

La Théorie de la Résonance Adaptative et les Moments de Zernike pour la Reconnaissance de Mots Arabes Manuscrits

Leila Chergui¹, Maamar Kef², Mohammed Benmohammed³

¹Université Larbi Ben Mhidi, Département d'informatique
Oum El Bouaghi – Algérie
pgliela@yahoo.fr

²Université Larbi Ben Mhidi, Département d'informatique
Oum El Bouaghi – Algérie
Lm_kef@yahoo.fr

³Université Mentouri Département d'informatique
Constantine – Algérie
Ibnm@yahoo.fr

Résumé. La reconnaissance de l'écriture arabe manuscrite est un domaine de recherche relativement récent et qui a connu ces dernières années des progrès remarquables. Il présente un intérêt indéniable dans l'accomplissement de tâches considérées fastidieuses dans certains domaines comme le tri postal, la lecture de chèques bancaires, la lecture des bordereaux, etc. Ce papier présente la conception, la réalisation et l'évaluation d'un système dédié à la reconnaissance automatique hors-ligne de mots manuscrits arabes représentant des noms de villes tunisiennes tirés de la base IFN/ENIT. Dans ce travail, nous nous pencherons sur une approche basée sur l'utilisation des moments de Zernike, et d'un classifieur neuronal utilisé pour la première fois dans ce domaine, à savoir le réseau Fuzzy ART. Nous montrerons, à travers les différentes étapes considérées, l'apport de notre technique dans la résolution des problèmes liés au traitement de l'écriture arabe. Par ailleurs, nous retenons les limitations enregistrées. Les résultats obtenus sont prometteurs.

Mots clés : Mots arabes manuscrits, Reconnaissance, Squelettisation, Moments de Zernike, Réseau Fuzzy ART.

1 Introduction

La reconnaissance automatique de caractères, au sens large du terme, est une discipline qui a vu le jour dès l'apparition des premiers ordinateurs. Reconnaître de l'écriture manuscrite consiste à associer une représentation symbolique à une séquence de symboles graphiques : on parle aussi de lecture automatique. Le but est de pouvoir utiliser cette représentation dans une application informatique. On distingue deux grands types d'utilisation :

- Traiter automatiquement des documents contenant de l'écriture manuscrite dont l'analyse par des individus prend trop de temps.

- Faciliter l'utilisation des ordinateurs pour des applications où un stylo est plus pratique qu'un clavier et une souris.

Contrairement au Latin, la reconnaissance de l'écriture arabe manuscrite reste encore aujourd'hui au niveau de la recherche et de l'expérimentation. Cependant et depuis quelques années elle a pris un nouvel essor et fait l'objet d'applications de plus en plus nombreuses. Parmi ces applications, nous citons le traitement automatique des dossiers administratifs, des formulaires d'enquêtes, des chèques bancaires, numérisation et sauvegarde du patrimoine culturel écrit, etc.

Notre article est organisé comme suit : la deuxième section abordera les caractéristiques de l'écriture arabe ainsi que les travaux effectués dans ce domaine. La troisième mettra le point sur l'architecture du réseau Fuzzy ART. L'implémentation du système sera détaillée dans la quatrième section. Le tout sera clôturé par une liste de perspectives.

2 L'écriture Arabe

L'écriture arabe a vu le jour aux alentours du VI^{ème} siècle avant l'apparition de l'écriture cursive nabatéenne, et s'est progressivement répandue avec l'existence de l'Islam et la révélation coranique. Les principales caractéristiques de la langue arabe sont :

- L'alphabet arabe comprend vingt-huit lettres fondamentales. Chacune a entre deux et quatre formes selon sa position dans le mot. La figure 1 donne toutes les formes possibles pour chaque lettre de l'alphabet arabe.

Début	Milieu	Fin et lié	Fin	Début	Milieu	Fin et lié	Fin
أ		أ	أ	ض	ض	ض	ض
ب	ب	ب	ب	ط	ط	ط	ط
ت	ت	ت	ت	ظ	ظ	ظ	ظ
ث	ث	ث	ث	ع	ع	ع	ع
ج	ج	ج	ج	غ	غ	غ	غ
ح	ح	ح	ح	ف	ف	ف	ف
خ	خ	خ	خ	ق	ق	ق	ق
د		د	د	ك	ك	ك	ك
ذ		ذ	ذ	ل	ل	ل	ل
ر		ر	ر	م	م	م	م
ز		ز	ز	ن	ن	ن	ن
س	س	س	س	ه	ه	ه	ه
ش	ش	ش	ش			و	و
ص	ص	ص	ص	ي	ي	ي	ي

Fig. 1. Les différentes formes possibles d'apparence des caractères de l'alphabet arabe.

- Quelques caractères arabes incluent dans leur forme un, deux ou trois points diacritiques. Ces points peuvent se situer au-dessus ou au-dessous du caractère mais jamais en haut et en bas simultanément [18].
- L'existence du "hamza" (le zigzag), qui se comporte, soit comme une lettre à part entière, soit comme un diacritique.
- Certaines formes de lettres ne peuvent dans aucun cas être rattachées à la lettre suivante, ce qui fait qu'un mot unique peut être entrecoupé d'un ou plusieurs espaces, lesquels sont aussi utilisés pour séparer les mots.
- Les voyelles "a", "i" et "ou" ne sont pas utilisées systématiquement dans l'écriture arabe ; des signes qui correspondent à des voyelles sont employés pour éviter des erreurs de prononciation.
- On trouve également des chevauchements et des ligatures dans l'écriture manuscrite ce qui complique la tâche de reconnaissance (Fig. 2.).

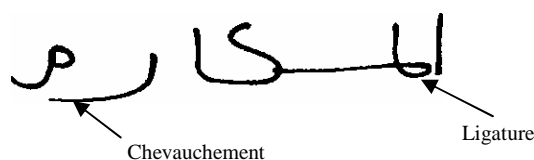


Fig. 2. Les ligatures et les chevauchements dans un mot arabe.

Plusieurs chercheurs ont conçu des systèmes de reconnaissance de l'écriture arabe, ils se diffèrent par le choix de type d'écriture ; imprimé ; manuscrit, en-ligne ou hors-ligne. Abd [1], Aburas [2], Benouareth [5], Farah [10], Farah [11], Khorsheed [16] et Mozaffari [20] ont préféré l'écriture manuscrite hors-ligne, tandis que Al-Muhtaseb [3], Ben Amor [4] et Khorsheed [17] ont choisi le type imprimé. L'écriture en-ligne quand à elle est discutée dans les systèmes de Biadisy [7], Elanwar [9] et Mezghani [19]. Les Classifieurs les plus utilisés pour l'écriture arabe sont les Chaînes de Markov Cachées (HMM) dans [3], [4], [5], [7], [16], et [17], les réseaux de neurones de type Perceptron Multi-couches (PMC) dans [4], [10], [11] et [19], on trouve également les SVMs dans [1], le classifieur bayésien dans [18] et les k-proches voisins (k-ppv) dans [18].

3 Architecture de Fuzzy ART

Les réseaux Fuzzy ART qui représentent une classe de la famille des réseaux ART (Adaptive Resonance Theory) est un modèle de réseau de neurones à architecture évolutive développé en 1987 par Carpenter et Grossberg [12]. C'est un réseau compétitif à deux couches de neurones complètement inter-reliées. Une couche de comparaison F_1 sert à coder les entrées avec un encodage dit complémentaire et une couche de compétition F_2 semblable à celle du réseau de Kohonen [14], ces deux couches sont activées par une entrée X (Fig.3.).

Le Fuzzy ART propose une catégorisation originale avec des classes représentées par des prototypes.

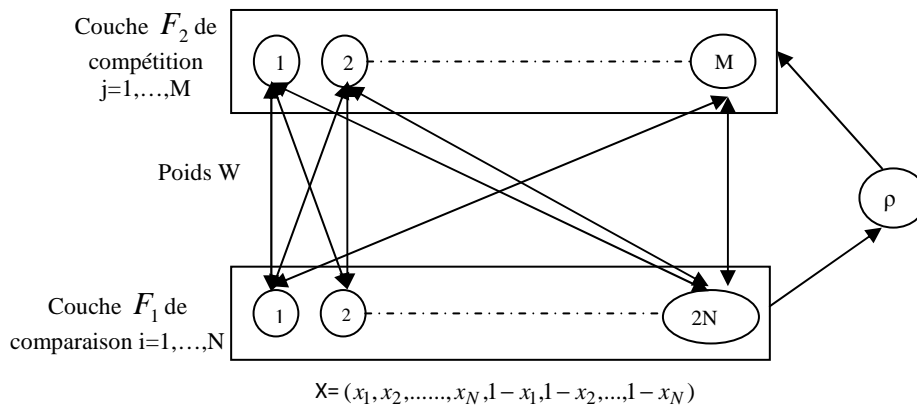


Fig. 3. Réseau de neurones Fuzzy ART.

Cependant, tout comme les autres architectures ART, le Fuzzy ART incorpore un mécanisme de rétroaction permettant de stabiliser les prototypes appris dans les vecteurs de poids qui relient les deux couches [13]. Ce mécanisme dit de résonance est contrôlé par un paramètre qui permet de réinitialiser au besoin la couche compétitive. Pour chaque entrée, les sorties du réseau spécifient une catégorie parmi les classes de sortie.

Le Fuzzy ART est un réseau constructif où de nouveaux neurones sont alloués lors de la phase d'apprentissage. Généralement, on fixe au départ un nombre maximum de neurones S , limitant ainsi le nombre maximum de catégories possibles [6]. Initialement, aucun neurone n'est actif. L'allocation subséquente de nouvelles catégories dépendra à la fois des entrées et des paramètres de l'algorithme. Parmi les avantages de Fuzzy ART on cite :

- L'algorithme de Fuzzy ART propose des calculs simplifiés pour la formation des classes sous forme d'hyper-boîtes, contrairement à des classes circulaires tels que retrouvés dans la plupart des algorithmes de réseaux de neurones [14].
- L'apprentissage se stabilise dans un nombre fini d'itérations.
- Le modèle Fuzzy ART exploite à fond un des avantages inhérents de l'approche neuronale ; le parallélisme [22].
- La forme prototype d'une classe sera immédiatement reconnue, même si elle n'a jamais été présentée, grâce aux caractéristiques pertinentes.

Ce classifieur n'a pas été exploité dans le domaine de reconnaissance de l'écriture arabe manuscrite ; notre système sera donc le premier à utiliser un réseau de neurones de type Fuzzy ART.

4 Implémentation du Système

Notre système qui réalise la reconnaissance d'écriture arabe manuscrite hors-ligne englobe plusieurs étapes décrites par la figure 4 :

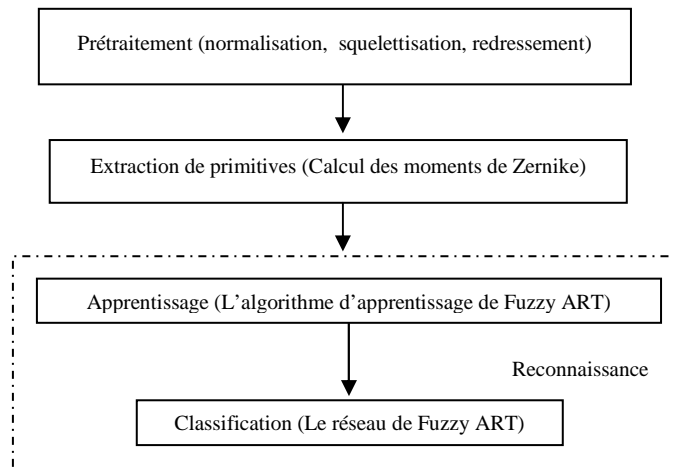


Fig. 4. Architecture du système.

4.1 Prétraitement

Qui est fait afin de réduire le bruit incluant les opérations de normalisation, de squelettisation et de redressement. La normalisation a été réalisée à travers une méthode de normalisation linéaire basée sur trois étapes [8] :

1. Calculer la matrice de dispersion de la forme.
2. Changer l'origine des axes des coordonnées vers le centre de la forme.
3. Changer l'échelle de base.

On a abouti à des images normalisées de taille 400×100 pixels.

La figure 5 illustre un exemple.

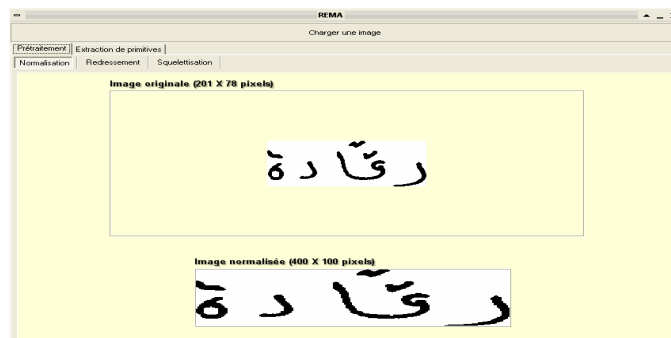


Fig. 5. Normalisation du mot "رقادة".

Pour la squelettisation on a appliqué quatre algorithmes ; de Rutovitz [8], de Zhang et Suen [24], de Deutch [8] et celui de Zhang et Wang [8] ce qui est montré par la figure 6.

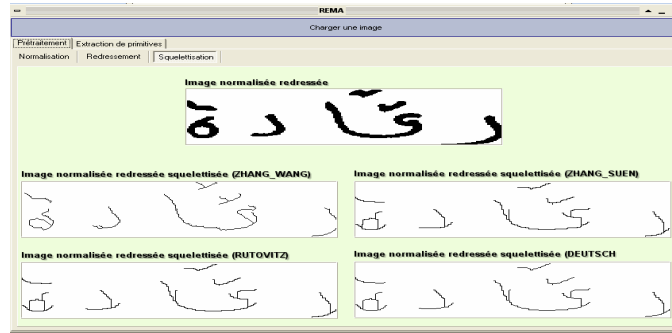


Fig.6. Application des algorithmes de squelettisation pour le mot "رئاسة".

L'utilisation de plusieurs algorithmes de squelettisation était dans le but de choisir le meilleur entre eux, c'est-à-dire celui préservant le plus la structure du mot.

Le redressement est effectué en utilisant les histogrammes de projection horizontale selon onze angles différents de rotation variant de -5° à $+5^\circ$, l'angle d'inclinaison correcteur sera celui de l'histogramme le plus dense. En calculant l'entropie de chacun des histogrammes obtenus, on pourra déterminer l'histogramme le plus dense représenté par la plus petite entropie. L'entropie est une mesure de l'information représentée par la formule suivante :

$$E = -\sum_i p_i \log(p_i). \quad (1)$$

$$p_i = \frac{N_i}{N}. \quad (2)$$

Où N_i est le nombre de pixels ayant l'ordonnée y_i dans le repère de projection et N est le nombre total des pixels ou de points de contour du mot. La probabilité p_i de l'histogramme désigne la fréquence d'occurrence de l'ordonnée y_i . La figure 7 illustre les histogrammes de rotation ; le plus dense (c'est-à-dire ayant la plus petite entropie) est celui coloré en jaune.

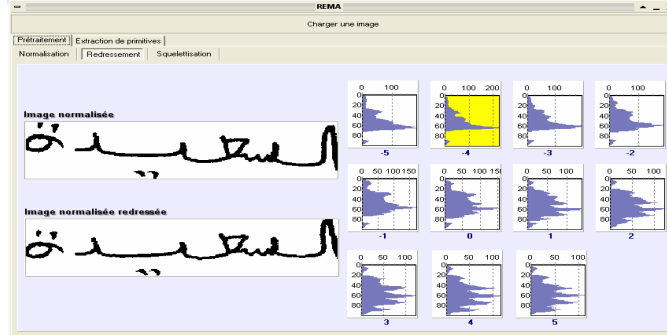


Fig.7. Histogrammes de rotation du mot "السعيدة".

4.1 Extraction de primitives

Les moments sont utilisés en physique pour décrire la répartition des masses dans un corps. En analyse d'image, on peut envisager la même démarche en associant le niveau de gris d'un point de l'image à la masse élémentaire en un point. Donc les moments fournissent des informations concernant l'arrangement spatial de l'image.

Les moments de Zernike sont introduits en 1934, ils sont définis par les équations suivantes :

$$Z_{pq} = \frac{(p+1)}{\pi(N-1)^2} \sum_{x=1}^N \sum_{y=1}^N V_{nm}^*(r, \theta) f(x, y). \quad (3)$$

Avec

$$V_{nm}^* = R_{nm}(r) \exp(im\theta) \quad (4)$$

Où

n : un entier positif ou nul

m : un entier tel que $|m| \leq n$

ρ : La longueur du vecteur entre l'origine et le pixel (x,y)

θ : L'angle entre le vecteur p et x

R_{nm} : un polynôme radial

$$R_{nm}(r) = \sum_{s=0}^{\left\lfloor \frac{n-|m|}{2} \right\rfloor} \binom{n-|m|}{2s} (-1)^s \frac{(n-s)!}{S! \left(\frac{n-2s+|m|}{2} \right)! \left(\frac{n-2s-|m|}{2} \right)!} \quad (5)$$

Le moment de Zernike est donc la projection d'une fonction $f(x,y)$ décrivant une image sur un espace de polynômes orthogonaux engendré par $V_{nm}^*(\rho, \theta)$ [25]. On désigne par moments de Zernike une série de calculs utilisée pour transformer une image en un vecteur de composantes réelles représentant les moments Z_{ij} .

L'avantage principal que présente les moments de Zernike est qu'ils sont insensibles à la translation, la rotation et au changement d'échelle ce qui permet de préserver l'information contenue dans les images [15]. Ils réduisent aussi le bruit et la redondance dans les vecteurs de caractéristiques. On a utilisé dans notre système les 47 premiers moments de Zernike où on présente un exemple dans la figure 8.

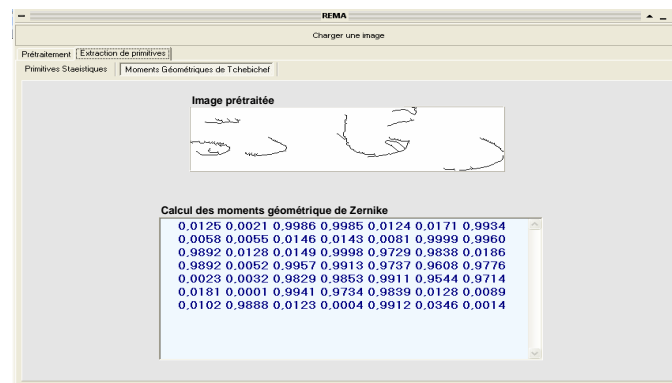


Fig.8. Calcul des 47 premier moments de Zernike du mot "رقادة".

4.1 Reconnaissance

Nous avons utilisé la base IFN/ENIT qui contient 26459 mots arabes représentant 964 classes de nom de villages tunisiens qui sont écrits par 411 scripteurs. Le processus de reconnaissance inclut les deux phases suivantes :

Apprentissage. Il est fait sur 4 ensembles de la base IFN/ENIT à savoir a, b, c, d. L'algorithme d'apprentissage non-supervisé utilisé est celui du réseau Fuzzy ART qui est détaillé par l'organigramme ci-dessous.

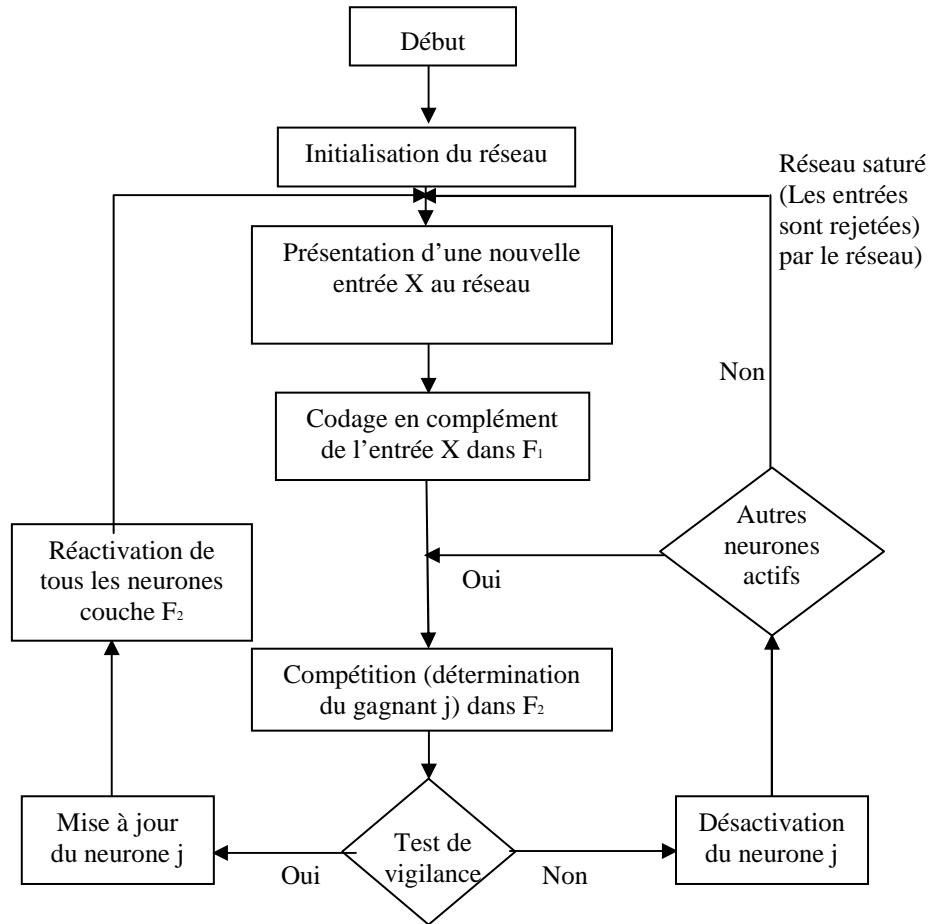


Fig. 9. Algorithme d'apprentissage de Fuzzy ART.

Classification et résultats. Elle est exécutée à travers le réseau Fuzzy ART. Les tests ont été effectués sur l'ensemble e de la base IFN/ENIT dont les résultats sont résumés dans le tableau 1. Le taux de reconnaissance obtenu avoisine les 88%, tandis que le taux de rejet est de 8% ; les 4% restants représentent le taux de mots mal-classés. Le taux d'erreur est principalement du à la mauvaise écriture où on illustre quelques échantillons dans la figure 10.

الدكتور المطار	الفيضان عينا	خديجة	عبرة الدول
بلاط الدض	الكبارج	لاستفطيمية	مناقص

Fig. 10. Echantillons de mots mal-classés.

Table 1. Résultats de classification.

Classes	Top1	Top2	Top5	Top8	Top10
150	88,66%	89,15%	90,52%	92,45%	93,48%
300	87,45%	88,25%	89,04%	90,44%	91,97%
450	85,52%	86,38%	88,79%	89,45%	90,86%
600	83,72%	84,75%	85,89%	87,40%	89,10%
750	81,15%	82,54%	83,75%	85,46%	88,79%
946	80,95%	81,68%	82,15%	83,25%	88,00%

Afin de mieux mesurer l'efficacité des résultats obtenus, des comparaisons avec des systèmes de reconnaissance de mots arabes manuscrits hors-lignes travaillant sur la base IFN/ENIT ont été effectuées. Les travaux concernés sont ceux de :

- **Le système ICRA (Intelligent Character Recognition for Arabic).** Conçu par Ahmed Abdulkader [18], qui a utilisé un classifieur neuronal de type Perceptron Multi Couche (PMC).
- **Le système TH-OCR.** Développé par Xiu, Wang, Jin, Jiang, Peng et Din [18]. Ce système a utilisé une méthode de reconnaissance statistique basée sur les réseaux de neurones en appliquant trois niveaux de segmentation, à savoir : la segmentation de texte en lignes, la segmentation de lignes en mots et la segmentation de mots en caractères.
- **Le système UOB.** Conçu par Mokbel et El Hajj [18] de l'université du Liban, où ils ont utilisé les chaînes de Markov cachées (HMM) comme classifieur et la densité des pixels comme primitives.
- **Le système REAM (Reconnaissance de l'Écriture Arabe Manuscrite).** C'est le résultat d'un travail de groupe de chercheurs tunisiens comprenant Masmoudi-Touj, Essoukhri-Ben Amara, Amiri et Ellouz [18], qui ont utilisé une chaîne de Markov de type planaire se basant sur des primitives statistiques.
- **Le système ARAB-IFN.** Conçu par Pechwitz [18], utilisant les chaînes de Markov cachées semi-continues d'ordre 1 comme classifieur. Ils ont extrait quelques primitives après la détection de la ligne de base.

Les résultats obtenus pour ces cinq systèmes de reconnaissance de l'écriture arabe manuscrite sont indiqués dans le tableau 2, où l'on remarque que le taux de reconnaissance enregistré par notre système REMA (Reconnaissance de L'Écriture Arabe Manuscrite) est nettement meilleur que celui obtenu par les systèmes; ICRA, SHOCRAN, et TH-OCR ; tandis que les systèmes, ARABE-IFN et UOB donnent des taux relativement meilleurs, ceci peut être expliqué par l'expérience cumulée à travers la massive exploitation des types de classifieurs utilisés par ces derniers, à savoir les HMMs et les PMCs.

Table 2. Comparaison des résultats.

Système	Type de Classifieur	Taux de reconnaissance
ICRA	Perceptron Multi Couche	87,75%
TH-OCR	Méthode statistique	50,14%
UOB	Chaînes de Markov Cachées	90,88%
REAM	Chaînes de Markov Cachées	49,86%
ARAB-IFN	Chaînes de Markov Cachées	89,77%
REMA	Fuzzy ART	88%

5 Conclusion

Dans cet article nous avons mis au point un système de reconnaissance d'écriture arabe manuscrite, utilisant un réseau Fuzzy ART comme classifieur et les moments de Zernike comme primitives. Nous nous sommes concentrés sur deux objectifs:

- Appliquer un classifieur Neuro-flou non-utilisé jusqu'à présent dans le domaine de la reconnaissance d'écriture arabe et analyser ses performances par rapport aux classifieurs déjà exploités.
- Démontrer l'efficacité des moments de Zernike et leur qualité en tant que primitives représentant des mots manuscrits arabes.

Les résultats obtenus sont encourageants comparés à d'autres systèmes travaillant eux aussi sur la base IFN/ENIT. Le taux d'erreur enregistré est principalement dû à la mauvaise écriture et aux problèmes de chevauchement des lettres dans les mots classés. Notons ici que le taux de reconnaissance peut être amélioré par une combinaison avec d'autres classifieurs et l'ajout d'une étape de poste-traitement.

References

1. Abd, M.A., Paschos, G.: Effective Arabic Character Recognition Using Support Vector Machines. *Innovations and Advanced Techniques in Computer and Information Sciences and Engineering*. Springer. 7--11 (2007)
2. Aburas, A.A., Rehiel, S.M.A.: Off-line Omni-Style Handwritten Arabic Character Recognition System based on Wavelet Compression. *CS&IT ARISER*. 4, 123--135 (2007)
3. Al-Muhtaseb, H.A., Mahmoud, S.A., Qahwaji, R.S.: Recognition of Off-Line Printed Arabic Text Using Hidden Markov Models. *Signal Processing*. Elsevier. 88, 2902--2912 (2008)
4. Ben Amor, N.: Multi-Font Arabic Characters Recognition Using Hough Transform and HMM/ANN. *Journal of Multimedia*. Academy Publisher. 1, 50--54 (2006)
5. Benouareth, A., Annajy, A., Sellami, M.: Semi-Continuous HMMs with Explicit State Duration for Unconstrained Arabic Word Modelling and Recognition. *Pattern Recognition Letters*. Elsevier. 29, 1742--1752 (2008)
6. Bezdek, J.C., Killer, J., Krisnapuram, R., Pal, N.R.: *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Springer, New York (2005)

7. Biadisy, F., El-Sana, J., Habash, N.: Online Arabic Handwritten Recognition Using Hidden Markov Models. *Pattern Recognition* (2004)
8. Cheriet, M., Kharma, N., Liu, C.L., Suen, C.Y.: *Character Recognition Systems*. Wiley-Interscience, New Jersey (2007)
9. Elanwar, R.I., ElRashwan, M.A., Mashali, A.: Simultaneous Segmentation and Recognition of Arabic Characters in an Unconstrained On-Line Cursive Handwritten Document. *International Journal of Computer and Information Science and Engineering (IJCISE)*. WASET. 1, 203--206 (2007)
10. Farah, N., Khadir, M.T., Sellami, M.: Artificial Neural Network Fusion: Application to Arabic Word Recognition. *ESANN Proceedings-European Symposium on Artificial Neural Networks*, pp. 151--156. Bruges (Belgium) (2005)
11. Farah, N., Souici, L., Farah, L., Sellami, M.: Arabic Words Recognition with Classifiers Combination : An Application to Literal Amounts. In *Proceeding of Artificial Intelligence: Methodology, Systems and Applications*, pp. 420--429. Bulgaria (2004)
12. Freeman, J.A., Skapura, D.M.: *Neural Networks: Algorithms, Applications and programming Techniques*. Addison-Wesley Publishing Company, United States of America (1991)
13. Graupe, D.: *Principles of Artificial Neural Networks*. World Scientific, Singapore (2007)
14. Hagan, M.T., Demuth, H.B., Beale, M.: *Neural Network Design*. PWS Publishing Company, United States of America (2002)
15. Kan, C., Srinath, M.D.: Invariant Character Recognition with Zernike and Orthogonal Fourier-Mellin Moments. *Pattern Recognition*, 35, 143--154 (2002)
16. Khorsheed, M.S.: Offline Recognition of Omni Font Arabic Text using the HMM Toolkit (HTK). *Pattern Recognition Letters*. Elsevier. 28, 1563--1571 (2007)
17. Khorsheed, M.S.: Recognition Handwriting Arabic Manuscripts Using Hidden Markov Model, *Pattern Recognition Letters*. Elsevier. 24, 2235--2242 (2003)
18. Lorgio, L.M.: Off-line Arabic Handwriting Recognition: A Survey. *IEEE on Pattern Analysis and Machine Intelligence*. 28, 712--724 (2006)
19. Mezghani, N., Mitiche, A., Cheriet, M.: A New Representation of Shape and its Use for High Performance in On-Line Arabic Character Recognition by an Associative Memory. *International Journal of Document Analysis*. Springer. 7, 201--210 (2005)
20. Mozaffari, S., Faez, K., Margner, V., El-Abed, H.: Lexicon Reduction Using Dots for Off-Line Farsi/Arabic Handwritten Word Recognition. *Pattern Recognition Letters*. Elsevier. 29, 724--734 (2008)
21. Nixon, M.S., Aguado, A.S.: *Feature Extraction and Image processing*. Newness, London (2002)
22. Park, S.S., Yoo, H.W., Lee, M.H., Kim, I.Y., Juang, D.S.: Clustering for Image Retrieval via Improved Fuzzy ART. Springer. 743--752 (2005)
23. Steinherz, T., Rivlin, E., Intrator, N.: Offline Cursive Script Word Recognition: a Survey. *IJDAR*, Springer. 2, 90--110 (1999)
24. Zhang, T.Y., Suen, C.Y.: A Fast Parallel Algorithm for Thinning Digital Patterns. *CACM*. 27, 236--239 (1984)
25. Zhenjiang, M.: Zernike Moment Based Image Shape Analysis and its Application. *Pattern Recognition*, 21, 169--177 (2000)