

# Classification des caractères manuscrits arabes par un perceptron multi-couches

**BEDDA Mouldi, RAMDANI Messaoud et DOGHMANE Nourredine (\*)**

## Résumé

*Cette communication présente une méthode de reconnaissance des textes arabes. L'étude est basée sur l'analyse optique du texte, un prétraitement de l'image permettant de réduire le bruit et de rendre l'image binaire. La Projection horizontale isole les lignes, la projection verticale isole les mots, divise le mot en caractères et détecte la position du caractère dans le mot? Le caractère est segmenté en parties primaire et secondaire. Le fait d'identifier séparément les secondaires réduit considérablement le nombre de classes. Les descripteurs de Fourier et les moments des projections verticale et horizontale du caractère sont le paramètres du vecteur de classement. La classification est réalisée par une simulation connexioniste à base de réseaux de neurones multi-couches. Les résultats trouvés montrent l'efficacité de la méthode pour la reconnaissance des caractères manuscrits.*

## Abstract :

*This communication presents a method for the recognition of arabic texts. The study is based on optical analysis of the text, makes the image binary. The horizontal projection isolates the lines. The vertical projection isolates the words in characters. The character is segmented in primary and secondary parts. The fact of identifying separately the secondaries reduces more the number of classes. The Fourier descriptors and the moments of the horizontal and vertical projections are the parameters of the classification vector. The classification is performed with a simulated multilayer perceptron. The results show the efficiency of this method for the recognition of handwritten characters.*

**Mots clés :** *Caractères arabes, reconnaissance des caractères, réseaux de neurones, descripteurs de Fourier, moments.*

**Key words :** *Arabic characters, characters recognition, neural networks, Fourier descriptors, moments.*

# 1. Introduction:

Dans les trois dernières décennies, la lecture optique des textes et la reconnaissance des caractères manuscrits Latins, Chinois, etc. ont été le sujet des domaines de recherches actifs et ont fait l'objet de développement de plusieurs méthodes (1,2,3), mais la recherche en reconnaissance des caractères arabes n'a pas reçu la même importance. Comme les caractéristiques de la langue arabe ne permettent pas l'application directe de ces méthodes, leur reconnaissance nécessite de développement des méthodes spécifiques aux caractères arabes (4,5,6.).

On cite quelques caractéristiques de la langue arabe ;

- . l'arabe s'écrit de la droite vers la gauche
- . chaque caractère peut se présenter sous plusieurs formes selon sa position dans le mot (début, intérieur du mot, fin du mot, isolé)
- . plusieurs caractères portent des points
- . chaque caractère peut porter un ("Chedda ")
- . quelques caractères peuvent porter un ("Hamza ")
- . un caractère peut porter un ("Medda ")

FM	IM	DM	ا	Nom	FM	IM	DM	ا	Nom
ل	ل	ل	ل	LA	ا			ا	ALIF
م	م	م	م	MA	ب	ب	ب	ب	BA
ن	ن	ن	ن	NA	ت	ت	ت	ت	TA
ص	ص	ص	ص	SAD	ث	ث	ث	ث	THA
ض	ض	ض	ض	DHA	ج	ج	ج	ج	DJIM
ع	ع	ع	ع	AYN	ح	ح	ح	ح	HA
غ	غ	غ	غ	GHYN	خ	خ	خ	خ	KHA
ف	ف	ف	ف	FA	د			د	DA
ق	ق	ق	ق	QAF	ذ			ذ	DHA
س	س	س	س	SA	ر			ر	RA
ش	ش	ش	ش	CHA	ز			ز	ZA
ه	ه	ه	ه	HA	ط	ط	ط	ط	TÄ
و	و	و	و	WAW	ظ	ظ	ظ	ظ	DHA
ي	ي	ي	ي	YA	ك	ك	ك	ك	KI

Figure 1: L'alphabet arabe sous ses différentes formes. I : caractère isolé; DM : caractère au début d'un mot; IM : caractère à l'intérieur d'un mot; FM : caractère à la fin d'un mot.

Nous présentons dans cette communication l'état actuel de nos travaux sur la reconnaissance des caractères

manuscrits arabes, la méthode proposée consiste à un prétraitement de l'image permettant d'extraire des paramètres utiles pour la classification qui est réalisée par un réseau de neurones.

# 2. Segmentation et prétraitement

L'image des caractères caractérisée par une haute résolution spatiale et un fond plus clair que les caractères est filtrée morphologiquement par une opération d'ouverture (érosion + dilatation) où l'élément structurant est un carré de taille 3x3 afin de réduire l'effet du bruit, puis elle est convertie en matrice binaire ("0" pour les caractères et "1" pour le fond) par une opération de seuillage des niveaux de gris (7). Les projections horizontales et verticales de l'image sont obtenues comme suit:

Projection horizontale :

$$h(i) = \sum_j g(i,j) \quad (1)$$

Projection verticales:

$$v(j) = \sum_i g(i,j) \quad (2)$$

où g(i,j) prend deux valeurs possibles "0" ou "1", i: étant l'indice des lignes et j celui des colonnes.

La figure 2 montre les projections horizontale et verticale du caractère DAH.

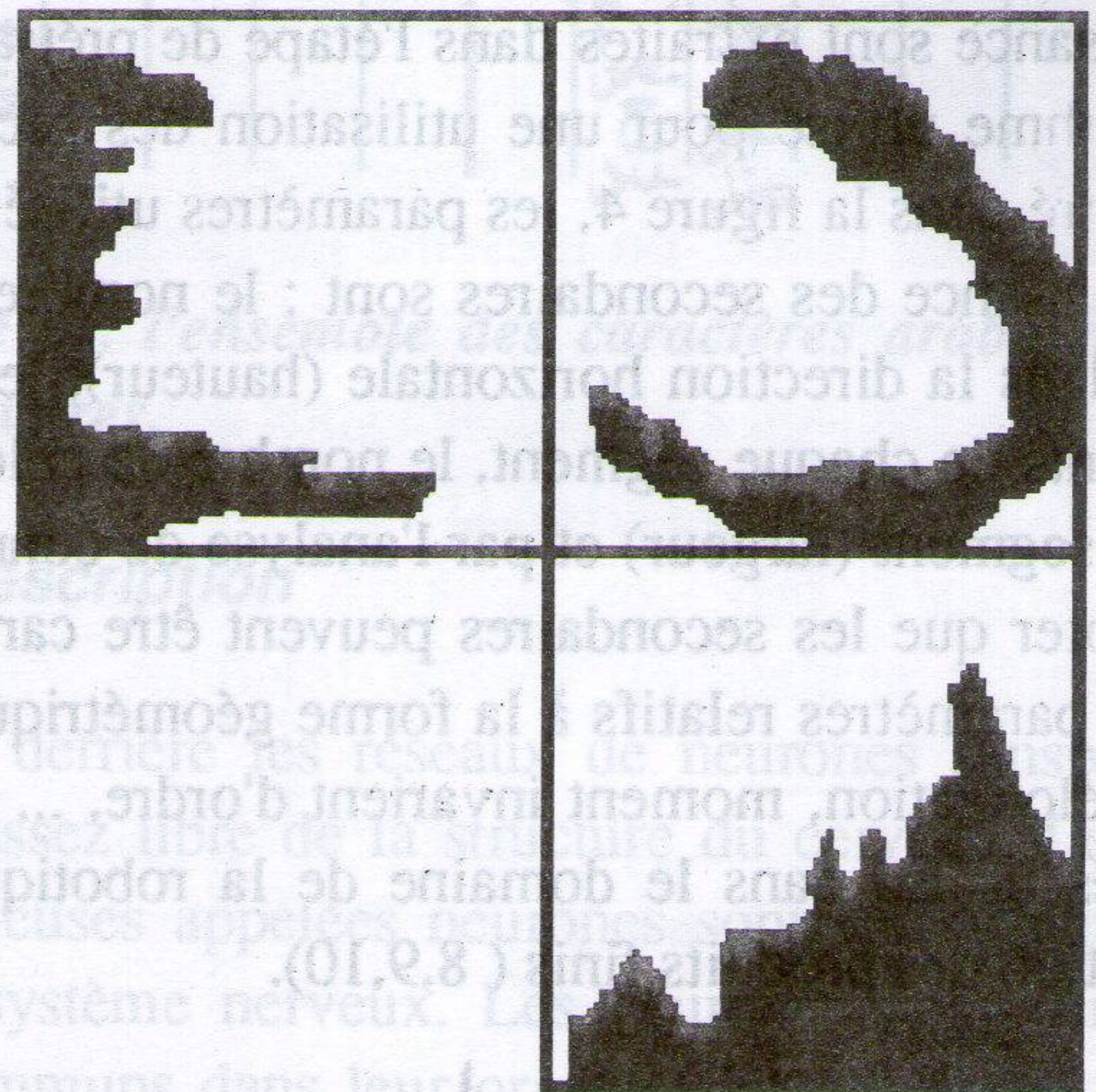


Figure 2 : Projections verticale et horizontale sur caractère DAH.

La figure 3 illustre le principe de segmentation des mots arabes en caractères en utilisant les projections verticale et horizontale comme le fait l'algorithme de Amin (6). Une fois les caractères sont isolés et localisés dans des fenêtres rectangulaires, un niveau de segmentation est appliqué en vue de confirmer l'existence et d'isoler les

secondaires ( , , , , , ) pouvant être associés à certains caractères.

Il faut signaler que les formes de base des caractères arabes sont relativement simples. Néanmoins, la position des secondaires rend leur reconnaissance de plus en plus compliquée. En effet, les parties secondaires peuvent être placées juste à proximité ou à l'intérieur en relation avec les primaires. L'objectif de la reconnaissance est donc d'isoler ces secondaires et de les identifier séparément.

Les projections horizontale et verticale étant calculées avec les formules (1) et (2), donc une simple scrutation des lignes à l'intérieur des fenêtres rectangulaires permet de séparer les segments tout en donnant une étiquette à chaque segment (1,2,3,4,5) par une simple analyse de connexité. Chaque segment sera caractérisé par le nombre de pixels noirs (PN) et le nombre de lignes (NL) pour une éventuelle classification. Ainsi le segment est classé comme un caractère primaire s'il possède un PN dépassant la moitié du nombre de PN du caractère.

### 2.1. Reconnaissance des secondaires.

Les secondaires associés aux caractères sont au nombre de six ( , , , , , ), ce qui facilite leur évaluation. Toutes les informations nécessaires à leur reconnaissance sont extraites dans l'étape de prétraitement. L'algorithme utilisé pour une utilisation des secondaires illustré dans la figure 4, les paramètres utilisés pour la reconnaissance des secondaires sont : le nombre de segments dans la direction horizontale (hauteur), le nombre d'éléments de chaque segment, le nombre de colonnes de chaque segment (largeur) et par l'analyse de connexité. Il est à noter que les secondaires peuvent être caractérisés par des paramètres relatifs à la forme géométrique (compacité, élongation, moment invariant d'ordre, ... etc), qui sont très utilisés dans le domaine de la robotique et de l'inspection des produits finis (8,9,10).

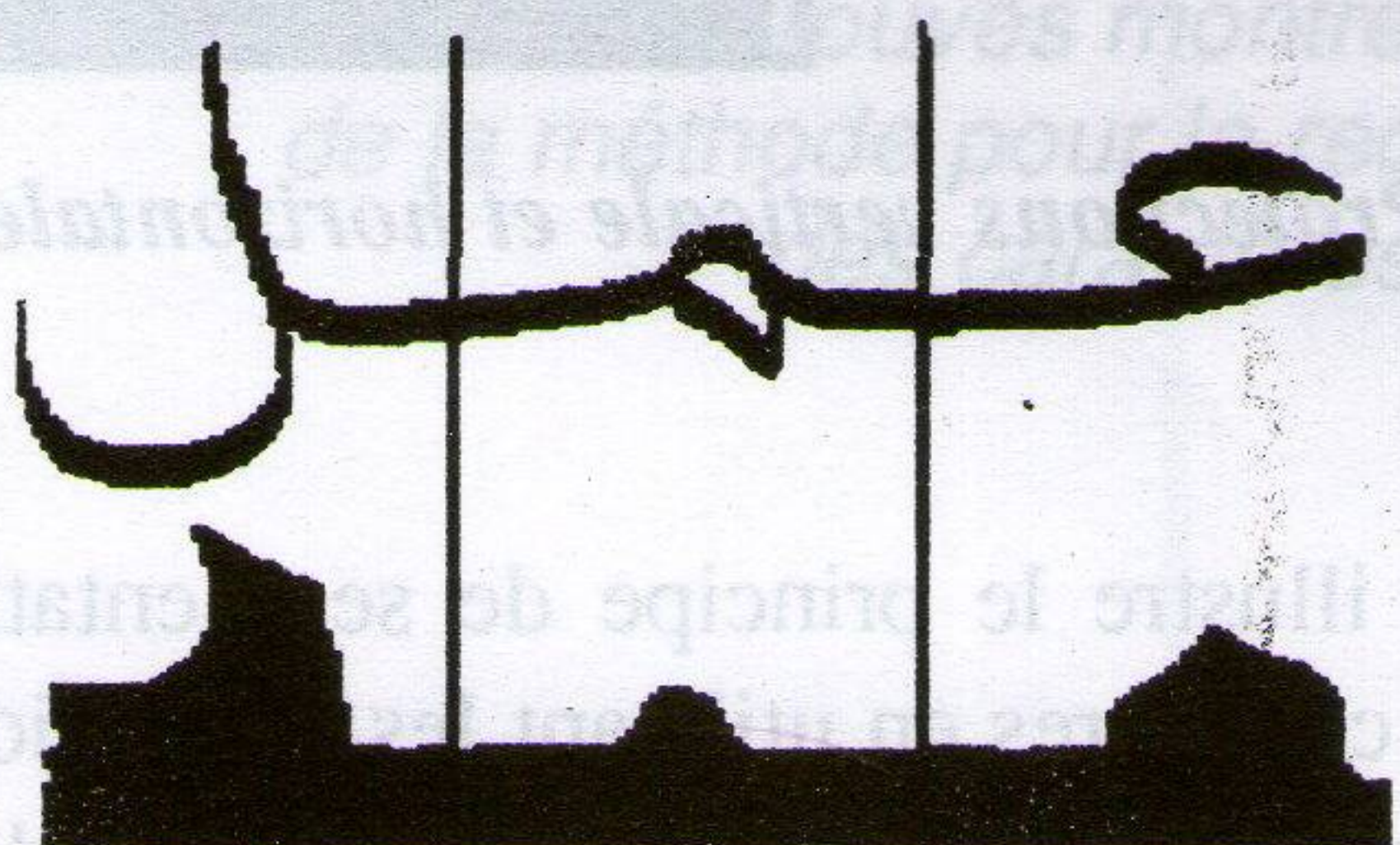


Figure 3: Segmentation d'un mot arabe en caractères par la méthode des histogrammes

Il devient très clair que l'idée d'isoler les secondaires pour les reconnaître séparément est d'une importance capitale puisqu'elle fournit une description structurale et réduit considérablement le nombre de formes à reconnaître, ce qui facilite la reconnaissance (figure 5).

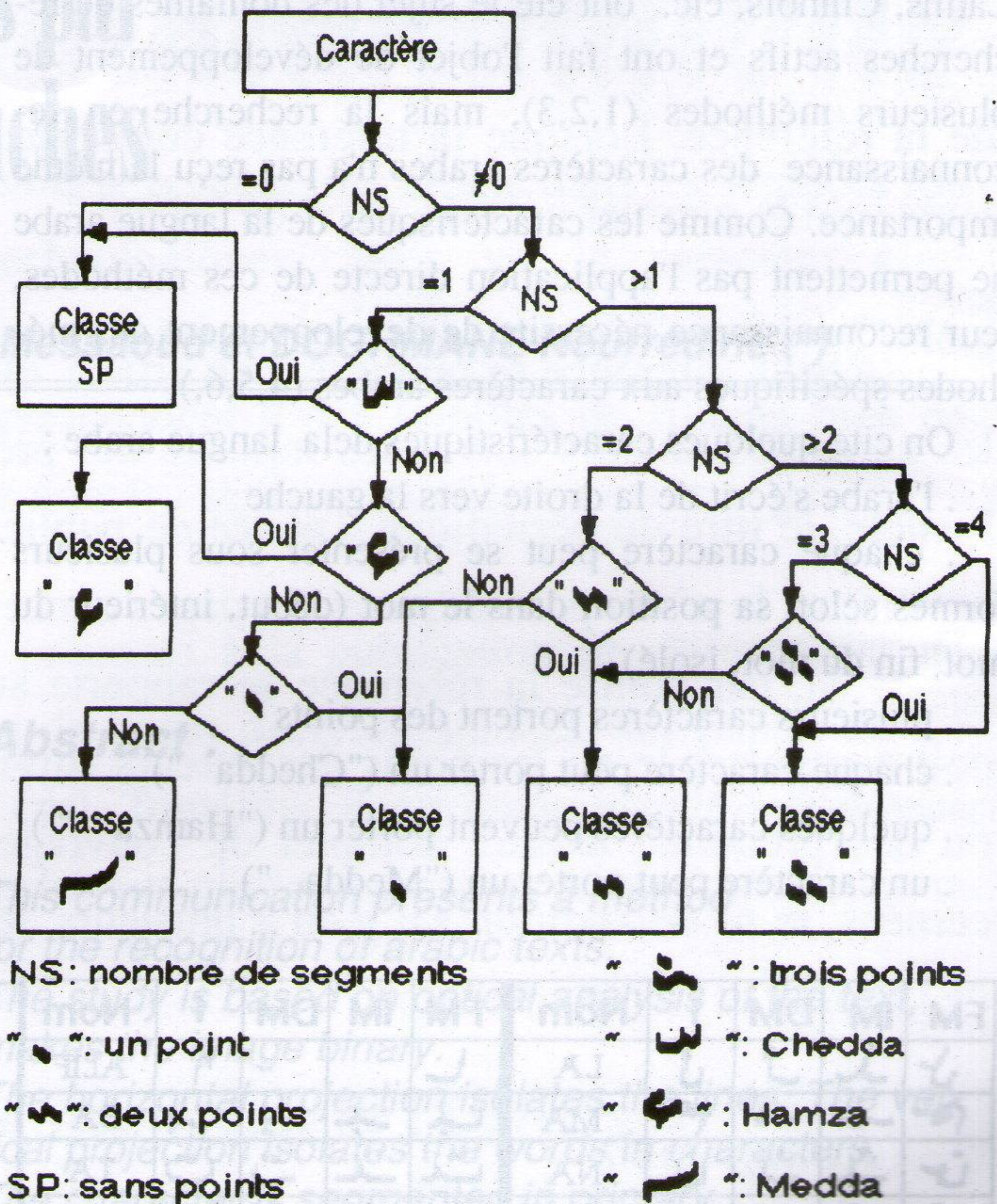


Figure 4: L'algorithme utilisé pour l'évaluation des secondaires.

Une fois les secondaires sont identifiés et, comme la position du caractère a été déjà déterminé dans l'étape de prétraitement, il reste à choisir les paramètres permettant une bonne discrimination entre les différentes classes.

### 2.2. Extraction des paramètres

Une simple revue des différentes techniques de reconnaissance de forme permet de constater l'intense utilisation des moments invariants dans le domaine de la reconnaissance de formes (8,9,11). Ainsi, des paramètres de discrimination sont extraits à partir des moments normalisés des projections horizontale ( $\mu^h$ ) et verticale ( $\mu^v$ ). Ces paramètres sont les suivants :

$$A1 = |\mu^v_1 / \mu^h_1|; A2 = |\mu^v_2 / \mu^h_2|;$$

$$A4 = |\mu^v_4 / \mu^h_4|$$

$$\text{avec } \mu_k = \sum (X_i - \bar{X})^k . P(X_i)$$

$$P(X_i) = \text{Prob} [X=X_i]$$

Par ailleurs, l'application de l'opérateur morphologique "Blur and minimum" (12), qui représente un opérateur très efficace pour la détection de contours, fournit l'image des contours des différents caractères. Ces contours qui contiennent l'essentiel de l'information sont utilisés pour faire une description des objets. Plusieurs techniques se prêtent pour décrire les courbes fermées, mais il y a des considérations théoriques et expérimentales qui montrent que les descripteurs de Fourier (DF) fournissent de bons attributs (13),(14).

Rappelons que les descripteurs de Fourier (13) sont définis comme suit : étant donné une courbe fermée  $\Gamma$  avec une représentation paramétrique  $(x(l), y(l)=z(l))$  où  $l$  est la taille d'un arc avec  $0 \leq l \leq L$  ( $L$  étant le périmètre de la courbe). Notons par  $\theta(l)$  l'angle de direction de  $l$  au point  $l$  et définissons  $\Phi(l)$  comme la fonction angulaire cumulative de telle sorte que  $\Phi(l)=\theta(l)-\theta(0)$ . On définit  $\Phi^*(l)$  par :

$$\Phi^*(l)=\Phi(L/2\pi)+1 \quad \text{avec } \Phi^*(0)=\Phi^*(2\pi)$$

où  $l$  varie de 0 à  $2\pi$ . La fonction  $\Phi^*(l)$  est ainsi invariante en translation, en rotation et en changement du périmètre  $L$ . Le développement de  $\Phi^*(l)$  en série de Fourier donne:

$$\Phi^*(l) = \mu_0 + \sum_{k=1}^{\infty} A_k \text{Cos}(k1-\alpha_k)$$

l'ensemble  $(A_k, \alpha_k, k=1, \dots, \infty)$  représente le descripteur de Fourier (DF) de la courbe  $\Gamma$ .

### 3. Reconnaissance du caractère primaire

La reconnaissance des textes manuscrits est un problème pour lequel un modèle de reconnaissance doit nécessairement prendre en compte un nombre important de variabilités, aussi ces techniques d'apprentissage peuvent elles apporter une certaine souplesse à la constitu-

tion d'un tel modèle, la technique utilisée dans cet article est les réseaux neuronaux

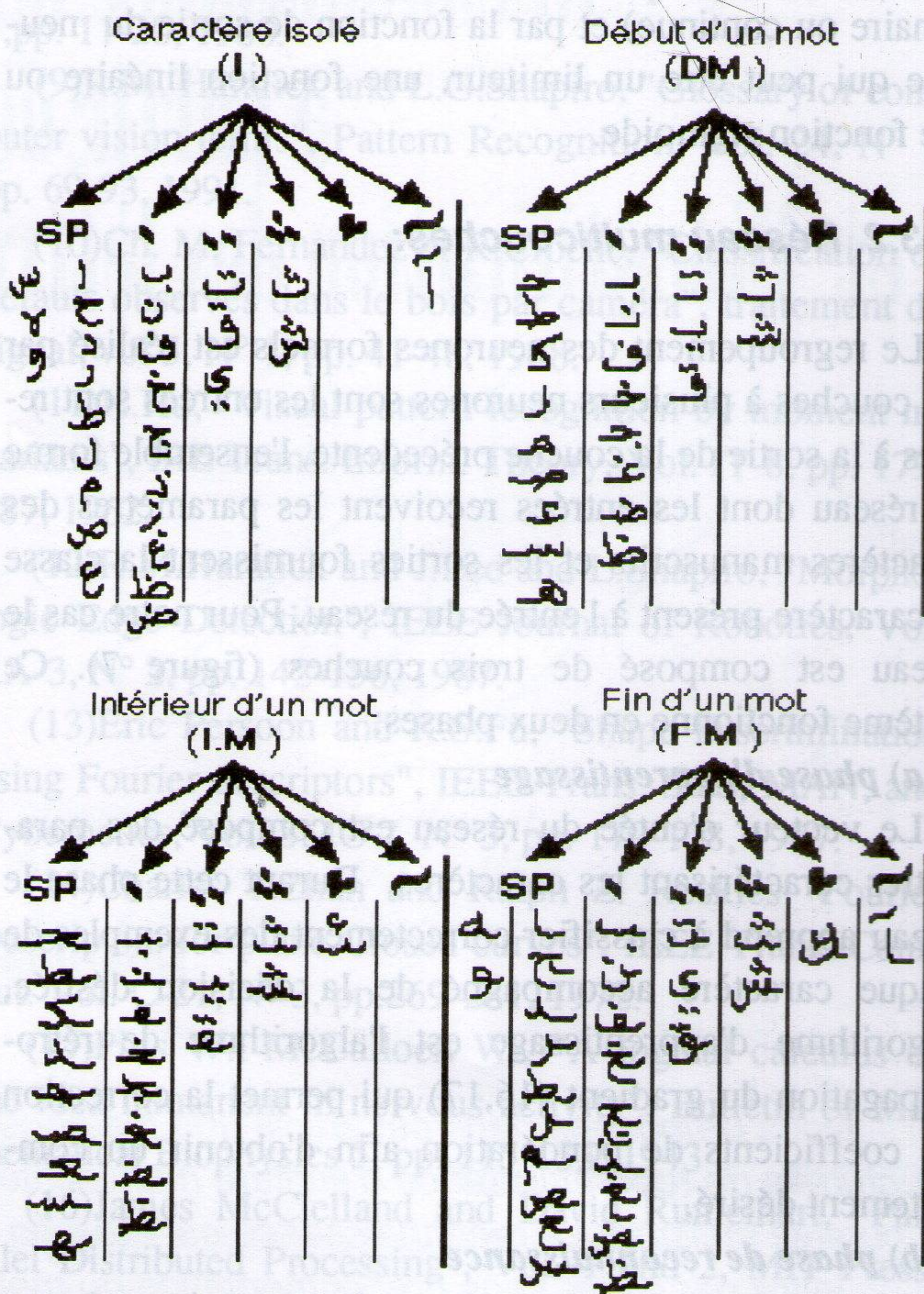


Figure 5 . l'ensemble des caractères arabes suivant les secondaires.

#### 3.1. Description

L'idée derrière les réseaux de neurones s'inspire de manière assez libre de la structure du cerveau. Les cellules nerveuses appelées neurones sont les éléments de base du système nerveux. Les neurones possèdent des points communs dans leur organisation générale, ils assurent les fonctions suivantes: recevoir les signaux et les intégrer, engendrer un flux nerveux et le conduire vers un neurone capable de le recevoir. La première modélisation d'un neurone date des années 40, elle a été proposée par MacCulloch et Pitts (15). S'inspirant de leurs travaux sur les neurones biologiques, ils ont proposé le modèle suivant : un neurone formel réalise la somme pondérée des potentiels d'actions qui lui parviennent (le potentiel est une valeur numérique représentant l'état du neurone émetteur), la sortie binaire dépend de la position de cette

somme par rapport à un seuil fixe (fig. 6)

Un neurone se caractérise par la nature de son entrée (binaire ou continue) et par la fonction de sortie du neurone qui peut être un limiteur, une fonction linéaire ou une fonction sigmoïde.

### 3.2. Réseau multicouches:

Le regroupement des neurones formels est réalisé par des couches à plusieurs neurones sont les entrées sont reliées à la sortie de la couche précédente, l'ensemble forme un réseau dont les entrées reçoivent les paramètres des caractères manuscrits et les sorties fournissent la classe du caractère présent à l'entrée du réseau. Pour notre cas le réseau est composé de trois couches (figure 7). Ce système fonctionne en deux phases:

#### a) phase d'apprentissage

Le vecteur d'entrée du réseau est composé des paramètres caractérisant les caractères. Durant cette phase le réseau apprend à classer correctement des exemples de chaque caractère accompagné de la décision désirée, l'algorithme d'apprentissage est l'algorithme de rétro-propagation du gradient (16,17) qui permet la correction des coefficients de pondération afin d'obtenir un comportement désiré.

#### b) phase de reconnaissance

Durant cette phase, on présente les paramètres du caractère à identifier à la couche d'entrée du réseau et on recueille à la sortie son affectation à l'une des M classes.

#### Sommation Seuillage

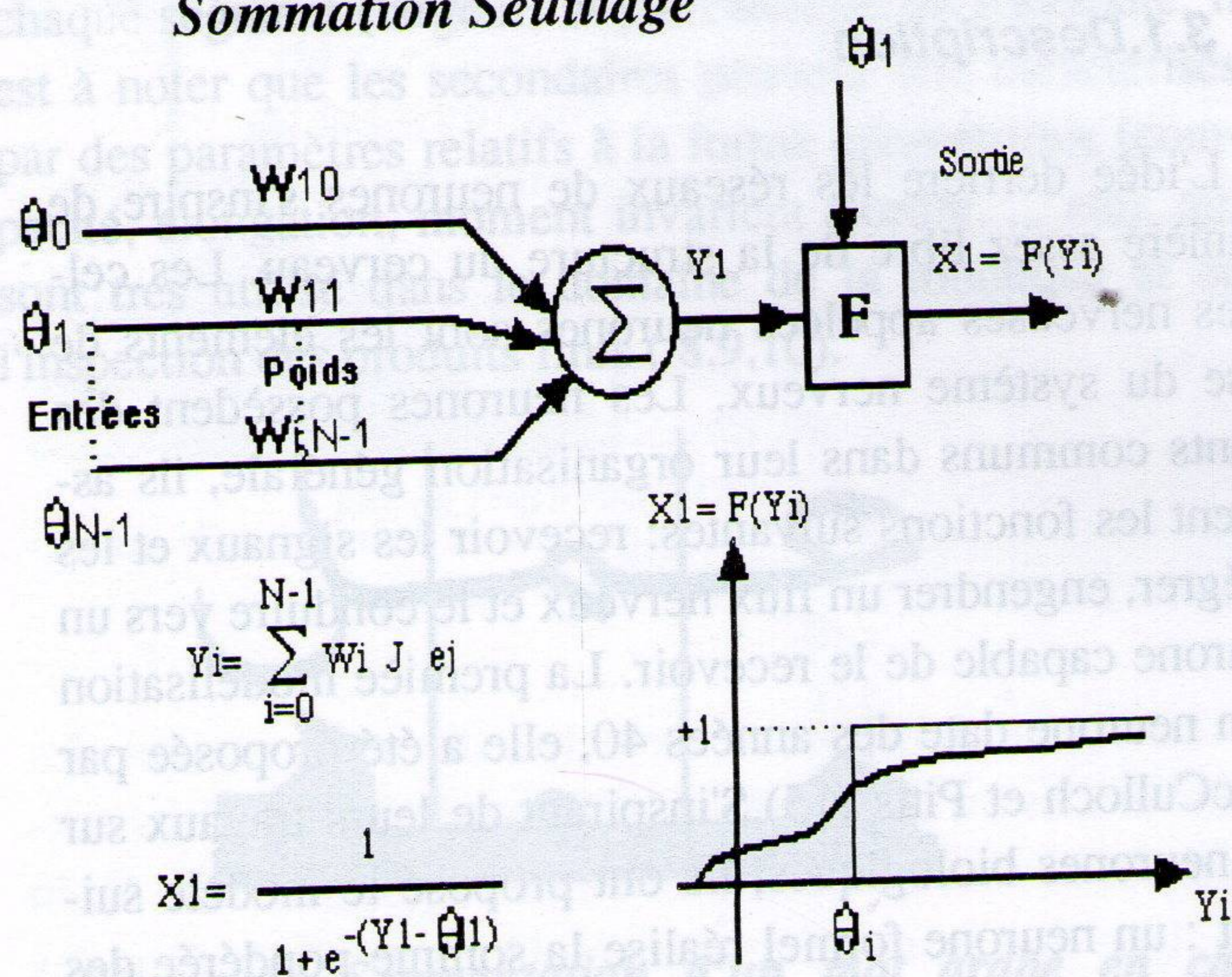


Figure 6: Modèle d'un neurone formel qui réalise une somme pondérée de N entrées analogiques

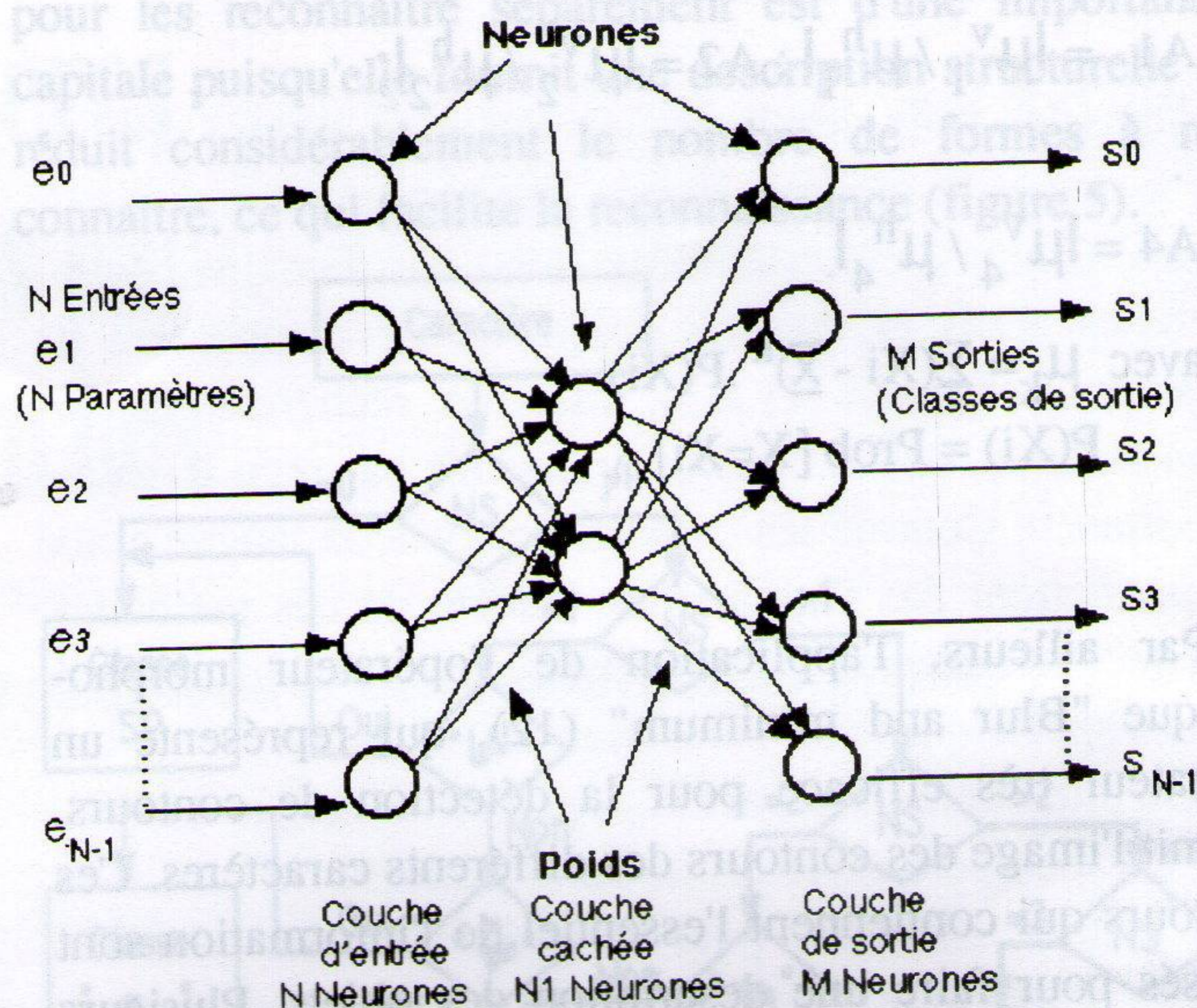


Figure 7: Réseau de neurone composé de trois couches avec N entrées analogiques et M neurones dans la couche de sortie

## 4. Résultats et discussion

Le vecteur du réseau contient 13 paramètres qui sont ; les 10 paramètres harmoniques du descripteur de Fourier et les 3 paramètres (A1, A2, A4) relatifs aux moments normalisés des projections horizontales et verticales. Le réseau retenu comporte 3 couches : 2 couches cachées chacune de 3 neurones, et une couche de sortie dont le nombre de neurones dépend de la position du caractère dans le mot et de la nature du secondaire (au maximum le nombre de neurones de la couche de sortie est de 17)

L'apprentissage de réseau a été effectué sur 18 échantillons par classe dont la qualité d'écriture est variable. Le résultat de classement de ce réseau avec des échantillons de test nous a donné un taux de reconnaissance de 93,4 %. Il est à noter que les erreurs sont facilement analysables puisqu'il s'agit de confusion des formes proches (exemple ; " ;Ha" et " ;Da"). Compte tenu que le comportement du réseau est fortement lié à son architecture, et à l'initialisation des fonctions de pondération, un choix adéquat de ces paramètres permet d'améliorer les performances du classifieur qui entraîne une augmentation du taux de reconnaissance (18, 19).

## 5. Conclusion

Cette étude a été consacrée à la reconnaissance des ca-

ractères manuscrits arabes, le fait de commencer par la détection de la position du caractère dans le mot et d'isoler les secondaires pour les identifier séparément nous a permis de réduire considérablement le nombre de classes, par exemple pour un caractère au début du mot on a :

- 9 classes pour les caractères sans point.
- 8 classes pour les caractères avec 1 point.
- 3 classes pour les caractères avec 2 points.
- 2 classes pour les caractères avec 3 points.

Les réseaux de neurones semblent être de bons candidats pour la classification, un de leur avantages est l'extraction efficace de l'information. Cependant dans la plupart des cas le sens de l'information extraite est caché. Cela favorise les méthodes neuronales pour des opérations rapides de type "boîte noire".

Pour conclure, l'objectif de travaux actuels est le traitement automatique des textes écrits en arabe par l'utilisation des réseaux de neurones. Une bonne reconnaissance reste toujours liée au choix des paramètres représentant le caractère et à l'architecture du réseau.

## Bibliographie

- (1) C.Y.Suen, M.Berthod, S.Mori, "Automatic recognition of handprinted characters - the state of the Art", Proc. IEEE; 68,4, pp. 469-487, 1980.
- (2)C.Tappert, C.Suen, TWakahara, "The state of the Art in on line Handwriting Recognition", IEEE, Trans. Patt. Anal. and Mach Intell., Vol. 12, N° 8, pp787-808, 1990.
- (3)P.H.Cheng, W.H.Hsu, M.Y.Chen, "Recognition of Handwritten chinese character by Modified Hough Transform Techniques", IEEE Trans. Patt. Anal. and Mach. Intell., Pami-11, Vol. 4, pp. 429-439, 1989.
- (4)Mohamed S. el-Wakil and Amin A. Shoukry, "On line recognition of handwritten isoleted Arabic characters", Pattern Recognition, Vol. 22, N° 2, pp. 97-105, 1989.
- (5)H.Al-Yousefi and S.S.Upda, "Recognition of arabic Characters" IEEE Trans. Patt. Anal. and Mach Intell., Vol. 14, N° 8, pp. 853-857,1992.
- (6)A.Amin and G.Masini, "Machine recognition of multifont printed Arabid texts" in Proc. 8th. Int. Conf. Patt. Recogn. Paris, pp. 392-395, 1986.
- (7)T.Pavlidis, "Algorithms for graphics and image processing", Computer Science Press, 1982.
- (8)G.J.Again, "Computer vision systems for industrial inspection and assembly", IEEE Comput., Vol. 13, N° 5,pp. 11-20, 1980.
- (9)R.M. Haralick and L.G.Shapiro, "Glossary of computer vision terms", Pattern Recognition, Vol. 24, N° 1, pp. 69-93, 1991.
- (10)Ch. M. Fernandez et R.Croche, "Classification de défauts observés dans le bois par caméra", traitement du signal, vol 5, N° 1, pp. 41-48, 1988.
- (11)K.Hu, "Visual pattern recognition by moment invariants", IRE Trans. Inform. Theory, Vol. IT-8, pp. 179-187, 1962.
- (12)R.M.Haralick and J.Lee and L.Shapiro, "Morphologic Edge Detection", IEEE Journal of Robotics, Vol. RA-3, N° 2, pp. 142-156, 1987.
- (13)Eric Persoon and K.S.Fu, "Shape Discrimination using Fourier descriptors", IEEE Trans. SYS, MAN, and Cybernetics, Vol. SMC-7, N° 3, pp. 170-178, 1977.
- (14)Charles T.Zhan and Ralph Z. Roskies "Fourier Descriptors for plane Closed curves", IEEE Trans. Comput. Vol. C-21, N° 3, pp.269-281, 1972.
- (15)Pitts W., McCulloch W., "A logical calculus of the idea immanent in nervous activity", Bulletin of Mathematical Biophysics 5, pp. 115-133, 1943
- (16)James McClelland and David Rumelhart, "Parallel Distributed Processing", Vol. 1 and 2, MIT Press, Cambridge, MA, 1986.
- (17)R.Lippmann, "An introduction to computing with Neural Nets", IEEE ASSP Magazine, pp. 4-22, 1987.
- (18)J.A.Freeman and D.M.Skapura, "Neural Networks: Algorithms, Applications, nad Programming Techniques", Addition-Wesley Publishing Company, 1992.
- (19)O.Simula, "Neural Networks and adaptive Pattern Recognition", Helsinki University of Technology (Finland), 1991.

(\*Institut d'Electronique Université Badji Mokhtar Annaba