction du paramètre mis en jeu, l'algorithme utilisé étant celui décrit plus haut (classique), et les paramètres fixes à:

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

Les résultats sont présentés en figure 2 sous forme de courbes d'erreur et du nombre de cellules du réseau en fonction du paramètre étudié.

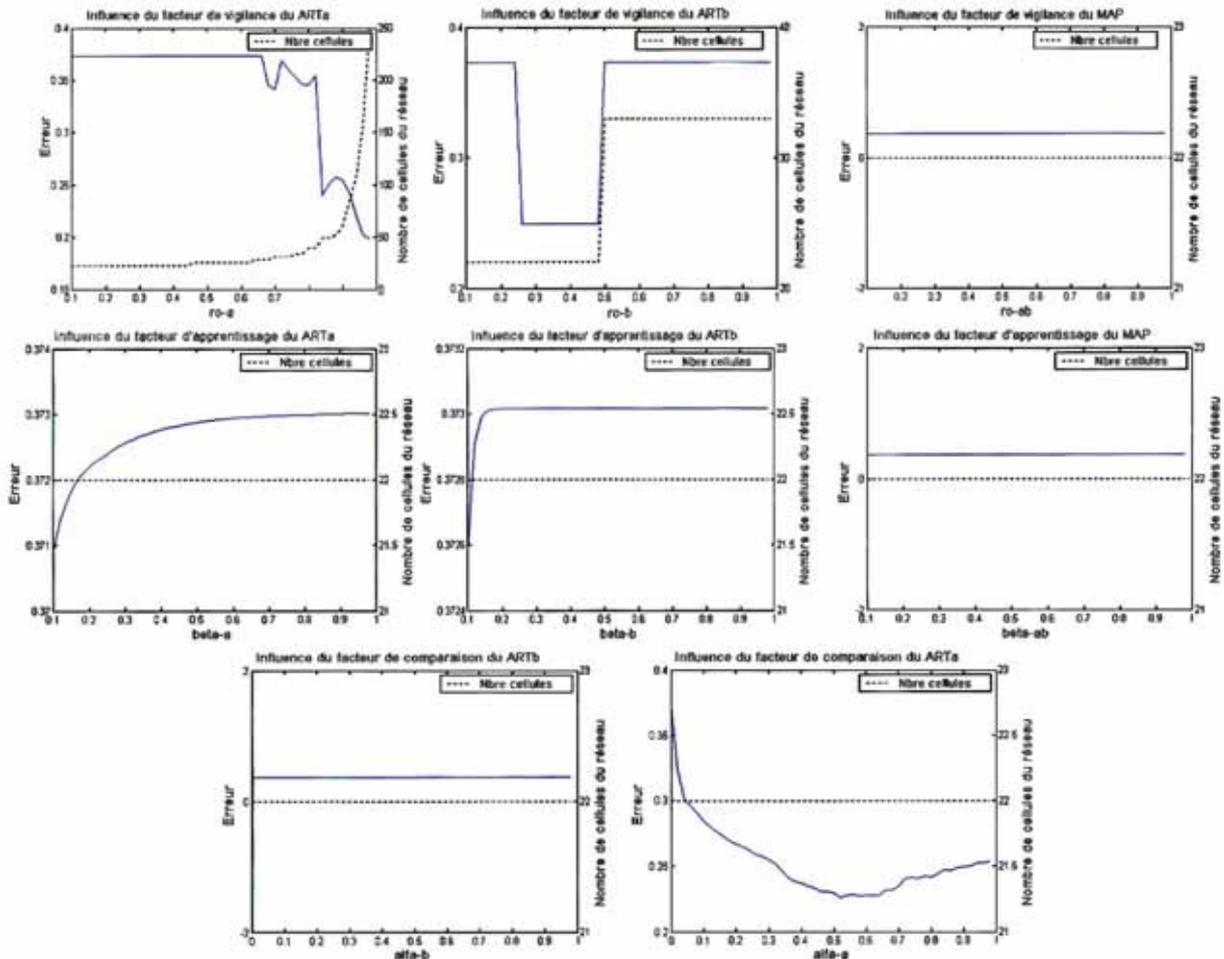


Fig. 2 Etude de l'influence d'un paramètre sur le taux d'apprentissage et sur le nombre de cellules du réseau.

Il apparaît clairement qu'il est très difficile de trouver les valeurs adéquates des paramètres pour faire converger le réseau au résultat escompté. Cette difficulté est encore plus ardue par l'existence du nombre important de combinaisons possibles. Dans cette optique une étude complémentaire consistant à faire varier deux paramètres à la fois en fixant les autres a été réalisée sans qu'elle donne des résultats probants. la figure 3 illustre à titre d'exemples l'étude de l'erreur et du nombre de cellules en fonction de $(\bar{\rho}^a, \rho^b)$, $(\bar{\rho}^a, \alpha^a)$ et (β^a, β^b) .

2°/ L'algorithme d'apprentissage du réseau ARTMAP flou tel que décrit en section 3.1, fait

passer les exemples un à un, et pour chaque exemple une mise à jour de l'architecture et/ou des poids est réalisée. Entre le passage d'un exemple se situant en début de la base et la fin de l'apprentissage le réseau sera fortement modifié si les paramètres sont mal choisis, cette modification influe négativement sur le degré d'apprentissage des premiers exemples.

La solution que nous préconisons dans cet article est de faire passer la base d'apprentissage autant de fois que l'erreur fixée n'est pas atteinte ou que l'architecture reste en évolution et cela sans se soucier des huit paramètres en les fixant une seule fois.

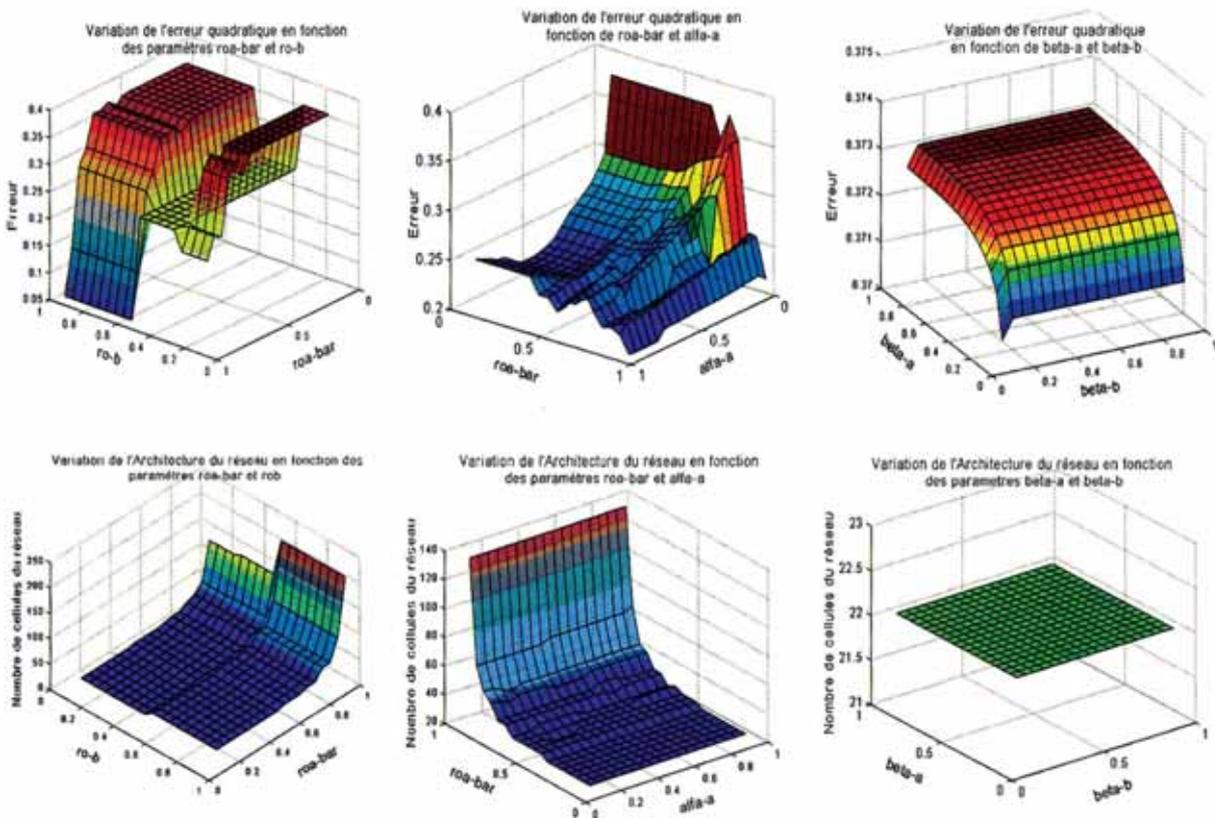


Fig. 3 Etude de l'influence de deux paramètres sur le taux d'apprentissage et sur le nombre de cellules du réseau.

L'étude menée sur la même base mais en la faisant passer plusieurs fois montre (figure 4) que l'erreur décroît après chaque passage. Nous avons fixé les facteurs de vigilance et d'apprentissage à de fortes valeurs égales à 0,95, par contre les facteurs de

comparaison sont choisis faibles (égaux à 10^{-4}). La stabilité architecturale est atteinte au bout de la sixième (6^{ème}) itération, alors que l'erreur quadratique moyenne qui est de l'ordre de $3,04 \cdot 10^{-10}$ n'est atteinte qu'à la neuvième (9^{ème}) itération.

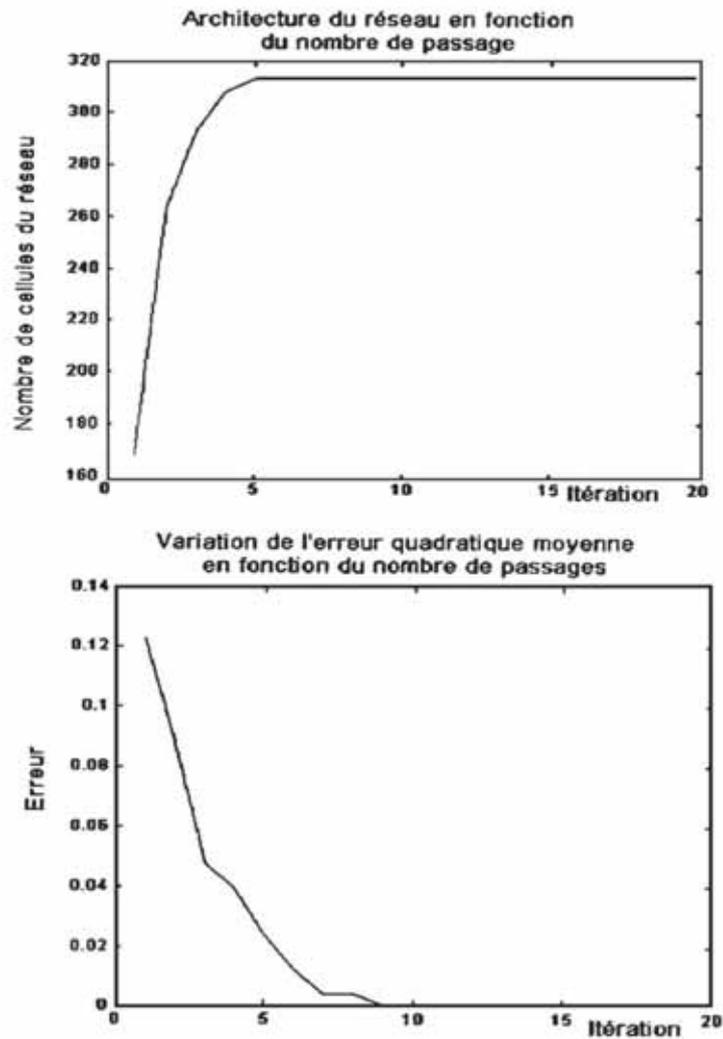


Fig. 4 Evolution de l'erreur quadratique moyenne et de l'architecture du réseau en fonction du nombre de passages.

4. Le réseau PMC à apprentissage rapide

Dans le réseau perceptron multi couches (PMC) (figure 5) à apprentissage par rétro propagation du gradient (RPG) [Zhou00, Rumelhart86], le gradient donne la direction vers laquelle il faut se déplacer pour trouver le minimum de l'erreur, mais ne donne en aucun cas le pas avec lequel on doit modifier les poids du réseau pour faire diminuer au plus vite cette erreur [Baldi95], en effet ce pas est un coefficient fixe ou adaptatif (variant pour chaque itération). L'algorithme de Levenberg-Marquardt [Hagan94, Baldi95, Wilamowski01] permet de déterminer ce pas en utilisant la dérivée seconde de l'expression de l'erreur quadratique moyenne.

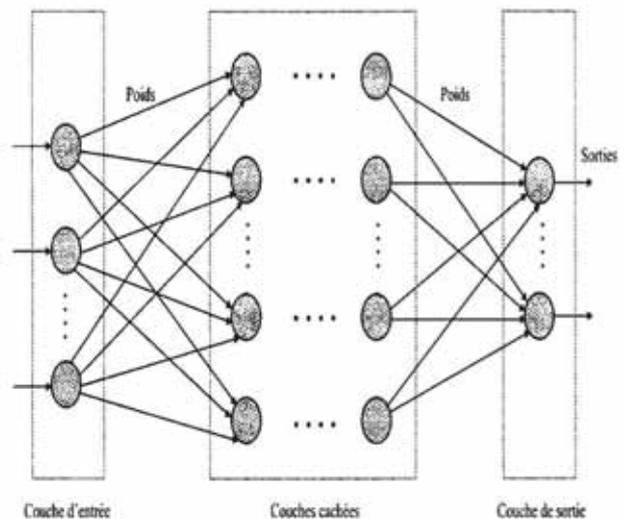


Fig. 5 Réseau perceptron multi couches PMC.

Cette erreur quadratique s'exprime par:

$$F(w) = \left\langle (g(a, w) - b)^2 \right\rangle \quad (6)$$

Où $\langle \rangle$ est la moyenne calculée sur l'ensemble des couples (a,b) et g une fonction de deux vecteurs a et w.

L'idée consiste en la détermination du nouveau vecteur poids W_{t+1} en fonction de l'actuel W_t , telle que la valeur de $F(W_{t+1})$ se rapproche le plus d'un minimum local de la fonction F. Pour cela, on calcule une approximation quadratique de F à partir d'une approximation linéaire de g autour du point w_t . En déterminant le point W pour lequel le gradient de l'approximation quadratique de F s'annule, on obtient:

$$w = w_t - H^{-1} d \quad (7)$$

Avec:

$$d = \left\langle (g(a, w_t) - b) \nabla g(a, w_t) \right\rangle \quad (8)$$

$$H = \left\langle \nabla g(a, w_t) \nabla g(a, w_t)^T \right\rangle \quad (9)$$

L'équation 7 pourrait servir dans la détermination de W_{t+1} à partir de W_t à la condition que g soit proche d'une droite autour de W_t .

L'utilisation de cette approche quadratique dans la région où g est quasi linéaire et une descente du gradient, dans le cas contraire elle représente le concept de Levenberg. Sa formulation est:

$$w_{t+1} = w_t - (H + \lambda I)^{-1} d \quad (10)$$

Effectivement, lorsque λ est faible, cette équation est équivalente à l'équation 7, et s'il est grand elle est équivalente à :

$$w_{t+1} = w_t - \frac{1}{\lambda} d = w_t - \frac{1}{2\lambda} \nabla F(a, w_t) \quad (11)$$

L'équation 11 est une descente du gradient. En pratique on procède comme suit:

$$F(w) = \sum_{p=1}^P \sum_{k=1}^K (d_{kp} - o_{kp})^2 \quad (12)$$

Où $w = [w_1 w_2 \dots w_N]^T$ est le vecteur poids du réseau entre les deux couches considérées, d_{kp} est la valeur désirée du k^{ème} nœud de la couche extérieure pour le p^{ème} exemple, O_{kp} est la valeur actuelle

ou celle obtenue par le réseau au nœud k pour l'exemple p ; P est le nombre d'exemples, et K le nombre de nœuds de la couche extérieure.

L'équation 6 peut s'écrire en fonction du vecteur d'erreur cumulative E comme suit:

$$F(w) = E^T E \quad (13)$$

Où $E = [e_{11} \dots e_{K1} e_{12} \dots e_{K2} \dots e_{1p} \dots e_{Kp}]^T$, $e_{pk} = d_{pk} - o_{pk}$,
 $k = 1, \dots, K$ et $p = 1, \dots, P$.

Les poids sont mis à jour comme suit:

$$w_{t+1} = w_t - (J_t^T J_t + \lambda I_t)^{-1} J_t^T E_t \quad (14)$$

Où J est la matrice Jacobien, W_{t+1} sont les nouveaux poids, W_t sont les anciens poids et I est la matrice identité.

Il est évident que ce procédé est particulièrement intéressant pour faire converger le réseau de neurones en un nombre d'itérations moindre, mais il est clair aussi que chaque itération nécessite plus de calcul du fait de l'inversion de la matrice.

5. Résultats

En plus de la modification du réseau de neurones ARTMAP flou, le but de ce papier est d'évaluer les performances du ARTMAP flou modifié. Nous avons donc effectué une comparaison avec un réseau PMC à apprentissage rapide utilisant l'algorithme de Levenberg-Marquardt.

Les deux réseaux étant à apprentissage supervisé, ils ont été entraînés sur une base de 252 exemples représentant quatre classes (la classe 1 pour l'eau, la classe 2 pour le bois, la classe 3 pour le bâtis et la classe 4 pour la terre nue ou agricole), l'erreur quadratique moyenne à atteindre lors de l'apprentissage étant de 10^{-3} .

Le réseau PMC choisi pour cette application est un réseau à une couche d'entrée de 3 cellules (chaque cellule reçoit un pixel fuzzifié d'une bande), une couche cachée à 9 cellules, et une couche de sortie à quatre cellules (exprimant les quatre classes).

Le réseau ARTMAP flou étant à architecture évolutive pendant la phase d'apprentissage, seuls les nombres de cellules des couches FO du ARTa et du ARTb sont fixés à six (les trois entrées fuzzifiées et leurs compléments à un) et à huit (les quatre classes fuzzifiées et leurs compléments à un) respectivement.

L'erreur fixée est atteinte par le réseau ARTMAP flou modifié après 9 itérations (figure 4), et au bout de 87 itérations pour le PMC choisi (figure 6), et en un temps de 4,69 fois plus important que le temps mis par le réseau ARTMAP flou modifié (3,375 s / 0,719 s).

Il est à remarqué qu'une erreur plus faible que celle fixée a été impossible à obtenir avec le PMC à apprentissage rapide par l'algorithme de Levenberg-Marquardt, ceci est dû au fait que dans cet algorithme une inversion de matrice doit être réalisée, et que pour cette exigence cette matrice devient singulière (impossible à inverser). A l'inverse avec l'ARTMAP flou modifié, une erreur quadratique moyenne de $3,04 \cdot 10^{-10}$ a pu être atteinte. Les deux classifieurs ont été contrôlés sur une base de contrôle de 217 points. Les résultats sont illustrés par la figure 7.

On constate que pour la classe 1, les deux classifieurs ont donné des résultats identiques avec un bon classement de tous les points, quant aux autres classes l'ARTMAP flou modifié présente de meilleurs résultats. En effet on a recensé 17 points mal classés pour l'ARTMAP flou modifié et 20 pour le PMC, ce qui donne un taux de bon classement de 92,17 % pour l'ARTMAP flou modifié et de 90,78 % pour le PMC à apprentissage rapide.

La généralisation de ces deux réseaux à l'image SPOT XS de la baie d'Alger de 1500x1000 points s'est faite en un temps égale à 205,734 s pour le ARTMAP flou modifié, et en 8031,70 s pour le PMC à apprentissage rapide, soit un rapport de 39 en faveur du ARTMAP flou modifié. Les résultats des classifications sont illustrés par la figure 8.

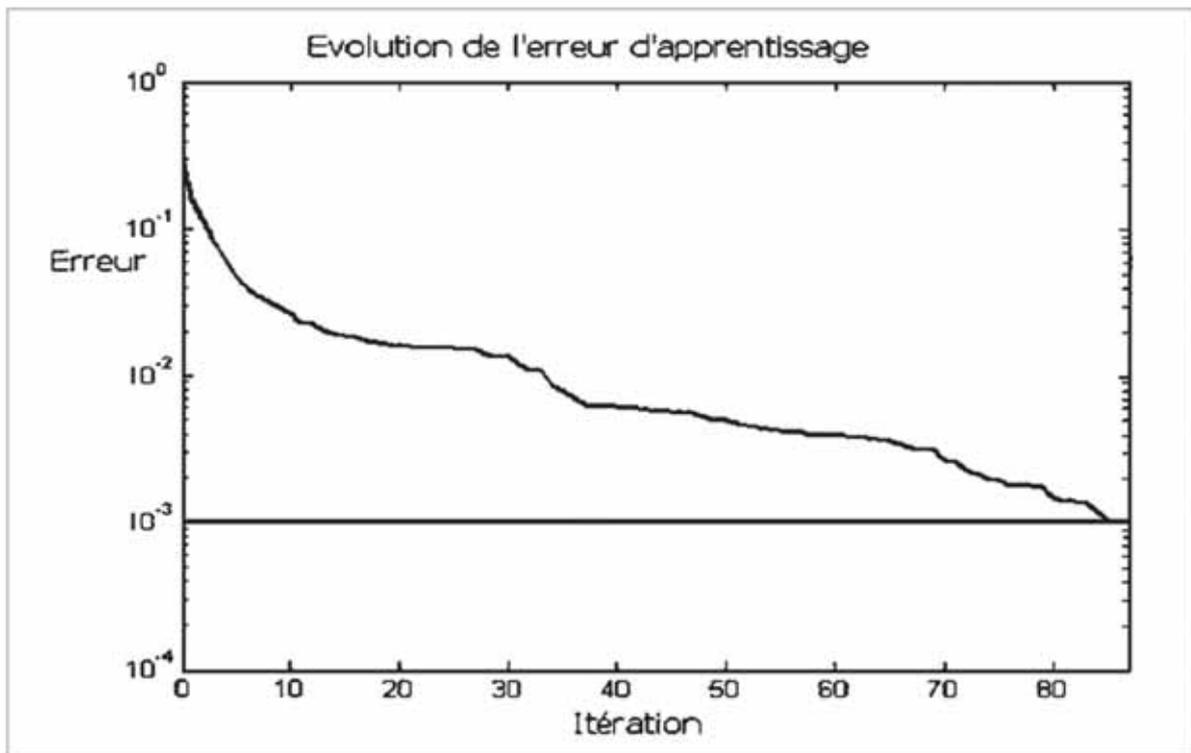


Fig. 6 Courbe de l'erreur quadratique moyenne du réseau PMC à apprentissage rapide.

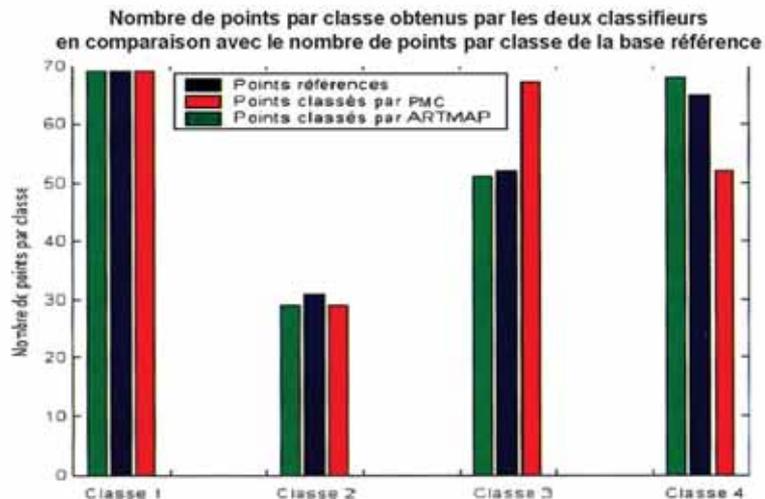


Fig. 7 Comparaison des classifications par l'ARTMAP flou modifié et le réseau PMC à apprentissage rapide par rapport à une base de contrôle.

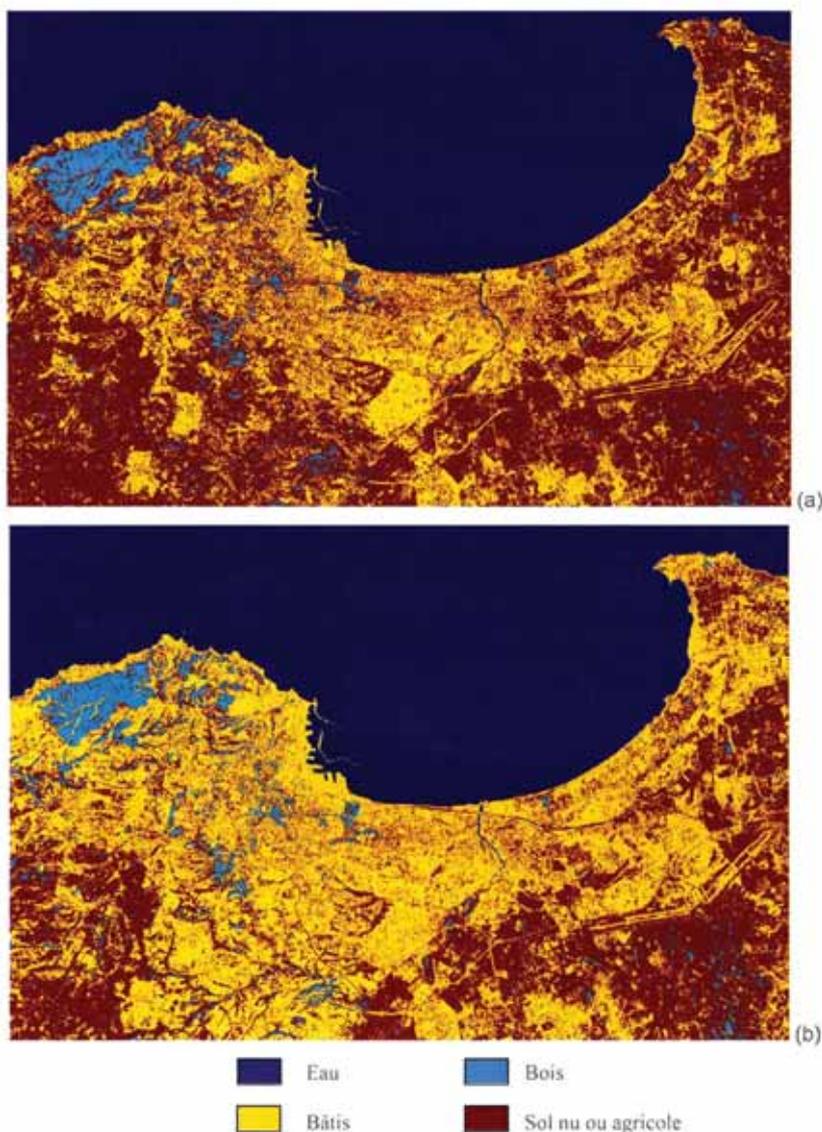


Fig. 8 Résultats de la classification de l'image SPOT XS de la baie d'Alger:
(a) image classée par l'ARTMAP flou modifié, (b) image classée par le réseau PMC à apprentissage rapide.

6. Conclusion

Le réseau ARTMAP flou possède beaucoup trop de paramètres à fixer pour atteindre un taux d'apprentissage raisonnable. La difficulté du choix de ces paramètres est encore plus accentuée par l'existence d'un nombre très important de combinaisons possibles. Une étude consistant à faire varier un ou deux paramètres à la fois en fixant les autres a été réalisée sans qu'elle donne des résultats probants.

De plus, l'algorithme d'apprentissage du réseau ARTMAP flou tel que défini jusqu'à présent, fait passer les exemples un à un, et pour chaque exemple une mise à jour de l'architecture et/ou des poids est réalisée. Entre le passage d'un exemple se situant en début de la base et la fin de l'apprentissage, le réseau sera fortement modifié si les paramètres sont mal choisis, et cette modification influe négativement sur le degré d'apprentissage des premiers exemples.

Ce travail est une contribution à l'amélioration de l'apprentissage du réseau ARTMAP flou, et l'utilisation du réseau ainsi modifié pour classifier une image multi spectrale SPOT XS de la baie d'Alger, ainsi que la comparaison de ses performances avec le classifieur connexionnistes PMC à apprentissage rapide à rétro propagation du gradient utilisant l'algorithme de Levenberg-Marquardt. L'amélioration proposée s'est montrée très utile par le fait que, d'une part nous nous soucions que peu des paramètres du ARTMAP flou que nous fixons une fois pour toutes aux valeurs indiquées, et d'autre part par le fait que cette technique a permis de réduire considérablement l'erreur d'apprentissage. La comparaison des performances des deux réseaux mis en compétition dans cette étude montre que l'ARTMAP flou modifié s'est montré plus rapide à apprendre et encore plus à généraliser pour les mêmes objectifs assignés. De plus le PMC, à l'inverse du ARTMAP flou, s'est montré incapable d'atteindre une erreur à l'apprentissage plus basse que celle fixée dans cette application.

Références

- Aggarwal R K., Xuan Q. Y., Johns A T., Li F., & Bennett A., 1999. A Novel Approach to Fault Diagnosis in Multicircuit Transmission Lines Using Fuzzy ARTmap Neural Networks. *IEEE Transactions on Neural Networks*, Vol. 10, N° 5.
- Baldi P., 1995 Gradient Descent Learning Algorithm Overview: A General Dynamical Systems Perspective. *IEEE Transactions on Neural Networks*, Vol. 6 N° 1.
- Baraldi A. & Alpaydın E., 2002. Constructive Feedforward ART Clustering Networks-Part I. *IEEE Transactions on Neural Networks*, Vol. 13, N° 3.
- Carpenter G. A., Grossberg S. & Rosen D. B., 1991. 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.
- Carpenter, G. A., Grossberg S., Markuzon N., Reynolds J. H & Rosen D. B., 1992. Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multi-dimensional Maps. *IEEE Transactions on Neural Networks*, Vol. 3 N° 5.
- Carpenter G. A., Grossberg S., Markuzon N. & Reynolds J. H., 1995. A Fuzzy ARTMAP Nonparametric Probability Estimator for Nonstationary Pattern Recognition Problems. *IEEE Transactions on Neural Networks*, Vol. 6 N° 6.
- Charalampidis D., Kasparis T., & Georgiopoulos M., 2001. Classification of Noisy Signals Using Fuzzy ARTMAP Neural Networks. *IEEE Transactions on Neural Networks*, Vol. 12, N° 5.
- Charalampidis D, Anagnostopoulos G. C., Georgiopoulos M. & Kasparis T., 2002. Fuzzy ART and Fuzzy ARTMAP with adaptively Weighted distances. *Proc. SPIE*, Vol. 4739 pp. 8697.
- Dagher I., Georgiopoulos M., Heileman G. L. & Bebis G., 1999. An Ordering Algorithm for Pattern Presentation in Fuzzy ARTMAP That Tends to Improve Generalization Performance. *IEEE Transactions on Neural Networks*, Vol. 10, N° 4.
- Frank T., Kraiss K. F. & Kuhlen T., 1998. Comparative Analysis of Fuzzy ART and ART-2A Network Clustering Performance. *IEEE Transactions on Neural Networks*, Vol. 9, N° 3.
- Gómez-Sánchez E., Dimitriadis Y. A., Cano-Izquierdo J. M., & López-Coronado J., 2002. μ ARTMAP: Use of Mutual Information for Category Reduction in Fuzzy ARTMAP. *IEEE Transactions on Neural Networks*, Vol. 13, N° 1.
- Hagan M. T. & Menhaj M. B., 1994. Training Feed forward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, Vol. 5 N° 6.
- Tan A. H., 1997. Cascade ARTMAP: Integrating Neural Computation and Symbolic Knowledge Processing. *IEEE Transactions on Neural Networks*, Vol. 8 N° 2.
- Rumelhart D., Hinton G., Williams R. 1986. Learning Internal Representations by Error Propagation. *Parallel Distributed Processing*, Vol. 1, pp. 318-362, Cambridge, MA, MIT Press.
- Wilamowski B. M., Iplikci S., Kaynak O. & Efe M. O., 2001. An Algorithm for fast Convergence in Training Neural Networks IJCNN'01 pp. 1778-1782 Washington C. c Juillet 15-19.
- Zhou Z., Chen S. & Chen Z., 2000. FANNC : A Fast Adaptive Neural Network classifier. *Knowledge and Information Systems*, Vol 2, N° 1, pp.115-129.