

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université de Tébessa  
Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie  
Département : Mathématiques et Informatique



MEMOIRE DE MASTER

Domaine: Mathématiques et Informatique  
Filière: Informatique  
Option: multimedia et système informatique

Thème



Reconnaissance hors ligne des chiffres manuscrite isolé  
(Base de donnée M.N.I.S.T)



Présenté par :  
MENASRIA Abdelaali  
ZEMOULI Saddam

Jury de soutenance :

|                    |                           |                       |           |
|--------------------|---------------------------|-----------------------|-----------|
| BENDIB Issam       | Maitre assistant classe A | Université de Tébessa | Président |
| GHRIB Nawel        | Maitre assistant classe A | Université de Tébessa | Examineur |
| GATTAL Abdeldjalil | Maitre assistant classe A | Université de Tébessa | Encadreur |

Date de soutenance: .....  
Note : ..... Mention : .....

## ملخص

يستند نظام التعرف على الأرقام المعزولة المكتوبة بخط اليد على الخطوات التالية: تجهيزها، واستخراج الميزات والتصنيف ويحتمل أن تليها خطوة مرحلة ما بعد المعالجة. في هذا العمل، نحن نركز على خطوتين : واستخراج الميزات والتصنيف استخراج الميزات يمكن وصفها بأنها وسيلة لتمييز رقم على الآخر ولذلك فمن الضروري تحديد الميزات الهامة خلال البرمجة.

في حالتنا، ونحن نركز على الجمع بين ثلاثة طرق : HOG ,Zernike moments و Gabor filter هذه الميزات تسمح بإعطاء فعالية كبيرة لمعرفة الأرقام باستخدام KNN .

الكلمات المفتاحية: الأرقام المعزولة المكتوبة بخط اليد، التصنيف، استخراج الميزات، HOG ،Gabor Filter ،Zernike moments.

# Abstract

An isolated handwritten digits recognition system is based on the following steps preprocessing, feature extraction and classification, possibly followed by a post-processing step.

In this work, we focus on two modules: feature extraction and classification. The features can be described as a way to distinguish one character from one class from another character from another class. So it is necessary to define significant features during development.

In our case, we focus on the combination of three descriptors: Zernike moments, the HOG histogram and Gabor filter. These features are provided an important recognition rate by using the KNN classifier.

**Key words :** isolated handwritten digits, recognition system feature, extraction, classification, Zernike moments, the HOG histogram , Gabor filter, KNN classifier.

# Résumé

Un système de reconnaissance des chiffres manuscrits isolés est fondé sur les étapes suivantes : prétraitements, l'extraction des caractéristiques et la classification, suivis éventuellement d'une phase de post-traitement.

Dans ce travail, nous nous intéressons à deux modules : l'extraction des caractéristiques et la classification.

Les caractéristiques peuvent être décrites comme un moyen permettant de distinguer un caractère d'une classe d'un autre caractère d'une autre classe. Alors, il est nécessaire de définir des caractéristiques significatives lors du développement.

Dans notre cas, nous avons concentré sur la combinaison de trois descripteurs : moments Zernike, l'histogramme HOG et filtre de Gabor. Ces caractéristiques sont fournies un taux de reconnaissance important en utilisant le classifieur KPPV.

**Mots-clés :** Un système de reconnaissance des chiffres manuscrits, l'extraction des caractéristiques, la classification, moments Zernike, l'histogramme HOG, filtre de Gabor, le classifieur KPPV.

---

---

## Table des matières

---

---

|                             |   |
|-----------------------------|---|
| Introduction générale ..... | I |
|-----------------------------|---|

### Chapitre1: La reconnaissance de l'écriture manuscrite

|  |    |
|--|----|
| 1-1 Introduction : .....                                     | 1  |
| 1.2 Différents aspects de reconnaissance .....               | 1  |
| 1.2.1. Reconnaissance en-ligne et hors-ligne .....           | 2  |
| 1.2.1.1 La reconnaissance En-Ligne .....                     | 2  |
| 1.2.1.2 La reconnaissance hors-ligne .....                   | 2  |
| 1.3 Processus de système de reconnaissance .....             | 3  |
| 1.3.1 Phase d'acquisition .....                              | 4  |
| 1.3.2 Phase de prétraitement .....                           | 5  |
| 1.3.2.1 La binarisation.....                                 | 5  |
| 1.3.2.2 Lissage :.....                                       | 6  |
| 1.3.2.3 La normalisation.....                                | 6  |
| 1.3.2.3.1 Normalisation de taille .....                      | 6  |
| 1.3.2.3.2 Normalisation de l'inclinaison (Redressement)..... | 7  |
| 1.3.2.4 Squelettisation (Thinning).....                      | 7  |
| 1.3.3 Phase de Segmentation.....                             | 8  |
| 1.3.3.1 Segmentation explicite.....                          | 9  |
| 1.3.3.2 Segmentation implicite .....                         | 9  |
| 1.3.4 Phase d'extraction de caractéristiques .....           | 9  |
| 1.3.4.1 Les primitives structurelles .....                   | 10 |
| 1.3.4.2 Les Primitives statistiques :.....                   | 10 |
| 1.3.4.3 Les primitives globales .....                        | 11 |
| 1.3.4.4 Primitives topologiques ou métriques.....            | 11 |
| 1.3.5 Phase de Classification.....                           | 12 |
| 1.3.5.1 Approche statistique.....                            | 13 |

|   |    |
|---|----|
| 1.3.5.2 Approche structurale .....                  | 13 |
| 1.3.5.3 Approche stochastique .....                 | 14 |
| 1.3.5.4 Approche hybride :.....                     | 14 |
| 1.4 Méthodes de classification statistiques .....   | 14 |
| 1.4.1 Réseaux de neurones .....                     | 15 |
| 1.4.2 Les Machines à Vecteur Support (SVMs) .....   | 15 |
| 1.4.3 Le classifieur du K plus proches voisins..... | 16 |
| 1.5 Conclusion.....                                 | 16 |

## **Chapitre 2: Etat de l'art**

|  |    |
|--|----|
| 2-1 Introduction :.....  | 23 |
| 2-2 LES TRAVEAUX ULTERIERS :.....                              | 23 |
| 2-2-1 Mayraz et Hinton. (2002) :.....                          | 23 |
| 2-2-2 Kussul et Baidyk (2004):.....                            | 24 |
| 2-2-3 Wu et Zhang (2010) : .....                               | 24 |
| 2-2-4 Ebrahimzadeh et Jampour (2014) : .....                   | 24 |
| 2-2-5 Wakahara et Yamashita (2014) :.....                      | 25 |
| 2-2-6 Lee et al (2015) :.....                                  | 25 |
| 2-2-7 Wakahara et Yamashita (2016) :.....                      | 26 |
| 2-2-8 Dundar et al (2016) :.....                               | 26 |
| 2-3 Quelque résultat expérimental obtenu par les auteurs ..... | 27 |
| 2-4 Conclusion .....   | 28 |

## **Chapitre 3: Extraction des caractéristiques**

|   |    |
|---|----|
| 3-1- Introduction :.....                    | 29 |
| 3-2- Histogramme de gradient Orienté :..... | 29 |
| 3-2-1- PRINCIPE : .....                     | 29 |
| 3-2-1-1- Description générale:.....         | 29 |
| 3-2-1-2- Calcul du gradient:.....           | 30 |

|   |    |
|---|----|
| 3-2-1-3- Construction de l'histogramme: .....   | 31 |
| 3-2-1-4- Formation des blocs:.....              | 31 |
| 3-2-1-5- Normalisation des blocs:.....          | 31 |
| 3-3- Les moments de Zernike :.....              | 32 |
| 3-4- Filtre de Gabor :.....                     | 33 |
| 3-5- K plus proches voisins K-NN (KPPV) : ..... | 34 |
| 3-5-1- Principe de classifieur KPPV .....       | 34 |
| 3-5-2- Choix du nombre K:.....                  | 35 |
| 3-5-2-1- Définition de la distance: .....       | 36 |
| 3-5-2-2- Sélection de la classe :.....          | 36 |
| 3-6- Conclusion .....                           | 37 |

#### **Chapitre 4: Résultats expérimentaux**

|                                   |     |
|-----------------------------------|-----|
| 4-1- Introduction .....           | 38  |
| 4-2- Bases de données .....       | 38  |
| 4-3- Critère d'évaluation .....   | 39  |
| 4-4- Protocole expérimental ..... | 40  |
| 4-4-1- Les expériences .....      | 40  |
| 4-5- Conclusion .....             | 41  |
| Conclusion Générale .....         | III |
| Références Bibliographiques.      |     |

---

---

## Liste des figures

---

---

| <b>N</b>   | <b>Titre</b>   | <b>Page</b> |
|------------|--|-------------|
| <b>1.1</b> | Schéma général d'un système de reconnaissance des chiffres manuscrits.     | 04          |
| <b>1.2</b> | Exemple d'un chiffre manuscrit normalisé                                   | 06          |
| <b>1.3</b> | Exemple de deux types de redressement                                      | 07          |
| <b>1.4</b> | Exemple de squelettisation   | 08          |
| <b>1.5</b> | Division d'une image en zone   | 10          |
| <b>1.6</b> | Exemple d'hybridation entre deux approches (statistiques et structurelles) | 14          |
| <b>3.1</b> | Méthode de K_pl us proches voisins   | 35          |
| <b>4.1</b> | Exemples de la base de données MNIST                                       | 39          |

---

---

## Liste des tableaux

---

---

| <b>N</b>   | <b>Titre</b>                                    | <b>Page</b> |
|------------|---|-------------|
| <b>2.1</b> | Quelque résultats rapportés dans la littérature | 27          |
| <b>4.1</b> | Répartition de la base MNIST pour les chiffres  | 39          |
| <b>4.2</b> | Résultats obtenus de chacune expérience         | 40          |
| <b>4.3</b> | matrice de confusion de l'expérience 5          | 41          |



# **Introduction Générale**

---

---

# INTRODUCTION GENERALE

---

---

La Reconnaissance de Chiffres Manuscrits (RCM), fait l'objet d'un nombre important de travaux de recherche, grâce à ses applications diverses et potentielles. Le recours à la RCM s'impose dans la plupart des domaines de la vie courante. A titre d'exemple, nous citons le tri postal qui est l'une des premières applications où tous les jours des milliers d'enveloppes sont automatiquement triés. A l'instar de cette application, on distingue la lecture du montant numérique des chèques bancaires et l'identification du numéro de sécurité sociale. Toutefois, malgré le progrès impressionnant des techniques utilisées ainsi que l'explosion dans la puissance de calcul des ordinateurs, la recherche sur la RCM avance avec une performance de reconnaissance qui reste loin de celle de l'œil humain.

A présent, nous nous intéressons en particulier aux méthodes d'extraction des caractéristiques. Avec le développement de nouvelles techniques d'extraction des caractéristiques, des améliorations graduelles ont été relevées sur les résultats obtenus par les systèmes de RCM. En effet, les techniques structurelles et statistiques ont été les premières méthodes utilisées.

Les performances se sont révélées insuffisantes pour confronter la variabilité de l'écriture. Ainsi, le choix de l'algorithme de classification est un problème qui suscite un grand nombre de chercheurs. Néanmoins, lorsqu'elles sont combinées avec un bon descripteur elles fournissent des précisions assez satisfaisantes.

Dans notre cas, nous avons concentré sur la combinaison de trois descripteurs : moments Zernike, l'histogramme HOG et filtre de Gabor. Ces caractéristiques sont fournies un taux de reconnaissance important contrairement d'autres caractéristiques.

## **Organisation de la mémoire :**

Le travail présenté dans cette thèse est organisé comme suit.

Dans le chapitre 1, nous présentons les étapes constituant un système de reconnaissance des chiffres manuscrits. Nous abordons la question de prétraitements, d'extraction de descripteurs ainsi que des méthodes de classification.

Dans le chapitre 2, présente un état de l'art sur différentes méthodes de reconnaissance des chiffres isolés en partant de la description d'une chacune méthode.

Dans Le chapitre 3, est consacré aux méthodes d'extraction des caractéristiques. Tout d'abord, nous revoyons quelques techniques comme les moments Zernike, l'histogramme HOG et filtre de Gabor. Nous introduisons ensuite, la méthode de classification KPPV.

Le chapitre 4, a été alloué pour présenter les résultats obtenus pour la validation du système.



# Chapitre 1:

## La reconnaissance des chiffre

---

# Chapitre 1

---

## La reconnaissance de l'écriture manuscrite

---

### 1-1 Introduction :

Depuis plusieurs décennies La reconnaissance de l'écriture manuscrite est une voie traditionnelle qui a attiré l'attention des chercheurs et reste un domaine de recherche très ouvert dû à son grand nombre d'applications pratiques. Aujourd'hui, les avancées réalisées dans ce domaine se concrétisent par un certain nombre d'applications comme la lecture automatique de chèques bancaires, d'adresses postales ou les adresses de formulaires. C'est un domaine très actif, d'autant plus que les interfaces homme-machine s'orientent de plus en plus vers la communication manuscrite comme les tablettes PC, les applications orientées stylo ...etc. La reconnaissance de l'écriture manuscrite est la transcription des données manuscrites au format numérique et puis de retrouver sa signification [1]. Cette tâche requiert toujours une étape de base dite la segmentation, c.-à-d. La séparation des graphèmes, des mots ou des chaînes numériques. Le problème de la segmentation des caractères manuscrits constitue le défi principal de la reconnaissance de l'écriture manuscrite ainsi plus que la segmentation est correcte plus que le système de reconnaissance est meilleur.

L'une des applications particulière de la reconnaissance est la reconnaissance du montant chiffre sur chèques bancaires du fait de la présence de bruits, le chevauchement de chiffres et même la fragmentation d'un chiffre en plusieurs parties difficilement identifiables [01],[02].

Nous présentons dans ce chapitre les différentes phases de reconnaissance des chiffres manuscrits en mettant l'accent sur toutes les phases pour la construction d'un système de reconnaissance fiable.

### 1.2 Différents aspects de reconnaissance

La première étape dans un système de reconnaissance qui permet de reconnaître n'importe quel caractère dans n'importe quel format, consiste à convertir l'écriture en grandeurs numériques adaptées au système de traitement avec un minimum de

dégradations possibles. Tout dépend du type dispositifs d'acquisition et de données traitées et bien évidemment de l'application visée. Il existe plusieurs modes des systèmes de reconnaissance de l'écritures manuscrits selon le mode d'acquisition [02],[03].

- Les systèmes de reconnaissance en ligne « online ».
- Les systèmes de reconnaissance hors-ligne « offline ».

### **1.2.1. Reconnaissance en-ligne et hors-ligne**

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

#### **1.2.1.1 La reconnaissance En-Ligne**

Ce mode s'appelle aussi dynamique, La reconnaissance d'écriture en-ligne est effectuée en temps réel, c'est-à-dire elle est effectuée pendant le traçage de caractère, ce qui permet d'obtenir une bonne correction et modification selon la réponse donnée à la phase de reconnaissance chevauchée à la phase d'acquisition [03].

Les moyens de saisie en ligne sont nombreux où la tablette graphique avec un stylo électronique et l'écran tactile sont couramment utilisés.

#### **1.2.1.2 La reconnaissance hors-ligne**

Ce mode s'appelle aussi statique où l'information se présente sous formes d'un ensemble de pixels qui représente l'image d'un texte déjà existant, obtenue par un scanner ou une caméra. La reconnaissance de l'écriture hors-ligne est plus complexe que celle qui est en-ligne due à la présence du bruit dans le procédé d'acquisition des images et la perte d'information temporelle telle que l'ordre d'écriture et la vitesse, ce mode introduit une difficulté supplémentaire relative à la variabilité du tracé en épaisseur et en connectivité, nécessitant l'application de techniques de prétraitement[03].

Les domaines d'application sont divers et principalement concernent le traitement automatique des adresses postales, des chèques bancaires, des formulaires, d'identifiants numériques, des feuilles de soins...etc.

### **1.2.2 Reconnaissance des chaînes de chiffres**

La reconnaissance des chaînes de chiffres numériques est une part importante dans le domaine de la reconnaissance de l'écriture manuscrite, elle est différente de celle des chiffres isolés en incluant la classification des chaînes de chiffres [01], la segmentation des chiffres chevauchés, la segmentation des chiffres connectés et la segmentation des chiffres chevauchés et/ou connectés. La segmentation des chaînes de chiffres est généralement une tâche difficile parce que les différents chiffres dans une chaîne peuvent chevaucher et/ou se connecter, ou encore un chiffre peut être constitué de plusieurs fragments [02].

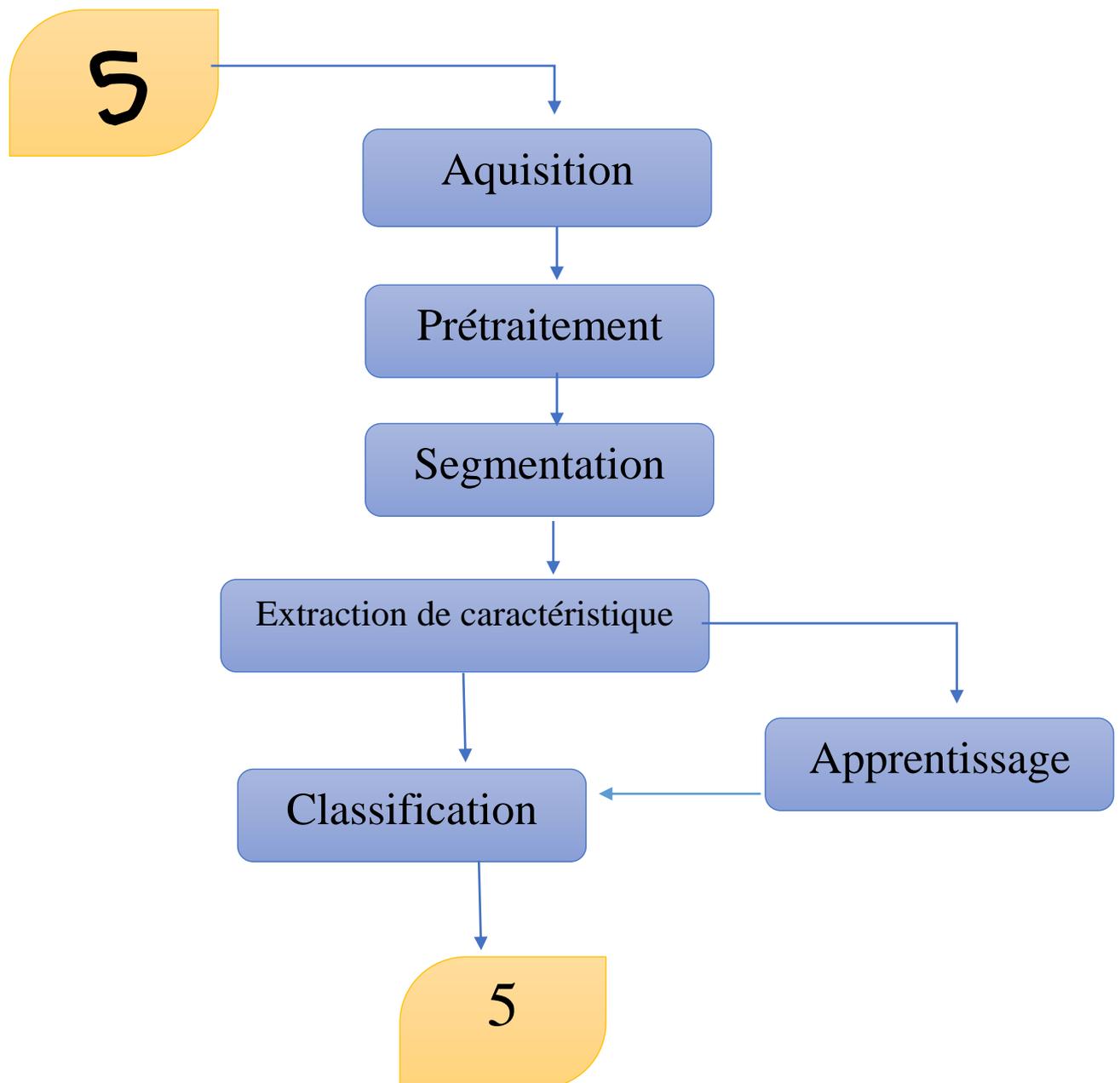
Il peut s'agir de la reconnaissance de numéros de téléphone, numéro de sécurité social, code client, et de code postal ...etc.

De nombreuses applications sont concernées : tri postale, traitement de chèques, traitement de formulaire, courrier.

### **1.3 Processus de système de reconnaissance**

Les systèmes de reconnaissance des chiffres manuscrits sont généralement basés sur les étapes principales suivantes :

Acquisition, prétraitements, segmentation, extraction des caractéristiques, suivis éventuellement d'une phase de classification (Figure. 1.1).



**Figure. 1.1** : Schéma général d'un système de reconnaissance des chiffres manuscrits

[03]

### 1.3.1 Phase d'acquisition

La phase d'acquisition base essentiellement sur la capture de l'image de l'écriture manuscrits au moyen des capteurs physiques et de la convertir en élévations numériques adaptés au système de traitement informatisés avec un minimum de dégradation possible.

Dans le cas où l'information est disponible sur un support souvent papier, les capteurs physiques sont des scanners ou des caméras numériques.

Durant cette phase, malgré la bonne qualité des systèmes d'acquisition, des bruits parasites peuvent apparaître et causer une hétérogénéité du fond [07]. Ceci est dû à la nature de la texture, l'aire de travail et de son éclairage.

### **1.3.2 Phase de prétraitement**

Lorsque l'acquisition est effectuée, la plupart des systèmes comportent une étape de prétraitement. Généralement, ces prétraitements ne sont pas spécifiques à la reconnaissance de l'écriture, mais sont des prétraitements classiques en traitement d'image. Le prétraitement a pour but de préparer l'image du tracé à la phase suivante d'analyse [03]. Il s'agit essentiellement de réduire le bruit superposé aux données et ne garder, autant que possible, que l'information significative de la forme présentée. Le bruit peut être dû au dispositif d'acquisition, aux conditions d'acquisition (éclairage, mise incorrecte du document...), ou encore à la qualité du document d'origine. Parmi les opérations de prétraitements généralement utilisées, citons : la binarisation, le lissage, la squelettisation, la normalisation, et la squelettisation.

#### **1.3.2.1 La binarisation**

Dans un système de reconnaissance d'écriture la numérisation des images est la première étape de traitement. Elle permet de passer d'une image de niveaux de gris à une image binaire composée de deux valeurs 0 et 1, plus simple à traiter. En général, on utilise un seuil de binarisation approprié qui traduit la limite des contrastes fort et faible dans l'image. Selon la méthode de calcul du seuil de binarisation, on distingue deux types de binarisation : le seuillage globale et le seuillage adaptatif [02]. Le premier type consiste à prendre un seuil ajustable mais identique pour toute l'image, cette méthode convient pour les documents de contraste fort et de bonne qualité. Dans le cas contraire, le seuillage adaptatif est plus approprié, dans ce cas le Calcul du seuil se fait localement en fonction du voisinage du pixel traité.

Par contre pour des images peu contrastées ou à contraste variable, il est difficile de fixer ce seuil à une valeur précise. Une méthode classique pour déterminer un seuil de binarisation consiste à calculer l'histogramme des niveaux de gris de l'image. Notons que, la valeur du seuil sera égale à la valeur du niveau de gris se trouvant dans la vallée entre les deux pics de l'histogramme. Les pixels ayant un niveau de gris supérieur à ce

seuil appartient au fond et ceux ayant une valeur inférieure appartiennent à l'objet [02][03].

### 1.3.2.2 Lissage :

Parfois l'image contient du bruit peut être dû au dispositif d'acquisition et souvent à la qualité du document, conduisant soit à des absences de points (trous) soit à des empâtements ou des excroissances et donc à une surcharge de points. Les techniques de lissage permettent de résoudre ces problèmes.

Il permet de réduire au maximum les discontinuités introduites dans l'image au cours des différentes transformations et ainsi de rétablir la régularité et la continuité du contour du mot ou du chiffre.

Le 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 [03].

### 1.3.2.3 La normalisation

L'un des problèmes rencontrés en la reconnaissance des chiffres manuscrits est la taille d'image et l'inclinaison des lignes, qui introduit des difficultés pour la segmentation donc l'étape de normalisation est nécessaire. Dans un système de reconnaissance d'écriture quand la taille varie avec le style de chiffre manuscrit pour chaque scripteur,

#### 1.3.2.3.1 Normalisation de taille

Cette étape permet de ramener les images des chiffres manuscrits à des tailles standard. La normalisation introduit généralement de légères déformations sur les images mais elle reste indispensable pour certains types de systèmes et de méthodes de reconnaissance qui sont sensibles aux petites variations dans la taille et la position comme c'est le cas dans les templates matching et les méthodes de corrélation [4].

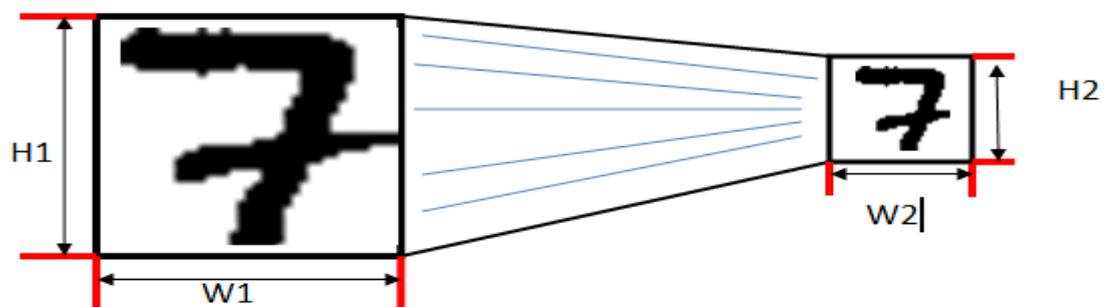


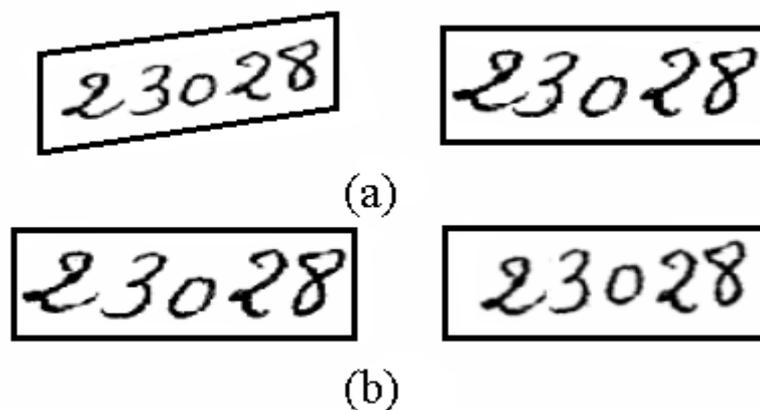
Figure 1 .2 Exemple d'un chiffre manuscrit normalisé

### 1.3.2.3.2 Normalisation de l'inclinaison (Redressement)

Le redressement (également appelée correction de « skew »), est une technique de correction de l'inclinaison qui a été développée pour corriger le défaut de positionnement du papier dans le scanner et aussi il dresse l'inclinaison interne de certaines parties du document qui peuvent apparaître.

En effet, la plupart des algorithmes de reconnaissance ne fonctionnent bien que si le document est bien redressé. Le redressement se fait selon deux niveaux [11]:

- Le redressement de la ligne de base : L'idée est de rendre horizontaux les chiffres à l'aide d'une transformation géométrique de type rotation isométrique des points de l'image.
- Le redressement des écritures penchées : Cette technique facilite la segmentation préalable des chiffres. L'idée est de trouver l'angle moyen d'inclinaison puis de faire une transformation géométrique de type cisaillement de l'image.



**Figure. 1.3:** Exemple de deux types de redressement [3]

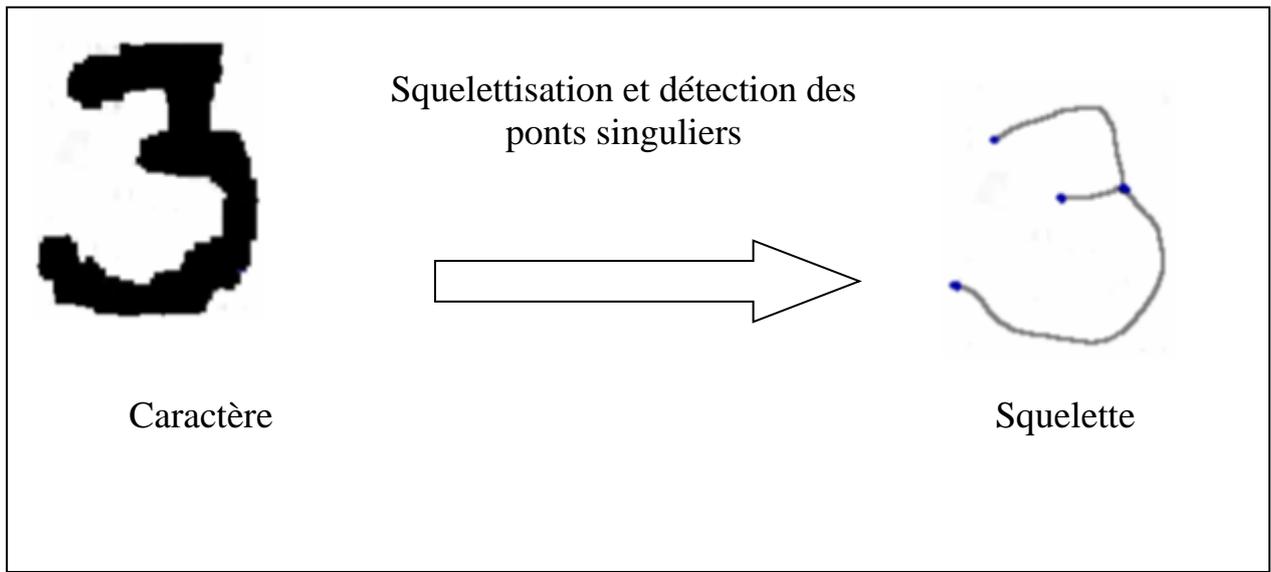
### 1.3.2.4 Squelettisation (Thinning)

Selon la qualité du document à traiter, le type d'écriture (manuscrite ou imprimée), l'image binaire se présente comme une succession de traits d'épaisseurs variables, mal définis, et généralement avec bruit, l'opération de squelettisation est appliquée pour simplifier l'image du caractère en une image à « ligne » plus facile à traiter en la réduisant au tracé du caractère.

Les points du tracé résultant constituent généralement les lignes centrales des points initiaux.

Elle est définie comme étant l'équivalent de la forme, mais avec une épaisseur très réduite qui égale à 1 pixel. La squelettisation ou Amincissement est une opération essentielle dans plusieurs systèmes de reconnaissance d'écriture manuscrite [2].

Les algorithmes de squelettisation se basent sur des méthodes itératives. Le processus s'effectue par passes successives pour déterminer si un tel ou tel pixel est essentiel pour le garder ou non dans le tracé.



**Figure. 1.4:** Exemple de squelettisation.

Selon la qualité du document à traiter, le type de l'écriture et la méthode d'analyse adoptée,

Une ou plusieurs techniques de prétraitement sont utilisées.

### 1.3.3 Phase de Segmentation

Dans un système de reconnaissance de l'écriture manuscrite, les données à traiter sont des images. La mise en œuvre d'une étape de segmentation permet de diviser l'image en différentes régions (imagettes) connexes présentant une homogénéité selon certain critère, et de taille moins importantes qui peuvent être des graphèmes, des lettres ou bien des sous-mots. Cependant une imagette reste une matrice de pixels [11].

L'objectif le plus potentiel de la segmentation des chaînes de chiffres, c'est la séparation des chiffres l'un de l'autre.

D'une manière générale, il existe deux approches de segmentation : l'approche implicite et l'approche explicite.

### **1.3.3.1 Segmentation explicite**

Dans l'approche explicite (appelée aussi discrète), une étape de sélection des points de segmentation les plus probables est effectuée. Certaines liaisons multiples ou à contact prolongé imposent de séparer les composantes selon un chemin. Les chemins de segmentation sont généralement obtenus par des points caractéristiques issus d'une analyse des contours de la forme [07], [08], du squelette ou d'un amincissement du fond, d'une analyse en deux dimensions du tracé [10], ou d'une combinaison analyse des contours/amincissement du fond.

### **1.3.3.2 Segmentation implicite**

Pour contourner le difficile problème du choix des points de segmentation, les approches à segmentation implicite (ou continues) considèrent tous les points du tracé comme points de segmentation potentiels.

Dans cette approche [08], La segmentation et la reconnaissance sont réalisées conjointement, d'où le nom parfois employé de segmentation-reconnaissance intégrée. Il s'agit de méthodes à fenêtres glissantes qui parcourent la séquence de chiffres à l'aide d'une fenêtre de taille fixe, en extrayant des caractéristiques de bas niveau. L'analyse des fenêtres est effectuée soit par un classifieur classique, soit par des modèles dynamiques.

Dans cette approche ne cherche pas à séparer les chiffres mais de préférence à les intégrer implicitement dans le module de reconnaissance [10].

### **1.3.4 Phase d'extraction de caractéristiques**

Dans un système de reconnaissance des chiffres manuscrits Le but ultime d'extraction de caractéristiques est obtenir le volume d'informations la plus pertinentes qui sera fourni au système. C'est une étape critique lors de la construction d'un système de reconnaissance. L'une des raisons pour laquelle cette phase pose un problème est qu'une plusieurs techniques d'extraction s'accompagne d'une perte d'information. De ce fait, il faut effectuer un compromis entre la quantité et la qualité de l'information [09].

La réduction du nombre de caractéristiques a de nombreux avantages : elle permet d'améliorer la visualisation et la compréhension des données, de réduire les temps d'apprentissage et de classification des systèmes, d'améliorer les performances en classification, et permet de réduire la taille des bases d'apprentissage.

Les types de caractéristiques peuvent être classés en quatre catégories principales de Primitives: structurelles, statistiques, globales (transformations globales) et topologiques ou Métriques [13].

### 1.3.4.1 Les primitives structurelles

Les primitives structurelles (ou primitives locales) basées sur une représentation linéaire du caractère (décomposition du caractère en segments de droites et courbes, contours du caractère, squelette) [09].

Les primitives structurelles sont généralement extraites non pas de l'image brute, mais via une représentation de la forme par le squelette ou par le contour [13]. Ainsi, parmi ces caractéristiques, il s'agit principalement des segments de droite, des arcs, boucles et concavités, des pentes, la hauteur et la largeur du caractère... etc.

### 1.3.4.2 Les Primitives statistiques :

Les primitives statistiques portant des informations concernant la distribution des pixels dans l'image du caractère ou chiffres, Les primitives statistiques décrivent une forme en termes d'un ensemble de mesures extraites à partir de cette forme. Les caractéristiques utilisées pour la reconnaissance des chiffres manuscrits sont :

- Le zonage (zoning) : consiste à répartir l'image en région ou zones. L'image source est une image binaire [3]. Pour chacune des régions résultantes, il s'agit de calculer la moyenne ou le pourcentage de points noirs.



**Figure. 1.5:** Division d'une image en zone

- L'histogramme : c'est un moyen de représentation du nombre de pixels sur chaque ligne ou colonne de l'image, en est un exemple classique et simple à calculer (histogramme directionnel, histogramme des transitions ...) [2].

Il est possible d'utiliser une autre primitive statistique basée sur un moyennage des pixels situés à l'intérieur d'un masque rectangulaire. La construction d'une matrice de masque recouvrant la totalité de la forme permet une représentation statistique à partir d'un nombre très réduit de valeurs correspondant à chaque masque.

### **1.3.4.3 Les primitives globales**

Les primitives globales basées sur une transformation globale de l'image. La caractéristique d'une primitive globale est de dépendre de la totalité des pixels d'une image, ces primitives sont donc dérivées de la distribution des pixels [09]. Ils dirigent trois familles de caractéristiques telles que : les moments invariants, les projections et les profils.

Une des transformations les plus simples est celle qui représente le squelette ou le contour d'un caractère sous forme d'une chaîne de codes de directions (code de Freeman). Parmi les méthodes de transformation, nous pouvons citer les transformées de Hough et de Fourier et d'ondelettes [02].

### **1.3.4.4 Primitives topologiques ou métriques**

Dans Primitives topologiques ou métriques basé essentiellement sur des densités de pixels.

Il s'agit d'effectuer une mesure sur l'échantillon au moyen d'une métrique. Parmi les mesures opère, nous pouvons note :

- Compter le nombre de trous,
- Evaluer les concavités,
- Mesurer des pentes, des courbures et évaluer des orientations principales,
- Mesurer la longueur, surfaces, les périmètres et l'épaisseur des traits,
- Détecter les croisements et les jonctions des traits,
- Déterminer le rectangle délimitant l'échantillon, ou le polygone convexe,
- Rendre compte de la disposition relative de ces primitives.

Toutes ces mesures peuvent être intégrées dans un seul vecteur de caractéristiques pour la Reconnaissance du chiffre manuscrit.

Lorsque le nombre de caractéristiques devient trop élevé, des méthodes de sélection de caractéristiques peuvent être mises en œuvre [09], [10].

### 1.3.5 Phase de Classification

Après la segmentation de caractères et l'extraction des attributs caractéristiques, une étape de reconnaissance basée classification de caractères est employée.

Les techniques de la classification reposent sur une stratégie de décision qui permet de catégoriser un objet le mieux possible selon certains critères d'optimisation. Et qui transforme les attributs caractérisant les formes en appartenance à une classe (passage de l'espace de codage vers l'espace de décision) [4].

Le type d'une méthode de classification se décline généralement en deux familles : Le mode supervisé et le mode non supervisé.

- **Classification supervisée** : cette technique est basée sur l'étiquetage des observations en affectant chaque observation à une classe (supervisés où la sortie correcte doit être fournie à l'avance).
- **Classification non supervisée** : aucune des observations n'est étiquetée (non supervisés où la sortie correcte n'est pas exigée à l'avance, elle résulte après une étape d'apprentissage).

Dans le cas au la reconnaissance des chiffres manuscrits, la classification supervisée a été adoptée car elle simplifie considérablement le problème.

Généralement, pour évaluer la performance d'un classificateur, on poursuit deux étapes Principales qui sont l'apprentissage et le test (Reconnaissance et décision).

- **Étape d'apprentissage** : Les données d'apprentissage sont les données utilisées pour construire un classificateur capable de reconnaître des formes inconnues pour caractériser les classes [5]
- **Étape de Reconnaissance et décision** : Elle cherche parmi les modèles de référence en présence, ceux qui lui sont les plus proches. Le problème revient à affecter une forme inconnue à l'une des classes obtenues pendant l'apprentissage [08]. Le résultat de la décision est un « avis » sur l'appartenance ou non de la forme aux modèles de l'apprentissage.

La reconnaissance peut conduire à :

- Un succès si la réponse est unique (un seul modèle répond à la description de la forme du caractère).
- Une confusion si la réponse est multiple (plusieurs modèles correspondent à la description).

- Un rejet de la forme si aucun modèle ne correspond à sa description.

Si toutes les données sont employées pour l'apprentissage et le test en même temps, il est possible que le classificateur soit incapable de reconnaître d'autres formes inconnues. C'est pourquoi il est important d'avoir trois ensembles de données pour améliorer la généralisation d'un classificateur : le premier ensemble est spécifié pour l'apprentissage, le 2<sup>ème</sup> pour la validation, et le dernier pour le test. Les données de validation sont utilisées pour continuer le processus d'apprentissage jusqu'à ce que les performances sur l'ensemble de validation ne s'améliorent plus. En ce moment, l'apprentissage devrait être arrêté afin d'éviter

Le sur-apprentissage (over-training) du classificateur où sa capacité de généralisation commence à diminuer.

Il existe trois approches principales et une approche hybride :

### **1.3.5.1 Approche statistique**

Dans cette approche La reconnaissance est basée sur l'étude statistique des mesures que l'on effectue sur les formes à reconnaître [10]. Mais elle a besoin d'un nombre élevé d'exemples pour effectuer 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 (la règle de Bayes, les réseaux de neurones, les chaînes de Markov...) [10] [19]
- Les méthodes non paramétriques (méthode des k plus proches voisins, ...) [8] [10]

### **1.3.5.2 Approche structurelle**

Les méthodes structurelles reposent sur l'extraction de primitives en prenant compte de l'information structurelle (la structure physique des caractères) [9]. De manière générale, les approches syntaxiques ou structurelles permettent la description de formes complexes à partir de formes 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 [13]. 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,...

### 1.3.5.3 Approche stochastique

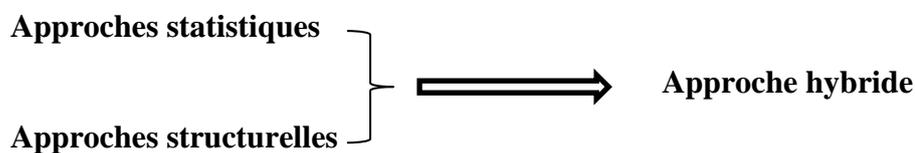
L'approche stochastique utilise un modèle pour la reconnaissance, prenant en compte la grande variabilité de la forme. Dans ce type d'approche, les modèles sont souvent discrets et de nombreux travaux reposent sur la théorie des champs de Markov et l'estimation bayésienne.

Les champs de Markov permettent de ramener des propriétés globales à des contraintes locales. Le modèle décrit ces états à l'aide de probabilités de transitions d'états et de probabilités d'observation par état [05]. La comparaison consiste à chercher dans ce graphe d'état, le chemin de probabilité forte correspondant à une suite d'éléments observés dans la chaîne d'entrée. [10]. Les méthodes les plus répandues dans cette approche sont les méthodes utilisant les modèles de Markov cachés (HMM).

### 1.3.5.4 Approche hybride :

Pour améliorer les performances de reconnaissance, la tendance aujourd'hui est de construire des systèmes hybrides qui utilisent différents types de caractéristiques, et qui combinent plusieurs classifieurs en couches. pour surmonter les faiblesses de chaque approche et obtenir des résultats plus précis plus meilleure que les résultats qui auraient été obtenus si la application de chaque approche séparément, comme les approche qui ont été fusionnés pour former celui intégré [03].

Un exemple de cette approche une approche hybride qui combine les approches statistiques et structurelles



**Figure 1.6.** Exemple d'hybridation entre deux approches (statistiques et structurelles)

## 1.4 Méthodes de classification statistiques

L'approche statistique a été retenue pour la reconnaissance des chiffres qui s'appuient sur des bases théoriques telles les méthodes paramétriques. Parmi les méthodes de classification qui ont été développées dans ce contexte (machine

d'apprentissage), nous distinguons par exemple les réseaux de neurones et les machines à vecteurs de support (SVM : Support Vector Machines) ainsi que K plus proches voisins (K Nearest Neighbors)[59].

### **1.4.1 Réseaux de neurones**

Le terme de réseau de Neurone peut prendre des significations différentes [06]

- Réseaux artificiels
- Neurosciences calculatoire
- Modélisation neuronale

Les réseaux de Neurones est une dénomination assez floue qui recouvre un ensemble de mécanismes de calcul inspirée initialement par des modèles issus du fonctionnement des systèmes nerveux que des techniciens y voient une source d'inspiration pour la construction des systèmes automatiques [04].

Un grand effort a été consacré au développement des réseaux neurones qui présentent une Alternative aux architectures classiques. Ceci est dû à leurs structures parallèles, à leurs Performances de classification et à leurs aptitudes de comprendre des phénomènes non linéaire.

Principalement un réseau de Neurone ne fonctionne qu'après avoir appris suffisamment de connaissances sur les sorties désirées à partir d'entrées données ; le développement de ces réseaux concerne essentiellement l'apprentissage et les lois de modification des poids de connexion. [06].

### **1.4.2 Les Machines à Vecteur Support (SVMs)**

Les SVMs sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de la classification binaire et de la régression. Les SVMs reposent sur deux idées.

Principales : la notion de la marge maximale et la notion de la fonction noyau. La marge maximale est employée pour les problèmes de la classification linéaire. Elle représente la distance entre la frontière de séparation et les échantillons d'apprentissages les plus proches. Ces derniers sont les vecteurs supports. Les fonctions noyau sont employées dans le cas des problèmes de la classification non-linéaire pour transformer l'espace de représentation des données d'entrées en un espace de plus grande dimension dans lequel il est probable qu'il existe de séparateurs linéaires [51].

Le classificateur SVM est un algorithme qui maximise la marge entre les classes du problème à résoudre et réduit au minimum l'erreur de classification. L'objectif de la marge maximale est de faire séparer deux classes par un hyperplan de telle sorte la distance par rapport aux vecteurs supports soit maximale [52].

Dans la tâche de classification, un SVM construit l'hyperplan optimal de séparation des attributs caractéristiques dans un espace de haute dimension [02]. Le calcul de cet hyperplan est fondé sur la maximisation de la marge entre les exemples d'apprentissages les plus proches qui appartiennent à différentes classes.

### **1.4.3 Le classifieur du K plus proches voisins**

Le classificateur du K plus proches voisins(K Nearest Neighbors) C'est un classificateur simple basé sur le calcul de distance entre les exemples d'apprentissage et les exemples de tests, généralement la norme euclidienne est souvent employée comme mesure de distance, dans chaque étape de l'apprentissage[64], l'algorithme mémorise les k meilleurs exemples de l'ensemble d'apprentissage ( $kppv(x)$ ) qui sont proches à l'exemple de test x. Cet algorithme est souvent performant s'il y a suffisamment d'exemples d'apprentissage, mais demande un temps de prédiction très long pour passer tous les exemples afin de trouver les K meilleures solutions [11][65].

## **1.5 Conclusion**

Dans ce chapitre, nous avons présenté les différents concepts généraux d'un système de reconnaissance des chiffres manuscrits. Ainsi nous avons démontré les problèmes liés à la complexité de la reconnaissance

Le module de segmentation de deux chiffre connecte constitue le principal objectif de notre travail. Aussi, nous présentons dans le chapitre suivant, les différentes méthodes de segmentations disponibles en citant les avantages et les inconvénients de chacune méthode.



# Chapitre 2:

## Etat de l'art

---

---

## Chapitre 2

---

---

### Etat de l'art

---

---

#### 2-1 Introduction :

Pour concevoir un système de reconnaissance d'un chiffre isolé manuscrite, il faut avoir les différents travaux ultérieurs qui 'ils ont réalisé sur la base de données M.N.I.S.T. Dans ce chapitre, nous décrivons de ces travaux. En précisant la performance et la résultat obtenus.

#### 2-2 LES TRAVEAUX ULTERIERS :

Dans cette section, on a présenté les différents travaux consacre la reconnaissance des chiffres manuscrits en fonction de l'approche utiliser et le résultat obtenu .

##### 2-2-1 Mayraz et Hinton. (2002) :

G. Mayraz et G. E. Hinton. 2002 utilisent le produit d'une procédure d'apprentissage d'experts qui permet de découvrir un ensemble de caractéristiques binaires stochastiques qui constituent un modèle génératif non linéaire d'images de chiffres manuscrites. La qualité des modèles génératifs formés peut être évaluée par l'apprentissage d'un modèle distinct pour chaque catégorie de chiffres, puis en comparant les probabilités non normalisées d'images de test sous les 10 différents modèles de classe spécifique. Pour améliorer les performances discriminatoires, une hiérarchie de modèles distincts peut être formée pour chaque classe de chiffres. Chaque modèle de la hiérarchie apprend une couche de détecteurs de caractéristiques binaires qui modélisent la distribution de probabilité des vecteurs d'activité des détecteurs de caractéristique dans la couche inférieure. Les modèles de la hiérarchie sont formés séquentiellement et chaque modèle utilise une couche de détecteurs de caractéristiques binaires pour former un modèle génératif des modèles d'activités caractéristiques dans la couche précédente. Après la formation, chaque couche de détecteurs de caractéristiques est produite séparée, non normalisée de score de probabilités. Avec trois couches de détecteurs de caractéristiques pour chacune des classes de 10 chiffres, une image de test produisent scores 30 qui peuvent être utilisés comme entrants dans un réseau de

classification supervisée et logistique qui est formé sur des données distinctes. Sur la base de données MNIST, ce système démontrant que le produit d'une procédure d'apprentissage d'experts peut produire des hiérarchies efficaces de modèles génératifs de données de grande dimension. Le taux de reconnaissance atteint 98.3% sur la base de donnée MNIST qui constitue 60 000 images d'apprentissage et 10 000 images de test[63].

#### **2-2-2 Kussul et Baidyk (2004):**

Kussul et Baidyk [] ont développé un nouveau classifieur neuronal Limited Réceptive Area (LIRA) pour la reconnaissance d'image. Le LIRA classifieur contient trois couches de neurones : entrée, associative et sortie. La couche d'entrée est reliée à la couche associative sans connexion aléatoire modifiable et la couche associative est reliée à la couche de sortie avec des connexions de l'apprentissage. Le processus d'apprentissage converge suffisamment rapide. Ce classificateur n'utilise pas d'opérations en virgule flottante et de multiplication. Le classifieur a été testé sur deux bases de données. La première base de données est la base de données MNIST. El contient 60.000 images de chiffres manuscrits pour l'apprentissage et 10.000 images de chiffres manuscrits pour le test. La seconde base de données contient 441 images de micro dispositif d'assemblage. Le classifieur LIRA donne un taux de reconnaissance de 99.59%[64].

#### **2-2-3 Wu et Zhang (2010) :**

Wu et Zhang ont utilisé les caractéristiques de direction extraites pour la réduction de la dimensionnalité. K-ième voisin le plus proche, les modèles de mélanges gaussiens et SVM se révélant les meilleurs candidats pour les caractéristiques. Le taux de reconnaissance atteint 98.81% sur la base de donnée MNIST qui constituer 60 000 images d'apprentissage et 10 000 images de test. En utilisant 3-NN[65].

#### **2-2-4 Ebrahim et Jampour (2014) :**

Ebrahim et Jampour ont proposé une approche à base de fonction de l'appariement qui traitent des données en utilisant l'Histogramme de gradient Orienté (HOG). HOG est un descripteur de caractéristique très efficace pour les chiffres

manuscrits, qui est stable sur des variations d'éclairément, car il est à base d'un descripteur de gradient. En outre, SVM linéaire a été employé comme classificateur qui a de meilleures résultats que les noyaux RBF et sigmoïde. Le taux de reconnaissance atteint 97,25% sur la base de données MNIST qui constitue 60 000 images d'apprentissage et 10 000 images de test[66].

#### **2-2-5 Wakahara et Yamashita (2014) :**

Global affine transformation (GAT) est proposée par Wakahara et al. C'est une méthode de modèle qui base sur l'appariement (matching) qui peut compenser la transformation affine incorporé dans un modèle d'entrée. La méthode de corrélation GAT a démontré une haute performance dans la reconnaissance de caractères et le matching des objets. Une nouvelle mesure d'adaptation appelée Nearest Neighbors Distance of Equi-Gradient Direction (NNDEGD) utilisé en coopération avec la méthode de corrélation GAT. Le NNDEGD est juste le paramètre de la fenêtre de la fonction gaussienne utilisée dans la méthode de corrélation GAT, qui est égale à la distance minimale moyenne entre un point dans une image et un autre point dans l'autre image avec la même direction de gradient. Ensuite ils utilisent cette valeur comme une nouvelle mesure d'appariement. Deuxièmement, ils ont étendu la méthode de corrélation GAT pour gérer le changement de largeur de trait en plus de la transformation affine. Enfin, ils appliquent la version originale et étendue de la méthode de corrélation GAT à la classification de FC-NN expériences en utilisant la base de données MNIST. Ces expériences sont réalisées de manière efficace pour la première fois parce que ils ont réduit considérablement la complexité et la mémoire de charge de calcul impliqués par rapport à la méthode de corrélation de GAT originale. Le taux de reconnaissance atteint 99.51% sur la base de donnée MNIST qui constituer 60 000 images d'apprentissage et 10 000 images de test[67].

#### **2-2-6 Lee et al (2015) :**

Dans cet article Lee et al ont amélioré les réseaux de neurones profonds en généralisant les opérations de pooling qui jouent un rôle central dans les architectures actuelles. Ils ont poursuivi une exploration minutieuse des approches pour permettre la pooling d'apprendre et de s'adapter à des modèles complexes et variables. Les deux directions principales se trouvent dans (1) l'apprentissage d'une fonction de pooling via

(deux stratégies) combinant des max et la moyenne de pooling, et (2) l'apprentissage d'une fonction de pooling sous la forme d'une fusion à structure arborescente des filtres de pooling eux-mêmes appris. Dans leurs expériences chaque opération de pooling généralisée, ils explorent à améliorer les performances lorsqu'il est utilisé à la place de max pooling. Le taux de reconnaissance atteint 99.71% sur la base de donnée MNIST qui constituer 60 000 images d'apprentissage et 10 000 images de test[68].

### **2-2-7 Wakahara et Yamashita (2016) :**

Le model de la reconnaissance basée sur l'apparement (matching) ne nécessite pas un processus d'apprentissage et le résultat de matching fournit des information intuitive et géométrique. Wakahara et al. Ont proposé Global Affine Transformation (GAT) pour la corrélation d'appariement, ce qui peut compenser les transformations affines imposées à un modèle. La corrélation d'appariement de GAT avec une méthode d'accélération et une nouvelle mesure d'apparement, appelée Nearest-Neighbor Distance of Equi-Gradient Direction (NNDEGD), atteint de hautes performances dans des expériences en utilisant la base de données MNIST. La mesure d'apparement GAT a été étendue à global projection transformation de mesure d'apparement (GPT) de permettre une déformation par transformation de projection 2D.

Leur travail consiste d'abord à développer une méthode d'accélération pour GPT d'apparement de corrélation. Deuxièmement, à améliorer les performances de reconnaissance, ils ont appliqué la courbure (curvature) des contours en traits à la mesure d'apparement. la Courbure (curvature) est souvent utilisée comme une caractéristique des caractères. Toutefois, ils avaient utilisé comme un poids dans le NNDEGD. Troisièmement, pour une enquête sur les performances des méthodes proposées, ils ont appliqué l'apparement de l'image et la reconnaissance des bases de données MNIST et IPTP pour  $k$ PPV ( $k$ -NN). Dans l'expérience avec la base donnée MNIST, l'apparement de corrélation GPT avec la mesure d'apparement de la courbure pondéré NNDEGD a achevé le taux de reconnaissance le plus élevée de 99.70% avec les méthodes  $k$ -NN [69] .

### **2-2-8 Dundar et al (2016) :**

Dundar et al Ont proposé de former un réseau de convolution profond basé sur une version améliorée des algorithme k-means, ce qui réduit le nombre de paramètres

corrélés dans la forme de filtres similaires, et ainsi augmente la précision de la catégorisation de test. Les expériences montrent que l'algorithme proposé surpasse les autres techniques qui apprennent les filtres non supervisés. Plus précisément, Le taux de reconnaissance atteint 98.6% sur la base de donnée MNIST qui constituer 60 000 images d'apprentissage et 10 000 images de test[70].

### 2-3 Quelque résultat expérimental obtenu par les auteurs

Dans la section précédente on a présenté différentes méthodes et algorithmes a compris le principe de mise en œuvre. Le tableau au-dessus, illustre les performances et l'approche utilisé pour chaque méthode.

| Année | EDITEUR                 | Taux (%) | METHODE                                |
|-------|-------------------------|----------|--|
| 2002  | Mayraz et Hinton        | 98.3     | Produit expert                         |
| 2004  | Kussul et Baidyk        | 99.59    | LIRA                                   |
| 2010  | Wu et Zhang             | 98.81    | Caractéristique de direction +SVM      |
| 2014  | Ebrahimzadeh et Jampour | 97.25    | HOG + SVM                              |
| 2014  | Wakahara et Yamashita   | 99.51    | GAT NNDEGD                             |
| 2015  | Lee et al               | 99.71    | Réseaux neurons                        |
| 2016  | Wakahara et Yamashita   | 92.5     | NNDEGD+GPT                             |
| 2016  | Dundar et al            | 98.6     | réseaux convolutional profond + Kmeans |

**Tableau 2.1** Quelque résultats rapportés dans la littérature.

## **2-4 Conclusion**

Dans ce chapitre nous avons présenté un aperçu sur état de l'art des méthodes et techniques des segmentations des chiffres manuscrits, notamment on définissant les approches les bien connues dans la littérature : l'approche de segmentation explicite, l'approche segmentation implicite et l'approche hybride. Ensuite nous avons présenté les différents algorithmes de segmentation a compris le principe de mise en œuvre, l'avantage et les inconvénients de chaque méthode et algorithme.



# Chapitre 3:

## *Extraction des caractéristiques*

---

## Chapitre 3

---

### *Extraction des caractéristiques*

---

#### **3-1- Introduction :**

Ce chapitre est consacré à la présentation des méthodes d'extraction des caractéristiques. La première partie présente les trois méthodes choisies accompagnées par leur définition détaillée. La deuxième partie sera accordée au classifieur utilisé. En conclusion ce chapitre par une conclusion.

#### **3-2- Histogramme de gradient Orienté :**

Histogramme de gradient Orienté (Histogram of Oriented Gradients HOG) a été proposé par Dalal et Triggs pour la détection de corps humain, mais il est maintenant l'un des descripteurs utilisés avec succès et populaires dans la vision d'ordinateur et reconnaissance de formes. Le descripteur HOG compte les occurrences d'orientation gradient en partie d'une image où il est un descripteur d'apparence. Il divise l'image d'entrée en petites cellules carrées (par exemple  $9 \times 9$ ), puis calcule l'histogramme des directions de gradient ou les directions de bord sur la base des différences centrales. Pour améliorer la précision, les histogrammes locaux ont été normalisés sur la base du contraste et ceci est la raison pour laquelle HOG est stable sur la variation d'éclairage. En le comparant avec SIFT et LBP, HOG est plus rapide et plus stable sur la variation d'illumination en raison de calculs simples., il a été également montré que les caractéristiques du HOG sont des descripteurs avec succès pour la détection.[71]

##### **3-2-1- PRINCIPE :**

###### **3-2-1-1- Description générale:**

L'idée importante derrière le descripteur HOG est que l'apparence et la forme locale d'un objet dans une image peuvent être décrites par la distribution de l'intensité du gradient ou la direction des contours. Ceci peut être fait en divisant l'image en des régions adjacentes de petite taille, appelées cellules, et en calculant pour chaque cellule l'histogramme des directions du gradient ou des orientations des contours pour les pixels à l'intérieur de cette cellule. La combinaison des histogrammes forme alors le descripteur HOG. Pour de meilleurs résultats, les histogrammes locaux sont normalisés en contraste, en calculant une mesure de l'intensité sur des zones plus larges que les cellules, appelées des blocs, et en utilisant cette valeur pour

normaliser toutes les cellules du bloc. Cette normalisation permet une meilleure résistance aux changements d'illuminations et aux ombres.

La pensée essentielle derrière l'histogramme des gradients orientés descripteur est que l'apparence de l'objet local et la forme d'une image peut être décrite par la distribution des gradients d'intensité ou les directions de bord. L'image est divisée en petites régions appelées cellules connectées et pour les pixels à l'intérieur de chaque cellule, un histogramme des directions de gradient est compilé. Le descripteur est la concaténation de ces histogrammes. Pour une meilleure précision, les histogrammes locaux peuvent être de contraste normalisé par le calcul d'une mesure de l'intensité dans une région agrandie de l'image, appelé bloc, puis en utilisant cette valeur de normaliser toutes les cellules à l'intérieur du bloc. Cette normalisation se traduit par une meilleure invariance aux changements dans l'éclairage et l'ombrage.

Le descripteur de HOG a quelques avantages clés par rapport aux autres descripteurs. Comme il fonctionne sur les cellules locales, il est invariant aux transformations géométriques et photométriques, sauf pour l'orientation de l'objet. De tels changements n'apparaissent dans les régions spatiales plus grandes. En outre, comme Dalal et Triggs découvert, l'échantillonnage spatiale grossière, fin échantillonnage d'orientation, et la normalisation forte photométrique locale permet la circulation des piétons du corps individuel à être ignoré aussi longtemps qu'ils maintiennent une position sensiblement verticale.

Dalal et Triggs ont testés de nombreuses combinaisons de pré et post-traitements ou de valeurs de paramètres. On ne décrit ici que les valeurs et les méthodes retenues. [71]

### **3-2-1-2- Calcul du gradient:**

La première étape de calcul dans de nombreux détecteurs de caractéristiques dans l'image prétraitement est d'assurer la couleur et gamma valeurs normalisées. En ce Dalal et Triggs remarquer, cependant, cette étape peut être omise dans le descripteur HOG calcul, en tant que descripteur de normalisation qui s'ensuit permet d'obtenir essentiellement le même résultat. Image prétraitement fournit donc peu d'impact sur les performances. Au lieu de cela, la première étape de calcul est le calcul des valeurs de gradient. La méthode la plus courante consiste à appliquer le 1-D centrée, point discret masque dérivé dans l'un ou l'autre des directions horizontale et verticale. Plus précisément, cette méthode exige le filtrage des données de couleur ou d'intensité de l'image avec Les masques suivants :

$[-1, 0, 1]$  et  $[-1, 0, 1]^T$

Dans le cas des images couleurs, le gradient est calculé séparément pour chaque composante, et on retient pour chaque pixel le gradient de plus grande norme. [71]

### **3-2-1-3- Construction de l'histogramme:**

La deuxième étape de calcul est la création d'histogrammes de la cellule. Chaque pixel à l'intérieur de la cellule émet un vote pondéré pour un canal d'histogramme fondée sur l'orientation en fonction des valeurs trouvées dans le calcul du gradient. Les cellules elles-mêmes peuvent être soit rectangulaire ou radiale en forme, et les canaux d'histogramme sont uniformément répartis sur 0 à 180 degrés ou 0 à 360 degrés, selon que le gradient est « non signé » ou « signé ». En ce qui concerne le poids de vote, la contribution de pixel peut être soit le gradient grandeur elle-même, ou une fonction de l'ampleur. Dans les tests, le gradient grandeur se produit généralement les meilleurs résultats. D'autres options pour le poids de vote pourraient inclure la racine carrée ou carré de la grandeur de gradient, ou une version écrêtée de l'ampleur. [71]

### **3-2-1-4- Formation des blocs:**

Pour tenir compte des changements dans l'éclairage et le contraste, les forces de gradient doivent être normalisées au niveau local, ce qui nécessite le regroupement des cellules ensemble dans de plus grands blocs, spatialement connectés. Le descripteur de HOG est alors le vecteur concaténé des composants des histogrammes normalisés de cellules provenant de toutes les régions du bloc. Ces blocs se chevauchent généralement, ce qui signifie que chaque cellule contribue plus d'une fois au descripteur final. Deux principales géométries de blocs existent: des blocs rectangulaires R-HOG et circulaires blocs C-HOG. R-blocs HOG sont généralement des grilles carrées, représentée par trois paramètres: le nombre de cellules par bloc, le nombre de pixels par cellule et le nombre de canaux par l'histogramme de la cellule. [71]

### **3-2-1-5- Normalisation des blocs:**

Dalal et Triggs exploré quatre méthodes différentes pour le bloc de normalisation. Soit  $v$  le vecteur non normalisé contenant tous les histogrammes dans un bloc donné,  $\|v\|_k$  être son k-norme par  $K=1,2$  et  $\epsilon$  est une petite constante. Ensuite, le facteur de normalisation peut être l'un des éléments suivants :

$$L2 - norme: f = \frac{v}{\sqrt{\|v\|_2^2 + \epsilon^2}}$$

$$L1 - norme: f = \frac{v}{(\|v\|_1 + \epsilon)}$$

$$L1 - racine: f = \sqrt{\frac{v}{(\|v\|_1 + \epsilon)}}$$

Une quatrième norme L2-hys, consistant à calculer v tout d'abord par la L2-norme, puis limiter les valeurs maximales de v à 0.2, et ensuite à normaliser. Les normes L2-Hys, L2-norme, et L1-racine obtiennent des performances similaires, tandis que L1-norme obtient de moins bons résultats, mais toutefois bien meilleurs que l'absence de normalisation. [71]

### 3-3- Les moments de Zernike :

Les polynômes de Zernike ont été proposées en 1934 dans le cadre de la théorie de la diffraction optique. Dérivés de ces polynômes, les moments ont été utilisés par de nombreux auteurs en reconnaissance de caractères. Plusieurs études montrent également la supériorité de ces descriptions par rapport à d'autres approches (En termes de résilience de bruit, redondance de l'information et de la capacité de reconstruction). Les moments  $A_{nm}$  de Zernike correspondent à la projection de la forme  $f(x, y)$  sur une base *Zernik Polimomial* (ZP) de fonctions orthogonales  $V_{nm}(x, y)$  et se définissent par :

$$A_{nm} = \frac{n+1}{\pi} \sum_{(x+y) \in I} V_{nm}(x, y)^* f(x, y)$$

Définissant le complexe conjugué. La base ZP est définie sur le cercle unité par :

$$ZP = \{V_{nm}(x, y) | (x^2 + y^2) \leq 1\}$$

où le polynôme complexe  $V_{nm}$  d'ordre  $n$  et de répétition  $m$  est défini avec  $n \in \mathbb{N}^+$  et  $m \in \mathbb{N}$  tel que

$$n - |m| \text{ soit pair et } |m| \leq n :$$

$$V_{nm}(x, y) = R_{nm}(r) e^{jm\theta} \text{ En coordonnées polaires (2.45)}$$

Avec :

$$R_{nm}(r) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)! r^{n-2s}}{s! \left(\frac{n+|m|}{s} - s\right)! \left(\frac{n-|m|}{s} - s\right)!} \quad (2.46)$$

Les moments sont invariants par rotation, translation et changement d'échelle (après normalisation de la taille de la forme). De plus, grâce à l'exploitation d'une base de fonctions orthogonales, ces moments sont peu corrélés. [72]

Cette représentation est inversible, l'image peut être reconstruite de la manière suivante :

$$f(x, y) = \lim_{n \rightarrow +\infty} \sum_{n=0}^N \sum_m A_{nm} V_{nm}(x, y) \quad (2.47)$$

L'ordre des moments possède une grande influence sur la conservation de l'information angulaire.

Plus l'ordre est élevé et plus les variations angulaires décrites sont fines ([73], [74]).  
CHONG & AL [75].

### 3-4- Filtre de Gabor :

Dans le traitement d'image, un filtre de Gabor, nommé d'après Dennis Gabor, est un filtre linéaire utilisé pour la détection de bord. Fréquence et orientation représentations de filtres de Gabor sont semblables à ceux du système visuel humain, et ils se sont avérés être particulièrement approprié pour la représentation de la texture et de la discrimination. Dans le domaine spatial, un filtre 2D Gabor est une fonction noyau gaussienne modulée par une onde sinusoïdale. Des cellules simples du cortex visuel du cerveau des mammifères peuvent être modélisés par des fonctions de Gabor. Ainsi, l'analyse de l'image avec des filtres de Gabor est considéré comme similaire à la perception du système visuel humain.

Sa réponse impulsionnelle est définie par une onde sinusoïdale (une onde plane pour les filtres 2D Gabor) multipliée par une fonction gaussienne. En raison de la propriété multiplication-convolution (convolution théorème), la transformée de Fourier de la réponse impulsionnelle d'un filtre de Gabor est la convolution de la transformée de Fourier de la fonction harmonique et la transformée de Fourier de la fonction gaussienne. Le filtre a une composante réelle et une composante imaginaire représentant des directions orthogonales. Les deux composantes peuvent être formés en un nombre complexe ou utilisés individuellement.

Complexe :

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \exp\left(i\left(2\pi \frac{x^1}{\lambda} + \psi\right)\right)$$

Réel :

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x^1}{\lambda} + \psi\right)$$

Imaginaire :

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x^1}{\lambda} + \psi\right)$$

Quand :

$$x' = x \cos \theta + y \sin \theta$$

Et :

$$y' = -x \sin \theta + y \cos \theta$$

Dans cette équation,  $\lambda$  représente la longueur d'onde du facteur sinusoïdal,  $\theta$  représente l'orientation de la normale aux bandes parallèles d'une fonction de Gabor,  $\psi$  est la phase de décalage,  $\sigma$  est écart type de l'enveloppe gaussienne et  $\gamma$  est le rapport d'aspect spatial, et détermine l'ellipticité de l'appui de la fonction de Gabor.

Dans le traitement de l'image des documents, les caractéristiques de Gabor sont idéales pour identifier le script d'un mot dans un document multilingue. Gabor filters avec des fréquences différentes et avec des orientations dans des directions différentes ont été utilisées pour localiser et extraire les régions de texte uniquement à partir d'images de documents complexes (à la fois gris et couleur), puisque le texte est riche en composants à haute fréquence, alors que les images sont relativement douces dans la nature. Il a également été appliquée pour la reconnaissance des expressions faciales, les applications d'analyse de motif, la reconnaissance optique de caractères, la reconnaissance de l'iris et la reconnaissance d'empreintes digitales.

### **3-5- K plus proches voisins K-NN (KPPV) :**

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). [76]

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

#### **3-5-1- Principe de classifieur KPPV**

Etant 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 :

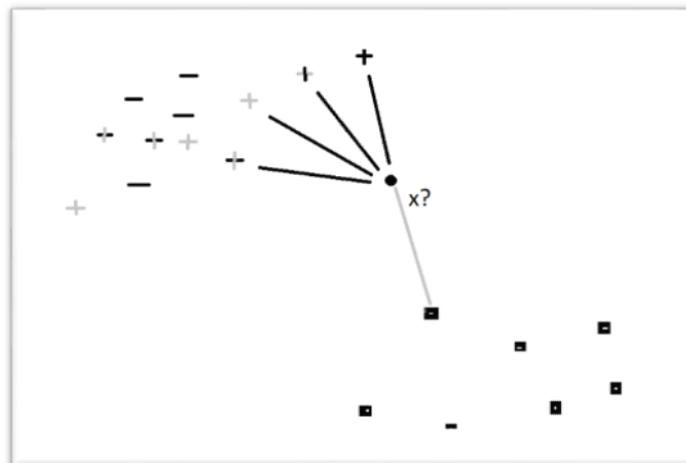
1. Choix d'un entier k (souvent  $k = \text{nombre d'attributs} + 1$ , ou  $k = \sqrt{n}$ ).
2. Calcul des distances (ex : distance euclidienne, c'est la distance la plus populaire).

3. Retenir les  $k$  observations pour lesquelles ces distances sont les plus petites (les  $k$  plus proches voisins d'un cas de).
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.[77]

Dans l'exemple de la figure 3.1. Et pour cinq voisins on classerait  $x$  dans la classe  $w_1$  :



**Figure 3.1.** Méthode de  $K$  plus proches voisins

### 3-5-2- 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.[78]

### 3-5-2-1- Définition de la distance:

Le choix de la distance est primordial au bon fonctionnement de la méthode. Quoique les distances les plus simples permettent d'obtenir des résultats satisfaisants.[76]

Une distance doit avoir quatre propriétés pour tous les vecteurs a, b et c : [79]

- $D(a,b) \geq 0$  (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. On choisit alors la distance donnant la meilleure erreur réelle estimée. Plusieurs types pour les distances:

- Distance Euclidienne.
- Distance “city block” ou ‘Manhattan’ (somme des valeurs absolues).
- Distance de Tchebycheff.
- Distance de Mahalanobis.
- .....

Dans notre cas nous avons utilisé la distance Euclidienne (la distance la plus populaire), sa formule en « d » dimensions est la suivante :

$$D(a,b) = \left( \sum_{k=1}^d (a_k - b_k)^2 \right)^{\frac{1}{2}} \quad (3-1)$$

### 3-5-2-2- Sélection de la classe :

La méthode la plus simple est de rechercher le cas le plus proche et de prendre la même décision. C'est la méthode 1-PPV (1-NN) du plus proche voisin. 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.[76]

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 voisin 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 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 voisin 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.[76]

### **3-6- Conclusion :**

Ce chapitre a été consacré à la présentation de principe et de formule mathématique des *différentes caractéristiques* utilisé. Celle-ci est fondée sur l'utilisation de classifieur K-PPV(K-NN). Notre principe repose sur l'utilisation de trois méthodes d'extraction en utilisant le classifieur K-PPV pour fournir la décision de la classification. Le prochain chapitre sera consacré aux résultats expérimentaux de notre méthode.



# Chapitre 4:

## Résultats expérimentaux

---

## Chapitre 4

---

# Résultats expérimentaux

---

### 4-1- Introduction:

Dans ce chapitre nous avons présenté les résultats obtenus pour la validation du système de reconnaissance des chiffres manuscrite isolé mais tout d'abord on va présenter une description de la base de données utilisée, ensuite plusieurs expériences sont effectuées pour l'étude de classification des images avec les combinaisons des différentes caractéristiques sur les résultats obtenus. Ces résultats permettent de mesurer les performances obtenues lors de l'utilisation des caractéristiques.

### 4-2- Bases de données :

Les expériences ont été conduites sur une base de données composée de chiffres manuscrits isolés sous forme d'images binaires (figure 4.1). Cette base, appelée la base de données MNIST (Mixed National Institute of Standards and Technology) est une grande base de chiffres manuscrits qui est couramment utilisée pour divers systèmes de traitement d'image.[80]

La base de données est également largement utilisée pour l'apprentissage et les tests. Elle a été créée par "re-mélange" les échantillons provenant des ensembles de données d'origine du NIST. En outre, les images en noir et blanc du NIST ont été normalisées et insérées dans une boîte 28x28 pixel englobante.[81]

La base de données MNIST contient 60.000 images d'apprentissage et de 10.000 images de test. **Tableau 4.1.**

| <i>Classes</i> | <i>APPRENTISSAGE</i> | <i>TEST</i> | <i>TOTALE</i> |
|----------------|----------------------|-------------|---------------|
| 0              | 5923                 | 980         | 6903          |
| 1              | 6742                 | 1135        | 7877          |
| 2              | 5958                 | 1032        | 6990          |
| 3              | 6131                 | 1010        | 7141          |
| 4              | 5842                 | 982         | 6824          |
| 5              | 5421                 | 892         | 6313          |
| 6              | 5918                 | 958         | 6876          |
| 7              | 6265                 | 1028        | 7293          |
| 8              | 5851                 | 974         | 6825          |
| 9              | 5949                 | 1009        | 6958          |
| <i>Totale</i>  | 60000                | 10000       | 70000         |

**Tableau 4.1** : Répartition de la base MNIST pour les chiffres.



**Figure 4.1** : Exemples de la base de données MNIST

#### 4-3- Critère d'évaluation:

La matrice de confusion est un critère d'évaluation de la classification. Cette matrice permet d'estimer les taux de bonne classification (ou de bonne reconnaissance) pour chaque modèle et également le taux global du système. Nous avons choisi de mettre les classes de référence en lignes et les résultats de test en colonnes pour chaque classe . Dans ce travail, nous utilisons deux taux : taux de reconnaissance par classe (TRC) (équation 4.1) et le taux moyen de bonne reconnaissance (TMBR) qui est défini comme étant le nombre de chiffres reconnus sur le nombre totale du chiffres exprimé en % (équation 4.2).

$$TRC (\%) = \frac{\text{nombre de chiffres reconnus d'une classe}}{\text{nombre de chiffres de la classe}} \quad (4.1)$$

$$TMBR (\%) = \frac{\text{nombre de chiffres reconnus}}{\text{nombre totale de chiffres}} \quad (4.2)$$

#### 4-4- Protocole expérimental :

Les expérimentations sont réalisées avec la base de donnée MNIST en respectant le même protocole de l'évaluation proposer par créateur de cette base .

Dans ce travail , nous avons choisis quelques méthodes d'extraction de caractéristique comme moments Zernike , histogramme de HOG et filtre de GABOR en utilisant le classifieur KPPV avec K=1.

##### 4-4-1- Les expériences:

**Expérience 1** :Nous avons utilisés les moments zernike avec deux paramètres différents, le moments Zernike avec un ordre de 20 appliqué sur l'image entière, et l'autre de l'ordre 8, appliqué sur des différentes régions de l'image fragmentée en 2x2 (mz 8 2/2).

**Expérience 2** : Nous avons utilisés l'histogramme de HOG sur l'image entière .

**Expérience 3**:Nous avons utilisés le filtre de GABOR avec les paramètres( l'échelle = 5, nombre d'orientation =8, filtre de la taille 37x37 ) sur l'image entier

**Expérience 4** :Nous avons concaténés les vecteurs de caractéristiques obtenu par les expériences 1,2et3.

**Expérience 5**:Nous avons utiliser la même expérience 4 en changeant le K=1 par K=5 de KPPV.

Le tableau suivant présente les résultats obtenus après l'évaluation automatique de chacune expérience sur la base MNIST, ce tableau contient le Taux de Reconnaissance (TR) et la Taille de Vecteur (TV).

|       | <b>Expérience 01</b> | <b>Expérience 02</b> | <b>Expérience 03</b> | <b>Expérience 04</b> | <b>Expérience 05</b> |
|-------|----------------------|----------------------|----------------------|----------------------|----------------------|
| TR(%) | <b>95.5</b>          | <b>92.98</b>         | <b>97.27</b>         | <b>97.53</b>         | <b>97.86</b>         |
| TV    | <b>765</b>           | <b>144</b>           | <b>1960</b>          | <b>2869</b>          | <b>2869</b>          |

**Tableau 4.2** :Résultats obtenus de chacune expérience

La bonne expérience est effectuée avec la concaténation des vecteurs de caractéristiques obtenu par les expériences 1,2et3 en utilisant un classifieur KPPV de K=5 ;

D'une manière générale ; le vecteur de caractéristiques final est composé de **2869** composantes.

Cette évaluation est illustrée dans la matrice de confusion reportée dans le Tableau 4.2

|                   |   | CLASSES DE REFERENCE |         |         |         |         |         |         |         |         |         |
|-------------------|---|----------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|                   |   | 0                    | 1       | 2       | 3       | 4       | 5       | 6       | 7       | 8       | 9       |
| CLASSES CALCULEES | 0 | 99.4898              | 0.1020  | 0.6122  | 0       | 0.2041  | 0.3061  | 0.4082  | 0.1020  | 0.4082  | 0.2041  |
|                   | 1 | 0.0881               | 99.5595 | 0.0881  | 0       | 0       | 0       | 0.2643  | 0.3524  | 0       | 0.4405  |
|                   | 2 | 0                    | 0.1938  | 97.9651 | 0.4845  | 0       | 0       | 0       | 0.5814  | 0.0969  | 0       |
|                   | 3 | 0                    | 0       | 0       | 96.5347 | 0       | 1.0891  | 0       | 0.2970  | 0.3960  | 0.2970  |
|                   | 4 | 0                    | 0       | 0       | 0       | 95.8248 | 0.1018  | 0       | 0       | 0.4073  | 0.3055  |
|                   | 5 | 0                    | 0       | 0       | 0.8969  | 0       | 97.0852 | 0.1121  | 0       | 0.2242  | 0.4484  |
|                   | 6 | 0.1044               | 0.1044  | 0.2088  | 0       | 0.5219  | 0.4175  | 99.0605 | 0       | 0.2088  | 0       |
|                   | 7 | 0.1946               | 0       | 0.9728  | 0.5837  | 0       | 0       | 0       | 97.8599 | 0.1946  | 0.3891  |
|                   | 8 | 0.1027               | 0.1027  | 0.2053  | 0.9240  | 0       | 0.4107  | 0.1027  | 0.1027  | 97.5359 | 0.5133  |
|                   | 9 | 0                    | 0       | 0       | 0.6938  | 3.3697  | 0.2973  | 0       | 0.6938  | 0.4955  | 97.4232 |
| TMBR              |   |                      |         |         |         |         |         |         |         |         | 97.8600 |

**Tableau 4.3 :** matrice de confusion de l'expérience 5.

On remarque que TMBR est important relativement dans cette expérience. Ce fait est attribuable à une confusion entre certains chiffres particuliers, puisqu'ils est vraiment difficile de faire distinguer entre elles par exemple ( 4 et 9, 5 et 3 et 7 et 9 ...). Sans oublier que le taux d'erreur dans cette expérience est de (2.14%) .

#### 4-5- Conclusion :

Ce chapitre a été consacré à l'évaluation des méthodes d'extraction des caractéristiques utilisés en utilisant classifieur KPPV pour la reconnaissance des chiffres manuscrits. Les résultats obtenus montrent que l'expérience 05 qui est donné le meilleur résultat par rapport les autres expériences, et le classifieur joue un rôle important pour sélectionnée le meilleur résultat.



# **Conclusion Générale**

---

---

## CONCLUSION GENERALE

---

---

L'objectif de ce travail est la mise en œuvre et l'évaluation de plusieurs méthodes d'extraction de caractéristiques pour la reconnaissance des chiffres manuscrits non normalisés en utilisant le classifieur KPPV.

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

Dans le domaine de la reconnaissance de chiffres manuscrits, les caractéristiques peuvent être décrites comme un moyen permettant de distinguer un caractère d'une classe d'un autre caractère d'une autre classe. Dès lors, il est nécessaire de définir des caractéristiques significatives lors du développement d'un système de reconnaissance. Nous avons présentés plusieurs méthodes d'extraction des caractéristiques pertinentes. Soit les moments Zernike, l'histogramme HOG et filtre de Gabor. L'évaluation de ces caractéristiques a été faite par le classifieur KPPV.

Notre principale contribution était d'achève a une bonne combinaison de plusieurs méthodes d'extraction des caractéristiques. Nous avons constaté que chaque méthode présente des points forts et des points faibles. Afin de résoudre ces problèmes, nous avons proposé de combiner ces méthodes pour améliorer les taux de reconnaissance. Les résultats obtenus montrent que la combinaison des méthodes d'extraction des caractéristiques permet d'améliorer efficacement le taux de bonne reconnaissance.

Avant d'aborder les perspectives nouvelles, ce système est considéré comme une étape préliminaire pour réaliser un système de reconnaissance des chiffres isolés puissant et performant. Au-delà un certain nombre d'amélioration peut être envisagé :

- Améliorer le system par utilisation le module de l'extraction de caractéristiques.
- Essayer d'agrandir le nombre des échantillons de bases de données pour tester l'efficacité de l'approche proposée.
- Améliorer le system par combinaison des plusieurs classifieurs.
- Tester d'autres descripteurs multidirectionnels pour une meilleure caractérisation des chiffres.



# **Références Bibliographiques**

---

## Référence bibliographique

---

- [01]. S.Nebti, «Reconnaissance de Caractères Manuscrits par Intelligence Collective», Thèse de doctorat, Université Ferhat Abass -Sétif, mars 2013.
- [02]. D.Abdelhakim, 2011, «La reconnaissance des chiffres manuscrits par les machines à vecteurs de support(SVMs)»; Thèse de Master, Université de Tébessa
- [03]. G.Abdeldjalil, 2011, «Segmentation automatique pour la reconnaissance numérique des chèques bancaires Algériens » ; Thèse de MAGISTER, centre Universitaire de Khanchela
- [04]. 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.
- [05]. A. El-Yacoubi, R. Sabourin, M. Gilloux, and C.Y. Suen. Improved model architecture and training phase in an off-line HMM-based word recognition system. In Proc. of the 13th International Conference on Pattern Recognition, Brisbane, Australia, 1998, pages 1521–1525.
- [06]. C. Touzet,. « Les réseaux de neurones artificiels : introduction au connexionnisme»,150 pages, Edition EC2, Paris. 1992,pp.6-35
- [07]. Jean-Pierre « reconnaissance de l'écriture manuscrite », Département Images, ENSTParis et Guy LORETTE RISA, CNRS UPRES-A 6074, Université de Rennes 1.
- [08]. P.M Lallican, C. Viarp-Gaudin and S. Knerr,. « From off-line to on-line handwriting recognition ». Proc. 7th workshop on frontiers in handwriting recognition, Amsterdam, 2000 ,pp. 303-312
- [09]. Frédéric Grandidier doctorat en génie ph.d. « un nouvel algorithme de sélection de caractéristiques – application à la lecture automatique de l'écriture manuscrite » montréal, le 24 janvier 2003
- [10]. C. Chatelain. « Extraction de séquences numériques dans des documents manuscrits quelconques », Thèse de doctorat, Université de Rouen, France. , 2006 p. 196.
- [11]. F. Menasri, «Contributions `a la reconnaissance de l'écriture arabe manuscrite», Thèse de doctorat, université´ paris des cartes, juin 2008.

- [12]. Benamara, 1999, « Utilisation des modèles de Markov cachés planaires en reconnaissance de l'écriture arabe imprimée » ; Thèse de doctorat, Université des sciences, des Techniques et de médecine de Tunis II, pp. 05-100
- [13]. Dargenton, «Contribution à la segmentation et la reconnaissance de l'écriture manuscrite» ; Thèse de Doctorat. Institut National des Sciences Appliquées de Lyon, 174 pages. 1994
- [14]. Oliveira, 2003, «Automatic recognition of handwritten numerical strings»; Thèse de Doctorat. Ecole de technologie supérieure, Université de Québec. Canada, 184 pages.
- [15]. P.R. Cavalin, A.Britto, F.Bortolozzi, R. Sabourin & L.Oliveira. An implicit segmentation-based method for recognition of handwritten strings of characters. ACM symposium on Applied computing, pages 836–840, 2006.
- [16]. A. Britto, R. Sabourin, F. Bortolozzi & C. Suen. Recognition of handwritten numeral strings using a two-stage HMM based method. IJDAR, vol. 5, pages 102–117, 2003.
- [17]. A. Britto, R. Sabourin, E. Lethelier, F. Bortolozzi & C. Suen. Improvement Handwritten numeral string recognition by slant normalization and contextual information. IWFHR, pages 323–332, 2000.
- [18]. Q. Xu, L. Lam & C.Y. Suen. Automatic segmentation and recognition system for handwritten dates on Canadian bank cheques. ICDAR, pages 704–708, 2003.
- [19]. Y. Lei, C. S. Liu, X.Q. Ding & Q. Fu. A Recognition Based System for Segmentation of Touching Handwritten Numeral Strings. IWFHR, pages 294–299, 2004.
- [20]. L.S. Oliveira, R. Sabourin, F. Botolozzi & C.Y. Suen. Automatic Recognition of Handwritten Numeral Strings : A Recognition and Verification Strategy. IEEE Trans. On PAMI, vol. 24, pages 1438–1454, 2002.
- [21]. U. Pal, A. Belaïd & C. Choisy. Water Reservoir Based Approach for Touching Numeral Segmentation. ICDAR, pages 892–897, 2001.
- [22]. M.Morita, R. Sabourin, F. Bortolozzi & C.Y. Suen. Segmentation and Recognition of Handwritten Dates : An HMMML Hybrid approach. IJDAR, to appear, 2006.
- [23]. K.K. Kim, J.H. Kim & C. Y. Suen. Segmentation-based recognition of handwritten touching pairs of digits using structural features. Pattern Recognition Letters, vol. 23, no. 1, pages 13–24, 2002.

- [24]. Z. Lu, Z. Chi, W.C. Siu & P. Shi. A background-thinning based approach for separating and recognizing connected handwritten digit strings. *Pattern Recognition*, vol. 32, pages 921–933, 1999.
- [25]. J. Sadri, C. Y. Suen & T. D. Bui. Automatic Segmentation of Unconstrained Handwritten Numeral Strings. *IWFHR*, pages 317–322, 2004.
- [26]. M. Koga, R. Mine, H. Sako & H. Fujisawa. A recognition method of machine-printed monetary amounts based on the two-dimensional segmentation and the bottom-up parsing. *ICDAR*, pages 968–971, 2001.
- [27]. Y. Chen & J.F. Wang. Segmentation of Single- or Multiple-Touching Handwritten Numeral String Using Background and Foreground Analysis. *IEEE Trans. on PAMI*, vol. 22, no. 11, pages 1304–1317, 2000.
- [28]. C.L. Liu & H. Sako. Class-specific feature polynomial classifier for pattern classification and its application to handwritten numeral recognition. *Pattern Recognition*, vol. 39, no. 4, pages 669–681, 2006.
- [29]. L. Heutte, P. Barbosa-Perreira, O. Bougeois, J.V. Moreau, B. Plessis, P. Courtellemont & Y. Lecourtier. Multi-bankcheck recognition system : consideration on the numeral amount recognition module. *IJPRAI*, vol. 11, pages 595–618, 1997.
- [30]. C.L. Liu, H. Sako & H. Fujisawa. Effects of Classifier Structures and Training Regimes on Integrated Segmentation and Recognition of Handwritten Numeral Strings. *IEEE Trans. on PAMI*, vol. 26, pages 1395–1407, 2004.
- [31]. L.Q. Zhang & C.Y. Suen. Recognition of courtesy amount on bank checks based on a segmentation approach. *FHR02*, pages 298–302, 2002.
- [32]. R. Palacios & A. Gupta. A system for processing handwritten bank checks automatically. *rapport interne*. 1997.
- [33]. H. Fujisawa, Y. Nakano & K. Kurino. Segmentation Methods for Character Recognition : From Segmentation to Document Structure Analysis. *Proc. of the IEEE*, vol. 80, no. 7, pages 1079–1092, 1992.
- [34]. T.M. Ha, M. Zimmermann & H. Bunke. Off-line Handwritten Numeral String Recognition by Combining Segmentation-based and Segmentation-free Methods. *Pattern Recognition*, vol. 31, no. 3, pages 257–272, 1998.
- [35]. Clément Chatelain. Extraction de séquences numériques dans des documents manuscrits quelconques. *Signal and Image processing*. Université de Rouen, 2006. French.

- [36]. Fenrich R, Krishnamoorthy S (1990) Segmenting diverse quality handwritten digit strings in near real-time. In Proceedings of 5th USPS Advanced Technology Conference, pp 523–537.
- [37]. Fujisawa H, Nakano Y, Kurino K (1996) Segmentation Methods for Character Recognition: From Segmentation to Document Structure Analysis. Proceedings of the IEEE, vol 80, no 7, pp 1079-1091, pp 21- 28.
- [38]. Shi Z, Govindaraju, V (1997) Segmentation and recognition of connected handwritten numeral strings. Pattern Recognit.vol 30, no 9, pp 1501–1504.
- [39]. Congedo G, Dimauro G, Impedovo S, Pirlo G (1995) Segmentation of Numeric Strings, Proc. of Third Int. Conf. on Document Analysis and Recognition. IEEE Computer Society, Montreal, Canada, pp.1038-1041.
- [40]. Pal U, Belaid, A., Choisy, C (2003) Touching numeral segmentation using water reservoir concept. Pattern Recognit. Lett. 24, pp 261–272.
- [41]. Chen Y K, Wang J F (2000) Segmentation of Single or Multiple-Touching Handwritten Numeral String Using Background and Foreground Analysis. IEEE Transactions on Pattern Analysis Machine Intelligence, vol 22, no 11, pp 1304-1317.
- [42]. Oliveira L S, Lethelier E, Bortolozzi F, Sabourin R (2000) A new approach to segment handwritten digits, Proc. 7th IWFHR, pp 577-582.
- [43]. Oliveira L S, Sabourin R, Bortolozzi F, Suen C Y (2002) Automatic Recognition of Handwritten Numerical Strings: A Recognition and Verification Strategy. IEEE Transactions on Pattern Analysis and Machine Intelligence - PAMI, vol 24, no 11, pp 1438-1454.
- [44]. Vellasques E, Oliveira L S, Britto Jr A S, Koerich A L, Sabourin R (2008) Filtering segmentation cuts for digit string recognition. Pattern Recognition, vol 41, no 10, pp 3044-3053.
- [45]. Suwa M (2005) Segmentation of connected handwritten numerals by graph representation. Document Analysis and Recognition, 2005.Proceedings. Eighth International Conference on Document Analysis and Recognition (ICDAR'05), vol. 2, pp 750-754.
- [46]. Elnagar A, Alhadj R (2003) Segmentation of connected handwritten numeral strings. Pattern Recognit.vol 36, no 3, pp 625–634.
- [47]. B. L. Everton, A. B. M. Carlos, Segmentation of connected handwritten digits using Self-Organizing Maps, Expert systems with applications 40(2013) 5867-5877.

- [48]. A. Britto- Jr., R. Sabourin, F. Bortolozzi, and C. Y. Suen. A string length predictor to control the level building of HMMs for handwritten numeral recognition. In Proc. of International Conference on Pattern Recognition (ICPR), volume 4, pages 31-34, Quebec City, Canada, 2002. IEEE Computer Society.
- [49]. Y. Guermur, A. Eliseeff, H. Paugam Moisy. A new multiclass svm based on a uniform convergence result, IJCNN, vol. IV, pages 183–188, 2000.
- [50]. J. Weston, C. Watkins, Multi-class support vector machines for multi-class pattern recognition. ESANN, pages 219–224, 1999.
- [51]. C-L. Huang, C-J. Wang, A GA-based feature selection and parameters optimization for support vector machine, Expert systems with application, vol. 31, pages 231-240, 2006.
- [52]. I. Guyon, J. Weston, S. Barnhill, V. Vapnik, Gene Selection for Cancer Classification using Support Vector Machines, Machine Learning, vol. 46, no. 3, pages 389-422, 2002.
- [53] J. Weston, S. Mukherjee, O. Chapelle, M. Pontil, T. Poggio, V. Vapnik, Feature Selection for SVMs, NIPS, pages 668–674, 2000.
- [54] L. Hermes, J.M. Buhmann. Feature Selection for Support Vector Machines. ICPR, vol. 2, pages 2712-2715, 2000.
- [55] J. Weston & C. Watkins. Multi-class support vector machines for multi-class pattern recognition. ESANN, pages 219–224, 1999.
- [56]. Ratzer, Vangheluwe, 2006, « The Implementation of Support Vector Machines Using the Sequential Minimal Optimization Algorithm»; Université des Sciences de, McGill, Canada.
- [57]. F. Denis and R. Gilleron, Apprentissage à partir d'exemples, Technical report, Grappa - Université de Lille 3, 1999.
- [58]. A. Nedjem Eddine, «sélection de modèle automatique des machines à vecteurs de support: application à la reconnaissance d'images de chiffres manuscrits», thèse de doctorat, université du québec, janvier 2004
- [59]. Burges, C. J. C., 1998. A Tutorial on support vector machines for pattern recognition, Data Mining and Knowledge Discovery (Edited by Ussama Fayyad), Vol. 2, 121-167.
- [60]. 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

- [61]. M. Milgram, Reconnaissance des formes méthodes numériques et connexionnistes, Edition Armand colin, Paris, 1993.
- [62]. G. Saporta, Probabilités, Analyse des données et Statistique, Édition Technip, Paris, France, 1990.
- [63]
- [64] Ernest Kussul, Tatyana Baidyk : Improved Method of Handwritten Digit Recognition, Glushkov Institute of Cybernetics, Kiev, Ukraine, 2004.
- [65] Ming Wu Zhen Zhang, Handwritten Digit Classification using the MNIST Data Set, ResearchGate publications, 2010.
- [66]
- [67] Yukihiro Yamashita ; Toru Wakahara; k-NN Classification of Handwritten Characters Using a New Distortion-Tolerant Matching Measure, ., Pattern Recognition (ICPR), 2014, 2nd International Conference on; Stockholm.
- [68] Lee et al, Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree, NSF awards, 2015.
- [69] Yukihiro Yamashita, Toru Wakahara: Affine-transformation and 2D-projection invariant  $k$ -NN classification of handwritten characters via a new matching measure, *Pattern Recognition, Volume 52, 2016, PP 459-470*.
- [70] Dundar et al, (2016).
- [71] Dalal, N. and B. Triggs. "Histograms of Oriented Gradients for Human Detection", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1 (June 2005), pp. 886–893.
- [72] F. ZERNIKE : Diffraction theory of the cut procedure and its improved form, the phasecontrast method. *Physica*, 1:689–704, 1934.
- [73] R. MUKUNDAN et K.R. RAMAKRISHNAN : Fast computation of legendre and zernikemoments. *PR*, 28(9):1433–1442, septembre 1995.
- [74] S.O. BELKASIM, M. AHMADI et M. SHRIDHAR : Efficient algorithm for fast computation of zernike moments. *Journal of the Franklin Institute*, 333:577–581, juillet 1996.
- [75] C.W. CHONG, P. RAVEENDRAN et R. MUKUNDAN : A comparative analysis of algorithms for fast computation of zernike moments. *Pattern Recognition*, 36(3):731–742, mars 2003.
- [76] <http://www.grappa.univ-lille3.fr/polys/fouille/sortie005.html#toc13>
- [77] [docs.happycoders.org/orgadoc/artificial\\_intelligence/classification\\_documents/](https://docs.happycoders.org/orgadoc/artificial_intelligence/classification_documents/)

classification.pdf

[78] M. Milgram, « *Reconnaissance des formes méthodes numériques et connexionnistes* », Edition Armand colin, Paris, 1993.

[79] R. O. Duda, P. E. Hart, D. G. Stork, « *Pattern classification* », Second edition, 1992.

[80] L. Deng, « The MNIST Database of Handwritten Digit Images for Machine Learning Research (Best of the Web) », *IEEE Signal Processing Magazine*, vol. 29, n° 6, novembre 2012, p. 141-142

[81] Platt, John C.. "Using analytic QP and sparseness to speed training of support vector machines". *Advances in neural information processing systems*,(1999), PP 557–563.