

UN RESEAU DE NEURONES MULTICOUCHES POUR LA RECONNAISSANCE HORS LIGNE DES CARACTERES MANUSCRITS ARABES

Reçu le 20/05/2000 – Accepté le 03/06/2002

Abstract

In this paper, we present an off line method of Arabic Handwritten Characters Recognition. The study is based on the analysis of several performances of feature vectors. It is hoped that the results of the evaluation contribute to the conception of operational systems. The futures are the projection moments, Barr-features, and Fourier descriptors. The classification is achieved by a standard multi-layer perception.

Key Words: Optical character recognition, Neural networks, Barr-features, Image processing.

Résumé

Dans cet article, nous proposons une méthode connexioniste hors ligne de reconnaissance des caractères manuscrits arabes. L'étude est basée sur l'analyse des performances de plusieurs vecteurs d'attributs afin que les résultats d'évaluation contribuent à la conception des systèmes opérationnels. Les techniques utilisées pour extraire les paramètres caractérisant le caractère sont: les moments de projection horizontale, verticale et, selon les deux diagonales, les paramètres en barre et les descripteurs de Fourier. La classification est réalisée par une simulation connexioniste à base d'un réseau de neurones multicouches.

Mots clés: Reconnaissance optique des caractères, réseau de neurones, paramètres en barre, traitement d'image.

S. OUCHTATI

Département d'Electrotechnique
Université de Skikda
Route El-Hadaik
Skikda, Algérie

M. RAMDANI M. BEDDA

Département d'Electronique
Université Badji Mokhtar
Annaba, Algérie

Durant ces dernières années, les recherches et les travaux dans le domaine de la lecture optique d'une façon générale et la reconnaissance des caractères d'une façon particulière ont été l'un des domaines de recherche actif et ont fait l'objet de plusieurs publications [1,2], ce qui a donné naissance à des systèmes opérationnels dans divers domaines tels que les systèmes de lecture des montants de chèques postaux et bancaires, d'interprétation des adresses postales et d'analyse des documents [1-4]. Si pour les caractères imprimés les principales difficultés ont été surmontées, la situation est complètement différente en ce qui concerne la reconnaissance des caractères manuscrits, à cause de différents problèmes liés principalement aux variations de la forme des caractères de la même classe (variation des styles, des instruments et des fontes d'écriture) et lors de la séparation des caractères dans le mot.

Aujourd'hui, la recherche dans ce domaine s'est accrue grâce à la disponibilité des échantillons d'images ou bases de données, aux développements continus des moyens d'acquisition, de stockage et de calcul et à l'apparition des paradigmes des réseaux de neurones (RN) qui ont permis la construction des systèmes rapides et adaptatifs pour la reconnaissance des formes [5-7]. Malgré les propriétés intéressantes qu'ils possèdent (capacité d'apprentissage, d'adaptation, de généralisation et de non linéarité), peu de travaux ont été consacrés à l'évaluation des performances d'un classifieur connexioniste alimenté par plusieurs vecteurs d'attributs, extraits à partir de plusieurs représentations des caractères (image binaires ou même à niveau de gris, les contours et les squelettes). Une telle étude permet de déterminer le vecteur d'attributs pertinents pour les systèmes d'application spécifiques et de développer des procédures de sélection automatique des paramètres de discrimination pour les systèmes de plusieurs applications.

Dans cette article, nous décrivons une évaluation des performances d'un classifieur connexioniste, alimenté par plusieurs vecteur d'attributs, extraits à partir de l'image binaire du caractère et de l'image du contour.

ملخص

خلال هذا العمل، قمنا بدراسة طريقة للتعرف على الحروف العربية المكتوبة باليد، وترتكز دراستنا على تحليل نتائج عدة أشعة متغيرات التقنيات المستعملة لاستخراج المتغيرات هي: عزوم الإسقاطات الأفقية، العمودية، و حسب القطرين، المتغيرات الخطية، والمعاملات الوصفية "لفوربي". من أجل تصنيف الحروف استعملنا شبكة عصبونات متعددة الطبقات. **الكلمات المفتاحية:** القراءة الضوئية للحروف، الشبكات العصبونية، معالجة الصورة.

La base de données utilisée contient tous les caractères arabes, quelles que soient leurs positions dans le mot, avec des tailles et des épaisseurs variables et avec 20 échantillons pour chaque caractère. Cette base de données est divisée en deux parties: l'une pour *l'apprentissage* et l'autre pour *le test*. Notons que les conclusions concernant les performances du classifieur sont encore partielles, puisque celui-ci n'est pas parfaitement optimisé en termes de paramètres et de comportement désiré. Avec la disponibilité des ressources de calcul de plus en plus efficace et l'intégration du contexte, et avec une analyse des ambiguïtés entre les classes, nous préconisons l'optimisation du classifieur. Cet article est organisé comme suit: le paragraphe 1 expose l'utilisation des réseaux de neurones (RN) dans les systèmes OCR (Optical Character Recognition). Le paragraphe 2 donne les caractéristiques générales des textes arabes. Le paragraphe 3 présente les différentes composantes du système de classification. Les résultats expérimentaux sont analysés dans le paragraphe 4.

1. LES SYSTEMES OCR ET LES RESEAUX DE NEURONES

La lecture optique des caractères est la conversion de l'information écrite sur un support physique à une forme électronique. Elle souvent divisée en trois tâches: acquisition et prétraitement de l'image, extraction des paramètres et enfin classification. L'apparition du domaine de l'OCR remonte au début des années soixante où les différentes recherches ont été largement influencées par le type d'écriture à identifier et par les difficultés rencontrées dues principalement aux nombreuses variations observées sur la forme des caractères manuscrits et aux problèmes de séparation des lettres dans le mot.

L'idée à l'origine des réseaux de neurones s'inspire de manière assez libre de la structure du cerveau humain. Les cellules nerveuses appelées neurones sont les éléments de base du système nerveux, les neurones possédant 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 des réseaux de neurones date des années quarante, elle a été proposée par Pitts et MacCulloch [8].

Aujourd'hui, et malgré la multitude d'algorithmes disponibles pour la normalisation, l'extraction des paramètres et la classification, l'apparition des réseaux de neurones ont permis à la recherche en OCR de prendre de l'ampleur grâce aux propriétés intéressantes qu'ils possèdent [9-11].

2. CARACTERISTIQUES DE LA LANGUE ARABE

Contrairement à d'autres langues, peu de travaux ont été menés concernant la reconnaissance de l'écriture arabe, et les recherches n'ont commencé en fait qu'au début des années quatre-vingt [12-15]. Ce retard est dû principalement aux caractéristiques spécifiques de cette langue qui ne permettent pas l'application directe des algorithmes développés pour d'autres langues. Parmi les

caractéristiques de la langue arabe, on peut citer:

1) La langue contient 28 caractères de base, chaque caractère a 2 à 4 formes en fonction de sa position dans le mot (Fig. 1a).

EF	MF	BF	IF	Nom	EF	MF	BF	IF	Nom
ض	ضد	ضد	ض	Dahd	ا			ا	Alef
ط	ط	ط	ط	Ttaa	ب	ب	ب	ب	Baa
ظ	ظ	ظ	ظ	Tthaa	ت	ت	ت	ت	Taa
ع	ع	ع	ع	Ain	ث	ث	ث	ث	Thaa
غ	غ	غ	غ	Ghain	ج	ج	ج	ج	Jeem
ف	ف	ف	ف	Faa	ح	ح	ح	ح	Haa
ق	ق	ق	ق	Qaaf	خ	خ	خ	خ	Khaa
ك	ك	ك	ك	Kaaf	د			د	Daal
ل	ل	ل	ل	Laam	ذ			ذ	Thaal
م	م	م	م	Meem	ر			ر	Raa
ن	ن	ن	ن	Noon	ز			ز	Zain
ه	ه	ه	ه	Hhaa	س	س	س	س	Seen
و			و	Waw	ش	ش	ش	ش	sheen
ي			ي	Yaa	ص	ص	ص	ص	Saad

Figure 1a: L'ensemble des caractères arabes utilisés.

IF : Isolated form (Forme isolée).

BF : Beginning form (Forme de début).

MF : Middle form (Forme de fin).

EF : End form (forme de fin).

2) L'écriture s'effectue de droite à gauche.

3) Elle est par nature cursive, que ce soit pour l'imprimé ou pour le manuscrit.

4) La majorité des caractères ont des secondaires (Fig. 1b) qui sont: un point (.) deux points (..) trois points (:), Zigzag (ε) qui peuvent jouer un rôle très important lors de la classification (réduction du nombre de classes). Ces secondaires sont localisés au-dessus, au-dessous ou au milieu de la partie primaire.

Point	•
Deux points	••
Trois points	•••
Zigzag	ε

Figure 1b: Les secondaires.

5) Les mots sont composés de un ou plusieurs sous-mots. Un sous-mot peut être aussi composé de un ou plusieurs caractères, par exemple le mot (ولد) est composé de deux sous-mots: (و) qui représente un seul caractère et (لد) composé de deux caractères.

6) La langue possède un ensemble d'accessoires orthographiques ou diacritiques (Fig. 1c) qui peuvent changer la signification et la prononciation d'un mot, qui sont:

a/ Les voyelles brèves qui se placent au-dessus ou au-dessous d'une lettre, pour lui attribuer un son déterminé.

b/ Le *Soukoun* qui se place au-dessus de la consonne pour marquer l'absence de voyelle.

c/ Le *Tanwine* qui est un redoublement de voyelle brève.

d/ La *Shedda* qui se place au-dessus d'une lettre pour indiquer son redoublement.

Fat-hah ع	Dammah د	Kasrah ك	Sukoun س	Shadda ص	Maddah م
--------------	-------------	-------------	-------------	-------------	-------------

Figure 1c: Les signes diacritiques.

7) En plus des 28 lettres principales, la langue arabe possède des caractères additionnels tels que TAA_MARBOUTA (ة) ou LAM_ALIF (لا).

8) Plusieurs caractères arabes possèdent exactement la même partie primaire. Cependant, ils sont distincts les uns des autres à partir de la position et du type des secondaires. Par exemple, les caractères BAA (ب), TAA (ت), THAA (ث), NOON (ن), YA (ي), possèdent la même partie primaire (ـ), mais ils se différencient en fonction de la position et du type des secondaires.

La figure 1a montre une liste des caractères arabes dans leurs différentes formes d'écriture en fonction de leur position dans le mot. Le fait d'isoler les secondaires et faire leur reconnaissance séparément réduit considérablement le nombre de classes, de 100 à 57 classes. La figure 1b montre les secondaires et la figure 1c, les signes diacritiques de la langue arabe.

3. LES ETAPES DE TRAITEMENT

Habituellement un système typique de reconnaissance des manuscrits se divise en trois phases:

- Acquisition et prétraitement.
- Extraction des caractéristiques.
- Classification.

3.a. Acquisition et prétraitement

Avant d'analyser les différents traitements envisagés pour traiter les données, rappelons que nous nous intéressons particulièrement au traitement hors ligne. Dans ces conditions, les systèmes d'acquisition actuels les plus courants sont donc les scanners ou les caméras linéaires. L'acquisition est effectuée avec un scanner numérique de résolution 300 dpi avec 8 bits/pixels. Les échantillons utilisés sont des textes qui contiennent tous les caractères arabes quelles que soient leurs positions dans le mot, avec des tailles et des épaisseurs variables, et avec 20 échantillons pour chaque classe. Afin d'améliorer les résultats de la reconnaissance, l'image doit subir une opération de réduction du bruit, puis une opération de binarisation par seuillage des niveaux de gris [15]. L'étape suivante est l'isolation des caractères manuscrits. Pour cela, une série de projections horizontales et verticales est appliquée sur l'image. Les projections sont:

Projection horizontale:

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

Projection verticale:

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

où $g(i, j)$ est la valeur de l'image binaire au pixel de coordonnées (i, j) ,
i: l'indice des lignes.

j: l'indice des colonnes.

Une fois les caractères isolés 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. Les projections horizontales et verticales sont calculées à partir des expressions (1) et (2); une simple scrutation des lignes à l'intérieur des fenêtres rectangulaires permet donc de séparer les segments tout en donnant à chacun une étiquette par simple analyse de connexité, chaque segment étant caractérisé par le nombre de pixels noirs (PN). Ainsi, le segment est classé comme un caractère primaire s'il possède un PN dépassant la moitié du nombre de PN du caractère. Enfin, l'image de la partie primaire est normalisée à une taille de 64*64 pixels (Fig. 2).

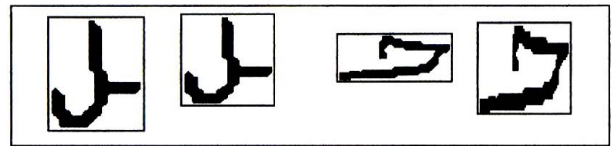


Figure 2: Caractères arabes et normalisation.

3.b. Paramètres de discrimination

Parmi les problèmes les plus importants dans la reconnaissance de forme, il y a l'extraction et le choix des paramètres de discrimination, étapes primordiales car l'image initiale est ignorée pour ne considérer que ses paramètres, les performances du système de classification dépendant directement de leurs choix. Un bon choix du vecteur d'attributs se présente comme un vecteur de nombre minimal de paramètres pour éliminer la redondance et l'ambiguïté, mais qui contienne un maximum d'informations utiles pour la classification. Parmi les techniques utilisées pour extraire les paramètres de discrimination, on trouve celles basées sur l'image binaire du caractère (utilisation des moments, techniques d'histogramme, techniques de distribution des pixels [16, 17] ou même à niveau de gris [18], ou bien celles qui sont basées sur des informations extraites à partir de l'image du contour du caractère (descripteurs de Fourier, codage des contours) [16, 19].

Dans notre travail, nous nous intéressons particulièrement aux images binaires des caractères et aux images des contours. Les méthodes utilisées pour extraire les paramètres de discrimination sont les suivantes:

3.b.1. Les paramètres en barre

Les paramètres en barre ont été utilisés avec succès dans plusieurs travaux [20, 21] et sont calculés sur des images binaires des caractères. Huit images de paramètres sont générées, chaque image de paramètres correspondant à l'une des directions Est (e), Nord (n), Nord-Est (ne), Nord-Ouest (nw) pour le fond et pour les objets, et possède une valeur entière représentant la longueur de la barre dans la direction en question. Un algorithme récursif à deux passes est utilisé pour calculer ces images. Dans la première passe, chacune des images paramètres est adaptée de gauche à droite et de haut en bas. Dans la deuxième passe, en arrière, le maximum est propagé dans le sens inverse (de bas en

haut et de droite à gauche). Un pseudo-code pour le calcul des paramètres en barre du fond est le suivant:

```

|*passe directe*|
Pour j = 1,2,...,nlig faire
Pour j = 1,2,...,ncols faire
e(i, j) = e(i, j-1) + 1
ne(i, j) = ne(i-1, j+1) + 1
n(i, j) = n(i-1, j) + 1
nw(i, j) = nw(i-1, j-1) + 1
|*passe indirecte*|
pour i = nlig, nlig - 1, ..., 1 faire
pour j = nclos, nclos - 1, ..., 1 faire
e(i, j) = max(e(i, j), e(i, j+1))
ne(i, j) = max(ne(i, j), ne(i+1, j-1))
n(i, j) = max(n(i, j), n(i+1, j))
nw(i, j) = max(nw(i, j), nw(i+1, j+1))

```

où nlig et ncols représentent les nombres de lignes et de colonnes respectivement (Fig. 3).

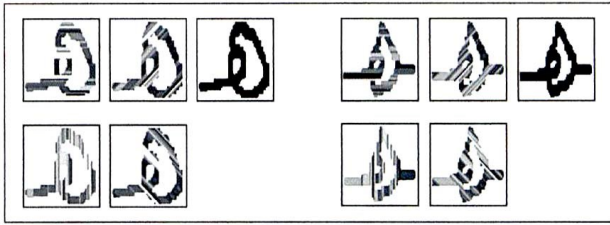


Figure 3: Le caractère Hhaa (au début et au milieu du mot) et ses quatre images paramètres.

Les vecteurs paramètres sont calculés à partir des images paramètres en utilisant des zones en chevauchement pour assurer un certain degré de lissage. Quinze zones rectangulaires sont arrangées en cinq lignes avec trois zones pour chaque ligne; chaque zone est de taille $(h/3 \times w/2)$ où h est w sont respectivement la hauteur et largeur des images. Les coins de haut à gauche des zones sont aux positions $\{(r,c) : r = 0, h/6, 2h/6, 3h/6, 4h/6 \text{ et } c = 0, w/4, 2w/4\}$. Les valeurs dans chaque zone des images de paramètres sont sommées, les sommes sont normalisées entre 0 et 1, en divisant par le maximum possible de la somme dans une zone. Le vecteur paramètre est de dimension $15 \times 8 = 120$. Soient f_1, f_2, \dots, f_8 les images paramètres associées à une forme en entrée, et soit Z_i une zone rectangulaire de taille $(h/3 \times w/2)$ avec le coin de haut à gauche (r_0, c_0) ; la valeur du $k^{\text{ème}}$ paramètre associé à la zone Z_i est:

$$Z_{ik} = \frac{1}{N} \sum_{r=r_0}^{r_0+h/3} \sum_{c=c_0}^{c_0+w/2} f_k(r, c) \quad (3)$$

où $\begin{cases} N = w & \text{si } f_k = e, ne, nw \\ N = h & \text{si } f_k = n \end{cases}$

Le vecteur de 120 valeurs ($i=1,2,\dots,15$ et $k=1,2,\dots,8$) représentent les paramètres en barre qui peuvent être utilisés comme vecteur d'entrée pour le module de reconnaissance.

4.b.2. Les moments de projection

Dans ce cas, les paramètres de discrimination sont

extraits à partir des moments normalisés des projections verticales, horizontales et selon les deux diagonales (Fig.4).

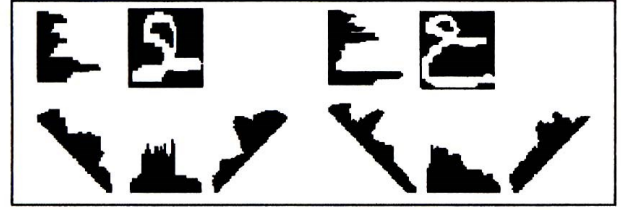


Figure 4: Le caractère Ain (à la fin et au milieu du mot) et ses différentes projections.

Les projections horizontales et verticales sont données par les expressions (1) et (2). L'algorithme utilisé pour le calcul des projections selon les deux diagonales est:

```

|* calcul de d1 *|    k = 0
pour j=0,1,..., ncols-1 faire
{ j1=j ; i=0 ;
  tant que (j1 >= 0) faire
  { d1(k) = d1(k) + g(i, j1)
    i++ ; j1-- ;
  } k++ ;
}
pour i=1,2,..., nlig-1 faire
{ i1=i ; j=ncols-1 ;
  tant que (i1 <= nlig-1) faire
  { d1(k) = d1(k) + g(i1, j)
    i1++ ; j-- ;
  } k++ ;
} |* calcul de d2 *|
k = 0
pour i=0,1,2,..., nlig-1 faire
{ i1=i ; j=ncols-1 ;
  tant que (i1 >= 0) faire
  { d2(k) = d2(k) + g(i1, j)
    i1-- ; j-- ;
  } k++ ;
} pour j=ncols-2, ..., 0 faire
{ i=nlig-1 ; j1=j ;
  tant que (j1 >= 0) faire
  { d2(k) = d2(k) + g(i, j1)
    i-- ; j1-- ;
  } k++ ;
}

```

où:

d2, d1: projections diagonales,

nlig: nombre de lignes.

Ncols: nombre de colonnes.

$g(i, j)$: valeur de l'image binaire au pixel de i, j .

Les paramètres extraits à partir des projections sont:

$$A1 = \frac{u_1^v}{u_1^h}, \quad A2 = \frac{u_2^v}{u_2^h}, \quad A3 = \frac{u_3^v}{u_3^h} \quad (4)$$

$$B1 = \frac{u_1^{d1}}{u_1^{d2}}, \quad B2 = \frac{u_2^{d1}}{u_2^{d2}}, \quad B3 = \frac{u_3^{d1}}{u_3^{d2}} \quad (5)$$

$$C1 = \frac{u_1^{d1}}{u_1^h}, \quad C2 = \frac{u_2^{d1}}{u_2^h}, \quad C3 = \frac{u_3^{d1}}{u_3^h} \quad (6)$$

$$D1 = \frac{u_1^{d1}}{u_1^v}, \quad D2 = \frac{u_2^{d1}}{u_2^v}, \quad D3 = \frac{u_3^{d1}}{u_3^v} \quad (7)$$

$$E1 = \frac{u_1^{d2}}{u_1^h}, \quad E2 = \frac{u_2^{d2}}{u_2^h}, \quad E3 = \frac{u_3^{d2}}{u_3^h} \quad (8)$$

$$F1 = \left| \frac{u_1^{d2}}{u_1^v} \right|, \quad F2 = \left| \frac{u_2^{d2}}{u_2^v} \right|, \quad F3 = \left| \frac{u_3^{d2}}{u_3^v} \right| \quad (9)$$

avec $u_k = \sum (x_i - \bar{x})^k \cdot (px_i)$ le moment d'ordre k de la projection en question:

$$p(x_i) = \text{prob}[x = x_i] \quad (10)$$

$$\bar{x} = \sum_i x_i p(x_i) \quad (11)$$

Les indices v , h , $d1$ et $d2$ désignent respectivement les grandeurs relatives aux projections verticale, horizontale et les deux diagonales.

3.b.3. Les descripteurs de Fourier

L'application de l'opérateur morphologique logique "Blur and minimum" [22], qui représente un opérateur très efficace pour la détection de contour, fournit l'image de contour des différents caractères. Ces contours, qui contiennent l'essentiel de l'information, sont utilisés pour la description des objets. Plusieurs techniques se prêtent pour décrire les courbes fermées, mais il existe des considérations théoriques et expérimentales qui montrent que les descripteurs de Fourier fournissent de bons attributs [23, 24].

Les descripteurs de Fourier sont définis comme suit: étant donné une courbe fermée avec une représentation paramétrique $(x(l), y(l) = z(l))$ où l est la taille d'un arc avec $0 < l < L$, L étant le périmètre de la courbe. Notons par $\theta(l)$ l'angle de direction de \hat{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 $\Phi^*(l) = \Phi(l/2\pi) + 1$ avec $\Phi^*(0) = \Phi^*(2\pi)$ où l varie de 0 à 2π . 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:

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

L'ensemble $(A_k, k, k=1, \dots, \infty)$ représente les descripteurs de Fourier (DF) de la courbe Γ .

3.c. Reconnaissance du caractère

Rappelons que les formes de base des caractères arabes sont relativement simples. Or, l'existence des secondaires rend leur reconnaissance de plus en plus compliquée; dès lors, l'isolation et l'identification séparément des secondaires réduisent considérablement le nombre de classes et rend le processus de reconnaissance plus simple. La reconnaissance du caractère arabe se divise donc en deux parties:

3.c.1. Reconnaissance de la partie primaire

La reconnaissance des caractères manuscrits est un problème pour lequel un modèle de reconnaissance doit nécessairement intégrer la prise en compte d'un nombre

important de variabilité. Dès lors, les méthodes basées sur les techniques d'apprentissage peuvent apporter une certaine souplesse à la construction d'un tel modèle. Dans notre travail, le choix a porté sur l'algorithme d'apprentissage par rétro-propagation du gradient, associé à un réseau de neurones multicouches dont la structure est fixée par tâtonnement. Ce choix trouve sa justification dans la capacité de ce type de modélisation capable de comprendre des phénomènes non linéaires et de complexités considérables [25, 26]. Rappelons que le modèle est généré seulement pour les parties primaires des caractères (caractères possédant des formes d'écriture similaires, mais dont les parties secondaires sont différentes). Ce modèle fonctionne en deux phases:

1. Phase d'apprentissage

L'apprentissage est le processus d'adaptation des facteurs de pondération pour une classification optimale. Le réseau utilise un apprentissage supervisé pour lequel l'algorithme de rétro-propagation du gradient minimise l'erreur quadratique moyenne entre les valeurs de la sortie actuelle et celles de la sortie désirée. Le plus important dans cette phase est de savoir quand il faut arrêter l'apprentissage, tout en s'assurant que le réseau est capable de généraliser, en évitant les problèmes causés par l'apprentissage (sur-apprentissage, oubli catastrophique).

2. Phase de test

Durant cette phase, on présente les paramètres du caractère à identifier (partie primaire) à la couche d'entrée du réseau, et on recueille à la sortie son affectation à l'une des classes possibles. Les résultats de cette phase peuvent être:

Caractère reconnu: le système arrive à associer un et un seul prototype au caractère à reconnaître.

Caractère ambigu: le système propose plusieurs prototypes avec des confiances de même ordre.

Caractère rejeté: le système ne prend aucune décision de classification.

3.c.2. Reconnaissance des secondaires

Les secondaires associés aux caractères arabes sont au nombre de quatre ('.', '...', '...', '...'), ce qui facilite leur évaluation. La reconnaissance des secondaires s'effectuent par le biais des paramètres suivants: le nombre de segments dans la direction horizontale (hauteur) et le nombre d'éléments et de colonnes de chaque segment, ainsi que par l'analyse de connexité. Notons qu'il est possible d'estimer le nombre des points par deux méthodes:

- Estimation du nombre des contours extrêmes moins un (en écartant le contour de la partie primaire).
- Estimation de la surface totale des points d'un caractère en divisant par la surface moyenne d'un point (qui est estimé dans une phase préliminaire en moyennant toutes les surfaces de tous les caractères utilisés dans la phase d'apprentissage).

Le nombre de points est considéré comme le maximum des deux estimés. Cette étape est nécessaire, puisque les points de quelques caractères sont liés à cause du bruit ou de l'effet de l'opération de lissage. De plus, quelques fontes

arabes utilisent une ligne au-dessus des caractères pour indiquer deux points et une forme en V inversé pour montrer trois points. Notons aussi que la présence ou l'absence d'un trou, le nombre de trous et l'élongation d'un trou peuvent apporter une information complémentaire permettant de résoudre certaines ambiguïtés.

RESULTATS EXPERIMENTAUX

Dans cet article, la base de données utilisée contient 1140 caractères arabes manuscrits dont la taille, l'épaisseur et la forme sont variables (Fig. 5). Une partie de cette base (10 échantillons de chaque classe dont la qualité d'écriture est variable), choisie aléatoirement, est utilisée pour l'apprentissage et l'entraînement du module de reconnaissance (le réseau de neurones) et l'autre partie, pour l'évaluation des performances.

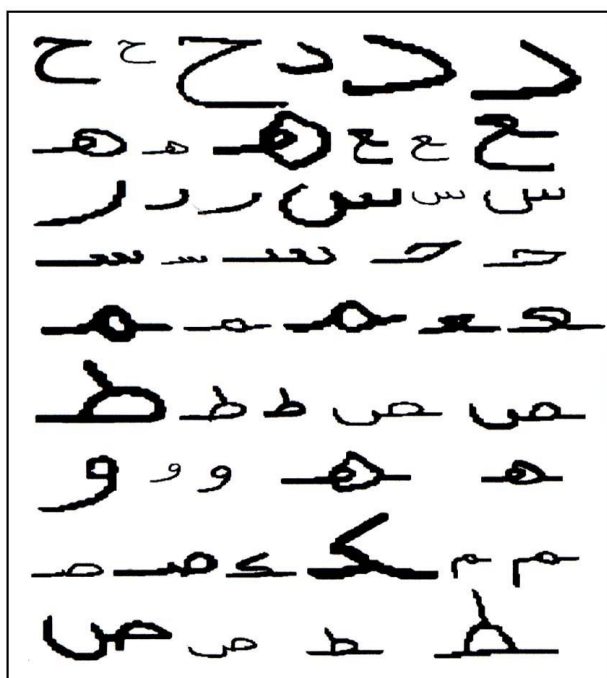


Figure 5: Quelques échantillons de la base de données utilisée.

Le système commence par identifier le type du caractère (début, fin, intérieur, ou bien isolé) par la détection de ligatures. Les vecteurs d'attributs sont composés des paramètres extraits par les méthodes décrites précédemment (Tab. 1). Le réseau de neurones utilisé est composé de trois couches: une couche cachée dont le nombre de neurones est choisi avec des heuristiques, puis affiné selon les résultats; une couche d'entrée dont le nombre de neurones dépend de la taille du vecteur d'entrée utilisé; enfin, une couche de sortie dont le nombre de neurones dépend du position du caractère dans le mot (au maximum, le nombre de neurones de la couche de sortie est de 18 en considérant le caractère LAM_ALIF (ﻻ) comme un caractère isolé). Les résultats de classement (Tab. 2) de ce réseau avec les échantillons d'apprentissage et des échantillons test ont donné des taux de reconnaissance relativement élevés (94,06 pour le taux de reconnaissance le plus élevé qui correspond au vecteur d'attributs V3). Il à

noter que les erreurs sont facilement analysable, puisqu'il s'agit de confusions entre quelques formes proches, ou bien de rejet des caractères dont la forme est dégradée.

Finalement, le taux de reconnaissance et le taux d'ambiguïté sont calculés par rapport à la partie de la base qui n'a pas été rejetée.

paramètre	V1	V2	V3	V4
	nombre	nombre	nombre	nombre
P-b	30	20	20	15
M-P	10	10	10	10
D-F	0	0	0	20

Tableau 1: Vecteurs d'attributs V1, V2 et V3.

P-b : paramètres en barre.

D-F : descripteurs de Fourier.

M-P : moments des projections.

	V1	V2	V3	V4
T-R (en %)	89.35	92.30	94.06	93.08
T-J (en%)	7.71	3.07	3.94	1.31
T-A (en%)	10.64	7.69	5.93	6.92

Tableau 2: Les taux pour les différents vecteurs l'attributs.

T-R : taux de reconnaissance.

T-J : taux de rejet.

T-A : taux d'ambiguïté.

CONCLUSION

Nous avons présenté une étude dans le domaine de la reconnaissance des caractères manuscrits arabes, basée sur l'analyse et l'évaluation des performances d'un classifieur connexionniste alimenté par plusieurs vecteurs d'attributs dont les paramètres sont extraits par plusieurs méthodes. Pour chaque vecteur d'attributs, le réseau de neurones prend une décision basée sur les caractéristiques globales et locales qu'il exploite. Notons que dans cette article, la comparaison des différentes méthodes a été faite dans le même formalisme neuronal. Nous prévoyons plusieurs possibilités d'évolution de ces travaux:

- Elargir la base de données en tenant compte d'un plus grand nombre de scripteurs et d'instruments d'écriture.
- Envisager la possibilité de rendre les phases de reconnaissance et d'extraction des paramètres interconnectées, c'est-à-dire que l'intégration d'un tel ou tel ensemble de paramètres est conditionnée par une certaine mesure de reliabilité.
- Améliorer les performances et les taux de reconnaissance de chaque méthode en faisant un choix convenable des paramètres du réseau (architecture du réseau, initialisation des paramètres de pondération, fonction d'activation, pas d'apprentissage), et par l'utilisation d'autres lois d'apprentissage.
- Envisager d'autres méthodes de classification.

REFERENCES

- [1]- Mori S., Suen C.Y. and Yamamoto K., "Historical Review of OCR Research and Development", *Proceedings of the IEEE*, Vol. 80, n°7, (1992), pp. 1029-1058.

- [2]- Suen C.Y. *et al.*, "Building a new of generation of handwritten recognition systems", *Pattern Recognition Letters*, Vol. 14, (1993), pp. 303-315.
- [3]- Suen C.Y., Naddal, Legault R., Mai T.A., and Lam L., "Computer Recognition of Unconstrained Handwritten Numerals", *hocc. IEEE*, Vol. 80, n°7, July (1992), pp. 1162-1180.
- [4]- Srihari S.N., "Recognition of handwritten and machine printed text for postal interpretation", *Pattern recognition letters*, Vol. 14, (1993), pp. 291-302.
- [5]- Lippmann R.P., "An introduction to computing with neural nets", *IEEE ASSP Magazine*, Apr. (1987), pp. 4-22.
- [6]- Richard M.D. and Hippman R.P., "Neural network classifiers estimate Bayesian a posteriori probabilities", *Neural computation*, Vol. 3, (1991), pp. 461-483.
- [7]- Sung-Bae Cho and Kim Jim H., "Multiple network fusion using fuzzy logic", *IEEE Trans. Neural networks*, Vol. 6, (1995), pp. 497-501.
- [8]- Pitts W., McCulloch W., "A logical calculus of the idea immanent in nervous activity", *Bulletin of Mathematical Biophysics*, 5, (1943), pp. 115-133.
- [9]- Jain A.L., Jianchiang Mao, Mohiuddin K.M., "Artificial neural networks: a tutorial", *Computer*, Vol. 29, n°3, (1996), pp. 31-44.
- [10]- Harrien Hausen-Muhlbauer B., Koop A., "Script a prototype for the recognition of continuous, cursive, handwritten input by means of neural network simulator", *IEEE International Conference on Neural Networks*, Vol. 3, (1993), pp. 1672-1677.
- [11]- Yanikoghu B.A. Sandon P.A., "Recognizing off-line cursive handwritten", *Proceeding IEEE computer society conference on computer vision and pat. Rec.*, (1994), pp. 397-403.
- [12]- Abuhaiba I.S.I., Mahmoud S.A., Geen R.J., "Recognition of handwritten cursive Arabic characters", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, pp. 664-672, 1994.
- [13]- Al-Badr B., Mahmoud S.M., "Survey and bibliography of Arabic optical text recognition", *Signal processing*, Vol. 41, (1995), pp. 49-77.
- [14]- Mahmoud S.A., "Arabic character recognition using Fourier descriptors and character contour encoding", *Pattern recognition*, Vol. 27, n°6, (1994), pp. 815-824.
- [15]- Amin A., Al-Sadoum H. and Fischer S., "Hand-hinted Arabic Character Recognition System using an artificial network", *Pattern Recognition*, Vol. 29, n°4, (1996), pp. 663-375.
- [16]- Pavalidis T., "Algorithms for graphics and image processing", *Computer Science Press Rockville, Maryland*, (1982).
- [17]- Bedda M., Ramdani M., Doghmane N., "Classification des caractères arabes manuscrits par une perception multicouches", *Revue Synthèse*, Université de Annaba, Vol. 1, N°1, Juin (1996).
- [18]- Bedda M., Ramdani M. et Ouchtati S., "Sur le choix d'une représentation des caractères manuscrits arabes", 2^{ème} Conférence Internationale Signaux, Systèmes et Automatiques SSA2'99, Université de Blida, Algérie, 10-12 Mai (1999), pp. 73-84.
- [19]- Pettier J.C., Camille Rapp J., "Reconnaissance statique des mots manuscrits", *Reconnaissance automatique de l'écriture*, Bigre 68, Le Havre, Mai (1990), pp. 136-147.
- [20]- Gader P.D., Mohamed M., and Chiang J., "Comparison of Crispand Fuzzy Character Networks in Handwritten Word Recognition", *IEEE Trans. Fuzzy Systems*, Vol. 3, n°3, (1995), pp. 357-364.
- [21]- Magdi M. and Gader P., "Handwritten Word Recognition Using Segmentation-Free Hidden Markow Modeling and Segmentation Based Dynamic Programming Techniques", *IEEE PAMI*, Vol. 18, n°5, May (1996), pp. 548-554.
- [22]- Haralick R., Lee J. and Shapiro L., "Morphologic edge detection", *IEEE Journal of Robotics*, Vol. RA-3, n°2, (1987), pp. 142-156.
- [23]- Eric Person and K.S.Fu, "Shape Discrimination using Fourier Descriptors", *IEEE Trans. SYS, MAN, and Cybernetics*, Vol. SMC-7, n°3, (1977), pp. 170-178.
- [24]- Charles T. Zhan and Ralph Z. Roskies, "Fourier Descriptors for plane closed curves", *IEEE Trans. Comput.*, Vol. C-21, n°3, (1972), pp 261-281.
- [25]- Lemarié B., "Practical Implementation of Radial Basis Function Network for Handwritten Digit Recognition", *In: Second International Conference on Document Analysis and Recognition (ICDAR, 93)*, Tsukuba, Japan, (1993), pp. 412-415.
- [26]- Sung-Bae Cho, "Neural network classifier for recognizing totally unconstrained handwritten numerals", *IEEE Trans. Neural network*, Vol. 8, n°1, (1997). pp. 43-53. □