



N° d'ordre :

UNIVERSITE DE M'SILA

**FACULTE DES MATHÉMATIQUES ET DE
L'INFORMATIQUE**

Département d'Informatique

MEMOIRE

Présenté pour l'obtention du diplôme de Magistère

Spécialité : Informatique

Option : SIC

Par :

BOUMEDIENE Salah

SUJET

**Implémentation d'un système de reconnaissance automatique des
montants numériques des chèques bancaires algériens**

Soutenu publiquement le : 26/05/2011 devant le jury composé de :

Mr. BOUDERAH Brahim	Professeur	Université de M'sila	Président
CHIBANI Youcef	Professeur	USTHB	Rapporteur
CHIKOUCHE Djamel	Professeur	Université de M'sila	Examineur
BOURAHLA Mustafa	MC (A)	Université de M'sila	Examineur

Promotion : 2006 /2007

À mes parents

à toute ma famille

Remerciement

Je remercie Allah de m'avoir donné la volonté et le courage qui m'ont permis de réaliser ce travail.

Je tiens à remercier Mr Y. CHIBANI pour avoir dirigé le travail présenté dans cette thèse, et pour la grande disponibilité pendant toute la durée de ce travail.

Je remercie aussi les membres de jury pour avoir accepté de juger notre travail.

Pour terminer, je remercie toute personne ayant participé de près ou de loin pour la réalisation de ce travail.

Sommaire

Introduction Générale	01
Chapitre 1: Généralités	
1.1.Introduction	03
1.2.Structure d'un système de reconnaissance de caractère	04
1.2.1. Phase d'acquisition	04
1.2.1.1.La résolution	05
1.2.1.2.Le contraste et la luminosité	05
1.2.1.3. La qualité de support	05
1.2.1.4.L'inclinaison	05
1.2.2.Phase de Prétraitement	05
1.2.2.1. Lissage et élimination du bruit	05
1.2.2.2. Détection et redressement de l'écriture	05
1.2.2.3.Normalisation	06
1.2.2.4.Squelettisation	06
1.2.3.Segmentation	06
1.2.3.1.Segmentation primaire	06
1.2.3.2.Segmentation secondaire	07
1.2.3.3.Chevauchement des caractères	07
1.2.3.4.Caractères connectés	07
1.2.4.Extraction des caractéristiques	07
1.2.5.Classification	08
1.2.5.1.Méthodes Statistiques	09
1.2.5.2.Réseaux de neurones	10
1.2.5.3.SVM Machines à vecteurs de support	10
1.2.5.4.Méthodes structurelles	11
1.2.5.5.Combinaison de classifieurs	11

1.2.6.Post traitement.....	12
1.3.Conclusion.....	13
 Chapitre 2 : Réseaux de Neurones	
2.1.Introduction.....	14
2.2.Neurone formel.....	14
2.2.1. Apprentissage.....	16
2.3.Modèle Perceptron	17
2.3.1.Utilisation.....	14.....17
2.4.Perceptron multicouche	19
2.5.Algorithme du rétropropagation du gradient.....	20
2.6.Conclusion.....	22
 Chapitre 3 : Extraction des Caractéristiques	
3.1.Introduction.....	23
3.2.Primitives topologiques ou métriques.....	23
3.3.Primitives structurelles.....	24
3.4.primitives statistiques.....	24
3.5.Techniques Implémentées :	26
3.5.1.Histogrammes.....	26
3.5.2.Profils du contour.....	26
3.5.3. Zonage.....	27
3.5.4. Ridgelets (Crêtelettes)	27
3.5.4.1. Transformée en Ridgelets Continue.....	28
3.5.4.2. Transformée de Radon.....	30
3.5.4.3. Transformation de Radon 2D par Fourier Slice Theorem	31
3.5.4.4. Transformée en Ondelette.....	32
3.5.4.5. La Multirésolution.....	34
3.6. Sélection des Caractéristiques	35
3.6.1. Sous-ensemble optimal.....	35

3.6.2. Pertinence.....	35
3.6.3. Redondance.....	36
3.6.4. Méthodes de sélection.....	36
3.6.4.1. Filtres (Filter).....	36
3.6.4.2. Méthodes internes (embedded method)	36
3.6.4.3. Méthodes enveloppantes (wrapper)	37
3.6.5. Algorithme de sélection de caractéristiques	37
3.7. Conclusion.....	38

Chapitre 4 : Résultats Expérimentaux :

4.1. Introduction.....	39
4.2. Protocole Expérimental	39
4.2.1. Base de données USPS	39
4.2.2. Base de données des cheques bancaires algériens	40
4.2.3. Critères d'évaluation	41
4.2.4. Configuration du réseau de neurone	41
4.3. Evaluation de la base de données USPS	42
4.3.1. Evaluation des données brutes	42
4.3.2. Evaluation des descripteurs	42
4.3.2.1. Histogrammes de projection verticale et horizontale	43
4.3.2.2. Zonage	44
4.3.2.3. Profils : gauche, droit, haut et bas	45
4.3.2.4. Ridgelets	46
4.3.2.5. Combinaison par concaténation de descripteurs	47
4.4 Evaluation des descripteurs pour les données des cheques bancaires algériens	49
4.4.1. Evaluation des données brutes	49
4.4.2. Evaluation des descripteurs	49

4.3.2.1. Histogrammes de projection verticale et horizontale	49
4.3.2.2. Zonage	50
4.3.2.3. Profils : gauche, droit, haut et bas	51
4.3.2.4. Ridgelets	52
4.3.2.5. Combinaison par concaténation de descripteurs	53
4.5. Réduction des caractéristiques	54
4.5.2 Application sur les données brutes	55
4.5.3. Application sur les ridgelets.....	56
4.5.4. Concaténation des ridgelets avec les profils	57
4.5.5. Discussion	58
4.6. Conclusion.....	59
Conclusion générale	60
Annexe	61
Références bibliographiques	

Liste des figures

Figure	Titre	Page
Chapitre 1 :		
Figure 1.1.	Architecture d'un système de reconnaissance de caractère	4
Figure 1.2.	Segmentation primaire	6
Figure 1.3.	Segmentation secondaire (a) Chiffres connectés (b) Chiffres chevauchés (c) Chiffres chevauchés et connectés	7
Figure 1.4.	Hyperplan, marge et supports vecteurs	11
Figure 1.5.	Combinaison de classifieurs : approche parallèle, approche séquentielle et approche hybride.	12
Chapitre 2 :		
Figure 2.1.	Neurone Formel.	15
Figure 2.2.	Fonction de transfert sigmoïde standard	15
Figure 2.3.	Fonction de transfert tangente hyperbolique	16
Figure 2.4.	Fonction de transfert : fonction gaussienne	16
Figure 2.5.	Fonction de transfert : fonction à seuil	16
Figure 2.6.	Représentation d'un perceptron pour la reconnaissance de M classes.	17
Figure 2.7.	Échantillon linéairement séparable et non linéairement séparable.	18
Figure 2.8.	Réseau de type perceptron multicouches	19

Chapitre 3 :

	Histogramme de Projection	26
Figure 3.1.	Profils : Haut, bas, droit et gauche	27
Figure 3.2.	Division d'une image en 4 zones (4 caractéristiques)	27
Figure 3.3.	Transformée en Ridgelet	28
Figure 3.4.	Ridgelets	29
Figure 3.5.	Calcul de la transformée de Radon	30
Figure 3.6.	Transversalement à cette ridgelet c'est une ondelette.	34
Figure 3.7.		

Chapitre 4 :

	Quelques images de chiffres de la base USPS	39
Figure 4.1.	Modèles de cheques algériens	40
Figure 4.2.	Montant numérique des cheques bancaires	40
Figure 4.3.		

Liste des Tableaux

<i>Tableau</i>	<i>Titre</i>	<i>Page</i>
	Chapitre 3 :	
Tableau 3.1.	Méthodes de génération de caractéristiques selon la représentation de l'image	25
	Chapitre 4 :	
Tableau 4.1.	Matrice de confusion calculée pour les données brutes (256 caractéristiques).	42

Tableau 4.2.	Matrice de confusion calculée pour l’histogramme de projection horizontale et verticale (a) 32 caractéristiques (b) 64 caractéristiques (4 zones).	43
Tableau 4.3.	Matrice de confusion calculée pour le zonage : 16 caractéristiques (zone de 4x4)	44
Tableau 4.4.	Matrice de confusion calculée pour les profils : droit, gauche, haut et bas : 64 caractéristiques.	44
Tableau 4.5.	Matrice de confusion calculée pour les ridgelets.	45
Tableau 4.6.	Combinaison de descripteurs.	46
Tableau 4.7.	Matrice de confusion calculée pour la concaténation des différents descripteurs.	46
Tableau 4.8.	Matrice de confusion calculée pour les données brutes.	47
Tableau 4.9.	Matrice de confusion calculée pour l’histogramme de projection horizontale et verticale.	47
Tableau 4.10.	Matrice de confusion calculée pour le zonage.	48
Tableau 4.11.	Matrice de confusion calculée pour le profils : droit, gauche, haut et bas.	48
Tableau 4.12.	Matrice de confusion calculée pour les ridgelets :	49
	192 caractéristiques.	50
Tableau 4.13.	Combinaison de descripteurs.	
Tableau 4.14.	Matrice de confusion calculée pour la concaténation des ridgelets avec les profils.	50
Tableau 4.15.	Réduction des caractéristiques pour les données brutes	52
Tableau 4.16.	Réduction des caractéristiques pour les ridgelets	53

Introduction générale

Aujourd'hui, la prolifération des ordinateurs dans notre société conduit à une dématérialisation du support de l'écriture, au profit de sa forme numérique. Ainsi sont nés les concepts de société sans papier et de bibliothèque virtuelle. Paradoxalement, un nombre considérable de documents écrits est produit : lettres, enveloppes, chèques, formulaires, . . . Il n'est alors pas surprenant de constater le développement de nombreuses techniques visant à la conversion de ces documents vers la forme numérique de l'écrit. Le résultat obtenu peut alors être diffusé plus largement (via le mass médium qu'est devenu internet), être stocké sous une forme numérique de durée de vie prétendue plus longue que le papier et évidemment permettre le traitement automatique de l'information par l'intermédiaire des ordinateurs.

Le travail que nous allons présenter dans ce mémoire, concerne la reconnaissance des montants numériques des chèques bancaires algériens. Malgré le progrès impressionnant des techniques utilisées ainsi que l'explosion dans la puissance de calcul des ordinateurs, la recherche sur la reconnaissance des chiffres manuscrits avance lentement avec une performance de reconnaissance qui reste loin de celle de l'œil humain.

La solution apportée à ce problème est le développement de nouvelles techniques différentes et complémentaires à celles déjà disponibles pour améliorer la performance de chaque étape du système de reconnaissance.

Un système de reconnaissance de chiffres manuscrits est composé de trois modules :

- Prétraitements et segmentations : qui a pour but d'analyser l'image digitalisé du montant numérique du cheque bancaire, enlever le bruit présent et localiser et isoler les chiffres les uns des autres.
- Génération des caractéristiques : afin de distinguer un chiffre d'un autre, il est nécessaire de définir des caractéristiques discriminantes lors du développement du système de reconnaissance. Il a été observé que certaines primitives extraites sont non pertinentes ou redondantes pour l'étape de classification. Notre travail analyse les méthodes et techniques associées à cette phase avec plus de détail.
- Classification : c'est l'élaboration d'une règle de décision qui transforme les caractéristiques extraites d'un chiffre en appartenance à une classe.

Le présent travail est organisé en quatre chapitres : Le premier chapitre constitue une introduction à la reconnaissance de l'écriture manuscrite. Nous présentons les différents modules constituant un système standard de reconnaissance des chiffres manuscrits en mettant l'accent sur le rôle de chacun.

Le deuxième chapitre décrit le classifieur utilisé fondé sur les réseaux de neurones avec ses différentes architectures.

Le troisième chapitre est consacré à la présentation des différentes méthodes de génération des caractéristiques, et de la sélection des caractéristiques pertinentes pour la réduction du vecteur caractéristique.

Le dernier chapitre décrit les différentes expériences réalisées sur la base de données USPS, et sur la base de données des chiffres manuscrits des montants du chèque de la Banque Nationale d'Algérie.

Nous terminerons ce mémoire par une conclusion, pour rappeler les résultats obtenus ainsi que les perspectives de ce travail.

1.1. Introduction :

La reconnaissance des caractères est une des applications les plus connues dans le domaine de la reconnaissance de formes, et s'avère un problème extrêmement complexe qui n'a pas de solution satisfaisante à ce jour. Bien qu'il existe actuellement un certain nombre de systèmes de reconnaissance sur le marché, ces systèmes n'ont encore pu atteindre les qualités humaines dans le fait de reconnaître plusieurs types d'écritures imprimées ou manuscrites, mono ou omni scripteurs ([1], [2]).

Selon la façon dont l'écriture est générée, on peut distinguer deux classes de système différentes : la reconnaissance en ligne et la reconnaissance hors-ligne [3]. Dans le mode en ligne la reconnaissance se fait en temps réel (pendant l'écriture), les symboles sont reconnus au fur et à mesure qu'ils sont écrits à la main. L'écriture est représentée comme un ensemble de points dont les coordonnées sont en fonction du temps. L'acquisition de l'écrit est assurée par une tablette graphique munie d'un stylo électronique. Dans le mode hors ligne l'acquisition des données est faite après que le processus de l'écriture soit terminé ; Les informations transmises au système seront les pixels d'une image acquise. Ce mode peut être considéré comme le cas le plus général de la reconnaissance de l'écriture.

La reconnaissance hors ligne peut être scindée en deux domaines : en fonction du nombre de personnes dont l'écriture doit pouvoir être reconnue. Lorsque ce nombre est limité, le système de reconnaissance peut subir un entraînement spécifique sur base de l'écriture propre aux personnes concernées. On parle alors de reconnaissance mono scripteur. Lorsqu'un tel traitement n'est pas possible, parce que les écritures peuvent provenir d'un très grand nombre de personnes distinctes, on parle de reconnaissance multi scripteur.

En général, l'objectif de la reconnaissance de l'écriture manuscrite est de développer un système qui se rapproche le plus de l'être humain dans sa capacité de lire. Cependant ce système comprend plusieurs étapes. Le cheque bancaire doit être numérisé avec une résolution relativement haute pour garantir une bonne qualité de représentation des chiffres. La seconde étape est la segmentation des chiffres qui a pour but la localisation et l'isolation des caractères les uns des autres. Le résultat de cette étape est utilisé pour extraire les caractéristiques discriminantes de chaque image. La dernière est de décider de l'appartenance de chaque chiffre à une classe.

1.2. Structure d'un système de reconnaissance de caractère :

D'une manière générale, la reconnaissance d'un caractère englobe les phases suivantes :

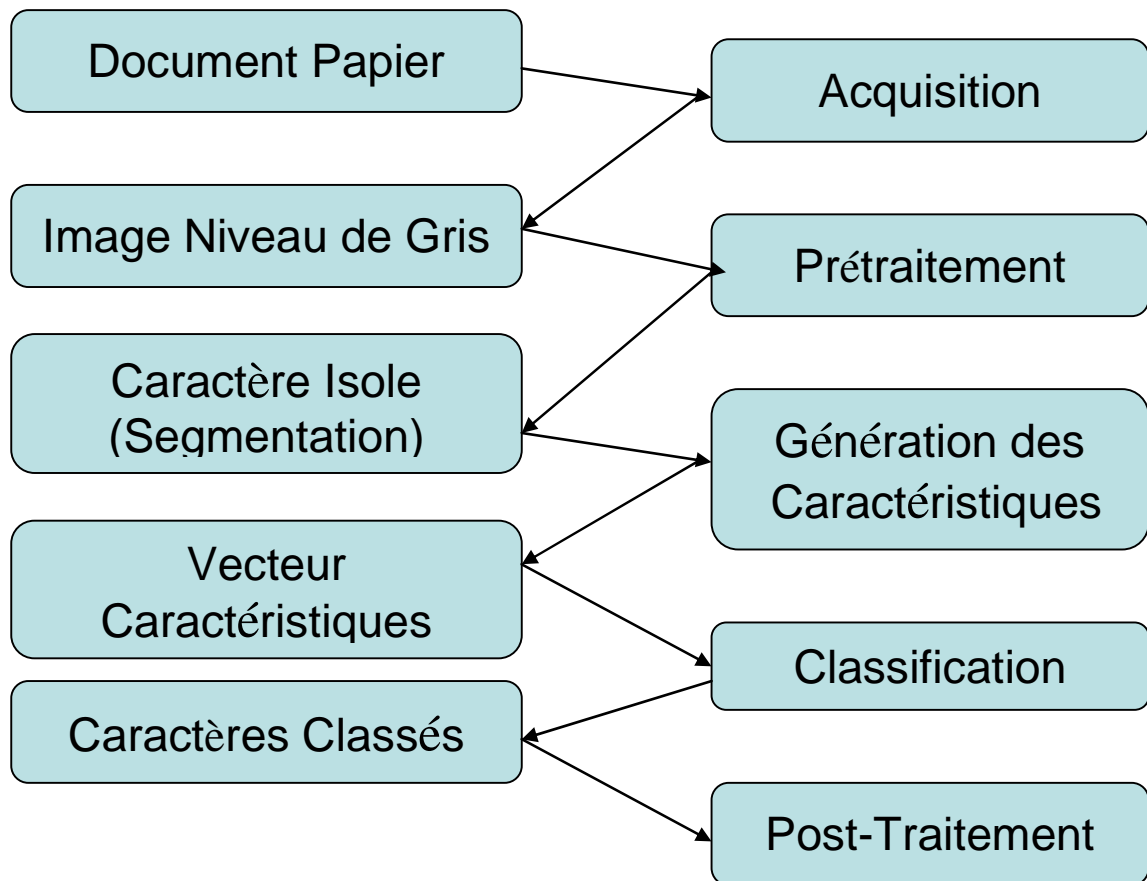


Figure 1.1. Architecture d'un système de reconnaissance de caractère

Nous décrivons dans les sections suivantes les différentes phases pour la reconnaissance des chiffres manuscrits.

1.2.1. Phase d'acquisition :

Cette phase permet la conversion du document papier sous la forme d'une image numérique en utilisant un capteur physique (scanner, camera). Cette étape est importante car elle se préoccupe de la préparation des documents à saisir, du choix et du paramétrage du matériel de saisie (scanner), ainsi que du format de stockage des images. La qualité de la saisie est tributaire de plusieurs facteurs :

1.2.1.1. La résolution : Le choix de la résolution de numérisation est fonction de la qualité du contenu en termes typographiques [4]. Le mauvais choix de la résolution, peut conduire, en sous échantillonnage, à un manque d'information (cassure de caractères, élimination de parties de graphiques, etc.), et en sur échantillonnage, à un surplus d'information, souvent traduit par la présence d'un bruit plus abondant. Un choix optimal est de numériser à 300 DPI (dot per inch ou point par pouce).

1.2.1.2. Le contraste et la luminosité : La luminosité permet de jouer sur l'éclairage du document à capturer : de plus clair à plus sombre. Le contraste permet de faire varier l'accentuation ou l'atténuation des transitions Noir/Blanc. Ces deux paramètres sont souvent corrélés entre eux et jouent un grand rôle dans la qualité de reconnaissance.

1.2.1.3. La qualité de support : peut influencer la qualité du résultat. Cependant le choix de la résolution peut permettre de corriger certains défauts.

1.2.1.4. L'inclinaison : est une source d'erreur pour l'extraction de lignes de texte et pour la modélisation de la forme des caractères.

1.2.2. Phase de prétraitement :

La conversion des documents papier vers le format numérique est une étape importante pour la diffusion, la préservation des documents et pour d'autres applications. L'image scannée contient généralement du bruit, et une suite d'opérations de prétraitement est nécessaire pour mettre l'image dans un format exploitable, réduire la quantité d'information à traiter et ne garder que les informations les plus significatives. Le bruit peut être dû aux conditions d'acquisition (éclairage, mise incorrecte du document, ...) ou encore à la qualité du document d'origine.

Durant cette phase l'image peut subir plusieurs opérations :

1.2.2.1. Lissage et élimination du bruit : Cette opération permet de mettre un effet de flou, et de réduire les pixels bruités pour les images à niveau de gris. L'effet de flou est utilisé pour l'élimination de petites taches. Pour des images binaires le lissage permet d'enlever le bruit, et de redresser les bords des caractères en procédant par le bouchage des trous interne au contour du caractère. L'opération de lissage consiste à examiner le voisinage d'un pixel et de lui attribuer la valeur 1 si le nombre de pixel noir dans cette zone est supérieur à un seuil s et zéro pour toute valeur inférieure à s [5].

1.2.2.2. Détection et redressement de l'écriture : l'inclinaison de l'écriture est souvent due à un mauvais positionnement du document sur le scanner. Cependant la détection de

l'inclinaison et le redressement de l'écriture constituent un problème qui doit être examiné durant les premières tâches d'analyse de l'image.

1.2.2.3. Normalisation : après l'opération de segmentation, les tailles des images des caractères segmentés sont variées. C'est l'opération la plus importante dans le processus de prétraitement. Elle permet de ramener les images des caractères à des tailles standard. Le but de la normalisation est de réduire la variation intra classes des caractères, et de faciliter la tâche d'extraction des caractéristiques. Cependant cette opération introduit généralement de légères déformations sur les images des chiffres. Pour cela, différentes méthodes de normalisation peuvent être utilisées pour garder l'aspect initial du chiffre [6].

1.2.2.4. Squelettisation : dans la plus part des cas, la forme à reconnaître ne dépend pas géométriquement de l'épaisseur du tracé de l'objet. Cette opération peut être considérée comme conversion de l'écriture en mode hors ligne vers l'écriture on ligne. La squelettisation sert à réduire l'épaisseur du trait d'écriture à 1 pixel, et de se ramener ainsi à une écriture linéaire.

1.2.3. Segmentation permet d'isoler les éléments textuels, mots et caractères, pour la reconnaissance. Elle se base sur des mesures de plages blanches (interlignes et inter caractères) pour faire la séparation.

1.2.3.1. Segmentation primaire : la segmentation primaire est une procédure de séparation de caractères qui se déroule en deux phases [3] :

- La détection des passages par zéro de l'histogramme de projection horizontale permet d'isoler les lignes du texte les une des autres.
- La détection des passages par zéro de l'histogramme de projection verticale calculée sur une ligne pour isoler les caractères qui composent la ligne.

L'inconvénient de cette procédure est qu'elle ne permet pas de détecter les caractères lorsqu'ils sont connectés ou en chevauchement, de plus elle est sensible à la rotation.

Figure 1.2. Segmentation primaire

1.2.3.2. Segmentation secondaire : pour pallier aux insuffisances de la segmentation primaire cette étape est nécessaire afin de détecter la présence de plusieurs caractères connectés ou en chevauchement.

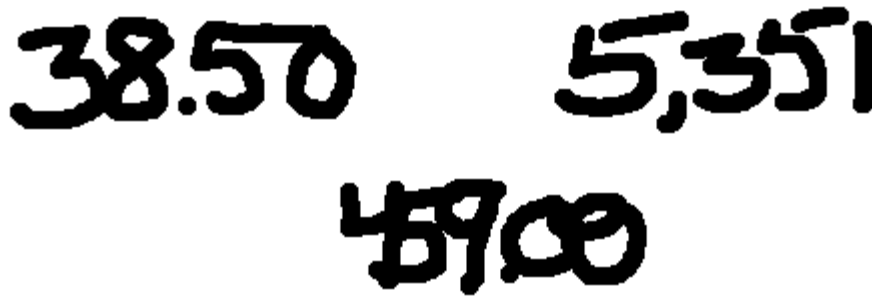


Figure 1.3. Segmentation secondaire

(a) Chiffres connectés (b) Chiffres chevauchés (c) Chiffres chevauchés et connectés

1.2.3.3. Chevauchement des caractères : ce phénomène est souvent dû à une écriture trop inclinée, et dépend généralement du soin apporté à l'écriture.

1.2.3.4. Caractères connectés

Pour extraire les chiffres connectés, il est nécessaire de trouver un pixel noir, et ainsi appliquer un algorithme pour rechercher tous les pixels connectés à ce dernier. Une fois l'ensemble de ces pixels ayant été localisés, ils sont regroupés dans un seul élément. Pour déterminer si l'élément trouvé est composé de plusieurs caractères il suffit de voir la taille et la position de cet élément. Le ratio de la hauteur sur la largeur du cadre contenant cet élément est un indicateur puissant pour déterminer si l'élément est connecté ou pas. L'isolation des chiffres connectés repose sur la recherche d'un chemin de séparation.

1.2.4. Génération des caractéristiques :

Pour la prise de décision, un système de reconnaissance n'a besoin que de l'information pertinente pour différencier un objet d'un autre. Dans ce but une étape de génération de caractéristiques est réalisée. C'est une phase critique lors de la construction d'un système de reconnaissance. L'une des raisons pour laquelle cette étape

pose un problème est qu'une grande majorité des techniques de génération de caractéristiques s'accompagne d'une perte d'information irrémédiable.

Pour un problème de classification donné, la principale qualité recherchée pour un ensemble de caractéristiques est sa faculté de rassembler les objets appartenant à une même classe dans une même partition de l'espace de représentation, tout en éloignant autant que possible les autres. Cette qualité est communément appelée *pouvoir discriminant* de l'ensemble de caractéristiques [7].

Afin d'obtenir une bonne qualité de représentation des caractères, et donc de garantir qu'une reconnaissance efficace puisse être menée, l'acquisition des caractères doit s'effectuer à relativement haute résolution. Un caractère digitalisé est alors représenté par une matrice de dimensions élevées, ce qui peut rendre complexe l'entraînement des paramètres d'un classifieur. Une procédure efficace de prétraitement des caractères, capable de générer des caractéristiques robustes et résistantes au bruit, est souvent requise pour un système de reconnaissance.

La génération des caractéristiques est une étape de grande importance: si elle est mal conçue, il sera difficile, voire impossible, d'effectuer une reconnaissance efficace. Pour la reconnaissance de caractères manuscrits, les caractéristiques recherchées doivent présenter une certaine invariance *géométrique* d'une part, et une certaine invariance *statistique* d'autre part. L'invariance géométrique signifie une tolérance vis-à-vis des opérateurs de translation, de rotation, et de changement d'échelle, tandis que l'invariance statistique signifie une tolérance vis-à-vis du bruit inévitablement présent sur chaque caractère.

1.2.5. Classification

La classification est l'élaboration d'une règle de décision qui transforme les caractéristiques représentant les formes en appartenance à une classe (passage de l'espace de codage vers l'espace de décision). Avant qu'un modèle de décision ne soit intégré dans un système de reconnaissance de l'écriture, il faut avoir procédé auparavant à deux étapes : l'étape d'*apprentissage* et l'étape de *test*.

L'étape d'apprentissage consiste à analyser les ressemblances entre les éléments d'une même classe et les dissemblances entre les éléments de classes différentes pour en déduire la

meilleure partition de l'espace des représentations. C'est une étape clé dans le système de reconnaissance [1].

Le type d'apprentissage se décline généralement en deux familles : le mode supervisé et le mode non supervisé. Si l'on dispose d'un ensemble de classes étiquetées on parlera de classification supervisée. Dans le cas contraire, on effectue une classification non supervisée appelée généralement classification automatique [8]. Dans le domaine de la reconnaissance de l'écriture, ce sont les méthodes basées sur un apprentissage supervisé qui sont le plus souvent utilisées ; et plus particulièrement pour les caractères manuscrits isolés car les classes naturelles sont connues et en nombre limité [1].

Les méthodes de classification sont regroupées dans plusieurs catégories : méthodes statistiques, Réseaux de neurones, machines à support de vecteurs, les méthodes structurelles et la combinaison de plusieurs méthodes de classification. Les méthodes statistiques, les réseaux de neurones et les machines à support de vecteurs nécessitent un vecteur de caractéristique de taille fixe en entrée [5].

1.2.5.1. Méthodes Statistiques :

Ces méthodes reposent essentiellement sur la règle de Bayes qui permet d'évaluer des probabilités a posteriori (c'est-à-dire après l'observation effective de certaines grandeurs) connaissant les distributions de probabilité conditionnelles a priori (c'est-à-dire indépendantes de toute contrainte sur les variables observées).

La règle de décision bayésienne est une théorie clé en classification qui permet d'estimer la probabilité à posteriori à partir de la probabilité conditionnelle et d'émettre un vote d'appartenance de la forme traitée. Elle s'écrit :

$$p(c_i / x) = p(x / c_i) * p(c_i) / p(x)$$

Pour K classes, la décision bayésienne cherche la classe c_i qui maximise la probabilité à posteriori. Elle s'écrit :

$$c(x) = \arg \max p(c_i / x)$$

Le terme $p(c_i)$ représente la probabilité a priori de la classe c_i .

Les méthodes statistiques ont besoin d'un nombre élevé d'exemples afin de réaliser un apprentissage correct des lois de probabilité des différentes classes. L'étude de leur répartition dans un espace métrique et la caractérisation statistique des classes permet de prendre une décision du type : plus forte probabilité d'appartenance à une classe. Cette approche bénéficie des méthodes d'apprentissage automatique qui s'appuient sur des bases théoriques connues telles que les méthodes paramétriques et les méthodes non paramétriques. ([5], [9], [10], [11])

1.2.5.2. Réseaux de neurones :

Un **réseau de neurones artificiel** est un modèle de calcul dont la conception est très schématiquement inspiré du fonctionnement des neurones biologiques. C'est un système composé de plusieurs unités de calcul simples fonctionnant en parallèle, dont la fonction est déterminée par la structure du réseau, la solidité des connexions, et l'opération effectuée par les éléments ou noeuds. Le principe de fonctionnement est le suivant. Nous disposons initialement d'une base de connaissances constituée de couples de données entrées / sorties et nous souhaitons utiliser cette base de données pour "entraîner" un algorithme à reproduire les associations constatées entre les entrées et les sorties de l'échantillon. Cette méthode sera détaillée dans le chapitre suivant [12].

1.2.5.3. SVM Machines à vecteurs de support (Séparateur à vaste marge):

C'est une méthode de classification binaire par apprentissage supervisé, elle fut introduite par Vapnik en 1995. Cette méthode est donc une alternative récente pour la classification. Cette méthode repose sur l'existence d'un classificateur linéaire dans un espace approprié. Puisque c'est un problème de classification à deux classes, cette méthode fait appel à un jeu de données d'apprentissage pour apprendre les paramètres du modèle. Elle est basée sur l'utilisation de fonction dite noyau (kernel) qui permettent une séparation optimale des données [13].

La plupart des techniques d'apprentissage possèdent un (trop) grand nombre de paramètres d'apprentissage à fixer par l'utilisateur (structure d'un réseau de neurones, coefficient de mise à jour du gradient, . . .). De plus, avec ces méthodes, le nombre de paramètres à calculer par l'algorithme d'apprentissage est en relation linéaire, voir exponentielle, avec la dimension de l'espace d'entrée. La formulation élégante de SVM laisse très peu de place aux paramètres utilisateurs et le nombre de paramètres est linéaire en la taille des données d'entraînement (apprentissage). SVM est donc une méthode de

classification particulièrement bien adaptée pour traiter des données de très haute dimension telles que les textes et les images [14].

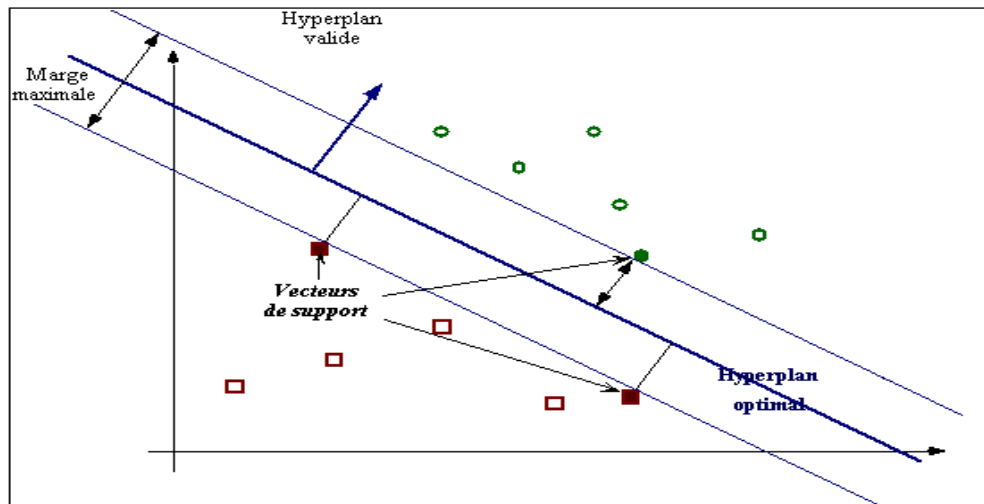


Figure 1.4. Hyperplan, marge et supports vecteurs

1.2.5.4. Méthodes structurelles :

Les méthodes de reconnaissances structurelles sont utilisées généralement dans la reconnaissance en ligne des caractères que dans la reconnaissance hors ligne. Cette approche cherche à structurer l'information en décrivant l'organisation topologique (la structure) de la forme à partir de ses composantes les plus élémentaires. La différence principale entre ces méthodes et les méthodes statistiques est que ces caractéristiques sont des formes élémentaires (les primitives sont de type topologiques) et non pas des mesures. On distingue plusieurs techniques telles que les structures de graphes, les structures syntaxiques, le calcul de distance d'édition entre deux chaînes et la programmation dynamique, les méthodes de tests, la comparaison de chaînes [2], [10].

1.2.5.5. Combinaison de classifieurs :

La combinaison de classifieurs a été utilisée avec succès en reconnaissance de formes et en particulier de caractères manuscrits afin de bénéficier de leur complémentarité. Il existe trois schémas de combinaisons de classifieurs :

- La combinaison parallèle dans laquelle le caractère à reconnaître est présentée à plusieurs classifieurs indépendants dont les sorties sont combinées pour donner la décision finale.
- La combinaison séquentielle de classifieur où les classifieurs sont disposés en niveaux successifs de décision permettant de réduire progressivement le nombre de classes possibles.
- Les approches hybrides consistent à combiner les architectures séquentielles et parallèles. Ce type d'approche est généralement dédié à un problème précis et est difficilement généralisable.

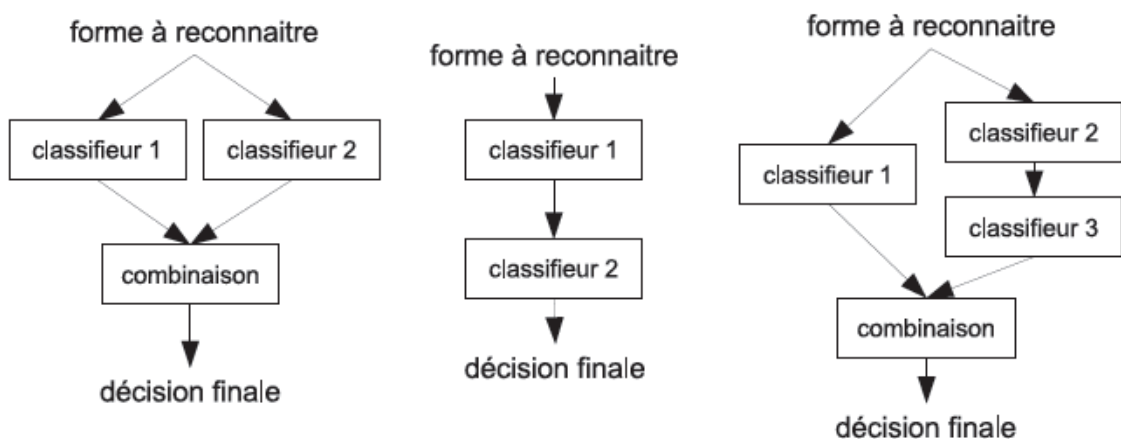


Figure 1.5. Combinaison de classifieurs : approche parallèle, approche séquentielle et approche hybride.

Parmi ces trois approches, la combinaison parallèle est la plus utilisée car contrairement aux deux autres approches, elle ne nécessite pas de connaître précisément le comportement des classifieurs. Les approches parallèles sont ainsi plus facilement généralisables et plus simples à mettre en oeuvre puisqu'elles nécessitent simplement de développer une étape de combinaison des sorties [15].

1.2.6. Post traitement :

Le but principal de cette étape est d'améliorer le taux de reconnaissance. *Le post-traitement* est effectué quand le processus de reconnaissance aboutit à la génération d'une liste de lettres, de chiffres, ou de mots possibles. Comme la classification peut aboutir à plusieurs candidats possibles, 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. Cette étape permet de réduire considérablement les erreurs [16].

1.3. Conclusion :

Ce chapitre nous a permis d'avoir un aperçu global sur la structure d'un système de reconnaissance hors ligne de caractères manuscrits à savoir : les prétraitements, la segmentation, l'extraction des caractéristiques et les différents types de classifieurs, ainsi que la combinaison de ces derniers.

Dans le cadre de notre projet, nous nous sommes intéressés à l'étape de l'extraction des caractéristiques qui a pour but de maximiser la performance du système de reconnaissance. Nous aborderons en détails les réseaux de neurones au chapitre suivant.

2.1. Introduction :

Aujourd'hui, les réseaux de neurones sont solidement implantés dans diverses industries : dans les milieux financiers pour la prédiction des fluctuations de marché ; en pharmaceutique pour analyser les relations entre les diverses molécules organiques ; dans le domaine bancaire pour la détection de fraudes sur les cartes de crédit ; dans les départements de marketing des compagnies de diverses industries pour prévoir les comportements des consommateurs ; et dans notre cas la reconnaissance des chiffres manuscrits [17].

Les réseaux de neurones sont composés d'éléments simples (ou neurones) fonctionnant en parallèles. Ces éléments ont été fortement inspirés du système nerveux biologique ; et comme dans la nature, le fonctionnement du réseau est fortement influencé par les connections des éléments entre eux. Les non linéarités présentes dans ces systèmes connexionnistes, ainsi que l'apprentissage discriminant qu'ils subissent, rendent ces derniers particulièrement bien adaptés aux tâches de classification.

« Un réseau de neurones est un processus distribué de manière massivement parallèle, qui a une propension naturelle à mémoriser des connaissances de façon expérimentale et de les rendre disponibles pour utilisation. Il ressemble au cerveau en deux points: [3]

1. la connaissance est acquise au travers d'un processus d'apprentissage;
2. les poids des connections entre les neurones sont utilisés pour mémoriser la connaissance. »

2.2. Neurone formel :

Le modèle de neurone formel est un modèle mathématique très simple dérivé d'une analyse de la réalité biologique.

Dans ce modèle, le traitement consiste à affecter à sa sortie α_i le résultat d'une fonction d'activation F de la somme pondérée de ses entrées :

$$\alpha_i = f(\sum w_j * x_j + b) \quad (2.1)$$

α_i : sortie du neurone.

$x_1 \dots x_n$: désignent les entrées du neurones, qui peuvent être des sorties d'autres neurones.

Elles peuvent être discrétisées grâce à un ensemble tel que $\{0,1\}$ ou bien être des valeurs réelles.

$w_1 \dots w_n$: poids synaptiques associées aux entrées x_j et qui déterminent la contribution de chaque entrée sur la sortie du neurone. Le but d'un algorithme d'apprentissage pour

un réseau de neurones est de déterminer le meilleur ensemble de poids pour le problème en considération. Trouver l'ensemble optimal est souvent un compromis entre temps de calcul, minimisation de l'erreur et conservation de la capacité de généralisation pour le réseau.

b : Le biais qui permet d'ajouter de la flexibilité au réseau en permettant de varier le seuil de déclenchement du neurone. Il est souvent négligé puisqu'il n'effectue pas un changement dans le fonctionnement du neurone.

F : fonction de transfert ; intègre toutes les entrées et le biais et calcul la sortie du neurone selon une fonction d'activation ou de seuillage. Elle sert à introduire une non linéarité dans le fonctionnement du neurone.

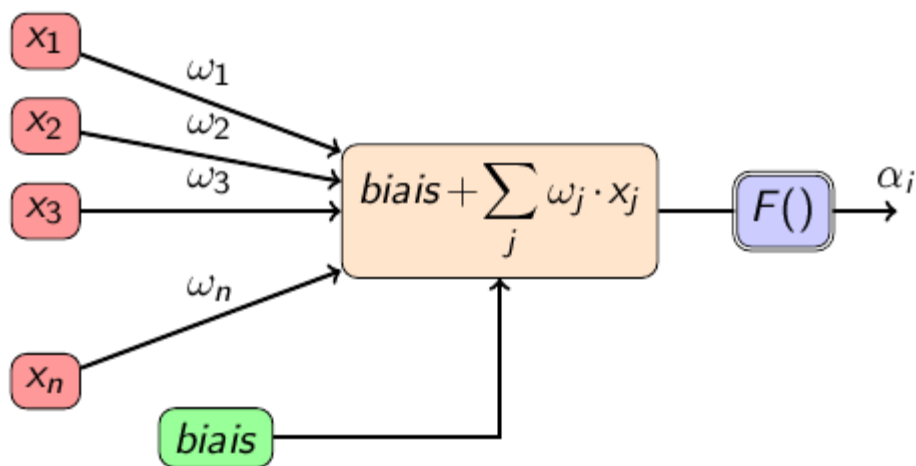


Figure 2.1. Neurone Formel.

L'entrée d'un neurone contient le biais et la somme des ses poids.

La sortie d'un neurone dépend de l'entrée du neurone et de sa fonction de transfert F .

Le tableau suivant présente les fonctions de transfert couramment utilisées.

➤ Sigmoïde standard (appelée fonction logistique) :

$$F(x) = \frac{e^x}{e^x + 1}$$

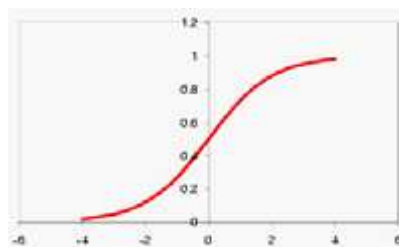


Figure 2.2. Fonction de transfert sigmoïde standard

- Tangente hyperbolique :

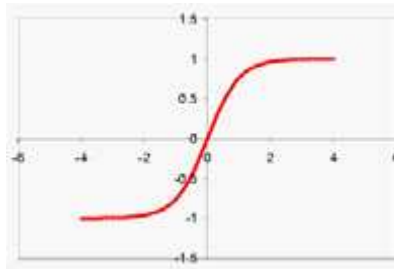


Figure 2.3. Fonction de transfert tangente hyperbolique

- Fonction Gaussienne :

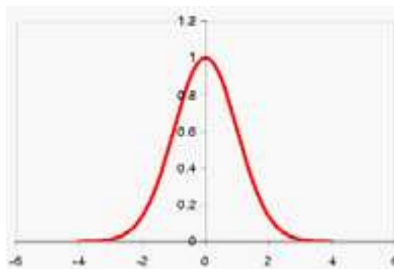


Figure 2.4. Fonction de transfert : fonction gaussienne

- Fonction à seuil :

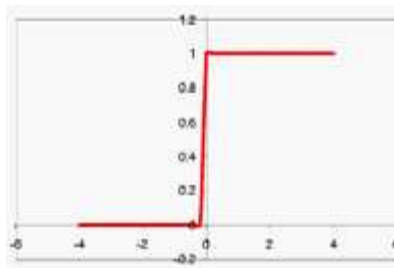


Figure 2.5. Fonction de transfert : fonction à seuil

2.2.1. Apprentissage :

Une caractéristique des réseaux de neurones est leur capacité à apprendre (par exemple à reconnaître une lettre, un son...). Mais cette connaissance n'est pas acquise dès le départ. La plupart des réseaux de neurones apprennent par l'exemple en suivant un algorithme d'apprentissage. Il y a deux algorithmes principaux : l'apprentissage supervisé et l'apprentissage non supervisé.

Lors d'un **apprentissage supervisé**, les résultats corrects (c'est-à-dire les valeurs que l'on désire que le réseau obtienne en sortie) sont fournis au réseau, si bien que celui-ci peut ajuster ses poids de connexions pour les obtenir. Après l'apprentissage, le réseau est testé en lui

donnant seulement les valeurs d'entrée mais pas les sorties désirées, et en regardant si le résultat obtenu est proche du résultat désiré. Le principe d'apprentissage consiste en fait à trouver les poids qui minimisent l'erreur entre la sortie obtenue et la sortie désirée.

Lors d'un **apprentissage non supervisé**, on ne fournit pas au réseau les sorties que l'on désire obtenir. On le laisse évoluer librement jusqu'à ce qu'il se stabilise.

2.3. Le modèle Perceptron :

C'est un des premiers réseaux de neurones, conçus en 1958 par Rosenblatt. Il est linéaire et monocouche. Il est inspiré du système visuel. La première couche (d'entrée) représente la rétine. Les neurones de la couche suivante (unique, d'où le qualificatif de monocouche) sont les cellules d'association, et la couche finale les cellules de décision. Les sorties des neurones ne peuvent prendre que deux états (-1 et 1 ou 0 et 1). La couche d'entrée contient N entrées, et la couche de sortie est composée de M sorties.

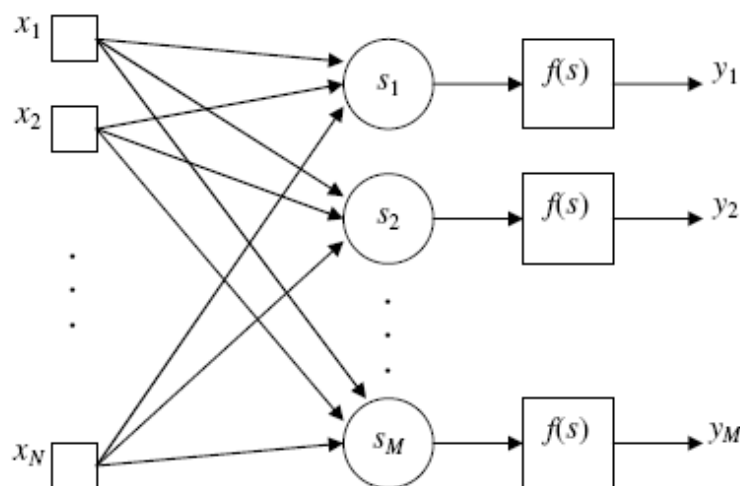


Figure 2.6. Représentation d'un perceptron pour la reconnaissance de M classes.

2.3.1. Utilisation :

L'algorithme consiste à présenter au perceptron un certain nombre de fois les vecteurs d'entrée. Lorsqu'un vecteur est présenté au perceptron, les poids sont modifiés à chaque fois que la sortie est différente de la valeur de la classe. L'apprentissage est terminé une fois que tous les vecteurs ont été présentés sans modification des poids.

Il convient de remarquer que les poids ne sont modifiés que si la sortie calculée est différente de la sortie attendue.

La règle de modification des poids est la suivante :

$$W_i = W_i + (C - Y) * K \quad (2.2)$$

- Où :
- W_i représente le poids de l'entrée i
- C la classe
- Y la sortie calculée
- K une constante positive.

La méthode de la descente du gradient permet de minimiser pour un ensemble de N exemple donné, l'erreur quadratique calculée par la formule suivante :

$$E = \frac{1}{2} \sum_{k=1}^N (y_k - s_k)^2 \quad (2.3)$$

Tels que :

Y_k : sortie attendue (classe)

s_k : sortie calculée pour le k ième exemple.

Le gradient des erreurs est calculé en dérivant l'erreur quadratique E par rapport aux poids, ce qui nécessite une fonction d'activation dérivable.

Ce modèle exige un échantillon linéairement séparable, ce qui est rarement le cas des données qu'on veut apprendre au réseau. Dans le cas où l'échantillon n'est pas linéairement séparable l'algorithme ne s'arrêtera pas. C'est ainsi que les réseaux multicouches sont apparus pour résoudre ce problème.

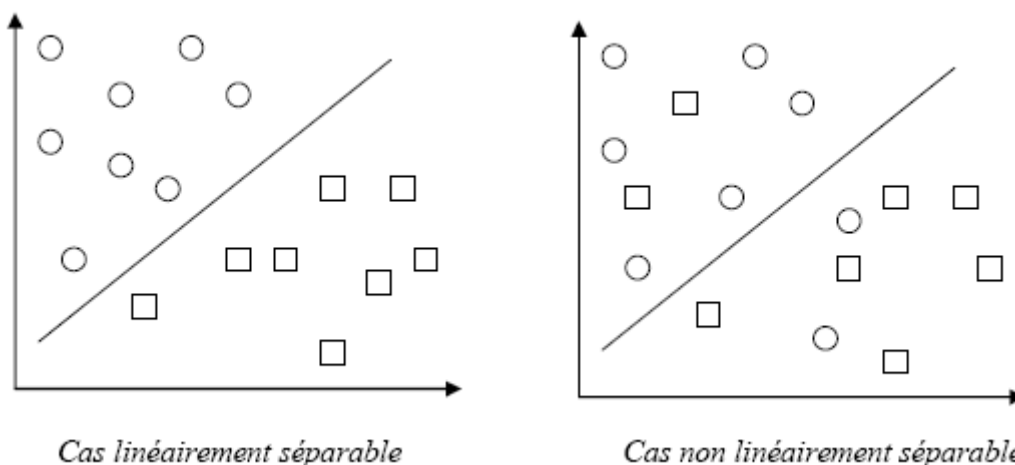


Figure 2.7. Échantillon linéairement séparable et non linéairement séparable.

Pour améliorer l'apprentissage du perceptron, Widrow et Hoff ont proposé une nouvelle règle d'apprentissage plus proche des mécanismes d'optimisation.

Il s'agit non seulement de trouver la frontière séparant les classes lorsqu'elle existe, mais aussi de trouver la meilleure solution qui minimise l'erreur lorsque cette séparation linéaire n'est pas possible. Cette règle consiste à modifier les poids après chaque exemple et non pas après que tous les exemples aient défilé. Ceci va donc minimiser l'erreur de manière précise et ce sur chaque exemple.

2.4. Perceptron multicouche :

C'est une amélioration du perceptron monocouche, comprenant une ou plusieurs couches intermédiaires dites couches cachées entre l'entrée du réseau et sa sortie. Chaque neurone n'est relié qu'aux neurones des couches directement suivante ou précédente, mais à tous les neurones de ces couches.

Un perceptron multicouche peut posséder un nombre de couche quelconques, et un nombre de neurone (ou d'entrées) par couche également quelconque. Les neurones sont reliés entre eux par des connections pondérées. Ce sont les poids de ces connections qui gouvernent le fonctionnement du réseau et programment une application de l'espace des entrées vers l'espace des sorties à l'aide d'une transformation non linéaire. Pour calculer ces poids un algorithme dit de rétropropagation du gradient est utilisé.

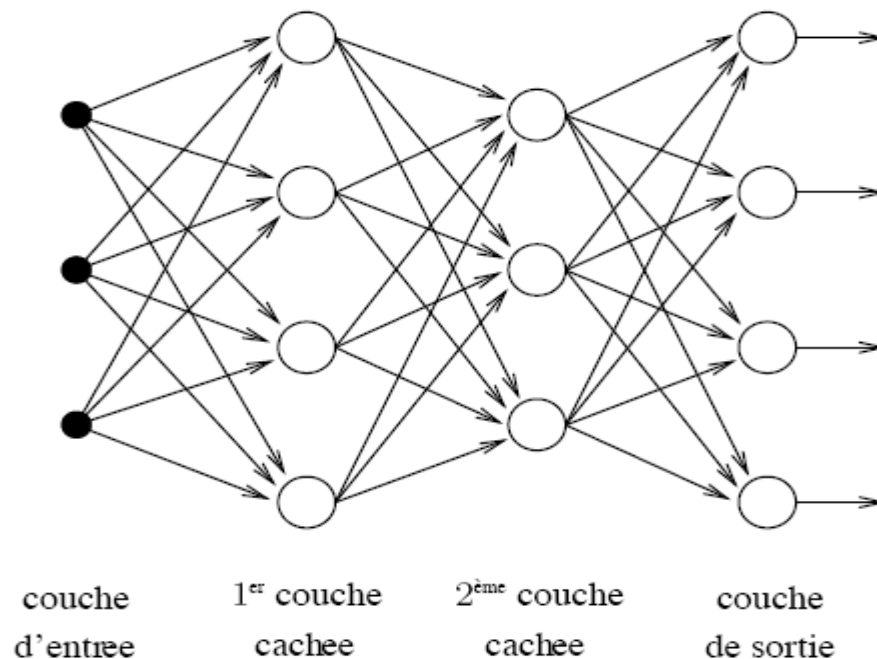


Figure 2.8. Réseau de type perceptron multicouches

2.5. Algorithme du rétropropagation du gradient :

L'algorithme consiste à déterminer l'erreur commise par chaque neurone puis à modifier la valeur des poids pour minimiser cette erreur. Il faut effectuer des retropropagations jusqu'à ce que l'erreur quadratique moyenne devienne inférieure à un certain seuil.

L'erreur quadratique est définie comme la somme sur les sorties des carrés des différences entre l'entrée obtenue et l'entrée désirée soit

$$E = \frac{1}{2} \sum (s - \gamma)^2 \quad (2.4)$$

On effectue tout d'abord une propagation de l'entrée à travers le réseau afin de déterminer l'erreur commise par chaque neurone de sortie, puis l'algorithme consiste à rétropropager cette erreur depuis les sorties jusqu'à l'entrée.

Cet algorithme nécessite une fonction continue, non linéaire et différentiable comme fonction de transfert du neurone.

Pour cela il faut modifier les poids du réseau afin d'obtenir la sortie désirée en suivant l'algorithme ci-dessous :

1. initialisation des poids à des valeurs aléatoires.
2. sélection d'un exemple d'apprentissage $(E, C)_1$ dans la base d'apprentissage.
3. présentation de la forme d'entrée (E) sur la couche d'entrée du réseau ;
4. calculer les sorties observées y_i en propageant les entrées vers l'avant ;
5. ajuster les poids en rétropropageant l'erreur observée :

$$\mathbf{w}_{ji}(\mathbf{n}) = \mathbf{w}_{ji}(\mathbf{n} - 1) + \Delta \mathbf{w}_{ji}(\mathbf{n}) = \mathbf{w}_{ji}(\mathbf{n} - 1) + \eta \delta_j(\mathbf{n}) y_i(\mathbf{n}) \quad (2.5)$$

Où le « gradient local » est défini par :

$$\delta_j(\mathbf{n}) = \begin{cases} e_j(\mathbf{n}) \cdot y_j(\mathbf{n}) \cdot [1 - y_j(\mathbf{n})] & \text{si } j \text{ appartient à la couche de sortie} \\ y_j(\mathbf{n}) [1 - y_j(\mathbf{n})] \sum \delta_k(\mathbf{n}) \cdot \mathbf{w}_{kj}(\mathbf{n}) & \text{si } j \text{ appartient à une couche cachée} \end{cases}$$

Avec $0 \leq \eta \leq 1$ représentant le taux d'apprentissage et $y_i(\mathbf{n})$ représentant soit la sortie du neurone i sur la couche précédente, si celui-ci existe, soit l'entrée i autrement.

6. Répéter les étapes 2 à 5 jusqu'à un nombre maximum d'itérations ou jusqu'à ce que la racine de l'erreur quadratique moyenne (EQM) soit inférieure à un certain seuil.
 - \mathbf{n} : représente la donnée d'entraînement c'est-à-dire le couple vecteur (entrée, sortie désirée).

- w_{ji} : poids de la connexion entre le neurone i de la couche précédente et le neurone j de la couche courante.

Lorsque le perceptron multicouches est utilisé en tant que classificateur, la couche de sortie comporte un neurone par classe. En phase d'apprentissage, les valeurs désirées pour ces sorties sont alors:

□ 1 pour le neurone qui correspond à la classe de l'objet présenté à l'entrée du réseau;

□ 0 pour tous les autres neurones.

Il existe plusieurs critères d'arrêts. Ces critères peuvent être combinés entre eux. Le premier critère est basé sur l'amplitude du gradient de la fonction d'activation puisque par définition le gradient sera à zéro au minimum. Le second critère d'arrêt est de fixer un seuil que l'erreur quadratique ne doit pas dépasser. Le troisième critère consiste à déterminer un nombre fixe de cycles à atteindre.

Malgré ses succès, la méthode de rétropropagation du gradient présente la difficulté dans la détermination du nombre de couches cachées, le nombre de neurones dans chacune des couches et les poids des connexions entre les différentes couches (pente de la sigmoïde, taux d'apprentissage...). De ce fait, pour une application donnée, la construction du classifieur de type MLP utilise des règles empiriques et nécessite un certain nombre d'essais afin d'obtenir des performances de généralisation intéressantes. Un autre inconvénient majeur réside dans le manque d'interprétabilité (a posteriori) des connaissances réparties dans les liaisons inter neurones complexes ainsi que la grande quantité nécessaire de données pour leur apprentissage.

2.6. Conclusion :

Ce chapitre était consacré à la présentation du classifieur basé sur les réseaux de neurones.

Nous avons décrit le neurone formel, le perceptron monocouche, le perceptron multicouches, les étapes de l'apprentissage du réseau et les différents critères d'arrêt de l'algorithme de rétropropagation.

L'élément essentiel des réseaux de neurones est qu'ils peuvent capter les dépendances non linéaires de haut niveau entre les variables explicatives (vecteur caractéristiques), ce qui est possible grâce à la présence d'une transformation, elle-même non linéaire, dans le calcul de la valeur prédite. Les réseaux de neurones sont de puissants outils de modélisation et de prédiction.

Un réseau de neurones ne peut être performant que si les données présentées à ce dernier sont pertinentes. Nous abordons dans le chapitre suivant, les différentes méthodes d'extraction des caractéristiques.

3.1. Introduction

Le savoir faire accumulé grâce aux travaux avancés dans le domaine de recherche a montré que la fiabilité des systèmes de reconnaissance de caractères manuscrits est conditionnée par celle de l'étape de génération de caractéristiques [2]. En effet pour la prise de décision, un système de reconnaissance n'a besoin que de l'information pertinente pour différencier un objet d'un autre. Dans ce but une étape de génération de caractéristiques est réalisée. C'est une phase critique lors de la construction d'un système de reconnaissance. L'une des raisons pour laquelle cette étape pose un problème est qu'une grande majorité des techniques de génération s'accompagne d'une perte d'information irrémédiable. De ce fait, il faut effectuer un compromis entre quantité et qualité de l'information.

Pour un problème de classification donné, la principale qualité recherchée pour un ensemble de caractéristiques est sa faculté de rassembler les objets appartenant à une même classe dans une même partition de l'espace de représentation, tout en éloignant autant que possible les autres. Cette qualité est communément appelée *pouvoir discriminant* de l'ensemble de caractéristiques.

En effet l'objectif commun de toutes les caractéristiques est de caractériser au mieux la forme des caractères afin de pouvoir distinguer si deux images appartiennent à deux classes différentes ou à la même classe, c'est-à-dire qu'elles doivent diminuer la variabilité intra classe et augmenter la variabilité interclasses. [7]

Suivant les applications et techniques utilisées pour le système de reconnaissance, les caractéristiques extraites peuvent être très différentes.

Dans la littérature, les caractéristiques sont classifiées de plusieurs façons différentes. Une première distinction peut être effectuée entre les caractéristiques globales et les caractéristiques locales. Les caractéristiques globales cherchent à représenter au mieux la forme générale d'un caractère et sont donc calculées sur des images relativement grandes (ex : transformée de Fourier et transformée de Hough). Les caractéristiques locales sont calculées lors d'un parcours des pixels de l'image avec un pas d'analyse qui dépend de la modélisation, du type de caractéristique et de la taille de l'image. Une seconde distinction est effectuée entre les caractéristiques topologiques, structurelles ou statistiques :

3.2. Caractéristiques topologiques ou métriques : elles consistent à compter dans une forme le nombre de trous, évaluer les concavités, mesurer des pentes et autres paramètres de courbures et évaluer des orientations, mesurer la longueur et l'épaisseur des traits, détecter les croisements et les jonctions des traits, mesurer les surfaces et les périmètres, ... [18]

3.3. Caractéristiques structurelles : elles ressemblent beaucoup aux caractéristiques topologiques. La différence est qu'elles sont généralement extraites non pas de l'image brute, mais à partir du squelette ou du contour de la forme. Ainsi, on ne parle plus de trous, mais de boucles ou de cycles dans une représentation filiforme du caractère [18].

3.4. Caractéristiques statistiques : l'histogramme, qui représente le nombre de pixels sur chaque ligne ou colonne de l'image, en est un exemple classique et simple à calculer. On peut citer également l'approche basée sur un moyennage des pixels situés à l'intérieur d'un masque rectangulaire : on construit une matrice de masques recouvrant la totalité de la forme qui permet une représentation statistique des valeurs correspondant à chaque masque [18].

Les caractéristiques structurelles et statistiques sont complémentaires et doivent être utilisées en même temps, car les caractéristiques extraites sont différentes.

La génération des caractéristiques est une étape cruciale dans un système de reconnaissance de caractère. Cependant, les autres étapes doivent être optimisées pour avoir un système optimisé car ces étapes ne sont pas indépendantes. Le choix des méthodes de génération des caractéristiques limite ou oblige à prendre en considération des opérations spécifiques dans la phase de prétraitement. Trier et al. [19] ont réalisé une taxonomie des méthodes de génération selon le format de représentation de l'image que nous présentons dans le tableau.

Le tableau 3.1 montre quatre formats de représentation d'image arrangés en colonnes qui sont : image de gris, image binaire, le contour et le squelette. Pour chaque format, des méthodes de génération de caractéristiques raisonnables sont indiquées. Notons que le choix des caractéristiques et du format de représentation ne sont pas totalement dépendants ni tout à fait indépendants.

Pour la reconnaissance de caractères manuscrits, les caractéristiques recherchées doivent présenter une certaine invariance *géométrique* d'une part, et une certaine invariance *statistique* d'autre part. L'invariance géométrique signifie une tolérance vis-à-vis des opérateurs de translation, de rotation, et de changement d'échelle, tandis que l'invariance statistique signifie une tolérance vis-à-vis du bruit inévitablement présent sur chaque caractère.

caractéristiques	Image en niveau de gris	Image binaire	Contour	Squelette
Appariement	X	X		X
-----	-----	-----		-----
Motifs déformables	X			X
-----	-----	-----		-----
Transformations unitaires	X	X		
-----	-----	-----		-----
Description graphique				X
-----	-----	-----		-----
Histogrammes de projection		X		
-----	-----	-----	-----	
Profil du contour			X	
-----	-----	-----	-----	-----
Zonage	X	X	X	X
-----	-----	-----	-----	-----
Moments géométriques	X	X		
-----	-----	-----		
Moments de Zernike	X	X		
-----	-----	-----	-----	
Spline			X	
-----			-----	-----
Descripteurs de Fourier			X	X

Tableau 3.1. Méthodes de génération de caractéristiques selon la représentation de l'image, d'après [2].

3.5. Techniques implémentées

3.5.1. Histogrammes de projection

Cette technique est très efficace pour la segmentation des caractères, des mots et des lignes du texte. Appliqués sur les images binaires, cette technique consiste à balayer l'image binaire et compter le nombre de pixels qui se trouvent dans la même direction.

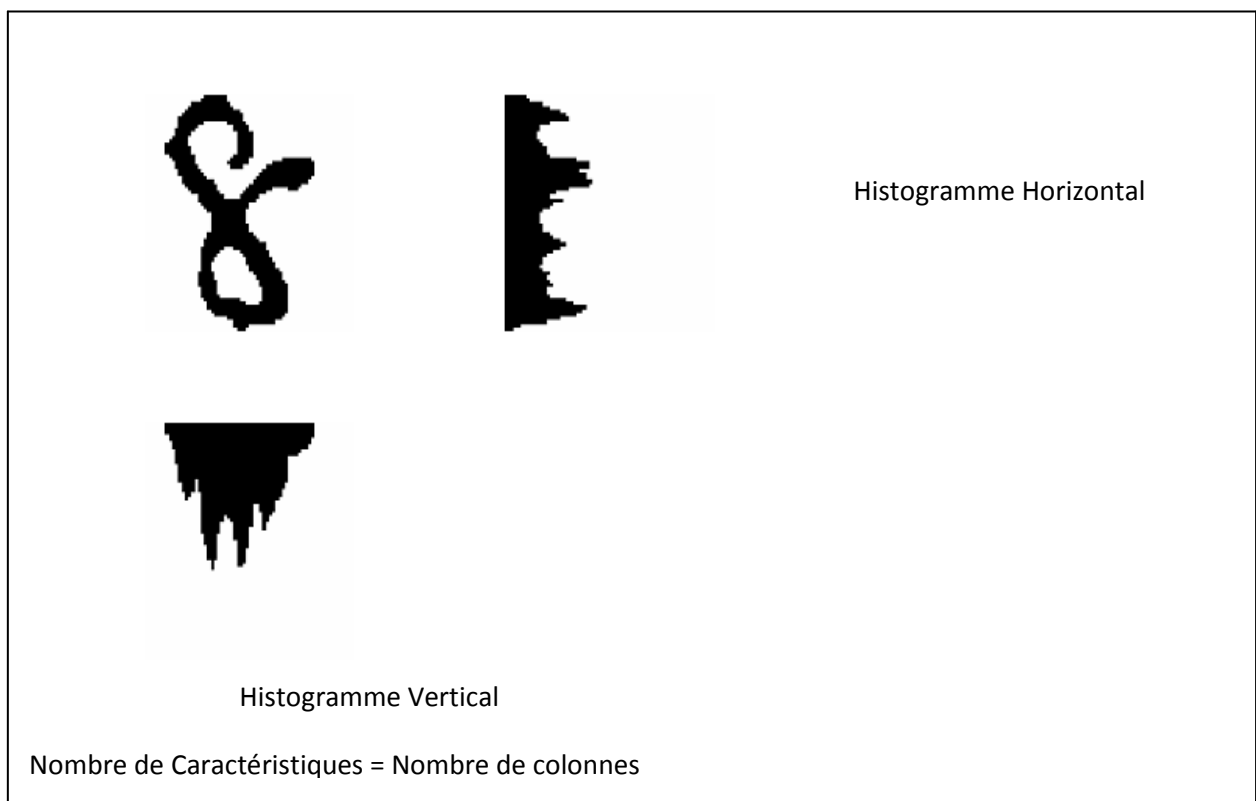


Figure 3.1. Histogramme de Projection

L'histogramme de projection est très sensible à la rotation et à la variation du style d'écriture.

3.5.2. Profils du contour

Permettent d'extraire le contour d'un caractère. On peut avoir le profil haut, bas, droit et gauche. Pour calculer le profil gauche il suffit de calculer la distance entre le bord gauche de l'image jusqu'à la rencontre du premier pixel appartenant au caractère, et ceci pour toutes les lignes de la matrice représentant le caractère.

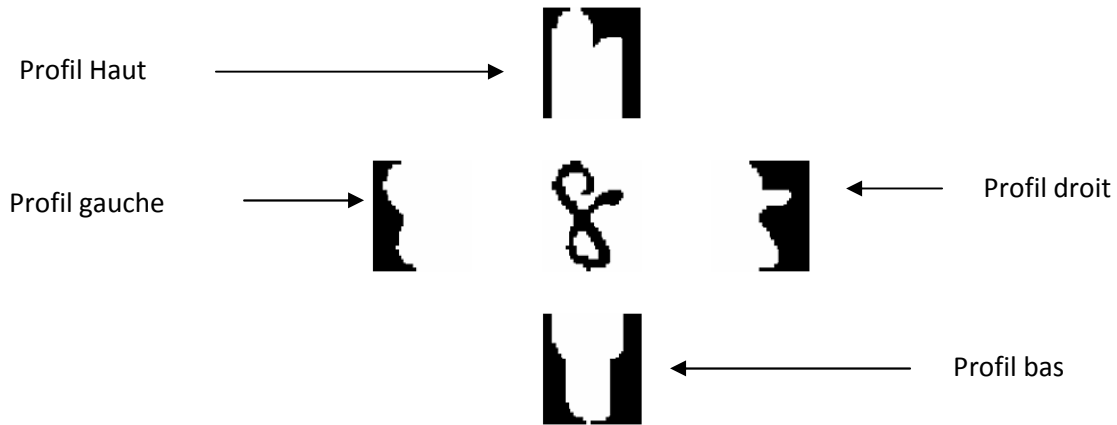


Figure 3.2. Profils : Haut, bas, droit et gauche

3.5.3. Zonage (Zoning)

La matrice représentant l'image du caractère (chiffre) est divisée en plusieurs zones de taille $p * q$. Pour chaque zone, on calcule la moyenne ou le pourcentage de point en niveau de gris (nombre de pixel appartenant au caractère dans une représentation binaire).

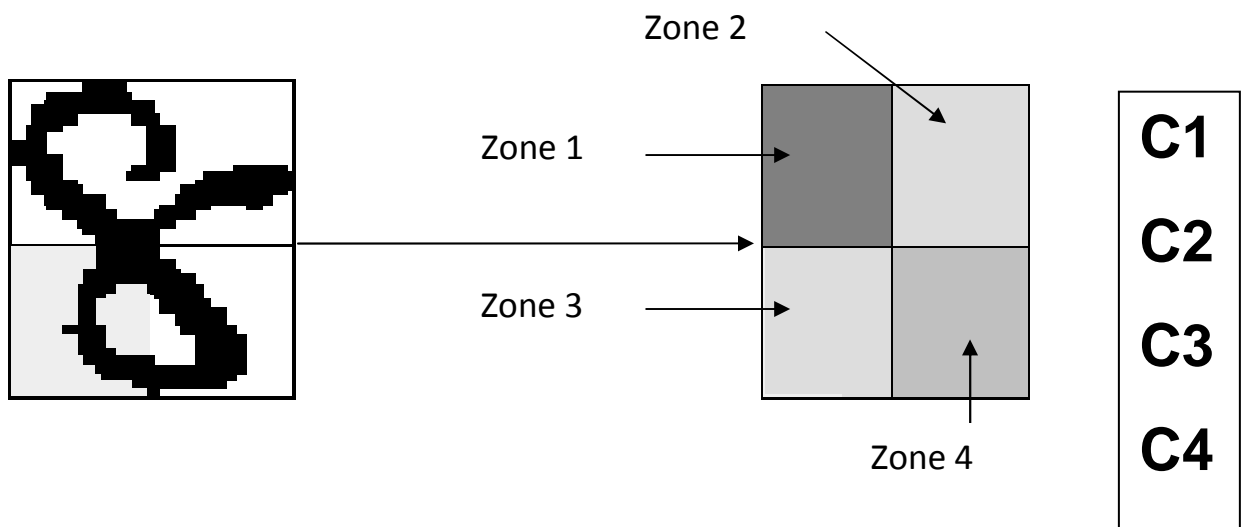


Figure 3.3. Division d'une image en 4 zones (4 caractéristiques)

La taille du vecteur caractéristique obtenu est $p*q$ (le nombre de zones).

3.5.4. Ridgelets (« Crêtelettes »)[20][21][22][23]

La transformée en ondelette est un outil très efficace dans l'analyse et le débruitage de signaux non stationnaires présentant des discontinuités ponctuelles. Néanmoins, les discontinuités présentes dans les images naturelles sont principalement générées par les

contours et par conséquent sont distribuées le long d'une courbe lisse dans le plan spatial. Pour cette catégorie de signaux, les ondelettes s'avèrent par nature sous optimales.

Afin de palier au problème des ondelettes en dimension supérieur, Candès et Donoho ont élaboré un nouveau système de représentations nommé ridgelets qui traite efficacement les singularités linéaire en deux dimensions.

La transformée en ridgelets est un nouveau outils d'analyse qui permet de trouver une solution au problème d'approximation efficace des objets en deux dimensions et plus présentant des singularités. L'idée est de représenter les singularités linéaires en singularités ponctuelles en utilisant la transformée de Radon.

La transformée en ridgelets est l'application de la transformée en ondelette sur la transformée de Radon de l'image.

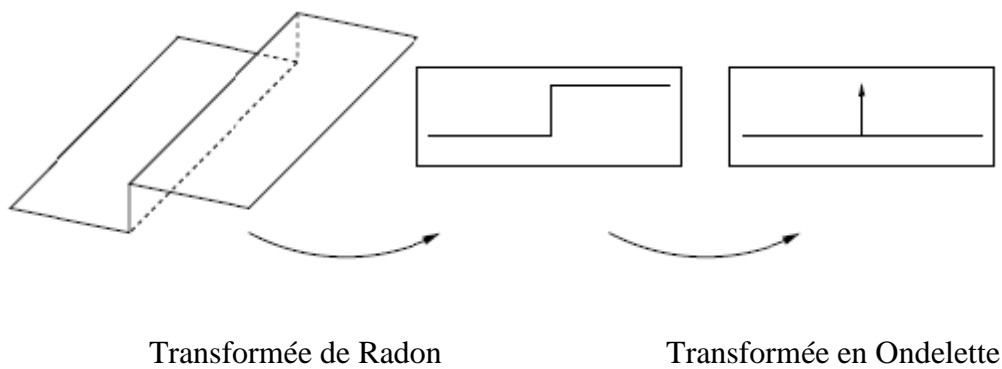


Figure 3.4. Transformée en Ridgelet

3.5.4.1. Transformée en Ridgelets continue

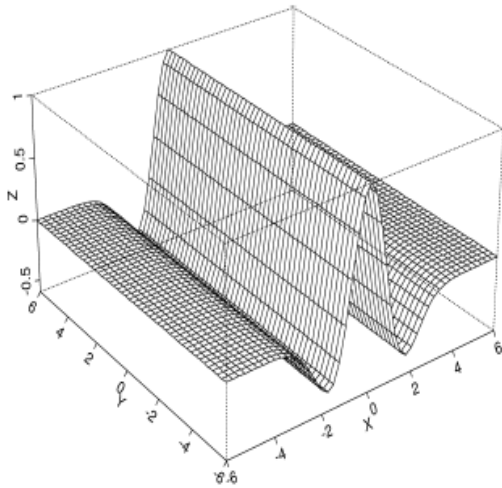
Etant donnée une fonction intégrable $f(x)$, la transformée en ridgelets de f (CRT) est définie par :

$$\text{CRT}(f) = \int_{\mathbb{R}^2} \psi_{a,b,\theta}(x) f(x) dx \quad (3.1)$$

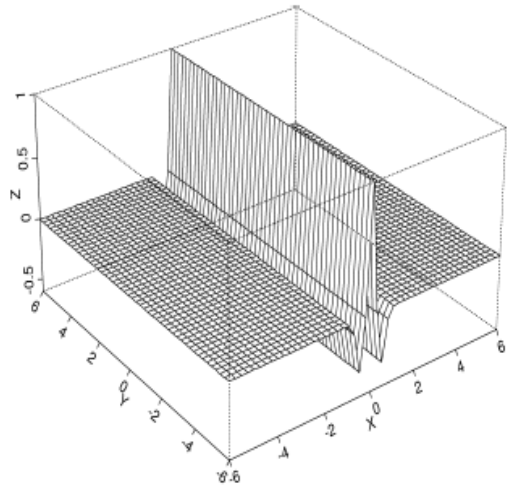
Où la ridgelet $\psi_{a,b,\theta}(x)$ en 2-D est définie à partir de l'ondelette en une dimension $\psi(x)$ par :

$$\Psi_{a,b,\theta}(x) = a^{-1/2} \psi((x_1 \cos \theta + x_2 \sin \theta - b)/a). \quad (3.2)$$

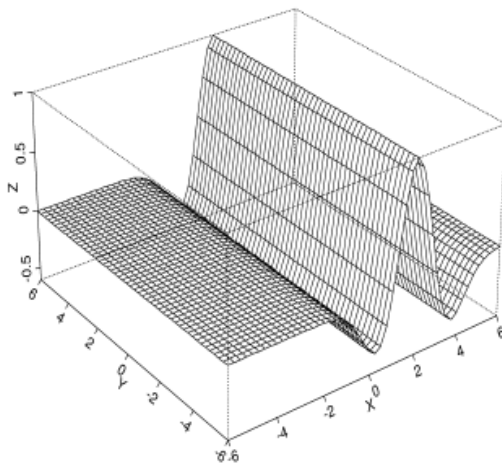
a est un paramètre d'échelle, et b paramètre de décalage (translation)



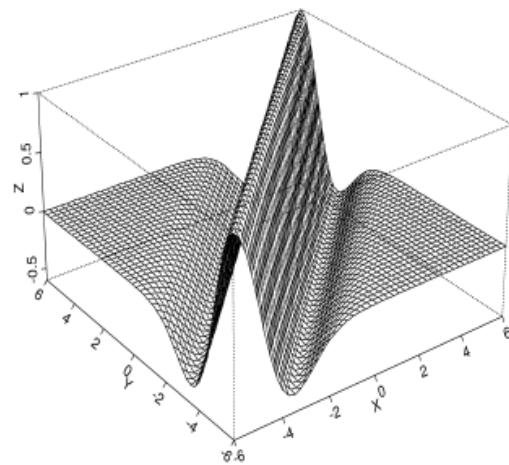
Ridgelet originale



Après changement d'échelle



Après décalage



Après Rotation

Figure 3.5. Ridgelets

La ridgelet est orientée avec un angle θ , et est constante le long de la ligne :

$$x_1 \cos \theta + x_2 \sin \theta = \text{const}$$

Dans un espace à deux dimensions, les points et les lignes sont liés via la transformée de Radon, et par conséquent les ondelettes et les ridgelets sont liés via la transformée de Radon.

3.5.4.2. Transformée de Radon [24][25]

La transformée de Radon consiste à projeter l'image sur un certain nombre d'orientations en intégrant l'image le long de la direction orthogonale à la projection (Fig. 3.6).

Appliquer la transformée de radon sur une image selon une série d'angles revient à calculer les projections de l'image selon ces angles. Le résultat de ces projections est la somme des intensités des pixels selon chaque direction, qui représente l'intégral selon une ligne.

La transformée de radon permet de projeter l'image avec une distance et un angle selon l'axe radial et l'axe angulaire.

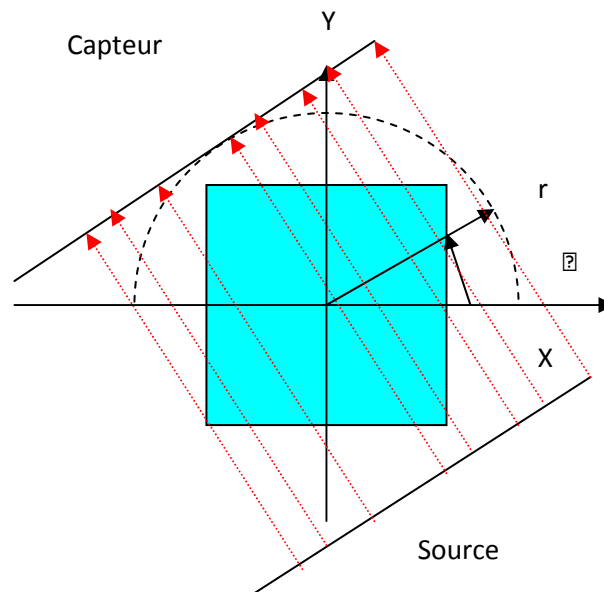
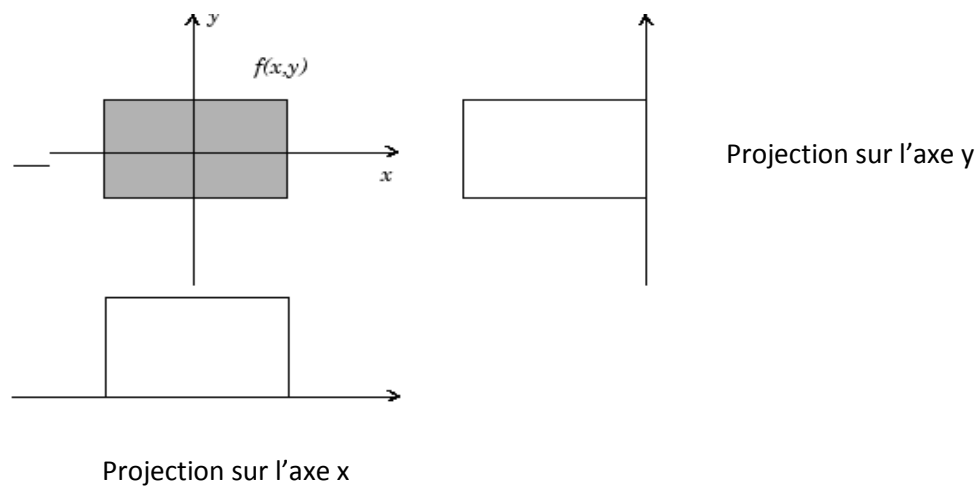


Figure 3.6. Calcul de la transformée de Radon

Le capteur et la source tourne autour du centre de l'image. Pour chaque angle θ , les intensités le long d'une droite sont accumulées au niveau du capteur. Ceci est répété pour l'ensemble des angles $\theta \in [0,180]$.



Soient (x, y) les coordonnées cartésiennes d'un point dans un espace à 2D, $f(x,y)$ une fonction continue. La transformée de Radon de $f(x, y)$, notée $Rf(\rho, \Theta)$ est définie par la relation suivante (Figure 3.6):

$$Rf(\rho, \Theta) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) \delta(\rho - x \cos \Theta - y \sin \Theta) dx dy \quad (3.3)$$

3.5.4.3. Transformation de Radon 2D par la transformée de Fourier d'une projection

Le théorème de transformation de Fourier d'une projection (Fourier slice Théorem) est une méthode permettant de réaliser la transformée de radon d'une fonction $f(x,y)$ au moyen de la transformée de Fourier.

Transformée de Fourier d'une projection : La transformée de Fourier d'une projection selon un angle Θ est égale à la transformée de Fourier bidimensionnelle de $f(x, y)$ selon une droite du plan de Fourier incliné de l'angle Θ .

Pour calculer la transformée de Radon à l'aide de la transformée de Fourier, on suit les étapes suivantes :

1. Appliquer la transformée de Fourier 2D d'une image $f(x, y) \rightarrow F(u,v)$.
2. Couper le plan de Fourier $F(u,v)$ selon une droite inclinée de l'angle Θ .
3. Appliquer la transformée de Fourier 1D inverse, pour obtenir une projection d'angle Θ .
4. Répéter les étapes 2 et 3 pour tous les angles

La reconstruction s'obtient en plaçant, pour chaque orientation de projection choisie, les coefficients de Fourier obtenus le long de cette même orientation, dans le domaine fréquentiel. On obtient l'image reconstruite en effectuant ensuite une transformée de Fourier 2D inverse.

La reconstruction parfaite pour cette transformée continue s'obtient pour un nombre de projections infini, parcourant l'ensemble des orientations possibles. La transformée de Hough est un cas particulier de transformée de Radon pour une image à valeurs binaire, et s'utilise principalement pour la reconnaissance de formes.

Le **théorème de projection de Radon** établit la possibilité de reconstituer une fonction réelle à deux variables (assimilable à une image) à l'aide de la totalité de ses projections selon des droites concourantes.

La transformée de Radon est très utilisée en tomographie, où les données capturées correspondent précisément à des projections du contenu de l'objet dont on cherche à obtenir

une image. Une reconstruction approximative de l'image recherchée, d'autant plus précise que le nombre de directions de projection est élevé, est obtenue par transformée de Radon inverse. Notons que pour une image discrète carrée dont la taille est un nombre premier, la discrétisation proposée dans peut s'appliquer pour obtenir une transformée de Radon discrète à reconstruction parfaite peu redondante.

3.5.4.4. Transformée en Ondelette

La transformée en Ondelette est capable de fournir une représentation simultanée temps fréquence du signal. Elle permet de mesurer les variations des composantes fréquentielles (spectrales) d'un signal dans le temps.

La transformée en ondelettes est, comme celle de Fourier, un passage d'une représentation à une autre ; mais comme la transformée à fenêtres, elle permet aussi de mesurer les variations dans le temps des composantes fréquentielles (spectrales) d'un signal. Néanmoins, la résolution temps fréquences de la transformée en ondelettes est différente.

La transformée en ondelette d'une fonction f en un point $(t; \omega)$ du plan temps fréquences ne dépend donc que des valeurs de $f(t)$ et $f(\omega)$ dans le rectangle de Heisenberg centré en $(t; \omega)$. L'avantage de faire varier ces largeurs devient alors évident : on minimise le nombre de translations en temps et en fréquences de la fenêtre en optimisant la largeur de celle-ci. Ainsi dans les basses fréquences, une grande largeur en fréquences n'est pas nécessaire, on peut donc utiliser des rectangles plus larges en temps.

Aux hautes fréquences, on va utiliser des rectangles plus larges en fréquences et plus localisés en temps. On peut voir cela comme une adaptation de l'ondelette à l'échelle qu'on lui impose : plus la fenêtre est petite dans le temps, plus l'ondelette va être compressée et osciller rapidement. Le contraire se produira lorsque la fenêtre est dilatée. Ainsi, les petites et grandes fenêtres enregistreront respectivement les variations rapides et moyennes du signal [26][27].

Soit $x(t)$ le signal à être analysé. L'*ondelette mère* $\psi(t)$ est choisie pour servir de prototype à toutes les fenêtres dans le processus. Toutes les fenêtres qui sont employées (ondelettes filles) sont des versions dilatées (ou comprimées) et décalées de l'ondelette mère.

$$\Psi_x(\tau, s) = \int x(t) \psi^*_{s, \tau}(t) dt \quad (3.4)$$



La Transformée Continue en Ondelette du signal $x(t)$ en utilisant translation l'ondelette mère d'analyse $\Psi(\cdot)$ Complexe conjugué de l'ondelette fille. Tous les noyaux (kernels) sont obtenus par une ou mise à l'échelle de l'ondelette mère

Le signal transformé est une fonction de deux variables, τ et s , les paramètres de *translation* (*translation*) et d'*échelle* (*scale*). $\Psi_s, \tau(t)$ est la fonction de transformation appelée l'ondelette fille (*baby functions, daughter wavelets*). L'**ondelette mère** $\psi(t)$ est choisie pour servir de prototype à toutes les fenêtres dans le processus. Toutes les fenêtres qui sont employées (ondelettes filles) sont des versions dilatées (ou comprimées) et décalées de l'ondelette mère.

Il y'a un certain nombre de fonction qui sont utilisées à cette fin. L'ondelette de Morlet, l'ondelette de Haar, et l'ondelette de Daubechies sont des exemples d'ondelettes utilisées pour l'analyse d'ondelette.

Le fait que la transformée utilise des fonctions bien localisées dans le plan temps fréquence lui donne beaucoup d'avantages :

- La résolution en fréquence de la transformée dépend du facteur de dilatation s par le principe d'Heisenberg, on peut donc choisir arbitrairement celle-ci suivant ce que l'on désire analyser.
- Pour des signaux physiques présentant des variations très rapides, des sauts, des *marches*, bref des discontinuités ; l'analyse en ondelettes est adaptée car l'ondelette va détecter ces variations et analyser celles-ci. Cette particularité rend l'analyse en ondelettes complémentaire à l'analyse de Fourier. En effet, avec l'analyse de Fourier, les discontinuités d'un signal ne sont pas facilement analysables, car les coefficients des fréquences correspondantes sont étalés dans toute la transformée.
- La localisation en temps est précieuse pour nombre d'applications.
- La transformée en ondelette peut représenter complètement et efficacement un signal quelconque en peu de coefficients.

3.5.4.5. La Multirésolution

Dans la multirésolution, on examine le signal à hautes fréquences, pour en prélever les détails, ensuite on analyse le signal à une résolution deux fois moins fine, et on réitère l'opération en grossissant son échelle d'un facteur deux, sans se préoccuper des échelles intermédiaires, jusqu'à obtenir une description complète du signal ; on parle alors d'une étude multirésolution du signal, d'où l'appellation de cet algorithme. Cette méthode nécessite deux fonctions en plus du signal analysé. Une fonction appelée fonction d'échelle et une ondelette

associée à cette fonction d'échelle. La fonction d'échelle donne une suite d'images du signal, chacune à des résolutions qui diffère de la précédente par un facteur de deux, quand on augmente la résolution, les images successives approximent le signal de mieux en mieux, quand on la diminue, la quantité d'information contenue dans les images diminue, puis s'annule. Elle détermine donc la résolution la plus fine à laquelle on étudie le signal [28],[26].

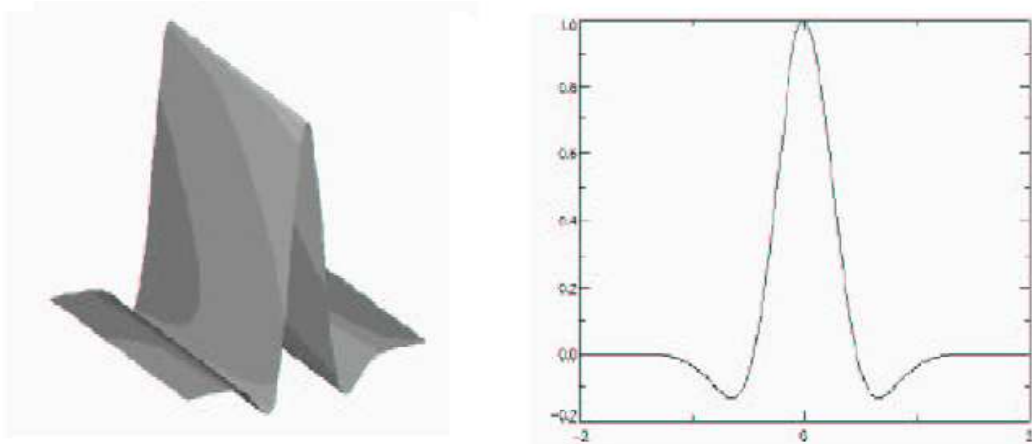


Figure 3.7. Transversalement à cette ridgelet c'est une ondelette.

3.6. Sélection des Caractéristiques

La sélection de caractéristiques est une technique permettant de choisir les caractéristiques, variables ou mesures les plus intéressantes, pertinentes ou informantes, à un

système donné, pour la réalisation de la tâche pour laquelle il a été conçu. Cette phase est généralement un module important d'un système complexe.

Les domaines d'application des techniques de sélection de caractéristiques sont variés tels que la modélisation, la classification, l'apprentissage automatique (*Machine Learning*) et l'analyse exploratoire de données (*Data Mining*). Dans ce mémoire nous nous intéressons plus particulièrement à la sélection de variables pour la classification.

Dans un problème de reconnaissance de formes, la performance du système de reconnaissance devient mauvaise et le temps de calcul augmente à mesure que le nombre de caractéristiques augmente. Dans certains cas, il existe quelques caractéristiques qui n'aideront pas à discriminer entre les classes. Autrement dit, il existe des caractéristiques redondantes ou non pertinentes. Ces caractéristiques seront inutiles dans la classification de l'objet d'où l'utilité d'effectuer la sélection des caractéristiques les plus pertinentes [1], [7].

la sélection des caractéristiques est le fait de sélectionner un sous-ensemble de m variables à partir de l'ensemble en contenant n , $m < n$, tel que la fonction critère choisie soit optimisée pour tous les sous-ensembles de taille m .

3.6.1. Sous-ensemble optimal

Soit un inducteur I , et un ensemble de données d'entraînement D avec les attributs A_1, A_2, \dots, A_n . Un sous-ensemble d'attributs optimal, A_{opt} est le sous-ensemble qui maximise la qualité du classifieur induit $Cl=I(D)$, mesurée par le pourcentage de prédiction correcte sur un ensemble test.

3.6.2. Pertinence

Un attribut A_i est pertinent si et seulement si il existe une valeur a_i de cet attribut et une valeur de la classe Y , y_i vérifiant $p(A_i = a_i) > 0$ et tels que : $p(Y=y_i / A_i = a_i) \neq p(Y=y_i)$. Avec $p(A_i = a_i)$ qui exprime le fait que les données contiennent au moins une instance ayant la valeur a_i pour l'attribut A_i .

Cette formule nous permet de dire que A_i est pertinent si la probabilité de prédire Y , en connaissant la valeur de A_i , est différente de la probabilité de prédire Y sans connaître la valeur de A_i .

3.6.3. Redondance [29][30][31][32]

Un attribut non-pertinent est défini comme celui n'affectant pas la structure fondamentale des données et un attribut redondant comme celui n'apportant rien de nouveau pour décrire la structure fondamentale des données.

Une méthode de sélection de variables est composée généralement des composantes suivantes:

- un critère d'évaluation pour comparer les différents sous-ensembles de variables et en retenir le meilleur,
- une procédure de recherche pour explorer différentes combinaisons de variables,
- un critère d'arrêt pour stopper la procédure de recherche ou de choisir l'ensemble des variables à sélectionner.

3.6.4. Méthode de sélection [6][32][34]

La sélection des caractéristiques peut être réalisée selon trois approches :

3.6.4.1. Filtres (Filter)

Si la sélection est en début d'algorithme, c'est-à-dire avant la recherche du classifieur à proprement parler, on parle de filtre car cette partie de l'algorithme va agir comme un filtre entre les données fournies et celles réellement utilisées par l'algorithme.

La caractéristique principale de cette idée est que l'on sépare la question du choix des attributs de celle de leur agencement pour définir le classifieur. La question traitée étant plus simple on peut espérer trouver des algorithmes plus rapides. Ces algorithmes peuvent alors en théorie s'appliquer avant n'importe quel algorithme d'apprentissage de classifieur pour diminuer le nombre d'attributs des données.

3.6.4.2. Méthodes internes (embedded method)

On parle de méthode interne si l'algorithme choisit les attributs pertinents en même temps qu'il définit son classifieur. Cette méthode est donc peu coûteuse et est adaptée à l'algorithme d'apprentissage considéré. En pratique, la quasi-totalité des algorithmes d'apprentissage ont en interne une méthode de sélection d'attributs, mais cette sélection peut être parfois une conséquence indirecte de l'apprentissage et non pas un mécanisme clairement conçu pour la sélection.

Dans le cas des réseaux neuronaux par exemple, l'algorithme va donner au cours de son déroulement plus ou moins d'importance aux différents attributs. Mais il n'y a pas de critère extérieur pour la sélection ou non d'un attribut.

3.6.4.3. Méthodes enveloppantes (wrapper)

On parle de wrapper lorsque l'algorithme est considéré comme une boîte noire fournissant un classifieur en fonction des exemples donnés et que la sélection s'effectue "autour" de cette boîte noire. C'est-à-dire que l'on va relancer plusieurs fois l'algorithme avec différents

ensembles d'attributs et que l'on conservera l'ensemble d'attributs le meilleur selon un score dépendant du classifieur appris.

Le principal reproche fait à cette méthode est son coût puisqu'elle implique de relancer plusieurs fois de suite l'algorithme d'apprentissage.

3.6.5. Algorithme de sélection de caractéristiques

Le but de notre travail étant d'augmenter les performances du système de reconnaissance des montants numériques des chèques bancaires, nous proposons d'améliorer ce dernier en optimisant l'étape de classification en utilisant l'information mutuelle comme critère d'évaluation des caractéristiques.

L'information mutuelle est définie pour une variable x comme la différence entre l'incertitude *a priori* et celle *a posteriori*, c'est-à-dire avant et après sélection de la variable x .

Parmi deux variables alternatives, la variable sélectionnée est bien sûr celle qui apporte le plus d'information.

La sélection des caractéristiques est réalisée selon les étapes suivantes :

1. calculer l'information mutuelle pour l'ensemble des couples classes-caractéristique.
2. éliminer la caractéristique si toutes les valeurs de l'information mutuelle sont positives ou nulles pour toutes les classes avec cette caractéristique. L'idée est de garder les caractéristiques qui ont des valeurs négatives pour certaines classes ; ceci permettra d'exclure ces classes et par conséquent réduire le nombre de choix possibles.
3. itérer le processus pour l'ensemble des caractéristiques.

3.7. Conclusion

Dans ce chapitre nous avons traité de la génération de caractéristiques pour la reconnaissance de l'écriture manuscrite. Le but de cette étape est d'extraire les caractéristiques les plus pertinentes pour la tâche de reconnaissance, car la représentation de l'information résultante conditionnera tout le reste du processus. La revue bibliographique de

ce domaine nous a permis de constater la grande variété de techniques permettant d'extraire l'information de l'image de l'écriture.

Afin de garantir une génération efficace des caractéristiques, les composantes du vecteur caractéristique sont obtenues à partir de plusieurs techniques de génération. En effet la complémentarité des différentes représentations possibles est exploitée de manière à augmenter le pouvoir discriminant du vecteur caractéristique. Il est alors naturel d'adopter l'approche qui consiste à extraire plusieurs ensembles de caractéristiques en utilisant différentes techniques, puis d'effectuer une sélection des caractéristiques les plus pertinentes. La sélection des caractéristiques est une étape visant à réduire la taille du vecteur caractéristique et d'améliorer le taux de classification.

Les résultats obtenus avec ces différentes techniques sont discutés dans le chapitre suivant.

4.1. Introduction

Les recherches dans le domaine de la reconnaissance de l'écriture nécessitent d'évaluer les résultats sur des bases de données représentatives des problèmes réels.

Nous présentons dans ce chapitre les résultats des expériences réalisées, pour évaluer les différentes méthodes proposées et discutées dans le chapitre précédent. Les expériences ont

été réalisées sur la base de données USPS, ainsi que sur des chiffres manuscrits extraits des montants des chèques de la Banque Nationale d’Algérie. Les expériences réalisées ont été faites sous environnement matlab7.0 (voir annexe pour le logiciel).

4.2. Protocole Expérimental

4.2.1. Base de données USPS

La base de données USPS (de l’anglais «US Postal Service») est une base de données bien connue au sein de la communauté de reconnaissance de formes. Cet ensemble contient 9298 images de chiffres manuscrits dont 7291 images d’apprentissage et 2007 images de test.

Ces images ont été saisies à partir d’images d’enveloppes collectées au centre CEDAR à Buffalo (États Unis). Chaque image de chiffre est représentée par 16×16 pixels de niveau de gris allant de 0 à 255, et est filtrée, normalisée et centrée. Cette base de données contient plusieurs images difficiles à classifier car l’erreur humaine se situe autour de 2.5% [20].

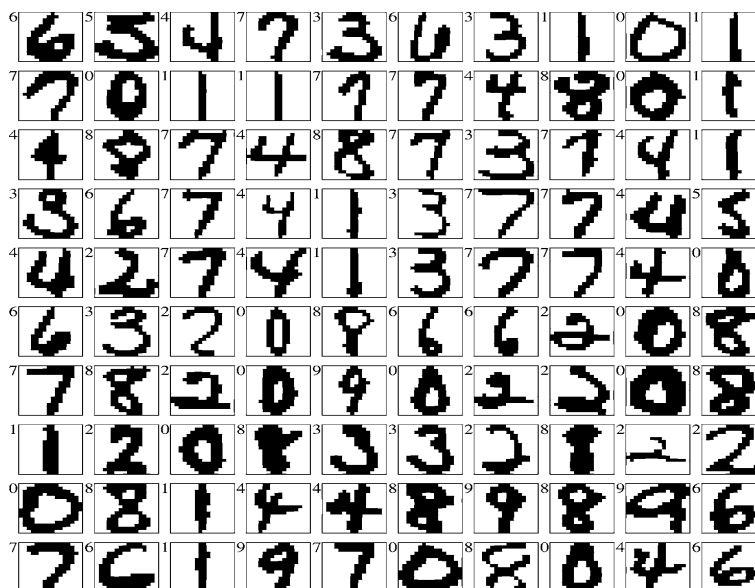


Figure 4.1 : Quelques images de chiffres de la base USPS

Les chiffres manuscrits isolés présentent l’avantage d’être abondamment disponibles sous la forme de bases de données étiquetées. Cette grande disponibilité permet d’ajuster le nombre d’exemples d’apprentissage selon les besoins tout en conservant de larges bases pour le test.

4.2.2. Base de données des chèques bancaires algériens

Cette base de données contient 600 images de chiffres manuscrits binarisées, de taille 16*16 pixels dont 450 images pour l'apprentissage, et 150 images pour la phase de test.



Figure 4.2 : Modèles de chèques algériens

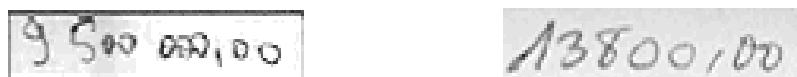


Figure 4.3 : Montant numérique des chèques bancaires

4.2.3. Critères d'évaluation

Les performances d'un système de reconnaissance peuvent être représentées par la matrice de confusion. Cette matrice permet d'évaluer le taux de bonne reconnaissance pour chaque

classe, le taux de confusion ainsi que le taux de reconnaissance globale. Les lignes de la matrice de confusion représentent les classes de référence, et les colonnes contiennent les valeurs de test pour chaque classe.

Le taux de reconnaissance globale est défini comme étant le nombre de chiffres reconnu sur le nombre total de chiffres multiplié par cent (100) :

$$TRG = \frac{\text{nombre de chiffres reconnus}}{\text{nombre totale de chiffres}} \times 100$$

De la même façon, on peut définir le taux de reconnaissance par classe comme étant le rapport entre le nombre d'exemples reconnus d'une classe et le nombre total d'exemples de cette classe multiplié par cent:

$$TRC = \frac{\text{nombre d'exemples bien reconnus pour la classe}}{\text{nombre totale d'exemples de la classe}} \times 100$$

Le taux de reconnaissance par classe se place sur la diagonale de la matrice de confusion.

4.2.4. Configuration du réseau de neurones

Pour bien apprendre le réseau de neurone plusieurs paramètres doivent être réglés : le nombre de couches cachées, et le nombre de noeuds dans chaque couche, le momentum, le nombre d'itérations et le pas d'apprentissage.

4.3. Evaluation de la base de données USPS

4.3.1. Evaluation des données brutes

Le tableau 4.1 illustre le taux de reconnaissance de chaque classe obtenue pour la meilleure configuration du réseau de neurone. Le réseau de neurone reçoit en entrée un vecteur de 256 entrées représentant l'image de taille 16x16 pixels.

Configuration du réseau de neurones :

Nombre de couches cachées	1	Nombre d'itérations	300000	Méthode d'apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	50	Momentum	0.9	Fonction de performance	Mse
Fonction d'activation	Logsig	Pas d'apprentissage	0.1	Approximation de l'erreur	0.00001

Classes de références	Classes de tests									
	0	1	2	3	4	5	6	7	8	9
0	96.66	0	0.28	1.11	0.28	0.84	0.56	0	0.28	0
1	0	94.70	0	0.38	2.27	0.75	1.52	0.38	0	0
2	1.01	0	83.84	2.02	3.54	1.51	1.01	3.03	4.04	0
3	1.01	0	1.81	86.75	0.6	4.82	0	0.6	1.4	3.01
4	0.5	1	2	0	88.00	1	2.5	1	0	4
5	1.25	0.63	0.63	5	1.25	87.49	0.63	1.25	1.25	0.63
6	0.59	0	1.36	0	1.36	1.98	94.71	0	0	0
7	0	0	1.36	0.68	2.72	0.68	0	91.84	0	2.72
8	1.2	0.6	1.2	1.81	2.41	1.81	1.2	1.2	84.34	4.22
9	0.56	0.56	0	0.56	1.13	0.56	0	1.13	0.56	94.93
TRG =										90.32 %

Tableau 4.1. Matrice de confusion calculée pour les données brutes.

On remarque que l'ensemble des classes sont reconnues avec des taux de reconnaissance variant entre 83.84% et 96.66%. Le taux de reconnaissance global est de 90.32%.

4.3.2. Evaluation des descripteurs

La section suivante présente les résultats obtenus pour les différents descripteurs. On notera que la meilleure configuration du réseau de neurone est prise pour chaque descripteur, et par conséquent elle n'est pas la même pour ces derniers.

4.3.2.1. Histogrammes de projection verticale et horizontale

L'entrée du réseau de neurone est constituée de 32 entrées : 16 entrées pour l'histogramme vertical, et 16 pour l'histogramme horizontal

Configuration du réseau de neurones :

Classes de tests					
Nombre de couches cachées	1	Nombre d'itérations	450000	Méthode d'apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	50	Momentum	0.9	Fonction de performance	Mse
Fonction d'activation	Logsig	Pas d'apprentissage	0.5	Approximation de l'erreur	0.00001

Classes de références	0	1	2	3	4	5	6	7	8	9	
	0	96,66	0,00	0,00	0,28	0,56	0,56	0,84	0,00	1,11	0,00
	1	0,00	95,83	0,00	0,38	1,52	0,38	1,52	0,00	0,00	0,38
	2	2,02	0,00	82,83	1,52	1,52	4,55	2,02	1,01	4,55	0,00
	3	1,81	0,00	3,61	74,70	0,60	10,24	0,00	0,00	7,23	1,81
	4	0,50	1,00	2,50	0,00	87,50	0,50	1,00	0,50	1,00	5,50
	5	6,88	0,00	6,88	14,38	0,63	60,00	0,00	1,25	7,50	2,50
	6	4,71	0,00	1,76	0,00	1,18	1,18	91,18	0,00	0,00	0,00
	7	0,68	0,00	0,68	0,00	4,08	0,00	0,00	88,44	1,36	4,76
	8	10,24	1,20	0,60	2,41	1,81	4,22	1,20	1,20	75,30	1,81
9	0,00	1,13	0,00	0,56	6,78	0,56	0,00	1,69	2,26	87,01	
TRG =										83,94 %	

(a)

Le taux de reconnaissance global a diminué par rapport aux données brutes. Les chiffres 3, 5 et 8 sont les moins reconnus.

Pour améliorer le taux de reconnaissance, nous avons divisé l'image en quatre zones, chacune de taille (8*8) pixels et on appliquant l'histogramme de projection sur chaque zone, on aura un vecteur de 64 caractéristiques en entrée du réseau de neurones (16 caractéristiques pour chaque zone). La configuration du réseau de neurones est la suivante :

Nombre de couches cachées	1	Nombre d'itérations	200000	Méthode d'apprentissage	Learngdm
---------------------------	----------	---------------------	---------------	-------------------------	-----------------

Nombre de nœuds dans la couche cachée **50** Momentum **0.9** Fonction de performance **Mse**
 Fonction d'activation **Logsig** Pas d'apprentissage **0.5** Approximation de l'erreur **0.00001**

Classes de tests

Classes de références	0	96,94	0,00	0,28	0,56	0,56	0,56	0,00	0,00	1,11	0,00
	1	0,00	94,70	0,00	1,52	0,76	0,00	1,52	0,00	1,14	0,38
	2	3,54	0,00	85,35	2,53	3,54	0,51	0,51	0,51	3,03	0,51
	3	1,20	0,00	1,81	87,95	0,60	6,02	0,00	0,60	1,20	0,60
	4	1,00	0,50	2,50	0,50	84,00	0,50	2,00	1,00	1,50	6,50
	5	3,13	0,00	0,63	4,38	2,50	83,75	0,00	0,00	3,13	2,50
	6	1,18	0,00	0,59	0,00	1,76	1,18	95,29	0,00	0,00	0,00
	7	0,00	0,00	0,68	0,00	4,08	0,00	0,00	89,80	0,68	4,76
	8	6,02	0,60	1,81	1,81	0,00	0,60	0,60	0,60	85,54	2,41
	9	0,00	0,56	0,00	0,00	2,82	0,00	0,00	2,26	2,82	91,53
TRG =											89,48 %

(b)

Tableau 4.2 Matrice de confusion calculée pour l'histogramme de projection horizontale et verticale

(a) 32 caractéristiques

On remarque une amélioration considérable pour les chiffres 3, 5, 6, 8 et 9 ce qui a amélioré le taux de reconnaissance global.

Zonage Le réseau de neurones est configuré avec les paramètres suivants :

Nombre de couches cachées **1** Nombre d'itérations **50000** Méthode d'apprentissage **Learngdm**
 Nombre de nœuds dans la couche cachée **70** Momentum **0.9** Fonction de performance **Mse**
 Fonction d'activation **Logsig** Pas d'apprentissage **0.5** Approximation de l'erreur **0.00001**

Le tableau 4.3 présente les résultats obtenus.

Classes de tests

0	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---

Classes de références	0	93.87	0.28	0.56	0.84	0.28	0.28	0.56	0	3.34	0
	1	0	93.56	0.38	0	1.57	0	0.38	0.38	3.73	0
	2	1.51	1.01	88.38	1.01	0.6	0.51	1.01	0.51	5.05	0
	3	0.6	0	0.6	86.14	1.2	4.22	0	1.81	3.61	1.81
	4	0	0.5	1.5	0	89.50	0	0.5	0	0.5	7.5
	5	1.23	0	0	10.63	0.63	85.00	0.63	0.63	0.63	0.63
	6	4.12	0	1.76	0	1.18	1.18	91.18	0	0.59	0
	7	0	0	0	1.36	2.04	0.68	0	88.44	1.36	6.12
	8	4.22	10.24	3.01	2.41	2.41	3.61	1.20	1.20	69.88	1.81
	9	0	0.56	0	0.56	2.26	0	0	4.52	2.26	89.83
TRG =											87.58 %

Tableau 4.3. Matrice de confusion calculée pour le zonage.

16 caractéristiques (zone de 4x4)

Le taux de reconnaissance globale s'est amélioré par rapport à l'histogramme de projection, mais il reste toujours inférieur à celui obtenu par les données brutes. Le chiffre 8 est le moins reconnu, et est confondu avec le chiffre 1.

Profils : gauche, droit, haut et bas

Configuration du réseau de neurones :

Comparativement à l'histogramme et le zonage, le profil donne de meilleurs résultats, et au

Nombre de couches cachées	1	Nombre d'itérations	450000	Méthode d'apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	50	Momentum	0.9	Fonction de performance	Mse
Fonction d'activation	Logsig	Pas d'apprentissage	0.5	Approximation de l'erreur	0.00001

dessus des résultats obtenus avec les données brutes. Les chiffres 5 et 7 sont confondus avec d'autres chiffres.

Classes de tests

	0	1	2	3	4	5	6	7	8	9
0	95,82	0,56	1,11	0,28	0,00	0,28	0,84	0,00	1,11	0,00
1	0,00	95,45	0,38	0,38	1,14	0,00	1,52	0,76	0,38	0,00
2	2,53	0,00	89,90	4,04	1,01	0,51	0,51	0,51	1,01	0,00
3	1,20	0,00	3,61	89,76	0,00	4,22	0,00	0,60	0,00	0,60
4	0,00	1,00	1,00	0,00	91,00	0,50	1,50	2,00	0,50	2,50
5	3,13	0,00	1,25	4,38	0,00	85,63	0,63	0,00	3,13	1,88
6	3,53	0,00	0,59	0,00	1,18	1,76	92,94	0,00	0,00	0,00
7	0,68	0,68	1,36	0,68	2,04	0,68	0,00	87,76	0,00	6,12
8	2,41	0,00	0,60	1,81	0,00	0,60	2,41	1,20	89,76	1,20
9	0,56	0,56	0,56	0,00	0,00	0,00	0,00	2,82	1,13	94,35

TRG = **91.24 %**

Tableau 4.4. Matrice de confusion calculée pour les profils : droit, gauche, haut et bas. 64 caractéristiques.

4.3.2.2. Ridgelets

Configuration du réseau de neurones :

Classes de tests

	0	1	2	3	4	5	6	7	8	9
0	97,77	0,00	0,56	0,56	0,56	0,00	0,28	0,00	0,00	0,28
1	0,00	97,73	0,00	0,76	0,76	0,00	0,76	0,00	0,00	0,00
2	2,02	0,51	88,89	1,01	2,53	1,01	0,51	1,01	2,53	0,00
3	0,00	0,00	3,01	86,75	0,60	6,63	0,00	0,00	2,41	0,60
4	0,50	1,00	2,00	0,00	90,00	1,00	1,50	1,00	0,00	3,00
5	1,25	0,00	0,63	5,00	1,25	90,00	0,00	0,00	0,63	1,25
6	1,18	0,00	1,18	0,00	2,35	0,59	94,71	0,00	0,00	0,00
7	0,00	0,00	0,68	0,00	6,12	0,00	0,00	91,16	1,36	0,68
8	2,41	0,00	1,20	2,41	0,60	3,01	0,60	0,00	87,35	2,41
9	0,00	1,13	0,00	0,56	1,13	0,56	0,00	0,56	1,13	94,92

TRG = **91.93 %**

Tableau 4.5. Matrice de confusion calculée pour les ridgelets.

Le taux de reconnaissance global est bien meilleur que celui obtenu pour les techniques précédentes, néanmoins le chiffre 3 ainsi que les chiffres 2 et 8 restent les moins reconnus alors que le taux de reconnaissance du chiffre 5 s'est nettement amélioré.

4.3.2.5. Combinaison par concaténation de descripteurs

Le tableau 4.6 montre le résultat obtenu après concaténation des différents descripteurs.

Descripteur	Taux de Reconnaissance
Ridg-Profil-Hist-Zone	93,98
Profils - Ridgelets	93,62
Histogramme - Profils	93,50
Ridgelets – profils - zonage	93,38
Zonage - Profils	93,00
Zonage - Ridgelets	92,42
Ridgelets	91,93
Histogramme - Ridgelets	91,29
Profils	91,24
Histogramme - Zonage	89,48
Zonage	87,58
Histogramme	83,94

Tableau 4.6. Combinaison par concaténation de descripteurs.

Le graphe suivant résume les résultats du tableau 4.6.

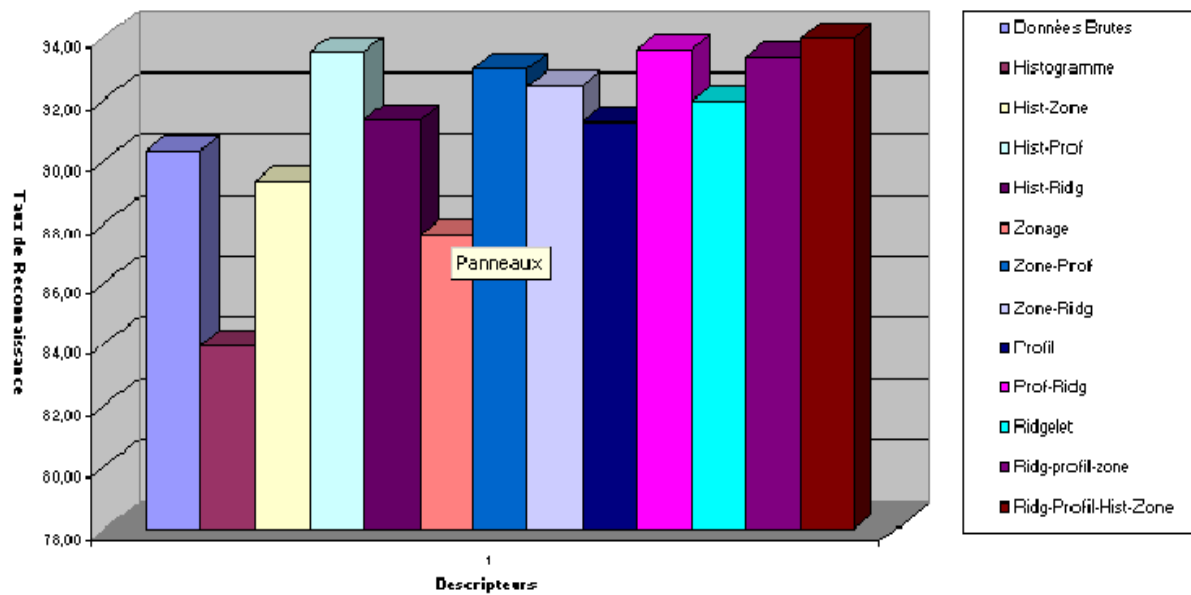


Tableau 4.7. Combinaison de descripteurs.

Le meilleur résultat obtenu est la combinaison des quatre descripteurs : histogramme de projection, zonage, ridgelets avec les profils. Le tableau 4.8 expose ces résultats.

Nombre de couches cachées	1	Nombre d'itérations	400000	Méthode d'apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	50	Momentum	0.9	Fonction de performance	Mse
Fonction d'activation	Logsig	Pas d'apprentissage	0.5	Approximation de l'erreur	0.00001

	0	1	2	3	4	5	6	7	8	9	
Classes de références	0	97,77	0,00	0,28	0,28	0,56	0,00	0,28	0,28	0,56	0,00
	1	0,00	96,21	0,38	0,38	1,14	0,00	1,89	0,00	0,00	0,00
	2	1,01	0,00	92,93	1,52	0,51	1,01	0,00	0,51	2,02	0,51
	3	0,60	0,00	1,81	90,96	0,00	3,61	0,00	0,00	2,41	0,60
	4	0,50	1,00	1,50	0,00	90,50	1,50	1,00	0,50	0,00	3,50
	5	1,25	0,00	0,00	2,50	0,00	93,75	0,00	0,00	1,25	1,25
	6	1,18	0,00	1,18	0,00	1,18	1,18	95,29	0,00	0,00	0,00
	7	0,00	0,00	0,68	0,68	4,76	0,00	0,00	91,84	0,00	2,04
	8	2,41	0,00	0,00	1,81	0,00	1,20	0,00	0,00	93,37	1,20
	9	0,00	0,56	0,00	0,56	0,56	0,00	0,00	0,56	1,13	96,61

TRG = **93,98 %**

Tableau 4.8. Matrice de confusion calculée pour la concaténation des différents descripteurs.

4.4. Evaluation des descripteurs pour les données des chèques bancaires algériens

4.4.1. Evaluation des données brutes

Le tableau 4.9 illustre le taux de reconnaissance de chaque classe obtenue pour la meilleure configuration du réseau de neurone. Le réseau de neurone reçoit en entrée un vecteur de 256 entrées représentant l'image de taille 16x16 pixels.

Classes de tests

	0	1	2	3	4	5	6	7	8	9
0	73.33	73.33	60.00	93.33	73.33	86.67	73.33	53.33	40.00	53.33

TRG = **68.00 %**

Tableau 4.9. Matrice de confusion calculée pour les données brutes.

Le taux de reconnaissance global a considérablement diminué. Les classes 7, 8 et 9 sont les moins reconnues, contrairement à la classe 3.

4.4.2. Evaluation des descripteurs

La section suivante présente les résultats obtenus pour les différents descripteurs. La meilleure configuration du réseau de neurone est prise pour chaque descripteur.

4.4.2.1. Histogrammes de projection verticale et horizontale

L'entrée du réseau de neurone est constituée de 32 entrées : 16 entrées pour l'histogramme vertical, et 16 pour l'histogramme horizontal.

Configuration du réseau de neurones :

	0	1	2	3	4	5	6	7	8	9
Classes de références	0	86.67	6.67	0	0	0	0	0	0	0
	1	0	80.00	6.67	0	0	0	0	13.33	0
	2	6.67	0	73.33	6.67	0	0	13.33	0	0
	3	0	0	6.67	86.67	0	6.67	0	0	0
	4	0	0	0	0	100	0	0	0	0
	5	0	0	6.67	0	0	73.33	6.67	0	6.67
	6	0	0	20.00	0	6.67	0	73.33	0	0
	7	0	0	0	6.67	0	0	0	93.33	0
	8	6.67	6.67	6.67	0	0	26.67	6.67	0	46.67
	9	0	6.67	6.67	13.33	0	0	0	0	0
TRG =										78.67 %

Tableau 4.10. Matrice de confusion calculée pour l’histogramme de projection horizontale et verticale.

Vue que la base d’apprentissage est faible (450 exemples) par rapport à la base USPS (7291 exemples) le taux de reconnaissance global a nettement diminué. Le taux de reconnaissance du chiffre 8 a diminué considérablement, tandis que celui du chiffre 4, 3 et 5 s’est amélioré.

4.4.2.2. Zonage

Configuration du réseau de neurones :

Nombre de couches cachées	1	Nombre d’itérations	50000	Méthode d’apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	40	Momentum	0.9	Fonction de performance	Mse
Fonction d’activation	Logsig	Pas d’apprentissage	0.5	Approximation de l’erreur	0.00001

Classes de tests

	0	1	2	3	4	5	6	7	8	9	
Classes de références	0	80.00	0	0	0	6.67	6.67	0	6.67	0	
	1	0	86.67	0	0	0	6.67	0	6.67	0	
	2	0	0	100	0	0	0	0	0	0	
	3	0	0	0	93.33	0	0	0	6.67	0	
	4	0	0	0	0	100	0	0	0	0	
	5	0	0	0	6.67	6.67	80.00	0	0	6.67	
	6	0	0	0	0	6.67	0	86.67	6.67	0	
	7	0	0	0	0	0	0	0	100	0	
	8	0	0	20.00	6.67	0	13.33	0	6.67	53.33	0
	9	0	0	0	53.33	0	0	0	0	0	46.67

TRG = **82.67 %**

Tableau 4.11. Matrice de confusion calculée pour le zonage.

Les chiffres 2, 4 et 7 sont bien reconnus alors que les chiffres 8 et 9 ne le sont pas. Le taux de reconnaissance globale a diminué par rapport à celui de la base USPS.

4.4.2.3. Profils : gauche, droit, haut et bas

Configuration du réseau de neurones :

Nombre de couches cachées	1	Nombre d'itérations	40000	Méthode d'apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	50	Momentum	0.9	Fonction de performance	Mse
Fonction d'activation	Logsig	Pas d'apprentissage	0.5	Approximation de l'erreur	0.00001

Classes de tests

	0	1	2	3	4	5	6	7	8	9
0	80.00	0	0	0	13.33	0	6.67	0	0	0
1	0	86.67	0	0	0	0	6.67	0	6.67	0
2	0	0	93.33	0	0	0	6.67	0	0	0
3	0	0	0	100	0	0	0	0	0	0
4	0	0	0	0	100	0	0	0	0	0
5	0	0	0	0	0	93.33	0	0	0	6.67
6	0	0	0	0	0	0	93.33	0	6.67	0
7	0	0	0	0	0	0	0	100	0	0
8	0	0	6.67	0	0	0	6.67	0	86.67	0
9	0	6.67	0	46.67	0	0	0	0	0	46.67

TRG = **88.00 %**

Tableau 4.12. Matrice de confusion calculée pour le profils : droit, gauche, haut et bas.

On constate une baisse légère du taux de reconnaissance global. Les chiffres 3, 4 et 7 sont bien reconnus. Le chiffre 9 est mal reconnu, et est considérablement confondu avec le chiffre 3.

4.4.2.4. Ridgelets

Configuration du réseau de neurones :

Nombre de couches cachées	1	Nombre d'itérations	10000	Méthode d'apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	50	Momentum	0.9	Fonction de performance	Mse
Fonction d'activation	Logsig	Pas d'apprentissage	0.5	Approximation de l'erreur	0.00001

Classes de tests

	0	1	2	3	4	5	6	7	8	9
0	80.00	13.33	0	0	0	0	0	0	6.67	0
1	0	86.67	0	0	0	0	13.33	0	0	0
2	0	0	80.00	13.33	0	0	6.67	0	0	0
3	0	0	0	100	0	0	0	0	0	0
4	0	0	0	0	100	0	0	0	0	0
5	0	0	0	0	0	93.33	0	0	0	6.67
6	0	0	0	0	0	0	100	0	0	0
7	0	0	0	0	0	0	0	100	0	0
8	0	13.33	6.67	6.67	0	6.67	0	13.33	46.67	6.67
9	6.67	0	6.67	20	0	0	0	0	0	66.67

TRG = **85.33 %**

Tableau 4.13. Matrice de confusion calculée pour les ridgelets.

192 caractéristiques.

Le chiffre 9 est toujours confondu avec le chiffre 3 ; le chiffre 8 est le moins reconnu. Les chiffres 3, 4, 6 et 7 sont reconnus sans aucune ambiguïté.

4.4.2.5. Combinaison par concaténation de descripteurs

Le graphe suivant montre le résultat obtenu après concaténation des différents descripteurs pour les données des chèques bancaires algériens.

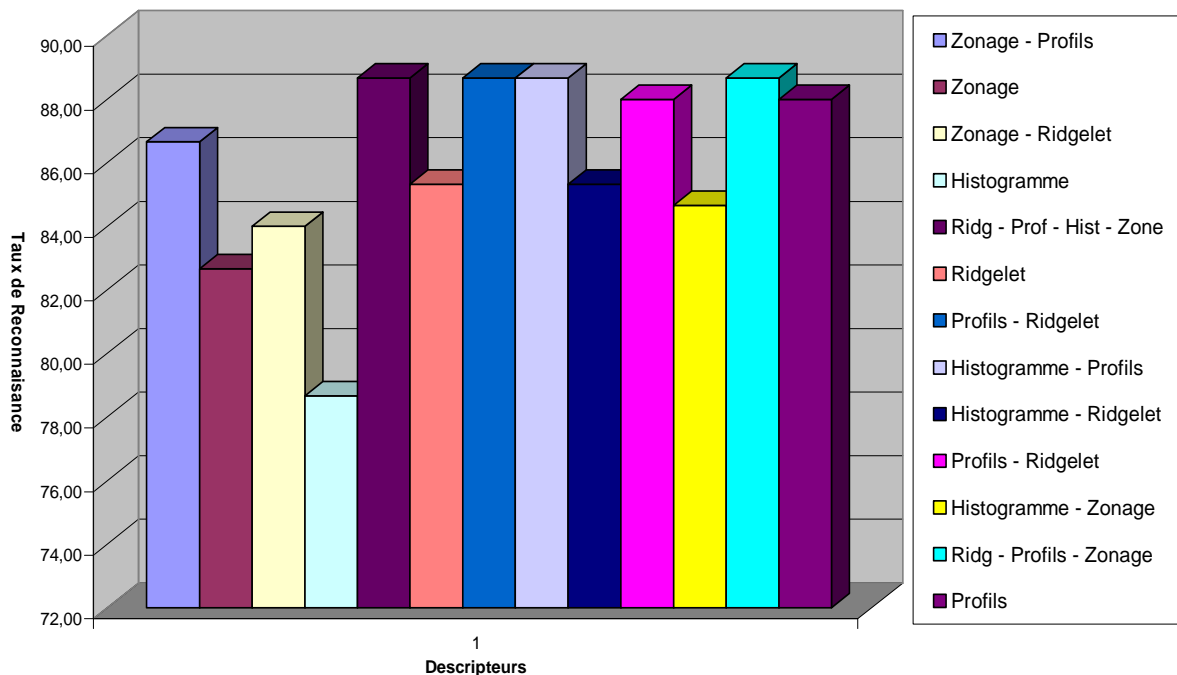


Tableau 4.14. Combinaison de descripteurs.

Le meilleur résultat obtenu est la combinaison des ridgelets avec les profils. Le tableau 4.15 expose ces résultats. Le réseau de neurones est configuré avec les paramètres suivants :

Nombre de couches cachées	1	Nombre d'itérations	10000	Méthode d'apprentissage	Learngdm
Nombre de nœuds dans la couche cachée	50	Momentum	0.9	Fonction de performance	Mse
Fonction d'activation	Logsig	Pas d'apprentissage	0.5	Approximation de l'erreur	0.00001

Classes de tests

	0	1	2	3	4	5	6	7	8	9
Classes de références	0	93,33	0,00	0,00	0,00	0,00	0,00	0,00	6,67	0,00
	1	0,00	86,67	0,00	0,00	0,00	0,00	6,67	0,00	0,00
	2	0,00	0,00	86,67	0,00	0,00	0,00	13,33	0,00	0,00
	3	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00	0,00
	4	0,00	0,00	0,00	0,00	100	0,00	0,00	0,00	0,00
	5	0,00	0,00	0,00	0,00	0,00	93,33	0,00	0,00	6,67
	6	0,00	0,00	0,00	0,00	0,00	0,00	100,00	0,00	0,00
	7	0,00	0,00	0,00	0,00	0,00	0,00	0,00	100	0,00
	8	0,00	6,67	13,33	6,67	0,00	0,00	0,00	0,00	73,33
	9	93,33	0,00	0,00	0,00	0,00	0,00	0,00	6,67	0,00
TRG =										88.67 %

Tableau 4.15. Matrice de confusion calculée pour la concaténation des ridgelets avec les profils.

4.5. Réduction des caractéristiques

La sélection des caractéristiques les plus pertinentes pour la tâche de classification est une étape nécessaire pour avoir une classification robuste, réduire le nombre de caractéristiques et maintenir ou améliorer le taux de reconnaissance.

La procédure de sélection est composée de deux éléments fondamentaux :

- une mesure de pertinence pour mesurer l'influence d'un sous-ensemble de variables sur la variable à expliquer,
- un algorithme de recherche pour parcourir l'ensemble des sous-ensembles de variables.

Si nous voulons quantifier la corrélation entre deux événements, il faut se demander comment la réalisation de l'un d'entre eux « $Y=y$ » va modifier l'incertitude sur l'autre « $X=x$ ».

La quantité suivante mesure l'incertitude :

$$h(x) = \log(1/ p(x))$$

La probabilité à priori $p(x) = \Pr(X = x)$ va devenir la probabilité à posteriori :

$$p(x/ y) = \Pr(X = x/ Y = y)$$

L'information mutuelle entre x et y représente la différence entre les deux quantités d'incertitude.

$$I(x, y) = h(x) - h(x/ y)$$

$$I(x, y) = \log_2(\Pr[X = x / Y = y] / \Pr(X = x))$$

$$I(x, y) = \log(p(x / y) / p(x))$$

$I(x;y) \geq 0$ ssi $p(x/y) \geq p(x)$ ie la réalisation de y augmente l'occurrence de x.

$I(x;y) \leq 0$ ssi $p(x/y) \leq p(x)$ ie la réalisation de y diminue l'occurrence de x.

$I(x;y) = 0$ ssi $p(x/y) = p(x)$ ie x et y sont indépendants.

Notre approche est basée sur les valeurs négatives de l'information mutuelle entre la classe et la caractéristique. Si toutes les valeurs de l'information mutuelle sont positives ou nulles entre une caractéristique et toutes les classes alors cette caractéristique est éliminée.

L'idée est d'écartier les caractéristiques redondantes et d'ignorer le bruit présent dans l'image du chiffre (un pixel qui apparaît rarement pour une classe aura une grande valeur de l'information mutuelle, alors qu'il n'a pas de pouvoir discriminant pour cette classe).

Le calcul de l'information mutuelle se base sur le calcul des densités conjointes de la classe et la caractéristique. Ceci étant n'est pas toujours aisé pour des variables multivariées, nous avons utilisés des distributions gaussiennes afin d'estimer ces densités.

4.5.1. Application sur les données brutes

Le calcul de l'information mutuelle entre classes et caractéristiques est assez simple, puisque on prend les fréquences d'apparition des classes, des caractéristiques et les fréquences d'apparition des classes et caractéristiques.

L'apprentissage a été réalisé sur l'ensemble de la base de données USPS (7291 exemples).

Les paramètres du réseau de neurones sont les suivant :

Nombre de Nœuds : 50

Momentum : 0,9

Pas d'apprentissage : 0,5

Nombre d'époques : 100000

Taille du vecteur caractéristique avant réduction : 256

Taille du vecteur caractéristique après réduction : 243

Le tableau 4.16 montre les résultats obtenus.

Classe	Taux de Bonne Reconnaissance Avant la Réduction	Taux de Bonne Reconnaissance Après Réduction
0	96,66	98,05
1	94,70	95,83
2	83,84	86,87
3	86,75	86,14
4	88,00	87,50
5	87,50	84,38
6	94,71	92,94
7	91,84	91,84
8	84,34	86,75
9	94,92	93,22
Taux Global :	90,32	90,35

Tableau 4.16. Réduction des caractéristiques pour les données brutes

On remarque que le taux de reconnaissance a peu augmenter, et le nombre de caractéristiques a diminuer de 5.08 %, c'est l'idéal.

4.5.2. Application sur les ridgelets

Le calcul de l'information mutuelle n'est pas aussi simple que celui pour les données brutes. La distribution gaussienne est utilisée pour estimer les densités conjointes. La base de données utilisée est l'USPS.

Le réseau de neurone est configuré de la manière suivante :

Nombre de Nœuds : 50

Momentum : 0,9

Pas d'apprentissage : 0,5

Nombre d'époques : 200000

Taille du vecteur caractéristique avant réduction : 192

Taille du vecteur caractéristique après réduction : 178

Les résultats obtenus sont exposés dans le tableau 4.17.

Classe	Taux de Bonne Reconnaissance Avant la Réduction	Taux de Bonne Reconnaissance Après Réduction
0	94,22	96,94
1	96,92	95,45
2	93,30	87,37
3	88,60	86,75
4	91,41	92,50
5	89,90	86,25
6	95,48	92,35
7	93,49	89,80
8	79,34	87,35
9	87,58	93,79
Taux Global :	91,02	90,85

Tableau 4.17. Réduction des caractéristiques pour les ridgelets

Le vecteur de caractéristique a diminué de 7.3 %, mais le taux de reconnaissance a lui aussi diminué de 0.17 %.

4.5.3. Concaténation des ridgelets avec les profils

Les probabilités conjointes sont calculées selon une distribution gaussienne et la base de données utilisée est l'USPS.

Le réseau de neurone est configuré de la manière suivante :

Nombre de Nœuds : 50

Momentum : 0,9

Pas d'apprentissage : 0,5

Nombre d'époques : 400000

Taille du vecteur caractéristique avant réduction : 256

Taille du vecteur caractéristique après réduction : 232

Les résultats obtenus sont montrés dans le tableau 4.18.

Classe	Taux de Bonne Reconnaissance Avant la Réduction	Taux de Bonne Reconnaissance Après Réduction
0	97,77	98,33
1	97,35	96,97
2	92,93	91,41
3	90,96	86,75
4	92,50	91,50
5	92,50	93,75
6	94,71	94,71
7	90,48	90,48
8	90,96	92,77
9	96,05	96,05
Taux Global :	93,62	93,27

Tableau 4.18. Réduction des caractéristiques pour les ridgelets

Le taux de reconnaissance a diminué de 0.35 %, de même que le nombre de caractéristiques a diminué de 9.4 %.

4.5.4. Discussion

On constate que pour les caractéristiques multivariées utilisées pour le calcul de l'information mutuelle, et donc ceux qui estiment les densité conjointes des classes et caractéristiques, le taux de reconnaissance a légèrement diminué avec le nombre de caractéristiques. Pour les données brutes, les probabilités conjointes sont calculées sans erreur et donc l'information mutuelle a un meilleur pouvoir discriminant.

On note que pour la base des cheques bancaires algériens les expériences faites non pas donner de résultats.

4.6. Conclusion

Notre travail étant consacré à l'évaluation des méthodes de génération des caractéristiques, Nous avons présenté, dans ce chapitre, les résultats expérimentaux réalisée sur deux bases données : l'USPS et les montants numériques extraient des cheques bancaires algériens. Ces expériences montrent que les caractéristiques issues des profils et des ridgelets sont les plus robustes comparativement aux autres caractéristiques. La concaténation des descripteurs permet d'améliorer la performance du système de reconnaissance.

Les expériences faites pour la sélection des caractéristiques pertinentes, montrent que le taux de reconnaissance n'est pas amélioré même si le nombre de caractéristiques est réduit. Cela est du sans doute à l'estimation des probabilités conjoints des caractéristiques et classes pour calculer l'information mutuelle. Dans le cas où les probabilités conjointes sont connues, on a une amélioration du taux de reconnaissance, et une baisse du nombre de caractéristiques.

Conclusion Générale

Le travail présenté dans ce mémoire a pour objectif l'implémentation d'un système de reconnaissance des chiffres manuscrit capable de traiter automatiquement les montants numériques des chèques bancaires algériens. Ce système est constitué de trois modules: les prétraitements et segmentation, la génération des caractéristiques et la classification. Notre travail est focalisé sur le module de génération des caractéristiques où nous avons tenté de proposer une méthode d'optimisation pour la sélection des caractéristiques pertinentes du système de reconnaissance des chiffres. En effet pour la prise de décision, un système de reconnaissance n'a besoin que de l'information pertinente pour différencier un objet d'un autre. Plusieurs méthodes ont été étudiées et comparées. Une fois l'étape de génération des caractéristiques terminée, la sélection des caractéristiques pertinentes du système est effectuée. L'information mutuelle est le critère d'évaluation des caractéristiques qui a été utilisée pour la sélection et réduction des caractéristiques. Les expériences ont été réalisées sur la base de données USPS et la base données des chèques bancaires algériens. Le classifieur utilisé est un réseau de neurone de type perceptron multicouche.

Les tests effectués montrent que les caractéristiques issues des profils et des ridgelets donnent de meilleurs résultats par rapport aux autres méthodes. On constate que la combinaison de ces méthodes améliore le taux de reconnaissance.

Pour la sélection des caractéristiques les tests réalisés montrent qu'on peut réduire le nombre de caractéristiques et augmenter le taux de reconnaissance si les densités conjointes des classes et caractéristiques sont connues.

Pour compléter ce travail, nous proposons pour l'étape de génération des caractéristiques, l'implémentation de nouveaux descripteurs plus robustes basés sur les mesures de similarités tel que la mesure de Jaccard. Pour la sélection des caractéristiques pertinentes ainsi que pour la réduction de ces dernières, nous estimons que l'utilisation des noyaux pour l'estimation de la fonction de distribution apporte de meilleurs résultats.

L'étape de classification sera plus performante en utilisant la combinaison de plusieurs classifieurs tels que les classifieurs basés sur les méthodes statistiques et ceux de types machines à vecteurs de support (SVM).

Annexe

Présentation de la plate forme

Nous présentons dans cette annexe la plate forme de reconnaissance de chiffres manuscrits implémentée sous environnement Matlab. Les sections suivantes décrivent les fonctions essentielles implémentées.

Fenêtre Principale

La fenêtre principale permet de sélectionner la base de données des chiffres manuscrits sur laquelle les expériences seront réalisées.

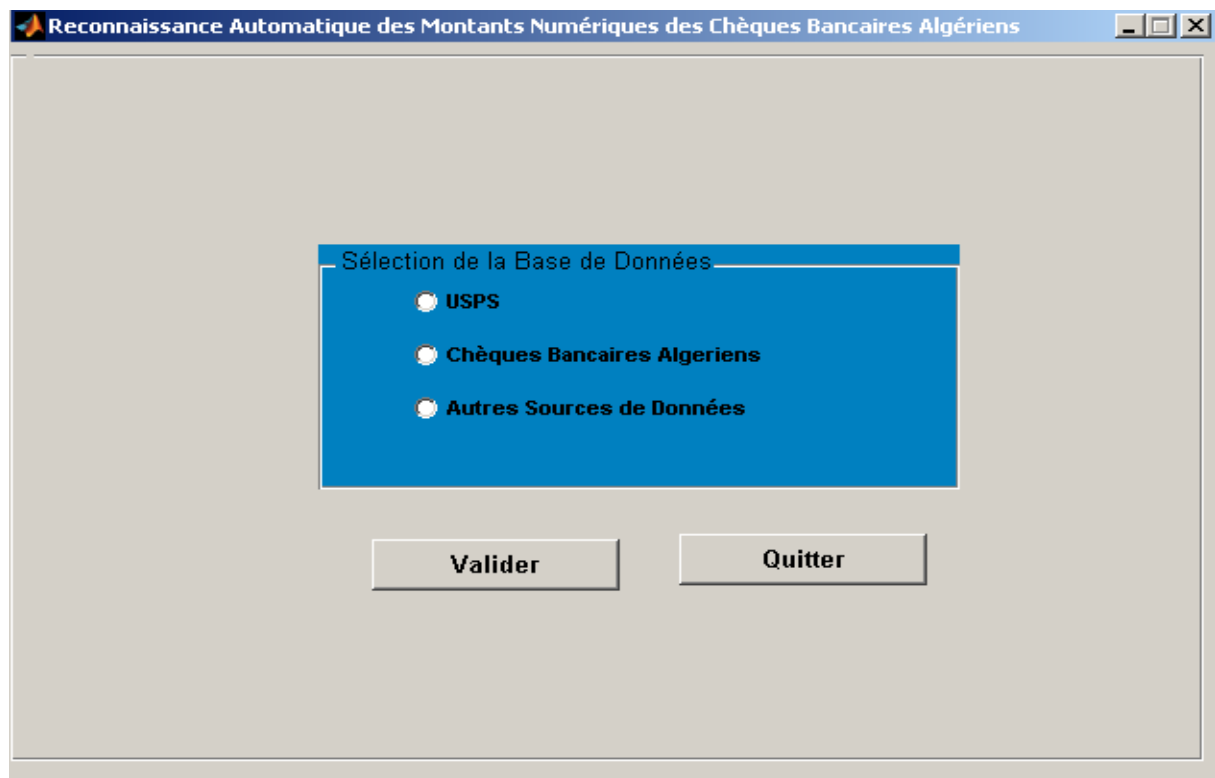


Figure 5.1 : Fenêtre principale

Le bouton valider permet de passer à la fenêtre suivante (extraction des caractéristiques).

Le bouton quitter permet de quitter la plate forme de reconnaissance.

Fenêtre extraction des caractéristiques

Cette fenêtre permet de travailler avec des données brutes, choisir une méthode d'extraction des caractéristiques, ou bien choisir plusieurs méthodes d'extraction des caractéristiques.

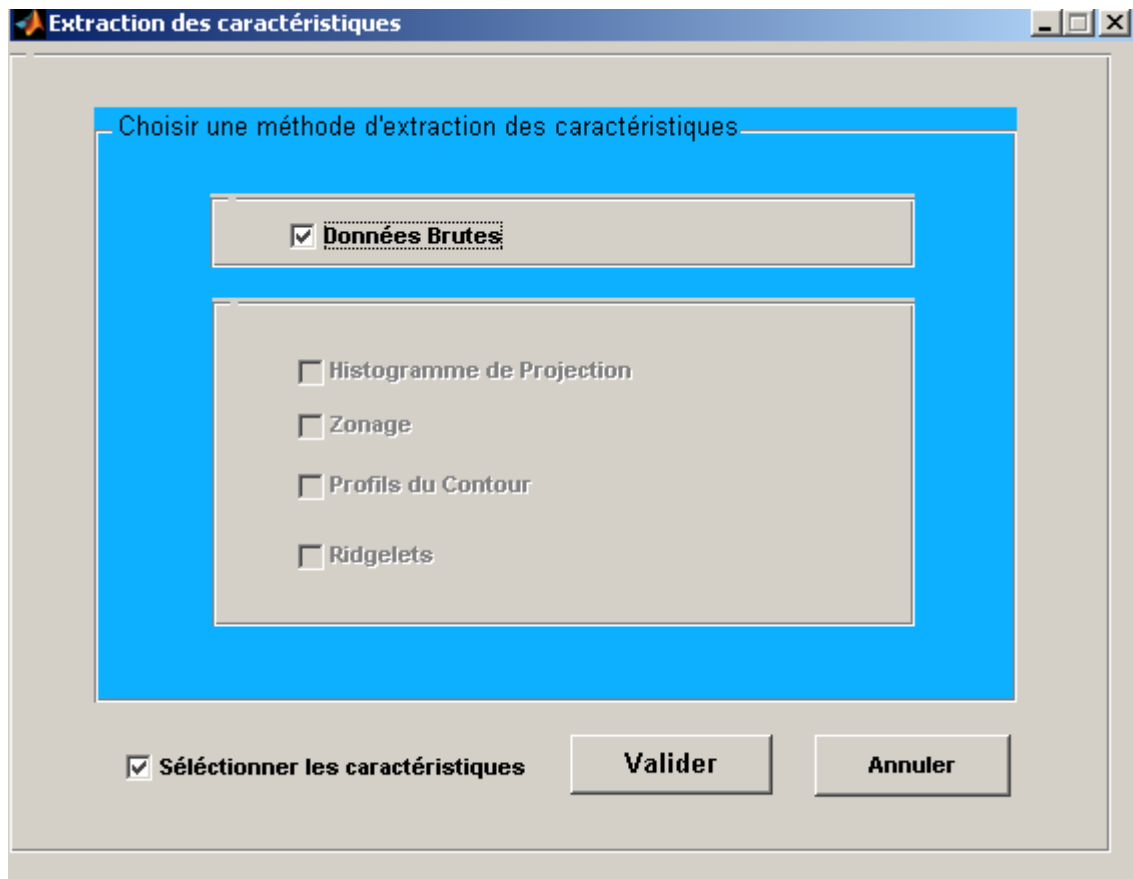


Figure 5.2 : Extraction des caractéristiques (données brutes sont choisies)

Dans cette fenêtre, on aperçoit que les données brutes ne peuvent être combinées avec d'autres méthodes d'extraction et que si les données brutes sont sélectionnées, le panel contenant les autres méthodes d'extraction est désactivé et vice versa.

Si une méthode d'extraction des caractéristiques est choisie (données brutes ne sont pas choisies), une autre fenêtre s'ouvre et permet de sélectionner les critères de sélection pour cette méthode la figure 5.3 montre un exemple sur la méthode de contour des profils.

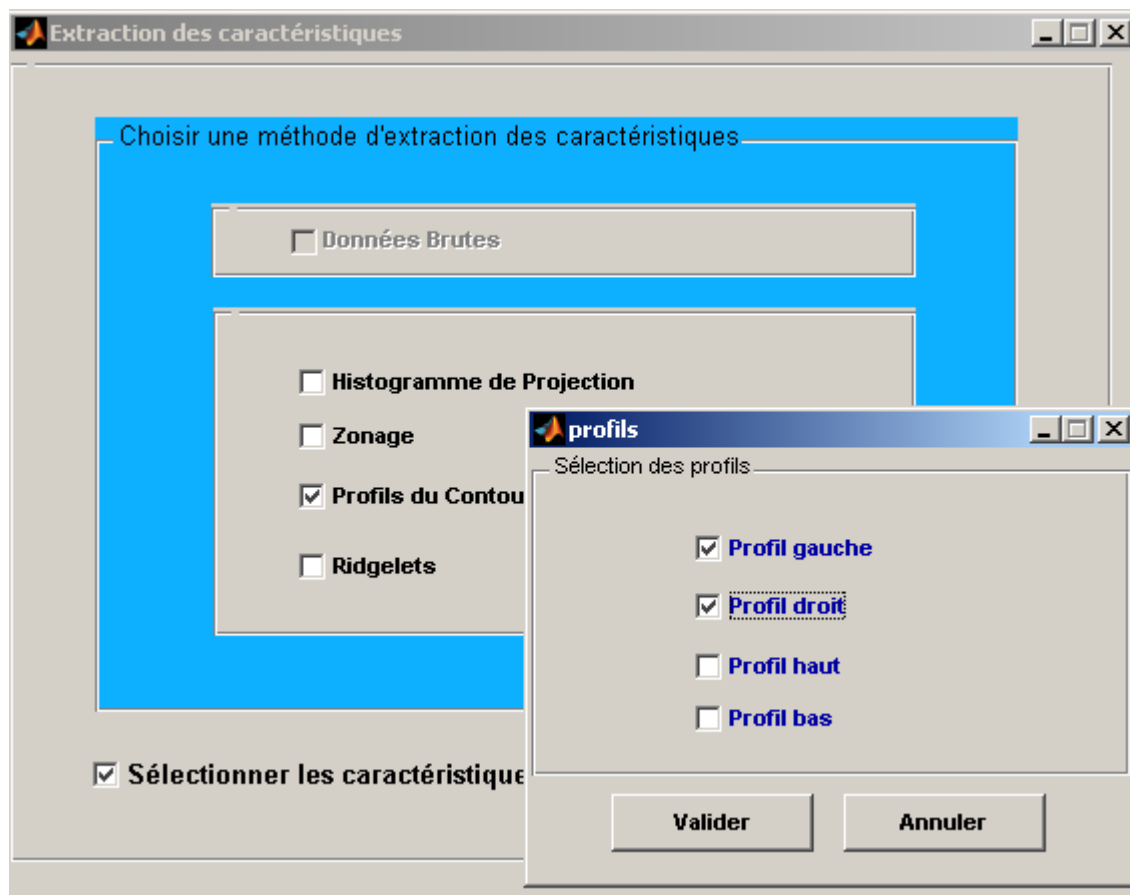


Figure 5.3 : Extraction des caractéristiques

(Fenêtre profils qui apparaît pour déterminer les critères de sélection)

- Le bouton valider permet de confirmer la sélection des critères.
- Le bouton annuler permet d'annuler la sélection et revenir à la fenêtre précédente.

La fenêtre extraction des caractéristiques permet en outre de passer par l'étape de sélection de caractéristiques (et ainsi calculer l'information mutuelle et appliquer les algorithmes de sélection des caractéristiques) avant de procéder à la reconnaissance si la case à cocher 'sélectionner les caractéristique' en bas de la fenêtre est cocher.

Fenêtre réseau de neurones

La fenêtre réseau de neurone permet d'initialiser les paramètres essentiels du réseau de neurones, avant de lancer son apprentissage. La figure 5.4 présente un exemple sur l'initialisation de ces paramètres :

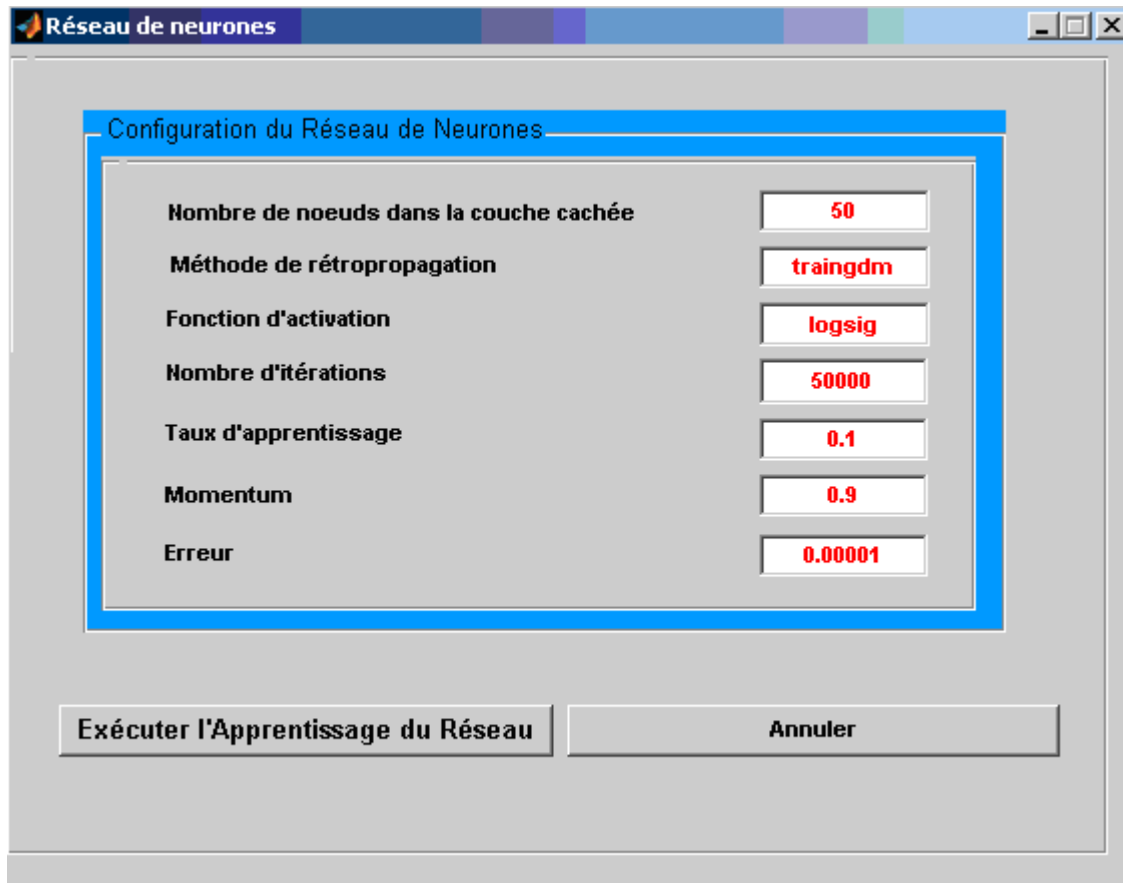


Figure 5.4 : Configuration du réseau de neurones

(Initialisation des différents paramètres)

- Le bouton exécuter l'apprentissage du réseau permet de lancer l'apprentissage.
- Le bouton annuler permet d'annuler l'apprentissage et de revenir à la fenêtre extraction des caractéristiques.

Fenêtre visualisation des résultats :

Une fois l'apprentissage du réseau de neurones est terminé, la fenetre de visualistion des résultats d'apprentissage apparaisse automatiquement et les taux de bonne reconnaissance ainsi que les taux de confusion s'afficheront dans la matrice de confusion.

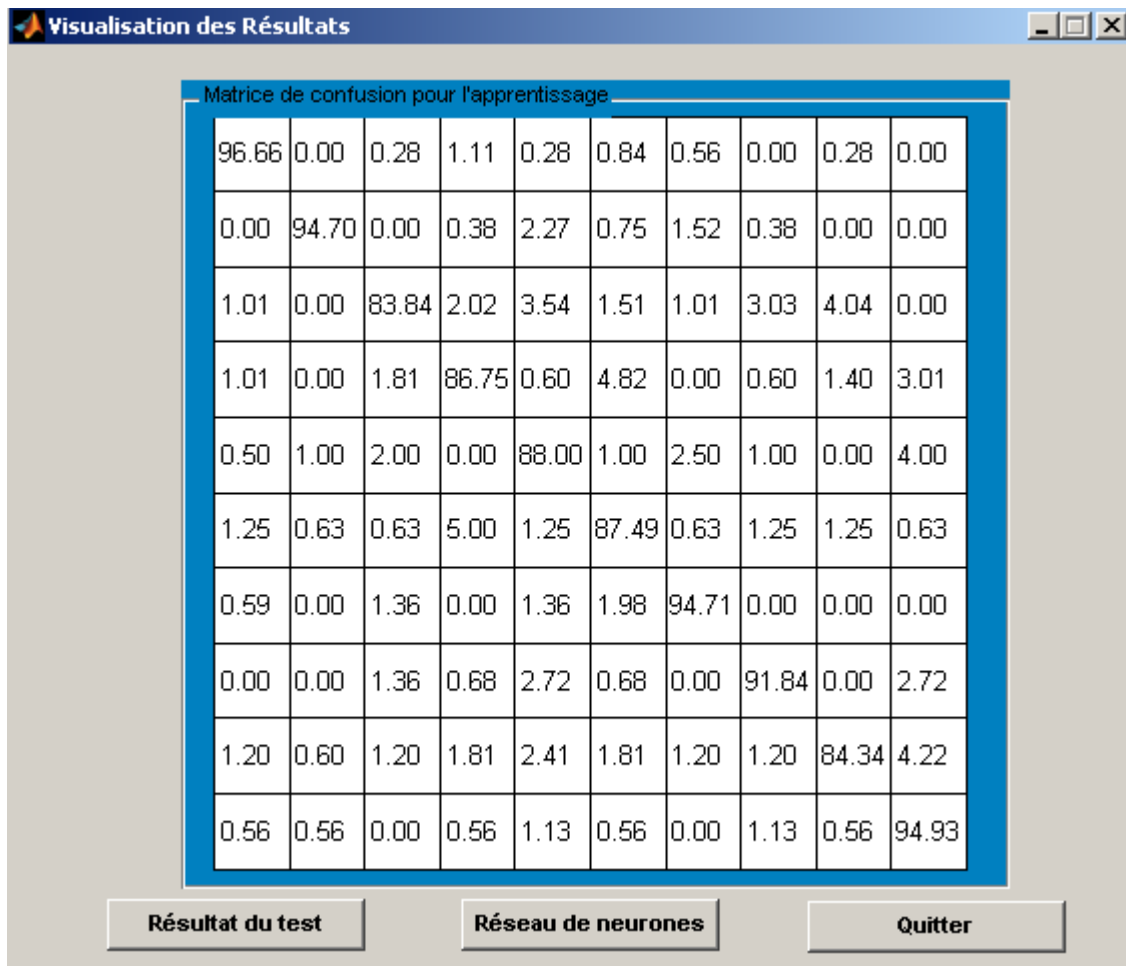


Figure 5.5 : Visualisation des résultats

(Matrice de confusion pour les résultats de la phase apprentissage)

- Le bouton résultat du test permet de voir les résultats du test du réseau de neurones.
- Le bouton réseau de neurones permet de revenir à la fenêtre réseau de neurones.
- Le bouton quitter permet de quitter la plate forme de reconnaissance.

Références Bibliographiques

[1] **N. Benahmed, 2002.** Optimisation des réseaux de neurones pour la reconnaissance de chiffres manuscrits isolés : Sélection et pondération des primitives par algorithmes génétiques, Thèse de Master, École de technologie supérieure, MONTRÉAL, pp.4-70.

[2] **Trier, O.D., Jain, A.K., et Taxt, T., 1996.** Feature extraction methods for character recognition: a survey, *Pattern Recognition Journal*, Vol. 29, 641-662.

[3] **B. Gosselin, 1996.** Application de réseaux de neurones artificiels a la reconnaissance automatique de caractères manuscrits », Thèse de doctorat, Faculté Polytechnique De Mons, pp.21- 28

[4] **A. Belaïd et H. Cecotti.** La numérisation de documents : Principe et évaluation des Performances. Université Nancy 2 – LORIA.

[5] **M. Cheriet, N. Kharma, C.L. Liu and C.Y. Suen, 2004.** Character Recognition Systems, a Textbook for Students and Practitioners; Edition John Wiley & Sons Inc. p.326.

[6] **A. Belaïd, 2003.** Reconnaissance automatique de l'écriture et du document. Cours de LORIA-CNRS.

[7] **F. Grandidier, 2003.** Un nouvel algorithme de sélection de caractéristiques : application à la lecture automatique de l'écriture manuscrite», Thèse de doctorat en génie PH.D, ETS MONTRÉAL. pp.30-120

[8] **C. Saint-jean, 2001.** Classification paramétrique robuste partiellement supervisée en reconnaissance des formes. Thèse de doctorat, Université de La Rochelle - UFR Sciences. Laboratoire d'Informatique et d'Imagerie Industrielle.

[9] **Ayat, N. E., 2004.** Sélection de modèle automatique des machines à vecteurs de support: application à la reconnaissance d'images de chiffres manuscrits, Thèse de doctorat : Ecole de technologie supérieure : Université du Québec, 220 pages.

[10] **A. Gattal, 2009.** Segmentation automatique pour la reconnaissance numérique des chèques bancaires algériens, mémoire de magister : centre universitaire de kenchela.

[11] <http://lithwww.epfl.ch/teaching/rdf/>

[12] <http://www.sylbarth.com/nn.php>

[13] **Mohamadally Hasan, Fomani Boris 2006.** SVM : Machines à vecteurs de support ou séparateurs à vastes marges. Rapport technique.

[14] **Jérôme CALLUT 2003.** Implémentation efficace des support vector machines pour la classification. Université libre de Bruxelles.

[15] **Clément Chatelain 2006.** Extraction de séquences numériques dans des documents manuscrits quelconques. Thèse de doctorat, université de Rouen.

[16] **A. Belaïd, 2003.** Reconnaissance automatique de l'écriture et du document. Cours de LORIA-CNRS.

[17] <http://www.apstat.com>.

[18] **Sylvain Chevalier, Mélanie Lemaître, Edouard Geoffrois.** Étude de primitives spectrales pour la reconnaissance de caractères manuscrits dans le cadre d'une approche markovienne 2D. Rapport technique.

[19] **Trier, O.D. et Taxt, T., 1994,** Evaluation of binarization methods for document images. IEEE transactions on pattern analysis and machine intelligence, volume 17, number 3, march 1995.

[20] **Hou Biao, Liu Fang, Jiao Licheng.** Linear feature detection based on ridgelet. Science in china (Series E), Vol. 46 No. 2, April 2003.

[21] **Emmanuel J. Candes.** Monoscale Ridgelets for the Representation of Images with Edges. Department of Statistics Stanford University Stanford, California 94305-4065.

[22] **Emmanuel J. Candes 1998.** Ridgelets: Theory and applications. Thèse de doctorat (Phd), université de Stanford.

[23] **Samuel Foucher.** Application de la transformation en ondelette dans un espace géométrique pour la réduction du chatoiement. Centre de recherche en informatique de Montréal.

[24] **Chen, G.Y., Bui, T.D., et Krzyzak, A., 2006.** Rotation invariant feature extraction using ridgelet and Fourier transforms; Pattern Analysis and Application Journal, Vol. 9, 83-93.

[25] **H. Carsten, 2007.** The radon transform. Technical report. Aalborg university.

[26] **V. Chapelier 2005.** Codage progressif d'images par ondelettes orientées. Thèse de doctorat université de Rennes.

[27] **Po-Shen Loh. Mentor: Dr. Emmanuel J. Candes 2003.** Image reconstruction with ridgelets. Rapport de recherche.

[28] **T. Oriol Ramos, 2006.** Linear combination of multiresolution descriptors: Application to graphics recognition. Thèse de doctorat, université de Barcelone.

[29] Minimum redundancy feature selection from microarray gene expression data. Journal of Bioinformatics and Computational Biology Vol. 3, No. 2 (2005)

[30] **V. Michel, R. Fabrice, et F. Damien.** Advances in feature selection with mutual information. Similarity-Based Clustering, Villmann, Th.; Biehl, M.; Hammer, B.; Verleysen, M. (Ed.) (2009) 52-69.

[31] **Avrim L. Blum, Pat Langley.** Selection of relevant features and examples in machines learning. Artificial intelligence archive. Volume 97, Issue 1-2 (December 1997)

[32] **Mark A. Hall.** Correlation-based feature selection for machine learning. Thèse de doctorat. Université de Waikato, Newzealand 1999.

[33] **P. Hanchuan, L. Fuhui, and D. Chris.** Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE transactions on pattern analysis and machine intelligence, vol 27, no. 8, January 2005.

[34] F. Fleuret. Binary Feature Selection with Conditional Mutual Information. Rapport de recherche. INRIA.2003.

[35] Reconnaissance de formes par réseau de neurones 2002, rapport de recherche, école polytechnique, France.