

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE



UNIVERSITE MOHAMED BOUDIAF - M'SILA
FACULTE DE TECHNOLOGIE



DEPARTEMENT D'ELECTRONIQUE

Présenté pour l'obtention du diplôme de MASTER

Domaine : Sciences et Technologie Master

Filière : Électronique

Spécialité : Électronique des Systèmes Embarqués

Par :

Djerdali Chahra

Dif Khaoula

THEME

**Reconnaisances des chiffres arabes par
les moments invariants**

Soutenu publiquement le /06/2020 , devant le jury composé de :

Mr.	Univ M'sila	Président
Mm. Khedidja Derdour	Univ M'sila	Encadreur
Mr.	Univ M'sila	Examineur
Mr.	Univ M'sila	Examineur

L'année universitaire 2019/2020

Remerciement

Avant toute chose, nous remercions dieu tout puissant de nous avoir aidé et éclairé le chemin pour la réalisation de ce mémoire.

Nous remercions, notre encadreur Mme Khadidja Derdour, pour son aide et soutien précieux durant cette année, et nous remercions également tous ceux qui ont participé de loin ou de près à la réalisation de ce mémoire.

Dédicace

Je dédie cet humble travail à:

La seule femme et homme que j'aime du fond du

cœur, que Dieu les bénisse

Sauve-les pour moi, pour leur grand sacrifice et pour

leur soutien

Et leur amour: maman et papa.

Mes frères et sœurs

Tous mes amis.

Khaoula

Dédicace

Je dédie ce modeste travail,

- ❖ *A mon père qui s'est tant donné et sacrifié.*
- ❖ *A ma très chère mère qui m'a toujours soutenue.*

Dans mes joies comme mes peines.

- ❖ *A mon mari.*
- ❖ *A toute la famille sans exception.*
- ❖ *A tous mes amis.*

Chahra

Résumé

Un système de reconnaissance des chiffres arabes est fondé sur les étapes suivantes : prétraitements, l'extraction de caractéristiques et la classification.

Les moments invariants sont des descripteurs qui ont montré leur efficacité dans plusieurs domaines. Dans ce travail nous avons utilisé les moments Hu pour tirer profit de leur invariance à la translation, rotation et changement d'échelle. Ces caractéristiques ont fournis un taux de reconnaissance important en utilisant le classifieur K-Plus Proches Voisins.

Mots-clés : Reconnaissance des chiffres arabes, l'extraction des caractéristiques , la classification, moments de Hu, le classifieur KPPV .

Abstract

An Arabic numeral recognition system is based on the following steps: preprocessing, feature extraction and classification.

The invariant moments are descriptors which have shown their effectiveness in several fields. In this work we have used the Hu moments to take advantage of their invariance in translation, rotation and change of scale. These characteristics are provided a high recognition rate using the K- Neareat Neighbors classifier.

Key words: Arabic numerals recognition system, feature extraction, classification, Hu moments, the classifier K-Nearest Neighbors (KNN).

ملخص

يعتمد نظام التعرف على الأرقام العربية على الخطوات التالية: المعالجة المسبقة واستخراج الميزات والتصنيف.

العزوم الثابتة هي ميزات أظهرت فعاليتها في عدة مجالات. في هذا العمل، استخدمنا عزوم Hu للاستفادة من عدم تغييرها في حالة الحركة والدوران وتغير الحجم. هذه الميزات أعطت معدل تعرف عال باستخدام المصنف KPPV

الكلمات المفتاحية : نظام للتعرف على الأرقام العربية، استخراج الميزات، التصنيف، عزوم Hu، المصنف KPPV

TABLE DES MATIERES

REMERCIEMENT.....	2
DÉDICACE.....	3
DÉDICACE.....	4
RÉSUMÉ.....	1
ABSTRACT.....	1
ملخص.....	1
TABLE DES MATIERES	1
LISTE DES FIGURES.....	3
LISTE DES TABLEAUX	3
LISTE DES ABREVIATIONS.....	4
INTRODUCTION GENERALE.....	5
RECONNAISSANCES DE FORMES.....	10
1 RECONNAISSANCE DE FORMES	11
1.1 INTRODUCTION :	11
1.2 DÉFINITION :	11
1.3 APPLICATIONS TYPIQUES DE LA RECONNAISSANCE DES FORMES :.....	12
1.4 RECONNAISSANCE DE L'ÉCRITURE ET DOMAINE D'APPLICATION :	13
1.4.1 LA RECONNAISSANCE DE L'ÉCRITURE MANUSCRITE :	14
1.4.1.1 Les systèmes en lignes :	14
1.4.1.2 Les systèmes hors lignes :	15
1.4.2 ÉCRITURE IMPRIME :	16
1.5 LE SCHÉMA GÉNÉRAL D'UN SYSTÈME DE RDF :	16
1.5.1 PRETRAITEMENT :	17
1.5.2 EXTRACTION DES CARACTERISTIQUES :	19
1.5.3 CLASSIFICATION :	21
1.5.3.1 L'apprentissage non supervisé :	21
1.5.3.2 L'apprentissage supervisé :	22
1.6 LES APPROCHES DE LA RECONNAISSANCE DES FORMES :	24
1.6.1 LES APPROCHES SYNTAXIQUES :	24
1.6.2 LES APPROCHES STATISTIQUES :	25
1.7 CONCLUSION :	25
LES MOMENTS INVARIANTS	26
2 LES MOMENTS INVARIANTS.....	27
2.1 INTRODUCTION :	27
2.2 CONSTRUCTION DE MOMENTS INVARIANTS:	27
2.3 LES MOMENTS DE TCHEBICHEF :	29
2.4 LES MOMENTS DE ZERNIKE :	30
2.5 LES MOMENTS DE HU :	31

2.6	INVARIANTS DU MOMENT AFFINE :	33
2.7	LES MOMENTS DE LEGENDRE :	34
2.8	LES MOMENTS DE KRAVTCHOUK :	35
2.9	LES MOMENTS DE HAHN :	37
2.10	CONCLUSION :	38
3	IMPLÉMENTATION ET RÉSULTATS	41
3.1	INTRODUCTION :	41
3.2	RESSOURCES MATÉRIELLES ET LOGICIELLES :	41
3.3	DESCRIPTION DE NOTRE SYSTÈME DE RECONNAISSANCE :	42
3.3.1	DESCRIPTION DE LA BASE :	42
3.3.2	PRETRAITEMENTS :	44
3.3.3	EXTRACTION DES CARACTERISTIQUES :	44
3.3.4	CLASSIFICATION :	45
3.3.4.1	<i>Principe de classifieur KPPV :</i>	46
3.3.4.2	<i>Choix du nombre K :</i>	47
3.3.4.3	<i>Sélection de la classe :</i>	47
3.3.4.4	<i>Définition de la distance :</i>	48
3.3.4.5	<i>L'apprentissage :</i>	49
3.4	RÉSULTATS ET DISCUSSION :	50
3.4.1.1	<i>Images grises</i>	50
3.4.1.2	<i>Images binaires</i>	51
3.5	CONCLUSION :	54
	CONCLUSION GENERALE	56
	BIBLIOGRAPHIE	57

LISTE DES FIGURES

Figure 1: Dichotomie des types d'écriture.	14
Figure 2: Reconnaissance de l'écriture manuscrite en ligne et hors-ligne.....	15
Figure 3: Schéma général d'un système de reconnaissance des formes.	17
Figure 4: Représentation de l'apprentissage supervisé et non supervisé.	22
Figure 5 : Exemple d'un réseau de neurones à trois couches.....	23

LISTE DES TABLEAUX

Tableau 1: les distance et leur formule.	49
Tableau 2: Le taux de classification pour la base de données (50% apprentissage / 50% Test).....	50
Tableau 3: Le taux de classification pour la base de données après le binarisation (50% apprentissage / 50% test).	51
Tableau 4: Le taux de classification pour la base de données (70% apprentissage / 30% test).	52
Tableau 5: Le taux de classification pour la base de données (60% apprentissage / 40% test).	53
Tableau 6: le taux de classification de différent valeur de k	54

LISTE DES ABBREVIATIONS

- EEG : Electro Encéphalo Gramme.
- ECG : Electro Cardio Gramme.
- RDF : Reconnaissance Des Formes.
- OCR : Optical Character Recognition.
- HTR : Handwritten Tescte Recognition.
- LDA : Linear Discriminant Analysés.
- ANN : Artificial Neural Networks.
- KNN : K-Neareat Neighbor.
- SVM : Support Vector Machine.
- K-PPV : K-Plus Proches Voisins.
- TRS : Translation, Rotation and Scaling.
- ZM : Zernike moments.
- CBIR : Content Based Image Retrieval.
- ZOA : Approsimation d'Ordre Zéro.
- DSP : Digital Signal Processor.



INTRODUCTION GENERALE

INTRODUCTION GENERALE

L'un des principaux objectifs de l'ordinateur est d'automatiser les tâches généralement effectuées par des humains. La plupart des activités humaines reposent sur une faculté importante de notre cerveau : voir distinguer les formes. Par exemple, reconnaître une pomme à partir d'une orange, un bon diagnostic d'un mauvais, ou un sous-marin d'une baleine. Pour pouvoir automatiser ce genre de tâche, l'ordinateur doit acquérir la capacité de reconnaître des formes.

La reconnaissance a besoin d'un modèle de l'objet. Pour l'homme, ce modèle correspond à une représentation mentale de l'objet qui s'apprend en conservant les caractéristiques les plus discriminantes de l'objet. Les caractéristiques peuvent être toutes sortes d'attributs de l'objet : forme, couleur, texture, taille, volume, etc.

Cette discipline prend aujourd'hui un essor considérable dans l'industrie, car les progrès de l'électronique permettent désormais de fournir des systèmes opérationnels à forte puissance de calcul à des prix raisonnables. Il rassemble un certain nombre de techniques allant de la simple identification d'objets isolés à l'analyse de scènes complexes ou à la compréhension de la parole. Dans ces derniers cas, il faut à la fois segmenter en objets, identifier ces objets et établir un réseau de relations entre objets.

La notion d'apprentissage est au cœur de la plupart des techniques développées. Ces techniques peuvent être utilisées soit seules dans le cas de systèmes dont la mission principale est d'identifier des objets, soit en lien étroit avec les techniques d'intelligence artificielle dans le cas de systèmes devant à la fois percevoir et raisonner des choses perçues. .

Parallèlement aux travaux sur les méthodes de reconnaissance, le traitement d'image et la vision par ordinateur ont été développés. Ces domaines ont focalisé le problème de la reconnaissance sur des données spécifiques, mais en revanche ils ont permis de situer la reconnaissance dans un processus plus large d'interprétation d'images impliquant des niveaux de perception et de connaissances spécifiques au domaine.

L'objectif de la RDF est de réaliser des systèmes informatisés qui simulent les activités humaines de perception, de reconnaissance et de compréhension : reconnaissance de l'écrit, de la parole, interprétation de scènes, robotique, reconnaissance de signaux médicaux EEG (électroencéphalogramme), ECG (électrocardiogramme). Cela implique aussi une certaine pluridisciplinarité pour comprendre l'aspect physique des capteurs, les aspects mathématiques de la classification, ceux relatifs à l'informatique.

Le processus de reconnaissance des chiffres est réalisé en deux phases : la première phase correspond à la conception du classifieur, qui est achevée par un apprentissage automatique de ce dernier. Tandis que la deuxième phase consiste à évaluer les performances du système à partir d'un ensemble de données appelé base de teste.

Les moments d'image et leurs fonctions ont été utilisés comme caractéristiques dans de nombreuses applications de traitement d'image, à savoir la reconnaissance de formes, la classification d'images, l'identification de cibles et l'analyse de formes. Les moments d'une image sont traités comme des descripteurs de forme basés sur la région.

Pour mener à bien ce travail, nous l'avons divisé en trois chapitres :

- Le premier chapitre contient une description complète du problème de la reconnaissance de formes, avec une présentation des différentes étapes d'un système de reconnaissance de l'écriture.
- Le deuxième chapitre présente les différentes méthodes d'extraction des moments invariants et leurs équations.
- Le dernier chapitre concerne notre application de reconnaissance des chiffres arabes ; nous présentons l'approche suivie pour sélectionner les paramètres de moment invariant et les résultats obtenus.



CHAPITRE I
RECONNAISSANCES DE FORMES

1 Reconnaissance de formes

1.1 Introduction :

La reconnaissance de formes (ou RDF) provient de différentes disciplines que sont les mathématiques (probabilités et statistiques), les sciences de l'ingénieur, l'informatique et l'intelligence artificielle. C'est à partir des années 1960 que la reconnaissance des formes est devenue une discipline spécifique. Le développement extraordinaire des ordinateurs ces dernières années a donné une impulsion au RDF en permettant des applications en temps réel, en particulier dans le domaine des applications visuelles et auditives. Les processus d'acquisition tels que les caméras et les scanners sont très accessibles, tout comme les ordinateurs à la fois puissants et peu coûteux. Ils permettent le traitement d'un grand nombre de données dans un délai raisonnable comme cela est souvent nécessaire en RDF.

Le problème que la reconnaissance de formes cherche à résoudre est d'associer une étiquette à des données qui peuvent se présenter sous la forme d'une image brute ou d'un signal. Différentes données peuvent recevoir la même étiquette ; ces données sont les réalisations ou les copies de la classe identifiées par l'étiquette. Par exemple, l'écriture du caractère A varié d'un écrivain à l'autre, mais le lecteur identifiera le caractère A pour chacune de ces réalisations.

Dans ce chapitre, nous allons passer en revue quelques techniques des plus connues permettant la reconnaissance de formes pour en choisir une que nous intégrerons dans notre application.

1.2 Définition :

On désigne par reconnaissance des formes (ou parfois reconnaissance de motifs) un ensemble de techniques et méthodes visant à identifier des motifs à partir de données brutes afin de prendre une décision dépendant de la catégorie attribuée à ce motif.

On considère que c'est une branche de l'intelligence artificielle qui fait largement appel aux techniques d'apprentissage automatique et aux statistiques. (Djabeur, 2017)

On considère souvent la Reconnaissance de formes comme un problème de classification : trouver la fonction qui affecte à toute forme inconnue sa classe la plus pertinente. Elle est partie intégrante de tout système intelligent destiné à la prise de décision (Theodoridis, 2003)

Les systèmes de reconnaissance de formes intègrent toute la chaîne perception reconnaissance depuis l'acquisition des données brutes jusqu'à la compréhension élaborée de ces données. Ces dernières ont entre temps subi de nombreuses transformations.

1.3 Applications typiques de la reconnaissance des formes :

- **Marketing** : La reconnaissance des formes est souvent utilisée pour classer les consommateurs selon les produits qu'ils sont susceptibles d'acheter. Elle est aussi utilisée par les sociétés de vente pour classer les clients selon qu'ils soient de bons ou mauvais payeurs, ou encore selon qu'ils vont oui ou non passer à la concurrence
- **Finances** : les systèmes de reconnaissance des formes sont utilisés pour la détection de transactions bancaires frauduleuses ainsi que la prédiction des banqueroutes
- **Usinage** : la qualité des produits dépend souvent de paramétrisation correcte, et les relations exactes entre la qualité et les valeurs des paramètres n'est pas claire. Les systèmes de reconnaissance des formes sont utilisés pour classer les paramètres selon la qualité des produits qu'ils sont susceptibles de générer. Ils permettent ainsi de réduire le nombre d'essais ce qui fait gagner du temps et de l'argent.
- **Energie** : les systèmes de reconnaissance des formes sont utilisés pour prévoir la consommation électrique (réduite, normale, élevée), permettant ainsi aux clients de réduire si nécessaire leur consommation, et aux producteurs de mieux gérer leurs unités de production.
- **Lecture automatisée** : les systèmes de reconnaissance des formes permettent de numériser les anciens documents ainsi que les archives, non pas sous la forme d'images, mais plutôt sous une forme textuelle.
- **Sécurité** : la reconnaissance vocale et rétinienne sont un exemple d'applications typiques de la reconnaissance des formes pour l'authentification. La vérification des signatures est aussi très populaire. (Djabeur, 2017)

1.4 Reconnaissance De L'écriture et Domaine D'application :

La reconnaissance de l'écriture est mieux connue sous le nom d'O.C.R (Optical Character Recognition). L'O.C.R connaît plusieurs applications pratiques dans plusieurs domaines parmi lesquels on cite (Bendjeddou *et al.* 2008) :

- Domaine bancaire
- Assistance à l'éducation
- Lecture des adresses postales
- La police et la sécurité

Il n'existe pas un système universel d'O.C.R qui peut traiter tous les cas de l'écriture, mais plutôt des voies d'approche dépendant du type de données traitées et bien évidemment de l'application visée (Belaid, 1992)

On utilise autrefois le terme O.C.R ou reconnaissance de caractère qui appartient au domaine lecture automatique de document qui couvre trois axes :

- a) Lecture optique des caractères imprimés
- b) La reconnaissance de caractères manuscrits
- c) La reconnaissance de l'écriture manuscrite cursive.

La reconnaissance de caractères signifie le décodage de n'importe quel texte imprimé ou manuscrit ou information symbolique, par exemple les formules mathématiques...etc. On présente dans la figure 1 une dichotomie des types d'écriture faite selon le média, le support saisi et l'application considérée. La reconnaissance des caractères manuscrits est considérée plus difficile que la reconnaissance de l'imprimé pour lequel des systèmes opérationnels commerciaux existent. (Bendjeddou *et al.* 2008)

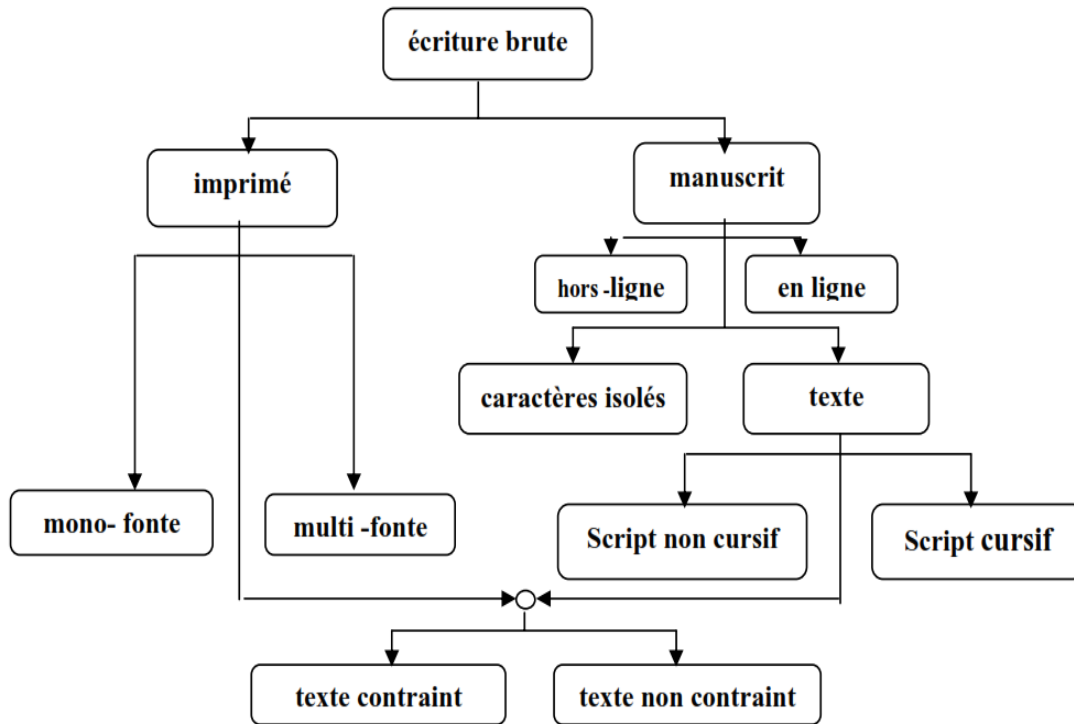


Figure 1: Dichotomie des types d'écriture.

1.4.1 La reconnaissance de l'écriture manuscrite :

En anglais, *handwritten text recognition* ou *HTR* : est un traitement informatique qui a pour but de traduire un texte écrit en un texte codé numériquement.

Deux Systèmes de reconnaissances distinctes, avec des problématiques et des solutions différentes.

1.4.1.1 Les systèmes en lignes :

Les systèmes en lignes reconnaissent le texte en temps réel. Les symboles sont reconnus en même temps qu'ils sont écrits à la main. L'écriture est présentée par un vecteur (x, y) dont les coordonnées sont en fonction du temps. L'acquisition de l'écriture est assurée par une tablette graphique équipée d'un stylo électronique. Ce système est généralement conçu pour la reconnaissance du manuscrit. Il opère en temps réel pendant l'écriture. (Amara, 2012)

1.4.1.2 Les systèmes hors lignes :

Les systèmes hors lignes démarrent suite à l'acquisition. Ils servent à la reconnaissance des textes écrits sur papier après leurs numérisations sous forme d'image. Les systèmes hors lignes jouent un rôle très important dans le développement de la société d'information, puisque la plupart des textes qui nous intéressent existent sur papier.

Dans la littérature, il existe des systèmes hors ligne qui reconnaissent le manuscrit et d'autres qui reconnaissent l'imprimé. Dans le cas du manuscrit, la plupart des caractères sont ligaturés d'où la nécessité de l'emploi de techniques de délimitation spécifiques pour guider la lecture. Les variations inter et intra-scripteurs sont accentuées par la nature calligraphique de l'écriture arabe (Fethi, 1999).

Nous distinguons trois types de systèmes de reconnaissance d'écriture qui dépendent de ces variations (Amara, 2012) :

- **Mono scripteur** : Le système de reconnaissance peut reconnaître l'écriture d'un seul scripteur après apprentissage de son écriture ;
- **Multi scripteur** : Le système de reconnaissance est capable de reconnaître les écritures d'un groupe restreint de scripteurs soit par apprentissage de leurs écritures, soit sans apprentissage ;
- **Omni scripteur** : Le système est capable de reconnaître toutes les écritures. Dans ce type de système, la variabilité intra-scripteur s'ajoute à la variabilité inter-scripteur

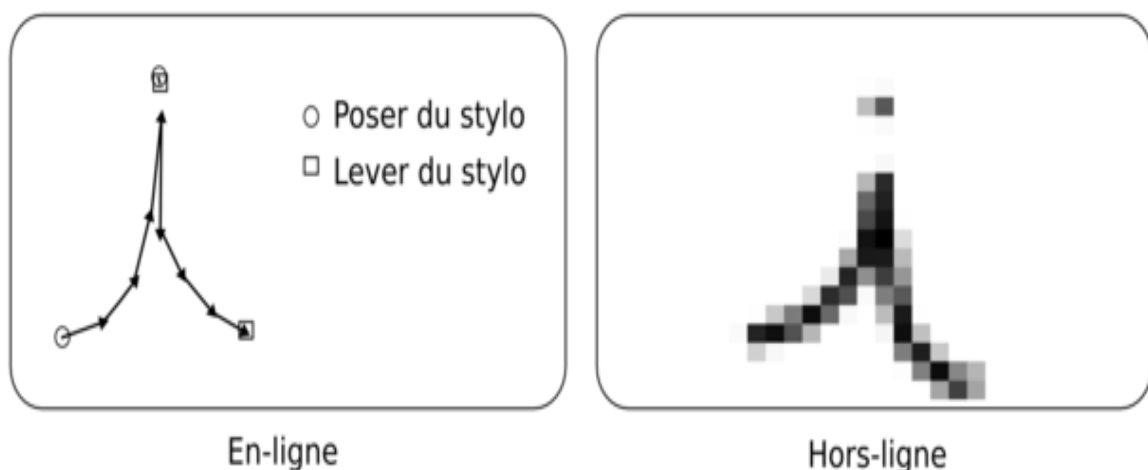


Figure 2: Reconnaissance de l'écriture manuscrite en ligne et hors-ligne.

La reconnaissance de l'écriture manuscrite fait appel à la reconnaissance de forme, mais également au traitement automatique du langage naturel. Cela veut dire que le système, tout comme le cerveau humain, reconnaît des mots et des phrases existant dans un langage connu plutôt qu'une succession de caractères. Ceci améliore grandement la robustesse.

1.4.2 Ecriture Imprimé :

Écriture imprimée, aussi appelé écriture script, est une écriture manuscrite propre aux systèmes d'écriture latin et, par extension, grec et cyrillique, dans lequel les lettres ne sont pas liées les unes aux autres et dont le tracé correspond approximativement, en le simplifiant, à celui des caractères typographiques utilisés en imprimerie. Elle s'oppose habituellement à l'écriture cursive d'aujourd'hui, aussi appelée écriture courante, où les traces simplifiées des caractères sont liées (cependant toutes les écritures dites cursives ne sont pas ou n'ont toujours été liées).

Dans le cas de l'imprimé, les caractères sont souvent séparés verticalement, ce qui simplifie la lecture. Certaines fontes présentent des accolements qu'il faut défaire. Dans le cas de l'imprimé la reconnaissance peut être (Amara, 2012) :

- **Monofonte** : Le système ne traite qu'une fonte à la fois. L'apprentissage est simple puisque l'alphabet représenté est réduit.
- **Multi fonte** : Le système es peut reconnaître un mélange de fontes parmi un ensemble préalablement apprises. Le prétraitement doit, alors, réduire les écarts entre les caractères (taille, épaisseur et inclinaison). De même, l'apprentissage doit gérer les ressemblances de caractères des différentes fontes.
- **Omni fonte** : Le système est capable de reconnaître toutes fontes sans l'avoir apprise

1.5 Le schéma général d'un système de RDF :

Du signal écriture sous ses différentes formes, à la prise de décision par un système, il existe un certain nombre d'étapes à mettre en œuvre.

La figure 3 représente globalement le processus général de reconnaissance de l'écriture imprimé. Dans un premier temps une phase de prétraitement est réalisée sur l'image acquise. Elle permet de réduire au maximum la variabilité intrinsèque à l'écriture ainsi que les bruits possiblement introduits lors de l'acquisition. Après, optionnelle, est celle de la segmentation. L'écriture étant une concaténation de

caractères, il est normal lors de la reconnaissance d'essayer de segmenter l'écriture à reconnaître en caractères.

La Deuxième étape à être réalisée directement sur les données présentées en entrée du système est l'extraction de caractéristiques. Son but est la réduction de la quantité d'information et l'extraction des caractéristiques les plus pertinentes pour la reconnaissance.

La prochaine étape et la dernière étape dans un système de reconnaissance de texte est la classification, c'est le stade de décision dans un système de reconnaissance de texte. A cette étape les primitives extraites dans l'étape précédente sont utilisées pour identifier le segment de texte selon des règles établies préalablement. Généralement, à ce niveau, on utilise des modèles obtenus dans une phase d'apprentissage pour classifier les données de test. Le figure 3 représentée le schéma général d'un système de reconnaissance des formes. (Mehennaoui, 2006).

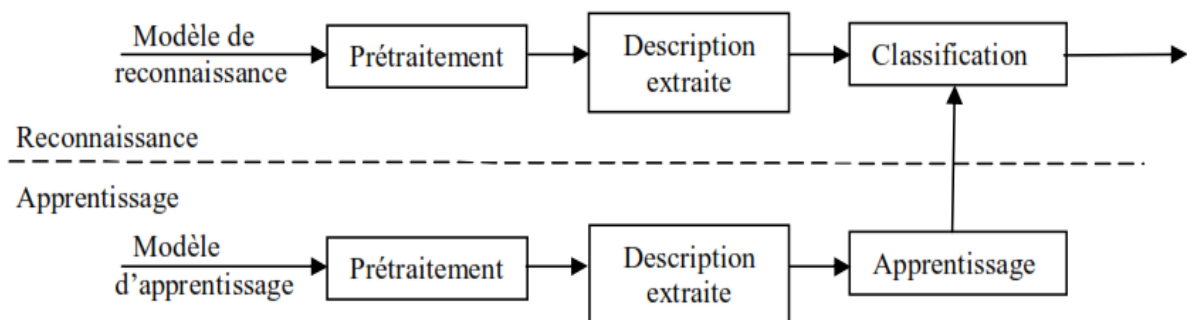


Figure 3: Schéma général d'un système de reconnaissance des formes.

1.5.1 Prétraitement :

Les données brutes issues des capteurs sont les représentations initiales des données à partir desquelles des traitements permettent de construire celles qui seront utilisées pour la reconnaissance.

Les données brutes sont bruitées, elles contiennent des informations parasites, et elles n'explicitent pas les informations utiles pour la reconnaissance.

La plupart des systèmes comportent une étape de prétraitement. Généralement, ces prétraitements ne sont pas spécifiques à la reconnaissance de texte, mais sont des prétraitements classiques basé sur le traitement d'images qui désigne un ensemble de méthodes dont l'objectif est de transformer des images (pour en améliorer l'apparence ou pour les coder de manière plus compacte en vue d'une transmission), soit d'en extraire de

l'information (par exemple lorsqu'il s'agit de reconnaître automatiquement l'écriture). (Mehennaoui, 2006)

Son objectif est de faciliter la caractérisation de la forme (lettre, chiffre, mot) ou de l'entité à générer soit en nettoyant la forme (en éliminant le bruit), soit en réduisant le bruit chevauchant les données et en les sauvegardant, autant que possible un grand nombre d'informations du formulaire présenté. Ainsi, le but du prétraitement est de préparer l'image d'histogramme pour la prochaine étape d'analyse.

Parmi les opérations de prétraitements généralement utilisées, citons : binarisation, lissage, normalisation de la taille, squelettisation, redressement, et la segmentation.

- **Binarisation :**

La binarisation est une opération qui permet le passage d'une image de niveaux de gris à une image binaire composée de 2 valeurs 0 et 1, plus simple à traiter. (Megar, 2016)

- **Lissage :**

L'image des caractères peut être entachée de bruit introduit durant l'acquisition et au cours des différentes transformations. Ce bruit correspond soit à des absence de points (trous), soit à des empâtements et donc à une surcharge de points. Le lissage consiste à examiner le voisinage d'un pixel et éliminer les pixels isolés d'une part (nettoyage), et à boucher les trous vide d'autres part (bouchage) (Belaid, 1992)

- **Normalisation de la taille :**

La taille d'un caractère peut varier d'une police à une autre et même au sein d'une même police après l'avoir agrandie ou réduite. Ce qui peut conduire à une instabilité des paramètres. Une technique de prétraitement naturelle consiste à réduire les caractères à la même taille. Par conséquent, la normalisation de la taille permet de réduire les images à la même taille. (Mehennaoui, 2006)

- **Squelettisation :**

La squelettisation permet d'obtenir une épaisseur égale à un pixel de la ligne d'écriture et donc de la réduire à une écriture linéaire. Le squelette doit conserver la forme, la connectivité, la topologie et les extrémités du chemin, et ne doit pas introduire d'éléments parasites ; et donc simplifier la représentation. (Megar, 2016)

- **Redressement :**

Un positionnement incorrect du document dans le scanner entraîne une inclinaison de l'image. Pour une bonne utilisation des documents, il est donc nécessaire de détecter l'angle d'inclinaison et de redresser les images de sorte que les lignes de texte soient parallèles aux bords. L'angle du document (ou de la ligne), est celui qui correspond à l'histogramme entropie maximale. Pour corriger ce biais, appliquez simplement un angle de rotation à l'image.

Segmentation :

Fondamentalement, la segmentation est un processus qui consiste à découper une image en régions connexes présentant une homogénéité selon un certain critère, comme par exemple la couleur. L'union de ces régions doit redonner l'image initiale. La segmentation est une étape importante pour l'extraction d'informations qualitatives de l'image.

1.5.2 Extraction 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 d'extraction de caractéristiques est réalisée. C'est une phase critique lors de la construction d'un système de reconnaissance. L'une des raisons pour lesquelles cette étape pose un problème est qu'une grande majorité des techniques d'extraction s'accompagne d'une perte d'information irrémédiable. (Mehennaoui, 2006)

La description sera évidemment beaucoup plus simple pour une reconnaissance de bas niveau de type global que pour une reconnaissance symbolique (structurel).

Dans le premier cas, par exemple, on pourra se contenter de détecter les intersections du caractère avec une grille fixe, dans le deuxième on pourra décrire avec un véritable langage, les segments, la concavité et les boucles du caractère analysé. Dans le premier cas l'image est transformée en un vecteur de valeurs numériques, dans le deuxième en une suite de symboles, obtenues par une analyse de caractère en formes élémentaires.

Dans le domaine de la reconnaissance des formes, il est intéressant d'essayer d'extraire, à partir des images, des informations "non visibles". Elles utilisent une transformation globale de manière à changer d'espace de représentation et ainsi faciliter l'extraction de caractéristiques pertinentes, il y'a plusieurs techniques utilisée par exemple :

•La transformée de Fourier :

C'est l'une des méthodes les plus utilisées en reconnaissance de formes et de caractères (Filatov *et al.* 1998). Les caractéristiques extraites sont en fait les descripteurs de Fourier basés sur les coefficients complexes des séries de Fourier. Elles sont invariantes aux rotations et aux changements d'échelle (Mahmoud, 1994). La propriété d'invariance aux rotations implique des problèmes de reconnaissance de certains caractères comme "6" et "9". Il faut donc rajouter d'autres caractéristiques au vecteur de manière à régler ce problème (Mehennaoui, 2006).

•Les ondelettes :

Une autre transformation globale, assez proche de celle de Fourier, Le principal intérêt des ondelettes est que ces dernières permettent d'obtenir une information fréquentielle localisée concernant un signal ou une fonction de base choisie. Malgré certains avantages, cette technique est peu utilisée en reconnaissance de formes. La raison en est que les caractéristiques extraites ne sont pas invariantes à la translation (Mehennaoui, 2006).

•La transformée de Hough :

Cette méthode est utilisée pour représenter le squelette d'un caractère comme un ensemble de segments de ligne puis utilisent la longueur, la position, et la pente des segments de ligne au squelette comme des caractéristiques.

•Les moments invariants :

Une autre grande famille appartenant à cette catégorie, L'invariance recherchée est liée à la rotation, à la translation et au Changement d'échelle. Les caractéristiques extraites par ces techniques sont considérées comme le résultat d'une transformation globale appliquée uniquement aux pixels de la forme analysée. Il existe plusieurs formulations des moments invariants, comme celle de Hu. Cependant, les plus utilisées actuellement sont celles dérivées des polynômes de Zernik (Grandidier, 2003). La raison est que ces derniers ont des performances supérieures en termes d'invariance (Mehennaoui, 2006).

Ces moments sont des mesures statistiques de la distribution des points dans l'image et ils peuvent être exprimés directement à partir des moments centraux normalisés d'ordre 2 et 3.

Les caractéristiques peuvent être classées en deux catégories :

- ❖ **Les caractéristiques locales** : qui sont souvent géométriques (par exemple : concavité/convexité, les fins de traits, les jonctions (en T ou Y), ainsi que les intersections en (X)).
- ❖ **Les caractéristiques globales** : qui sont souvent topologiques (connectivité, nombre de composantes connexes, ...etc.) ou statistiques (transformé de Fourier, moments invariants...etc.).

1.5.3 Classification :

Le rôle du classifieur est de se prononcer sur l'appartenance d'une forme à chacune des classes à partir du vecteur de caractéristiques par un critère de décision.

L'apprentissage a pour but de généraliser l'information portée par les exemplaires de l'ensemble d'apprentissage pour faire émerger les représentations des classes. Ces représentations doivent permettre de s'affranchir de la variabilité des exemplaires d'une classe lors de la prise de décision. La validation de l'apprentissage est faite par les taux de reconnaissance obtenus sur un ensemble test constitué d'exemplaires différents de ceux de l'ensemble d'apprentissage. Pour les démarches qui excluent cet apprentissage, le critère de décision (une mesure de ressemblance) devra supporter seul la variabilité intra-classe.

D'une manière générale, nous distinguons Deux types de techniques d'apprentissage : supervisé et non supervise.

1.5.3.1 *L'apprentissage non supervisé :*

La classification non supervisée ne nécessite pas une base d'images étiquetées pour faire l'apprentissage (training) du classifieur utilisé. Ce type d'apprentissage consiste à doter le système d'un mécanisme automatique lui permettant de trouver les classes de référence avec une assistance minimale. Des échantillons sont introduits en grand nombre par l'utilisateur sans indication de leur classe.

L'objectif de ce type de classification consiste à créer des regroupements d'objets (contenus dans l'image) homogènes (des classes) en utilisant uniquement des informations propres à ces objets (les caractéristiques extraites à partir des objets). Pour cela, la classification non supervisée se base essentiellement sur la similarité entre les objets considérés en supposant que ceux qui sont proches les uns des autres font partie de la même classe (Augereau, 2013)

Ce type d'apprentissage est intéressant car il permet de renseigner l'utilisateur sur les ambiguïtés entre les caractères afin d'agir en conséquence, par exemple, en ajoutant des échantillons pour renforcer la représentative d'une classe. Ce type d'apprentissage est appelé apprentissage automatique, dans le figure 4 suivante nous avons présentée l'apprentissage supervisé et non supervisé.

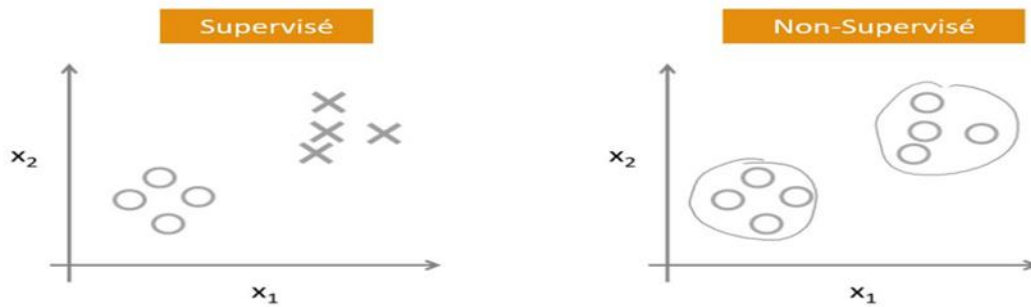


Figure 4: Représentation de l'apprentissage supervisé et non supervisé.

1.5.3.2 L'apprentissage supervisé :

A l'inverse de la classification non supervisée, la classification supervisée nécessite une base d'images labellisée qui sera utilisée pour effectuer un apprentissage du classifieur (Jacques-Henri, 2012). Cet apprentissage permet de déterminer les paramètres du classifieur qui seront par la suite utilisés pour la classification de nouvelles images.

Il est réalisé lors d'une étape préliminaire à la reconnaissance, en introduisant un grand nombre d'échantillons de référence. L'utilisateur joue le rôle de professeur pour indiquer le nom de chaque échantillon. Ce type d'apprentissage est appelé apprentissage manuel.

Parmi les méthodes de classification supervisées, on cite : l'analyse discriminante linéaire, les réseaux de neurones artificiels, l'arbre de décision, le classifieur de Bayes, les K plus proches voisins et les machines à vecteurs de support.

- **L'analyse discriminante linéaire :**

Le principe de l'analyse discriminante linéaire (en anglais Linear Discriminant Analysis : LDA) consiste à séparer les classes d'objets en construisant les limites de décision par optimisation du critère d'erreur (Cheng *et al.*, 2006).

- **Les réseaux de neurones artificiels :**

Les réseaux de neurones artificiels (en anglais Artificial Neural Networks ANN) sont très utilisés pour la détection. Ils tirent leur inspiration du système nerveux biologique. Ainsi, leur architecture est notamment semblable à celle du cerveau humain. Un ANN est donc composé d'un ensemble de couches : une couche d'entrée qui est alimentée par les valeurs de descripteurs, une couche de sortie qui représente le résultat de la classification et une ou plusieurs couches cachées (Cheng et *al.*, 2006).

La figure 5 montre l'architecture d'un réseau de neurones à trois couches. L'avantage des ANNs réside dans leur capacité d'auto-apprentissage, ce qui rend cette méthode adaptée à la résolution de problèmes complexes. Néanmoins, l'inconvénient majeur relatif à l'application des ANNs réside dans le fait qu'il est difficile de choisir le nombre de couches cachées ainsi que le nombre de neurones composant chaque couche.

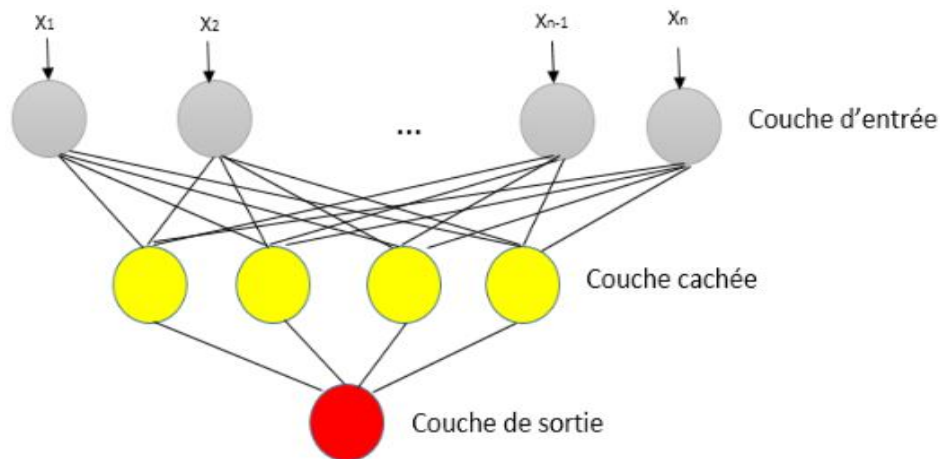


Figure 5 : Exemple d'un réseau de neurones à trois couches

- **L'arbre de décision:**

L'arbre de décision (en anglais Decision tree) est une méthode de classification efficace. Au lieu d'utiliser la totalité de l'espace de caractéristiques, cette méthode procède par divisions successives en sous-ensembles de caractéristiques. Ces sous-ensembles seront par la suite utilisés à des niveaux d'arbre différents (Karthikeyan et *al.*, 2013). En particulier, l'arbre de décision binaire divise récursivement l'espace de caractéristiques en deux sous espaces en sélectionnant à chaque fois un seuil permettant de séparer les données d'entrée en deux classes. Ce processus est itéré jusqu'à ce que les données de chaque nœud fils soient plus homogènes que les données dans le nœud père.

Comparée au réseau de neurones, l'arbre de décision est plus simple et plus rapide (Cheng *et al.*, 2006)

- **Le classifieur de Bayes:**

Il s'agit d'une méthode statistique qui fonctionne selon le principe du théorème de Bayes. Le but de ce classifieur est de maximiser la probabilité de bonne classification et, en contre partie, minimiser l'erreur de classification (Santhi, 2015)

- **Les K plus proches voisins :**

Le principe de base des K plus proches voisins (en anglais K-Nearest Neighbor KNN) consiste à attribuer une classe à un nouvel exemple en choisissant celle de la majorité de ses voisins qui lui sont les plus proches dans la base d'apprentissage. La recherche des K plus proches voisins est basée sur la similarité entre les caractéristiques. Cette similarité peut être mesurée par plusieurs métriques de distance telle que la distance euclidienne ou la somme des différences de distances (EOsman *et al.*, 2009).

- **Les machines à vecteurs de support :**

Le principe des machines à vecteurs de support (en anglais Support Vector Machine SVM) consiste à séparer les classes dans un espace de caractéristiques de grande dimension en construisant un hyperplan optimal qui maximise la marge de séparation entre ces classes (Kharroubi, 2002).

1.6 Les approches de la reconnaissance des formes :

La reconnaissance de formes se base sur la définition d'algorithmes permettant de classer des formes en les comparant à des formes-types. Elle intervient dans de nombreux domaines tels que la reconnaissance vocale, la reconnaissance de caractères, l'automatisation industrielle, le diagnostic médical, la classification de documents, etc. De manière générale on distingue deux types de reconnaissance des formes :

1.6.1 Les approches syntaxiques :

Les approches syntaxiques, structurelles et la mise en correspondance des formes (Template Matching), utilisent les hypothèses sur les distributions des données à l'intérieur des classes. Les procédures de classification, dans ce cas, sont construites à l'aide d'hypothèses probabilistes (Ex. classement Bayésien). Elles se basent sur une représentation

des formes à l'aide de grammaires et nécessitent des moyens de calculs importants. De ce fait, ce type d'approche n'est pas approprié aux applications en temps réel (Amar, 2008).

1.6.2 Les approches statistiques :

La reconnaissance de formes statistique, contrairement à la précédente, s'appuie sur une représentation numérique des formes en se basant sur des méthodes paramétriques ou non paramétriques. Les méthodes non paramétriques ne posent pratiquement aucune hypothèse sur la forme des distributions (Ex. K-plus proches voisins). Parmi les méthodes les plus utilisées, on cite les méthodes connexionnistes qui constituent un sous-ensemble des méthodes statistiques (Amar, 2008).

1.7 Conclusion :

La reconnaissance de formes dans les images est un des problèmes les plus difficiles en vision par ordinateur puisqu'il est indispensable pour toutes sortes d'applications.

La vision par ordinateur ne cherche pas à comprendre ou à reproduire la vision humaine, mais à construire un modèle algorithmique qui vu de l'extérieur, possède des propriétés semblables. D'où la vision artificielle est d'extraire des caractéristiques à partir des images. C'est une mise en correspondance image/reconnaissance. Connaître la forme nécessite une bonne acquisition.



CHAPITRE II

LES MOMENTS INVARIANTS

2 Les moments invariants

2.1 Introduction :

L'extraction des caractéristiques est l'une des étapes les plus importantes de tout système de reconnaissance. Elle vise à extraire de l'image d'une classe donnée information pertinente permettant de la distinguer plus facilement des autres classes (Chevalier, 2004).

Les moments géométriques ont largement été utilisés pour la description d'images ou de formes. L'ensemble des moments calculés pour une image représente généralement ses caractéristiques globales, et fournit plus précisément une quantité d'information importante sur les primitives géométriques contenues dans l'image (Jaeger, 2006).

Leur principe consiste à projeter une distribution (image) sur des fonctions de base. Les moments d'ordre faible représentent les caractéristiques globales de la forme, tandis que les moments d'ordre élevé caractérisent les détails de l'image. (Chergui, 2013).

Dans ce chapitre nous allons présenter quelques types de moments à savoir : les moments de Tchebichef, les moments de Zernike, les moments de Hu, les moments Affine, les moments de Legendre, les moments de Kravtchouk et les moments de Hahn. On va expliquer le principe de chaque type.

2.2 Construction de moments invariants:

Une question essentielle dans le l'analyse des motifs est la reconnaissance des objets et des caractères indépendamment de leur position, taille et orientation, comme illustré dans la figure6. (Muharrem et Gulez, 2005)

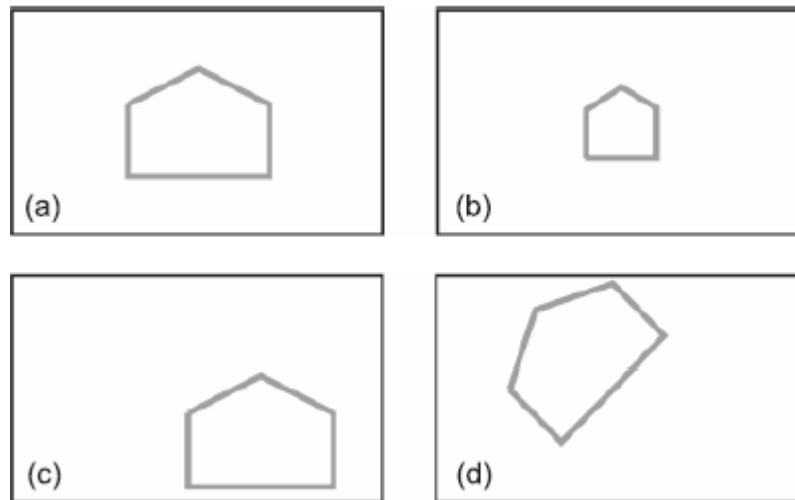


Figure 6 : (a) Objet 2D, (b) changement de taille, (c) changement de position, (d) changement d'orientation.

Les transformations la translation, la rotation et la mise à l'échelle (TRS : translation, rotation and scaling), sont des transformations à quatre paramètres, qui peut être représentée comme (Flusser et *al.*, 2016) :

$$x' = sR \cdot x + t$$

Où s est un coefficient d'échelle positif, t est un vecteur de translation et R est une matrice de rotation de la forme pratique :

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Où θ est l'angle de rotation

Des moments invariants concernant la translation, la rotation et la mise à l'échelle sont essentiellement dans presque toutes les applications pratiques. Plus, le modèle TRS est une approximation appropriée de la déformation de l'image existante si la vue est plate et perpendiculaire à l'axe visuel. Par conséquent, une plus grande attention a été accordée aux invariants TRS. (Alenezi1 et *al.*, 2018)

2.3 Les moments de Tchebichef :

Les polynômes de Tchebichef font partie des bases les plus couramment utilisées pour l'approximation de fonctions à valeurs réelles. Les fonctions de base de ces moments sont orthogonales sur l'espace des coordonnées de l'image. (Chergui, 2013)

Les moments de Tchebichef sont inclus dans la famille des moments orthogonaux discrets, Les polynômes de Tchebichef d'ordre n sont définis sur l'espace discret des coordonnées

$x \in [0, 1, 2, \dots, N-1]$ comme suit :

$$t_n(x) = (1-N)_n {}_3F_2(-n, -x, 1+n; 1, 1-N, 1) \quad (1)$$

Où $n, x = 0, 1, 2, \dots, N-1$.

La fonction ${}_3F_2$ est une fonction hypergéométrique définie par:

$${}_3F_2(a_1, a_2, a_3; b_1, b_2; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k (a_2)_k (a_3)_k z^k}{(b_1)_k (b_2)_k k!} \quad (2)$$

Où $(a_i)_k$ et $(b_i)_k$ sont des symboles de Pochhammer donnés par :

$$(a)_k = a(a+1)(a+2) \dots (a+k-1) \quad (3)$$

$$(b)_k = b(b+1)(b+2) \dots (b+k-1) \quad (4)$$

Les moments de Tchebichef satisfont la relation d'orthogonalité suivante :

$$\sum_{i=0}^{N-1} t_m(i) \cdot t_n(i) = \rho(n, N) \delta_{mn} \quad (5)$$

Pour $0 \leq m, n \leq N-1$, la fonction $\rho(n, N)$ est la norme carrée donnée par :

$$\rho(n, N) = (2n)! \binom{N+n}{2n+1} \quad (6)$$

La valeur du polynôme $t_n(x)$ peut être facilement calculée en utilisant la relation récurrente suivante:

$$(n+1)t_{n+1}(x) - (2n+1)(2x-N+1)t_n(x) + n(N^2-n^2)t_{n-1}(x) = 0 \quad (7)$$

Avec comme conditions initiales :

$$t_0(x) = 1$$

$$t_1(x) = 2x - N + 1 \quad (8)$$

Les moments de Tchebichef à une dimension sont donnés par la formule :

$$T_p = \frac{1}{\tilde{\rho}(r,N)} \sum_{x=0}^{N-1} \tilde{t}_p(x) f(x) \quad (9)$$

Où :

$$\tilde{t}_n(x) = \frac{t_n(x)}{\beta(n,N)} = \frac{1}{\beta(n,N)} n! \sum_{k=0}^n (-1)^{n-k} \binom{N-1-k}{n-k} \binom{n+k}{n} \binom{x}{k} \quad (10)$$

Les moments de Tchebichef à deux dimensions peuvent être exprimés par la formule

$$T_{p,q} = \frac{1}{\tilde{\rho}(p,N)\tilde{\rho}(q,N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \tilde{t}_p(x) \tilde{t}_q(y) f(x,y) \quad (11)$$

2.4 Les moments de Zernike :

Les moments Zernike (ZM) sont des moments orthogonaux et peuvent être utilisés pour représenter le contenu de la forme d'une image avec un minimum de redondance d'information. Les moments orthogonaux permettent une reconstruction précise de l'image et optimisent l'utilisation des informations de forme (Srinivasa et al., 2010).

Les moments orthogonaux permettent une reconstruction précise de l'image et optimisent l'utilisation des informations de forme Les moments de Zernike (ZM) sont largement utilisés dans le CBIR (Content Based Image Retrieval) comme descripteurs de forme. ZM possède de nombreuses propriétés souhaitables, à savoir l'invariance de rotation et la robustesse au bruit. Les ZM complexes sont dérivés en projetant la fonction d'image sur un polynôme orthogonal à l'intérieur d'un cercle unitaire $x^2 + y^2 = 1$ comme suit :

$$V_{nm}(x,y) = V_{nm}(p,\theta) = R_{nm}(\rho) \exp(jm\theta) \quad (12)$$

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{(n-|m|)}{2}} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s} \quad (13)$$

Où, n est un entier non négatif, m est un entier tel que $n-|m|$ est pair et $|m| < n$

$$\rho = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1}\left(\frac{y}{x}\right).$$

Projection de la fonction image sur l'ensemble de base, résultats Moments de Zernike d'ordre n avec répétition m donnés par :

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V_{nm}(\rho, \theta) \quad (14)$$

Où, $x^2 + y^2 \leq 1$. définit le complexe conjugué L'ordre des moments possède de grande influence sur la conservation de l'information angulaire. Plus l'ordre est élevé et plus les variations angulaires décrites sont fines. La Figure 7 en donne une illustration. (Kermi, 1999).

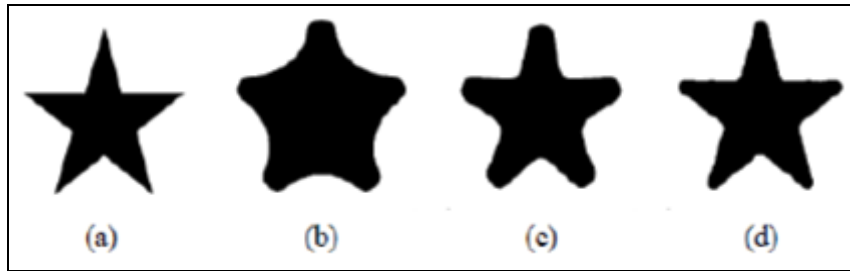


Figure 7 : Exemple de reconstructions à partir des descripteurs de Zernike, (a) image d'origine, (b) reconstruction d'ordre 10, (c) reconstruction d'ordre 20, (d) reconstruction d'ordre 40.

Les moments Zernike sont bien connus pour être invariants à la rotation. Une normalisation d'image est nécessaire pour rendre ces moments invariants à la translation et au facteur d'échelle.

2.5 Les moments de Hu :

L'idée d'utiliser des moments en reconnaissance de formes gagnait en importance lorsque Hu (1962) dérivait d'un ensemble de variables utilisant des invariants algébriques (Muharrem, Gulez, 2005) Moments bidimensionnels d'un image $M \times M$ échantillonnée numériquement ayant une fonction grise $f(x,y)$, $(x,y = 0, \dots, M-1)$ est donné comme:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} (x)^p \cdot (y)^q f(x,y) \quad p, q = 0, 1, 2, 3, \dots \quad (15)$$

Les moments $f(x,y)$ traduits par une quantité (a,b) sont définis comme suit :

$$\mu_{pq} = \sum_x \sum_y (x+a)^p \cdot (y+b)^q f(x,y). \quad (16)$$

Ainsi, les moments centraux m_{pq} ou μ_{pq} peuvent être à partir de (16) en substituant

$a = -\bar{x}$ et $b = -\bar{y}$ comme :

$$\bar{x} = \frac{m_{10}}{m_{00}} \text{ et } \bar{y} = \frac{m_{01}}{m_{00}}, \mu_{pq} = \sum_x \sum_y (x-\bar{x})^p \cdot (y-\bar{y})^q f(x,y). \quad (17)$$

Avec $\bar{x} = \frac{m_{10}}{m_{00}}$ et $\bar{y} = \frac{m_{01}}{m_{00}}$ sont les coordonnées d'un centre de masse d'un image.

Les moments centraux peuvent être exprimés en termes de moments géométriques comme :

$$\mu_{nm} = \sum_{k=0}^n \sum_{j=0}^m \binom{n}{k} \binom{m}{j} (-1)^{k+j} \bar{x}^k \bar{y}^j M_{n-k, m-j} \quad (18)$$

Cette relation a une importance pour la considération théorique, elle est utilisée pour calculer les moments centraux (Alenezi1 et al 2018) Quand une mise à l'échelle la normalisation est appliquées les moments centraux changent comme :

$$\eta_{pq} = \eta_{pq} / \eta_{00}^\gamma, \quad \gamma = [(p+q)/2] + 1 \quad (19)$$

En particulier, Hu (1962) (HU, 1961) définit sept valeurs, calculées en normalisant les moments centraux jusqu'à l'ordre trois, qui sont invariables à l'échelle, à la position et à l'orientation de l'objet. En terme des moments centraux (Muharrem, Gulez, 2005), les sept moments sont donnés comme :

$$M_1 = (\eta_{20} + \eta_{02}), \quad (20)$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \quad (21)$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \quad (22)$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \quad (23)$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \quad (24)$$

$$+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2],$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (25)$$

$$\begin{aligned}
& +4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \\
M_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
& -(\eta_{30} + 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].
\end{aligned} \tag{26}$$

Les six premiers moments sont invariants aux translations, aux rotations et aux changements d'échelle par contre le septième moment n'est pas invariant aux réflexions Ce dernier change de signe lorsqu'une telle transformation est appliquée à l'image et permet donc de détecter celle-ci (Chergui, 2013).

2.6 Invariants du moment affine :

Les moments sont l'un des paramètres qui décrivent l'image ou l'objet d'intérêt. Les moments invariants sont des moments qui ne changent pas sous un groupe de transformations. La normalisation de l'image signifie apporter l'image à une position dans laquelle l'effet de transformation est éliminé (Hosny, 2008). La transformation affine est représentée par la forme matricielle suivante :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{27}$$

Pour atteindre la normalisation, la transformation affine est décomposée en un groupe de transformation simples à un paramètre (Suk, 2005). Ce groupe comprend la translation, la mise à l'échelle uniforme, la première rotation, l'étirement et la deuxième rotation

$$x' = x - x_0, \quad y' = y - y_0, \tag{28}$$

$$x' = \alpha x, \quad y' = \alpha y, \tag{29}$$

$$x' = x \cos \theta - y \sin \theta, \quad y' = x \sin \theta + y \cos \theta, \tag{30}$$

$$x' = \delta x, \quad y' = \frac{1}{\delta} y, \tag{31}$$

$$x' = x \cos \emptyset - y \sin \emptyset, \quad y' = x \sin \emptyset + y \cos \emptyset, \tag{32}$$

Où (x_0, y_0) est le centroïde ; $\alpha, \delta > 0$; θ, \emptyset sont les angles de rotation.

2.7 Les moments de Legendre :

Les moments d'ordre 2 dimension Legendre (p+q) pour la fonction d'intensité d'image $f(x,y)$ sont définis comme :

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} * \int_{-1}^1 \int_{-1}^1 P_p(x)P_q(y) dx dy \quad (33)$$

Le polynôme de Legendre d'ordre $P_p(x)$ est défini comme :

$$P_p(x) = \sum_{\substack{k=0 \\ p-k = \text{even}}}^p B_{k,p} x^k, \quad (34)$$

Où $x \in [-1, 1]$, et la matrice de coefficients $B_{k,p}$ définie comme :

$$B_{k,p} = (-1)^{\binom{p-k}{2}} \frac{1}{2^p} \frac{(p+k)!}{\binom{p-k}{2}! \binom{p+k}{2}! k!}. \quad (35)$$

Le polynôme de Legendre $P_p(x)$ obéit à la relation récursive suivante :

$$P_{p+1}(x) = \frac{(2p+1)}{(p+1)} x P_p(x) - \frac{p}{(p+1)} P_{p-1}(x), \quad (30)$$

Avec $P_0(x) = 1$, $P_1(x) = x$ et $p > 1$. L'ensemble des polynômes de Legendre $P_p(x)$ forme une base orthogonale complète établie sur l'intervalle $[-1, 1]$. Une image numérique de $M \times N$ pixels avec fonction d'intensité $f(x_i, y_j)$ est mappé dans le carré $[-1, 1] \times [-1, 1]$ où, $1 \leq i \leq M$ et $1 \leq j \leq N$. Pour cette version de l'espace discret de l'image (Hosny K. M., 2010), Eq (33) est généralement approximée en utilisant l'approximation d'ordre zéro (ZOA) comme suit :

$$\tilde{L}_{pq} = \lambda_{pq} \sum_{i=0}^M \sum_{j=0}^N P_p(x_i) P_q(y_j) f(x_i, y_j), \quad (31)$$

Où :

$$\lambda_{pq} = \frac{(2p+1)(2p+1)}{MN}. \quad (32)$$

2.8 Les moments de Kravtchouk :

Les moments de Kravtchouk sont un ensemble formé à partir des polynômes de Kravtchouk, ceux-ci étaient introduits par Mikhail Kravtchouk (Krawtchouk, 1929), qui sont basés sur les distributions polynomiales (Batioua et *al.*, 2014). On peut contrôler la forme des polynômes de Kravtchouk avec la variation de la variable de probabilité p .

Dans (Krawtchouk, 1929), la définition des polynômes classiques de Kravtchouk de degré n est donnée par :

$$K_n(x; p, N) = \sum_{k=0}^N a_{k,n,p} x^k = {}_2F_1 \left(-n, -x; -N; \frac{1}{p} \right) \quad (33)$$

Avec $x, n = 0, 1, 2, \dots, N-1$, $N > 0$ et $p \in [0, 1]$. ${}_2F_1$ Représente la fonction hypergéométrique donnée par :

$${}_2F_1(a, b; c; z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k z^k}{(k!) (c)_k} \quad (40)$$

$(a)_k$ représente le symbole de Pochhammer donné par la formule :

$$(a)_k = a(a+1)(a+2) \dots (a+k-1) \quad k \geq 1, \quad a_0 = 1$$

Etant donné que la fonction de poids est :

$$\omega(x, p, N-1) = \binom{N-1}{x} p^x (1-p)^{N-1-x}. \quad (41)$$

La condition d'orthogonalité est donnée par :

$$\sum_{x=0}^{N-1} \omega(x, p, N-1) k_n(x, p, N-1) k_m(x, p, N-1) = \rho(x; p, N-1) \delta_{nm}. \quad (42)$$

Avec $n, m = 1, 2, \dots, N-1$, δ_{nm} est la fonction de Kronecher et le facteur de normalisation $\rho(x; p, N-1)$ présenté par :

$$\rho(x; p, N-1) = (-1)^x \left(\frac{1-p}{p} \right)^x \frac{n!}{(-N+1)_x}. \quad (43)$$

Afin d'éviter les erreurs de calcul prévues par la fonction hypergéométrique, ils ont présenté une relation récurrente pour les polynômes de Kravtchouk classiques, la formule est donnée par :

$$K_n(x, p, N-1) = \frac{(N-1)_p - 2(n-1)_p + n-1}{p(N-n)} K_{n-1}(x, p, N-1) \quad (34)$$

$$- \frac{1-pN-1}{p} \frac{K_{n-2}(x, p, N-1)}{N-n}$$

$$K_0(x, p, N-1) = 1 \quad (45)$$

$$K_1(x, p, N-1) = 1 - \frac{x}{(N-1)_p} \quad (46)$$

Après la définition des polynômes de Kravtchouk, nous introduisons la formule de calcul de moments de Kravtchouk, pour l'ordre (n+m) qui est exprimée en fonction de polynômes de Kravtchouk et des valeurs d'intensité de pixel exprimé par $f(x, y)$, la formule est donnée comme suit :

$$A_{mn} = \frac{1}{\rho(n, p_1, N-1) \rho(m, p_2, N-1)} \quad (35)$$

$$\times \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} K_n(x, p_1, N-1) K_m(y, p_2, N-1) f(x, y)$$

Les paramètres N et M représente la taille de l'image suivant les deux axes X et Y

Afin de rectifier le problème des erreurs de calcul au niveau des moments basés sur les polynômes classiques données par (44), Dans (Pew-Thian et al. 2003) ils ont présenté les polynômes de Kravtchouk normalisés. Ces derniers sont obtenus à partir de facteur de normalisation donnée par l'équation (43), les polynômes normalisés sont donnés par :

$$\bar{K}_n(x, p, N-1) = \frac{K_n(x, p, N-1)}{\sqrt{\rho(n, p, N-1)}} \quad (36)$$

Pour encor diminuer les erreurs de calcul, les auteurs ils ont présenté une nouvelle forme des polynômes de Kravtchouk pondérés-normalisés donnée par :

$$\bar{K}_n(x, p, N-1) = K_n(x, p, N-1) \sqrt{\frac{\omega(x, p, N-1)}{\rho(n, p, N-1)}} \quad (37)$$

Avec cette modification la condition d'orthogonalité sera modifiée comme suit :

$$\sum_{x=0}^{N-1} \bar{K}_n(x, p, N-1) \bar{K}_m(x, p, N-1) = \delta_{nm}. \quad (50)$$

2.9 Les moments de Hahn :

Pour tout entier $x \in [0, N-1]$ (N est un entier positif donné), polynôme de Hahn d'ordre n . $n=0,1,\dots,N-1$, est défini comme :

$$h_n^{(\mu, v)}(x, N) = (N+v-1)_n (N-1)_n \times \sum_{k=0}^n (-1)^k \frac{(-n)_k (2N+\mu+v-n-1)_k}{(N+v-1)_k (N-1)_k k!} \quad (51)$$

Où $(a)_k = a \cdot (a+1) \dots (a+k-1) = \frac{\Gamma(a+k)}{\Gamma(a)}$ est le symbole de Pochhammer et

$\mu, v (\mu > -1, v > -1)$ sont des paramètres ajustables contrôlant la forme des polynômes.

Les polynômes de Hahn discrets satisfont aux orthogonaux suivants état :

$$\sum_{x=0}^{N-1} \rho(x) h_m^{(\mu, v)}(x, N) h_n^{(\mu, v)}(x, N) = d_n^2 \delta_{nm}, \quad 0 \leq m, n \leq N-1, \quad (52)$$

Où δ_{mn} désigne la fonction de Dirac, $\rho(x)$ est une fonction dite de pondération qui est donné par :

$$\rho(x) = \frac{1}{\Gamma(x+1)\Gamma(x+\mu+1)\Gamma(N+v-x)\Gamma(N-n-x)} \quad (53)$$

et la norme carrée d_n^2 a l'expression suivant :

$$d_n^2 = \frac{\Gamma(2N+\mu+v-n)}{(2N+\mu+v-2n-1)\Gamma(N+\mu+v-n)} \times \frac{1}{\Gamma(N+\mu-n)\Gamma(N+v-n)\Gamma(n+1)\Gamma(N-n)}. \quad (54)$$

Pour éviter les fluctuations numériques du calcul des moments, nous modifions généralement polynômes de Hahn en utilisant la norme carrée et la fonction de pondération, c'est-à-dire :

$$\tilde{h}_n^{(\mu,\nu)}(x,N) = h_n^{(\mu,\nu)}(x,N) \sqrt{\frac{\rho(x)}{d_n^2}}, \quad n = 0,1,\dots,N-1. \quad (55)$$

Par conséquent, la propriété orthogonale des polynômes de Hahn normalisés peut être décrite comme :

$$\sum_{x=0}^{N-1} \tilde{h}_m^{(\mu,\nu)}(x,N) \tilde{h}_n^{(\mu,\nu)}(x,N) = \delta_{mn}, \quad 0 \leq m,n \leq N-1. \quad (38)$$

Etant donné une image numérisée $f(x,y)$ de taille $N \times M$, l'ordre $(m+n)$ de Hahn Moment de l'image est :

$$H_{mn} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) \tilde{h}_m^{(\mu,\nu)}(x,N) \tilde{h}_n^{(\mu,\nu)}(y,N), \quad m,n = 0,1,\dots,N-1. \quad (39)$$

En utilisant (57), l'équation (58) conduit à la transformation de moment inverse suivante :

$$f(x,y) = \sum_{m=0}^{N-1} \sum_{n=0}^{M-1} H_{mn} \tilde{h}_m^{(\mu,\nu)}(x,N) \tilde{h}_n^{(\mu,\nu)}(y,N) \quad (40)$$

Il indique que l'image peut être complètement reconstruite en calculant sa moments Orthogonaux discrets jusqu'à l'ordre $2N-2$. Cette propriété rend le discret moments orthogonaux supérieurs aux orthogonaux continus conventionnels des moments (Jian Zhou1, et al). Si les moments sont limités à un ordre M, on peut approximer f par \hat{f}

$$\hat{f}(x,y) = \sum_{m=0}^M \sum_{n=0}^M H_{m-n,n} \tilde{h}_{m-n}^{(\mu,\nu)}(x,N) \tilde{h}_n^{(\mu,\nu)}(y,N) \quad (41)$$

2.10 Conclusion :

Dans (Zhang et al. 2007) une étude comparative entre les moments de Tchebichef, de Legendre et de Zernike a conclu sur la supériorité de la représentation des formes par les moments de Tchebichef.

Le groupe des moments géométriques invariants de Hu a été souvent le premier choix pour de nombreux auteurs dans le domaine de reconnaissance de formes et la vision par ordinateur. Leur invariance à la translation, rotation et changement d'échelle leur donne la possibilité d'être utilisés pour des nouveaux domaines de traitement d'image et même les combiner avec d'autres moments pour avoir un meilleur résultat en profitant des points forts de chaque moment (Boumaza et al., 2015).



CHAPITRE III
IMPLEMENTATION ET
RESULTATS

3 Implémentation et résultats

3.1 Introduction :

Pour valider l'étude théorique de notre système il est nécessaire de passer maintenant à la partie pratique, et donc à l'implémentation d'un système de reconnaissance.

Dans ce chapitre nous allons présenter notre travail, en commençant par une présentation de la base de données utilisée, et les outils de développement appliqué, ainsi que l'apprentissage et le test des classificateurs individuels. Enfin nous allons présenter les résultats obtenus.

3.2 Ressources matérielles et logicielles :

Notre application a été développée sur un PC dont les caractéristiques sont les suivantes :

- *Lenovo*.
- *RAM* : 6.00 Go.
- *Processeur* : Intel® Core™ i7-2670QM CPU 2.20GHz 2.20GHz.
- *Système* : Windows 8.1 Professionnal.

On a utilisé MATLAB version 2020a .

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

3.3 Description de notre système de reconnaissance :

Nous présentons dans ce schéma une description détaillée de notre système de reconnaissance de chiffres Arabes, avec l'évaluation de performance et les différents résultats d'expérimentations.

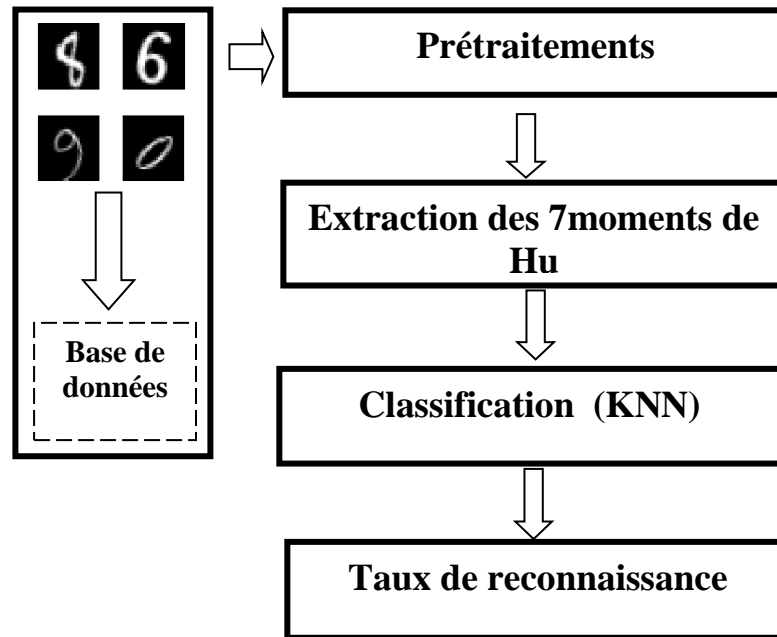


Figure 8: Les étapes majeures pour la réalisation d'une RF arabe.

3.3.1 Description de la base :

Cette base des données est entièrement synthétique, pour les apprentissages et les tests. Les images synthétiques ont été générées en appliquant des transformations affines aléatoires à des images numériques créées en utilisant différentes polices. On a 10000 images, chaque image contient des différents types de chiffres qui varient de 0 jusqu'à 9. Chaque image numérique est de 28 par 28 pixels.

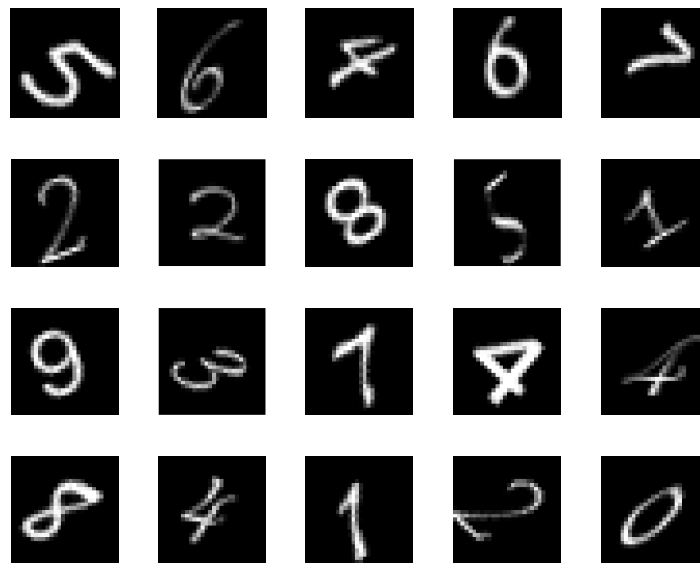


Figure 9 : Un échantillon d'images de la base de données d'apprentissage .

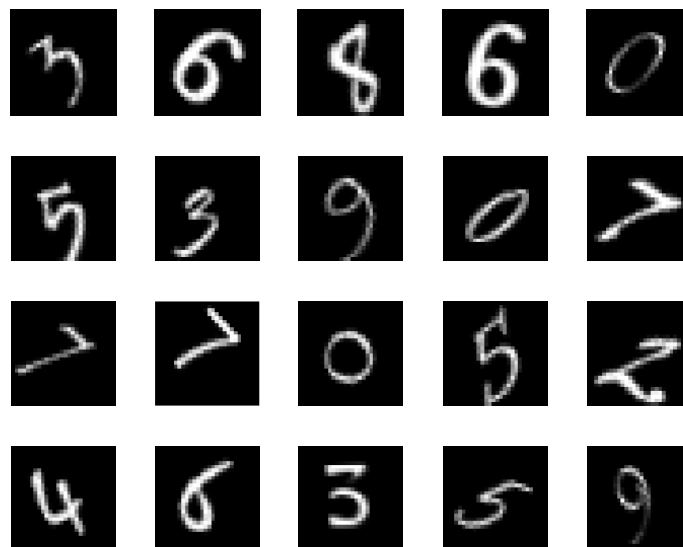


Figure 10: Un échantillon d'image de la base de données de Test .

3.3.2 Prétraitements :

il faut prétraiter les images pour qu'ils soient normalisées et facile à classifier.

- **Binarisation**

L'opération de binarisation (seuillage) consiste à mettre à zéro tous les pixels ayant un niveau de gris inférieur à une certaine valeur (appelée seuil, en anglais treshold) et à la valeur maximale les pixels ayant une valeur supérieure. Ainsi le résultat du seuillage est une image binaire contenant des pixels noirs et blancs, c'est la raison pour laquelle le terme de binarisation est employé.

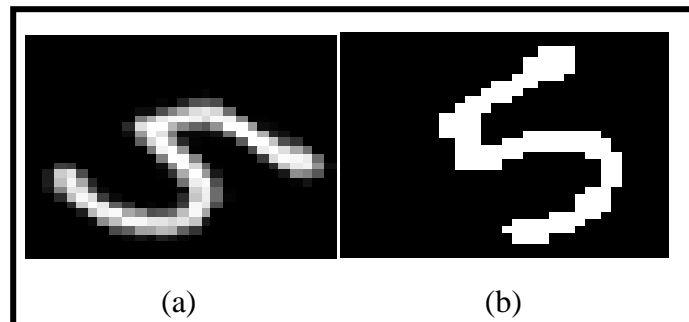


Figure 11 : (a) chiffre en niveau de gris , (b) chiffre en binaire.

3.3.3 Extraction des caractéristiques :

Après les prétraitements, les caractéristiques peuvent être extraites. on a choisi les moments de Hu, vu leur invariance aux translation, rotation et changement d'échelle. Les sept moments de Hu sont donnés par les équations suivantes:

$$M_1 = (\eta_{20} + \eta_{02}), \quad (1)$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \quad (2)$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \quad (3)$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \quad (4)$$

$$M_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \quad (5)$$

$$+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2],$$

$$M_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (6)$$

$$+ 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}),$$

$$M_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \quad (7)$$

$$- (\eta_{30} + 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].$$

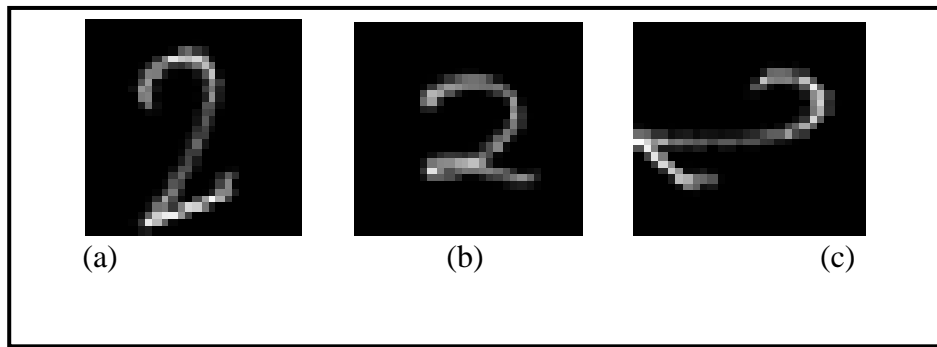


Figure 12 : Différentes invariances des moments de Hu d'une chiffer de la base, (a) chiffer originale, (b) chiffer à échelle, (c) chiffer avec rotation .

3.3.4 Classification :

Le système de reconnaissances des chiffres arabes est basée sur la comparaison de du vecteur de moments de Hu de chaque chiffre de la base de test avec tous les chiffres de la base d'apprentissage. Pour ce faire nous avons utilisé plusieurs distances (Euclidienne, Mahalanobis, Tchebychev, Cityblock, Cosine, Correlation, Spearman, Hamming, Jaccard, Minkowski) et le classifieur : le K-PPV .

K plus proches voisins est une méthode de raisonnement à partir de cas. Elle part de l'idée de prendre des décisions en recherchant un ou des cas similaires déjà résolus en mémoire. Contrairement aux autres méthodes de classification (Arbres de Décision, Réseaux de Neurones, Algorithmes Génétiques). (Abdelaali, 2016).

C'est l'une des méthodes les plus simples d'apprentissage automatique supervisé.

3.3.4.1 Principe de classifieur KPPV :

Étant donnée une base d'apprentissage d'images, pour prédire la classe d'un nouveau cas, le classifieur KPPV cherche les K plus proches voisins de ce nouveau cas et prédit la réponse la plus fréquente de ces K plus proches voisins. La méthode utilise donc deux paramètres : le nombre K et la fonction de similarité pour comparer le nouveau cas aux cas déjà classés. Le principe est donné par :

1. Choix d'un entier k (souvent $k = \text{nombre d'attributs} + 1$, ou $k = \sqrt{n}$).
2. Calcul des distances (ex : distance euclidienne, c'est la distance la plus populaire).
3. Retenir les k observations pour lesquelles ces distances sont les plus petites
4. Compter les nombres de fois où ces k observations apparaissent dans chacune des classes (déterminer les classes correspondantes).
5. Choisir la classe la plus représentée.

Ces étapes sont résumées comme suit :

Pour une forme inconnue x à classer, nous allons examiner la distance de x à tous les échantillons (qui définissent toutes les classes), puis nous sélectionnons les K plus proches échantillons et nous affectons x à la classe majoritaire parmi ces K échantillons. (Abdelaali, 2016).

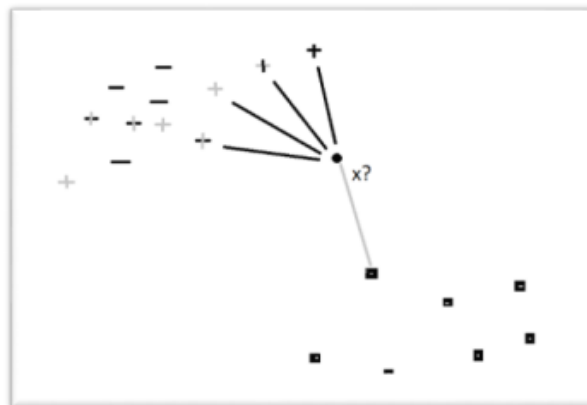


Figure 13 : Méthode de K-plus proches voisins.

3.3.4.2 *Choix du nombre K :*

Le choix du paramètre K dans la règle des KPPV a une influence directe sur le style du classifieur ainsi déterminé. Une faible valeur de K va donner un classifieur de bonne résolution (définir des frontières compliquées entre classes) mais très sensible au bruit sur les échantillons et sur le vecteur à classer. Une valeur grande de K aura un comportement inverse, lissant les frontières mais peu sensible au bruit. (Milgram, 1993). Dans ce travail nous avons utilisé différentes valeurs de k : $k=1$ jusqu'à 5, $k = \text{nombre d'attributs} + 1$, et $k = \sqrt{n}$

3.3.4.3 *Sélection de la classe :*

La méthode la plus simple est de rechercher le cas le plus proche et de prendre la même décision. Si cette méthode peut fournir de bons résultats sur des problèmes simples pour lesquels les points sont bien répartis en groupes denses de même classe, en règle générale, il faut considérer un nombre de voisins plus important pour obtenir de bons résultats.

Une première façon de combiner les k classes des k voisins les plus proches est le vote majoritaire. Elle consiste simplement à prendre la classe majoritaire.

Une seconde façon est le vote majoritaire pondéré. Chaque vote, c'est-à-dire chaque classe d'un des k voisins sélectionnés, est pondéré. Le poids est inversement proportionnel à la distance entre le cas à classer et les autres k plus proches voisins classes (Une telle pondération s'appelle un noyau).

Dans les deux cas précédents, il est possible de définir une confiance dans la classe attribuée égale au rapport entre les votes gagnants et le total des votes. (Abdelaali, 2016)

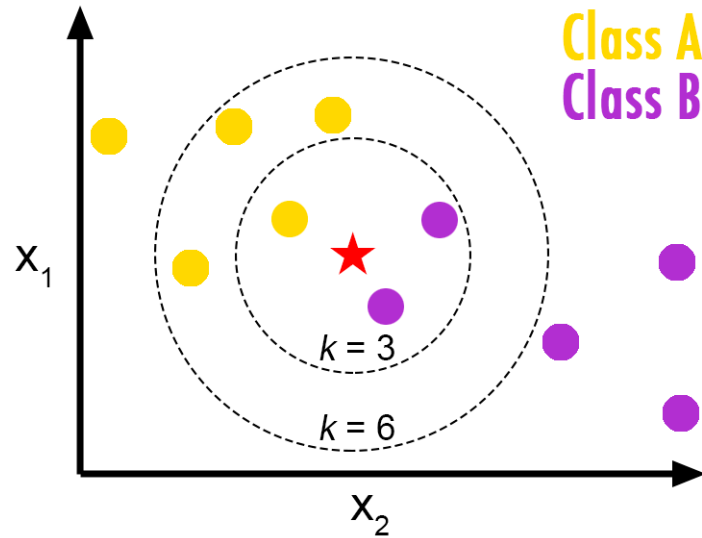


Figure 14 : KPPV avec $k=3$ et $k=6$.

3.3.4.4 Définition de la distance :

Le choix de la distance est primordial au bon fonctionnement de la méthode. Une distance doit avoir quatre propriétés pour tous les vecteurs a , b et c $D(a,b) \geq 0$ (non-négativité)

- $D(a,b)=0$ si et seulement si $a=b$ (réflexivité)
- $D(a,b)=D(b,a)$ (symmetries)
- $D(a,b)+D(b,c) \geq D(a,c)$. (Triangle inégalité)

On peut cependant noter qu'un point « a » peut avoir un plus proche voisin « b »

Tandis que « b » possède de nombreux voisins plus proches que « a ». Le choix de la distance se fait en fonction des connaissances préalables du problème. Il est possible de choisir la distance en faisant varier cette distance et, pour chacun des choix, estimer l'erreur réelle (Abdelaali, 2016). On choisit alors la distance donnant la meilleure erreur réelle estimée.

Nous avons utilisé les distance présentées dans le tableau :

Tableau 1: les distance et leur formule.

distance	formule	distance	formule
Euclidienne	$D(A,B) = \sum_{i=1}^n \sqrt{(A_i - B_i)^2}$	Minkowski	$d(x_i, x_j) = \left(\sum_{s=1}^p x_{is} - x_{js} ^q \right)^{\frac{1}{q}}$
cityblock	$D(A,B) = \sum_{i=1}^n A_i - B_i $	Jaccard	$d(x,y) = \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i)^2 + (y_i)^2 - [(x_i)(y_i)]}$
Tchebycheff	$D = \max(A_i - B_i), i = 1..n$	Correlation	$d(x,y) = \frac{\frac{1}{n} \sum_i x_i y_i - \mu_X \mu_Y}{\sigma_X \sigma_Y}$
Mahanalobis	$d(x_i, y_i) = \sqrt{(x_i - y_i)^T S^{-1} (x_i - y_i)}$	Cosine	$d(x,y) = \frac{x \cdot y}{\ x\ \ y\ } = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$
Hamming	$D_H = \sum_{i=1}^k x_i - y_i $ $x = y \Rightarrow D = 0$ $x \neq y \Rightarrow D = 1$	Spearman	$D(x,y) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2}$

3.3.4.5 L'apprentissage :

L'apprentissage a pour but de généraliser l'information portée par les exemplaires de l'ensemble d'apprentissage pour faire émerger les représentations des classes. Ces représentations doivent permettre de s'affranchir de la variabilité des exemplaires d'une classe lors de la prise de décision. La validation de l'apprentissage est faite par les taux de reconnaissance obtenus sur un ensemble test constitué d'exemplaires différents de ceux de l'ensemble d'apprentissage.

Dans notre travail, nous mettons 3 divisions de la base :

- 5000 exemples d'apprentissage et 5000 pour le test.
- 7000 exemples d'apprentissage et 3000 pour le test.
- 6000 exemples d'apprentissage et 4000 pour le test.

3.4 Résultats et discussion :

3.4.1.1 Images grises

Tableau 2: Le taux de classification pour la base de données (50% apprentissage / 50% Test).

La base de données	K distance		1	2	3	4	5
(50% apprentissage / 50% test)	Euclidean	appr	100	92.78	91.88	88.26	86.70
		test	85.08	80.34	80.20	78.50	76.94
	Mahalanobis	appr	100	97.06	96.50	94.84	93.98
		test	93.56	91.12	91.30	90.24	89.26
	Cosine	appr	100	94.32	93.22	90.92	89.30
		test	88.20	84.68	83.52	81.70	81.28
	Correlation	appr	100	94.84	93.90	91.32	89.92
		test	88.64	85.44	85.10	82.66	82.66
	Spearman	appr	43.42	44.44	45.30	44.64	44.80
		test	38.16	39.44	40.20	40.92	40.74
	Hamming	appr	100	40.00	10.00	10.00	19.96
		test	10	10	10	10	10
	Jaccard	appr	100	40.00	10.00	10.00	19.96
		test	10	10	10	10	10
	Minkowski	appr	100	92.78	91.88	88.26	86.70
		test	85.08	80.34	80.20	78.50	76.94
	Cityblock	appr	100	94.12	93.26	90.18	88.76
		test	87.74	82.96	83.70	81	80.06
	Chebychev	appr	100	91.64	90.06	86.94	84.74
		test	82.24	78.06	77.98	75.60	74.94

Le tableau 2 montre le taux de reconnaissance de la base d'apprentissage. Nous avons obtenu un score parfait (100%) dans le cas de $K = 1$ et toutes les distances sauf pour Spearman (43.42%). En revanche, le taux de reconnaissance de la base de test représente (93.56%) pour la distance de Mahalanobis et $K = 1$. Ce qui indique l'efficacité de l'apprentissage.

La figure15 suivant représente les performance du système en utilisée KNN, Nous avons Gardé $k = 1$ et en changeant les distances.

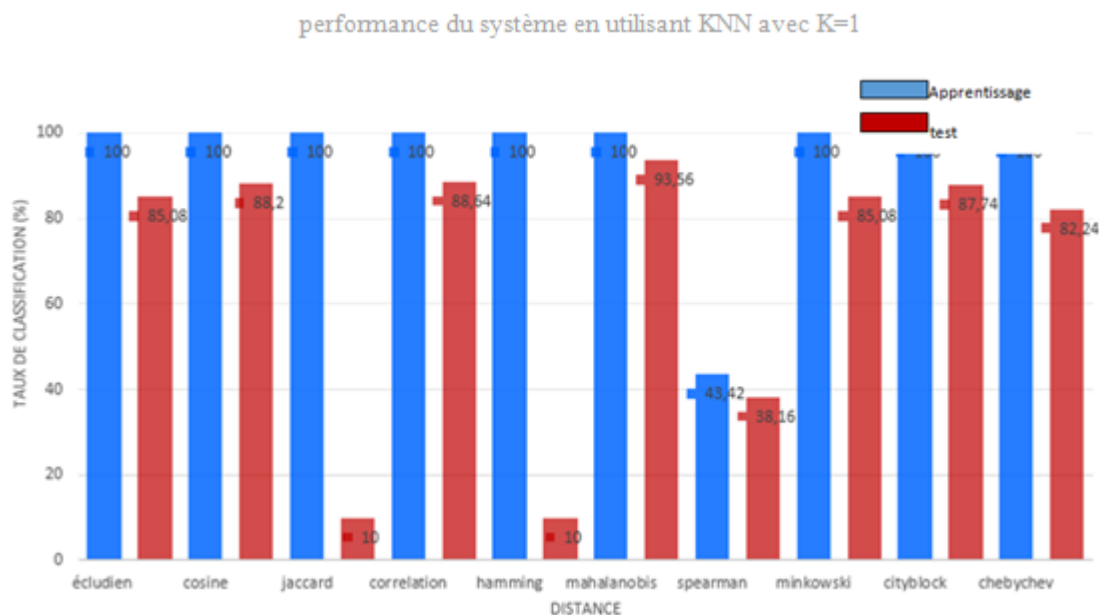


Figure 15: les performances du système en utilisant KNN avec K=1.

3.4.1.2 Images binaires

Tableau 3: Le taux de classification pour la base de données après le binarisation (50% apprentissage / 50% test).

La base de données	K Distance	1	2	3	4	5
		(50% apprentissage / 50% test)				
Euclidean	Appr	100	76.38	72.86	70.66	68.52
	Test	52.74	50.78	51.84	52.98	53.64
Mahalanobis	Appr	100	77.2	74.68	71.76	70.24
	Test	55.88	53.08	55.46	56.56	56.54
Cosine	Appr	100	74.96	71.66	68.48	66.32
	Test	50.8	48.42	50.76	50.92	51.84
Correlation	Appr	100	74.54	71.92	68.94	66.98
	Test	51.12	48.6	51.56	51.56	52.1
Spearman	Appr	42.12	39.52	40.72	40.82	41.90
	Test	30.78	32.2	33.56	34.54	35.58
Hamming	appr	100	79.56	11.32	10.48	10.48
	test	10.44	10.38	10.04	10.04	10.04
Jaccard	appr	100	79.56	11.32	10.48	10.48
	test	10.44	10.38	10.04	10.04	10.04
Minkowski	appr	100	76.38	72.86	70.66	68.52
	test	52.74	50.78	51.84	52.98	53.64
Cityblock	appr	100	76.42	73.02	70.86	68.68
	test	53.34	51.12	52.16	53.18	53.98
Chebychev	appr	100	75.32	71.56	68.76	67.10
	test	50.96	49.68	50.94	51.54	52.08

D’Après le tableau 3 nous remarquons que le meilleur résultat est (56.56%) à k = 4 avec la distance de Mahalanobis .

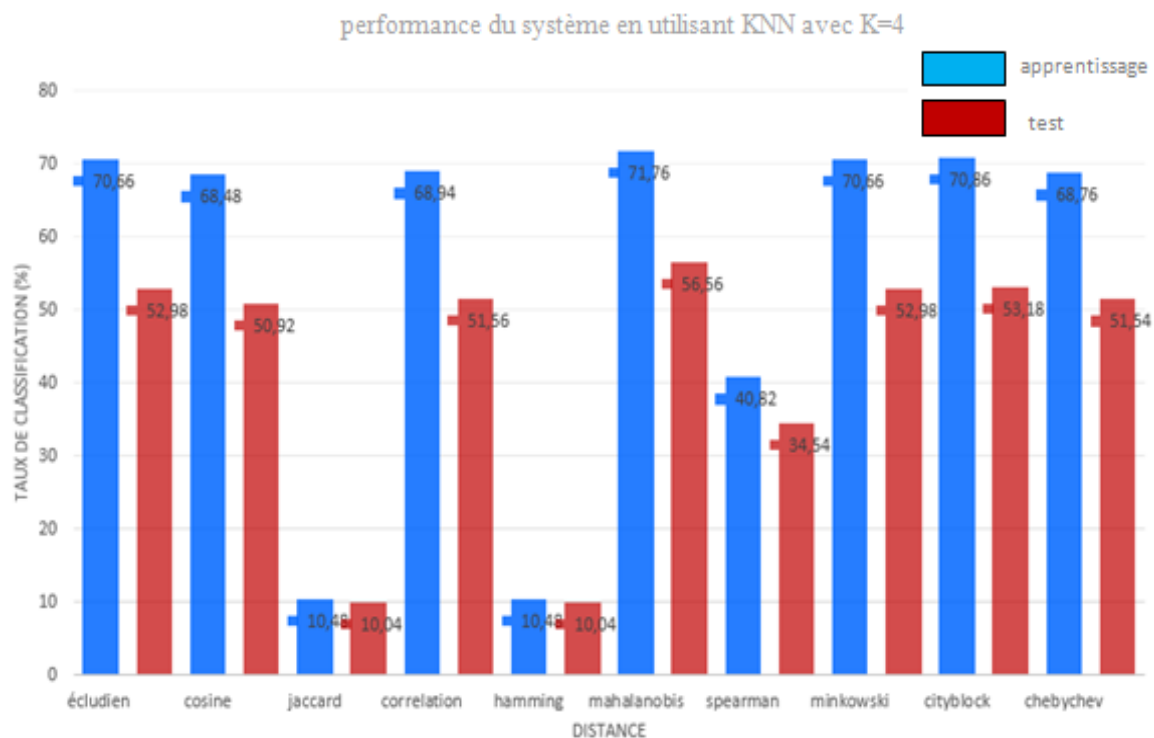


Figure 16: les performances du système en utilisant KNN avec K=4.

Tableau 4: Le taux de classification pour la base de données (70% apprentissage / 30% test).

La base de données	K distance		1	2	3	4	5
			(70% apprentissage / 30% test)	Euclidean	appr	100	85.80
test	70.33	67.56			68.43	67.80	67.53
Mahalanobis	appr	100		88.51	87.60	85.47	83.30
	test	77.16		73.70	74.66	73.40	73.40
Cosine	appr	100		85.17	82.57	80.24	78.15
	test	69.66		67.36	69.26	67.93	67.20
Correlation	appr	100		84.81	82.57	79.74	77.74
	test	68.93		65.76	68.30	67.03	66.90
Spearman	appr	35.18		34.47	38.28	37.78	38.15
	test	27.86		28.63	32.76	32.40	33.60
Hamming	appr	100		40.00	10.00	10.00	19.97
	test	10		10	10	10	10
Jaccard	appr	100		40.00	10.00	10.00	19.97
	test	10		10	10	10	10

	Minkowski	appr	100	85.80	84.22	81.71	79.30
		test	70.33	67.56	68.43	67.80	67.53
	Cityblock	appr	100	85.95	84.64	82.05	80.34
		test	71.30	71.30	68.76	68.63	68.06
	Chebychev	appr	100	85.32	83.84	80.84	78.80
		test	68.93	66.06	67.53	66.90	65.96

A partir du tableau 4, on remarque que le meilleur résultat 77.16% qui correspond à $k = 1$ et à la distance de mahalanobis, les plus petites valeurs enregistrées sont 10% dans les deux distance Hamming et Jaccard.

Tableau 5: Le taux de classification pour la base de données (60% apprentissage / 40% test).

La base de données		K distance	1	2	3	4	5
(60% apprentissage / 40% test)	Euclidean	appr	100	77.1	74.03	71.15	69.31
		test	52.75	51.05	53.02	53.06	54.07
	Mahalanobis	appr	100	77.38	75.9	73.5	71.65
		test	56.27	53.87	55.82	56.87	57.1
	Cosine	appr	100	76.1	72.48	68.83	66.86
		test	51.52	49.25	51.7	52.12	52.27
	Correlation	appr	100	75.43	72.53	69.83	67.45
		test	52.02	49.25	51.87	52.17	52.27
	Spearman	appr	40.45	39.1	40.36	40.46	42.11
		test	30.27	32.62	34.37	34.65	36
	Hamming	appr	100	40.51	49.78	20.93	11.06
		test	10.47	10.3	10.32	10.05	10.02
	Jaccard	appr	100	40.51	49.78	20.93	11.06
		test	10.47	10.3	10.32	10.05	10.02
	Minkowski	appr	100	77.1	74.03	71.15	69.31
		test	52.75	51.5	53.02	53.6	54.07
	Cityblock	appr	100	77.06	74.45	72	70.03
		test	53.5	51.47	53.02	53.87	53.87
	Chebychev	appr	100	75.76	72.93	70.43	67.86
		test	51.42	50.25	51.4	52.57	52.77

Dans le cas de la division de la base en 60% pour l'apprentissage et 40% pour test, on constate que le meilleur résultat (57,1%) est en $k = 5$ et la distance de Mahalanobis.

nous avons utilisé aussi différentes valeurs de k : $k = \text{nombre d'attributs} + 1$ (en binaire et en niveau de gris) et $k = \sqrt{10000}$, $k = \sqrt{7000}$. Les résultats sont présentés dans le tableau 6 :

Tableau 6: le taux de classification de différent valeur de k .

K	Le taux de classification de l'apprentissage (%)	Le taux de classification de test (%)
$\sqrt{10000}$	55.48	51.26
$\sqrt{7000}$	47.40	46.12
nombre d'attributs + 1=7+1(en niveau de gris)	64.16	53.04
nombre d'attributs + 1=7+1(en binaire)	64.16	53.04

d'après les résultats obtenus précédemment, on constate que :

- le taux de reconnaissance varie d'une distance à l'autre .
- Le meilleur résultat présenter dans le cas de la division de la base par 50% d'apprentissage et 50% en niveau de gris avec taux de classification de (93.56%) pour k=1.

3.5 Conclusion :

Nous avons présenté dans ce chapitre une description détaillée de notre système de reconnaissance des chiffres, avec l'évaluation des moment Hu en utilisant le classificateur KPPV avec différentes distances et différentes valeurs de k.

Les résultats obtenus ont montré que la distance Mahalanobis donnait le meilleur résultat (93.56%) .



CONCLUSION GENERALE

CONCLUSION GENERALE

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

Le travail présenté dans ce mémoire décrit les différentes étapes nécessaires à la construction d'un système de reconnaissance de chiffres arabes, à savoir : l'extraction des caractéristiques et la classification.

Dans le domaine de la reconnaissance de chiffres arabes, les caractéristiques peuvent être décrites comme un moyen permettant de distinguer un chiffre d'une classe d'un autre chiffre d'une autre classe. Dès lors, il est nécessaire de définir des caractéristiques significatives lors du développement d'un système de reconnaissance.

Nous avons essayé d'évaluer les performances des moments de Hu comme descripteurs des chiffres arabes en utilisant différents types de distances et le classifieur KPPV .

D'autres méthodes caractéristiques et classifieurs peuvent être utilisées dans des futures travaux. On insiste sur deux facteurs pour tirer profit des caractéristiques des moments de Hu , le premier est le besoin d'une étape de prétraitement très efficace vu que les moments de Hu sont sensibles aux bruits et le deuxième est la nécessité de combiner les moments de Hu avec d'autres types de descripteurs pour obtenir des meilleurs résultats.

Comme perspectives, nous suggérons aux futures promotions d'implanter des systèmes de reconnaissance des formes basés sur les machines d'apprentissage approfondi sur des cartes DSP afin de les mettre directement dans l'industrie.

BIBLIOGRAPHIE

- Abdelaali, M. (2016). Reconnaissance hors ligne des chiffres manuscrite isolé (Base de donnée M.N.I.S.T). memoire de master, Université de Tébessa. , Tébessa.
- Alenezi1. Ahmad, Siti Hasana. S, Rakhimov. I.S. (2018). On Geometric Moment Invariants. Universiti Putra Malaysia.
- Amar, Z. H. (2008). Analyse temps-échelle et reconnaissance des formes pour le diagnostic du système de guidage d'un tramway sur pneumatiques. Thèse de doctorat, Université Blaise Pascal, Clermont-Ferrand II.
- Amara, M. (2012). Reconnaissance des caractères arabes imprimés par l'approche neuro-génétique. Ecole nationale des science de l'informatique.
- Augereau.O. (2013). Reconnaissance et classification d'images de documents.phD.
- Batioua. I, Zenkouar. Kh, Zahi. A. (2014). Théorie des moments et classification : Application à la reconnaissance de forme. Master Système Intelligent et Réseaux, Université Sidi Mohammed Ben Abdellah.
- Belaid. A, Belaid. Y. (1992). reconnaissance des formes : méthodes et applications. interedition.
- Bendjeddou. F, Kherbéche. T. (1999). modélisation de l'écriture arabe en graphèmes. mémoire d'ingénieur d'état en électronique.
- Bendjeddou. T. (2008). Le choix de paramètres pour la reconnaissance des chiffres manuscrits .
- Boumaza. R, Beddiaf. H, Chergui. L. (2015). Reconnaissance d'empreintes digitales en utilisant les moments de Hu. Master, Université L'arbi Ben Mhidi-Oum El Bouaghi-, Oum El Bouagh.
- Cheng. HD, Shi. X, Rui. M,H. L, Cai. X, Du. H. (2006). Approaches for auto- mated detection and classification of masses in mammograms. Pattern recognition.

- Chergui. L. (2013). Combinaison de classifieurs pour la reconnaissance de mots arabes manuscrits. Thèse, Université Mentouri Constantine, Constantine .
- Chevalier. S. (2004). Reconnaissance d'écriture manuscrite pour des techniques markoviennes: une approche bidimensionnelle et générique. Thèse de Doctorat, Université René Descartes, France.
- Djabeur, Djeddar. M. R, Benkada. F.-a. (2017). Mise au Point d'une Application de Reconnaissance de Formes. Master en Informatique, Université Abou Bakr Belkaid–Tlemcen, Tlemcen.
- EOsman. M, Abdel Wahed. M, Mohamed. A, Kadah. Y. (2009). omputer aided diagnosis system for classification of microcalcifications in digital mammograms. In Radio Science Conference.
- Filatov. A, Nikitin. N, Volgunin. A, Zelinsky. P. (1998). The Address Script Recognition System for Handwritten Envelopes. Nagano.
- Flusser. J, Suk. T, Barbara. Z. (2016). 3D Moment Invariants to translation, rotation, and Scaling. 2D and 3D Image Analysis by Moments.
- Ganesan. K, Acharya. R, Chua. K, Min. L, Abraham. T, Ng. K-H. (2013). Computer-aided breast cancer detec- tion using mammograms : a review. IEEE Reviews in Biomedical Engineering.
- Grandidier. F. (2003). Un Nouvel Algorithme de Sélection de Caractéristiques : Application à la Lecture Manuscrite. thèse de Doctorat.
- Hu M.K. (1961). Visual Pattern Recognition by Moment Invariants. Ire Transactions On Information Theory, Vol. 49.
- Jacques-Henri. S. (2012). Classification non supervisée : de la multiplicité des données à la multiplicité des analyses.
- Jaeger. S, Doermann. D. (2006). Arabic and Chinese Handwriting Recognition. Springer, Germany.

- Kermi, S. (1999). « Classifieur neuronal base connaissances, application à la reconnaissance des caractères arabes isolés manuscrits ». Thèse de magister, université Badji Mokhtar, Annaba, Algérie.
- Khalid. M, H. (2008). On the computational aspects of affine moment invariants. agazig University, Zagazig, Egypt. Egypt .
- Khalid. M, H. (2010). Refined translation and scale Legendre moment invariants. Zagazig University. Egypt .
- Kharroubi. J. (2002). Etude de techniques de classement” Machines `a vecteurs supports” pour la vérification automatique du locuteur. thesis.
- Koutroumbas. T, Koutroumbas. K. (2003). “Pattern Recognition, Second Edition”. Academic Press, Elsevier.
- Krawtchouk. M. (1929). On interpolation by means of orthogonal polynomials. memoirs Agricultural Inst.
- Mahmoud. A.S. (1994). "Arabic Character Recognition Using Fourier Descriptors and Character Contour Encoding". Pattern Recognition, vol 27, Num 6.
- Megar. I, Adel. (2016). Contribution à la reconnaissance d'écriture arabe manuscrite en utilisant la classification neuronale.
- Mehennaoui. Z. (2006). Reconnaissance De L'écriture Arabe Manuscrite A Base Des Machines A Vecteurs De Supports.
- Mercimek. M, Gulez. K. (2005). Real object recognition using moment invariants. 34349 Besiktas-Istanbul, Turkey.
- Milgram. M. (1993). « Reconnaissance des formes méthodes numériques et connexionnistes ».
- Santhi. R, Nithya. B. (2015). Computer aided diagnosis system for mammogram analysis : a survey. Journal of Medical Imaging and Health Informatics.
- Srinivasa. Ch, Srinivas Kumar. S, Chandra Mohan. B. (2010). Content Based Image Retrieval Using Exact Legendre Moments And Supporte Vector Machine.

Suk. T, Flusser. F. (2005). Affine Normalization of Symmetric Objects.

Yap. P-T, Paramesra. R, Ong. S-H, Senior. M. (2003). Image Analysis by Krawtchouk Moments. Ieee transactions on image processing. Vol. 12.

Zhang. L, Gong-Bin. Q, Wei-Wei. X. (2007). Geometric distortions invariant blindsecond generation watermarking technique based on Tchebichef moment of original image. Journal of Software, vol. 18, no. .

Zhou. J, Shu. H, Zhu. H, Toumoulin. CH, Luo. L. Image Analysis by Discrete Orthogonal Hahn Moments. Southeast University.