



UNIVERSITE MOHAMED BOUDIAF - M'SILA
FACULTE DES MATHÉMATIQUES ET
DE L'INFORMATIQUE



DEPARTEMENT D'INFORMATIQUE

MEMOIRE de fin d'étude

Présenté pour l'obtention du diplôme de MASTER

Domaine : Mathématiques et Informatique

Filière : Informatique

Spécialité : Systèmes d'Informations Avancés

Par: Baza Halima Saadia

SUJET

La reconnaissance d'écriture arabe manuscrite

Soutenu publiquement le : / /2016 devant le jury composé de :

.....	Université de M'sila	Président
Kamel Mohamed	Université de M'sila	Rapporteur
.....	Université de M'sila	Examineur
.....	Université de M'sila	Examineur

Promotion : 2015 /2016

TABLE DES MATIERES

Introduction générale.....	6
----------------------------	---

Chapitre I : Reconnaissance de formes

1. Introduction:.....	8
2. Reconnaissance des formes :	8
2.1. La reconnaissance d'objet:.....	8
2.2. Similarité des formes et les taches de RF :	8
2.2.1. Tache de classification :.....	9
2.2.2. Taches de régression :	9
2.2.3. Taches de description :	10
2.3. Classes, formes et caracteristiques :	10
2.3.1. Classes :	10
2.3.2. Formes :	10
2.3.3. Caractéristiques:	11
2.4. Les approches du RF:.....	11
2.5. Projet du RF :	13
2.5.1. Taches du projet :	13
2.5.2. Apprentissage et test :	14
2.5.3. Les logiciels de la reconnaissance:	14
3. La reconnaissance d'écriture:.....	15
3.1. Les phases d'un système de reconnaissance d'écriture :	15
3.1.1. Acquisition d'image :	15
3.1.2. Prétraitements et normalisations :	16
3.1.3. Segmentation :	21
3.1.4. EXTRACTION DE CARACTERISTIQUES :	23
3.1.5. Post-traitements :	24
4. La reconnaissance de l'écriture arabe manuscrite :	24
4.1. Problématique :	24
4.1.1. Présentation de la langue arabe :	24

4.1.2. Difficultés inhérentes à la reconnaissance de l'écriture arabe	25
5. Conclusion :.....	30
<u>Chapitre II: Les systèmes de classification.</u>	
1. INTRODUCTION :.....	32
2. Définition formelle d'un classificateur :.....	32
2.1. Classificateur de type classe :.....	33
2.2. Classificateur de type rang :.....	33
2.3. Classificateur de type mesure :.....	33
3. Les réseaux de neurones :.....	34
3.1. Définition :.....	34
3.2. Neurone simple :.....	34
3.2.1. Le neurone de Mac Culloch-Pitts :.....	34
3.2.2. Le Perceptron :.....	35
3.3. Interprétation mathématique :.....	35
3.4. Types et architectures des RNA :.....	36
3.4.1. Le Perceptron Multicouches :.....	36
3.4.2. Fonctions à bases radiales :.....	38
3.5. L'apprentissage :.....	39
3.5.1. La loi d'apprentissage de Hebb [HEB49] :.....	39
3.5.2. La règle d'apprentissage du Perceptron :.....	40
3.5.3. Retro-propagation du gradient :.....	40
3.6. Avantages et inconvénients des réseaux neuronaux :.....	41
3.6.1. Avantages de l'approche neuronale :.....	41
3.6.2. Inconvénients de l'approche neuronale :.....	41
4. K plus proches voisins :.....	42
4.1. Formalisation mathématique :.....	42
4.2. ALGORITHME :.....	43
5. Modèles de markov caches :.....	44
5.1. Définitions :.....	44
5.1.1. Chaîne de Markov :.....	44
5.1.2. Chaîne stationnaire :.....	44
5.2. Un modèle de markov cache :.....	45

5.3.	<i>Interprétation mathématique :</i>	45
5.4.	<i>Les types des HMMs :</i>	46
5.4.1.	<i>HMM ergodique :</i>	46
5.4.2.	<i>HMM gauche-droite :</i>	46
6.	<i>Conclusion :</i>	47

Chapitre III : Conception et résultats.

1.	<i>Introduction :</i>	49
2.	<i>Conception de la solution proposée :</i>	49
2.1.	<i>Prétraitement :</i>	49
2.1.1.	<i>Binarisation :</i>	50
2.1.2.	<i>Squelettisation :</i>	50
2.1.3.	<i>Centrage :</i>	51
2.1.4.	<i>Extraction des caractéristiques :</i>	51
2.2.	<i>Apprentissage et classification :</i>	52
3.	<i>Base des caractères et outils de développement :</i>	52
3.1.	<i>Base des Caractères :</i>	52
3.2.	<i>Mathworks MATLAB :</i>	53
4.	<i>Présentation de l'application :</i>	54
4.1.	<i>Les étapes de reconnaissance d'un caractère :</i>	54
4.1.1.	<i>Chargement de l'image à partir de la base des caractères :</i>	54
4.1.2.	<i>La binarisation :</i>	55
4.1.3.	<i>Squelettisation :</i>	55
4.1.4.	<i>Centrage :</i>	55
4.1.5.	<i>Extraction des caractéristiques :</i>	56
4.1.6.	<i>Reconnaissance :</i>	56
5.	<i>Test et validation :</i>	56
5.1.	<i>Classification :</i>	56
5.1.1.	<i>Le premier RNA :</i>	57
5.1.2.	<i>Le deuxième RNA :</i>	57
5.1.3.	<i>Le troisième RNA :</i>	58
5.1.4.	<i>Le quatrième RNA :</i>	59
5.1.5.	<i>Le cinquième RNA :</i>	60

6. Combinaison :	61
7. Conclusion :	62
Conclusion et perspectives	61
Bibliographie	63
Liste des figures	64
Liste des tableaux	65

INTRODUCTION GENERALE

Les techniques liées au traitement de l'information connaissent actuellement un développement très actif en liaison avec l'informatique et présentent un potentiel de plus en plus important dans le domaine de l'interface homme machine. L'écriture restera l'un des grands fondements des civilisations et le mode par excellence de conservation et de transmission du savoir. La reconnaissance de l'écriture est un domaine vaste qui constitue un sous ensemble des systèmes de la reconnaissance des formes (RF). Ces systèmes sont la première étape d'un processus de compréhension de notre univers dans le cadre global de la communication homme machine. La reconnaissance de l'écriture manuscrite ou imprimée reste encore un sujet de recherche et d'expérimentation, le problème n'est pas encore entièrement résolu bien que l'on sache atteindre des taux assez élevés dans certaines applications pour lesquelles soit le vocabulaire est limité, soit la fonte est unique ou en nombre restreint.

Il existe cependant plusieurs domaines pour lesquels la reconnaissance de l'écriture est appliquée avec un certain succès : 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 est beaucoup plus complexe que celle de l'écriture imprimée due à son extrême variabilité: variabilité des formes, des espacements entre mots et caractères, fluctuation des lignes.

Les travaux de recherche en reconnaissance de l'écriture arabe manuscrite, bien que moins avancés que pour d'autres langues, deviennent plus intenses qu'avant.

Après cette introduction, dans le premier chapitre, nous présenterons rapidement les notions de bases de la reconnaissance des formes, la reconnaissance de l'écriture et la reconnaissance de l'écriture arabe manuscrite.

Dans le deuxième chapitre, nous présenterons un ensemble de classificateurs couramment utilisés en reconnaissance de l'écriture arabe.

Dans le troisième chapitre nous présenterons notre travail sur la reconnaissance des caractères arabes manuscrits, L'originalité de notre travail consiste à construire un classificateur neurone composé de cinq réseaux de neurones, avant d'utiliser une stratégie pour les combiner.

**Chapitre I:
Reconnaissance
des formes**

1. Introduction:

La reconnaissance de formes comporte actuellement un champ très large des méthodes soutenant le développement de nombreuses applications dans beaucoup de différents secteurs d'activités. Les méthodes et techniques de reconnaissance de formes se trouvent généralement dans le milieu de la simulation « intelligente » des tâches, qui a certainement infiltré notre vie quotidienne [MAR01]. La fabrication aidée par les robots, les systèmes de diagnostic médical, la reconnaissance automatique des documents ou caractères (OCR), la prévision des variables économiques, l'exploration des ressources terrestres, et l'analyse des données satellitaires sont juste quelques exemples des champs d'activités où cette tendance s'applique.

Alors dans ce chapitre nous allons présenter l'ensemble des généralités sur le domaine de reconnaissance de formes. Pour commencer, nous introduisons la reconnaissance de formes puis la reconnaissance de l'écriture en générale, en fin la reconnaissance de l'écriture arabe.

2. Reconnaissance de formes :

2.1. La reconnaissance d'objet:

La reconnaissance des objets est une tâche effectuée quotidiennement par les êtres vivants à leur aptitude et à la nécessité de traiter l'environnement. Il est interprété dans la plupart des circonstances (la navigation vers les sources de nourriture, de la migration, l'identification des prédateurs, la détermination des camarades, etc.) avec une efficacité remarquable. La reconnaissance des objets est considérée ici dans un sens large cognitif et peut consister en une tâche très simple.

2.2. Similarité de formes et les taches de RF :

Selon [MAR01], la similarité est une notion fondamentale dans la reconnaissance de formes, indépendamment de quelle approche nous allons suivre. Nous reconnaissons que deux objets sont similaires s'ils ont des valeurs des attributs communs similaires. Souvent, la similarité est liée à un sens plus abstrait, non pas entre des objets, mais entre un objet et un concept cible.

L'appréciation sur la similarité des formes est fortement liée à la tâche de reconnaissance de formes proposée comme décrit dans ce qui suit :

2.2.1. Tache de classification :

[CRO02]La classification est un processus d'association d'une classe à une observation

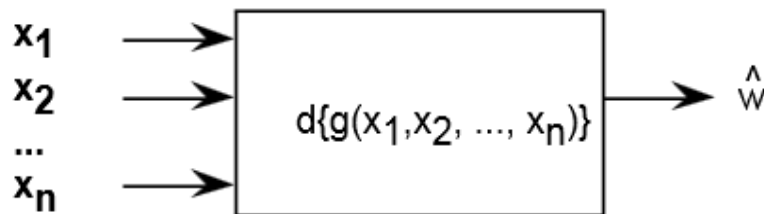


Figure I.1 : Tache de classification

Pour un vecteur de caractéristique il sort une estimation de la classe, w Les techniques de reconnaissance de formes statistiques fournissent une méthode Pour induire des tests d'appartenance à partir d'un ensemble d'échantillons.

La classification se résume à une division de l'espace de caractéristique en partition disjoint. Cette division peut-être fait par estimation de fonctions paramétrique ou par une liste exhaustive des frontières.

Le critère de division est la probabilité de commettre une erreur.

Cette probabilité est fournie par la règle de Bayes

$$p(\text{Classe}_k | \vec{\mathbf{X}}) = \frac{p(\vec{\mathbf{X}} | \text{Classe}_k) p(\text{Classe}_k)}{p(\vec{\mathbf{X}})}$$

2.2.2. Taches de régression :

Un système qui produit la solution de régression, généralement, faire des prévisions, un peu dévié de la valeur vraie (courbe). La différence entre la valeur prédite et la valeur réelle constitue une erreur de prédiction. Et l'objectif ici est une solution avec des valeurs prédites similaires à la cible, c'est-à-dire, avec des petites erreurs.

Parfois les tâches de régression sont également effectuées dans le cadre d'une classification.

2.2.3. Taches de description :

Dans cette tâche de description la similarité des objets (transitions, accélérations, décélérations, etc., dans cet exemple) est évaluée au moyen d'une règle structurelle. Deux objets sont similaires s'ils se conforment à la même règle. Par conséquent toutes les transitions sont semblables, toutes les accélérations sont semblables.

La situation peut être résumée comme suite :

- Il y a un ensemble de classes (états) dans lequel peut être trouvé une certaine entité étudiée.
- un certain ensemble de formes (représentations) (signaux, images, etc.), correspondant à chaque classe (état).
- À partir de chaque forme nous pouvons extraire l'information qui la caractérise, les caractéristiques.

2.3. Classes, formes et caractéristiques :

2.3.1. Classes :

Les classes sont les états de « nature » ou des catégories d'objets associés à des concepts ou prototypes. Dans ce qui suit, nous supposons classes dénotés $\in \Omega$, ($= 1, \dots$), où Ω est l'ensemble de toutes les classes, connu sous le nom espace d'interprétations. L'espace d'interprétation a des propriétés des concepts telles que les unions, les intersections et les arbres hiérarchiques de classes.

2.3.2. Formes :

Les formes sont des représentations « physiques » des objets. Habituellement signaux, images ou tables simples des valeurs. Souvent on réfère aux formes avec objets, cas ou échantillons.

Dans ce qui suit nous allons utiliser la lettre n pour indiquer tout le nombre de formes disponibles afin de concevoir un système de RF, on appelle cette ensemble des formes l'ensemble d'apprentissage.

2.3.3. Caractéristiques:

Les caractéristiques sont des mesures, des attributs ou des primitifs dérivés des formes, qui peuvent être utiles pour leur caractérisation.

Une forme est représentée par un vecteur de caractéristiques:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \in X \subset R^d$$

Figure I.2 : vecteur de caractéristique

Où X est le domaine à d dimensions des vecteurs de caractéristiques.

2.4. Les approches du RF:

Il y a une multiplicité d'approches de RF et aucun consensus clair sur comment les catégoriser. L'objectif d'un système de RF est d'accomplir une correspondance entre l'espace de représentation et l'espace d'interprétation. Une telle correspondance, qui est une classification, une régression ou une description, s'appelle également une hypothèse.

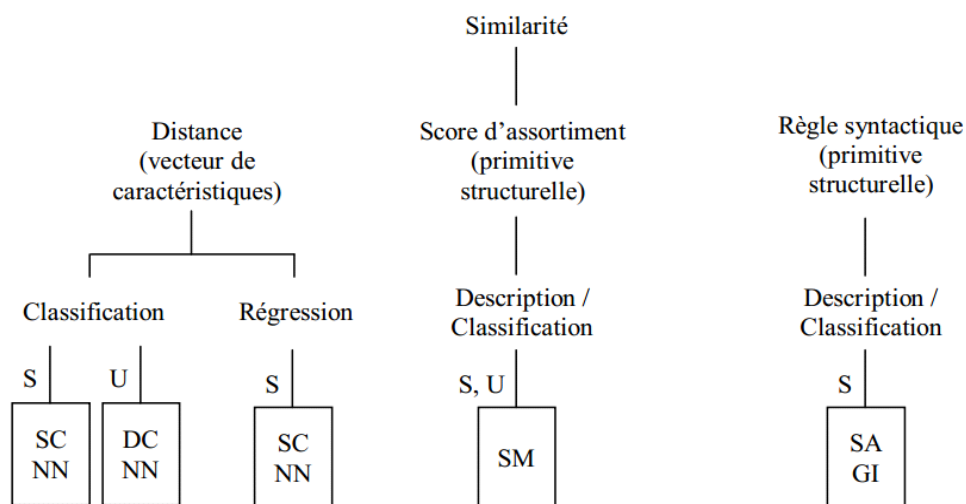


Figure I.3: Les approches de RF - S: supervisé, U: non supervisé, SC: classification statistique, NN: réseaux neuronaux, DC: groupement de données, SM : classification structurale, SA: analyse syntaxique, GI: inférence grammaticale [MAR01].

Il y a deux manières distinctes que telles hypothèses peuvent être obtenues :

- Supervisée :

Guidée par les concepts (concept-driven) ou hypothèses inductives : trouver dans l'espace de représentation une hypothèse correspondante à la structure de l'espace d'interprétation.

- Non supervisée :

Guidée par les données (data -driven) ou hypothèses déductives : trouver une structure dans l'espace d'interprétation correspondant à la structure dans l'espace de représentation.

L'approche non supervisée tente de trouver une hypothèse utile basée seulement sur les relations de similarité dans l'espace de représentation.

L'hypothèse est dérivée par l'utilisation des méthodes d'apprentissage qui peuvent être statistiques, approximative (minimisation d'erreur) ou structurale.

Tenant compte de comment l'hypothèse est dérivée et la similarité de forme est mesurée, nous pouvons établir la catégorisation hiérarchique montrée dans Figure I.1 [MAR01].

2.5. Projet du RF :

2.5.1. Taches du projet :

Les systèmes de reconnaissance de formes, indépendamment de l'approche suivie pour les concevoir, ont des unités facultatives spécifiques comme montré dans la Figure I.2. Quelques systèmes n'ont pas de l'unité de post-traitement et/ou de prétraitement.

Les unités d'un système de reconnaissance de formes et les tâches du projet correspondantes sont :

1. Acquisition de formes, qui peut prendre plusieurs formes : acquisition de signal ou d'image, collecte de données.

2. Prétraitement et Segmentation, en cas d'acquisition d'image, généralement l'image n'est pas prête à l'extraction des caractéristiques et des étapes de prétraitement et segmentation sont nécessaires, dans le prétraitement on parle de seuillage, de squelettisation, etc.

3. Extraction des caractéristiques sous forme de mesures, extraction des primitifs, etc.

4. Prétraitement des caractéristiques, dans certains cas les valeurs de caractéristiques ne sont pas directement introduites dans le classificateur ou le descripteur. Par exemple dans les réseaux neuronaux il est habituel de normaliser les caractéristiques d'une certaine manière.

5. L'unité de classification, régression ou de description : est l'unité noyau du système de reconnaissance de formes.

6. Post-traitement : Parfois la sortie obtenue à partir de l'unité noyau de reconnaissance de formes ne peut pas être directement utilisée. Il peut être qu'elle a besoin, par exemple, d'une certaine opération de décodage. Ceci s'appelle un post-traitement.

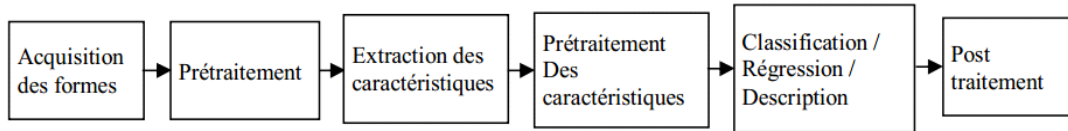


Figure I.4:Un système de reconnaissance de formes avec ses unités.

Quelques systèmes n'ont pas de l'unité de prétraitement et/ou de post-traitement.

2.5.2. Apprentissage et test :

Comme mentionné dans la section précédente le développement d'une application de reconnaissance de formes commence par l'évaluation du type de caractéristiques à utiliser et de adéquate de reconnaissance de formes pour le problème à résolu. À ce but un ensemble initial de formes est habituellement disponible. Dans les approches supervisées cet initial représenté par n vecteurs d -dimensionnels de caractéristiques ou n chaînes établis avec d primitifs, est utilisé pour développer le noyau de reconnaissance de formes. Il constitue l'ensemble d'apprentissage.

La performance d'un système de reconnaissance de formes est habituellement évaluée en termes de taux d'erreur pour chaque classe et un taux d'erreur globale. Quand cette évaluation de performance est basée sur les formes de l'ensemble d'apprentissage nous obtenons, en moyenne, des vues optimistes. Afin d'obtenir de meilleures évaluations d'une performance de système de reconnaissance de formes il est indispensable pour l'évaluer en utilisant un indépendant de formes, c.-à-d., formes non utilisées dans son conception ou son apprentissage. Cet ensemble indépendant de formes s'appelle un ensemble de test. Les évaluations de performance de système de reconnaissance de formes à l'aide d'ensemble de test nous donnent une idée d'à quel point le système est capable de généraliser ses capacités de reconnaissance sur des nouvelles formes.

Pour des systèmes de classification et de régression le degré de confiance qu'on peut avoir sur des évaluations de performance de système de reconnaissance de formes, aussi bien sur ses possibilités de généralisation, dépend fortement du rapport n/d , le rapport de dimensionnalité.

2.5.3. Les logiciels de la reconnaissance:

Il y a beaucoup de produits pour développer les applications de reconnaissance de formes, qui peuvent guider la conception d'un système de reconnaissance de formes des les

premier étapes jusqu'à l'évaluation finale. Une seule recherche par l'internet révélera plusieurs de ces produits et outils, gratuits ou payants. Beaucoup de ces produits sont spécialisés, par exemple dans le secteur de réseaux neuronaux. D'une façon générale, les types suivants de logiciels peuvent être trouvés [MAR01]:

- Bibliothèques d'outil (par exemple dans C).
- Outils fonctionnant sous d'autres logiciels (par exemple Microsoft Excel ou le *Matlab de Math Works*).

- Produits pour la conception des applications de RF en utilisant une méthode spécifique.

- Produits pour la conception des applications de RF en utilisant une panoplie de différentes méthodes.

3. La reconnaissance d'écriture:

3.1. Les phases d'un système de reconnaissance d'écriture :

Un système de reconnaissance fait appel généralement aux étapes suivantes: Acquisition, prétraitements, segmentation, extraction des caractéristiques, classification, suivis éventuellement d'une phase de post-traitement.

3.1.1. Acquisition d'image :

L'acquisition hors ligne et l'acquisition en ligne ce sont deux modes différents d'un système de reconnaissance des mots, ayant chacun ses outils propres d'acquisition et ses algorithmes correspondants de reconnaissance.

- Acquisition hors ligne :

Dans le cas hors ligne, il s'agit de reconnaître des textes manuscrits à partir de documents écrits au préalable. L'image du texte écrit est numérisée à l'aide d'un scanner,

les informations recueillies se présentent sous la forme d'une image discrète constituée d'un ensemble de pixels Figure I.3.

- Acquisition en ligne :

Dans le cas en ligne, il s'agit de reconnaître l'écriture au fur et à mesure de son tracé. Le texte est saisi avec un stylo et une tablette à numériser, les informations recueillies sont constituées par une suite ordonnée de points (définis par leurs coordonnées) échantillonnés à cadence fixe Figure I.3. La reconnaissance hors ligne ne peut pas a priori s'appuyer sur l'information temporelle du tracé qui est perdue, mais elle peut tenir compte de l'épaisseur du tracé (les pleins et les déliés).

La reconnaissance en ligne peut disposer de l'information temporelle (vitesse, accélération, levés de stylo, retours en arrière, barres de t, points diacritiques), mais d'aucune information sur l'épaisseur du tracé si on ne dispose pas d'un signal de pression de la pointe du stylet sur le support.

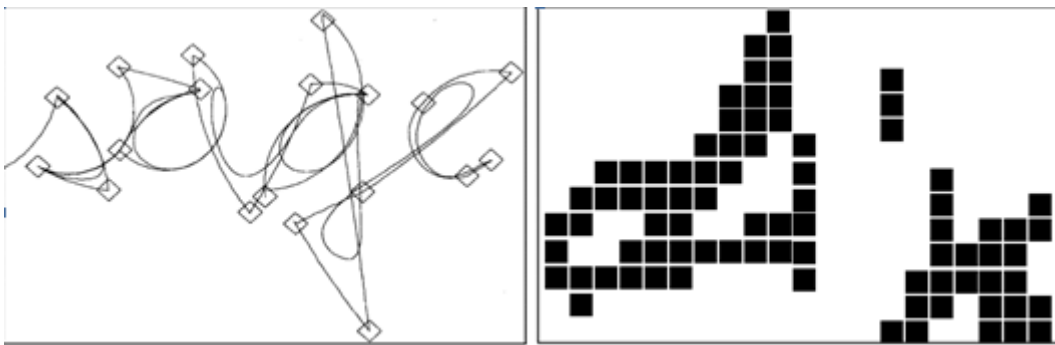


Figure I.5: Différents modes de capture de mots tracés du mot à gauche "sage" et image du mot "dix" à droite.

3.1.2. Prétraitements et normalisations :

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.

3.1.2.1. Seuillage :

On distingue en général deux approches : les approches à seuillage global, et les approches à seuillage adaptatif [MEN08].

Pour une évaluation de différentes méthodes de binarisation, on pourra consulter le travail [TRI95].

- Seuillage global :

Le seuillage global consiste à prendre un seuil ajustable, mais identique pour toute l'image. Chaque pixel de l'image est comparé à ce seuil et prend la valeur blanc ou noir selon qu'il est supérieur ou inférieur. Cette classification ne dépend alors que du niveau de gris du pixel considéré.

Cette méthode convient pour les documents simples et de bonne qualité. Néanmoins, elle n'est plus applicable lorsque la qualité d'impression du texte n'est pas constante dans toute la page, des caractères peuvent être partiellement perdus. Des problèmes surviendront également si le fond est bruité ou non homogène, dans ce cas des taches parasites peuvent apparaître.

Les histogrammes de la répartition des niveaux de gris des pixels de ces images contiennent deux pics nets : l'un pour la forme, l'autre pour le fond. Un seuil global peut aisément être sélectionné entre les deux pour classer les pixels selon qu'ils appartiennent aux formes ou au fond.

- Seuillage adaptatif :

Dans les documents pour lesquels l'intensité du fond et l'intensité de la forme peuvent varier au sein du document, un seuillage global est inadapté. Il devient nécessaire de choisir le seuil de binarisation de manière locale. On calcule un seuil de binarisation pour chaque pixel de l'image, en fonction de son voisinage.

3.1.2.2. Squelettisation :

La squelettisation est une opération qui permet de passer d'une image à sa représentation en "fil de fer". Le squelette a un pixel d'épaisseur. C'est une manière de représenter l'information indépendamment de l'épaisseur initiale de l'écriture.

Il permet d'extraire des caractéristiques importantes, comme les intersections et le nombre de tracés, leurs positions relatives. Il est également possible de renormaliser l'épaisseur de l'écriture à partir du squelette.

Il n'existe pas de définition unique du squelette. Le squelette doit seulement remplir trois conditions:

- Il doit être aussi fin que possible (typiquement, 1 pixel d'épaisseur).
- Il doit respecter la connexité.
- Il doit être centré dans la forme qu'il représente.

3.1.2.3. Normalisations :

Les normalisations ont pour objectif de rendre l'écriture la plus indépendante possible du scripteur.

Nous présenterons ici trois techniques de normalisation :

- Correction de l'inclinaison des lignes :

La correction de l'inclinaison des lignes de texte (également appelée correction de "skew"), consiste à redresser horizontalement les lignes d'écriture obliques.

Plusieurs méthodes sont disponibles. Les deux plus populaires sont la transformée de Hough (appliquée sur les centres de gravité des composantes connexes), et les histogrammes de projection.

La méthode de l'histogramme consiste à parcourir l'image selon des directions proches de l'horizontale, et à compter le nombre de pixels noirs selon ces directions pour chaque ligne [MEN08].

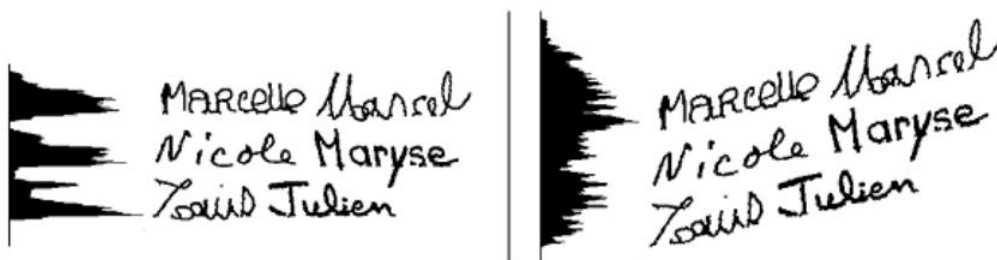


Figure I.6: Correction de l'inclinaison des lignes à l'aide d'histogrammes de projection horizontale. Les pics les plus prononcés correspondent à la rotation correcte

L'angle du document (ou de l a ligne), θ , est celui qui correspond à l'histogramme d'entropie maximale.

Pour corriger cette inclinaison, il suffit d'appliquer une rotation de l'image d'angle θ

[MEN08]:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

- Correction de l'inclinaison des lettres :

Certains scripteurs écrivent leurs lettres de façon inclinée par rapport à l'axe vertical. Cette inclinaison de l'écriture est également appelée «slant ». Les lettres peuvent être inclinées vers la droite ou vers la gauche. Pour la même raison que dans le paragraphe précédent, il convient de corriger cette inclinaison de l'écriture pour la rendre la plus indépendante possible des spécificités d'écriture du scripteur.

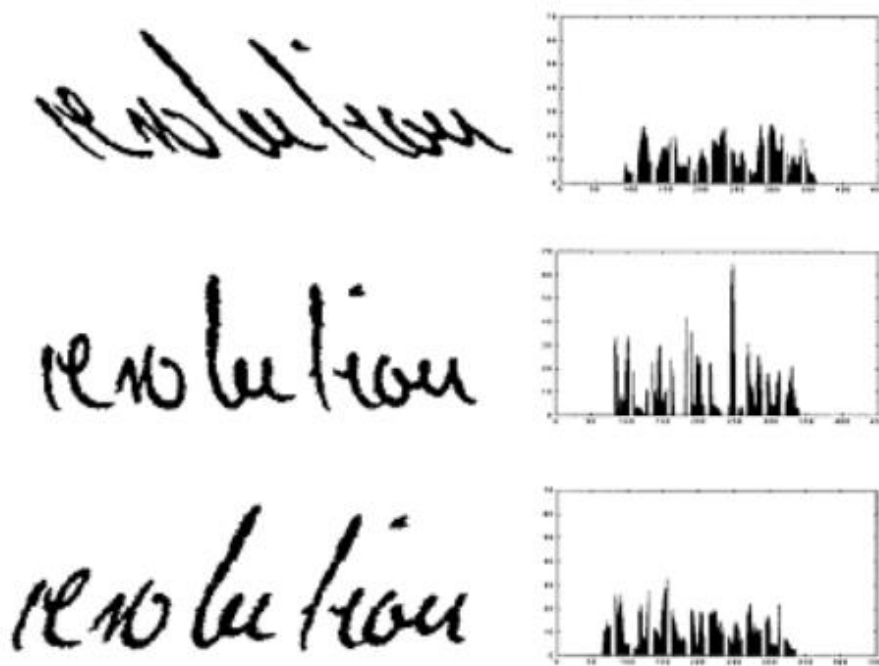


Figure I.7:Inclinaison de l'écriture, et histogrammes de projection verticale correspondants

- Lissage du contour :

Dans certains cas, la numérisation, binarisation ou les prétraitements peuvent introduire des bruits dans l'image, qui se traduisent en particulier par la présence d'irrégularités le long des contours des lettres. Ces bruits peuvent dégrader les performances de reconnaissance.

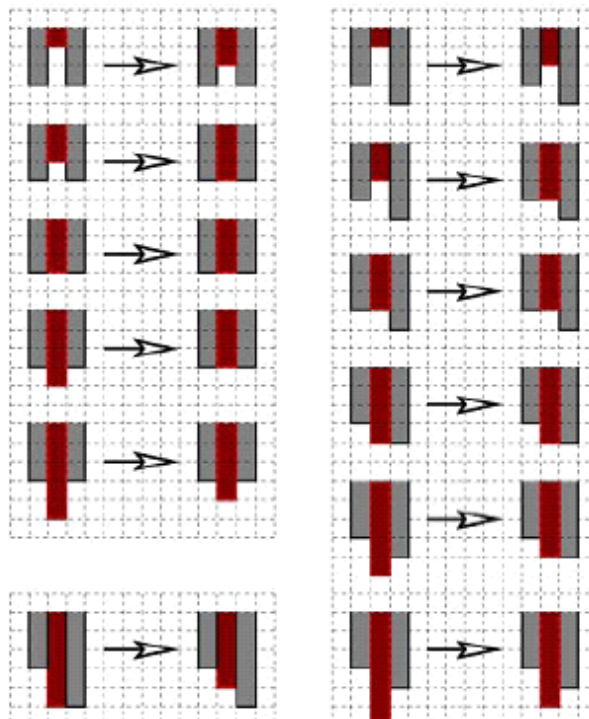


Figure I.8: Masques qui permettent de lisser le contour des formes. Les masques complémentaires S'obtiennent par rotation [MEN08].

- La ligne de base :

Les lignes de base encadrent la bande des minuscules et délimitent les zones contenant les ascendants et descendants. Ces lignes sont importantes en reconnaissance de l'écriture. Cette information est utilisée pour détecter ascendants et descendants, et également pour normaliser les primitives, les rendant ainsi moins dépendantes de la hauteur de l'écriture.

3.1.3. Segmentation :

Dans cette phase les différentes parties logiques d'une image sont extraites. A partir d'une image acquise il y'a d'abord séparation des blocs de texte et des blocs graphiques, puis à partir d'un bloc de texte il y'a extraction des lignes, ensuite à partir de ces lignes sont extraits les mots puis les caractères ou graphèmes [BAD95].

3.1.3.1. 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 [BEL97].

3.1.3.2. 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 composants les mots. La plupart des études proposées dans ce domaine s'appuient sur une décomposition de l'image en composantes connexes [BEN99b].

D'autres par contre utilisent des techniques s'appuyant en grande partie sur les histogrammes de la projection horizontale [BAD95]. Et certains auteurs optent pour des méthodes spécialisées telle que celle utilisée par Bennisri et Al dans [BEN99b] pour la segmentation en lignes de l'écriture arabe manuscrite.

3.1.3.3. Segmentation des 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 des squelettes ou la détermination des composantes connexes... .

3.1.3.4. Segmentation des mots

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 [CAS96].

3.1.4. EXTRACTION DE CARACTERISTIQUES :

Dans [OLI02], J. J. Oliveira et al présentent un certain nombre de primitives couramment utilisées en reconnaissance de l'écriture. Ils distinguent trois catégories:

- Les primitives perceptuelles
- Les primitives directionnelles
- Les primitives topologiques

Les auteurs indiquent que les primitives perceptuelles sont celles qui permettent d'obtenir les meilleurs résultats, et leur combinaison avec les autres est utile pour améliorer la reconnaissance.

La ligne de base est utile pour extraire et pour normaliser ces caractéristiques.

3.1.4.1. Les primitives perceptuelles :

Dans la liste des primitives perceptuelles usuelles, on retrouve : la position des ascendants et descendants, leur hauteur, la position et la taille des boucles, des primitives extraites sur l'enveloppe convexe du mot, le nombre de transitions noir/blanc dans la zone centrale,

3.1.4.2. Les primitives directionnelles :

Ces primitives sont généralement basées sur le squelette, ces types de primitives donnent la direction générale du tracé, longueurs et angles.

3.1.4.3. Les primitives topologiques :

Ce type de caractéristiques est basé sur des densités de pixels. On peut par exemple projeter des images de tailles différentes (les graphèmes) dans une matrice de taille fixe. Les caractéristiques extraites sont les valeurs des cellules de cette matrice.

Dans ce type de primitives, on compte également les profils et histogrammes. Pour maintenir un vecteur de taille fixe, on divise l'image en un nombre fixe de bandes horizontales et verticales. Les caractéristiques sont les moyennes des valeurs sur ces bandes.

3.1.4.4. Reconnaissance :

Ou classification, C'est l'étape principale. On l'a abordé dans le chapitre précédant (chapitre I) où nous avons présenté un ensemble de classificateurs couramment utilisés en reconnaissance de l'écriture.

3.1.5. Post-traitements :

Des vérifications contextuelles classiques telles que la recherche dans un dictionnaire, les probabilités d'occurrence de bi-gramme et de trigramme..., sont appliquées dans les différents travaux qui prévoient un post-traitement.

La méthode du dictionnaire est traditionnellement simplifiée pour accélérer la recherche et réduire la complexité du calcul : le dictionnaire est construit à partir de mots réduit à leurs racines, les suffixes et les préfixes sont éliminés. Cependant des modèles sont élaborés afin de spécifier la relation racine-suffixe-préfixe [BEN99a].

4. La reconnaissance de l'écriture arabe manuscrite :

4.1. Problématique :

4.1.1. Présentation de la langue arabe :

L'arabe littéral, arabe moderne unifié ou encore classique est le nom que l'on donne à une variante de la langue arabe, utilisée comme langue officielle dans tous les pays arabes, et comme langue commune entre pays arabes. Elle est également employée dans la plupart

des écrits et, à l'oral, dans les situations officielles ou formelles (discours religieux, politiques, journaux télévisés).

L'arabe littéral se distingue ainsi de l'arabe dialectal, qui est la langue vernaculaire parlée au quotidien et ce depuis l'expansion de l'islam. Cette variété de la langue recouvre plusieurs dialectes locaux pouvant varier assez fortement d'un pays à l'autre. Dans tous les pays arabes, un dialecte national composé par plusieurs dialectes locaux est parlé. Aucun d'entre ces dialectes n'est identique complètement à l'arabe classique ou littéraire.

Notons également que le farsi (persan), utilisé principalement en Iran et en Afghanistan, partage un grand nombre de points communs avec l'écriture arabe. Cet aspect est hors du cadre de cette thèse : nous nous restreignons ici à l'étude de l'écriture arabe manuscrite sur des données en provenance de scripteurs tunisiens. Sauf mention contraire, les données utilisées proviennent de la base IFN/ENIT [PEC02], une base de noms de villes tunisiennes.

4.1.2. Difficultés inhérentes à la reconnaissance de l'écriture arabe

- Alphabet :

L'alphabet arabe comporte 28 lettres (voir Tableau I.1) La forme des lettres dépend de leur position dans le mot. Certaines lettres prennent jusqu'à 4 formes différentes : par exemple (ع ع ح ح) ou (ه ه ه ه).

Mais pour la plupart des lettres, les formes début/milieu et fin/isolé sont identiques à la ligature près. La présence d'une ligature avec la lettre précédente ou avec la lettre suivante ne modifie pas la forme de la lettre de manière significative (pas plus que dans l'écriture manuscrite cursive latine).

En arabe, les ligatures se situent toujours au niveau de la ligne d'écriture, c'est-à-dire qu'il n'existe pas de lettre à liaison haute comme le 'o' ou le 'v' en alphabet latin, il existe toutefois des ligatures verticales.

- Signes diacritiques :

Le terme « signe diacritique » peut porter à confusion : dans certains travaux, seules les voyelles arabes sont appelées diacritiques. Dans d'autres travaux, en revanche, tous les signes secondaires sont appelés diacritiques, qu'il s'agisse des voyelles, des points ou des autres signes (chaddah, hamzah, ...).

C'est cette deuxième terminologie que nous employons ici : un signe diacritique est une composante secondaire d'une lettre, qui vient la compléter ou en modifier le sens. Dans la suite de cette thèse, les « signes diacritiques » désigneront à la fois points, voyelles et autres signes secondaires.

- Des points nécessaires pour différencier les lettres

Dans l'alphabet arabe, 15 lettres parmi les 28 possèdent un ou plusieurs points. Ces signes diacritiques sont situés soit au-dessus, soit en dessous de la forme à laquelle ils sont associés, mais jamais les deux à la fois.

N°	A la fin	Au milieu	Au début	Isolée
1	ى	ل	ا	ا
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	د	د
9	ذ	ذ	ذ	ذ
10	ر	ر	ر	ر
11	ز	ز	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	ع	ع	ع	ع
20	ف	ف	ف	ف
21	ق	ق	ق	ق
22	ك	ك	ك	ك
23	ل	ل	ل	ل
24	م	م	م	م
25	ن	ن	ن	ن
26	ه	ه	ه	ه
27	و	و	و	و
28	ي	ي	ي	ي

Tableau I.1:L'alphabet arabe

•	◌	◌		
◌◌	◌	◌	◌◌	◌
◌◌	◌	◌	◌	

Tableau I.2:Points en arabe : un, deux ou trois points.

La Figure Tableau I.2 illustre la variabilité des styles d'écriture des points ou groupes de points en écriture arabe manuscrite. Un groupe de deux points peut ainsi s'écrire sous forme d'une seule, ou de deux composantes connexes. On remarque la très forte similarité entre deux points reliés par un trait, et une voyelle de type « A » ou « I » dont les exemples sont donnés par Figure I.7.

Un groupe de trois points peut donner lieu à une, deux ou trois composantes connexes, en fonction du style d'écriture.

- Les voyelles

En arabe, les voyelles ne sont pas des lettres, mais des signes diacritiques associés aux lettres sur lesquelles ils s'appliquent Figure I.7.

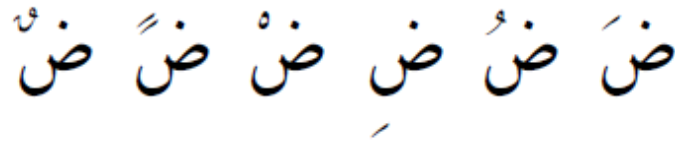


Figure I.9:Voyelles en arabe

L'absence de voyelles peut toutefois être source de confusions, un mot peut avoir plusieurs voyelles possibles et par conséquent plusieurs catégories grammaticales. Dans certains cas, une phrase peut donc avoir deux voyelles différentes, ce qui nous donne deux structures syntaxiques possibles. Par exemple, dans le cas :

يخشى الأستاذ الطلبة

Qui peut se voyeller des deux manières suivantes :

L'enseignant craint les étudiant

يَخْشَى الْأَسْتَاذُ الطَّلَبَةَ

Les étudiants craignent l'enseignant

يَخْشَى الطَّلَبَةُ الْأَسْتَاذَ

- Ascendants et descendants

Comme dans l'écriture latine, l'écriture arabe contient des ascendants et des descendants (voir Figure I.8). En arabe, les descendants 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.

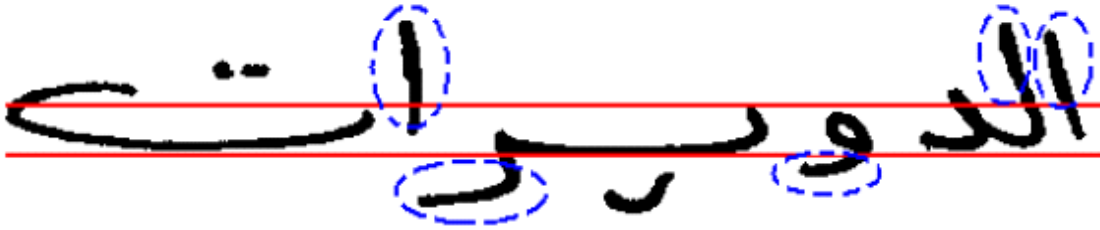


Figure I.10: Les ascendants et descendants sont entourés

- Une ou plusieurs composantes connexes par mot :

6 lettres ne sont pas liées à leur successeur : و, ز, ر, ذ, د, ا. Ces lettres introduisent donc un en coupure dans le mot. Un pseudo-mot est une unité connexe regroupant une ou plusieurs lettres sous forme d'une séquence. Un mot peut être composé d'un ou plusieurs pseudo-mots (voir Figure I.9).



Figure I.11: Un mot peut être composé de plusieurs composantes connexes (pseudo-mots).

- Ligatures verticales

En écriture arabe, il n'y a pas de liaisons hautes comme le « v » ou le « o » en latin : les ligatures se situent au niveau de la ligne support de l'écriture (ligne de base). En revanche, les scripteurs sont libres de constituer certains groupes de deux ou trois lettres liées verticalement en début de pseudo-mot. Ce sont les ligatures verticales (Figure II.10 et Figure II.11). En général très complexes à segmenter, nous choisissons de les reconnaître telles quelles.



Figure I.12: Sous forme d'une ou de deux composantes connexes



Figure I.13:Ligatures verticales et inversion de l'ordre du tracé.

5. Conclusion :

Jusqu'à ici nous avons obtenu un ensemble de notions de bases sur le domaine de reconnaissance de formes, nous pouvons poursuivre vers un domaine plus spécialisé, vers les systèmes de classifications.

Chapitre II :
Les Systèmes de
Classifications

1. INTRODUCTION :

Ce chapitre sera consacré à une présentation de quelques propriétés des réseaux de neurones qui ont été utilisées dans ce mémoire, puis un rappel de l'algorithme de k plus proches voisins qui a été aussi utilisé, enfin une brève description des Modèles de Markov Cachées à cause de son large utilisation dans le domaine de la reconnaissance de l'écriture arabe manuscrite enfin une centration sur les réseaux de neurones qui ont été appliqués dans notre projet .

2. Définition formelle d'un classificateur :

Définir un problème de classification, c'est d'abord de choisir une taxonomie c'est-à-dire une répartition des formes à reconnaître dans un ensemble de classes $\omega_i, i \in [1, \dots, c]$. Dans le cas le plus général, on peut associer à la forme à reconnaître un vecteur de degré d'appartenance $D(x)$ tel que :

$$D(x) = \begin{bmatrix} D_1(x) \\ D_2(x) \\ \vdots \\ D_N(x) \end{bmatrix}$$

Avec $D_i(x) = D\{\omega_i\}$ Dans ce cas, peut appartenir à plusieurs classes si $D_i(x) \neq 0$.

Toutefois, dans la majorité des problèmes de classification, on a affaire à une classification exclusive dans laquelle une forme ne peut appartenir qu'à une seule classe.

On a alors $D_i(x) = \delta_{i,j}$ tel que :

$$\delta_{i,j} = \begin{cases} 1 & \text{si } j = i \\ 0 & \text{si } j \neq i \end{cases}$$

ω_i est alors souvent appelée « la vraie classe »:

$$D(x) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Selon Xu [XUL92] La réponse d'un classificateur peut être divisée en trois catégories suivantes :

2.1. Classificateur de type classe :

$$e_j(x) = \omega_i, i \in \{1, \dots, c\}$$

ω_i est la classe attribuée par le classificateur e_j à x . Dans ce cas, l'avis du classificateur est binaire. On peut alors représenter la réponse du classificateur par un vecteur binaire dans lequel l'indique la classe proposée par le classificateur. Un classificateur peut aussi produire un ensemble de classes. Il considère alors que la forme x appartient à une des classes de cet ensemble sans donner d'autres informations permettant de discriminer les classes.

2.2. Classificateur de type rang :

$$e_j(x) = (r_{1,j}, r_{2,j}, \dots, r_{N,j})$$

$R_{i,j}$ est le rang attribué à la classe ω_i par le classificateur . Il s'agit d'un classement sur les classes. Le classificateur indique ce classement en fournissant en sortie un vecteur des rangs de taille. La classe placée au premier rang de la liste proposée par le classificateur est considérée comme la plus probable pour la forme et la classe du dernier rang est la moins probable.

2.3. Classificateur de type mesure :

$$e_j(x) = (m_{1,j}, m_{2,j}, \dots, m_{N,j})$$

$M_{i,j}$ est la mesure attribuée à la classe par le classificateur . Elle indique le niveau de confiance du classificateur dans sa proposition. La sortie du classificateur est donc un vecteur de mesures de taille. Cette mesure, normalisée ou non, peut être une distance, une probabilité a posteriori, une valeur de confiance, un score, une fonction de croyance, une possibilité, une crédibilité, une mesure floue, etc... .

3. Les réseaux de neurones :

3.1. Définition :

Selon [HAY94]:

Un réseau de neurones est un processeur massivement distribué en parallèle qui a une propension naturelle pour stocker de la connaissance empirique (experiential knowledge selon l'auteur) 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.

Selon [NIG93]:

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. Bilan

3.2. Neurone simple :

3.2.1. Le neurone de Mac Culloch-Pitts :

Le premier neurone artificiel apparu est le neurone de [MCC43], il se caractérise principalement par une fonction d'activation de type limitation 0 ou 1. Il est montré dans la Figure II.1.

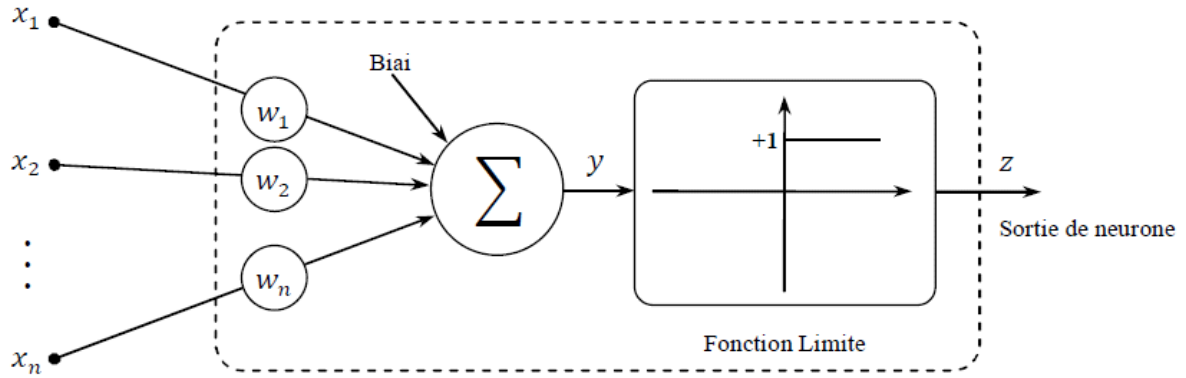


Figure II.1:Le Neurone de Mac Culloch-Pitts [MCC43]

Le neurone formel est donc une modélisation mathématique qui reprend les principes du fonctionnement du neurone biologique, en particulier la sommation des entrées. Sachant qu'au niveau biologique, les synapses n'ont pas toutes la même «valeur» (les connexions entre les neurones étant plus ou moins fortes), les auteurs ont donc créé un algorithme qui pondère la somme de ses entrées par des poids synaptiques (coefficients de pondération). De plus, les 1 et les -1 en entrée sont là pour figurer une synapse excitatrice ou inhibitrice.

3.2.2. Le Perceptron :

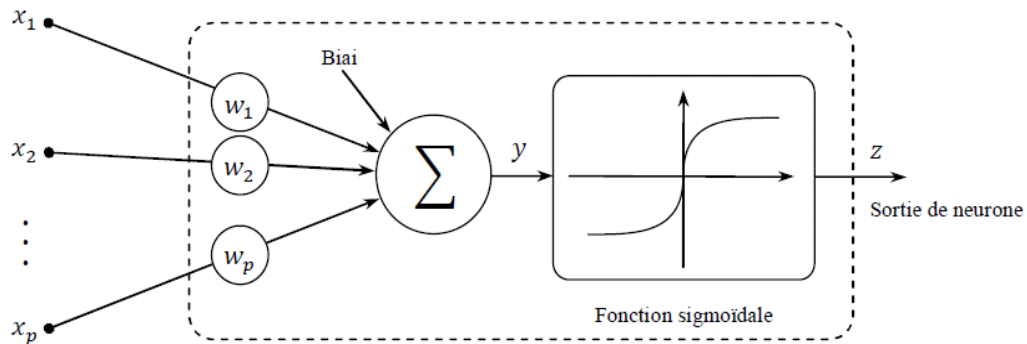


Figure II.2 : Le perceptron

Le perceptron se caractérise par une fonction d'activation continue et différentiable :

$$z = \frac{2}{1+e^{-By}} - 1$$

3.3. Interprétation mathématique :

D'un point de vue mathématique, le neurone formel peut être représenté de la manière suivante:

Un neurone à entrée est une fonction $\mathbb{R}^{p+1} \rightarrow \mathbb{R}$ définie par :

1. $g: \mathbb{R} \rightarrow \mathbb{R}$
2. $W \in \mathbb{R}^{p+1}, W = (w_1, w_2, \dots, w_{p+1})$
3. $\forall x \in \mathbb{R}^p, f(W, x) = g(\sum_{i=1}^p w_i x_i + w_{p+1})$
avec $x = (x_1, x_2, \dots, x_p)$

Cette définition est inspirée du neurone biologique, les poids jouant le rôle de synapses, le vecteur celui des entrées et w celui des coefficients ou poids. Le coefficient w_{p+1} est appelé le biais et souvent noté b . La fonction est appelée fonction de d'activation ou de transfert ou encore fonction de seuil.

3.4. Types et architectures des RNA :

3.4.1. Le Perceptron Multicouches :

Les perceptrons multicouches sont les réseaux de neurones les plus populaires, du fait de leurs bonnes performances et de leur facilité de mise en œuvre.

Ce sont des réseaux de neurones à propagation directe sans cycles, avec au moins une couche cachée. La fonction de transfert des neurones est généralement de type sigmoïde ou \tanh à valeurs dans $[0; 1]$ ou dans $[-1; 1]$.

Il a été prouvé par [CYB89] qu'une fonction continue quelconque peut être arbitrairement bien approximée par un réseau feed-forward à couche cachée unique, ou chaque neurone dans la couche cachée a comme fonction d'activation une fonction non linéaire sigmoïdale (le perceptron).

Un réseau de neurones multicouches à n sorties, p entrées et C couches est une liste de couches (C_1, \dots, C_C) connectées les unes aux autres de telle sorte que :

1. $\forall i \in \{1, \dots, C\}$, chaque couche C_i possède n_i neurones et p_i entrées
2. $\forall i \in \{1, \dots, C-1\}$, $n_i = p_{i+1}$, de plus $p_1 = p$ et $n_C = n$

3. les coefficients de la couche C_i sont notés $(W_1^i, \dots, W_{n_i}^i)$ cette couche définit une fonction F_i
4. soit la suite $(Z_i)_{0 \leq i \leq C}$ définie par : $Z_0 \in \forall i \in \{1, \dots, C\}, Z_i = F_i(W_1^i, \dots, W_{n_i}^i, Z_{i-1})$

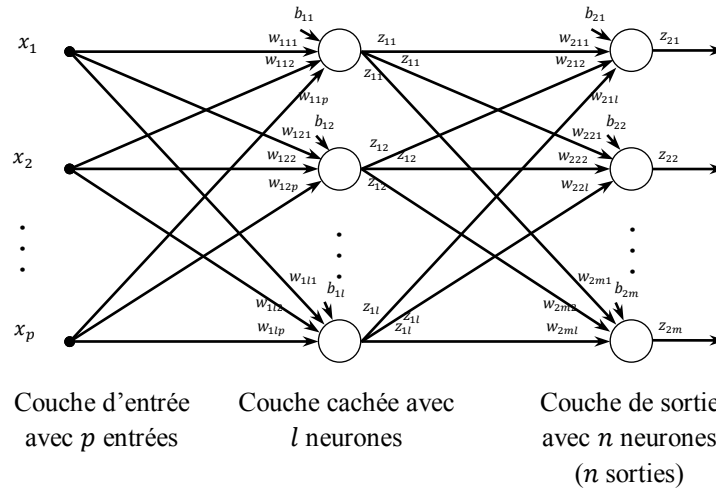


Figure II.3: Un perceptron à une couche cachée unique

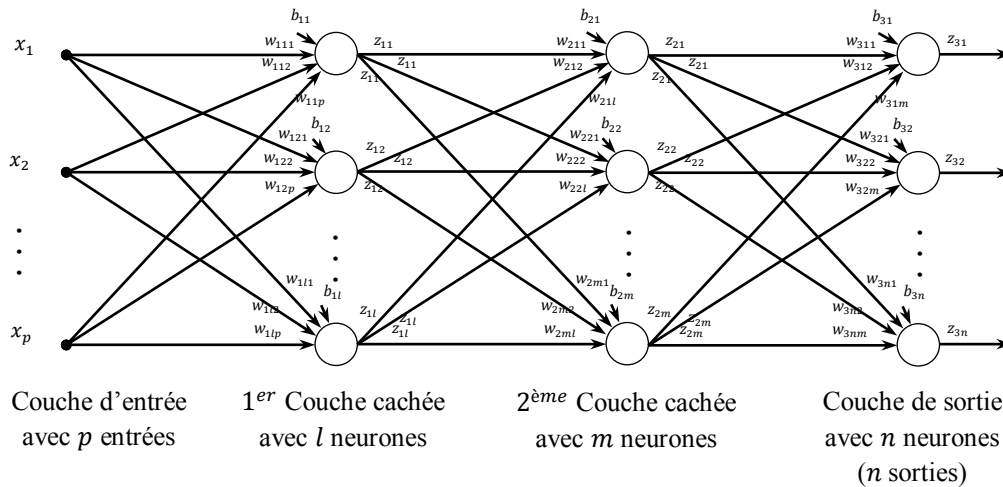


Figure II.4: Un perceptron à deux couches cachées

Souvent, on considère que les entrées forment la couche C_0 de manière à simplifier les écritures. Ainsi, chaque couche C_i du perceptron a pour entrées les sorties de la couche C_{i-1} . Cette définition est plus facile à illustrer qu'à énoncer (Figure II. et Figure II.).

Le mécanisme qui permet de calculer les sorties d'un réseau de neurones sachant ses poids est appelé *propagation*.

(x_1, \dots, x_p)	Les entrées
C	Nombre des couches
C_i	Nombre de neurones sur la couche i ($C_0 = p$)
$z_{c,i}$	Sortie du neurone i de la couche c par extension, $z_{0,i} = x_i$
$y_{c,i}$	Potentiel du neurone i de la couche c
$w_{c,i,j}$	Coefficient associé à l'entrée j du neurone i de la couche c
$b_{c,i}$	Biais du neurone i de la couche c
$f_{c,i}$	Fonction de seuil du neurone i de la couche c

3.4.1.1. La Propagation :

Cet algorithme s'applique à un réseau de neurones pour calculer les sorties de ce réseau connaissant ses poids ($w_{c,i,j}$) et ses entrées (x_j).

Etape A : initialisation

pour $i = 1$ à C_0 **faire**

$z_{0,i} \leftarrow x_i$

fin pour

Etape B : Propagation

pour $c = 1$ à C **faire**

pour $i = 1$ à C_c **faire**

$z_{c,i} \leftarrow 0$

pour $j = 1$ à C_{i-1} **faire**

$z_{c,i} \leftarrow z_{c,i} + w_{c,i,j} z_{c-1,j}$

fin pour

$z_{c,i} \leftarrow f(z_{c,i} + b_{c,i})$

fin pour

fin pour

Les perceptrons multicouches définissent une classe de fonctions qui permet d'approximer n'importe quelle fonction continue à support compact.

3.4.2. Fonctions à bases radiales :

Les réseaux de neurones de type RBF (Radial Basis Function) ont généralement une couche cachée. Les neurones de cette couche cachée sont de types gaussiens. Les neurones

de sorties sont une combinaison linéaire des neurones gaussiens de la couche cachée, parfois suivis d'une fonction de transfert sigmoïde.

Les caractéristiques des neurones gaussiens (centres et variances) sont généralement estimées au début de l'apprentissage. Ces paramètres sont ensuite figés. L'apprentissage du réseau de neurones consiste à apprendre la couche de combinaison linéaire de sortie, par descente de gradient.

3.5. L'apprentissage :

L'apprentissage peut être considéré comme le problème de la mise à jour des poids des connexions au sein du réseau afin de réussir la tâche qui lui est demandée. L'apprentissage est la caractéristique principale de RNA, il peut se faire de différentes manières et selon de différentes règles.

Au niveau des algorithmes d'apprentissage, il a été défini deux grandes classes selon que l'apprentissage est dit supervisé ou non supervisé. Cette distinction repose sur la forme des exemples d'apprentissage. Dans le cas de l'apprentissage supervisé, les exemples sont des couples (Entrée, Sortie associée) alors que l'on ne dispose que des valeurs (Entrée) pour l'apprentissage non supervisé. Remarquons cependant que les modèles à apprentissage non supervisé nécessitent avant la phase d'utilisation une étape de labellisation effectuée l'opérateur, qui n'est pas autre chose qu'une part de supervision.

3.5.1. La loi d'apprentissage de Hebb [HEB49] :

S'il y a une couche de neurones qui partagent P entrées, la sortie du neurone j est

$$z_j = f_j(y_j) = f_j\left(\sum_{i=1}^{p+1} w_{ij}x_i\right), w_{p+1} = \theta_j$$

Représente le biais, la modification des poids est définie comme suite :

$$\Delta w_{ij} = \mu x_i z_j$$

$$w_{ij(t+1)} = w_{ij(t)} + \Delta w_{ij(t)}$$

- z_j est la sortie du neurone j
- $w_{ij(t+1)}$ est le nouveau poids i du neurone j , $w_{ij(t)}$ est l'ancien poids i du neurone j
- μ représente le coefficient d'apprentissage

L'algorithme d'apprentissage modifie de façon itérative 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. La règle d'apprentissage de Hebb est une règle d'apprentissage non supervisé.

3.5.2. La règle d'apprentissage du Perceptron :

La règle d'apprentissage du perceptron [ROS58] s'applique sur un neurone de type signe, elle est définie comme suite :

$$\Delta w_{ij} = \mu x_i e_j$$
$$e_j = d_j - z_j = \text{signe} \left(\sum_{i=1}^{p+1} w_{ij} x_i \right)$$

Avec :

- e_j est l'erreur de la sortie j
- d_j est la sortie désirée j

L'apprentissage avec cette règle est un apprentissage supervisé.

3.5.3. Retro-propagation du gradient :

L'algorithme de rétro-propagation a été développé en particulier par [RUM86]. Cet algorithme repose sur la minimisation de l'erreur quadratique entre les sorties calculées et celles souhaitées. Le terme rétro-propagation du gradient provient du fait que l'erreur calculée en sortie est transmise en sens inverse vers l'entrée.

$$\Delta w_{c,i,j} = \mu f'(y_{c,i}) \delta_{c,i} z_{c-1,j}$$

Si c représente une couche de sortie alors : $\delta_{c,i} = d_i - z_{c,i}$

Si c représente une couche cachée alors : $\delta_{c,i} = \sum_{j=1}^{c_{c+1}} w_{j,i} \delta_{c-1,j}$

3.6. Avantages et 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 [DAV93], [MEI00].

3.6.1. Avantages de l'approche neuronale :

- 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 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).

3.6.2. Inconvénients de l'approche neuronale :

- 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 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.

4. K plus proches voisins :

L'algorithme des k plus proches voisins (KPPV) est certainement un des algorithmes les plus simples d'apprentissage automatique, il a été proposé par [COV91]. Il est motivé par l'idée que des entrées semblables devraient avoir des cibles semblables. Ainsi, pour bien définir un algorithme KPPV, il suffit de définir ce que veut dire "semblable" dans le contexte des entrées et de définir l'influence de ces voisins sur la prédiction de la cible pour une entrée de test. Donc, pour obtenir une prédiction de la cible pour une entrée de test, il suffit de trouver les k plus proches voisins selon une métrique déterminant jusqu'à quel point des entrées sont semblables et d'utiliser ces k plus proches voisins pour prédire la cible de x. Dans un problème de classification, la prédiction correspond à la classe majoritaire parmi les k plus proches voisins, i.e. que l'ensemble des k plus proches voisins votent pour la classe correspondant à leur cible respective et la classe recueillant le plus de vote est choisie en tant que prédiction par l'algorithme.

4.1. Formalisation mathématique :

Soit :

$T = \{(x_i, y_i) \mid x_i \in \{1, \dots, c\}\}$ un ensemble d'apprentissage, où $y_i \in Y = \{0, \dots, -1\}$

correspond à l'identité de la classe cible de l'entrée x.

- c le nombre de classes.

- $d(.,.)$ une fonction de distance
- une entrée de test.
- $V = (T, d(.,.), K)$ l'ensemble des k plus proches voisins de x parmi les entrées

De T ainsi que leur cible associée

La prédiction de classification par l'algorithme des k plus proches voisins est donc:

$$\operatorname{argmax}_{y \in Y} \sum_{(x_i, y_i) \in (x, T, d(.,.), k)} I_{y_i = y}$$

Une fonction de distance couramment utilisée est la distance euclidienne:

$$d(a, b) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

4.2. ALGORITHME :

On définit un algorithme d'apprentissage en précisant sa procédure d'entraînement et de prédiction pour une nouvelle entrée de test. Etant donné que la procédure d'entraînement de l'algorithme KPPV consiste simplement à mettre en mémoire l'ensemble d'apprentissage T , voici donc la procédure de KPPV [LAR07][COV91]:

```

pour  $i \leftarrow 0$  à  $C - 1$ 
  faire  $c[i] \leftarrow 0$ 

pour  $i \leftarrow 0$  à  $k - 1$ 
  faire  $\begin{cases} \text{voisin}[i] \leftarrow \text{NULL} \\ \text{dist}[i] \leftarrow \text{inf} \end{cases}$ 

pour  $i \leftarrow 0$  à  $k - 1$ 
  faire  $\begin{cases} dt \leftarrow d(x_i, x) \\ j \leftarrow \text{argmax}_i \text{dists}[i] \\ \text{si } dt < \text{dists}[j] \\ \text{alors } \begin{cases} (x_i, y_i) \leftarrow \text{voisins}[i] \\ c[y_i] \leftarrow c[y_i] + 1 \end{cases} \end{cases}$ 

retourner ( $\text{argmax}_i c[i]$ )
  
```

5. Modèles de markov caches :

5.1. Définitions :

5.1.1. Chaîne de Markov :

Une chaîne de Markov discrète d'ordre n est un processus stochastique discret $X = \{X_t | t = 1, \dots, T\}$ avec des variables aléatoires discrètes (les réalisations de ces variables sont appelées états) [HAL04].

$$P(X_t = q_{i_t} | X_{t-1} = q_{i_{t-1}}, \dots, X_1 = q_{i_1}) = P(X_t = q_{i_t} | X_{t-1} = q_{i_{t-1}}, \dots, X_{t-n} = q_{i_{t-n}})$$

$$\forall t \in [1, T]$$

Où $Q = \{q_1, \dots, q_n\}$ représente l'ensemble des états.

Dans la suite on n'utilise que les chaînes de Markov d'ordre 1.

5.1.2. Chaîne stationnaire :

Une chaîne de Markov d'ordre 1 est stationnaire si pour tout t et k on a :

$$P(X_t = q_i | X_{t-1} = q_j) = P(X_{t+k} = q_i | X_{t+k-1} = q_j)$$

Dans ce cas, on définit une matrice de probabilité de transition $A = (\alpha_{ij})$ telle que:

$$a_{ij} = P(X_t = q_j | X_{t-1} = q_i)$$

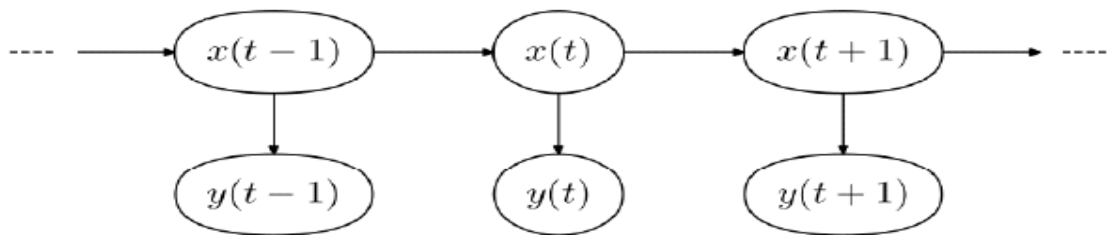
A un instant donné d'un processus quelconque.

5.2. Un modèle de markov cache :

Un modèle de Markov caché (HMM) est une chaîne de Markov cachée stationnaire où l'observation est une fonction probabiliste de l'état.

La le schéma d'un HMM.

X état caché



Y observation

Figure II.5:Modèle de Markov Caché.

5.3. Interprétation mathématique :

Selon [DUP04]

Un HMM discret du premier ordre est défini par:

N états représentés par $Q = \{q_1, q_2, \dots, q_N\}$. On rappelle que l'on désigne un état au temps t par $q_t \in Q$

M symboles observés dans chaque état $V = \{o_1, o_2, \dots, o_M\}$. On désigne une observation au temps t par $O_t \in V$

La matrice de probabilité de transition entre les états $A = (\alpha_{ij})_{1 \leq i, j \leq N}$ avec

$$a_{ij} = P(X_t = q_j | X_{t-1} = q_i) \quad \forall i, j \in [1, N]$$

$$a_{ij} \geq 0$$

$$\sum_{j=1}^N a_{ij} = 1 \quad \forall i \in [1, N]$$

5.4. Les types des HMMs :

Selon la topologie du réseau des états, il y a deux types de HMMs. Ce sont le modèle ergodique et le modèle gauche-droite [HAL04].

5.4.1. HMM ergodique :

C'est un modèle sans contraintes où toutes les transitions d'un état vers l'autre sont possibles, c'est-à-dire $a_{ij} \geq 0 \forall (i,j) \in [1, N]$.

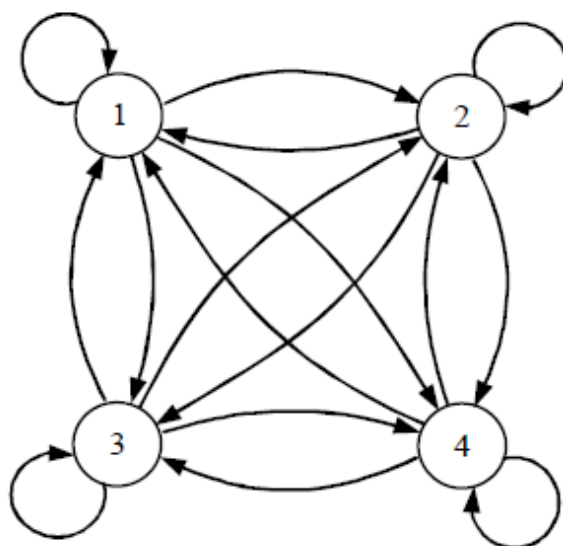


Figure II.6:Modèle ergodique

5.4.2. HMM gauche-droite :

C'est un modèle où il y a des contraintes sur des transitions : seulement la transition d'un état ayant un indice bas vers un état ayant un indice haut est acceptée. Il existe deux sous-types de ce modèle : le modèle parallèle et le modèle séquentiel.

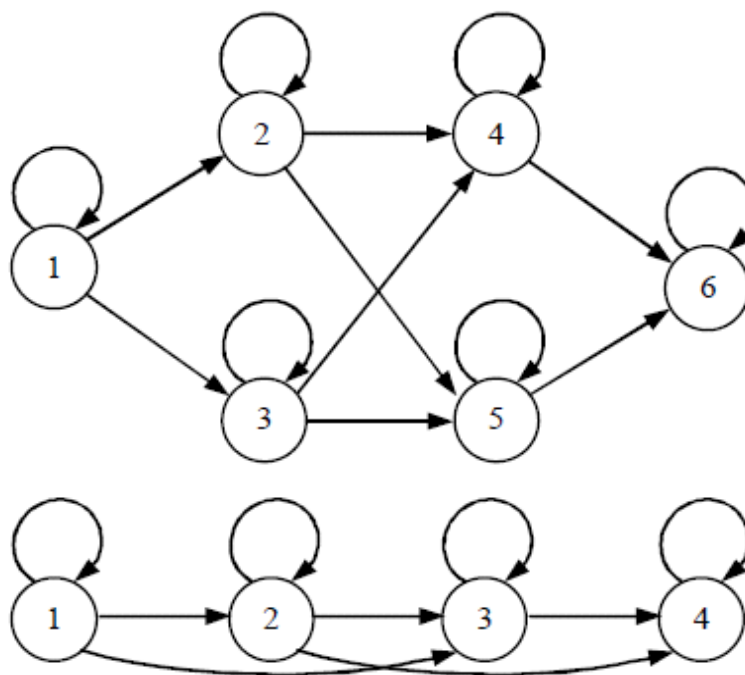


Figure II.7: Modèles gauche-droite

Le modèle séquentiel est souvent utilisé pour reconnaître la parole et l'écriture manuscrite car il est convenable à la modélisation des signaux et il y a moins de paramètres à estimer dans la phase de l'apprentissage. Dans ce modèle, la densité de probabilité initiale n'est pas nécessaire [HAL04].

Selon les types des symboles des observations on a les HMMs discrets et les HMMs semi-continus.

6. Conclusion :

Nous avons vu dans ce chapitre les réseaux de neurones et leurs principes de bases, l'algorithme de KPPV et les modèles de Markov Cachés, on peut passer maintenant au troisième chapitre où nous présenterons les tests et la validation des résultats obtenus.

**CHAPITRE III :
CONCEPTION
ET RESULTATS**

1. Introduction :

Dans ce chapitre nous allons présenter notre travail, en commençant par une présentation de la base de mots et les outils de développement utilisés, ainsi que l'apprentissage et le test des classificateurs individuels. Enfin nous allons voir si la combinaison des cinq réseaux va augmenter le taux de reconnaissance moyen.

2. Conception de la solution proposée :

Dans la suite on va présenter les différents étapes et méthodes mis en œuvre lors de la réalisation de notre projet. Pour réaliser un système de reconnaissance de caractères arabes, on passe par les étapes illustrées sur la figure suivante.

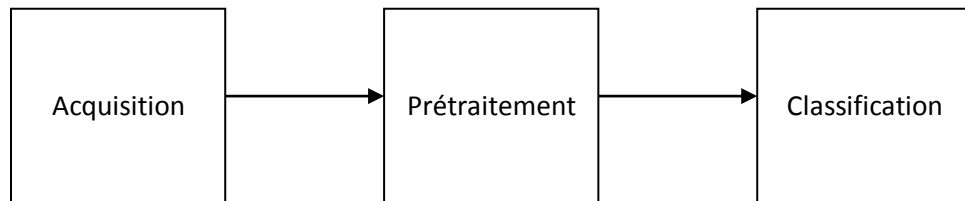


Figure III.1: Les étapes majeures pour la réalisation d'une RF arabe.

Dans l'étape d'acquisition on doit acquérir une base de données représentant les différentes formes des caractères manuscrits (arabes). L'acquisition de cette base se fait par la méthode suivante :

- Acquisition des caractères à l'aide d'un Scanner (pour le mode statique ou off-line).
- Acquisition des images dessinés dans la zone du dessin via la souris (pour le mode on-line)

2.1. Prétraitement :

Un classificateur ne peut pas effectuer efficacement sur les images originales. C'est pourquoi, on a besoin de prétraiter les images pour qu'ils soient normalisées et facile à classifier.

Le prétraitement passe par quatre étapes :

2.1.1. Binarisation :

L'opération de binarisation (seuillage) consiste à mettre à zéro tous les pixels ayant un niveau de gris inférieur à une certaine valeur (appelée seuil, en anglais treshold) et à la valeur maximale les pixels ayant une valeur supérieure. Ainsi le résultat du seuillage est une image binaire contenant des pixels noirs et blancs, c'est la raison pour laquelle le terme de binarisation est employé. La binarisation permet de mettre en évidence les caractères ou des objets dans une image. Cette étape est nécessaire dans le cas d'acquisition par un scanner ou une caméra

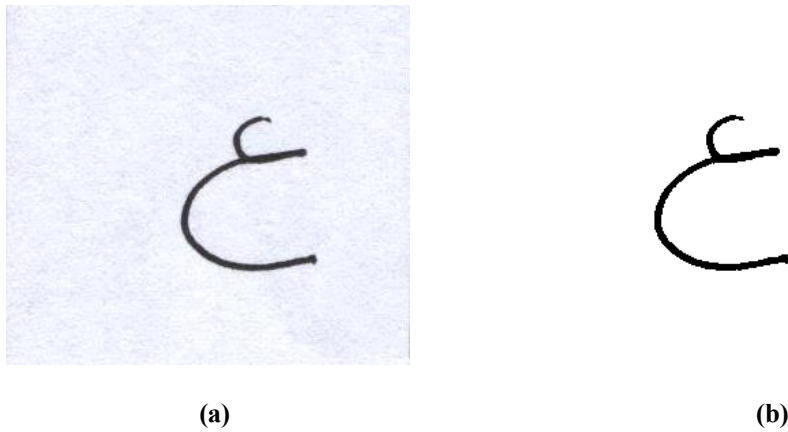


Figure III.2: (a) Image colorée, (b) Image en noir et blanc.

2.1.2. Squelettisation :

La squelettisation est une procédure qui a pour but de réduire l'épaisseur du tracé d'un caractère à un pixel seulement. L'amincissement des lignes se fait par la suppression de tous les pixels redondants jusqu'à ce que l'épaisseur reste un seul point.



Figure III.3: (a) Image avant la squelettisation, (b) Image après la squelettisation

2.1.3. Centrage :

Puisqu'un caractère peut être de différentes tailles, rien ne l'empêche d'être positionné n'importe où dans la fenêtre qui lui est consacrée, comme le montre la figure III.19(a).

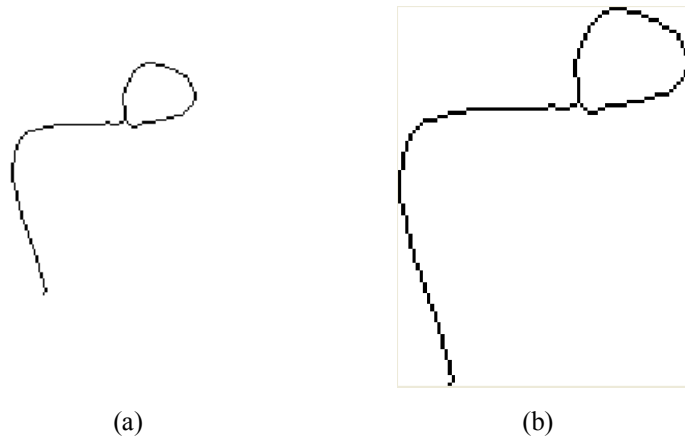


Figure III.4: (a) Image normale, (b) Image centrée

Pour des raisons d'homogénéité entre la base d'apprentissage et les images des caractères que le réseau aura à reconnaître, il est nécessaire de ramener le caractère au centre du carré, voir figure III.19(b).

2.1.4. Extraction des caractéristiques :

Normalement, l'image d'objet est représentée par une matrice de pixels et le classificateur ne fonctionne pas bien lorsque l'entrée est une matrice de grande dimension. L'extraction de caractéristiques est une procédure très importante pour extraire des caractéristiques. A partir d'ici, chaque image sera représenté par son vecteur de caractéristiques.

Parmi les différents descripteurs d'objet, nous avons choisi la méthode des pixels moyennés.

2.1.4.1. Méthode des pixels moyennés :

Dans cette méthode-là, la matrice de l'image est divisée en $n \times m$ (dans notre cas 6×6) zones de même taille. Pour chaque zone, on calcule la valeur moyenne de niveaux de gris de tous les pixels dans cette zone et le résultat obtenu sera utilisé comme un élément du vecteur de caractéristique. Donc le vecteur de caractéristique est de taille $6 \times 6 = 36$.

2.2. Apprentissage et classification :

L'étape d'apprentissage et de test n'existe qu'à la première fois. Après, elle va être remplacée par la phase de classification.

Dans notre cas on va utiliser d'abord l'apprentissage supervisé guidé par les but (target), les exemples sont des couples (Entrée, Sortie associée).

La classification se fait par un réseau de neurones d'une seule couche cachée, en utilisant l'algorithme de rétro-propagation d'erreur. Dans le cadre de ce projet on va implémenter en Matlab tous les algorithmes du projet.

3. Base des caractères et outils de développement :

3.1. Base des Caractères :

Les images des caractères arabes utilisées dans ce projet sont des caractères isolés, Cette base des caractères est notre propre travail, et elle contient les 28 lettres arabes illustrées par le tableau suivant:

N°	Caractère arabe	Caractère a.manuscrit	N°	caractère arabe	Caractère a.manuscrit
1	أ	ا	15	ض	ض
2	ب	ب	16	ط	ط
3	ت	ت	17	ظ	ظ
4	ث	ث	18	ع	ع
5	ج	ج	19	غ	غ
6	ح	ح	20	ف	ف
7	خ	خ	21	ق	ق
8	د	د	22	ك	ك
9	ذ	ذ	23	ل	ل
10	ر	ر	24	م	م
11	ز	ز	25	ن	ن
12	س	س	26	ه	ه
13	ش	ش	27	و	و
14	ص	ص	28	ي	ي

Tableau III.1 : Des caractères extraits de la base utilisée.

3.2. Mathworks MATLAB :

MATLAB est à la fois un langage de programmation et un environnement de développement, il est développé et commercialisé par la société américaine The MathWorks. MATLAB est utilisé dans les domaines de l'éducation, de la recherche et de l'industrie pour le calcul numérique mais aussi dans les phases de développement de projets.

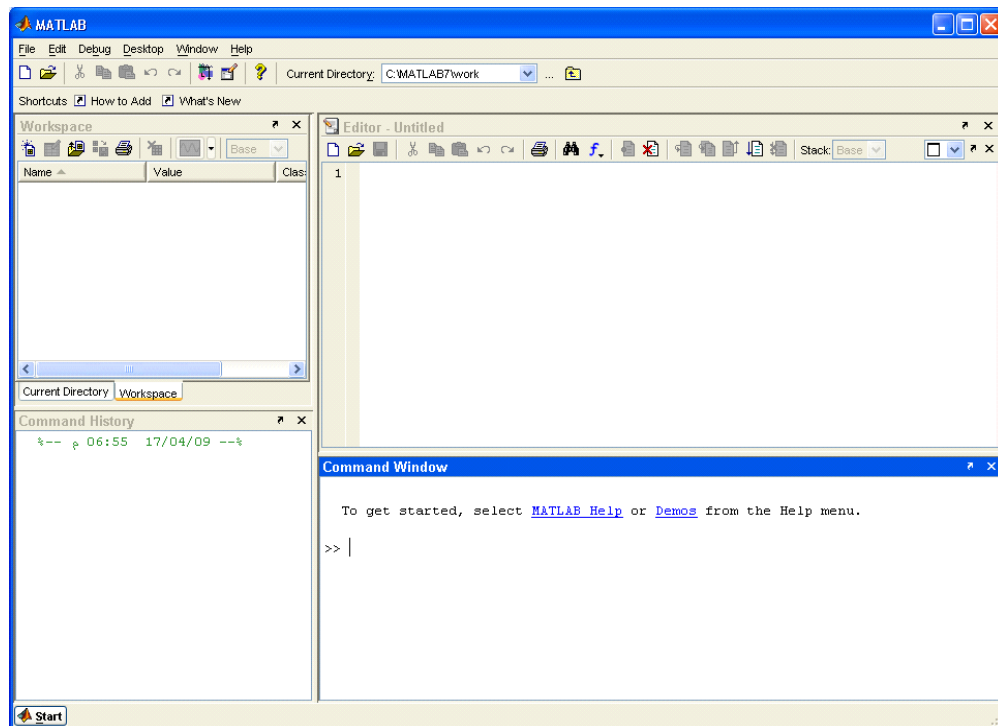


Figure III.5: L'interface principale du Matlab.

4. Présentation de l'application :

4.1. Les étapes de reconnaissance d'un caractère :

Le caractère peut être reconnu selon deux modes différents: mode en-ligne et mode hors-ligne mais dans notre application on a utilisé que le mode hors- ligne. Ce mode convient aux caractères manuscrits déjà rédigés.

4.1.1. Chargement de l'image à partir de la base des caractères :

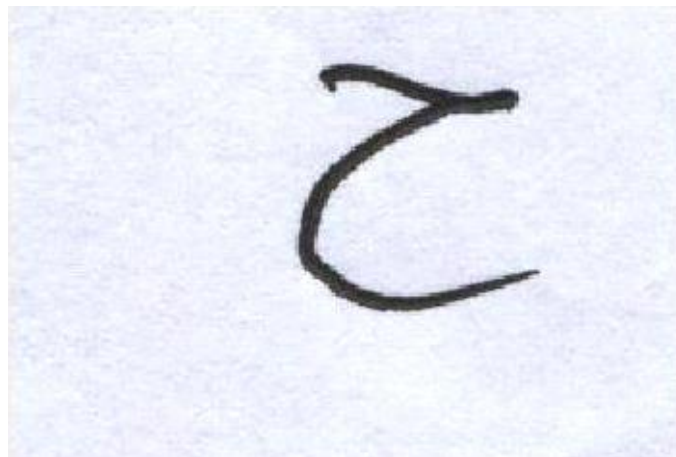


Figure III.6: Image chargée de la base.

4.1.2. La binarisation :

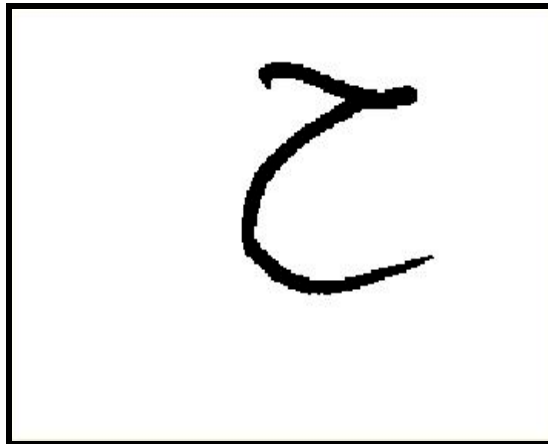


Figure III.7: Image après la binarisation.

4.1.3. Squelettisation :

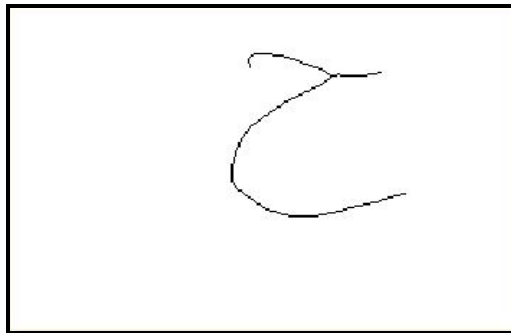


Figure III.8: Image après la squelettisation

4.1.4. Centrage :

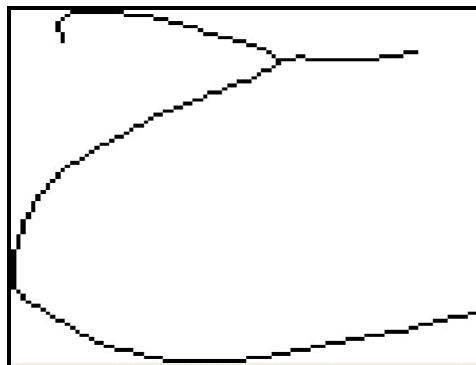


Figure III.9 : Image après le centrage

4.1.5. Extraction des caractéristiques :

0,046875	0,0625	0,0625	0,070313	0,0625	0,011719
0	0,027344	0,0625	0,019531	0	0
0,035156	0,035156	0	0	0	0
0,0625	0	0	0	0	0
0,0625	0	0	0	0	0
0,046875	0,0625	0,0625	0,0625	0,0625	0,0625

Tableau III.2 : Extraction des caractéristiques.

4.1.6. Reconnaissance :



Figure III.10 : Reconnaissance

5. Test et validation :

Cette étape d'apprentissage et de test n'existe qu'à la première fois. Après, elle va être remplacée par la phase de classification.

Nous avons utilisé la base des caractères qui contient 28 caractères, elle contient en global 168 caractères à partir des différents scripteurs, cet ensemble a été divisé en deux sous ensembles, un pour l'apprentissage (75%), et l'autre pour le test (25%).

5.1. Classification :

La classification se fait par un ensemble de 5 réseaux de neurones avec une seule couche cachée créée sous Matlab avec la fonction **newff**, chaque réseau est détaillé comme suit :

- Une couche d'entrée avec 36 entrées correspondantes aux caractéristiques extraites.
- Une couche cachée avec 12 neurones, ce nombre de neurones a été trouvé expérimentalement.
- Une couche de sortie avec 28 neurones correspondants au nombre des caractères arabes.

5.1.1. Le premier RNA :

L'apprentissage du premier réseau a donné un taux de reconnaissance de 32,1429 % de l'ensemble du test, la courbe de performance du réseau durant l'apprentissage est montrée par la figure suivante.

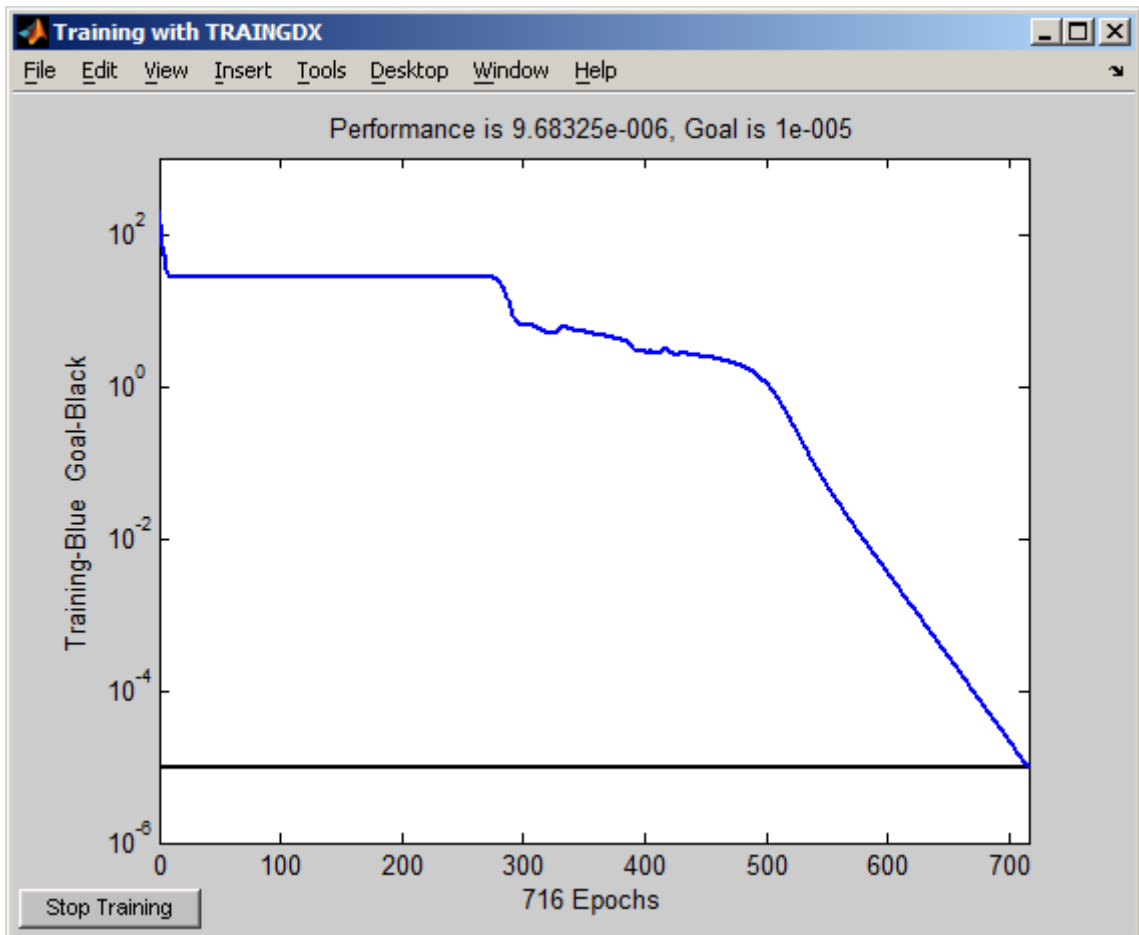


Figure III.11: la progression d'apprentissage du premier RNA.

L'erreur moyenne quadratique minimale obtenue par l'apprentissage du premier réseau de neurones est $9.68325e-006$ après 716 époques (voir Figure III.11).

5.1.2. Le deuxième RNA :

L'apprentissage du deuxième réseau a donné un taux de reconnaissance de 41,66% de l'ensemble du test, la courbe de performance du réseau durant l'apprentissage est montrée par la figure suivante :

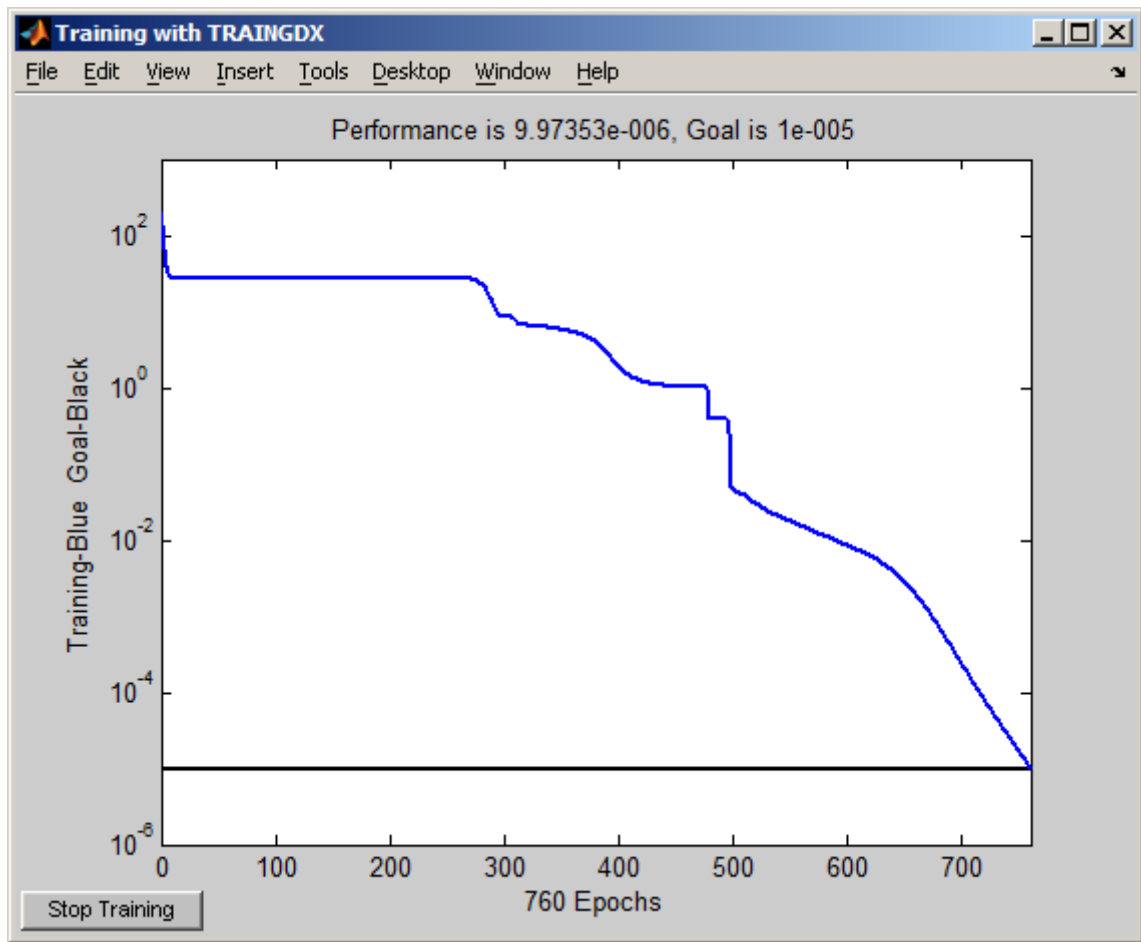


Figure III.12: la progression d'apprentissage du deuxième RNA.

L'erreur moyenne quadratique minimale obtenue par l'apprentissage du deuxième réseau de neurones est 9.97353×10^{-6} après 760 époques (voir Figure III.12)

5.1.3. Le troisième RNA :

L'apprentissage du troisième réseau a donné un taux de reconnaissance de 50,00 % de l'ensemble du test, la courbe de performance du réseau durant l'apprentissage est montrée par la figure suivante :

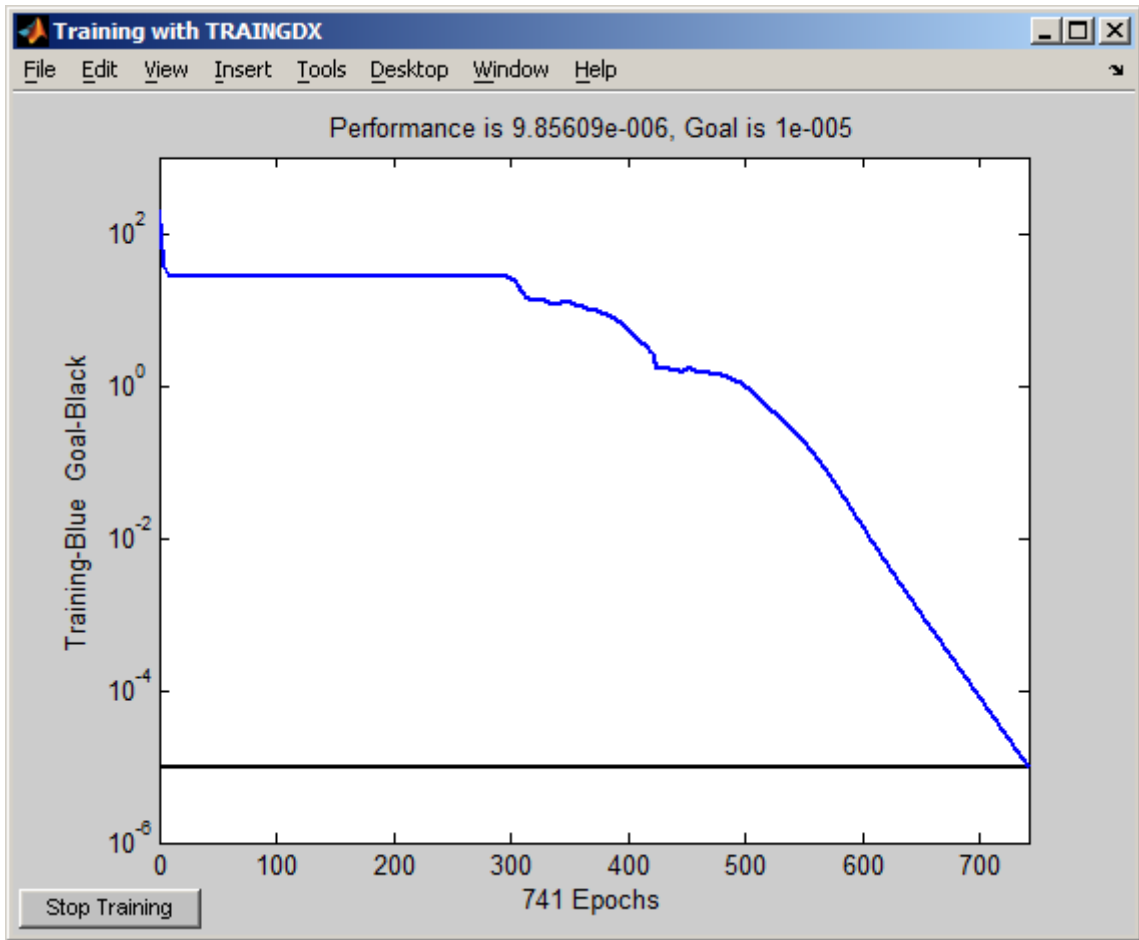


Figure III.13 : La progression d'apprentissage du troisième RNA.

L'erreur moyenne quadratique minimale obtenue par l'apprentissage du premier réseau de neurones est $9.85609e-006$ après 741 époques (voir Figure III.13).

5.1.4. Le quatrième RNA :

L'apprentissage du quatrième réseau a donné un taux de reconnaissance de 41,66 % de l'ensemble du test, la courbe de performance du réseau durant l'apprentissage est montrée par la figure suivante.

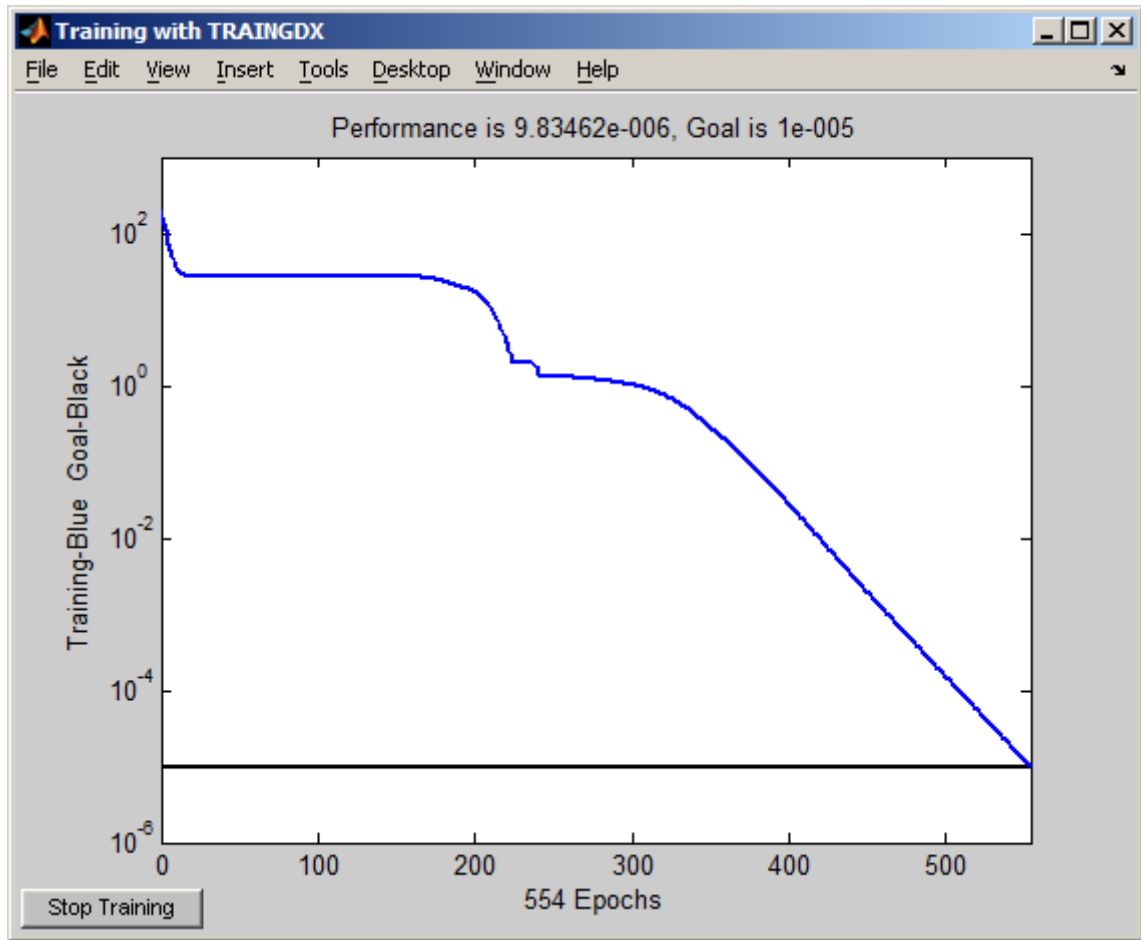


Figure III.14: La progression d'apprentissage du quatrième RNA.

L'erreur moyenne quadratique minimale obtenue par l'apprentissage du premier réseau de neurones est 9.83462×10^{-6} après 554 époques (voir Figure III. 14).

5.1.5. Le cinquième RNA :

L'apprentissage du cinquième réseau a donné un taux de reconnaissance de 50,00% de l'ensemble du test, la courbe de performance du réseau durant l'apprentissage est montrée par la figure suivante.

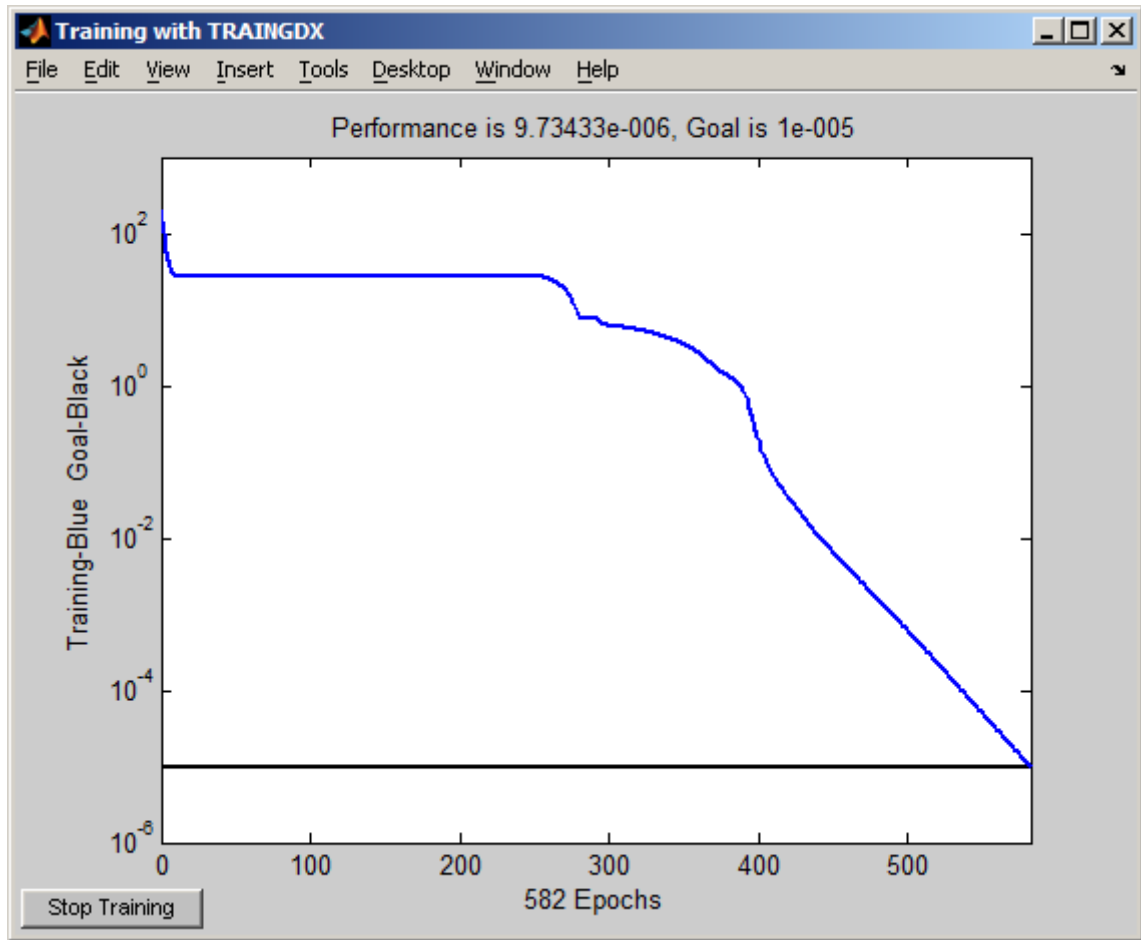


Figure III.15: La progression d'apprentissage du Le cinquième RNA.

L'erreur moyenne quadratique minimale obtenue par l'apprentissage du premier réseau de neurones est $9.73433e-006$ après 582 époques (voir Figure IV.15).

6. Combinaison :

Après avoir vu les résultats des cinq réseaux individuels, on les a combiné parallèlement en utilisant une combinaison par somme, nous avons obtenu un taux de reconnaissance égale à 67,85%, la structure du système est montrée par la figure III.16.

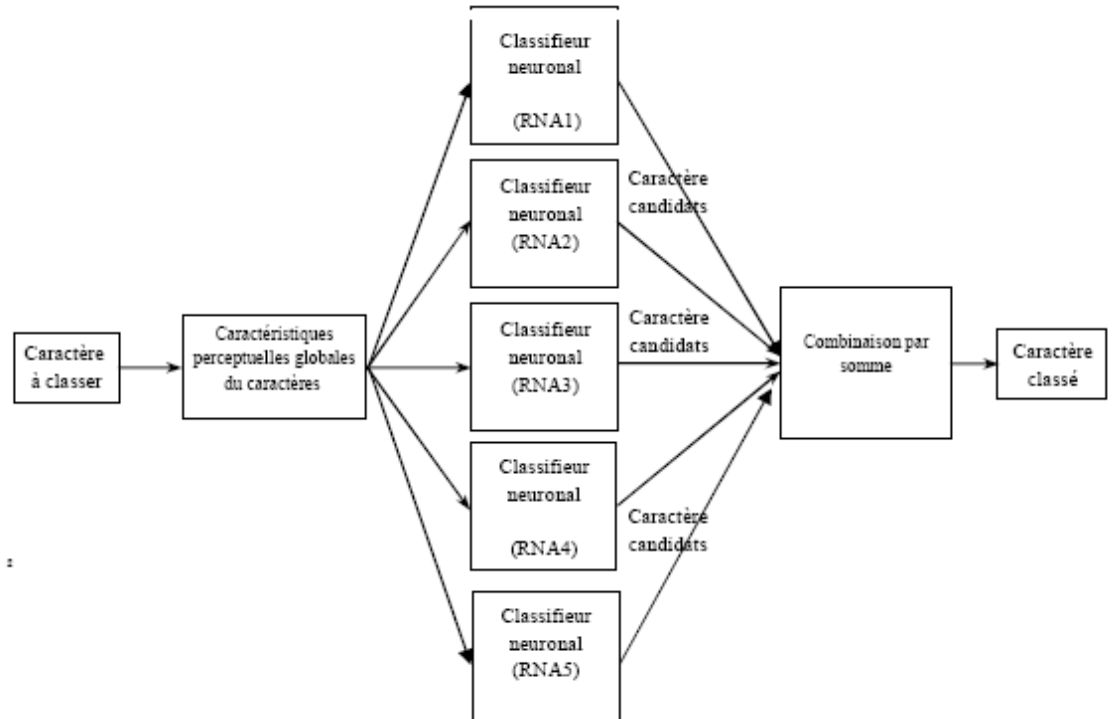


Figure III .16 : L'architecture du système de classification proposé.

Les résultats de performances des réseaux individuels et les réseaux combinés sont résumés dans le Tableau III.3.

Classificateur	Taux de reconnaissance
RNA1	32 ,1429 %
RNA2	41,66 %
RNA3	50 %
RNA4	41,66 %
RNA5	50 %
Les cinq combinés	67,85 %

Tableau III.3 : Les résultats de performance avant et après la combinaison.

7. Conclusion :

Le taux de reconnaissance a été augmenté avec la combinaison des cinq réseaux par 17.85% par rapport au meilleur réseau et par 67.85% par rapport à la moyenne des cinq réseaux, on déduit qu'avec l'utilisation de la combinaison des réseaux on bénéficie des puissances individuelles des classificateurs, pour un taux de reconnaissance plus élevé.

**CONCLUSION
ET
PERSPECTIVES**

Le domaine de la reconnaissance de la langue arabe est très vaste et riche en connaissance, que ce soit sur la langue elle-même ou ses différentes tournures et ses ambiguïtés.

Le problème de la reconnaissance des caractères arabes est un axe de recherche en plein d'expansion et attire les chercheurs.

Ce mémoire porte sur un système intelligent de classification basée sur les Réseaux de Neurones pour la reconnaissance des caractères arabes manuscrits avec des scripteurs multiples.

Un système de classification a été réalisé pour un ensemble de caractères réduit et fixé, et pour des réseaux neuronaux multiples. Ce système obtient des performances meilleures qu'avec un seul réseau comme montré dans le dernier chapitre.

Par ce travail nous espérons que nous avons couvert une partie concernant le domaine de recherche de la combinaison des classificateurs appliquée sur l'écriture arabe manuscrite, et pu contribuer à l'évolution des recherches, malgré que les efforts de nos jours s'intensifient dans ce domaine et chaque jour de nouveaux articles sont publiés, traitant du sujet.

Dans l'avenir il est possible d'intégrer un ou plusieurs classificateurs basé sur K-Plus Proches Voisins dans le système de classification et essayer de bénéficier les puissances des K-Plus Proches Voisins montrées dans les travaux d'aujourd'hui.

Nous espérons dans l'avenir pouvoir intégrer un groupe de recherche dans ce domaine pour pouvoir contribuer dans la recherche ainsi que développer notre algorithme pour couvrir plus de caractères arabes et plus de scripteurs et contribuer dans l'avancement des recherches dans le domaine de reconnaissance de l'écriture arabe manuscrite.

BIBLIOGRAPHIE

[BAD95]	B. Al-Badr and S.A. Mahmoud, "Survey and bibliography of Arabic optical text recognition," <i>Signal processing</i> , vol. 41, pp. 49-77, 1995.
[BEL97]	A. Belaid. (1997) Analyse de documents: de l'image à la représentation par les normes de codage.
[BEN99b]	A. Bennis, A. Zahour, and B. Taconet, "Extraction des lignes d'un texte manuscrit arabe," <i>Vision interface</i> , 1999.
[CAS96]	R.G. Casey and E. Lecolinet, "A survey of methods and strategies in character segmentation," <i>IEEE Transactions on pattern analysis and machine intelligence</i> , vol. 18, no. 7, july 1996.
[CRO02]	James L. Crowley, "Reconnaissance de formes dans les images", DESS GI - Option IHM, 22, janvier 2002.
[DUP04]	Xavier Dupré, "Contributions à la reconnaissance de l'écriture cursive à l'aide de modèles de Markov cachés," Paris V, 2004.
[HAL04]	Khalid Hallouli, "Reconnaissance de caractères par méthodes markoviennes et réseaux baésiens," Paris, Thèse de doctorat 2004.
[HAY94]	S. Haykin, <i>Neural Networks: A Comprehensive Foundation</i> . New York: Macmillan, 1994.
[HEB49]	D. Hebb, <i>The Organization of Behavior</i> . New York: Wiley, 1949.
[MAR01]	J. P. Marques de Sa, <i>Pattern Recognition: Concepts, Methods and Applications.</i> : Springer, 2001.
[MCC43]	W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," <i>Bulletin of Mathematical Biophysics</i> , vol. 5, pp. 115-133, 1943.
[MEN08]	Farès Menasri, "Contributions à la reconnaissance de l'écriture arabe manuscrite," UNIVERSITE PARIS DESCARTES, Thèse de doctorat 2008.
[NIG93]	A. Nigrin, <i>Networks for Pattern Recognition</i> . Cambridge: MA: The MIT Press, 1993.
[OLI02]	J. J. Oliveira, J. de Carvalho, C. Freitas, and R. Sabourin, <i>Feature sets evaluation for handwritten word recognition.</i> , 2002.
[PEC02]	M. Pechwitz, S. Snoussi Maddouri, V. Maergner, N. Ellouze, and H. Amiri, "IFN/ENIT database of handwritten arabic words," <i>CIFED</i> , 2002.
[ROS58]	F. Rosenblatt, "The Perceptron: a probabilistic model for information storage and organization in the brain," <i>Psychological Review</i> , vol. 65, pp. 386-408, 1958.
[XUL92]	L. Xu, A. Krzyzak, and C. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," <i>IEEE Transactions on Systems, Man, and Cybernetics</i> , vol. 22, no. 3, pp. 418-435, 1992.

LISTE DES FIGURES

<u>Figure I.1</u> Tache de classification.....	9
<u>Figure I.2</u> Vecteur de caractéristique.....	11
<u>Figure I.3</u> Les approches de RF.....	11
<u>Figure I. 4</u> Un Système de reconnaissance de formes avec ses unités.....	13
<u>Figure I.5</u> Différents modes de capture de mots tracé du mot à gauche ‘sage’ et image du mot ‘dix’ à droite	16
<u>Figure I. 6:</u> Correction de l’inclinaison des lignes à l’aide d’histogramme de projection horizontale. Les pics les plus prononcés correspondent à la rotation correcte.....	18
<u>Figure I. 7:</u> Inclinaison de l’écriture, et histogrammes de projection verticale correspondants.....	19
<u>Figure I. 8:</u> Masques qui permettent de lisser le contour des formes. Les masques complémentaires s’obtiennent par rotation.....	20
<u>Figure I.9:</u> Voyelles en arabe	26
<u>Figure I.10 :</u> Les ascendants et descendants sont entourés	27
<u>Figure I .11:</u> Un mot peut être composé de plusieurs composantes connexes (pseudo-mots).....	27
<u>Figure I.12:</u> Sous formes d’une ou de deux composantes connexes	28
<u>Figure I.13 :</u> Ligatures verticales et inversion de l’ordre du tracé.....	28
<u>Figure II.1 :</u> Le neurone de Mac Culloch-Pitts.....	33
<u>Figure II.2 :</u> Le perceptron.....	33
<u>Figure II.3 :</u> Un perceptron a une couche cachée unique.....	35
<u>Figure II.4 :</u> Un perceptron a deux couches.....	35
<u>Figure II.5 :</u> Modèles de Markov Caché.....	43
<u>Figure II.6 :</u> Modèles ergodique.....	44
<u>Figure II.7:</u> Modèles gauche droite.....	45
<u>Figure III.1:</u> Les étapes majeurs pour la réalisation d’une RF	47
<u>Figure III.2:</u> (a)Image colorée, (b) Image en noir et blanc	48
<u>Figure III.3:</u> (a) Image avant la squelettisation,(b) Image après la squelettisation	49
<u>Figure III.4:</u> (a) Image normale,(b) Image centrée	49
<u>Figure III.5:</u> L’interface principale du Matlab.....	51
<u>Figure III.6 :</u> Image chargée de la base.....	52
<u>Figure III.7 :</u> Image après la binarisation.....	52
<u>Figure III.8 :</u> Image après la squelettisation.....	52
<u>Figure III.9 :</u> Image après le centrage.....	53
<u>Figure III.10:</u> Reconnaissance.....	54
<u>Figure III.11:</u> La progression d’apprentissage du premier RNA.....	55
<u>Figure III.12:</u> La progression d’apprentissage du deuxième RNA.....	56
<u>Figure III.13:</u> La progression d’apprentissage du troisième RNA.....	57
<u>Figure III.14:</u> La progression d’apprentissage du quatrième RNA.....	58
<u>Figure III.15:</u> La progression d’apprentissage du cinquième RNA.....	59
<u>Figure III.16:</u> L’architecture du système de la classification proposée.....	60

LISTE DES TABLEAUX

<u>Tableau II.1 : L'alphabet arabe</u>	32
<u>Tableau II.2 : Points en arabe : un deux ou trois points</u>	32
<u>Tableau III.1 : Des caractères extraits de la base utilisée</u>	32
<u>Tableau III.2 : Extraction des caractéristiques</u>	32
<u>Tableau III.3 : Les résultats de performance avant et après la combinaison</u>	32

RESUME

Le présent travail porte sur une étude concernant le domaine de reconnaissance des formes (RF) appliqué sur l'écriture arabe manuscrite, il s'agit d'abord de faire une étude générale sur la RF, puis d'étudier les systèmes de classification couramment utilisés dans ce domaine, puis de faire une étude bibliographique sur les systèmes existants et les différentes recherches effectuées sur la reconnaissance d'écriture arabe. Pour enfin proposer un système intelligent de classification pour la reconnaissance des caractères arabes manuscrits à l'aide de réseaux de neurones.

MOTS-CLES : *RF, Reconnaissance des caractères arabes, caractère manuscrit, extraction des caractéristiques, classificateur, RNA, pré-traitement.*

التلخيص

هذا العمل يندرج ضمن دراسة مجال التعرف على الأشكال (RF) المطبقة على النصوص العربية المكتوبة يدويا ، في البداية سنقوم بدراسة عامة فيما يخص التعرف على الأشكال ، ثم دراسة مجموعة من أنظمة التصنيف المستخدمة عادة في هذا المجال ، ثم القيام بدراسة على الأنظمة القائمة والبحوث المختلفة في مجال التعرف على الكتابة العربية، وأخيرا اقتراح نظام ذكي للتعرف على الحروف العربية المكتوبة يدويا باستخدام شبكات الخلايا العصبية الاصطناعية.

الكلمات المفتاحية: التعرف الضوئي على الحروف، التعرف على الحروف العربية ، حرف مكتوب يدويا، واستخراج الخصائص ، المصنف ، المعالجة القبلية.

Abstract

This work is a study of Pattern Recognition (PR) field applied to the Arabic handwriting, first, we will make a general study of PR systems, then we study the classification systems commonly used in this field, then do a literature review on existing systems and various research on the Arabic writing recognition, finally, we will propose an intelligent classification system for *handwritten Arabic characters* using Artificial neural networks.

KEYWORDS: *PR, Arabic characters Recognition, handwritten character, feature extraction, classifier, ANN, pre-treatment.*