

République Algérienne Démocratique et Populaire Ministère
de l'Enseignement Supérieure et de la Recherche Scientifique
Université d'Adrar
Faculté des Sciences et de la Technologie
Département des Mathématiques et Informatique



Mémoire de Fin d'étude
En vue de l'Obtention du Diplôme de Master en Informatique

Option : Réseaux et systèmes intelligents

Thème

Reconnaissance de l'écriture arabe manuscrite en utilisant les réseaux neuronaux

Préparé par

M^{elle}. BELBALI Malika et M^{elle}. KADOURI Karima

Encadrées par

Mr. MAMOUNI Elmamoun

Année Universitaire 2015 / 2016



Dédicace

Je dédie ce mémoire à

*A mes très chers parents, Auxquels je dois ce que je suis. Que
Dieu vous protège*

*à tous ceux et toutes celles qui m'ont accompagné et soutenu
durant la réalisation de ce travail*

Malika



Dédicace

Je dédie ce travail à:

*Mes chers parents notamment mon père qui m'a beaucoup
soutenu.*

et l'esprit de ma mère; Source de vie

Mes frères et mes sœurs.

Mes oncles et mes cousins

Ma belle famille.

Mes amies

Tout ce qui porte le nom de KADOURI.

Karima

Remerciement

On remercions tout d'abord notre clément Dieu qui nous a donné la puissance pour que nous puissions terminer ce travail.

Nous adressons nos remerciements les plus sincères, à notre professeur Mr. MAMOUNI Elmamoun, qui a très volontiers accepté d'être le promoteur de ce projet. Sa grande connaissance dans le domaine, ainsi que son expérience, ont joué un rôle important dans la conception de ce travail. Il trouve ici, le gage de mon infinie gratitude.

Enfin, nous tenons à remercier toutes les personnes qui nous ont conseillé et relu lors de la rédaction de ce travail

Résumé

Le présent travail porte sur une étude concernant le domaine de reconnaissance optique de mots arabes manuscrits, particulièrement la reconnaissance hors-lignes. Une étude générale sur les systèmes de reconnaissance de l'écriture a été développée, puis elle a été affinée par un intérêt particulier à une phase considérée comme cruciale dans le procédé de reconnaissance c'est la phase de segmentation. Nous avons présenté la méthode de projection vertical comme une méthode de segmentation et les réseaux de neurones pour la reconnaissance basant sur l'apprentissage supervisé.

Mots-clés: reconnaissance, écriture manuscrite, segmentation, apprentissage, réseaux de neurones, perceptron et perceptron multicouches.

Abstract

This work focuses on a study of the field of optical recognition of handwritten Arabic words, particularly the recognition off-line. A general study of handwriting recognition systems has been developed and it has been refined with a special interest in a phase considered crucial in the process of recognition this segmentation phase. We presented the vertical projection method as a method of segmentation and neural networks for recognition based on supervised learning.

Keywords: recognition, handwriting, segmentation, learning, neural networks, perceptron and multilayer perceptron.

ملخص

يركز هذا العمل على دراسة مجال التعرف الضوئي على الكلمات العربية المكتوبة بخط اليد. وقد تم وضع دراسة عامة لأنظمة التعرف على الكتابة اليدوية وتم تكريره مع اهتمام خاص في مرحلة تعتبر حاسمة في عملية التعرف، هذه المرحلة تتمثل في تجزئة الكلمات إلى حروف. قدمنا طريقة الإسقاط الرأسية كوسيلة للتجزئة، فيما اعتمدنا على الشبكات العصبية للتعرف على أساس خوارزميات التعلم. الكلمات المفتاحية: التعرف، الكتابة بخط اليد، التجزئة، التعلم، الشبكات العصبية، المستقبلات والمستقبلات متعددة الطبقات

sommaire

Dédicace	I
Remerciement	III
Résumé	IV
Liste des Figures	VIII
Liste des tableaux	X
Introduction générale	1
Chapitre I: Reconnaissance de l'écriture arabe manuscrite	
I.1 Introduction	4
I.2 Caractéristiques de l'écriture arabe manuscrite.....	4
I.3 Différents aspects de l'OCR.....	8
3.1 Reconnaissance en-ligne et hors-ligne.....	9
3.2 Reconnaissance globale ou analytique	10
I.4 Processus de reconnaissance de mots arabes manuscrits	11
4.1 Phase d'acquisition.....	12
4.2 Phase de prétraitement.....	12
34 Phase de segmentation.....	15
4.4 Phase d'extraction des caractéristiques	17
3.5 Phase de classification.....	21
4.6 Phase de post traitement	22
I.5 Problèmes liés à l'OCR	22
I.6 Conclusion	23
Chapitre II: Etat de l'art de la segmentation	
II.1. Introduction	25
II.2 La segmentation de textes arabes (particularités et difficultés	25
2.1 L'ambiguïté vocalique des mots.....	25
2.2 L'ambiguïté dérivationnelle	26
2.3 L'ambiguïté structurelle	26
2.4 L'utilisation des signes de ponctuation	26

sommaire

2.5 L'agglutination	27
II.3 Les niveaux de segmentation	27
3.1 Segmentation de la page	28
3.2 Segmentation d'un bloc de texte en lignes.....	28
3.3 Segmentation de lignes en mots	28
3.4 Segmentation de mots en caractères.....	29
II.4 Les méthodes de segmentation d'écriture cursive	29
4.1 Segmentation à partir du squelette	30
4.2 Segmentation à partir du contour	30
4.3 Segmentation à partir des histogrammes.....	31
4.4 Segmentation basée sur des réservoirs	31
4.5 Segmentation basée sur les fenêtres glissantes.....	31
4.6 Segmentation en bandes verticales.....	32
II.5 Certains travaux de segmentation de mots arabes en caractères.....	32
5.1 Algorithme proposé par Zoubeir Mouelhi	32
5.2 Algorithme proposé par A.Hamid et R.Haraty	33
5.3 Algorithme proposé par S.T. Masmoudi et al « PHMM » (modèles Markoviens cachés planaires	33
5.4 Algorithme proposé par Sari et al « ACSA » (Arabic Character Segmentation Algorithm	34
II.6 Conclusion.....	34
Chapitre III: les réseaux de neurones	
III.1 Introduction.....	36
III.2 Historique.....	36
III.3 Le neurone biologique	37
3.1 Définition	37
3.2 Système nerveux biologique	38
3.3 Réseaux de neurones et neurosciences	38
III.4 Les réseaux de neurones artificiels	39
4.2 La mise en œuvre d'un réseau de neurones artificiel	40

sommaire

4.3 Les différentes configuration des réseaux	41
4.4.1 Modèle de McCullochs et Pitts (1943).....	43
4.4.2 Le Perceptron simple (Modèle de Rosenblatt, 1950.....	43
4.4.3 Le Perceptron Multi Couches (PMC	44
III.5 l'Apprentissage.....	47
5.2 Apprentissage par renforcement.....	48
5.4 Règles d'apprentissage	49
Widrow-Hoff (delta rule.....	50
III.6 L'apprentissage d'un PMC.....	51
III.8 Les domaines d'application de réseaux de neurones.....	56
Chapitre IV: Résultats expérimentaux et discussions	
IV.1 Introduction	58
IV.2 Environnement de développement	58
IV.3Description des bases de données.....	58
IV.4 Description de notre système de reconnaissance.....	61
IV.5 Tests et résultats.....	67
IV.6 Conclusion.....	72
Conclusion générale	74
Bibliographies	76

Liste des figures

Liste des Figures

Figure I.1 la cursivité de l'écriture arabe	05
Figure I.2 Exemple de pseudo-mots constituant des mots arabes	06
Figure I.3 L'article défini " ال " dans deux mots arabes	08
Figure I. 4 Schéma général d'un système de reconnaissance de caractères.....	12
Figure I.5 effets de certaines opérations de prétraitement	13
Figure I.6 Exemple de Binarisation adaptative.....	13
Figure I.7 : Exemple de normalisation de mots manuscrits.....	15
Figure II.1 Exemple d'une combinaison de mot arabe	26
Figure II.2 le processus de segmentation.....	27
Figure II.3 Détection des différentes zones d'une page de document	28
Figure II.4 Segmentation de texte en lignes	28
Figure II.5 Segmentation de Ligne en Mots	29
Figure II.6 Segmentation de mot en Caractères.....	29
Figure II.7 Segmentation à base de squelette: basée sur des motifs.....	30
Figure II.8 Extrema du contour supérieur et inférieur	30
Figure II.9 Segmentation à partir d'histogrammes de projection selon plusieurs directions.....	31
Figure II.10 Segmentation à base de fenêtre glissante : découpage du mot en bandes verticales	31
Figure II.11 Segmentation en bandes verticales uniforme et non-uniforme	32
Figure III.1 Schéma d'un neurone biologique	37
Figure III.2 Schéma d'un Système nerveux biologique	38
Figure III.4 Réseau à connexions locales	42
Figure III.5 Réseau à connexion complète	42
Figure III.6 Réseau à connexions récurrentes.....	43
Figure III.7: Le Neurone de Mac Culloch-Pitts.....	43
Figure III.9 Le perceptron de Rosenblatt(1950)	44
Figure III.10 Architecture d'un Perceptron Multi-couches une seule couche cachée	45
Figure III.11 Exemples de fonctions de transfert d'un PMC	46
Figure III.12 schéma représente quelque types de fonction de transfert	46

Liste des figures

Figure III.13 Schéma représente l'apprentissage supervisé	48
Figure III.14 Schéma représente l'apprentissage non supervisé	49
Figure IV.1 des échantillons de la base des caractères	59
Figure IV.2 des échantillons de notre base	60
Figure IV.3 des échantillons de la base IFN/ENIT	61
Figure IV.4 Schéma générale de notre système de reconnaissance.....	62
Figure IV.5 L'organigramme d'algorithme de segmentation.....	64
Figure IV.6 Résultat de segmentation de mot " سعيدة "	65
Figure IV.7 Exemple de matrice de distribution (7*7) de la lettre alphabet arabe «khaa	66
Figure IV.8 Interface de création des corpus.....	67
Figure IV.9 : Résumé des résultats de segmentation des mots de notre base.....	68
Figure IV.10 le taux de segmentation globale pour les mots de notre base	69
Figure IV.11 Résumé des résultats de segmentation des mots de la base IFN /ENIT	69
Figure IV.12 le taux de segmentation globale pour les mots de la base IFN /ENIT	70
Figure IV.13 le taux de reconnaissance pour les caractères	71
Figure IV.14 le taux de reconnaissance pour les mots de notre base	71
Figure IV.15 le taux de reconnaissance pour les mots de la base IFN /ENIT	71

Liste des tableaux

Liste des tableaux

Tableau I.1 L'alphabet arabe	05
Tableau I.2. Lettres arabes ayant des points diacritiques	07
Tableau I.3. Différents hampes et jambages dans les lettres arabes	08
Tableau III.1: Analogie entre le neurone biologique et le neurone formel	40
Tableau VI.1 Les lettres arabes utilisées avec ses forme(96classes	60
Tableau VI.2 Résumé des résultats de segmentation.....	68
Tableau IV.3 Résultats de la segmentation pour les différences(20, 40, 70)	70
Tableau IV.4 Résumé de l'évaluation de la reconnaissance	71

Introduction

générale

Introduction générale

Introduction générale

La reconnaissance de l'écriture manuscrite est le vieux rêve de tous ceux qui ont eu besoin d'entrer des données dans un ordinateur. Il remonte à plus d'une trentaine d'années. Aujourd'hui, il existe plusieurs domaines dans lesquels la reconnaissance de l'écriture manuscrite est attendue avec impatience, par exemple dans le tri automatique du courrier, le traitement automatique de dossiers administratifs, des formulaires d'enquêtes, ou encore l'enregistrement des chèques bancaires.

La reconnaissance de l'écriture manuscrite regroupe deux types de systèmes qui se distinguent par la nature du support d'écriture qu'ils ont à traiter :

- Les systèmes de reconnaissance d'écriture les plus utilisés et qui constituent l'objet principal de notre étude, sont qualifiés de systèmes hors-lignes. Après numérisation, le support informatique des données écrites, est une image. L'objectif n'est pas d'établir une interaction directe entre un scripteur et la machine, mais de traiter des documents classiques de manière automatique. Les domaines d'applications les plus typiques sont principalement associés au traitement automatique des adresses postales, des montants des chèques ou encore des bordereaux de livraisons.
- Le deuxième mode d'acquisition concerne les systèmes de reconnaissance d'écriture qualifiés de systèmes en-lignes. La reconnaissance de l'écriture s'effectue pour ces systèmes, le plus souvent, au moment où le scripteur écrit ; l'acquisition est alors réalisée au moyen d'un stylo et d'une tablette électronique, qui peut être assimilée directement à un écran (ex : papier électronique).

Les applications concernées par ces systèmes de reconnaissance sont principalement les interfaces orientées stylo qui cherchent à intégrer l'écriture manuscrite comme une nouvelle modalité d'interaction entre l'homme et la machine.

Les travaux effectués dans le cadre de ce mémoire portent sur la reconnaissance de mots arabes manuscrits hors-lignes. L'objectif est de présenter une nouvelle architecture d'OCR (Optical Character Recognition System) arabe basée sur la méthode de projection verticale pour la segmentation des mots en caractères et les réseaux de neurones de type d'un Perceptron Multi-Couches (PMC) pour la classification, et de l'évaluer sur une nouvelle base de mots arabes que nous avons

Introduction générale

créée et qui est aussi présentée dans ce travail. Cette thèse est subdivisée en quatre chapitres :

Le premier chapitre sera consacré essentiellement à la description de la reconnaissance de l'écriture arabe manuscrite. Premièrement, nous mettrons l'accent sur les caractéristiques de l'écriture arabe. Nous présenterons par la suite l'architecture générale d'un système de reconnaissance d'écriture manuscrite.

Le deuxième chapitre est spécifique à l'état de l'art de la segmentation des mots. Pour celui nous décrivons le processus de la segmentation. Nous mettons l'accent sur la méthode utilisée dans cette étape, en donnant les problèmes rencontrés dans le cas de l'écriture arabe.

Le troisième chapitre va mettre l'accent sur la technique des réseaux de neurones avec toutes ses notions de base. On y trouve une description de l'architecture générale d'un réseau et de son mode de fonctionnement, et en particulier pour un perceptron et un réseau multicouches. Ensuite, on y présente les différentes étapes de mise en œuvre d'un réseau de neurones.

Le quatrième chapitre présente notre contribution à la reconnaissance hors-ligne de l'écriture arabe manuscrite. Il regroupe la description des différentes expériences réalisées et l'analyse des résultats obtenus. Nous présentons dans un premier temps l'architecture de notre système de reconnaissance d'écriture arabe manuscrite à travers ses différentes étapes de prétraitement, d'extraction de caractéristiques et de classification. Les performances du système proposé seront évaluées sur deux bases de mots arabes manuscrits; la base IFN/ENIT et notre base nouvellement créée (base des wilayas d'Algérie).

Chapitre I

Reconnaissance de l'écriture
arabe manuscrite

I.1 Introduction

Ces dernières années, des progrès considérables ont été réalisés dans la mise en œuvre de systèmes pour la reconnaissance de l'écriture manuscrite, et cela grâce, d'une part aux nombreux travaux effectués dans ce domaine, et d'autre part, à la production évaluée à bas prix des micro-ordinateurs et des systèmes d'acquisition (scanner, tablette à digitaliser... etc.).

La reconnaissance de l'écriture manuscrite par ordinateur est un domaine très vaste, Les travaux de recherches sur l'écriture arabe, bien qu'ils soient moins nombreux en comparaison avec d'autres types d'écriture (ex. le Latin et le Japonais).

Nous présenterons dans ce chapitre les caractéristiques de l'OCR (Optical Character Recognition) ,ses différents aspects , ensuite le processus de reconnaissance de l'écriture manuscrite. Enfin nous présenterons quelques problèmes liés à ce domaine.

I.2 Caractéristiques de l'écriture arabe manuscrite

L'écriture arabe est née aux alentours du VI^{ème} siècle , Elle possède plusieurs caractéristiques parmi les on cite :

1. L'alphabet arabe comporte 28 lettres fondamentales. Contrairement à l'alphabet latin, chacune des lettres arabes se décline sous plusieurs formes suivant sa place dans le mot : début (D), milieu (M), fin (F), et isolée (I) comme il est montré dans le tableau 1.1. On distingue 22 lettres de l'alphabet, ayant quatre formes d'écriture.

Les six restantes ne peuvent être rattachées à leurs successeurs et donc elles n'ont que deux formes. Citons comme titre d'exemple la lettre « ا ». [1], le tableau suivant représente les 28 lettres fondamentales de l'alphabet arabe.

Nom de caractère		I	D	M	F	Nom de caractère		I	D	M	F
Alif	ألف	ا	ا	ا	ا	Dhad	ضاد	ض	ض	ض	ض
Ba'	باء	ب	ب	ب	ب	Tta'	طاء	ط	ط	ط	ط
Ta'	تاء	ت	ت	ت	ت	Dha'	ظاء	ظ	ظ	ظ	ظ
Tha'	ثاء	ث	ث	ث	ث	A'in	عين	ع	ع	ع	ع
Jeem	جيم	ج	ج	ج	ج	Ghain	غين	غ	غ	غ	غ
H'a'	حاء	ح	ح	ح	ح	Fa'	فاء	ف	ف	ف	ف
Kha'	خاء	خ	خ	خ	خ	Qaf	قاف	ق	ق	ق	ق
Dal	دال	د	د	د	د	Kaf	كاف	ك	ك	ك	ك
Thal	ذال	ذ	ذ	ذ	ذ	Lam	لام	ل	ل	ل	ل
Rai	راي	ر	ر	ر	ر	Meem	ميم	م	م	م	م
Zai	زاي	ز	ز	ز	ز	Noon	نون	ن	ن	ن	ن
Seen	سين	س	س	س	س	Ha'	هاء	ه	ه	ه	ه
Sheen	شين	ش	ش	ش	ش	Waw	واو	و	و	و	و
Sad	صاد	ص	ص	ص	ص	Ya'	ياء	ي	ي	ي	ي

Tableau I.1 L'alphabet arabe.

2. Contrairement aux écritures occidentales qui s'écrivent de gauche-à-droite, l'arabe s'écrit de droite-à-gauche.[1]
3. Il n'y a pas de différence entre les lettres manuscrites et les lettres imprimées.
Les notions de lettre capitale et lettre minuscule n'existent pas, l'écriture est donc monocamérale. [1]
4. La plupart des lettres s'attachent entre elles, même en imprimé, ce qui offre à l'écriture arabe la caractéristique de cursivité [2], la figure 1.1 illustre un exemple tiré du sacré Coron.

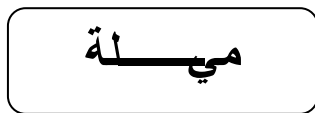
﴿ فَأَمَّا مَنْ أُوتِيَ كِتَابَهُ بِيَمِينِهِ فَيَقُولُ هَؤُلَاءِ أَقْرَبُوا كِتَابِي ﴾

Figure I.1 la cursivité de l'écriture arabe

5. Un caractère arabe peut contenir un trait vertical (ALIF(أ)), un trait oblique (TAA(ط)) ou un zigzag (KAF(ك)) [3].

6. Comme on a indiqué précédemment six lettres ne s'attachent jamais à la lettre suivante : "أ", "ز", "ر", "د", "ذ", et "و", de ce fait, un mot unique peut être entrecoupé d'un ou plusieurs espaces donnant plusieurs pseudo-mots ou composantes connexes ou également sous-mots [4], ce qui est le cas du mot représenté dans la figure I.2.

7. Un pseudo-mot est une unité connexe regroupant une ou plusieurs lettres sous forme d'une séquence. En manuscrit, l'espacement entre les différents pseudo-mots d'un même mot n'est pas forcément systématiquement supérieur à l'espacement entre deux mots différents, ce qui pose parfois des problèmes de segmentation.



1 pseudo-mots



4 pseudo-mots

Figure I.2. Exemple de pseudo-mots constituant des mots arabes.

8. Les caractères arabes sont majoritairement des consonnes, on trouve seulement trois caractères dans la langue arabe qui représentent les voyelles ; "و", "ا" et "ى". Une bonne partie de la vocalisation étant produite par les signes diacritiques [5].

9. Certains mots se différencient seulement par des signes diacritiques d'où leur importance pour la reconnaissance de mots. Ces signes peuvent se situer au-dessus ou au-dessous du caractère, mais jamais en haut et en bas simultanément [6]. Les signes diacritiques dans la langue arabe sont donc plus nombreux et variés que pour les caractères latins. On trouve quatre signes de voyellation qui peuvent se placer en dessous tels que Kasrah "ـِ" et des signes de voyellation qui peuvent se situer au dessous tels que Fatha "ـَ", Dhamma "ـُ", et Soukoun "ـْ". De plus trois tanwin peuvent être formés à partir d'un double Fatha "ـً", d'un double Dhamma "ـٌ" ou d'un double Kasrah "ـٍ".

10. La bande centrale est généralement la plus chargée au point de vue densité d'informations en pixels [7]. Elle correspond à l'emplacement des ligatures verticales, aux caractères centrés et aux boucles, on l'appelle ligne de base.

11. Dans l'alphabet arabe, 15 lettres parmi les 28 possèdent un ou plusieurs points. Ces points diacritiques sont situés soit au-dessus, soit en dessous de la forme à laquelle ils sont associés, mais jamais les deux à la fois. Le nombre maximal de points que peut avoir une lettre est de trois points au-dessus du caractère, ou deux points en dessous. Ces points permettent de différencier la prononciation des lettres arabes. [1] Le tableau I.2 représente les lettres ayant des points diacritiques ainsi que leurs nombres et positions.

Position \ NB de points	Au dessous	Au dessus
Un point diacritique	خ ز ض ظ غ ف ن	ب ج
Deux point diacritique	ت ق	ي
Trois point diacritique	ث ش	

Tableau I.2. Lettres arabes ayant des points diacritiques.

12. Les voyelles sont d'une invention postérieure aux consonnes. Dans l'arabe contemporain ordinaire, on écrit seulement les consonnes et les voyelles longues. Un même mot avec différentes voyelles courtes peut être compris comme verbe, nom ou adjectif. Par d'exemple « عرف » peut signifier «savoir : عَرَفَ» ou « définir : عَرَّفَ » ou encore «tradition : عُرْفُ», selon sa voyellation. [1]

13. Comme dans l'écriture latine, l'écriture arabe contient des ascendants et des descendants. En arabe, les descendants ou les jambages peuvent se prolonger horizontalement sous la bande de base, ce qui introduit une superposition verticale entre la lettre qui comprend le descendant et la lettre suivante. [1] Le tableau I.3 donne les différentes lettres arabes possédant des hampes et des jambages.

Hampes	
Lettres	ا ا ك ل ل لا ط
Jambage	
Lettres	ع, غ ر و م ن ق
	ى, ي س, ش ص, ض ج, ح, خ ع, غ

Tableau I.3. Différents hampes et jambages dans les lettres arabes.

18. Contrairement à l'écriture latine, les articles (le, la, les) font partie du mot auquel ils sont rattachés. La séquence "ال" (un pseudo mot qui contient la lettre " ا " isolée, suivi d'un autre pseudo-mot qui commence par la lettre " ل ") correspond nécessairement au début d'un mot [8]. La figure I.3 indique le commencement des deux mots arabes " الفردوس " et " الكوثر " par cette séquence.

الفردوس الكوثر

Figure I.3. L'article défini " ال " dans deux mots arabes.

I.3 Différents aspects de l'OCR (Optical Character Recognition)

Il n'existe pas de système universel d'OCR qui permet de reconnaître n'importe quel caractère dans n'importe quelle fonte. Cela dépend du type de données traitées et bien évidemment de l'application visée [50]. Il existe plusieurs modes de classification des systèmes OCR parmi lesquels on peut citer:

- Les systèmes qualifiés de « en-ligne » ou « hors-ligne » suivant le mode d'acquisition.
- Les approches globales ou analytiques selon que l'analyse s'opère sur la totalité du mot, ou par segmentation en caractères.
- Les approches statistiques, structurelles ou stochastiques relatives aux traits caractéristiques extraits des formes considérées.

I.3.1 Reconnaissance en-ligne et hors-ligne

Ce sont deux modes différents d'OCR, ayant chacun ses outils propres d'acquisition et ses algorithmes correspondants de reconnaissance.

a) La reconnaissance en-ligne (on-line) :

Ce mode de reconnaissance s'opère en temps réel (pendant l'écriture). Les symboles sont reconnus au fur et à mesure qu'ils sont écrits à la main.

Ce mode est réservé généralement à l'écriture manuscrite. C'est une approche « signal » où la reconnaissance est effectuée sur des données à une dimension.

L'écriture est représentée comme un ensemble de points dont les coordonnées sont fonction du temps [10][11].

La reconnaissance en-ligne présente un avantage majeur c'est la possibilité de correction et de modification de l'écriture de manière interactive vu la réponse en continu du système [12].

L'acquisition de l'écrit est généralement assurée par une tablette graphique munie d'un stylo électronique.

b) La reconnaissance hors-ligne (off-line) :

Démarre après l'acquisition. Elle convient aux documents imprimés et les manuscrits déjà rédigés. Ce mode peut être considéré comme le cas le plus général de la reconnaissance de l'écriture. Il se rapproche du mode de la reconnaissance visuelle. L'interprétation de l'information est indépendante de la source de génération [13].

La reconnaissance hors-ligne peut être classée en plusieurs types :

- **Reconnaissance de texte ou analyse de documents** : Dans le premier cas il s'agit de reconnaître un texte de structure limitée à quelques lignes ou mots. La recherche consiste en un simple repérage des mots dans les lignes, puis à un découpage de chaque mot en caractères [9].

Dans le second cas (analyse de document), il s'agit de données bien structurés dont la lecture nécessite la connaissance de la typographie et de la mise en page du document. Ici la démarche n'est plus un simple prétraitement, mais une démarche experte d'analyse de document il y'a localisation des régions, séparation des régions graphiques et photographique, étiquetage sémantique des zones textuelles à partir de modèles, détermination de l'ordre de lecture et de la structure du document [13].

• **Reconnaissance de l'imprimé ou du manuscrit** : Les approches diffèrent selon qu'il s'agisse de reconnaissance de caractères imprimés ou manuscrits. Les caractères imprimés sont dans le cas général alignés horizontalement et séparés verticalement, ce qui simplifie la phase de lecture [9]. La forme des caractères est définie par un style calligraphique (fonte) qui constitue un modèle pour l'identification. Dans le cas du manuscrit, les caractères sont souvent ligaturés et leur graphisme est inégalement proportionné provenant de la variabilité intra et inter scripteurs.

Cela nécessite généralement l'emploi de techniques de délimitation spécifiques et souvent des connaissances contextuelles pour guider la lecture [14].

Dans le cas de l'imprimé, la reconnaissance peut être monofonte, multifonte ou omnifonte.

Un système est dit monofonte s'il ne peut reconnaître qu'une seule fonte à la fois c'est à dire qu'il ne connaît de graphisme que d'une fonte unique.

C'est le cas le plus simple de reconnaissance de caractères imprimés [15].

Un système est dit multifonte s'il est capable de reconnaître divers types de fontes parmi un ensemble de fontes préalablement apprises [9].

Et un système omnifonte est capable de reconnaître toute fonte, généralement sans apprentissage préalable. Cependant ceci est quasiment impossible car il existe des milliers de fontes dont certaines illisibles par l'homme (sauf bien sûr pour celui qui l'a conçue) et avec un logiciel de création de fonte n'importe qui peut concevoir des fontes à sa guise [15]. Anigbogu [15] a présenté une autre définition pour ce terme c'est l'expression « polyfonte » et a qualifié un système polyfonte de système capable de reconnaître un très grand nombre de fontes.

I.3.2 Reconnaissance globale ou analytique

L'approche globale: considère le mot comme une seule entité et le décrit indépendamment des caractères qui le constituent. Cette approche présente l'avantage de garder le caractère dans son contexte avoisinant, ce qui permet une modélisation plus efficace des variations de l'écriture et des dégradations qu'elle peut subir.

Cependant cette méthode est pénalisante par la taille mémoire, le temps de calcul et la complexité du traitement qui croissent linéairement avec la taille du lexique considéré, d'où une limitation du vocabulaire [11], [16].

L'approche analytique: contrairement à l'approche globale, le mot est segmenté en caractères ou en fragments morphologiques significatifs inférieurs au caractère appelés graphèmes. La reconnaissance du mot consiste à reconnaître les entités segmentés puis tendre vers une reconnaissance du mot, ce qui constitue une tâche délicate pouvant générer différents types d'erreurs [16], [10]. Un processus de reconnaissance selon cette approche est basé sur une alternance entre deux phases : la phase de segmentation et la phase d'identification des segments.

Deux solutions sont alors possibles : la segmentation explicite (externe) ou la segmentation implicite (interne) [17]. Par ailleurs, les méthodes analytiques par opposition aux méthodes globales, présentent l'avantage de pouvoir se généraliser à la reconnaissance d'un vocabulaire sans limite à priori, car le nombre de caractères est naturellement fini. De plus l'extraction des primitives est plus aisée sur un caractère que sur une chaîne de caractères [18].

I.4 Processus de reconnaissance de mots arabes manuscrits

Un système de reconnaissance fait appel généralement aux étapes suivantes : acquisition, prétraitement, segmentation, extraction des caractéristiques, classification, suivies éventuellement d'une phase de post-traitement .la figure suivante représente ces étapes

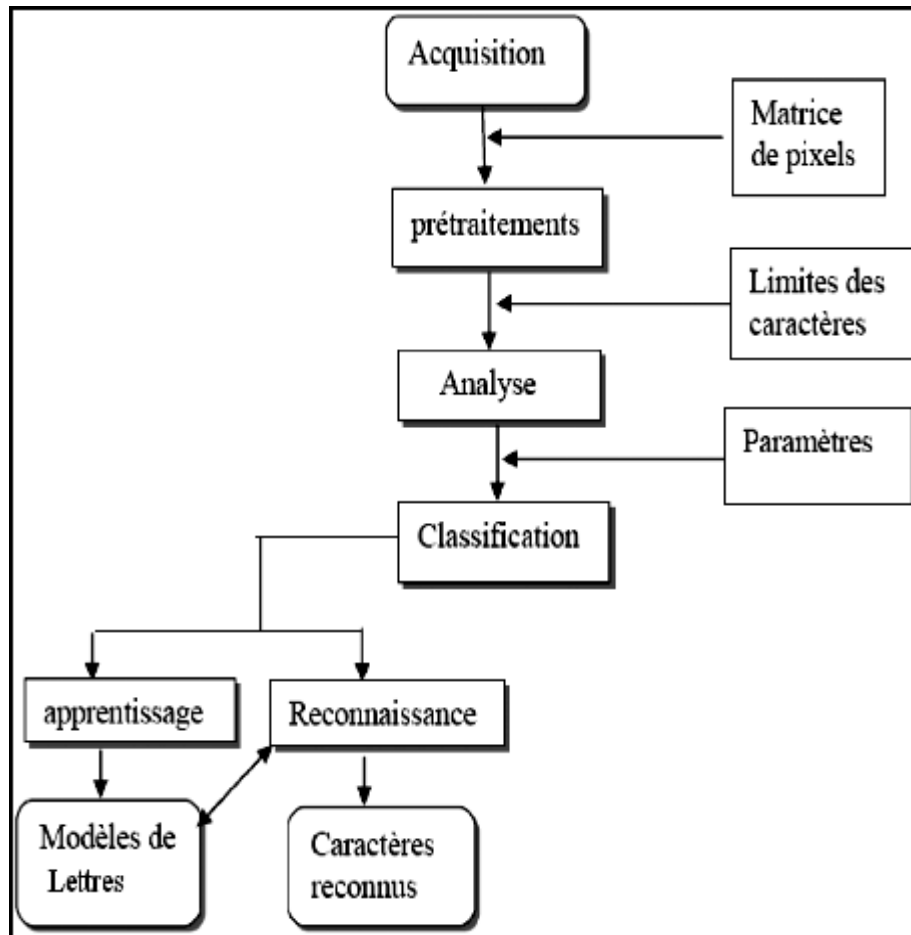


Figure I. 4 Schéma général d'un système de reconnaissance de caractères.

I.4.1 Phase d'acquisition

Elle consiste à capter l'image d'un texte au moyen des capteurs physiques (scanner, caméra,...) et de la convertir en grandeurs numériques adaptées au système de traitement, avec un minimum de dégradation possible. Elle consiste deux phases :

- Echantillonnage (numérisation) d'une image : est spatial, par découpage en pixels ;
- Quantification (codage) : est une valeur numérique donnée à l'intensité lumineuse, c'est un niveau de gris, appelé la dynamique de l'image.

I.4.2 Phase de prétraitement

Le prétraitement consiste à préparer les données issues du capteur à la phase suivante. Il s'agit essentiellement de réduire le bruit superposé aux données et essayer de ne garder que l'information significative de la forme représentée. Le bruit peut être dû aux conditions d'acquisition (éclairage, mise incorrecte du document, ...) ou encore à la qualité du document d'origine. Parmi les opérations de prétraitement généralement

utilisées on peut citer : l'extraction des composantes connexes, le redressement de l'écriture, le lissage, la normalisation et la squelettisation (figure I.5)

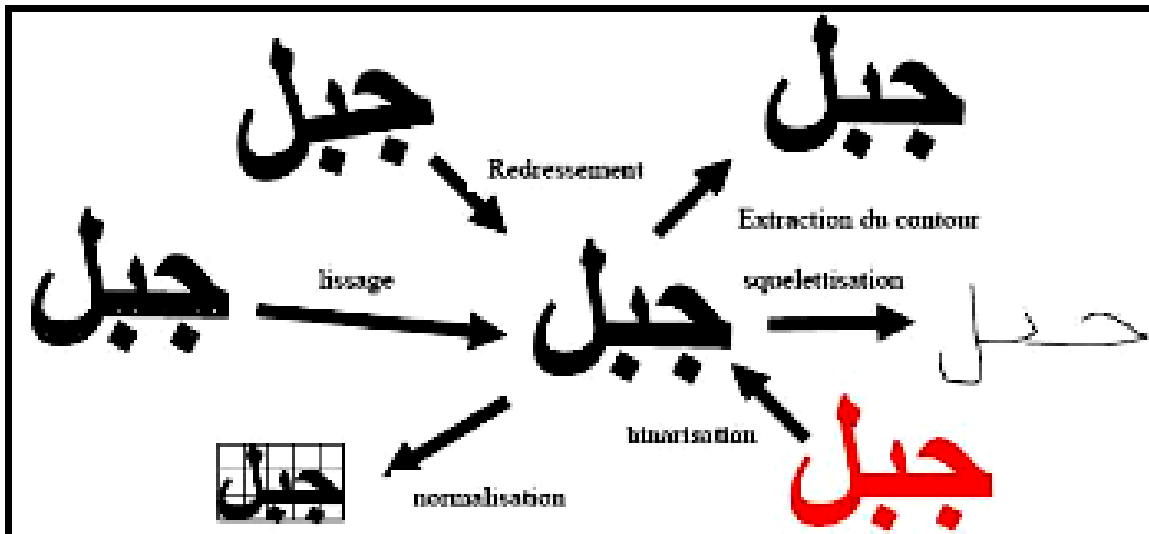


Figure I.5 Effets de certaines opérations de prétraitement.

a. La Binarisation

La binarisation c'est le passage d'une image en couleur ou définie par plusieurs niveaux de gris en image bitonale (composée de deux valeurs 0 et 1) qui permet une classification entre le fond (image du support papier en blanc) et la forme (traits des gravures et des caractères en noir).

Pour des images de niveaux de gris, on peut trouver dans [19] une liste des méthodes de binarisation, proposant des seuils adaptatifs (ex. s'adaptant à la différence de distribution des niveaux de gris). [16] proposent une solution pour les images d'adresses postales. La recherche du seuil passe par plusieurs étapes : binarisation préliminaire basée sur une distribution de mixture multimodale, analyse de la texture à l'aide d'histogrammes de longueurs de traits, et sélection du seuil à partir d'un arbre de décision.

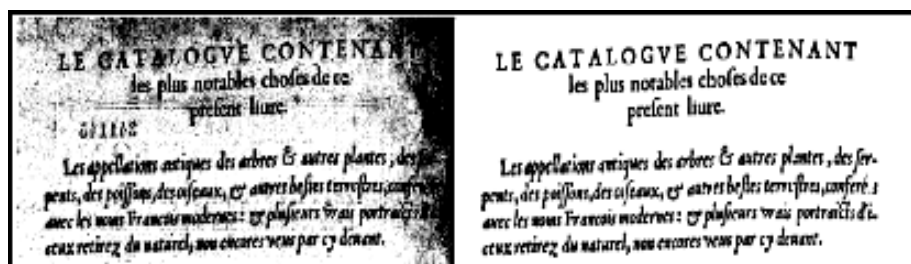


Figure I.6 Exemple de Binarisation adaptative [20].

b. Extraction de composantes connexes

Une composante connexe (CXX) est un ensemble de points dans le plan. Elle peut correspondre à un point diacritique, un accent, au corps d'un caractère ou d'une chaîne de caractères... Une fois localisés les CXX sont regroupées pour former les mots. Cette technique est utilisée pour le repérage des points diacritiques dans les images de textes arabes [9].

c. Redressement de l'écriture

L'un des problèmes rencontrés en OCR est l'inclinaison des lignes du texte, qui introduit des difficultés pour la segmentation. L'inclinaison peut provenir de la saisie, si le document a été placé en biais, ou être intrinsèque au texte. Il convient alors de le redresser afin de retrouver la structure de lignes horizontales d'une image texte. Si α est l'angle d'inclinaison, pour redresser l'image, une rotation isométrique d'angle α est opérée grâce à la transformation linéaire suivante [16] :

$$\begin{cases} x' = x \cos \alpha - y \sin \alpha \\ y' = y \cos \alpha + x \sin \alpha \end{cases} \quad \text{I.1}$$

d. Lissage

L'image des caractères peut être entachée de bruits dus aux artefacts de l'acquisition et à la qualité du document, conduisant soit à une absence de points ou à une surcharge de points. Les techniques de lissage permettent de résoudre ces problèmes par des opérations locales qu'on appelle opérations de bouchage et de nettoyage [21].

L'opération de nettoyage permet de supprimer les petites tâches et les excroissances de la forme. Pour le bouchage il s'agit d'égaliser les contours et de boucher les trous internes à la forme du caractère en lui ajoutant des points noirs.

Plusieurs autres techniques similaires sont utilisées dont la méthode statistique, une méthode basée sur la morphologie mathématique [9].

e. Normalisation

Après la normalisation de la taille, les images de tous les caractères se retrouvent définies dans une matrice de même taille. Pour faciliter les traitements ultérieurs (Figure I.7).

Le principe de la normalisation est d'essayer de normaliser localement différentes parties du mot, de manière à augmenter la ressemblance d'une image à une autre.

Cette opération introduit généralement de légères déformations sur les images. Cependant certains traits caractéristiques tels que la hampe dans les caractères (Ø Û á Ç par exemple) peuvent être éliminée à la suite de la normalisation, ce qui peut entraîner à des confusions entre certains caractères [22].

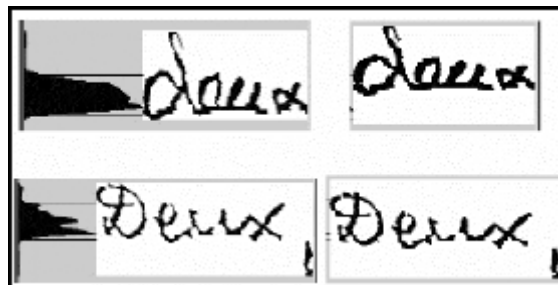


Figure I.7 : Exemple de normalisation de mots manuscrits [16].

f. Squelettisation

Le but de cette technique est de simplifier l'image du caractère en une image à « ligne » plus facile à traiter en la réduisant au tracé du caractère. Les algorithmes de squelettisation se basent sur des méthodes itératives. Le processus s'effectue par passes successives pour déterminer si un tel ou tel pixel est essentiel pour le garder ou non dans le tracé [22].

La squelettisation des caractères arabes peut induire en erreur : deux points diacritiques sont souvent confondus avec un seul [9].

I.4.3 La phase de Segmentation

La mise en œuvre d'une étape de segmentation permet de diviser l'image en différentes SOUS Images de taille moins importantes qui peuvent être des graphèmes,

des lettres ou bien des sous mots. Cependant une SOUS Image reste une matrice de pixels. Il existe deux techniques permettant la mise en œuvre de la segmentation :

a. segmentation explicite

Elle est antérieure à la reconnaissance et n'est pas remise en cause pendant la phase de reconnaissance. Les hypothèses des caractères sont déterminées à partir des informations de bas niveau présentes sur l'image. Ces hypothèses sont définitives, et doivent être d'une grande fiabilité car la moindre erreur de segmentation remet en cause la totalité des traitements ultérieurs. Les approches de segmentation explicite, s'appuient sur une analyse morphologique du mot manuscrit pour localiser des points de segmentation potentiels. Elles sont particulièrement adaptées à l'analyse de la représentation bidimensionnelle et donc plus souvent utilisées dans les systèmes de reconnaissance hors-ligne de mots. Certaines méthodes de segmentation explicite sont basées sur une analyse par morphologiques mathématiques, exploitent les concepts de régularité et singularité du tracé, analyse des contours supérieurs/inferieurs du mot. Les points de segmentation potentiels détectés sont confirmés à l'aide de diverses heuristiques [23].

b. Segmentation implicite

Les méthodes de segmentation implicite s'inspirent des approches utilisées dans le domaine de la parole, où le signal est divisé en intervalles de temps réguliers, et procèdent à une sur-segmentation importante de l'image du mot à pas fixe (un ou quelques pixels). Cela permet d'assurer un taux de présence important des points de liaison entre lettres considérées. La segmentation s'effectue pendant la reconnaissance qui assure son guide. Le système recherche dans l'image, des composantes ou des groupements de graphèmes¹ qui correspondent à ses classes de lettres [23]. Classiquement, il peut le faire de deux manières : par fenêtrage : le principe est d'utiliser une fenêtre mobile de largeur variable (qui n'est pas facile à déterminer) pour trouver des séquences de points de segmentation potentiels qui seront confirmés ou non par la reconnaissance de caractères. Elle nécessite deux étapes : la génération d'hypothèses de segmentation (séquences de points obtenus par le fenêtrage) ; la deuxième est le choix de la meilleure hypothèse de la reconnaissance (validation). par

recherche de primitives : il s'agit de détecter les combinaisons de primitives qui donneront la meilleure reconnaissance.

I.4.4 Phase d'extraction des caractéristiques

C'est l'une des étapes les plus délicates et les plus importantes en OCR. La reconnaissance d'un caractère passe d'abord par l'analyse de sa forme et l'extraction de ses traits caractéristiques (primitives) qui seront exploités pour son identification.

Les types de caractéristiques peuvent être classés en quatre groupes principaux: caractéristiques structurelles, caractéristiques statistiques, transformations globales, et superposition des modèles et corrélation [24] [11].

a. Caractéristiques structurelles

Les caractéristiques structurelles décrivent une forme en terme de sa topologie et sa géométrie en donnant ses propriétés globales et locales. Parmi ces caractéristiques on peut citer [24]:

- Les traits et les anses dans les différentes directions ainsi que leurs tailles. v Les points terminaux.
- Les points d'intersections.
- Les boucles.
- Le nombre de points diacritiques et leur position par rapport à la ligne de base. I Les voyellations et les zigzags (hamza).
- La hauteur et la largeur du caractère.
- La catégorie de la forme (partie primaire ou point diacritique, etc).
- Plusieurs autres caractéristiques peuvent être tirés, suivant qu'ils soient extraits d'une courbe, un trait ou un segment de contour.

b. Les caractéristiques statistiques:

Les caractéristiques statistiques décrivent une forme en termes d'un ensemble de mesures extraites à partir de cette forme. Les caractéristiques utilisés pour la reconnaissance de textes arabes sont : le zonage (zoning), les caractéristiques de lieu géométrique (Loci) et les moment [24].

- Le zonage consiste à superposer une grille $n \times m$ sur l'image du caractère et pour chacune des régions résultantes, calculer la moyenne ou le pourcentage de points en niveaux de gris, donnant ainsi un vecteur de taille $n \times m$ de caractéristiques.
- La méthode Loci est basée sur le calcul du nombre de segments blancs et de segments noirs le long d'une ligne verticale traversant la forme, ainsi que leurs longueurs [11].
- La méthode des moments : les moments d'une forme par rapport à son centre de gravité sont invariants par rapport à la translation et peuvent être invariants par rapport à la rotation [18]. Ils sont aussi indépendants de l'échelle. Ces caractéristiques peuvent être facilement et rapidement extraites d'une image de texte, ils peuvent tolérer modérément les bruits et les variations [25].

c. Les transformations globales

La transformation consiste à convertir la représentation en pixels en une représentation plus abstraite pour réduire la dimension des caractères, tout en conservant le maximum d'informations sur la forme à reconnaître.

Une des transformations les plus simples est celle qui représente le squelette ou le contour d'un caractère sous forme d'une chaîne de codes de directions [11]. La chaîne de code obtenue est souvent simplifiée pour réduire les redondances et les changements brusques de direction.

d. Superposition des modèles (template matching) et corrélation:

La méthode de 'template matching' appliquée à une image binaire (en niveaux de gris ou squelettes), consiste à utiliser l'image de la forme comme vecteur de caractéristiques pour être comparé à un modèle (template) pixel par pixel dans la phase de reconnaissance, et une mesure de similarité est calculée [24].

La classification dans un système OCR regroupe deux tâches : l'apprentissage et la reconnaissance et décision. A cette étape les caractéristiques de l'étape précédente sont utilisées pour identifier un segment de texte et l'attribuer à un modèle de référence [24].

a. L'apprentissage

Il s'agit lors de cette étape d'apprendre au système les propriétés pertinentes du vocabulaire utilisé et de l'organiser en modèles de références.

L'idéal serait d'apprendre au système autant d'échantillons que de formes d'écritures différentes, mais cela est impossible à cause de la grande variabilité de l'écriture qui conduirait à une explosion combinatoire de modèles de représentation. La tendance consiste alors à remplacer le nombre par une meilleure qualité des traits caractéristiques [9], [11]. L'apprentissage consiste en deux concepts différents : l'entraînement et l'adaptation. L'entraînement consiste à enseigner au système la description des caractères tandis que l'adaptation sert à améliorer les performances du système en profitant des expériences précédentes. Certains systèmes permettent à l'utilisateur d'identifier un caractère lorsqu'ils échouent à le reconnaître et ils utilisent l'entrée de l'utilisateur à chaque fois que le caractère est rencontré [11].

Les procédés d'apprentissage sont différents selon qu'il s'agisse de reconnaissance de caractères imprimés ou manuscrits ou de reconnaître des textes monofonte ou multifonte. D'une manière générale, on distingue deux types de techniques d'apprentissage : supervisé et non supervisé.

- L'apprentissage est dit supervisé s'il est guidé par un superviseur appelé professeur. Il est réalisé lors d'une étape préliminaire de reconnaissance en introduisant un grand nombre d'échantillons de référence. Le professeur indique dans ce cas le nom de chaque échantillon. Le choix des caractères de référence est fait à la main en fonction de l'application. Le nombre d'échantillons peut varier de quelques unités à quelques dizaines, voir même quelques centaines par caractère [9], [24].
- L'apprentissage non supervisé ou sans professeur consiste à doter le système d'un mécanisme automatique qui s'appuie sur des règles précises de regroupement pour trouver les classes de référence avec une assistance minimale. Dans ce cas les échantillons sont introduits en un grand nombre par l'utilisateur sans indiquer leur classe [9].

b. Reconnaissance et décision

La décision est l'ultime étape de reconnaissance. A partir de la description en paramètres du caractère traité, le module de reconnaissance cherche parmi les modèles de référence en présence, ceux qui lui sont les plus proches.

La reconnaissance peut conduire à un succès si la réponse est unique (un seul modèle répond à la description de la forme du caractère). Elle peut conduire à une confusion si la réponse est multiple (plusieurs modèles correspondent à la description). Enfin elle peut conduire à un rejet de la forme si aucun modèle ne correspond à sa description. Dans les deux premiers cas, la décision peut être accompagnée d'une mesure de vraisemblance, appelée aussi score ou taux de reconnaissance [9].

Les approches de reconnaissance peuvent être regroupées en trois groupes principaux: l'approche statistique, l'approche structurelle, et l'approche stochastique.

▪ Approche statistique

Elle est fondée sur l'étude statistique des mesures que l'on effectue sur les formes à reconnaître. L'étude de leur répartition dans un espace métrique et la caractérisation statistique des classes, permettent de prendre une décision de reconnaissance du type « plus forte probabilité d'appartenance à une classe » [9].

Les approches statistiques bénéficient des méthodes d'apprentissage automatique qui s'appuient sur des bases théoriques fondées, telles que la théorie de la décision bayésienne, les méthodes de classification non supervisées ... En reconnaissance, le problème revient à affecter une forme inconnue à l'une des classes obtenues pendant l'apprentissage [11].

Nous pouvons citer trois méthodes statistiques parmi celles les plus couramment utilisées :

- L'apprentissage est dit supervisé s'il est guidé par un superviseur appelé professeur. Il est réalisé lors d'une étape préliminaire de reconnaissance en introduisant un grand nombre d'échantillons de référence. Le professeur indique dans ce cas le nom de chaque échantillon. Le choix des caractères de référence est fait à la main en fonction de l'application. Le nombre

d'échantillons peut varier de quelques unités à quelques dizaines, voir même quelques centaines par caractère [9], [24].

- L'apprentissage non supervisé ou sans professeur consiste à doter le système d'un mécanisme automatique qui s'appuie sur des règles précises de regroupement pour trouver les classes de référence avec une assistance minimale. Dans ce cas les échantillons sont introduits en un grand nombre par l'utilisateur sans indiquer leur classe [9].

L'approche bavésienne: consiste à choisir parmi un ensemble de caractères, celui pour lequel la suite de primitives extraites a la plus forte probabilité à posteriori par rapport aux caractères préalablement appris [15].

La méthode du plus proche voisin

L'algorithme KNN (K Nearest Neighbors) affecte une forme inconnue à la classe de son plus proche voisin en la comparant aux formes stockées dans une classe de références nommée prototypes. Il renvoie les K formes les plus proches de la forme à reconnaître suivant un critère de similarité. Une stratégie de décision permet d'affecter des valeurs de confiance à chacune des classes en compétition et d'attribuer la classe la plus vraisemblable (au sens de la métrique choisie) à la forme inconnue [9, 21].

Cette méthode présente l'avantage d'être facile à mettre en œuvre et fournit de bons résultats. Son principal inconvénient est lié à la faible vitesse de classification due au nombre important de distances à calculer.

- **Les réseaux de neurones**

En OCR, les primitives extraites sur une image d'un caractère (ou de l'entité choisie) constituent les entrées du réseau. La sortie activée du réseau correspond au caractère reconnu. Le choix de l'architecture du réseau est un compromis entre la complexité des calculs et le taux de reconnaissance [26].

Par ailleurs, le point fort des réseaux de neurones réside dans leur capacité de générer une région de décision de forme quelconque, requise par un algorithme de classification, au prix de l'intégration de couches de cellules supplémentaires dans le réseau [16].

1.4.5 Phase de post traitement

L'objectif du post-traitement est l'amélioration du taux de reconnaissance des mots (par opposition au taux de reconnaissance du caractère). Cette phase est souvent implémentée comme un ensemble d'outils relatifs à la fréquence d'apparition des caractères dans une chaîne, aux lexiques et à d'autres informations contextuelles.

Comme la classification peut aboutir à plusieurs candidats possibles, le post-traitement a pour objet d'opérer une sélection de la solution en utilisant des niveaux d'informations plus élevés (syntaxiques, lexicales, sémantiques...) [5]. Le post-traitement se charge également de vérifier si la réponse est correcte (même si elle est unique) en se basant sur d'autres informations non disponibles au classifieur.

I.5 Problèmes Liés à L'OCR

La tâche de l'OCR n'est pas aisée, divers problèmes compliquent le processus de reconnaissance, parmi lesquels on peut citer [11], [9]:

- La qualité du document : un document télécopié ou photocopie plusieurs fois est plus difficile à traiter que la copie originale. L'écriture peut devenir plus mince ou au contraire plus épaisse, dégradée avec des parties du texte qui manquent ou de tâches qui apparaissent, des ouvertures ou des bouchages de boucles ...
- L'impression : un document composé est de meilleure qualité qu'un document dactylographié qui, à son tour, est plus clair qu'un texte issu d'une imprimante matricielle. Une imprimante à jet d'encre peut introduire des tâches d'encre et un étalement des caractères, une imprimante laser peut générer des lignes ou des fonds ...
- La discrimination de la forme : selon le style de la fonte utilisée, son corps et sa graisse..., le caractère change de graphisme. Le nombre de formes est d'autant plus important que le nombre de styles d'écriture est élevé. De plus, plusieurs caractères présentent une forte ressemblance tels que :
 - pour l'arabe : ه et ص, د et م, و et م
 - pour le Latin : U et V, O et 0, S et 5, Z et 2.
- Le support de l'information, tel que le papier, joue également sur les performances de la reconnaissance par sa qualité : son grammage, sa granulation et sa couleur.

- L'acquisition : la numérisation en temps réel introduit souvent des distorsions dans l'image. Dans le cas hors-ligne la qualité du texte numérisé est un compromis entre les variations de la position (inclinaison, translation, rétrécissement...), la propreté de la vitre du dispositif de numérisation et sa résolution.

1.4. Conclusion

La reconnaissance de l'écriture arabe est rattachée au vaste domaine de la reconnaissance de formes. Son but est de prendre une décision quant au contenu sémantique du message transmis à partir de sa représentation physique. Les applications de systèmes capables de remplir cette tâche sont nombreuses ; nous pouvons citer entre autres la lecture automatique de bons de commande, le traitement automatique des chèques, la vérification de signatures ou encore le tri automatique du courrier.

Chapitre II

Etat de l'art de la
segmentation

II.1. Introduction

La segmentation est définie, comme étant l'opération qui cherche à décomposer une image de texte en pseudo-images de symboles individuels. Le résultat de cette opération, est une forme isolée à partir d'une image et qui pourrait être un caractère ou un symbole. C'est une étape critique dans plusieurs systèmes de reconnaissance d'écriture, elle est dans la plupart des cas sujette aux erreurs et très pénalisante en temps de calcul.

on peut distinguer plusieurs niveaux de segmentation. A titre d'exemple, dans le traitement automatique de document, consiste à isoler les différentes parties logiques (graphiques, textes, photos, etc.) constituant une page de document. aussi, dans les applications où les informations traitées sont des pages de texte, la séparation des lignes de paragraphes, l'extraction des mots ou pseudo-mots d'une ligne, et enfin, la décomposition des mots ou pseudo-mots en caractères.

Dans ce chapitre nous allons exposer les différentes techniques de segmentation, et en particulier la segmentation du mot en caractères.

II.2 La segmentation de textes arabes(particularités et difficultés)

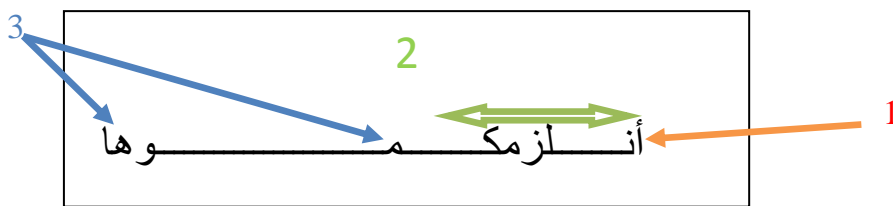
La segmentation automatique de textes arabes présente plusieurs difficultés spécifiques à la langue arabe. Nous présentons dans ce qui suit certaines ambiguïtés qui rendent la segmentation difficile à réaliser sans une étude approfondie sur un corpus à large couverture.

2.1 L'ambiguïté vocalique des mots

Un texte arabe non voyellé est fortement ambigu. La proportion des mots ambigus passe à plus de 90% si les comptages portent sur les voyellations globales de ces mots. Ainsi, un mot non voyellé peut avoir plusieurs caractéristiques morphologiques possibles. Par exemple le mot "فهم" peut être un nom, un verbe, ou un pronom personnel précédé d'une conjonction de coordination.

2.2 L'ambiguïté dérivationnelle

Le mot arabe n'est pas le résultat d'une simple concaténation de morphèmes comme c'est le cas pour l'anglais mais c'est à partir d'une racine, d'une combinaison de voyelles, de préfixes, d'infices, de suffixes et d'un schème morphologique qu'on obtient un mot. Ainsi, l'identification de la catégorie grammaticale de certains mots est ambiguë ce qui entraîne des difficultés au niveau de la segmentation automatique. par exemple:



1:outil d'interrogation ; 2:le verbe ; 3: compléments .

Figure II.1 Exemple d'une combinaison de mot arabe.

2.3 L'ambiguïté structurelle

La phrase arabe est relativement longue et complexe en comparaison avec d'autres langues, tels que le français ou l'anglais. Ainsi il n'est pas rare de trouver des phrases arabes composées de plusieurs dizaines de mots.

1. Exemple en arabe:

الابتزاز له أنواعه فمنه اللطيف ومنه العنيف, فالابتزاز هو أن يحاول الشخص الآخر عمل ما يريد هو سواء كنت تحب ما يريد أو لا تحب, وللابتزاز صور عدة, فمنه اللطيف الذي فيه التوسل أو بحكم القرب والصدقة, ومنه المتوسط الذي فيه اللوم كأن يقول الوالدان " لقد تعبتنا في تربيتك ولا تحقق لنا ما نريد ", ومنه الابتزاز العنيف الذي فيه تهديد, "ستخرج من المنزل إذا لم تفعل ما أمرك", وكمثال ابتزاز امرأة العزيز للنبي يوسف (ولئن لم يفعل ما أمره ليسجنن وليكونا من الصاغرين) فكان ابتزازا, وقد اختار ما يريد هو لا ما تريده هي (رب السجن أحب إلي مما يدعونني إليه) .

2. Exemple en anglais:

A protocol is the set of rules that governs data communications.

2.4 L'utilisation des signes de ponctuation

L'arabe n'est pas appuyé principalement sur les signes de ponctuations et les marqueurs typographiques; ces derniers ont généralement un rôle pausale. Ainsi, nous

pouvons trouver tout un paragraphe arabe ne contenant aucun signe de ponctuation à part un point à la fin de ce paragraphe. Comme un exemple:

في الماضي عندما اخترت طريقا واتخذت قرارا فقد كان أفضل قرار بالنسبة لك في ذلك الوقت وكذلك الآخرين فهم يتخذون أفضل الخيارات التي لديهم في الوقت الحالي.

2.5 L'agglutination

Les conjonctions de coordinations jouent un rôle important dans la segmentation de textes arabes. Cependant, elles sont toujours agglutinées aux mots qui les suivent. Ainsi par exemple la lettre "و" (w) dans le mot "وهم" peut représenter une lettre du mot en question (i.e., « wahmun » (imagination)) ou une conjonction de coordination suivie d'un pronom personnel (i.e., "و"+"هم" « wa+hum » (et + ils)).

II.3 Les niveaux de segmentation

On peut distinguer quatre niveaux de segmentation, comme suit:

- a) segmentation de la page ;
- b) segmentation de texte en lignes;
- c) segmentation de lignes en mots;
- d) segmentation de mots en caractères.

Nous pouvons présenter ces niveaux dans la figure suivante

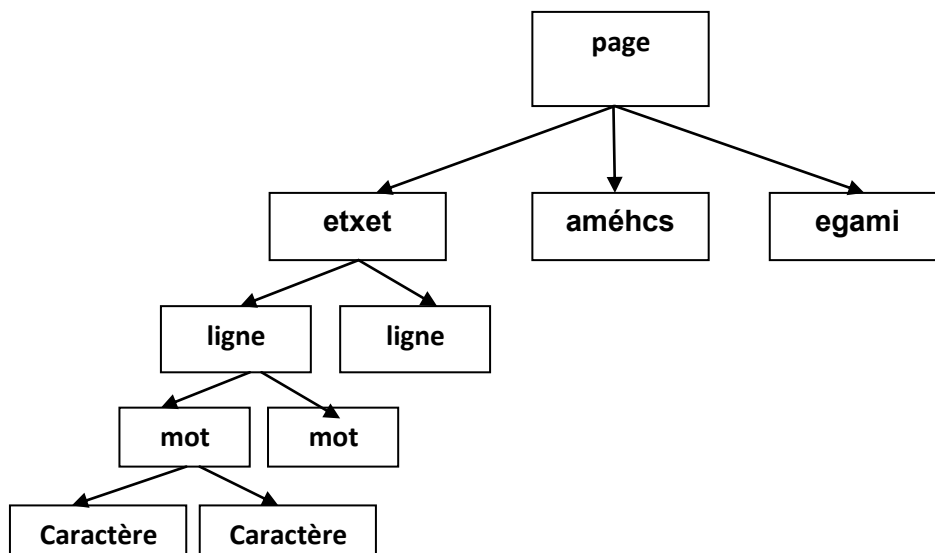


Figure II.2 le processus de segmentation.

3.1 Segmentation de la page

Cette étape permet de localiser dans chaque page, les zones d'information conformément à leur apparence physique. Elle est associée généralement à l'étiquetage logique qui consiste à déterminer la nature du media représenté dans chaque zone (texte, graphique, photographie etc.). Cette classification permet ensuite d'orienter la reconnaissance vers des systèmes spécialisés dans l'analyse de chaque type de media [28]. La figure suivante illustre cette étape.

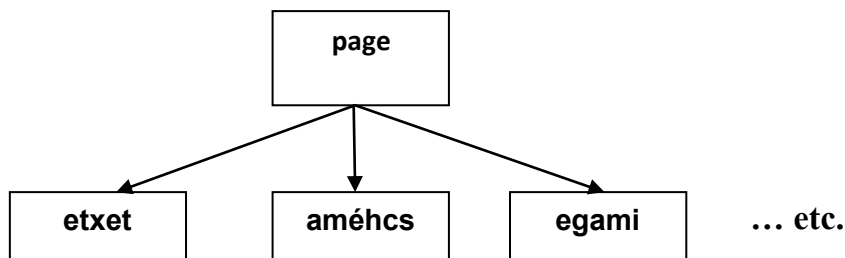


Figure II.3 Détection des différentes zones d'une page de document.

3.2 Segmentation d'un bloc de texte en lignes

Cette étape consiste à séparer les différentes lignes du texte pour en extraire les mots puis les caractères composant les mots. La plupart des études proposées dans ce domaine s'appuient sur une décomposition de l'image en composantes connexes.

D'autres par contre utilisent des techniques s'appuyant en grande partie sur les histogrammes des projections horizontales, et certains auteurs optent pour des méthodes spécialisées pour la segmentation en lignes de l'écriture arabe manuscrite. La figure suivante montre un exemple d'extraction des lignes d'un texte [28].

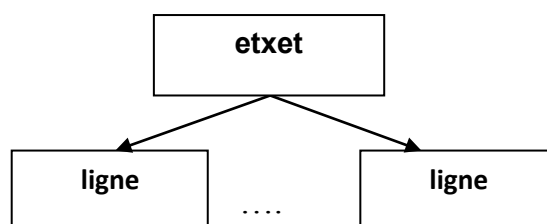


Figure II.4 Segmentation de texte en lignes.

3.3 Segmentation de lignes en mots

La segmentation en mots est réalisée en déterminant l'histogramme des projections verticales des lignes pour détecter les espaces entre les mots et pouvoir les séparer.

Cependant cette technique peut ne pas être efficace dans certains cas où les mots se chevauchent (cas par exemple de l'écriture arabe).

Dans ce cas d'autres techniques sont utilisées telles que: le suivi du contour, détermination du squelette ou la détermination des composantes connexes [28]. La figure II.4 montre un exemple de segmentation de ligne en mots.

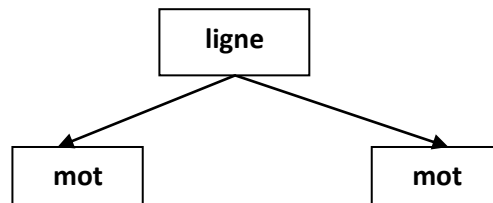


Figure II.5 Segmentation de Ligne en Mots.

3.4 Segmentation de mots en caractères

La segmentation des caractères est une opération qui tente de décomposer une image de séquence de caractères (mot) en sous-images de symboles individuels. C'est l'un des processus de décision dans un système de reconnaissance optique de caractères. Son but est de décider si un motif isolé d'une image (caractère ou autre entité identifiable du mot) est correct ou non [28]. La figure II.5 illustre cette étape.

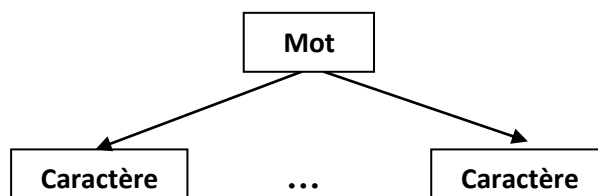


Figure II.6 Segmentation de mot en Caractères.

II.4 Les méthodes de segmentation d'écriture cursive

Les graphèmes sont des images extraites de l'image à segmenter. Passer d'une seule image à une séquence de graphèmes pose le problème de la taille de ces éléments. Ils ne doivent pas être trop petits afin d'être statistiquement significatifs, et pas trop gros afin de ne pas dépasser la taille d'une lettre. Il est en effet important qu'un graphème donné soit une sous partie d'une seule lettre : cette condition est nécessaire pour construire un modèle de mot comme étant la concaténation de modèles de lettres [29].

4.1 Segmentation à partir du squelette

A partir du squelette, on cherche à repérer certains motifs, pour en déduire les candidats de points de coupures. La détection de ces motifs introduit des calculs de courbures et d'angles, qui sont comparées à des seuils ajustés de manière à obtenir le résultat désiré. X.Dupré souligne que cette approche est erronée dans environ 10% des cas. Les configurations difficiles à segmenter sont celles pour lesquelles les lettres sont souvent enchevêtrées, comme les "tt", ou les lettres `a liaison haute ('b', 'o', 'v', 'w') avec leur successeur [29].

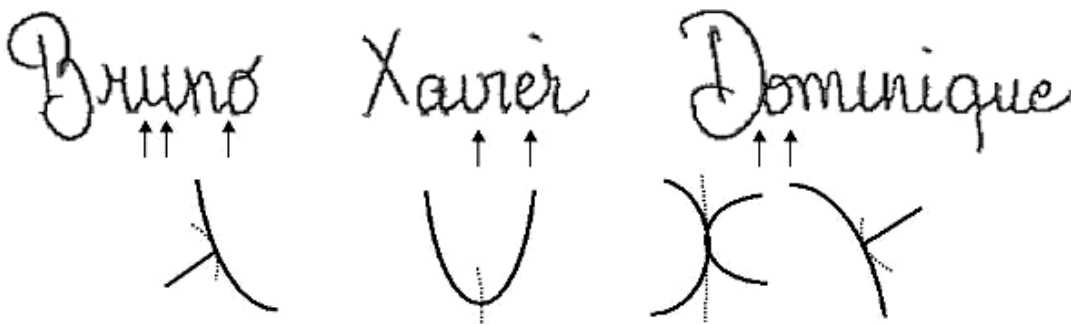


Figure II.7 Segmentation à base de squelette: basée sur des motifs.

4.2 Segmentation à partir du contour

La segmentation à partir du contour consiste à déterminer les meilleurs points candidats de coupure entre graphèmes, en s'appuyant sur les extrema locaux du contour, qui sont associées selon un critère de proximité.

Comme la segmentation en graphèmes à partir du contour nécessite de nombreux ajustements avant de trouver les critères de décision. Cette mise au point par tâtonnements est le point commun de nombreux traitements d'images liés à la reconnaissance de l'écriture manuscrite. Faciles à ajuster lorsque la qualité de l'écriture est bonne, ces prétraitements peuvent avoir des comportements tout à fait erratiques lorsque l'écriture est de mauvaise qualité [29].

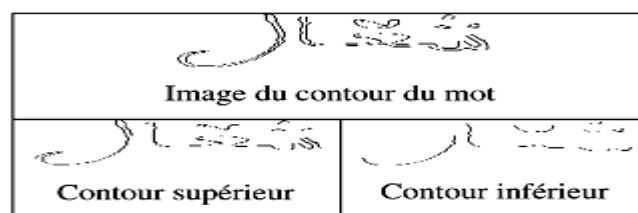


Figure II.8 Extrema du contour supérieur et inférieur [32]

4.3 Segmentation à partir des histogrammes

La segmentation en utilisant des histogrammes est une méthode proposée par B. Yanikoglu et P. Sandon. Elle consiste à calculer des histogrammes de projection dans plusieurs directions proches de la verticale. Les droites choisies sont celles qui interceptent le moins de pixels noirs, avec une contrainte d'espacement régulier dans l'image. Néanmoins, Cette méthode montre des limites lorsque les lettres sont très proches ou enchevêtrées [29].



Figure II.9 Segmentation à partir d'histogrammes de projection selon plusieurs directions.

4.4 Segmentation basée sur des réservoirs

X. Dupré étend à l'écriture cursive la technique à base de réservoirs initialement appliquée à la segmentation de chiffres liés. Il souligne que les règles de décision sont plus difficiles à mettre en place dans le cas des lettres, car ces dernières sont de tailles variables [29].

4.5 Segmentation basée sur les fenêtres glissantes

L'utilisation d'une fenêtre glissante revient à découper l'image en bandes verticales. Ce découpage peut être régulier ou non, éventuellement avec recouvrement partiel des bandes successives.

Cette technique présente l'avantage d'être simple, robuste au bruit, et est indépendante de la connexité. Le défaut de cette méthode est que la séquence générée contient beaucoup de bruit (recouvrement de deux lettres successives). C'est également vrai dans le cas des lettres superposées verticalement, mais qui ne se touchent pas nécessairement : une barre de 't' avec la lettre suivante, ou les descendants comme '’' ou ' ’' en arabe [29].

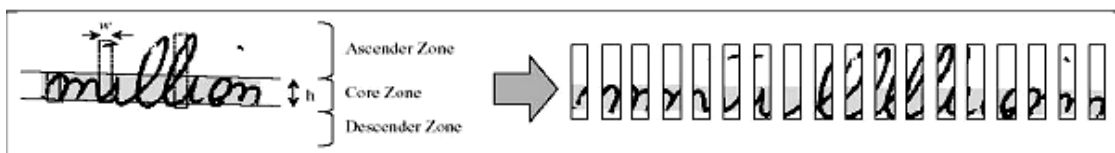


Figure II.10 Segmentation à base de fenêtre glissante : découpage du mot en bandes verticales.

4.6 Segmentation en bandes verticales

A. Benouareth et al utilisent deux stratégies de segmentation en bandes verticale : segmentation uniforme et non uniforme (figure II.9). La stratégie de segmentation uniforme consiste à utiliser la même largeur pour toutes les bandes (fixée empiriquement à 20 colonnes). Alors que dans une segmentation non uniforme, les bandes verticales n'ont pas forcément la même largeur, les limites de chaque bande sont déterminées par une analyse des minima et maxima de l'histogramme de projection verticale du mot [31].

Les auteurs signalent qu'une segmentation non uniforme en bandes verticales semble plus adéquate pour leur modélisation à base de HMM (Hidden Marcov Model).



Figure II.11 Segmentation en bandes verticales uniforme et non-uniforme [31].

II.5 Certains travaux de segmentation de mots arabes en caractères

Plusieurs méthodes de segmentation ont été développées en utilisant des différentes techniques, et dans le suite, on va présenter des travaux de différents auteurs sur la segmentation de textes arabes.

5.1 Zoubeir Mouelhi [30]

Écrit en Visual Basic et tournant sous Windows, AraSeg est un ségmenteur à la fois à grammaire et à dictionnaire. L'analyse de chaque séquence de caractères est basée sur le modèle d'analyse du mot graphique en arabe dans lequel un mot graphique est considéré comme une suite de constituants immédiats. Le rôle principal de ce ségmenteur est l'itémisation, la segmentation lexicale.

5.2 A.Hamid et R.Haraty dans [42]

proposent un algorithme neuro-heuristique pour la segmentation des mots arabes manuscrits. Le processus comporte six étapes. La première concerne l'acquisition du texte, la seconde sa Binarisation en utilisant un algorithme heuristique générant une matrice prenant pour valeur 1 pour un pixel noir et 0 pour un blanc. La troisième étape consiste à extraire les pseudo-mots. La quatrième étape consiste à extraire les caractéristiques à partir de chaque colonne de la matrice de pixels. Les caractéristiques sont de type topographique. La cinquième étape est une génération de points de pré-segmentation. C'est une sur-segmentation basée sur les caractéristiques extraites pour chaque colonne de la matrice et d'une largeur approximative des caractères. La sixième étape est une vérification de la validité des points de pré-segmentation et pour cela un réseau de neurones de 52 entrées, 4 couches cachées et une sortie. Les entrées sont les attributs de caractéristiques d'un point de présegmentation et la sortie est la validité de ce point.

5.3 S.T. Masmoudi et al dans [43]

Présentent l'étape de segmentation associée à un modèle basé PHMM (modèles Markoviens cachés planaires) pour la reconnaissance hors-ligne des noms de villes tunisiennes manuscrites. Le procédé consiste en trois étapes. La première est une segmentation horizontale qui consiste à diviser l'image du texte en cinq zones logiques : la zone de points diacritiques supérieurs, la zone de points diacritiques inférieurs, la zone de hampes, la zone de jambages et la zone médiane. Après cela il y'a détection des boucles qui caractérisent la zone médiane. La seconde étape est appelée segmentation naturelle. Elle consiste à séparer les différents pseudo-mots en détectant l'espace entre les pseudo-mots dans la zone médiane. La troisième étape est appelée segmentation verticale. C'est l'étape la plus difficile dans le processus de segmentation. Elle consiste à examiner le contour supérieur du pseudo-mot pour localiser les points minimaux locaux, et les considérer comme points de segmentation. Le résultat de cette étape est un ensemble de graphèmes qui peuvent correspondre à un caractère, un caractère sur-segmenté (une partie du caractère) ou à un caractère sous-segmenté (deux caractères ou plus comme dans le cas de caractères ligaturés verticalement).

5.4 Sari et al dans [44]

Proposent un algorithme pour la segmentation des caractères arabes manuscrits appelé ACSA (Arabic Character Segmentation Algorithm). Après acquisition de l'image du texte, les mots sont normalisés puis lissés pour réduire le bruit et régulariser le contour des mots. Les pseudo-mots sont extraits de chaque mot par suivi du contour extérieur. L'étape suivante est l'extraction des caractéristiques et là ils utilisent différents types de caractéristiques dont on peut citer les boucles, hamza, hampes, jambages ..., ces caractéristiques sont stockées dans une liste. Après détection de la ligne de base, en utilisant l'histogramme des projections horizontal. La ligne d'écriture est divisée en trois zones, une zone supérieure comportant les hampes, une zone médiane comportant les jambages et une zone médiane comportant le corps principal de l'écriture. Chaque pseudo-mot est ensuite considéré séparément pour être segmenté en caractères. Les points de segmentation sont considérés comme les points minimaux du contour extérieur inférieur du pseudo-mot. après détection de ces points, ils sont validés dans une autre étape en leur appliquant un ensemble de règles pour en générer un ensemble de points de segmentation valides (représentant une segmentation idéale en caractères). Ces segments sont remis au classifieur ainsi que leurs caractéristiques pour être reconnus.

II.6 conclusion

Nous avons vu dans ce chapitre différentes techniques utilisant pour la segmentation des textes, précisant sur les techniques de segmentation de l'écriture arabe manuscrite en présentant certaines ambiguïtés qui rendent la segmentation difficile à réaliser.

On peut dire que la performance d'une segmentation dépend généralement de plusieurs facteurs comme la nature et la qualité du document, l'outil d'acquisition (scanner), et les méthodes de prétraitement et algorithmes de segmentation choisis.

Chapitre III

les réseaux de neurones

III.1 Introduction

Puisque la pensée est produite par le cerveau ou en est une propriété, commencer par étudier comment celui-ci fonctionne. Ce qui a conduit à l'étude de réseaux de neurones formels. On désigne par la démarche consistant à vouloir rendre compte de la cognition humaine par des réseaux de neurones. dans le suivant on va représenter des notions de neurosciences utiles à notre travail, ceci nous donnons ensuite un rapide historique des réseaux de neurones avec quelques modèles célèbres de neurone formel.

Puisque l'apprentissage est vraisemblablement la propriété la plus intéressante des réseaux neuronaux, on va représenter les types d'apprentissage avec quelque règles de lui même .

III.2 Historique

Citons quelques étapes dans la formalisation des réseaux de neurones :

Première définition d'un neurone formel par McCulloch et Pitts en 1943;

Les percepts ou concepts sont physiquement représentés dans le cerveau par l'entrée en activité (simultanée) d'une de neurones (Donald Hebb, 1949);

L'hypothèse concurrente est la spécialisation de certains neurones dans des tâches cognitives complexes;

Deux neurones entrant en activité simultanément vont être associés (c'est-à-dire que leur contacts synaptiques vont être renforcés). On parle de Hebb et d'associationnisme;

Le perceptron de Frank Rosenblatt (1958) : le premier modèle pour lequel un processus d'apprentissage a pu être défini. De cette période, date également les travaux de Widrow et Hoff;

Le livre de Minski et Papert "Perceptrons" (1969). Cet ouvrage contient une étude critique très complète des perceptrons. On lui reproche parfois violemment d'avoir sonné le glas des recherches sur les réseaux neuronaux dans les années 70, ce que nient leurs auteurs. Ce livre a été réédité en 1980, avec des ajouts et corrections manuscrites dans les marges, sans doute pour qu'on ne puisse pas les accuser de camoufler la première version du texte ;

L'algorithme de rétropropagation du gradient dans les réseaux multi-couches découvert au début des années 80 par Rumelhart et McClelland, Parker et Hinto. Les "inventeurs" sont nombreux car l'idée de descente de gradient est naturelle. La plupart de ces travaux étaient associés à des études empiriques montrant la puissance du modèle;

Le modèle de Hopfield (1982) qui utilise des réseaux totalement connectés basés sur la règle de Hebb qui ont permis de définir la notion d'attracteurs et de mémoire associative;

les cartes de Kohonen (1984) avec un algorithme non supervisé basé sur l'auto-organisation;

la machine de Boltzman (1985), autre type de réseaux à attracteurs avec une dynamique de Monte-Carlo;

Aujourd'hui, les réseaux de neurones artificiels sont largement reconnus et utilisés dans la recherche et dans l'industrie.

III.3 Le neurone biologique

3.1 Définition: Un neurone est une cellule capable de transmettre des informations à d'autres neurones au travers de ses différentes connexions (synapses);

- Il existe plusieurs types de neurones avec des fonctionnements différents;
- Les neurone sont inter-connectés et forment des réseaux.

La figure III.1 représente la composition de neurone biologique

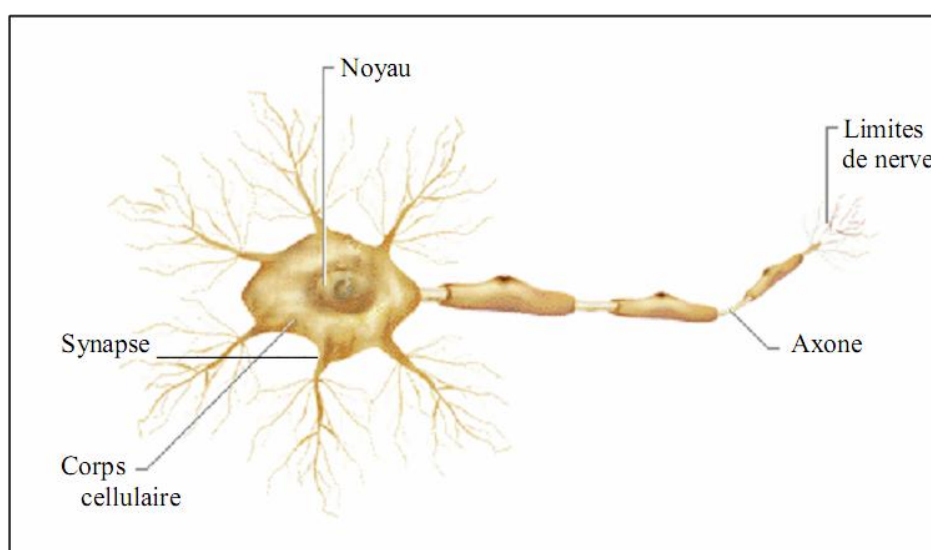


Figure III.1 Schéma d'un neurone biologique

Le neurone biologique est composé de quatre parties distinctes : **le corps cellulaire**, qui contient le noyau de la cellule nerveuse; c'est en cet endroit que prend naissance l'influx nerveux, qui représente l'état d'activité du neurone; **les dendrites**, ramifications tubulaires courtes formant une espèce d'arborescence autour du corps cellulaire; ce sont les entrées principales du neurone, qui captent l'information venant d'autres neurones;

l'axone, longue fibre nerveuse qui se ramifie à son extrémité; c'est la sortie du neurone et le support de l'information vers les autres neurones;

la synapse, qui communique l'information, en la pondérant par un poids synaptique, à un autre neurone; elle est essentielle dans le fonctionnement du système nerveux.

3.2 Système nerveux biologique

Un neurone utilise des dendrites pour rassembler des données d'entrée issues d'autres neurones. Ces données d'entrée sont combinées pour produire une réponse envoyée à d'autres neurones ou d'autres cellules. Comme la figure suivante montre .

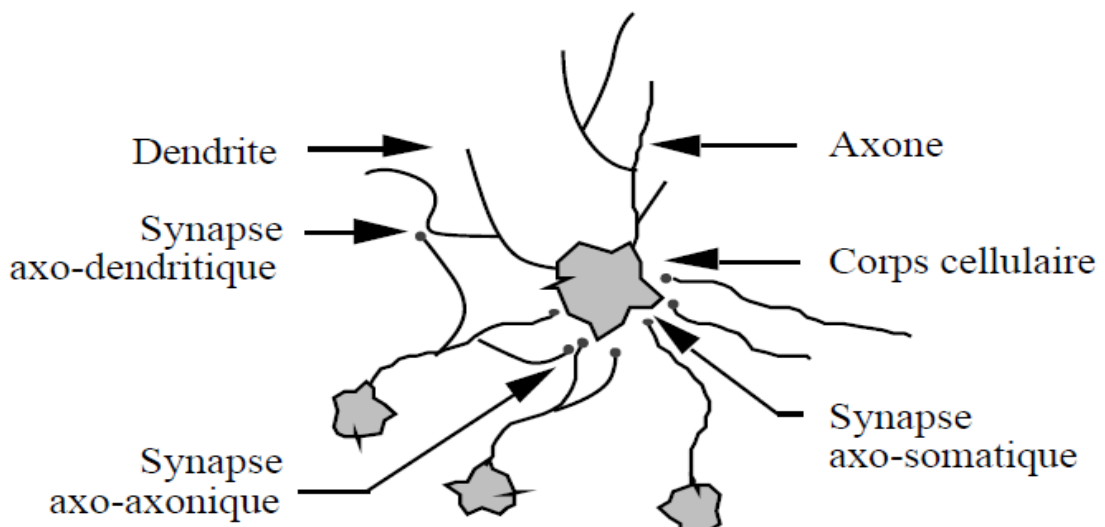


figure III.2 Schéma d'un Système nerveux biologique

3.3 Réseaux de neurones et neurosciences

○ Le cerveau contient environ 100 milliards de neurones. on compte de quelques centaines à plusieurs dizaines de milliers de contacts synaptiques par neurone. Le nombre total de connexions est estimé à environ 10^{15} .

- la vitesse de propagation des influx nerveux est de l'ordre de 100m/s. C'est à dire bien inférieure à la vitesse de transmission de l'information dans un circuit électronique.
- Propriétés collectives de réseaux, qui mettent en jeu des relations excitatrices ou inhibitrices.
- Dans le cerveau, on distingue des sous-ensembles, des connexions internes et externes, des entrées et des sorties, réalisant certaines tâches.
- Les recherches sur le fonctionnement du cerveau ont fait d'énormes progrès (exposition Cité des Sciences, 2002), tant du point de vue de la compréhension du fonctionnement qu'à l'échelle cellulaire et membranaire.
- La connectique du cerveau ne peut pas être codée dans un "document biologique " tel l'ADN pour de simples raisons combinatoires. La structure du cerveau provient donc en partie des contacts avec l'environnement. L'apprentissage est donc indispensable à son développement.
- Il semble que l'apprentissage se fasse par un double mécanisme: des connexions sont établies de manière redondantes et aléatoires puis seules les connexions entre des neurones simultanément actifs sont conservés (phase de sélection) tandis que les autres sont éliminés. On parle de stabilisation sélective.

III.4 Les réseaux de neurones artificiels

4.1 Définitions

Haykin propose la définition suivante [33]:

Un réseau de neurones est un processeur massivement distribué en parallèle qui a une propension naturelle pour stocker de la connaissance empirique et la rendre disponible à l'usage. Il ressemble au cerveau sur deux aspects:

- La connaissance est acquise par le réseau au travers d'un processus d'apprentissage ;
- Les connexions entre les neurones, connues sous le nom de poids synaptiques servent à stocker la connaissance[34].

Selon [41] Un réseau de neurones est un circuit composé d'un nombre très important d'unités de calcul simples basées sur des neurones. Chaque élément opère seulement

sur l'information locale. Chaque élément opère de façon asynchrone; il n'y a donc pas d'horloge générale pour le système [35].

On pourra résumer une modélisation par le tableau 1, qui nous permettra de voir clairement la transition entre le neurone biologique et le neurone formel.

Le neurone biologique	Le neurone artificiel
Dendrite	Fonction de combinaison
Corps du neurone	Fonction de transfert
Axone	Élément de sortie
synapse	Poids

Tableau III.1: Analogie entre le neurone biologique et le neurone formel[45]

4.2 La mise en œuvre d'un réseau de neurones artificiels

La conception d'un système de classification, qu'il soit un système de reconnaissance de formes conventionnel ou un réseau de neurones artificiels, nécessite les étapes suivantes :

A) Acquisition des données

Le choix des données initiales est très délicat car il conditionne en grande partie l'efficacité des performances du réseau de neurones. L'acquisition des données permet de convertir les données de leurs formes physiques (son, image, etc.) en une forme acceptable par l'ordinateur.

B) Prétraitement

Il consiste à localiser, segmenter et normaliser les représentations. Par exemple pour une image (en occurrence l'image d'un caractère), on cherchera à supprimer le bruit et à augmenter le contraste.

C) Le choix des attributs

Le choix d'un vecteur d'attributs pour caractériser un objet peut se révéler délicat. En effet, il faut faire un compromis entre la dimension du vecteur et le contenu des informations. Un vecteur de taille trop petite conduirait à de mauvaises performances du réseau de neurones. Avant le traitement des données nous avons donc à effectuer le

choix des objets, la définition des attributs caractérisant les objets et la construction de la base d'apprentissage. A la fin de cette phase on obtient un tableau de nombres à deux entrées : les données et les attributs les caractérisant.

A la fin de ces étapes on obtient une base de connaissance qui sert comme une base d'apprentissage.

D) La décision

Ayant une nouvelle observation, il s'agit de décider à quelle classe l'affecter. Pour que le système soit performant, il faut qu'il se comporte bien sur une base de données autre que sa base d'apprentissage, cette base s'appelle la base de généralisation.

4.3 Différentes configuration des réseaux:

Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle. Elle peut être quelconque, mais le plus souvent il est possible de distinguer une certaine régularité.

4.3.1 Réseau multicouche: les neurones sont arrangés par couche. Il n'y a pas de connexion entre neurones d'une même couche et les connexions ne se font qu'avec les neurones des couches avales (**Figure III.3**). Habituellement, chaque neurone d'une couche est connecté à tous les neurones de la couche suivante et celle-ci seulement. Ceci nous permet d'introduire la notion de sens de parcours de l'information (de l'activation) au sein d'un réseau et donc définir les concepts de neurone d'entrée, neurone de sortie. Par extension, on appelle couche d'entrée l'ensemble des neurones d'entrée, couche de sortie l'ensemble des neurones de sortie. Les couches intermédiaires n'ayant aucun contact avec l'extérieur sont appelés couches cachées.

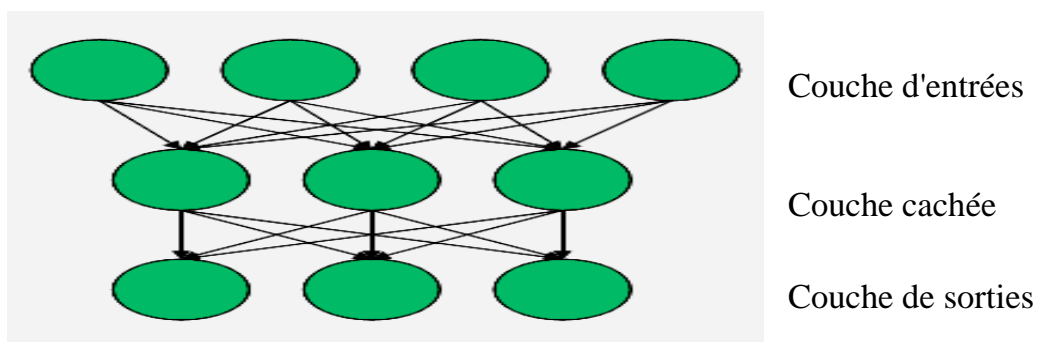


Figure III.3 Schéma présentant un réseau multicouche

4.3.2 Réseau à connexions locales: Il s'agit d'une structure multicouche, mais qui à l'image de la rétine, conserve une certaine topologie. Chaque neurone entretient des relations avec un nombre réduit et localisé de neurones de la couche avale (**Figure III.4**). Les connexions sont donc moins nombreuses que dans le cas d'un réseau multicouche classique.

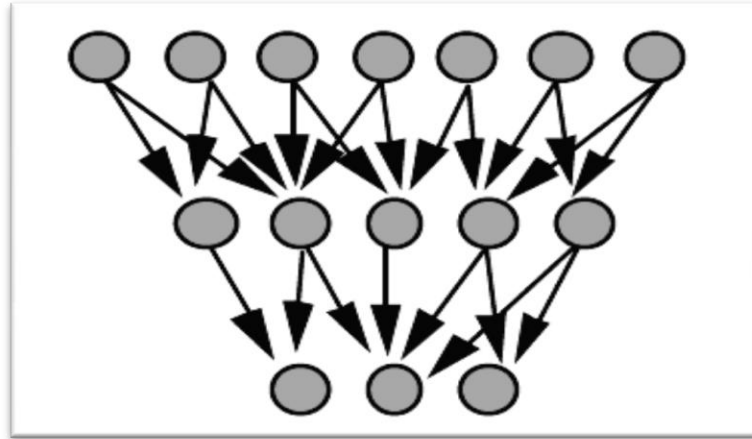


Figure III.4 Réseau à connexions locales

4.3.3 Réseau à connexion complète: c'est la structure d'interconnexion la plus générale (**Figure III.5**). Chaque neurone est connecté à tous les neurones du réseau (et à lui-même).

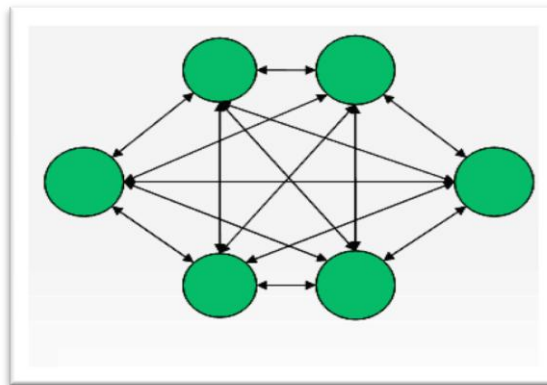


Figure III.5 Réseau à connexion complète.

4.3.4 Réseau à connexions récurrentes: les connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche. Ces connexions sont le plus souvent locales (**Figure III.6**).

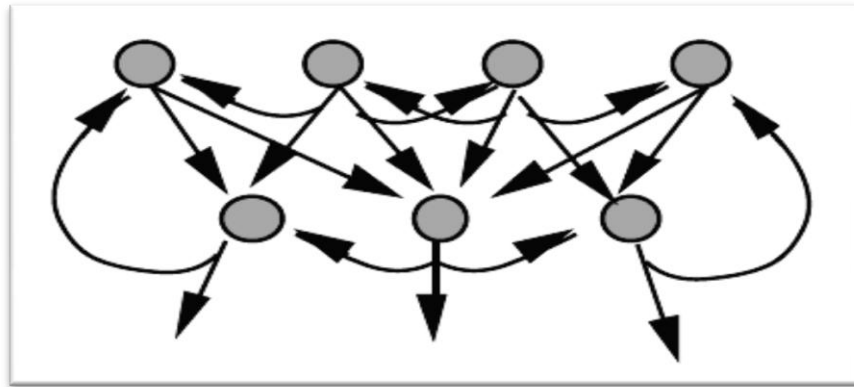


Figure III.6 Réseau à connexions récurrentes

4.4 Les modèles célèbres de réseaux neuronaux

4.4.1 Modèle de McCullochs et Pitts (1943)

Un modèle simple de neurone formel. Il s'agit d'un neurone binaire, c'est-à-dire dont la sortie vaut 0 ou 1. Pour calculer la sortie, le neurone effectue une somme pondérée de ses entrées puis applique une fonction d'activation à seuil. Si la somme pondérée dépasse une certaine valeur, la sortie du neurone est 1, sinon elle vaut 0.

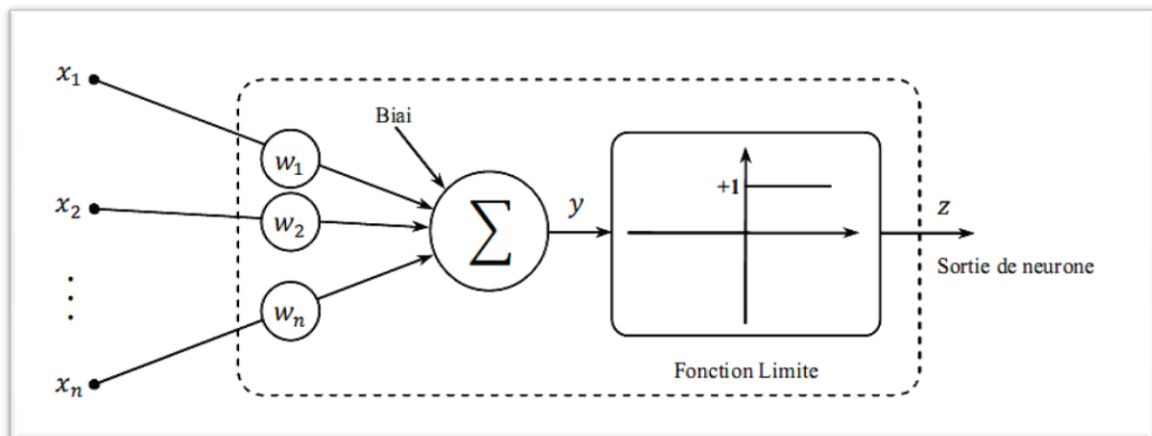


Figure III.7: Le neurone de Mac Culloch-Pitts [35]

4.4.2 Le Perceptron simple (Modèle de Rosenblatt, 1950)

L'origine du réseau PMC remonte la fin des Années 1950 lorsque F. Rosenblatt proposa un premier modèle de réseau de neurones artificiels. Extrêmement simple, ce modèle ne comportait alors qu'un seul neurone: le Perceptron simple. Il permet de classifier correctement des objets appartenant deux classes linéairement

séparables. Il consiste en un seul neurone qui possède un seuil ainsi qu'un vecteur de poids synaptiques ajustables. La structure ainsi que les diverses composantes d'un Perceptron simple sont illustrées dans la figure III.9.

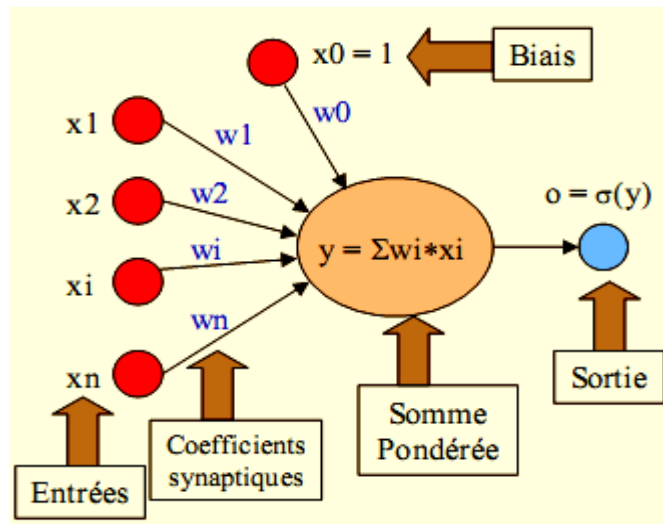


Figure III.9 Le perceptron de Rosenblatt(1950)

4.4.3 Le Perceptron Multi Couches (PMC)

a) Définition: Les perceptrons multicouches sont les réseaux de neurones les plus populaires et les plus simples. Ce sont des réseaux à propagation directe sans cycle, avec une couche d'entrée, une couche de sortie et au moins une couche cachée.

b) la couche d'entrée: Le nombre de neurone de cette couche dépend du nombre de paramètres que peut générer chaque méthode d'extraction des caractéristiques.

c) la couche cachée: Le nombre de neurones de cette couche est choisit après plusieurs essais. Le meilleur nombre étant celui qui donne le meilleur résultat dans le sens du pourcentage de reconnaissance.

d) la couche de sortie: Le nombre de neurones de sortie est dicté par le nombre de situation à séparer.

les neurones sont généralement complètement connectés et la fonction de transfert est de type sigmoïd (valeur dans $[0,1]$), tanh (valeur dans $[-1,1]$) ou softmax (une fonction associée à une fonction de cout J d'entropie relative (distance de Kullbak-Leibler) a été proposée par Bridle, pour évaluer des probabilités a posteriori des classes en sortie . La fonction Softmax force les sorties à vérifier les propriétés des

probabilités : les sorties sont comprises entre 0 et 1 et leur somme est égale à 1)[39]. Le schéma suivante représente un PMC trois couches. La couche d'entrée comporte quatre neurones. La couche cachée contient trois neurones et enfin la couche de sortie possède quatre neurones.

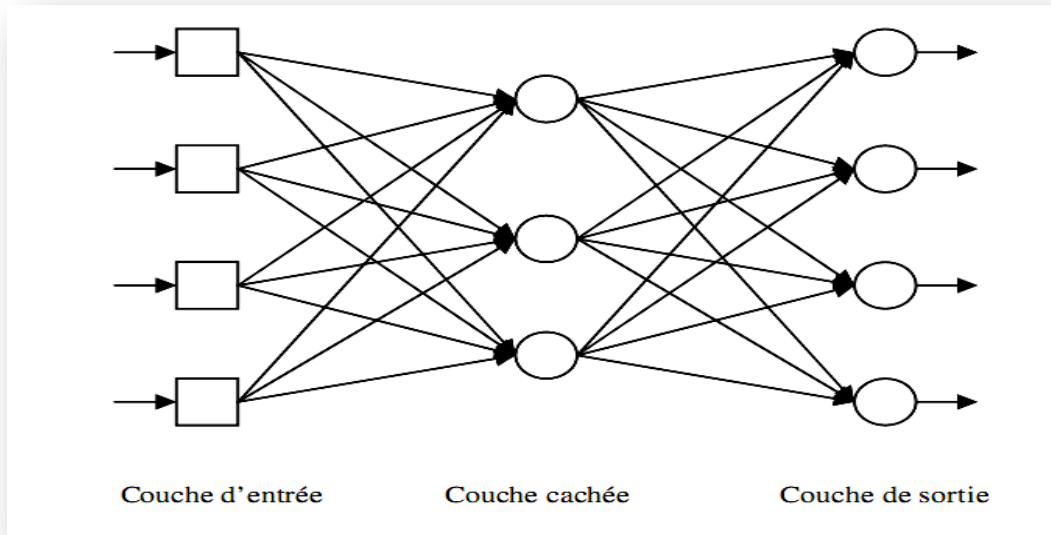


Figure III.10 Architecture d'un Perceptron Multi-couches une seule couche cachée.

Le choix de la fonction de transfert pour les couches cachées réside dans sa dérivabilité et sa simplicité de calcul pour l'apprentissage comme par exemple la fonction sigmoïde ou la fonction tangente hyperbolique :

$$\text{Sigmoïde}(\mathbf{x}) = \frac{1}{1+e^{-x}} \quad \text{entre } [0 \ 1] \quad \text{III.1}$$

$$\text{tangente hyperbolique}(x) = \frac{\sinh x}{\cosh x} = \frac{-1+e^{-2x}}{1+e^{-2x}} \quad \text{entre } [-1 \ 1] \quad \text{III.2}$$

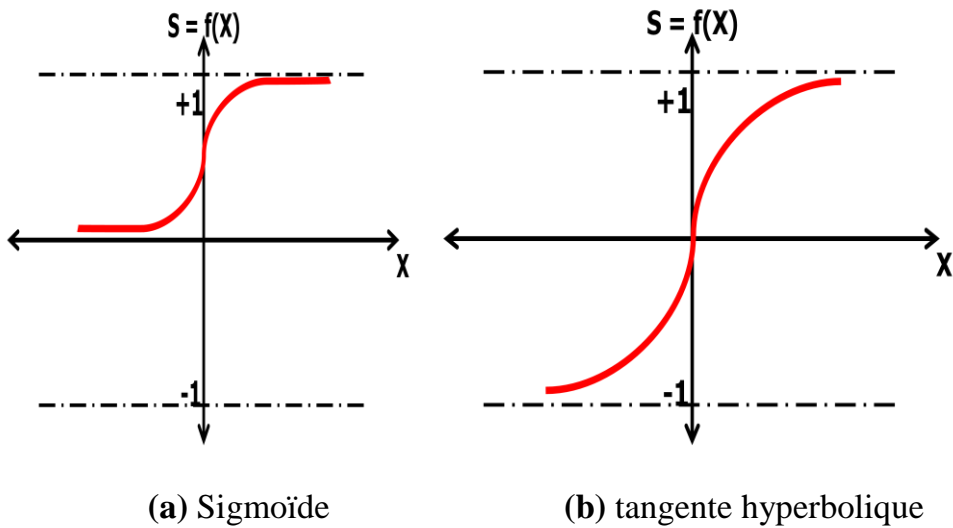
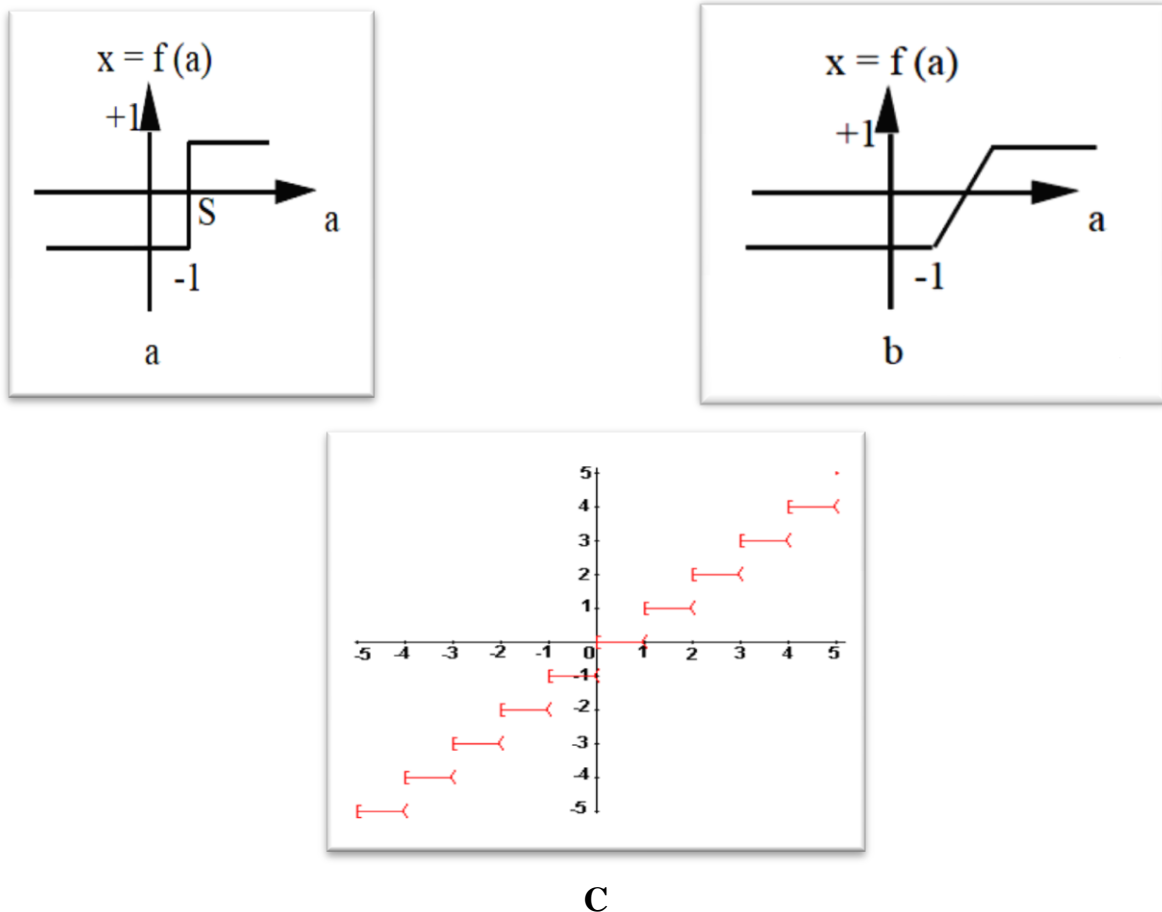


Figure III.11 Exemples de fonctions de transfert d'un PMC.

Il ya d'autre fonctions de transfert de réseau formel



a : fonction à seuil (S , la valeur du seuil), b : linéaire par morceaux, c:fonction à valeurs discrètes.

Figure III.12 schéma représente quelque types de fonction de transfert.

III.5 l'Apprentissage

L'apprentissage est vraisemblablement la propriété la plus intéressante des réseaux neuronaux. Elle ne concerne cependant pas tous les modèles, mais les plus utilisés.

L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. L'apprentissage neuronal fait appel à des exemples de comportement.

a) Définition : L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré [36].

L'apprentissage consiste à descendre le réseau de façon itérative en ajustant les poids à chaque passage selon le calcul d'erreur jusqu'à ce qu'il n'y ait plus d'amélioration. Pour cela, un algorithme de rétro-propagation (backpropagation) de l'erreur est mis en œuvre [40]. Principes de l'algorithme de rétro -propagation (backpropagation) de l'erreur :

- Initialiser la matrice des poids au hasard et Choisir un exemple en entrée;
- Propager le calcul de cette entrée à travers le réseau;
- Calculer la sortie de cette entrée;
- Mesurer l'erreur de prédiction par différence entre sortie réelle et sortie prévue;
- Calculer la contribution d'un neurone à l'erreur à partir de la sortie;
- Déterminer le signe de modification du poids W_i ;
- Corriger les poids des neurones pour diminuer l'erreur;

Le processus est itéré jusqu'à ce qu'un taux d'erreur minimal soit atteint (critère d'arrêt).

b) Base d'apprentissage : exemples représentatifs du comportement ou de la fonction à modéliser. Ces exemples sont sous la forme de couples (entrée ; sortie) connus.

c) Base d'essai : pour une entrée quelconque (bruitée ou incomplète), calculer la sortie. On peut alors évaluer la performance du réseau.

On distingue trois classes d'apprentissage :

- L'apprentissage supervisé
- Apprentissage par renforcement

– L'apprentissage non supervisé

5.1 L'apprentissage supervisé: est l'adaptation des coefficients synaptiques d'un réseau afin que pour chaque exemple, la sortie du réseau corresponde à la sortie désirée.

L'apprentissage supervisé est la forme la plus répandue d'apprentissage (dans le monde biologique aussi bien que technique). Association imposée entre un vecteur d'entrée et un vecteur de sortie (la réponse désirée). L'erreur est calculée à chaque essai afin de corriger les poids. Les poids sont modifiés jusqu'à l'erreur minimale, voire aucune erreur.

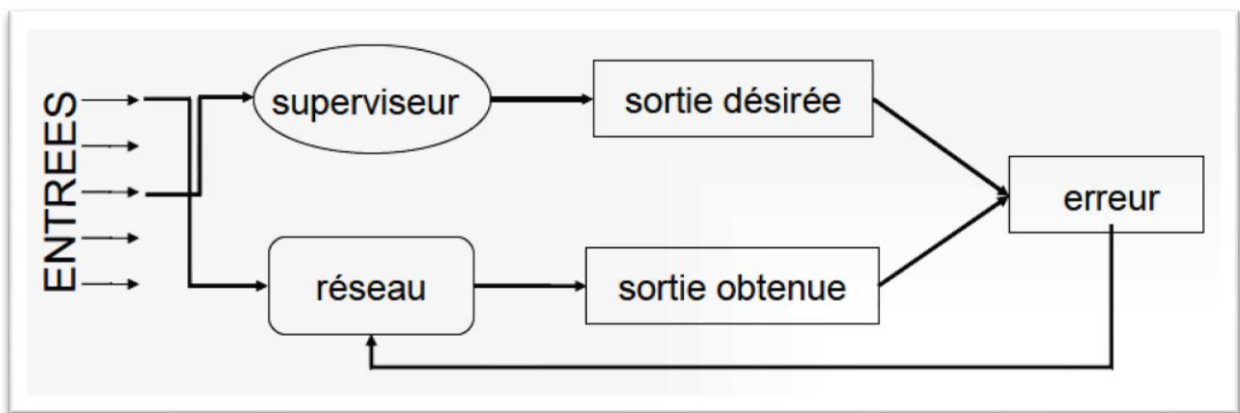


Figure III.13 Schéma représente l'apprentissage supervisé[51].

5.2 Apprentissage par renforcement : Dans ce cas, bien que les sorties idéales ne soient pas connues directement, il y a un moyen quelconque de connaître si les sorties du RNA s'approchent ou s'éloignent du but visé. Ainsi, les poids sont ajustés de façons plus ou moins aléatoire et la modification est conservée si l'impact est positif ou rejetée sinon.

5.3 L'apprentissage non supervisé : on ne dispose pas de base d'apprentissage. Les coefficients synaptiques sont déterminés par rapport à des critères de conformité : spécifications générales.

L'apprentissage non supervisé des réseaux de neurones consiste, comme dans le cas des apprentissages supervisé, à modifier les poids des connections des neurones. Dans ce cas les exemples de la base d'apprentissage sont des données seules :ils n'est pas possible de modifier les poids du réseau souhaitées, puisqu'aucune réponse n'est comme a priori.

Enfant, nous allons voir que ce sont les données elles même qui servent de réponse, car les poids du réseau sont modifier en fonction de leur composant. L'information utile se trouve donc uniquement dans les données, en particulier dans les redondance que peuvent exister dans l'ensemble des données d'apprentissage.

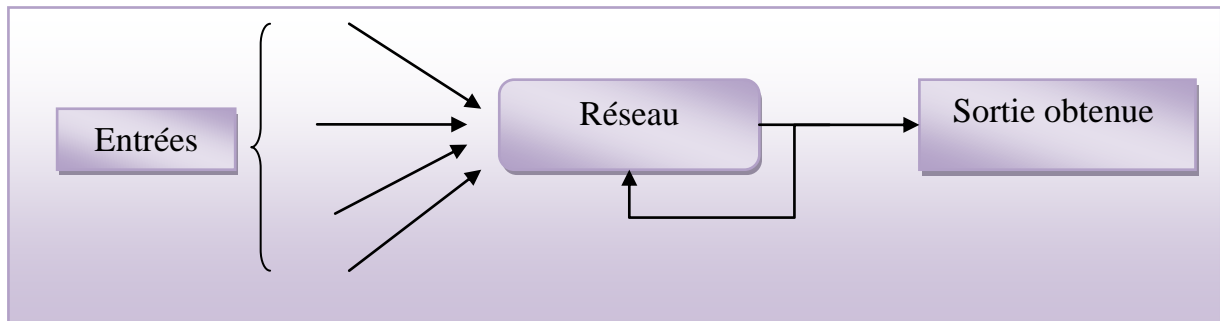


Figure III.14 La figure représente l'apprentissage non supervisé[51].

5.4 Règles d'apprentissage

L'apprentissage consiste à modifier le poids des connections entre les neurones.



Il existe plusieurs règles de modification : parmi les on présente Loi de Hebb, Règle de Widrow-Hoff (delta rule) et Règle de Grossberg .

5.4.1 Règle de Donald Hebb (1949)

Principe : Si deux neurones sont activés en même temps, alors la force de connexion augmente.

Base d'apprentissage :

On note S la base d'apprentissage S est composée de couples (e, c) où :

e : est le vecteur associé à l'entrée (e_1, \dots, e_n) ;

c : la sortie correspondante souhaitée.

Réseau de neurones :

- n entrées e_1, \dots, e_n
- m neurones N_1, \dots, N_m .
- w_{ij} le coefficient synaptique de la liaison entre les neurones N_i et N_j
- une sortie o

- un seuil S
- Fonction de transfert : fonction Signe
 - si $x > 0$: $\text{Signe}(x) = +1$
 - si $x \leq 0$: $\text{Signe}(x) = -1$



Algorithme :

L'algorithme d'apprentissage modifie de façon itérative (petit à petit) les poids pour adapter la réponse obtenue à la réponse désirée. Il s'agit en fait de modifier les poids lorsqu'il y a erreur seulement.

- 1/ Initialisation des poids et du seuil S à des valeurs (petites) choisies au hasard.
- 2/ Présentation d'une entrée $E_i = (e_1, \dots, e_n)$ de la base d'apprentissage.
- 3/ Calcul de la sortie obtenue x pour cette entrée :

$$a = \sum_1^n (w_i \cdot e_i) - s \quad \text{III.3}$$

(la valeur de seuil est introduite ici dans le calcul de la somme pondérée)

$x = \text{signe}(a)$ (si $a > 0$ alors $x = +1$ sinon $x = -1$)

- 4/ Si la sortie x est différente de la sortie désirée d_i pour cet exemple d'entrée E_i alors modification des poids (μ est une constante positive, qui spécifie le pas de modification des poids) :

$$w_{ij}(t+1) = w_{ij}(t) + \mu \cdot (x_i \cdot x_j) \quad \text{III.4}$$

- 5/ Tant que tous les exemples de la base d'apprentissage ne sont pas traités correctement (i.e. modification des poids), retour à l'étape 2. [36]

5.4.2 Loi de Widrow-Hoff (delta rule)

Cette loi se base sur la même hypothèse que la loi de Hebb mais ajoute une variable supplémentaire pour contrôler le taux de variation du poids entre les neurones avec une constante d'apprentissage qui assure à la fois la vitesse de convergence et la stabilité du RNA (Réseaux de Neurones Artificielle).

$$E = d_i - a_i \quad \text{III.5}$$

a_i : activation produite par le réseau;

d_i : réponse désirée par l'expert humain.

Par exemple si la sortie est inférieure à la réponse désirée, il va falloir augmenter le poids de la connexion à condition bien sûr que l'unité j soit excitatrice (égale à 1). On est dans l'hypothèse d'unités booléennes {0,1}.

$$\begin{aligned}
 & a_i = 0 \quad a_i = 1 \\
 d_i = 0 \quad & \Delta W_{ij} = 0 \quad \Delta W_{ij} = -R \\
 d_i = 1 \quad & \Delta W_{ij} = R \quad \Delta W_{ij} = 0
 \end{aligned}$$



$$\Delta W_{ij} = R(d_i - a_i)a_j \quad \text{III.6}$$

R: c'est la vitesse d'apprentissage.

5.4.3 Loi de Grossberg

Loi de Grossberg, qui est une version modifiée de la loi de Hebb. On augmente les poids qui entrent sur l'unité gagnante a_i s'ils sont trop faibles, pour les rapprocher du vecteur d'entrée a_j . C'est la règle d'apprentissage utilisée dans les cartes auto-organisatrices de Kohonen.

III.8



$$\Delta W_{ij} = R a_i (a_j - W_{ij}) \quad \text{III.7}$$

R: c'est la vitesse d'apprentissage.

III.6 L'apprentissage d'un PMC

L'apprentissage d'un réseau PMC est de type supervisé c'est-à-dire que l'on présente au réseau, en même temps, une forme et son modèle, ce qui consiste à appliquer des couples (entrées, sorties désirées) à l'entrée du réseau. L'algorithme le plus utilisé est celui de rétropropagation des erreurs qui consiste à lui présenter un ensemble de N données d'entraînement : $D = \{(x_1, s_1), (x_2, s_2), \dots, (x_N, s_N)\}$, ceci dans le but d'ajuster itérativement ses différents paramètres de manière à minimiser une certaine fonction d'erreur. Une donnée d'entraînement $(x, s) \in D$ est en fait un couple de vecteurs $(x, s) \in \mathbb{R}^n \times \mathbb{R}^m$ tel que x est un vecteur d'entrée qui est propagé à travers toutes les couches du PMC, jusqu'à la couche de sortie, et s est le vecteur des sorties désirées [46]. Rappelons que le vecteur s joue en quelque sorte le rôle du professeur et sert à superviser l'entraînement. Le processus d'apprentissage par rétropropagation est dit itératif dans le sens où les paramètres du réseau sont ajustés pour une seule donnée

d'entraînement à la fois, ceci tant que la fonction d'erreur n'a pas atteint un seuil acceptable. L'algorithme d'apprentissage comprend les étapes suivantes :

Etape 1 : la première étape de l'algorithme consiste à initialiser aléatoirement les coefficients synaptiques d'un réseau initial. La structure du réseau initial (le nombre de couches et le nombre de neurones) est déterminée par l'utilisateur. Habituellement, les valeurs des coefficients synaptiques et des biais sont fixées à de petites valeurs aléatoires distribuées uniformément [49]. Bien que cette initialisation aléatoire puisse avoir en théorie un impact sur la convergence d'un réseau, en pratique, elle n'est absolument pas critique.

Etape 2 : par la suite, un ordre aléatoire est construit pour la représentation des données. Cette permutation aléatoire vise à éviter que le réseau reste pris dans un minimum local durant la phase d'apprentissage.

Etapes 3 et 4 : l'apprentissage proprement dit d'un couple (x, s) est réalisé aux étapes 3 et 4. D'abord, le vecteur d'entrée x est propagé à travers tout le réseau à l'étape 3. À l'étape 4, l'erreur engendrée par x est mesurée en comparant les sorties observées du réseau au vecteur de sortie désirée. En pratique, l'erreur mesurée pour chaque neurone de sortie est la différence entre la valeur désirée s_k et la valeur observée y_k , $k = 1, 2, \dots, m$. Cette erreur mesurée est alors rétropropagée à travers le réseau de manière récursive, couche par couche, en commençant par la couche de sortie. Lors de la rétropropagation, chaque paramètre du réseau est ajusté proportionnellement au taux de changement de la fonction d'erreur par rapport au changement du paramètre concerné, c'est-à-dire proportionnellement au gradient de l'erreur [47]. Ce gradient s'obtient en dérivant partiellement la fonction d'erreur par rapport au paramètre. Notons qu'une condition nécessaire pour que la fonction d'erreur soit dérivable partiellement est que la fonction d'activation soit différentiable. Par conséquent, contrairement au Perceptron simple qui fait usage d'une fonction de seuil pour activer un neurone, l'algorithme de rétropropagation requiert, comme il est cité précédemment l'utilisation d'une fonction continue différentiable. Généralement, la fonction sigmoïde est utilisée à titre de fonction d'activation, défini par l'équation suivante:

$$\varphi(\vartheta) = \frac{1}{1+e^{-\vartheta}}$$

III.8

Mentionnons simplement qu'à l'étape 4, un facteur de gain η permet de contrôler le taux d'apprentissage du réseau (le taux de changement des paramètres proportionnellement au gradient de l'erreur) et un facteur de momentum α assure une certaine stabilité au processus de convergence.

Étapes 5 et 6 : à l'étape 5, le couple suivant qui apparaît dans l'ordre créé aléatoirement est choisi pour l'entraînement. Lorsque tous les couples ont été entraînés, nous disons qu'une passe d'apprentissage est accomplie et nous vérifions, à l'étape 6, si le réseau a convergé, c'est-à-dire s'il a trouvé une transformation $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ qui est optimale au sens d'un certain critère d'arrêt. Dans les faits, il n'existe pas de critère universel pour décider du moment opportun pour conclure un entraînement par rétropropagation des erreurs [48].

Des critères couramment employés consistent à terminer l'apprentissage lorsqu'un nombre maximal de passes est atteint, ou encore lorsque l'erreur mesurée est inférieure à un niveau jugé acceptable. Par contre, ces critères sont très subjectifs et requièrent de la part de, l'utilisateur une bonne connaissance a priori des résultats escomptés.

III.7 Les avantages et Les inconvénients des réseaux neuronaux

L'intérêt porté aux réseaux de neurones tient sa justification dans les propriétés intéressantes qu'ils possèdent; ils présentent aussi un certain nombre de limitations ou d'inconvénients. A partir de ces propriétés, les applications potentielles de l'approche neuronale peuvent être déduites [35].

A)Avantages

L'exploitation de connaissances empiriques : L'apprentissage à partir d'exemples (méthode d'apprentissage empirique) se fait d'une façon assez simple et permet d'obtenir de bons résultats par rapport aux autres techniques d'apprentissage automatique.

La robustesse

Dans les réseaux de neurones, la mémoire est distribuée, elle correspond à une carte d'activation de neurones. Cette carte est en quelque sorte un codage des faits mémorisés ce qui attribue à ces réseaux l'avantage de résister aux bruits (pannes) car la perte d'un élément ne correspond pas à la perte d'un fait mémorisé. Ainsi, un réseau

peut bien fonctionner même quand des unités sont en panne. De plus, de nombreux modèles de réseaux donnent de bons résultats même quand leurs entrées sont bruitées.

La dégradation progressive : Les réseaux, de par leur nature continue, ne fonctionnent pas en tout ou rien et leurs performances ont plutôt tendance à diminuer progressivement en cas de problème (bruit, panne, entrée inconnue...). Cette propriété est très recherchée car les systèmes cognitifs vivants montrent une telle faculté. Les réseaux permettent de bien généraliser les connaissances présentes dans la base d'apprentissage et sont moins sensibles aux perturbations que les systèmes symboliques. Le fait de travailler sur une représentation numérique des connaissances rend les réseaux plus adaptés pour manipuler des données quantitatives (valeurs continues). Les réseaux de neurones sont moins vulnérables aux données approximatives et à la présence de données incorrectes dans la base d'apprentissage.

Le parallélisme massif

Les réseaux sont composés d'un ensemble d'unités de traitement de l'information qui peuvent opérer en parallèle. Bien que la plupart des implémentations des réseaux connexionnistes soient réalisées sur des simulateurs séquentiels, il est possible de faire des implémentations (logicielles ou matérielles) exploitant la possibilité d'activer simultanément les unités. La plupart des implémentations des réseaux de neurones peuvent être facilement converties d'une version séquentielle vers une version parallèle.

La prise en compte du non linéarité et du temps

Les réseaux de neurones artificiels présentent l'intérêt de pouvoir prendre en compte la non linéarité (les fonctions d'activation sont en général non linéaires). Certains réseaux peuvent aussi prendre en compte les aspects temporels (cas des réseaux récurrents).

B) Inconvénients

La difficulté de choix de l'architecture et des paramètres

Il n'existe pas de méthode automatique pour choisir la meilleure architecture possible pour un problème donné. Il est assez difficile de trouver la bonne topologie du réseau ainsi que les bons paramètres de réglage de l'algorithme d'apprentissage. L'évolution du processus d'apprentissage est très influencée par ces deux éléments (l'architecture

du réseau et les paramètres de réglage) et dépend beaucoup du type de problème traité. Le simple fait de changer la base d'apprentissage utilisée, peut nous obliger à reconfigurer le réseau en entier.

Le problème d'initialisation et de codage

Les algorithmes d'apprentissage connexionniste sont en général très dépendants de l'état initial du réseau (initialisation aléatoire des poids) et de la configuration de la base d'apprentissage. Un mauvais choix des poids employés pour initialiser le réseau, de la méthode de codage des données, ou même de l'ordre des données, peut bloquer l'apprentissage ou poser des problèmes pour la convergence du réseau vers une bonne solution;

Le manque d'explicabilité

Les connaissances acquises par le réseau sont codées par l'ensemble des valeurs des poids synaptiques ainsi que par la façon dont les unités sont interconnectées. Il est très difficile pour un être humain de les interpréter directement. Les réseaux connexionnistes sont en général des boîtes noires, où les connaissances restent enfermées et sont inintelligibles pour l'utilisateur ou pour l'expert. Un réseau ne peut pas expliquer le raisonnement qui l'a amené à une solution spécifique.

Le manque d'exploitation de connaissances théoriques

Les réseaux classiques ne permettent pas de profiter des connaissances théoriques disponibles sur le domaine du problème traité. Ils sont dédiés à la manipulation de connaissances empiriques. Une façon simpliste de profiter des connaissances théoriques consiste à convertir des règles en exemples (prototypes). Cependant, cette méthode ne garantit pas que ces exemples aillent être bien représentés dans les connaissances du réseau à la fin de l'apprentissage, car nous sommes obligés à passer pour une phase d'apprentissage où se mélangent sans distinction des connaissances empiriques avec des connaissances théoriques codées par des exemples.

III.8 Les domaines d'application des réseaux de neurones:

On utilise des réseaux de neurones artificiels (RNAs) pour tenter de résoudre des problèmes complexes qui sont plus faciles à résoudre par le cerveau que par un ordinateur [52].

Approximation de fonctions : les fonctions trop compliquées peuvent être approximées, grâce au réseau, par une somme de fonctions plus simples comme des polynômes ou des sigmoïdes.

Optimisation de trajectoires : On peut, par exemple, déterminer quelle est la meilleure trajectoire pour un avion, une fusée...

Reconnaissance : un réseau peut servir à reconnaître des caractères. Cela est déjà utilisé à la poste pour lire les codes postaux, ou même dans certaines banques pour lire les chèques. Il est aussi possible de retrouver la prononciation des mots à partir d'un texte.

Prévision : on utilise de plus en plus les réseaux pour faire des prévisions en marketing (prédiction de comportement, de possibilité de vente d'un produit, ...) ou pour le trafic routier... Mais les prévisions en météo ou en bourse sont trop compliquées à réaliser.

Contrôle : on peut contrôler les produits dans une industrie.

Robotique : certains robots sont dotés de réseaux de neurones. Des entreprises japonaises se vantent déjà de leur utilisation, même pour des produits électroménagers ou informatiques.

III.9 Conclusion

Dans ce chapitre, on représentait une méthode de classification qui est les réseaux de neurone. En général, la plupart des méthodes d'apprentissage comme les réseaux de neurones possèdent un grand nombre de paramètres d'apprentissage fixés par l'utilisateur. Aussi, on abordera dans le prochain les résultats obtenus par cette méthode de classification appliquée à la reconnaissance des mots arabes manuscrits.

Chapitre IV

résultats expérimentaux et
discussions

IV.1 Introduction

Dans ce chapitre nous allons aborder la dernière partie qui représente la partie réalisation de notre projet, en se basant sur les mécanismes évoqués précédemment dans le chapitre de réseaux de neurones.

Ce chapitre est composé de deux parties : la première présente l'environnement notre logiciel, ainsi que les résultats des tests effectués.

IV.2 Environnement de développement

2.1 Environnement hard

- ❖ Hôte : HP Pavilion g series.
- ❖ Microprocesseur : Intel **I3** Inside.
- ❖ RAM : 4GO.

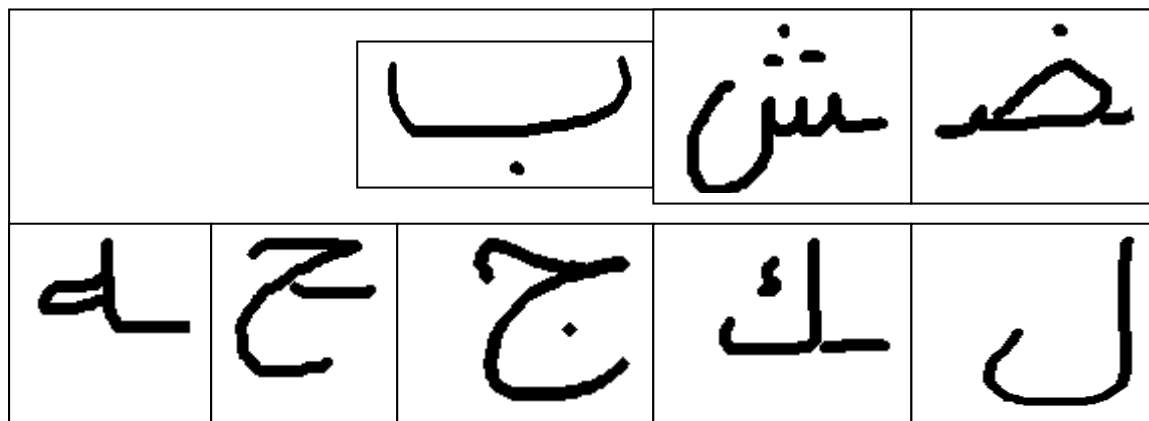
2.2 Présentation de l'environnement de programmation Embarcadero RAD Studio XE

Embarcadero Technologies annonce la disponibilité mondiale d'[Embarcadero RAD Studio XE](#), suite d'outils de développement incluant les nouvelles versions de Delphi, C++Builder, Delphi Prism et, désormais, RadPHP XE. La nouvelle version « 2011 », baptisée RAD Studio XE, propose de nouvelles fonctionnalités pour améliorer les performances des applications, rationaliser le processus de développement et réduire le délai nécessaire pour bâtir des applications Windows, .Net et Web.

IV.3 Description des bases de données

3.1 La base de données des caractères

Nous avons construit une base de données contenant 4840 images de caractères arabes manuscrits. Les lettres se trouvent dans les différentes positions (isolée, début, milieu et fin), pour les lettres isolées on a $100 \times 28 = 2800$ images et pour les autres il y a $30 \times 68 = 2040$ images, donc le nombre de classes est $28+68=96$ (Tableau 4.1). Cette base est divisée en deux parties une pour l'apprentissage contenant 3309 images est l'autre pour le test contenant 1531 images.



La figure IV.1 des échantillons de la base des caractères.

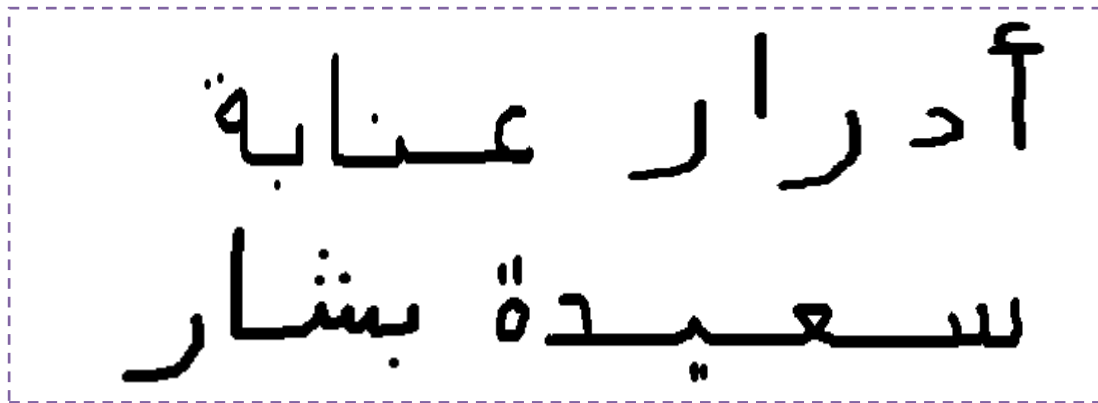
N°	Lettre isolé	Au début	Au milieu	A la fin
1	أ	/	أ	/
2	ب	ب	ب	ب
3	ت	ت	ت	ت
4	ث	ث	ث	ث
5	ج	ج	ج	ج
6	ح	ح	ح	ح
7	خ	خ	خ	خ
8	د	/	د	/
9	ذ	/	ذ	/
10	ر	/	ر	/
11	ز	/	ز	/
12	س	س	س	س
13	ش	ش	ش	ش
14	ص	ص	ص	ص
15	ض	ض	ض	ض
16	ط	ط	ط	ط
17	ظ	ظ	ظ	ظ
18	ع	ع	ع	ع
19	غ	غ	غ	غ
20	ف	ف	ف	ف

21	ق	قـ	قـ	قـ
22	ك	كـ	كـ	كـ
23	ل	لـ	لـ	لـ
24	م	مـ	مـ	مـ
25	ن	نـ	نـ	نـ
26	هـ	هـ	هـ	هـ
27	و	/	وـ	/
28	ي	يـ	يـ	يـ

Tableau VI.1 Les lettres arabes utilisées avec ses différentes formes (96 classes).

3.2 Notre base (base des mots)

C'est une base de données, qui contient des noms manuscrits des wilayas algériennes (48 wilayas) en arabes. Pour chaque wilaya on prend 50 exemples ; $50 \times 48 = 24000$ images. La figure suivante représente des échantillons de la base.

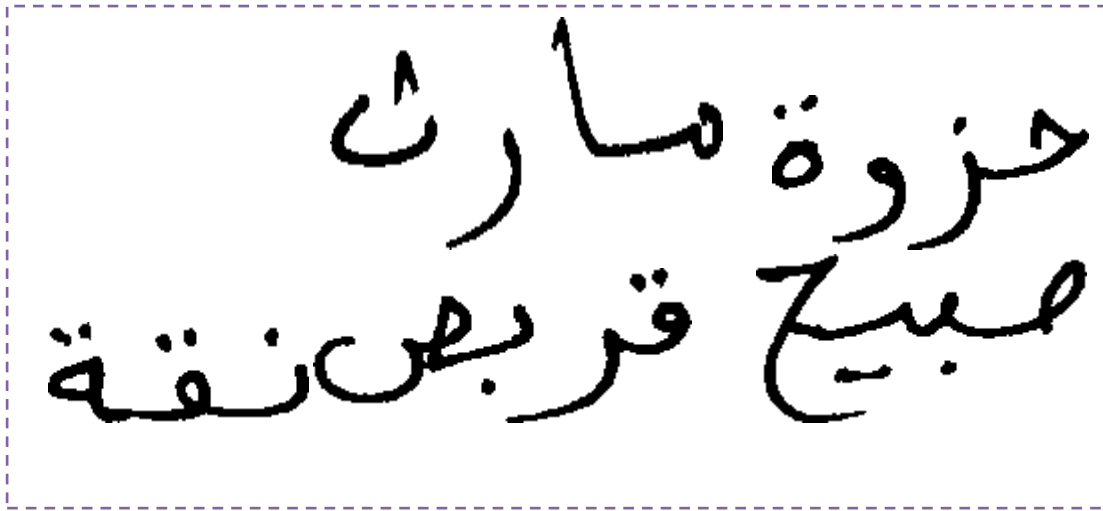


La figure IV.2 des échantillons de notre base.

3.3 Description de la base de données *IFN/ENIT DATABASE*

Lors de l'évaluation des performances d'un système de reconnaissance il est nécessaire de disposer d'une base de données. Il est possible d'obtenir auprès de l'IFN/ENIT <http://www.ifnenit.com/> la base de données IFN/ENIT DATABASE. C'est une nouvelle base de données, qui contient des noms manuscrits de villes/villages tunisiens en arabes. Pour chaque nom les informations de base, par exemple l'ordre des formes de caractère, les informations sur le style de l'écriture, et la ligne de base sont codées.

Plus de 400 auteurs ont rempli des formulaires. La figure suivante représente des échantillons de la base.



La figure IV.3 des échantillons de la base IFN/ENIT.

IV.4 Description de notre système de reconnaissance

4.1 Acquisition de l'image

Cette phase de travail ne pose pas de problème parce que les formes des images dans la base IFN/ENIT sont bien préparées, et sous format: bitmap ('.bmp') et en noir et blanc; car il est très facile de manipuler les objets bitmap par les langages de programmation.

4.2 Prétraitement: Dans cette étape on procède à suite d'opération, on utilise :

- ❖ Binarisation (pour le cas général), et des techniques pour éliminer le bruit par seuillage.
- ❖ Dilatation, cette opération est effectuée parfois pour agrandir l'épaisseur du tracé car l'écriture par la tablette graphique est avec une épaisseur très petite.
- ❖ Squelettisation, Le but de cette technique est de simplifier l'image du caractère en une image à « ligne » plus facile à traiter en la réduisant au tracé du caractère.
- ❖ Erosion, utilisé pour éliminer le bruit (les pixels éloignés) pour pouvoir extraire le cadre qui inclut le caractère.

La figure suivante représente le Schéma générale de système.

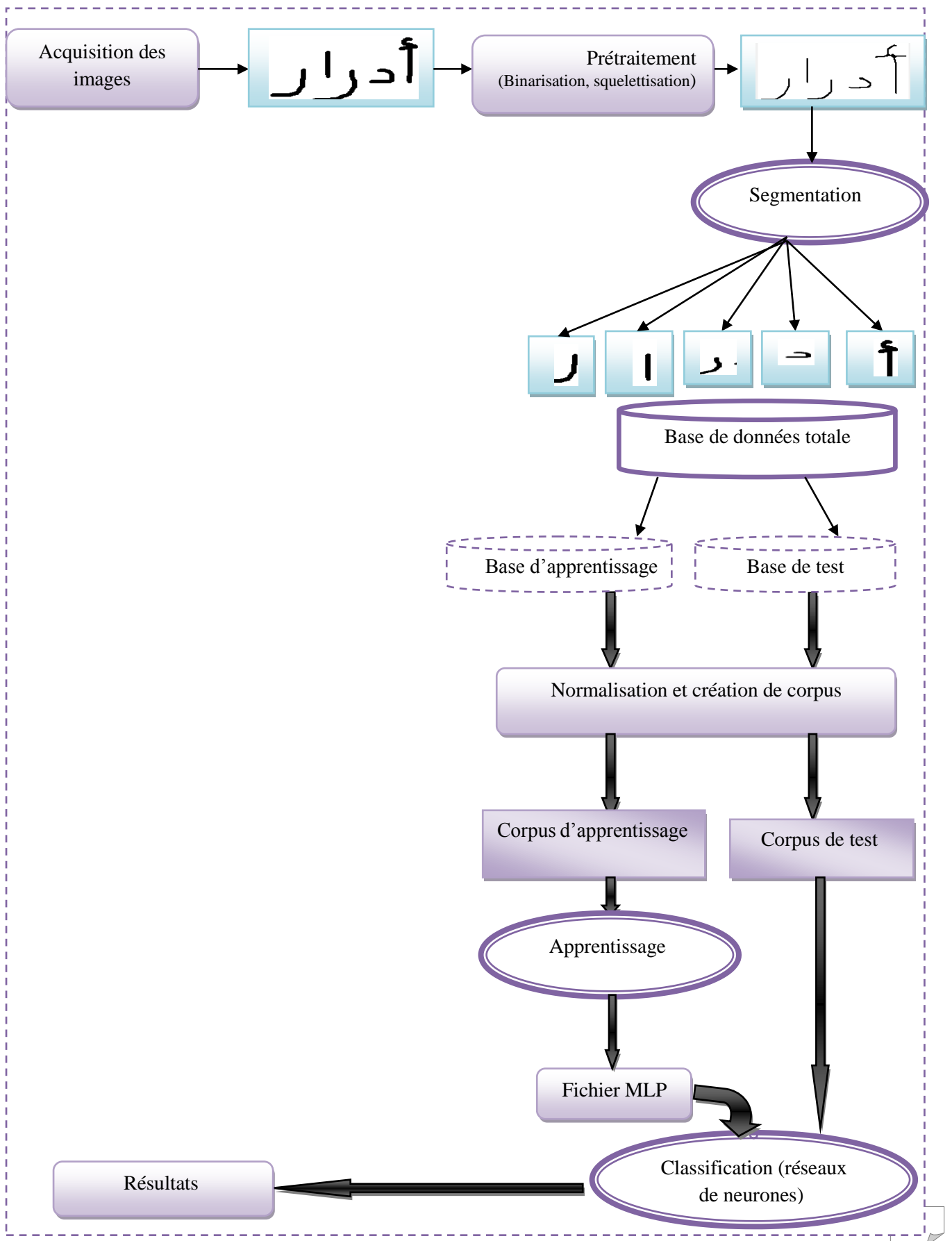


Figure IV.4 Schéma générale de notre système de reconnaissance.

4.3 La segmentation

Nous avons choisi comme méthode de segmentation la méthode de projection verticale, il s'agit de trouver un segment qui inclut un seul caractère dans sa surface de pixels.

Dans cette phase nous avons proposé un algorithme qui a les paramètres suivants :

- ❖ La taille de bloc (T_b);
- ❖ Le pas (P): qui représente le point de départ de prochain bloc;
- ❖ La différence($Diff$): on obtient par le calcul $S-SP$ tel que
 - S : le nombre de pixel noir dans le bloc actuel ;
 - SP : le nombre de pixel noir dans le bloc précédent .

4.3.1L'algorithme

```

Définir les paramètres :  $T_b$ ,  $P$  et  $Diff$ 
Initialisation des variables :
 $S=0$ ;  $Sp=0$ ;  $i=P$ ;  $j=0$ ;
Entrée:
Histogramme de projection vertical:  $His[]$ 
Sortie :
Vecteur de points de segmentation :  $Seg[]$ 
Début
  Pour(  $k=0$  ;  $k \leq T_b$ ;  $k++$ )
     $Sp=Sp+His[k]$  ;
  Fin pour
Tant que ( $i < \text{Taille d'image} - T_b$ ) faire
  Pour ( $k=i$  ;  $k \leq P+T_b$  ;  $k++$ )
     $S=S+His[k]$  ;
  Fin pour
  Si ( $Sp-S > Diff$ ) alors
     $Seg[j]=i+T_b$  ;
     $j=j+1$ ;
     $Sp=0$  ;
  Sinon
     $Sp=S$  ;
  Fin si
   $i=i+P$  ;
Fin tant que
Fin

```

4.3.2 L'organigramme

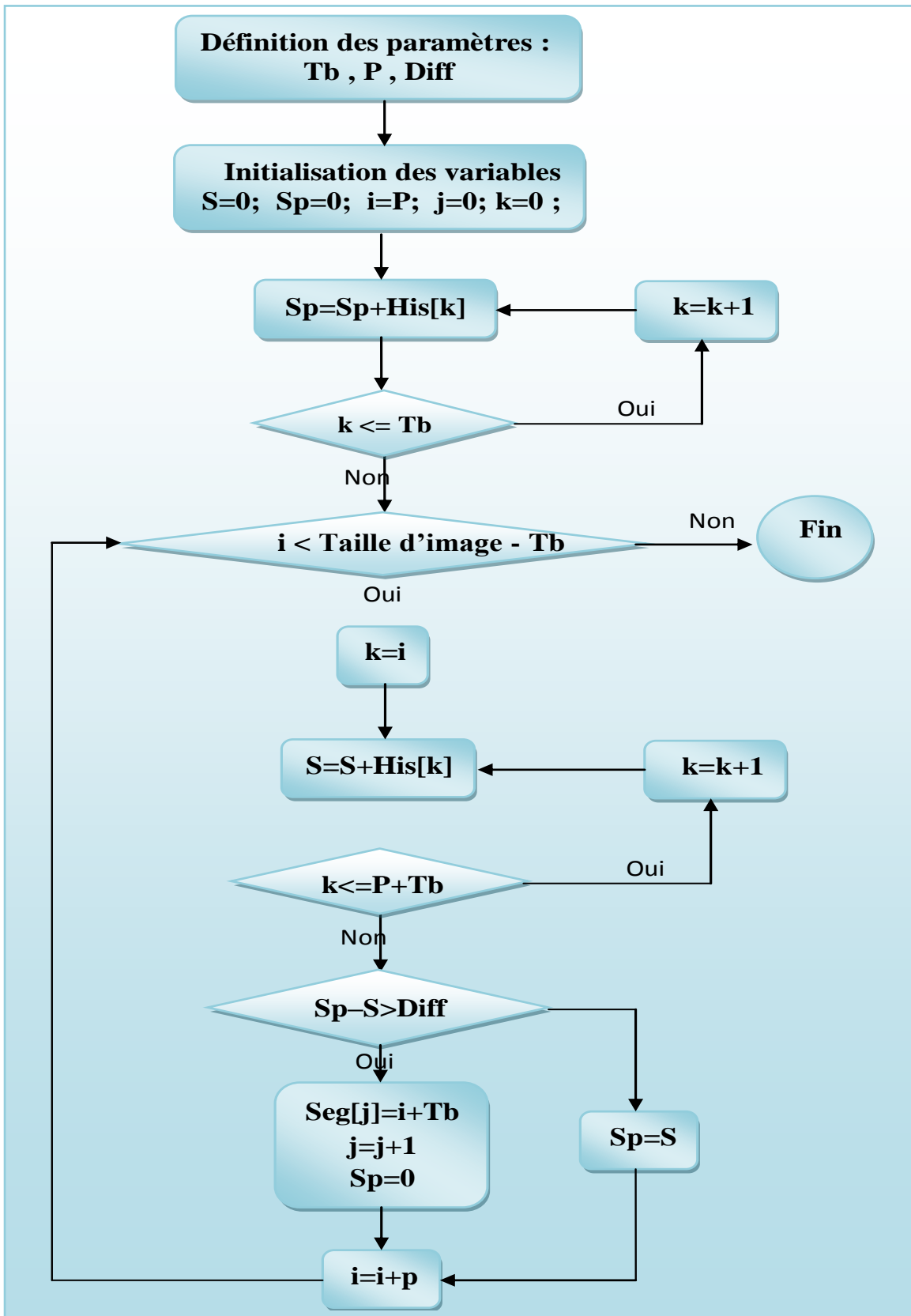


Figure IV.5 L'organigramme d'algorithme de segmentation.

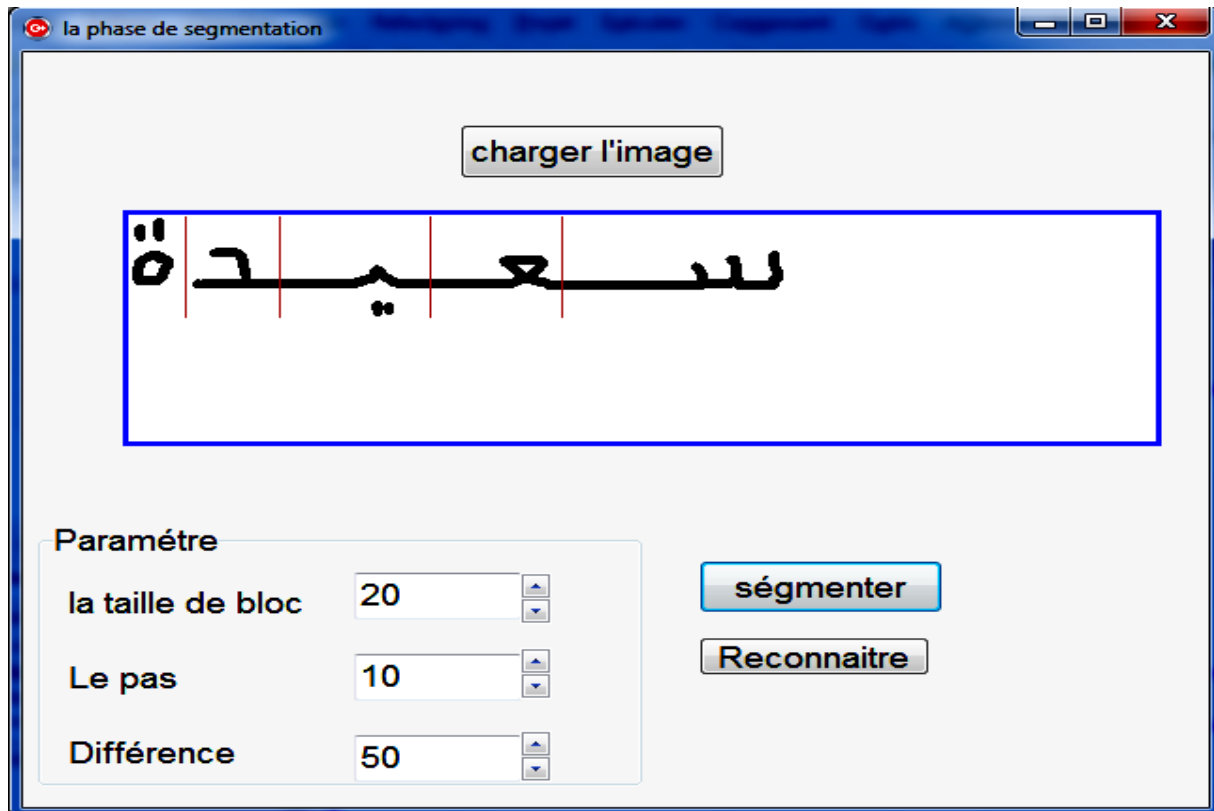


Figure IV.6 Résultat de segmentation de mot " سعيدة "

4.4 La normalization

Pour faciliter les traitements ultérieurs et pour éliminer les conditions qui peuvent fausser les résultats (comme la différence de taille), on a besoin d'effectuer l'opération de la normalisation. Dans ce cas on fournit un cadre de taille 4900 (Hauteur 70 / Largeur 70).

4.5 L'extraction des primitives

C'est une phase critique lors de la construction d'un système de reconnaissance. Pour faire la reconnaissance de mot, il faut d'abord analyser sa forme et extraire ses primitives, cette étape faite en deux parties: Construire la matrice de distribution et faire le vecteur de caractéristiques.

4.5.1 Matrice de distribution: La construction de la matrice de distribution est l'une des phases importantes dans notre système.

Pour une matrice de distribution de taille $N \times N$, le principe consiste à superposer une grille ($N \times N$) sur l'image du caractère et calculer pour chacune des cellules $[i,j]$

résultantes le nombre de pixels noir, puis attribuer ce nombre à la case[i,j] de la matrice de distribution.

Nous considérons par exemple la représentation de la lettre 'khaa', qui s'écrit en arabe 'خ' dans sa forme isolée, sur une matrice de distribution de 7x7 (Figure IV.7).

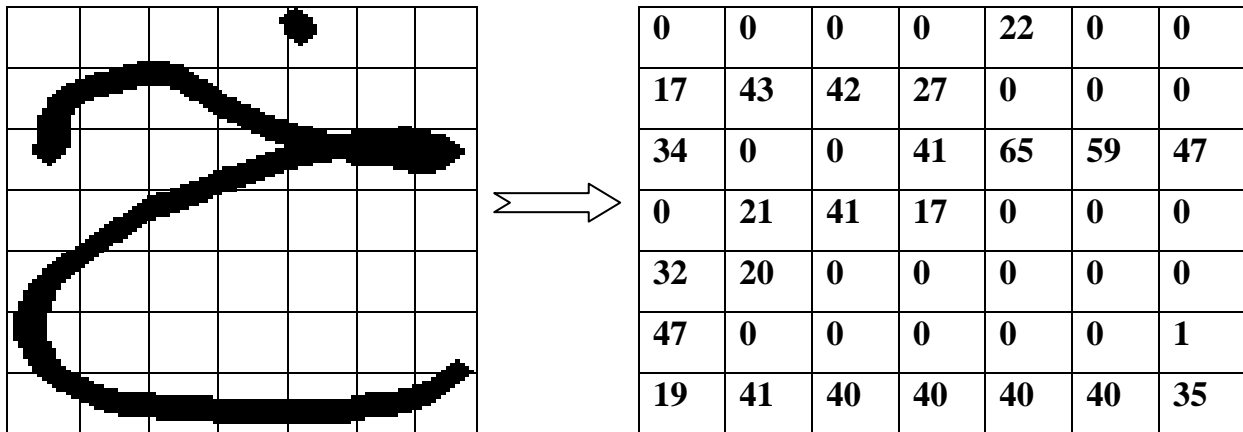


Figure IV.7 Exemple de matrice de distribution (7*7) de la lettre alphabet arabe «khaa ».

4.5.2 Vecteur des caractéristiques

Le vecteur proposé est de taille 49 représentant les valeur de la matrice de distribution(7*7), la figure suivants montre un exemple de cette opération pour le lettre "ص"

0	7	30	35	36	39	31	31	0	0	33	38	35	43	0	0	0
1	2	3	4	5	6	7		40	41	42	43	44	45	46	47	48	49

4.5.3 Corpus

Le corpus est un fichier texte possède une structure particulière où Chaque ligne du corpus représente le vecteur de caractéristique d'une image, la première valeur de la ligne est le numéro de la classe, ensuite chaque valeur de vecteur de caractéristique est précédé par un index (Index:Valeur).

Nous présentons ci-dessous une ligne de corpus d'apprentissage pour la lettre 'Ba' au debut « — » qui a le numéro de la classe 30.

30 1:0 2:0 3:0 4:0 5:0 6:0 7:25 8:0 9:0 10:0 11:0 12:0 13:0 14:50 15:6 16:10 17:10 18:10 19:10 20:10 21:62 22:67 23:90 24:90 25:90 26:90 27:90 28:56 29:0 30:0 31:0

32:0 33:0 34:0 35:0 36:0 37:0 38:0 39:0 40:0 41:1 42:0 43:9 44:10 45:10 46:10 47:10
48:47 49:14

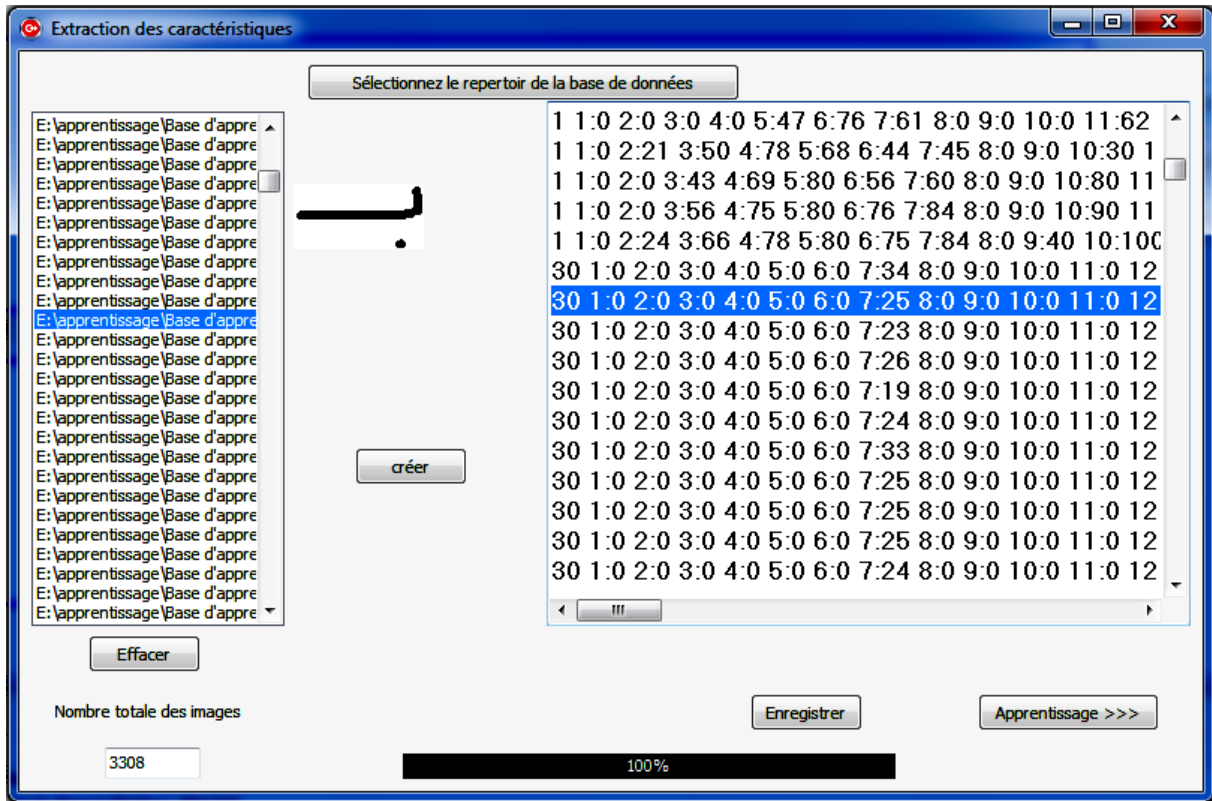


Figure IV.8 Interface de création des corpus.

IV.5 Test et résultat

Pour la phase segmentation, les premiers essais de segmentation des mots en caractères, appliqués sur notre base de données (base des wilayas Algériennes), montrent l'efficacité de cette technique où on a abouti à un taux de segmentation (82.02%).

Ce qui nous a poussé à optimiser et évaluer cette approche de segmentation pour la nouvelle base IFN /ENIT ; où on a abouti un taux de segmentation des mots en caractères de 66.20 %, sur les exemples choisis de cette base.

L'ensemble des résultats de segmentation sont présentés dans le tableau suivant:

La différence	Le pas	La taille de bloc	Notre base	la base IFN /ENIT
50	5	5	54,97	48,93
		10	79,34	58,60
		15	75,70	62,11
		20	72,04	53,84
	10	5	56,40	47,67
		10	79,06	66,20
		15	81,31	65,25
		20	82,02	52,09
	15	5	39,29	25,59
		10	60,15	39,67
		15	68,91	47,13
		20	48,76	37,83
	20	5	24,54	17,53
		10	41,36	29,02
		15	53,02	32,05
		20	51,45	37,75

Tableau VI.2 Résumé des résultats de segmentation

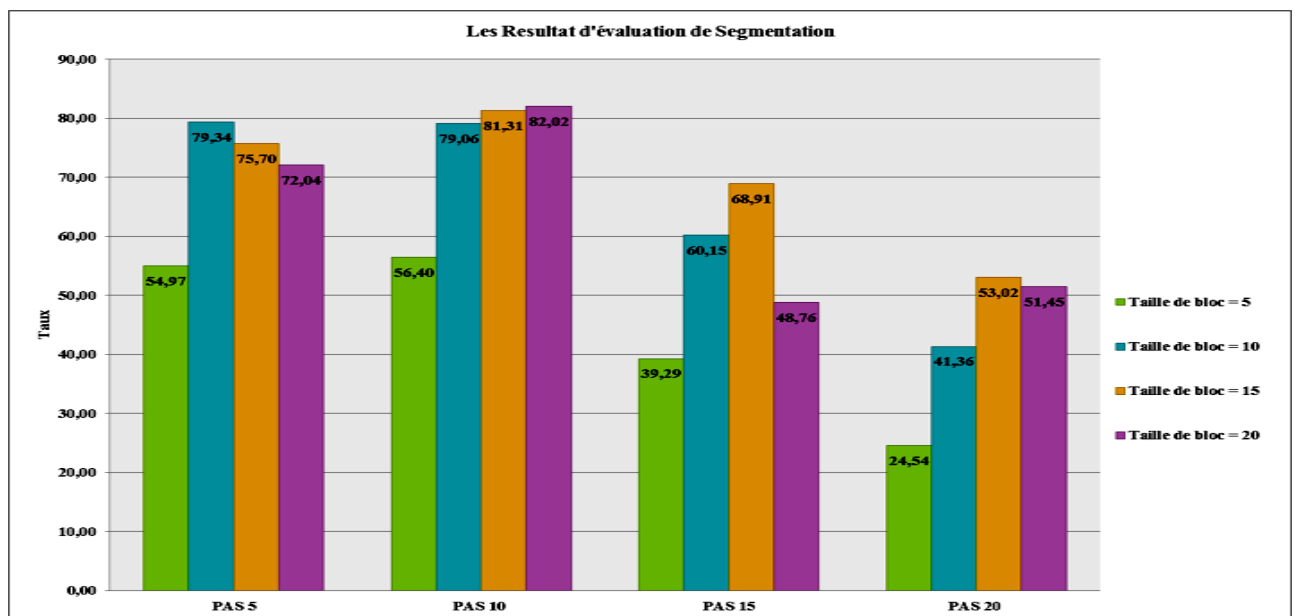


Figure IV.9 : Résumé des résultats de segmentation des mots de notre base.

La figure IV.9 représente les résultats d'évaluation de la méthode de segmentation de mots de base des wilayas en caractères en fonction de pas et de taille de bloc. On observe qu'on obtient le meilleur taux (82.02%) quand on prend les paramètres suivants:(pas=10 taille de bloc=20).la figure suivante résume le taux de segmentation globale pour les mots de notre base.

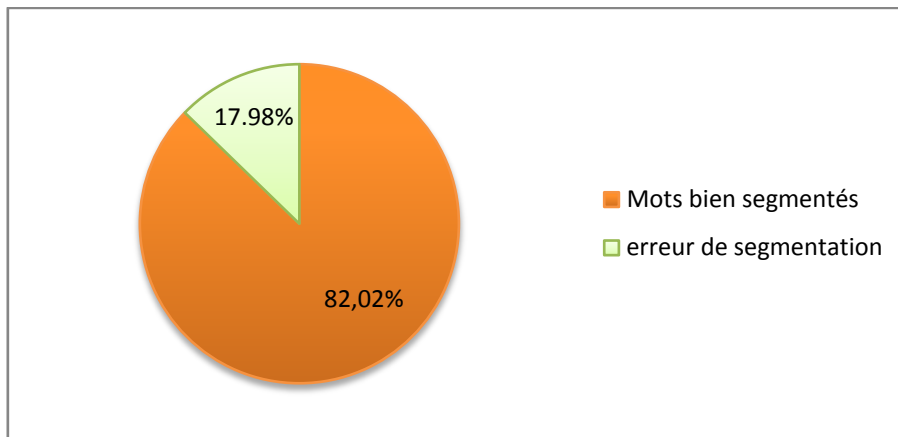


Figure IV.10 le taux de segmentation globale pour les mots de notre base.

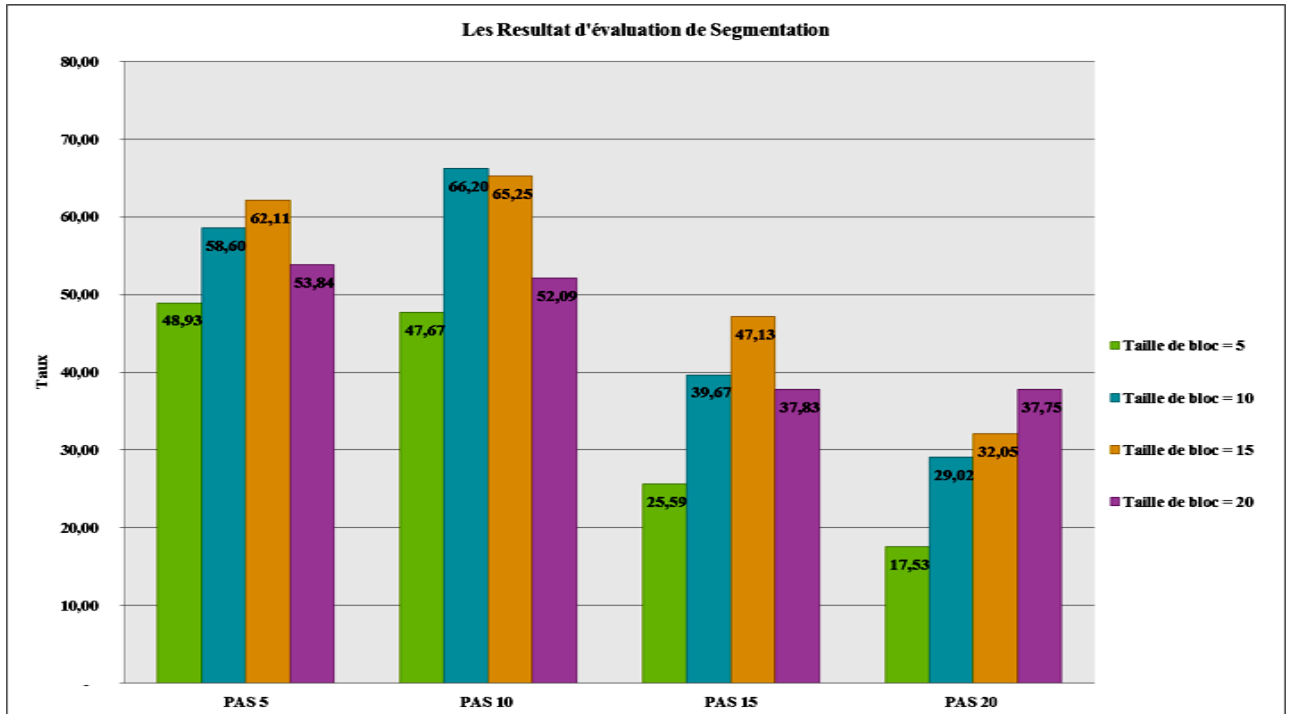


Figure IV.11 Résumé des résultats de segmentation des mots de la base IFN /ENIT.

La figure IV.11 représente les résultats d'évaluation de la méthode de segmentation pour la base IFN /ENIT en fonction de pas et de taille de bloc. On

observe qu'on obtient le meilleur taux (66.20%) quand on prend les paramètres suivants (pas=10 taille de bloc=10).la figure suivante résume le taux de segmentation globale pour les mots de la base IFN /ENIT.

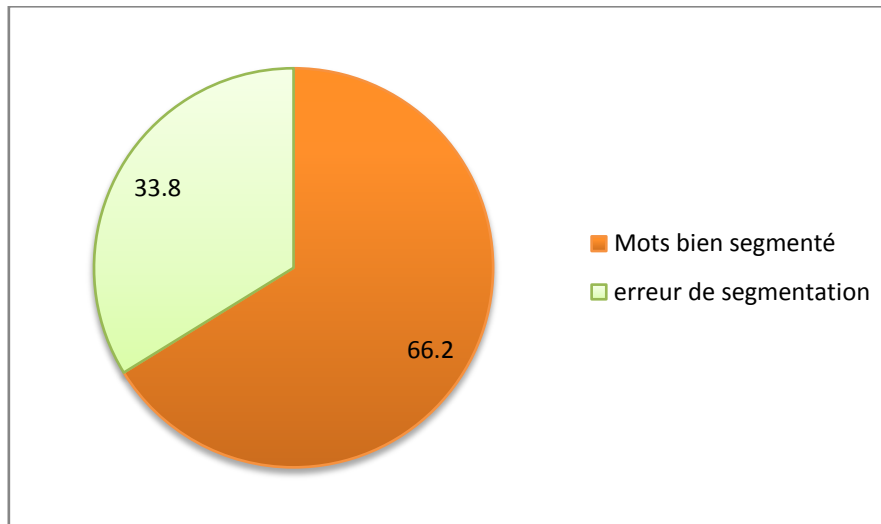


Figure IV.12 le taux de segmentation globale pour les mots de la base IFN /ENIT.

Et pour l'évaluation de segmentation pour les valeurs de différence(20, 40, 50, 70) on obtient les résultats montrés dans le tableau suivant.

La base	Le Pas	Taille de bloc	Différence	Taux %
Notre Base	10	20	20	55.62
			40	80.29
			50	82.02
			70	61.28
Standard base	10	10	20	43.08
			40	63.98
			50	66.20
			70	54.74

Tableau IV.3 Résultats de la segmentation pour les différences(20, 40, 50, 70).

Pour la phase d'apprentissage et de classification nous avons utilisé les réseaux de neurones, en particulier le perceptron multicouche (MLP), Chaque réseau est constitué de 49 neurones dans sa couche d'entrée, les valeurs d'entrée comprise entre 0 et 1 qui représente les valeurs de la matrice de distribution après la normalisation. Le nombre de neurone de la couche caché est un paramètre qui va être changé à chaque fois pour trouver la meilleure valeur. Et pour le nombre de neurones dans la couche sorti il est de 96 selon le nombre de classes .On obtient des résultats comme suite :

a) **Le cas des caractères isolés:** Pour évaluer les performances de notre système, nous avons traité 2800 caractères arabes manuscrits isolés tel que 2/3 des lettres pour l'apprentissage et 1/3 ont été utilisées pour le test. On obtient un taux de reconnaissance de (82%), pour un perceptron multi couches comprenant 28 neurones dans la couche cachées avec un nombre d'itération de 10000. Les résultats résumés dans la figure suivante.

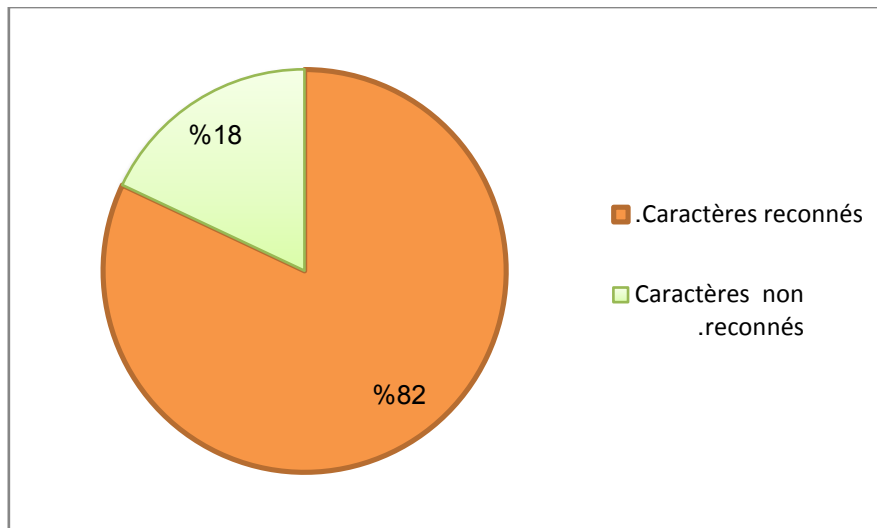


Figure IV.13 le taux de reconnaissance pour les caractères.

b) **Le cas des mots:**

Puisque les résultats de la reconnaissance par les réseaux de neurones ne sont pas fiables, on choisi d' intégrer les résultats de l'SVM, on obtient les résultats montrées dans le tableau suivant.

	Notre base	Base IFN /ENIT
Taux de reconnaissance	58.24%	30.50%

Tableau IV.4 Résumé de l'évaluation de la reconnaissance

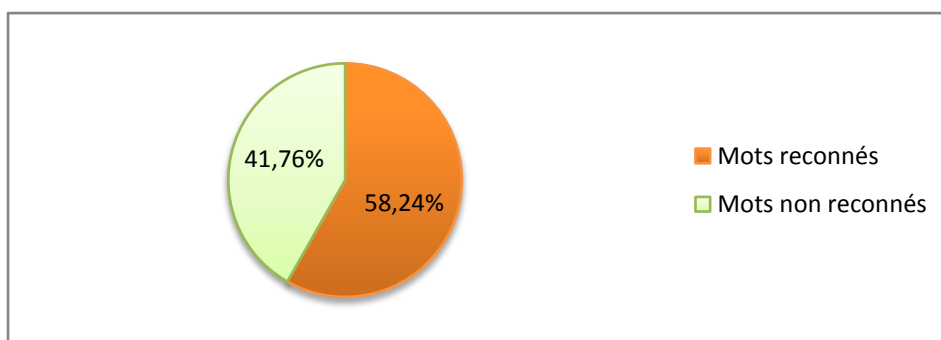


Figure IV.14 le taux de reconnaissance pour les mots de notre base.

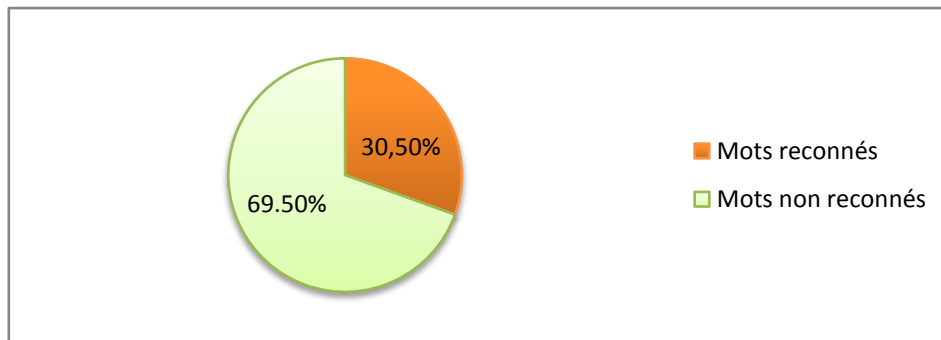


Figure IV.15 le taux de reconnaissance pour les mots de la base IFN /ENIT.

IV.6 Conclusion

D'après les résultats obtenus pendant tout le travail fait, on peut dire que l'objectif est atteint en grande partie pour la phase de segmentation et de l'extraction des primitives, mais pour la partie de reconnaissance les résultats sont insuffisants par rapport la littérature.

Pour l'amélioration de ce travail, nous proposons d'optimiser les techniques de segmentation, d'extraction et de reconnaissance en utilisant d'autres moyens, comme par exemple au niveau de la représentation des données, des transformations aptes à garder le maximum d'information des données.

Conclusion

générale

Conclusion générale

Conclusion générale

La reconnaissance de mots cursifs est une tâche très importante, utilisée dans plusieurs applications telles que la lecture d'adresses postales et la reconnaissance de montants littéraux. Actuellement, aucune solution optimale au problème de la reconnaissance de l'écriture cursive n'est encore connue. Bien que, des progrès importants dans ce domaine aient permis le développement de quelques systèmes pratiques; atteindre un taux de reconnaissance comparable aux performances humaines. Pour cette raison, ce domaine de recherche constitue encore un terrain fertile pour de futurs travaux.

Ce mémoire porte sur un système de réseaux de neurones pour la reconnaissance de l'écriture arabe manuscrite. Notre système a donné des bons résultats au niveau de la segmentation, cela montre l'efficacité de la méthode utilisée .

Nous espérons dans l'avenir pouvoir intégrer un groupe de recherche dans ce domaine pour pouvoir mettre l'épreuve notre algorithme en l'intégrant dans un système de reconnaissance de caractères et le voir contribuer à l'avancement des recherches dans le domaine de reconnaissance des caractères arabes.

Bibliographies

Bibliographies

- [1] CHERGUI Leila, "Combinaison de classifieurs pour la reconnaissance de mots arabes manuscrits", thèse de Doctorat: Informatique, l'Université de Mentouri - Constantine , p13, 2013.
- [2] F. Slimane, R. Ingold, A.M. Alimi, and J. Hennebert, "Duration models for Arabic text recognition using hidden Markov models", Proceedings of IEEE International Conference on Computational Intelligence for Modelling, Control and Automation, Vienne–Austria, IEEE, pp. 838-843, 2008.
- [3] I. Alkhoury, "Arabic handwritten word recognition based on Bernoulli mixture HMM", Thèse de Master, Université de Valencia, Spain, 2010.
- [4] Y. Kessentini."Modèles de Markov multi-flux pour la reconnaissance de l'écriture manuscrite multi-scripts", Thèse de Doctorat, Université de Rouen, France, 2009.
- [5] P. Ahmed and Y. Al-Ohali, "Arabic character recognition: progress and challenges", Journal of King Saudi university, vol. 12, pp. 85-116, 2000.
- [6] M. Cheriet, "Visual recognition of Arabic handwriting: challenges and new directions", Proceedings of the Conference on Arabic and Chinese Handwriting Recognition (SACH), Berlin-Germany, Springer, pp. 1-21, 2008.
- [7] S. Alma'adeed, "Recognition of off-line handwritten Arabic words", Thèse de Doctorat, Université de Nottingham, England, 2004.
- [8] F. Menasri, N. Vincent, E. Augustin, et M. Cheriet, "Un système de reconnaissance de mots arabes manuscrits hors-ligne sans signes diacritiques", Proceedings du Colloque International Francophone sur l'Écrit et le Document (CIFED), Rouen-France, pp. 121-126, 2008.
- [9] N. Ben Amara, "Utilisation des modèles de Markov cachés planaires en reconnaissance de l'écriture arabe imprimée". These de doctorat, specialite Genie Electrique, Universite des sciences, des Techniques et de medecine de Tunis II, 1999.
- [10] E.Lecolinet, O. Barrett, "Cursive word recognition : Methods and strategies ". In NATO/ASI , Fundamentals in handwriting recognition, Bonas, France June 21-july 3, 1993.

Bibliographies

- [11] B. Al-Badr , S.A. Mahmoud , " Survey and bibliography of Arabic optical text recognition ". Signal processing , vol. 41, pp. 49-77,1995.
- [12] P.M. Lallican, C. Viarp-Gaudin, S. Knerr , From off-line to on-line handwriting recognition ". Proc. 7th workshop on frontiers In handwriting recognition, pp. 303-312, Amsterdam 2000.
- [13] I.R. Tsang, "Pattern recognition and complex systems". Thèse de doctorat, université d'Anterwerpen, 2000.
- [14] M.M.M. Fahmy, S.Al Ali, "Automatic recognition of handwritten Arabic characters using their geometrical features ". Studies in informatics and control journal (SIC journal), vol. 10, No 2, 2001.
- [15] J. Anigbogu , " Reconnaissance de textes imprimés mutifontes à l'aide de modèles stochastiques et métriques ". thèse de doctorat, Université de Nancy I, 1992.
- [16] HAITAAMAR Schahrazed , "segmentation de textes en caracteres pour la reconnaissance optique de l'écriture arabe" ,thèse de Magister : informatique : universite el-hadj lakhdhar batna ,Algerie,2007.
- [17] R.G. Casey, E. Lecolinet, "Strategies in character segmentation : A survey ". IEEE. Proc. 3rd international conference on document Analysis and recognition (ICDAR'95), pp. 1028-1033, Montreal, Canada, 1995.
- [18] B. Al-Badr , R.M. Haralick, " Symbol recognition without prior segmentation ". Conference SPIE-EI 1994.
- [19] Y. Liu and S. Srihari, "Document image binarization on texture features, On Patter Analysis and Machine Intelligence ", vol. 19, n.5, pp. 540-544, May 1997.
- [20] H. Emptoz, F. Lebourgeois, V. Eglin, Y. Leydier. " La reconnaissance dans les images numérisées : OCR et transcription, reconnaissance des structures fonctionnelles et des métadonnées ", 2003.
- [21] P. Burrow , "Arabic handwriting recognition ", these de Master en informatique, université de Edinburg, England, 2004.

Bibliographies

- [22] T. Steinherz, E. Rivlin, N. Intrator, "Off-line cursive word recognition :a survey". International journal on document analysis and recognition, 2(2), pp. 90-110, 1999.
- [23] Dr Suici-Meslati, " Reconnaissance de formes et écriture arabe manuscrite ", support de cours 2008.
- [24] S. Kermi, "Classifieur neuronal base connaissances, application à la reconnaissance des caractères arabes isolés manuscrits ", Thèse de magister, université Badji Mokhtar, Annaba, Algérie 1999.
- [25] I.R. Tsang, "Pattern recognition and complex systems", Thèse de doctorat, université d'Anterwerpen, 2000.
- [26] L. Souici, Z. Zmirli, M. Sellami, "Système connexionniste pour la reconnaissance de l'arabe manuscrit", 1 ères journées scientifiques et techniques (JST FRANCIL), pp. 383-388, Avignon, France, 1997.
- [27] S. Snoussi maddouri, " Modèle perceptif neuronal à vision globale-locale pour la reconnaissance de mots manuscrits arabes ", 2002.
- [28] M. ZAIZ Faouzi, "Les Supports Vecteurs Machines (SVM) pour la reconnaissance des caractères manuscrits arabes", Thèse de magister, université Mohamed Khider – BISKRA, Algérie, 2010.
- [29] Farés Menasri, "Contributions `a la reconnaissance de l'écriture arabe manuscrite", thèse de doctorat en informatique, Université de Paris Descartes, juin 2008.
- [30] Zoubeir Mouelhi, AraSeg : un segmenteur semi-automatique des textes arabes In JADT 2008: 9es Journées internationales d'Analyse statistique des Données Textuelles.
- [31] Hanene BOUKERMA, "Combinaison de classifieurs flous pour la reconnaissance de l'écriture arabe manuscrite", thèse de magister en informatique, université 20 AOUT 1955 – SKIKDA, 2010.
- [32] Yousri Kessentini, Thierry Paquet et AbdelMajid Ben Hamadou, " Reconnaissance de l'écriture manuscrite arabe et latine par des modèles de Markov cachés multi-Flux ", Laboratoire LITIS EA 4108, université de Rouen, France décembre 2009.

Bibliographies

- [33] S. Haykin, "Neural Networks - A comprehensive Foundation", Macmillan College Publishing Company, New York, 1994.
- [34] Bernard GOSSELIN, "application de réseaux de neurones artificiels à la reconnaissance automatique de caractères manuscrits", thèse de doctorat, Faculté Polytechnique de Mon, Année académique 95-96.
- [35] KAMEL Mohamed, "reconnaissance de formes appliquée à l'écriture arabe manuscrite par des multiclassifieurs", thèse de Magister en informatique, Université Mohamed Khider, Biskra, 2010, p65.
- [36] Claude TOUZET, "les réseaux de neurones artificiels" introduction au connexionnisme, Juillet 1992.
- [37] Gérald PETITJEAN, "introduction aux réseaux de neurones", gerald.petitjean@eurodecision.com.
- [38] Marie Cottrell, "Les réseaux de neurones historiques, méthodes et applications", Université Paris 1- Sorbonne.
- [39] Aziez Lyes, Semeria Agostino, "Introduction à la reconnaissance de caractères par réseaux neuronaux", Coordinateur P. Collard 2002 / 2003.
- [40] Andrew D. Back, Neural Network Architectures with Time Delay Connections for Nonlinear Signal P.
- [41] A. Nigrin, Networks for Pattern Recognition. Cambridge: MA: The MIT Press, 1993.
- [42] A. Hamid, R.Haraty, "A neuro-heuristic approach for segmenting handwritten Arabic text". IEEE. Proc. Of International conference on computer systems and applications (ACS), pp. 110-113, 2001.
- [43] S.T. Masmoudi, N.E. Benamara, H. Amiri, " Segmentation stage of a PHMM-based model for off-line recognition of Arabic handwritten city names", IEEE. International conference on systems, Man and cybernetics, SMS 2002, vol. 4, 6-9 October 2002.
- [44] T. Sari, L. Souici, M. Sellami, " Off-line handwritten Arabic character segmentation algorithm : ACSA ". IEEE Proc. 8th international workshop on frontiers in handwriting recognition (IWFHR'02), 2002.

Bibliographies

- [45] IDIOU Ghania, " Régression et modélisation par les réseaux de neurones", thèse de magistère en Mathématique, UNIVERSITE MENTOURI - CONSTANTINE- ALGERIE, p.10, juin 2009.
- [46] S Haykin, " Neural Networks: A Comprehensive Foundation", Pearson Prentice Hall, India, 2005.
- [47] N. Kasapov, " Foundations of Neural Networks: Fuzzy Systems and Knowledge Engineering", The MIT Press, England, 1998.
- [48] P. Melin and O. Castillo, "Hybrid Intelligent Systems for Pattern Recognition using Soft Computing", An Evolutionary Approach for Neural Networks and Fuzzy Systems, Springer, Germany, 2005.
- [49] L.Rousseau, "Reconnaissance d'écriture manuscrite hors-ligne par reconstruction de l'ordre du tracé en vue de l'indexation de documents d'archives", Thèse de Doctorat, Institut National des Sciences Appliquées de Rennes, France, 2007.
- [50] MAMOUNI Elmamoun, " Sélection expérimentale de modèles SVM multi-classes application à la reconnaissance des caractères arabes", Université des Sciences et des Technologies d'Oran Mohamed BOUDIAF -Algérie, 10/2012.
- [51] Manuel Clergue, "Réseaux de Neurones Artificiels", support de cours.
- [52] Andrei Doncescu, "Les réseaux de neurones artificiels", support de cours.