



Université de Ghardaïa

N° d'ordre :  
N° de série :

Faculté des Sciences et Technologies  
Département des Sciences et Technologie

Mémoire présenté en vue de l'obtention du diplôme de

## MASTER

**Domaine :** *Sciences et Technologies*

**Filière :** Automatique

**Spécialité :** *Automatique et système*

**Par :** REZMA Ismail

**ZEHOUANI Abdelmadjid**

## Thème

**Proposition d'un système d'identification des chiffres numérique, application sur la reconnaissance des codes postaux**

**Soutenu publiquement le 20/06/2018**

**Devant le jury :**

<b>Fihakhir Mahdi amine</b>	M.A.A	Univ. Ghardaïa	<b>Président</b>
<b>Benaouicha Karim</b>	M.A.A	Univ. Ghardaïa	<b>Examinateur</b>
<b>Laadjal Boumedian</b>	M.A.A	Univ. Ghardaïa	<b>Examinateur</b>
<b>Mohamed ARIF</b>	M.A.A	Univ. Ghardaïa	<b>Encadreur</b>
<b>BECHOUAT Mohcen</b>	M.C.B	Univ. Ghardaïa	<b>Co-Encadreur</b>

**Année universitaire 2017/2018**

# Remerciement

❖ Nous d'abord remerciant et a louanges à dieu tout  
puissant

❖ Nous tenons tout d'abord a remercier **Dr,**  
**Mohammed ARIF, Dr mohcin bechouate** notre  
encadreur pour sons soutien tout au long de cette  
étude et ses conseils et suggestion.

❖ Sans oublier mes amis et toute personne qui besoin a  
ce travail

# DEDICACE

Avant tous, je remercie dieu le tout puissant de m'avoir donné le courage et la patience pour réaliser ce travail malgré toutes les difficultés rencontrées.

Je dédie ce modeste travail :

A mes très chers parents, que dieu les garde et les protège pour leurs soutien moral et financier, pour leurs encouragements et les sacrifices qu'ils ont endurés.

A mes frères

A mes sœurs

A mes grandes familles

Aux chers amis

A tous ce que j'aime et qui m'aiment  
A tous mes collègues d'études surtout  
ceux D'automatique

Abdelmadjid

# Dédicace

Je dédie ce modeste travail

A ma mère avec toute mon affection.

A mon père avec toute ma reconnaissance.

A et ma grand-mère que dieu les gardes.

A mes frères et mes sœurs.

A tous mes amis

*Smail*

## **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: prétraitements 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 SVM et coefficient de corrélation sont fournis après des tests un taux de

**Mots-clés :** Un système de reconnaissance des chiffres manuscrits, prétraitements des images, la classification, SVM, support vector machine, corrélation

## **Abstract**

A system for recognizing isolated manuscript digits is based on the steps following pretreatments, feature extraction and classification, followed possibly a post-treatment phase.

In this work, we are interested in two modules: preprocessing and classification. Characteristics can be described as a means of distinguishing a character of a class of another character of another class. So, it is necessary to define significant characteristics during development.

In our case, we focus on the combination of SVM and correlation coefficient are provided after testing a rate of

**Keywords:** A system for recognizing handwritten figures, image preprocessing, classification, SVM, victor machine support, correlation

## ملخص

يستند نظام التعرف على أرقام مخطوطة معزولة على الخطوات بعد العلاجات ، واستخراج ميزة والتصنيف ، تليها ربما مرحلة ما بعد العلاج. في هذا العمل ، نحن مهتمون بوحدين: المعالجة المسبقة والتصنيف. يمكن وصف الخصائص كوسيلة للتمييز شخصية لصف من شخص آخر من فئة أخرى. لذا ، من الضروري تحديد خصائص مهمة أثناء التطوير.

في حالتنا ، نركز على الجمع بين SVM ومعاملة الارتباط يتم توفيرها بعد اختبار معدل

**كلمات البحث:** نظام التعرف على الأرقام المكتوبة بخط اليد ، معالجة الصور المسبقة ، التصنيف ، SVM ، دعم آلة المنتصر ، الارتباط

---

## TABLE DE MATIERES

**Remerciement**

**Résumé**

**Liste des tableaux & figures**

### TABLE DE MATIERES

**Introduction générale** ..... 1

### CHAPITRE I : l'état d'art de reconnaissance des chiffres

I.1 Introduction sur le reconnaissance manuscrit ..... 3

I.2 Type de reconnaissance manuscrit ..... 3

    I.2.1 Reconnaissance de caractères ..... 3

    I.2.2 Reconnaissance de mots ..... 4

    I.2.3 Reconnaissance de documents ..... 4

    I.2.4 Ecriture arabe ..... 5

I.3 Historique de reconnaissance des chiffres ..... 5

    I.3.1 Mayraz et Hinton. (2002) ..... 5

    I.3.2 2- Kussul et Baidyk (2004) ..... 6

    I.3.3 3- Wu et Zhang (2010) ..... 7

    I.3.4 4- Ebrahim et Jampour (2014) ..... 7

    I.3.5 5- Wakahara et Yamashita (2014) ..... 7

    I.3.6 6- Lee et al (2015) ..... 8

    I.3.7 7- Wakahara et Yamashita (2016) ..... 8

    I.3.8 8- Dundar et al (2016) ..... 9

I.4 conclusion ..... 9

### CHAPITRE II : prétraitement de de reconnaissance des chiffres

    II.1 Introduction ..... 11

    II.2 Etapes de prétraitement ..... 11

        II.2.1 Filtrage ..... 11

            II.2.1.1 Elimination de bruit (filtrage) ..... 11

            II.2.1.2 Types de filtres ..... 11



II.2.1.2.1	Les filtres linéaires .....	11
II.2.1.2.2	Les filtres non linéaires .....	12
II.2.1.3	Les filtres utilisés .....	12
II.2.1.3.1	Filtre passe bas .....	12
II.2.1.3.2	Filtre passe haut .....	13
II.2.1.3.3	Filtre Moyenner .....	13
II.2.1.3.4	Filter Médian.....	14
II.2.2	Binarisation .....	15
II.2.2.1	Seuillage global .....	15
II.2.2.2	Seuillage adaptatif .....	16
II.2.3	La squelettisation .....	16
II.2.4	Lissage .....	17
II.2.5	La normalisation .....	18
II.2.6	Segmentation .....	18
II.3	CONCLUSION.....	18

### **CHAPITRE III : METHODE SVM**

III.1	Introduction .....	20
III.2	Principe des machines à vecteurs de support .....	20
III.2.1.	Notions de base .....	20
1.1	Hyperplan .....	20
1.2	Vecteurs de support .....	21
1.3	Marge .....	22
III.2.2.	Propriétés fondamentales .....	22
III.3	Fondement mathématiques.....	24
III.3.1.	Cas de données linéairement séparables .....	24
III.3.2.	Cas des données non – linéairement séparables.....	26
III.4	Architecture d'un classificateur SVM .....	27
III.4.1.	la fonction noyau.....	27
III.4.2.	Sélection de modèle SVM.....	28
III.5	Estimation de l'erreur de généralisation.....	28
III.6	Détection par SVM.....	29
III.7	Le coefficient de corrélation.....	30
III.8	Conclusion .....	31

**CHAPITRE IV: Test & Résultat**

IV.1	Introduction .....	33
IV.2	Description notre travail.....	33
IV.2.1	Base de donné .....	33
IV.2.2	<i>Critère d'évaluation</i> .....	33
IV.2.3	Résultat & discussion .....	33
IV.3	Conclusion.....	39
	<b>Conclusion générale .....</b>	<b>41</b>
	<b>Bibliographie .....</b>	<b>43</b>

---

## Liste des figures :

<b>Figure.1.1</b> types de communication Homme machine.	2
<b>Figure II .1</b> : Filtre passe-bas	12
<b>Figure II.2</b> : Filtre passe-haut	13
<b>Figure II.3</b> : filtre moyenné	14
<b>Figure II.4:</b> Minéen œuvre du filtre Médian	14
<b>Figure II .5:</b> image filtrée	15
<b>Figure II.6:</b> image binaire	16
<b>Figure II.7</b> : image squelettisé	17
<b>Figure III 1</b> : l'hyperplan H qui sépare les deux ensembles de points.	21
<b>Figure III 2</b> : Les vecteurs de support.	21
<b>Figure III 3:</b> hyperplan optimal, vecteurs de support et marge maximale.	22
<b>Figure III 4</b> : meilleur hyperplan séparateur.	23
<b>Figure III 5</b> cas linéarité et non linéarité spérable	23
<b>Figure III 6:</b> cas non linéaire de SVM	24
<b>Figure III 7</b> séparateur à vaste marge.	26
<b>Figure III 8</b> : Architecture d'une machine à vecteurs de support (d'un nombre N)	27
<b>Figure III 9</b> : Processus d'extraction de fenêtres à la phase de détection.	30

## Liste des tableaux

- Tableau 4.1** : résultat pour 1 image d'apprentissage et 3 images de test pour chaque code pour chaque personne. 34
- Tableau 4.2** : résultat pour 2 images d'apprentissage et 3 images de test pour chaque code pour chaque personne 35
- Tableau 4.2** : résultat pour 3 images d'apprentissage et 3 images de test pour chaque code pour chaque personne. 37

# **INTRODUCTION GÉNÉRALE**

## **Introduction général**

La reconnaissance hors ligne de l'écriture manuscrite présente un intérêt l'indéniable dans l'accomplissement des tâches fastidieuses comme celles que l'on rencontre dans certains domaines : le tri postal, la lecture de chèques bancaires, la lecture des bordereaux... Les premières études sur la reconnaissance de l'écriture manuscrite, antérieures aux années 1980, tentaient de généraliser les résultats obtenus pour l'écriture imprimée

La reconnaissance automatique de l'écriture est un domaine de recherche qui a trouvé une application à grande échelle dans le tri du courrier. En effet, une grande partie du courrier est actuellement triée de manière entièrement automatique. Les enveloppes passent devant une caméra, et chaque image est traitée automatiquement par une machine qui localise le code postal et le reconnaît.

La reconnaissance des chiffres du code postal est un problème difficile, surtout lorsqu'il s'agit d'écriture manuscrite. Les quelques images ci-dessous [Anexxe] montrent des exemples de chiffres écrits à la main. Chaque ligne de chiffres est écrite par une personne différente. On notera la grande variabilité de l'écriture, lorsque l'on passe d'une personne à l'autre.

Dans le cadre de cette étude nous nous sommes intéressés à la reconnaissance des chiffres des codes postaux basée plus précisément sur le coefficient de corrélation. Dans le premier chapitre on donne l'état de l'art sur la reconnaissance de l'écriture manuscrite.

Dans le deuxième chapitre on cite les différentes étapes d'analyse, et les différentes fonctions de prétraitement d'image.

Dans le chapitre trois, nous exposons une étude de SVM et corrélation dans le but d'améliorer le taux de reconnaissance

Dans le chapitre quatre est sacrée pour nos résultats et la discussion de de différents paramètres des tests.

# **CHAPITRE I**

**l'état d'art de  
reconnaissance des  
chiffres**

## I. Introduction sur la reconnaissance manuscrite

L'écriture manuscrite est un moyen naturel de communication qui présente l'avantage d'être familier à la majorité des gens. De ce fait elle présente un moyen d'interaction facile avec l'ordinateur. L'un des objectifs les plus recherchés est de doter les ordinateurs de capacités de l'être Humain (Figure 1.1).

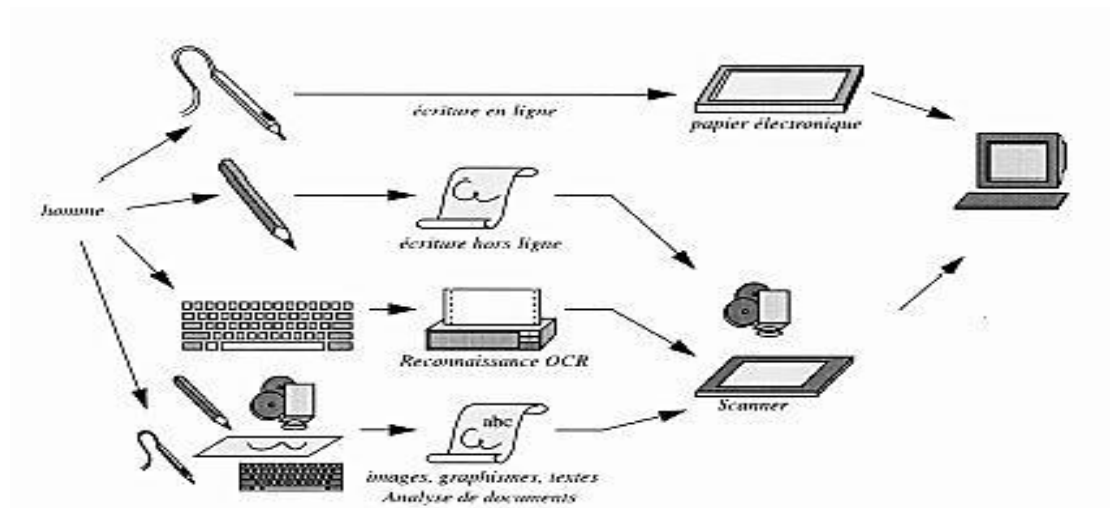


Figure.1.1 types de communication Homme machine.

Donc le but de la reconnaissance de l'écriture est de transformer un texte écrit en une représentation compréhensible par une machine et facilement reproductible par un traitement de texte. Cette tâche n'est pas triviale, car l'écriture possède une infinité de représentations dues au fait que chaque personne possède une écriture qui lui est propre avec de nombreux styles.

Ce qui rend la tâche difficile pour la reconnaissance, c'est la variabilité du type de l'écriture (manuscrite, cursive, ou imprimée).[1]

## II. Type de reconnaissance manuscrite

### II.1. Reconnaissance de caractères

La reconnaissance de caractères concerne la reconnaissance de caractères numériques ou Alphabétiques. Les caractères sont donnés par des images étiquetées, par exemple une image du caractère '5' aura une étiquette 5.

C'est la tâche la plus basique d'un système de reconnaissance de l'écriture. L'effort l'analyse est concentrée sur un seul élément à la fois du vocabulaire (vue comme une forme globale). Ces études portent sur la reconnaissance de caractères manuscrits provenant de la



segmentation de chiffres comme par exemple le montant numérique d'un chèque ou le code postal d'une adresse ou de la segmentation en lettres minuscules et/ou majuscules dans le cas par exemple du montant littéral d'un chèque ou du nom de la ville figurant sur une adresse. Ces caractères isolés présentent de fortes variations principalement provoquées par la position de la lettre dans le mot.

Deux types de reconnaissance de caractères numériques manuscrits : la reconnaissance des chiffres isolés : peut être considéré comme un problème de classification des images des caractères isolés dans un ensemble d'alphabet donné. La reconnaissance des chaînes numériques manuscrites : il est à noter que ce type des systèmes commence généralement par une phase de segmentation qui consiste à séparer la chaîne numérique à des entités (chiffres) isolées et reconnaître ces dernières par un module de reconnaissance des chiffres isolés[2]

### **II.2. Reconnaissance de mots**

Dans la présente situation, on s'intéresse à la reconnaissance du mot entier. Deux types de reconnaissance peuvent être considérés, la reconnaissance globale où le processus de la reconnaissance est ramené à un processus de reconnaissance de caractères avec des formes plus complexes que les caractères. Le deuxième type est une reconnaissance analytique où le mot est divisé en graphèmes. Elle cherche à identifier les caractères ou les sous-caractères (graphèmes) issus de la segmentation (séparation de mots, des caractères) pour reconstituer les mots.

La difficulté d'une telle approche a été clairement évoquée par Sayre en 1973 et peut être résumée par le dilemme suivant : "pour reconnaître les lettres, il faut segmenter le tracé et pour segmenter le tracé, il faut reconnaître les lettres" [2]. Cette approche est la seule applicable dans le cas de grands vocabulaires. Elle peut s'adapter facilement à un changement de vocabulaire. Elle permet théoriquement une discrimination plus fine des mots car elle se base sur la reconnaissance des lettres qui la composent et il est possible de récupérer l'orthographe du mot reconnu. Son inconvénient principal demeure la nécessité de l'étape de segmentation avec les problèmes de sous- ou de sur-segmentation que cela implique.

Certaines des approches actuelles se proposent de tirer avantage des deux méthodes, réduisant la complexité de l'approche globale en l'appliquant sur des entités plus petites (lettres). L'approche analytique recherche la séquence de lettres contenues dans l'image à reconnaître. Certains modèles permettent de combiner ces deux niveaux en un seul et peuvent ainsi s'affranchir de la segmentation préalable de l'image [3]

### **II.3. Reconnaissance de documents**

Cette approche a une vision générale du mot; elle se base sur une description unique de l'image du mot, vue comme une entité indivisible. Disposant de beaucoup d'informations, elle absorbe plus facilement les variations au niveau de l'écriture.

Cette approche présente l'avantage de garder le caractère dans son contexte avoisinant, ce qui permet une modélisation plus efficace des variations de l'écriture et des dégradations qu'elle peut subir. Cependant cette méthode est pénalisante par la taille mémoire, le temps de calcul et la complexité du traitement qui croient linéairement avec la taille du lexique considéré, d'où une limitation du vocabulaire souvent appliquée pour réduire la liste de mots candidats dans le contexte d'une reconnaissance à grands vocabulaires (mais limitée dynamiquement : pouvant être réduit de façon dynamique, comme l'ensemble des noms de rues associés à un bureau de poste distributeur). [3]

### **II.4. Ecriture arabe**

L'alphabet arabe comprend vingt-neuf lettres fondamentales (vingt-huit si l'on exclut la hamza, qui se comporte, soit comme une lettre à part entière, soit comme un diacritique). Le sens d'écriture va de droite à gauche. Il n'y a pas de différence entre les lettres manuscrites et les lettres imprimées ; les notions de lettre capitale et lettre minuscule n'existent pas. En revanche, la plupart des lettres s'attachent entre elles, même en imprimerie, et leur graphie diffère selon qu'elles sont précédées et/ou suivies d'autres lettres ou qu'elles sont isolées (variantes contextuelles). Certaines lettres, cependant, ne s'attachent jamais à la lettre suivante de ce fait, un mot unique peut être entrecoupé d'un ou plusieurs espaces, lesquels sont aussi utilisés pour séparer les mots. La longueur de cet espace inter-mot est généralement supérieure à l'espace intra-mot entre caractères non attachés [4,5].

L'écriture arabe est un cas très intéressant en reconnaissance automatique de l'écriture et du document suite à sa cursivité. De plus, les diacritiques rendent la reconnaissance encore plus délicate. Les techniques utilisées sont celles de la reconnaissance de mots du paragraphe même si le problème de la segmentation n'est pas encore totalement résolu. Enfin les systèmes actuels ne traitent que les écrits sans vocalisation donc sans diacritique.

### **III. Historique de reconnaissance des chiffres**

Dans cette partie , on a donne les différents travaux consacre la reconnaissance des chiffres manuscrits

#### **III.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 24 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[6]

#### **III.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. Il 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% [7]

### **III.3. Wu et Zhang (2010) :**

Wu et Zhang ont utilisé les caractéristiques de direction extraites pour la réduction de la dimensionnalité.  $K^{\text{ième}}$  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. [8]

### **III.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 résultat que le 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. [9]

### **III.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'apparement (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'apparement.

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 étendues 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. [10]

### **III.6. Lee et al (2015) :**

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. [11]

### **III.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 'apparement de l'image et la reconnaissance des bases de données MNIST et IPTP pour kPPV (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. [12]

### **III.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. [13]

## **IV. Conclusion**

Dans ce chapitre, nous expliquons les différents types de la reconnaissance de l'écriture avec des définitions de chaque type, puis nous donnons une brève historique sur récemment travaux dans le domaine de notre travail de recherche, la reconnaissance des chiffres.

Avant d'attaquer le travail essentiel de notre mémoire qui est la reconnaissance des images des chiffres, nous avons obligé de traiter ces image et préparer à cette phase qui est fait au deuxième chapitre prétraitement des images.

# **CHAPITRE II**

## **prétraitement pour la reconnaissance des chiffres**

## I. Introduction

Pour ce qui suit on introduit quelques définitions et notations empruntées au domaine du traitement d'image ; On désigne par traitement souvent appelé (prétraitement) un ensemble de méthodes dont l'objectif est : Soit de transformer les images en appliquant les opérations de manipulation de l'image produisant ainsi une nouvelle image ; Soit pour réduction de bruit par renforcement de l'homogénéité à l'intérieur de l'image ; améliorer la qualité d'image où préservation de la forme et renforcement de la séparation entre régions adjacentes.

## II. Etapes 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 [14]:

### II.1. Filtrage

Il existe plusieurs types de filtrage, on cite les suivants :

- Filtrage linéaire : sert au rehaussement, la détection de contour, lissage passe-bas ...
- Filtrage morphologique : l'élimination du bruit, la restauration des composantes connexes, et l'extraction du contour dans une image binaire peuvent se faire par des opérateurs morphologiques (éléments structurants).

#### II.1.1. Elimination de bruit (filtrage)

La Binarisation étant faite, on dit pose maintenant d'un objet image à deux niveaux. A Cette objet sont superposés des pixels «étranges» qu'il aurait été mieux de s'en débarrassé. Ces pixels sont appelés «bruit» et des techniques spéciales sont utilisées Pour nettoyer l'image de ces pixels. Il s'agit des techniques de filtrages qu'on essayera De voir dans cette partie.

##### II.1.1.1. Types de filtrages

###### a- Les filtres linéaires

Un filtre est linéaire s'il peut s'écrire de la manière suivante:

$$\text{Filtre}(c1.\text{image1}+c2.\text{image2})=c1\text{Filtre}(\text{image1})+c2\text{Filtre}(\text{image2})$$

Filtre signifie convoler une image  $I(x,y)$  avec une fonction  $f(x,y)$  appelée réponse impulsionnelle (ou opérateur de convolution) du filtre. Dans le cas continu, l'image filtrée est donnée par:



$$If(x, y) = (f * I)(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x', y') I(x - x', y - y') dx' dy' \quad (\text{II.1})$$

Dans le cas discret et pour prendre l'exemple simplifié d'une image carrée, les domaines de I et de f sont bornés. Le domaine de I est  $[N/2, +N/2]^2$  si N désigne la taille de l'image et le domaine de f est  $[K/2, +K/2]^2$  avec nécessairement  $K \leq N$ .

La convolution s'écrit alors:

$$If(i, j) = (f * I)(i, j) = \sum_{i'=-k/2}^{i'+k/2} \sum_{j'=-k/2}^{j'+k/2} f(i - i', j - j') I(i', j') \quad (\text{II.2})$$

On notera que le filtrage linéaire consiste simplement à remplacer chaque niveau de gris par une combinaison linéaire des niveaux de gris des points voisins.

Les filtres linéaires les plus connus sont les filtres passe-haut, passe-bas, Sorel, Prewitt et Laplace.

#### b- Les filtres non linéaires

Comme pour la convolution, ces filtres travaillent avec le voisinage du pixel courant. On prend alors le minimal, le maximal ou le médian du pixel et de son voisinage. Les filtres les plus connus sont les filtres Moyenne, Médiane, Nagao.

### **II.1.1.2. Les filtres utilisés**

#### a- Filtre passe bas

Il est utilisé pour diminuer le bruit sur une image. Le problème avec ce filtre est qu'il Atténue également tous les détails de hautes fréquences comme les contours. Ce filtre utilise le masque suivant

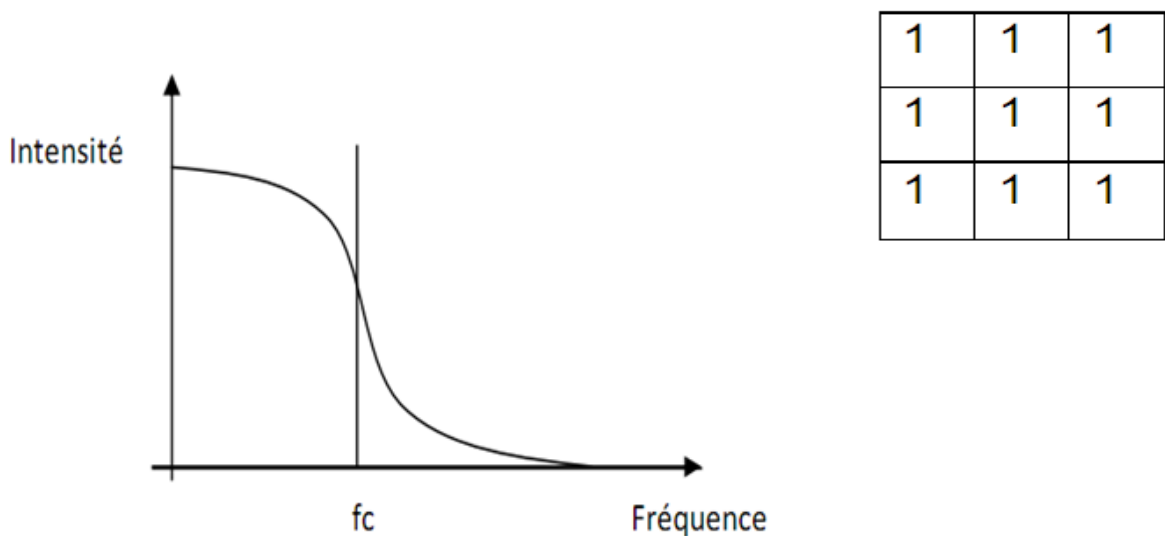


Figure II .1: Filtre passe-bas

### b- Filtre passe haut

Il est utilisé pour amplifier les détails de hautes fréquences. Il peut permettre par exemple de restaurer des images qui ont été dé focalisées et d'accentuer les contours. En fait, il fait ressortir les pixels compris entre des zones homogènes.

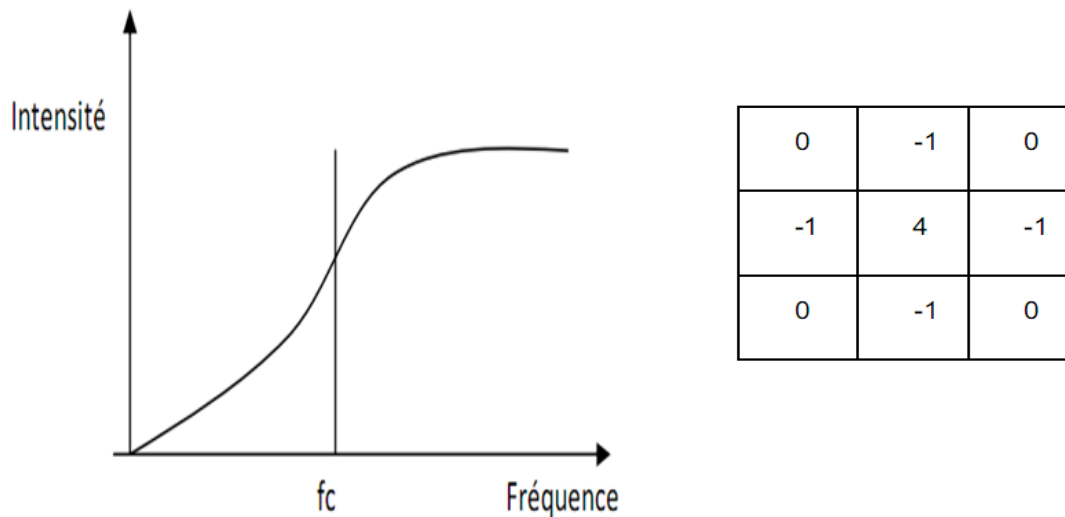


Figure II.2 : Filtre passe-haut

### c- Filtre Moyenner

Le filtrage des bruits blancs conduit à des filtres de moyenne (ou de lissage) qui remplacent le pixel courant par la valeur moyenne des pixels sur une fenêtre 3x3 ou 5x5. Ce filtre est bien compatible avec notre observation de départ selon laquelle le bruit est caractérisé par d'importantes hautes fréquences. Mais certaines fréquences (assimilées à de fortes variations dans les niveaux de gris) sont définitivement détruites, le lissage agissant comme un filtre passe bas.

L'image est non seulement lissée mais elle est aussi fortement altérée, les contours étant particulièrement dégradés par les filtres de moyenne. Le filtre de moyenne est une version d'une gamme plus vaste de filtres de lissages dits linéaires.

Ici, le masque représentant ce filtre est le suivant:

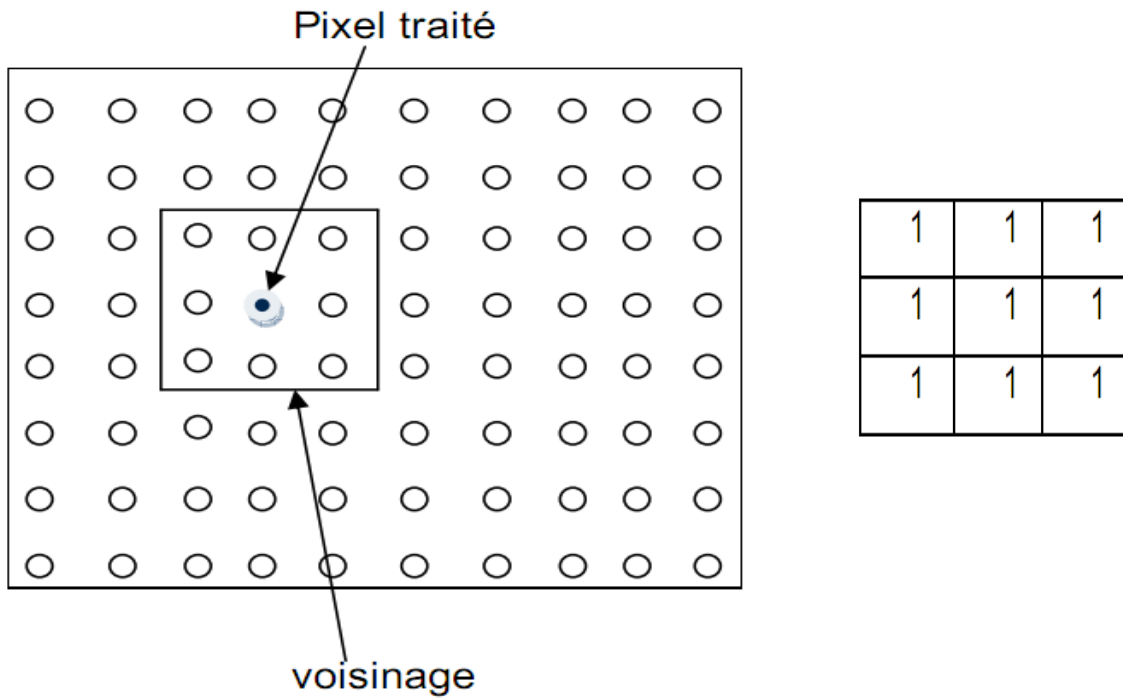


Figure II.3: filtre moyenner

Ce filtre, normalisé par un facteur  $1/9$  pour le ramener dans les valeurs normales, Génère un lissage fort.

Un lissage moyen peut-être réalisé par le biais du filtre suivant:

1	2	1
2	4	2
1	2	1

Le facteur de normalisation est  $1/16$  et le coefficient plus élevé du pixel Central permet d'atténuer l'effet de lissage.

d- Filtre Médian

Les inconvénients du filtre de Moyenne induisent une orientation vers le filtre dit Médian, qui préserve les contours. Ce filtre vise à remplacer la valeur du pixel central Par la valeur médiane de la répartition (on trie les luminances dans l'ordre croissant) des Niveaux de gris des pixels situés à l'intérieur de cette fenêtre (figure II.4)

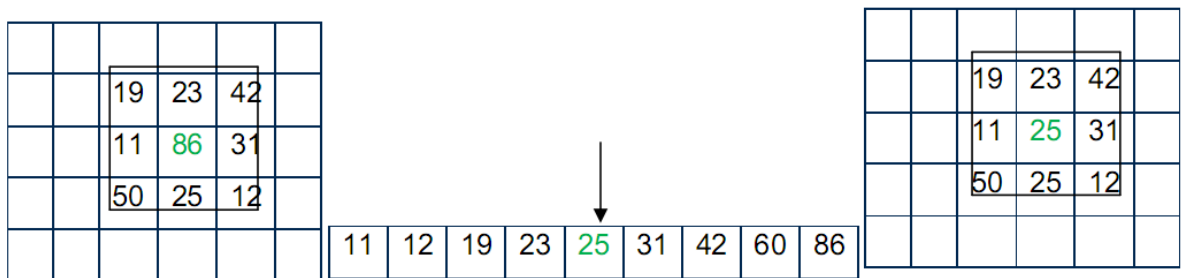


Figure II.4: Minéen œuvre du filtre Médian

Ce filtre - non linéaire - permet d'éliminer des pixels isolés ou les traînées fines Noyées dans des zones assez homogènes.

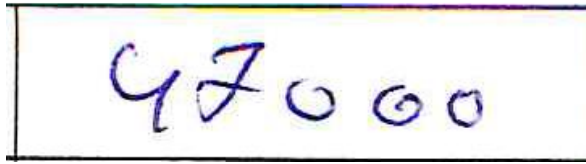


Figure II .5: image original de notre base de données

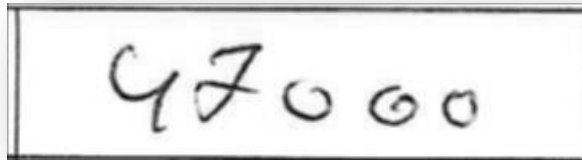


Figure II .6: image filtrée de notre base de données

### II.2. Binarisation

Binariser une image en niveaux de gris consiste à la transformer en une image en noir et blanc en associant à chaque pixel un label 0 ou 1 (0 pour blanc et 1 pour Noir). Ceci est illustré sur la Figure I.6.

On peut distinguer deux types de binarisation d'images, globale et locale. La Binarisation globale trouve un seuil valable pour toute l'image, tel que les pixels dont la valeur est au-dessus du seuil sont considérés comme l'arrière-plan (blanc) et les autres comme l'information utile (appartenant aux mots écrits) (noir). La Binarisation globale a l'avantage d'être rapide et, pour certains travaux comme ceux effectués sur la base Rimes, elle peut être amplement suffisante quand les images Utilisées sont très propres.

Un algorithme très connu et amplement utilisé est l'algorithme de séparation de classes d'Otsu présenté dans [126].

Pour appliquer cet algorithme, on calcule l'histogramme des valeurs des pixels en niveau de gris. L'algorithme Otsu sépare cet histogramme en deux classes (foncé et clair) à partir des Moments des deux premiers ordres.

On distingue en général deux approches : les approches à seuillage global, et les Approches à seuillage adaptatif.

a- Seuillage global

Le seuillage global consiste à prendre un seuil ajustable, mais identique pour toute l'image. Chaque pixel de l'image est comparé à ce seuil et prend la valeur blanc ou noir selon qu'il est supérieur ou inférieur.

Cette classification ne dépend alors que du niveau de gris du pixel considéré. Cette méthode convient pour les documents simples et de bonne qualité.

b- Seuillage adaptatif

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

Par exemple, dans [Ber 86], propose le calcul suivant pour déterminer le seuil de binarisation de chaque pixel:

$$F_{min}(x, y) = \min_{(xi, yi) \in v(x,y)} f(xi, yi) \quad (II.3)$$

$$F_{max}(x, y) = \max_{(xi, yi) \in v(x,y)} f(xi, yi) \quad (II.4)$$

$v(x, y)$  est le voisinage du pixel  $(x, y)$   $f(x, y)$  est la valeur de l'intensité du pixel  $(x, y)$  dans l'image  $F_{min}(x, y)$  est la valeur minimale de l'intensité dans le voisinage du pixel considéré  $F_{max}(x, y)$  est la valeur maximale de l'intensité dans le voisinage du pixel considéré La moyenne des valeurs  $F_{min}(x, y)$  et  $F_{max}(x, y)$  est utilisée pour déterminer le seuil de binarisation du pixel  $(x, y)$ :

$$g(x, y) = \frac{F_{max}(x,y) + F_{min}(x,y)}{2} \quad (II.5)$$

$$b(x, y) = \begin{cases} 1 & \text{si } f(x, y) < g(x, y) \\ 0 & \text{sinon} \end{cases} \quad (II.6)$$

La binarisation des documents bruités reste un domaine de recherche actif.

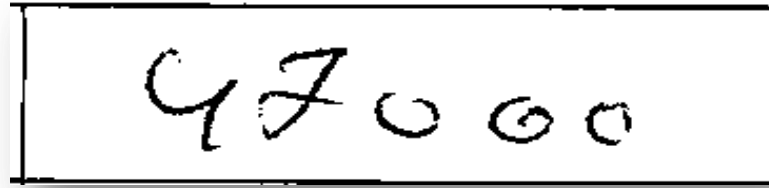


Figure II.7: Image binarisée

### II.3. La squelettisation

Sert à obtenir une épaisseur égale à 1 du trait d'écriture études ramené ainsi à une écriture linéaire. Le squelette doit préserver la forme, connexité, topologie et extrémités du tracé, et ne doit pas introduire d'éléments parasites.

La squelettisation est une opération qui permet de passer d'une image à sa représentation en "fil de fer".

Le squelette un pixel d'épaisseur. C'est une manière de représenter l'information indépendamment de l'épaisseur initiale de l'écriture. Il permet d'extraire des caractéristiques importantes, comme les intersections et Le nombre de tracés, leurs positions relatives.

Il est également possible de ré normalisé L'épaisseur de l'écriture à partir du squelette. Il n'existe pas de définition unique du squelette.

Le squelette doit seulement remplir Certaines conditions:

- Il doit être au **ssifin** que possible (typiquement, 1 pixel d'épaisseur)
- Il doit respecter la connexité
- Il doit être centré dans la forme qu'il représente;
- Il doit conserver les propriétés topologiques de la forme comme: Le nombre de partie et le nombre de trous;
- Il doit respecter les propriétés métriques de la forme.

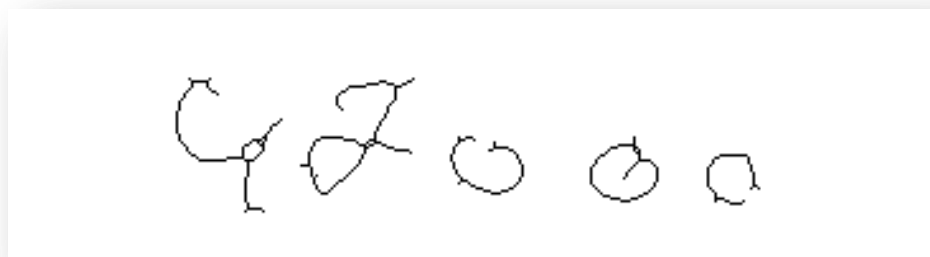


Figure II.8: squelettisation d'image

### **II.4. Lissage**

L'image des caractères peut être entachée de bruits dus aux artefacts de l'acquisition et à la qualité du document, conduisant soit à une absence de points ou à une surcharge de points. Les techniques de lissage permettent de résoudre ces problèmes par des opérations locales qu'on appelle opérations de bouchage et de nettoyage. L'opération de nettoyage permet de supprimer les petites tâches et les excroissances de la forme. Pour le bouchage il s'agit d'égaliser les contours et de boucher les trous internes à la forme du caractère en lui ajoutant des points noirs.

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

### **II.5. La normalisation**

Pour se donner toutes les chances d'une réussite pendant la phase de reconnaissance, les images «écriture» doivent être ramenées à une forme standard.

Cette standardisation peut concerner la taille de l'image comme elle peut concerner la correction de l'inclinaison des mots ou des lettres.

### **II.6. Segmentation**

Une image texte est un ensemble de lignes. La segmentation commence d'abord par la séparation de ces lignes les uns par rapport aux autres. On peut s'aider pour réaliser cette opération des histogrammes horizontaux. [AlBadr95].

Puis ces lignes sont séparées en mots. Pour cette opération, on peut se servir des histogrammes verticaux [AlBadr95].

La phase de segmentation est terminée à c'est il agit d'un traitement des mots entiers. Si c'est plutôt le traitement des caractères qui est envisagé une étape supplémentaire est encore à envisager: Le découpage des mots en lettres.

### **III. Conclusion**

Dans ce chapitre, on représenté quelque méthodes de prétraitement d'image par exemple de filtrage , binairisation et sequlletisation pour renforcer le contraste d'une image afin de pouvoir par la suite appliquer plus facilement des traitements avant l'analyse principale et l'extraction de l'information.



# **CHAPITRE III :**

**Méthode SVM**

## I. Introduction

Les machines à vecteurs de support sont un ensemble des techniques d'apprentissage destinées à résoudre les problèmes de discrimination , c'est-à-dire décider à quelle classe appartient un échantillon, ou de régression, c'est-à-dire prédire la valeur numérique d'une variable.

Le succès de cette méthode est justifié par les solides bases théoriques qui la soutiennent. Il existe en effet un lien direct entre la théorie de l'apprentissage statistique et l'algorithme d'apprentissage de SVM.

SVM est une méthode de classification particulièrement bien adaptée pour traiter des données de très hautes dimensions telles que les textes, les images et la voix...etc. Dans ce qui suit on présente les aspects théoriques de la méthode SVM.

## II. Principe des machines à vecteurs de support

Un SVM, comme un perceptron, trouve un séparateur linéaire entre les points de données de deux classes différentes. En général, il peut y avoir plusieurs séparateurs possibles entre les classes (en supposant le problème linéairement séparable) et qu'un perceptron n'a pas de préférence parmi celles ci. Dans les SVMs, cependant, nous faisons un choix particulier parmi tous les séparateurs possibles : nous voulons celui avec la “marge” maximale. [15]

### II.1. Notions de base

#### a) Hyperplan

Plaçons-nous dans le cas d'une classification binaire (i.e. les exemples à classer répartis en 2 classes). On appelle hyperplan séparateur un hyperplan qui sépare les deux classes figure III.1, en particulier il sépare leurs points d'apprentissage. Comme il n'est en générale pas possible d'en trouver un, on se contentera donc de chercher un hyperplan discriminant qui est une approximation au sens d'un critère à fixer (maximiser la distance entre ces deux classes). [16] [17]

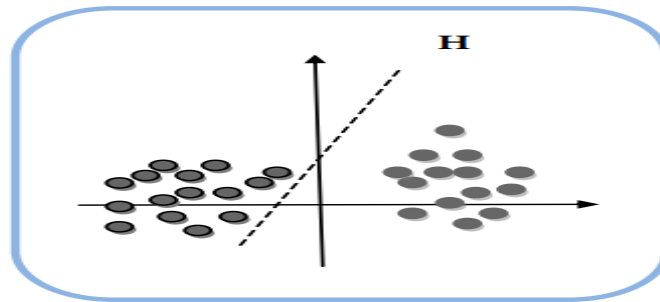


Figure III 1 l'hyperplan H qui sépare les deux ensembles de points.

*b) Vecteurs de support*

Pour une tâche de détermination de l'hyperplan séparable des SVM est d'utiliser seulement les points les plus proches (i.e. Les points de la frontière entre les deux classes des données) parmi l'ensemble total d'apprentissage, ces point sont appelés vecteurs de support figure III.2. [16] [17]

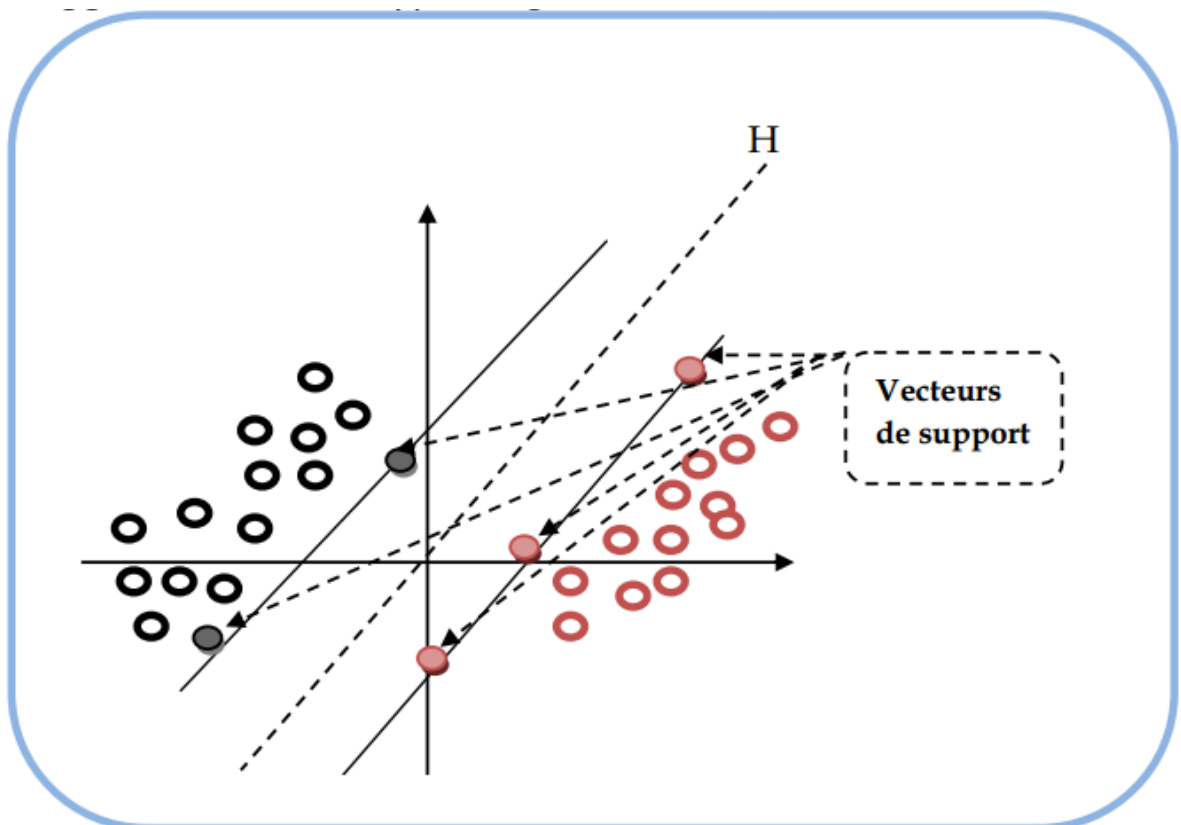


Figure III. 2 : Les vecteurs de support.

*c) Marge*

il existe une infinité d'hyperplans capable de séparer parfaitement les deux classes d'exemples. Le principe des SVM est de choisir celui qui va maximiser la distance minimale entre l'hyperplan et les exemples d'apprentissage (i.e. La distance entre l'hyperplan et les vecteurs de support), cette distance est appelée la marge (figure III.3). [16] [17]

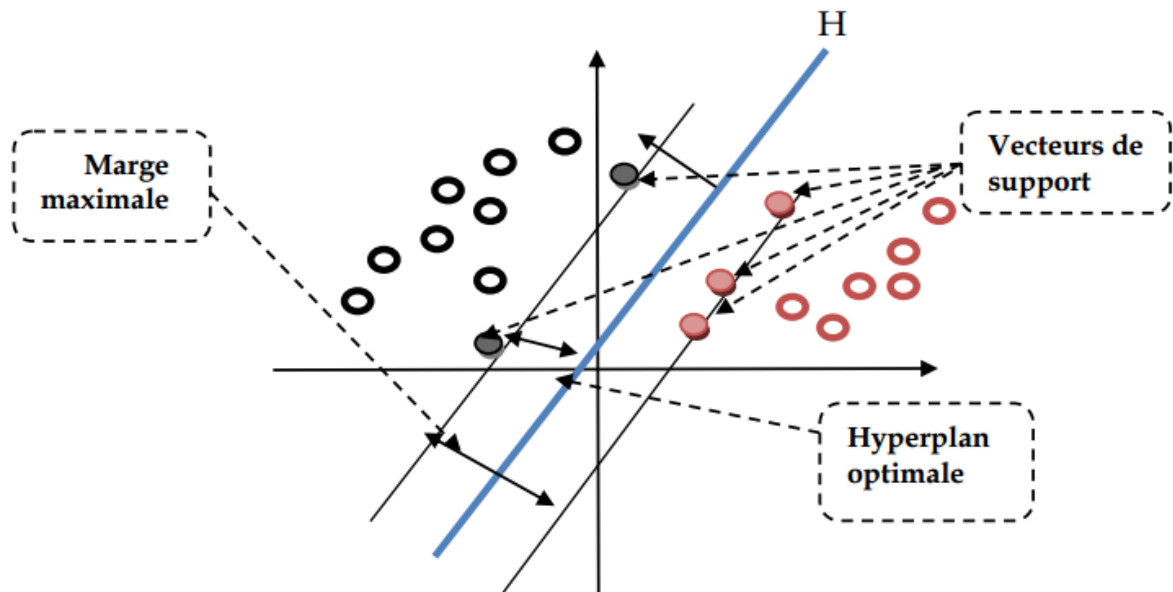


Figure III 3 hyperplan optimal, vecteurs de support et marge maximale.

## II.2. Propriétés fondamentales

*a) Pourquoi maximiser la marge ?*

Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsqu'on classe un nouvel exemple. De plus, si l'on trouve le classificateur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma figure III.4, la partie droite nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie gauche qu'avec une plus petite marge, l'exemple se voit mal classé. [16] [17]

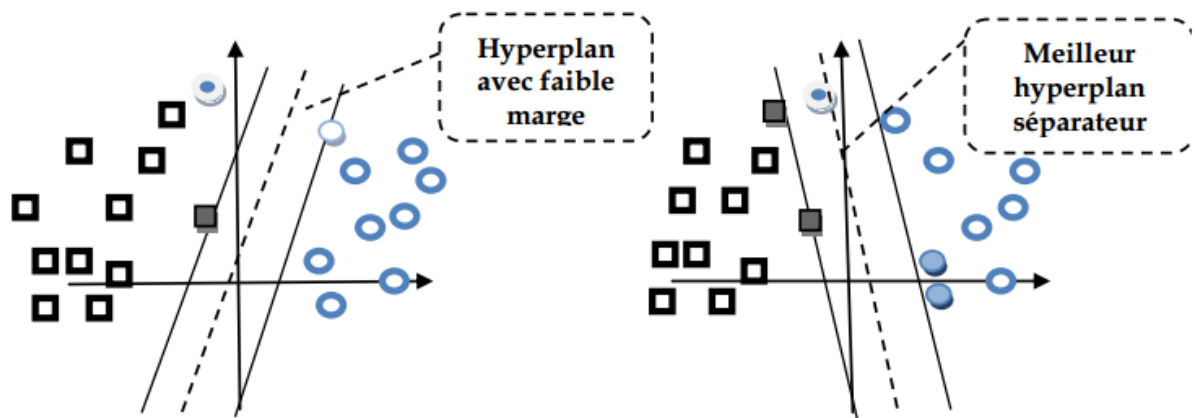
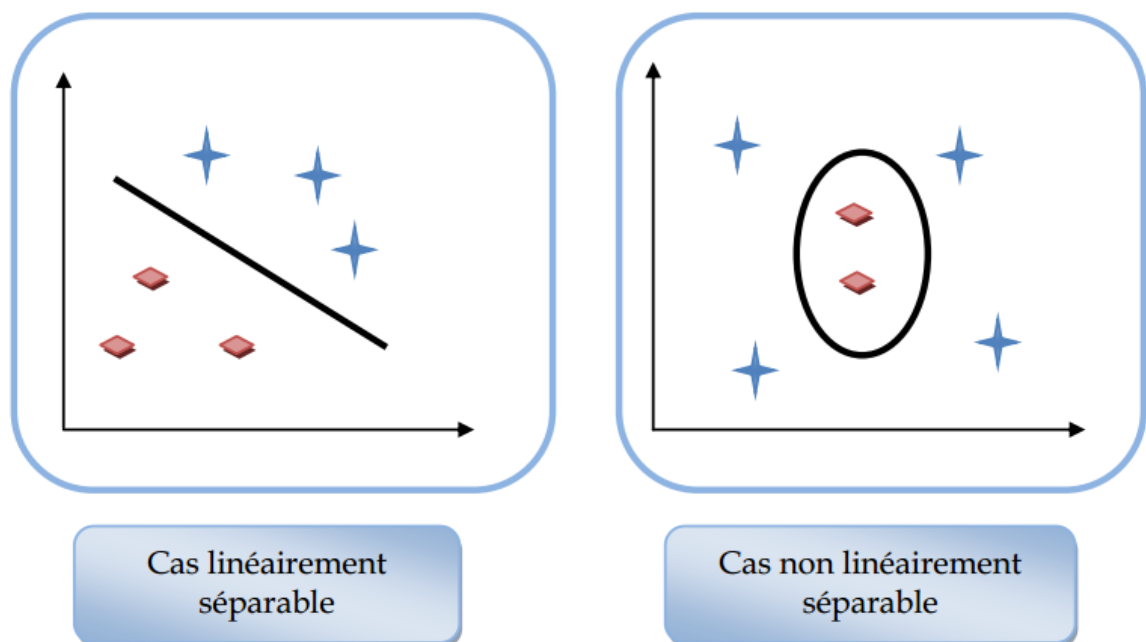


Figure III 4 meilleur hyperplan séparateur.

*b) Linéarité et non-linéarité*

Parmi les modèles des SVM, on constate les cas linéairement séparables et les cas non linéairement séparables. Les premiers sont les plus simples des SVM car ils permettent de trouver facilement le classificateur linéaire. Dans la plupart des problèmes réels il n'y a pas de séparation linéaire possible entre les données, le classificateur de marge maximale ne peut pas être utilisé car il fonctionne seulement si les classes de données d'apprentissage sont linéairement séparables. [16] [17]

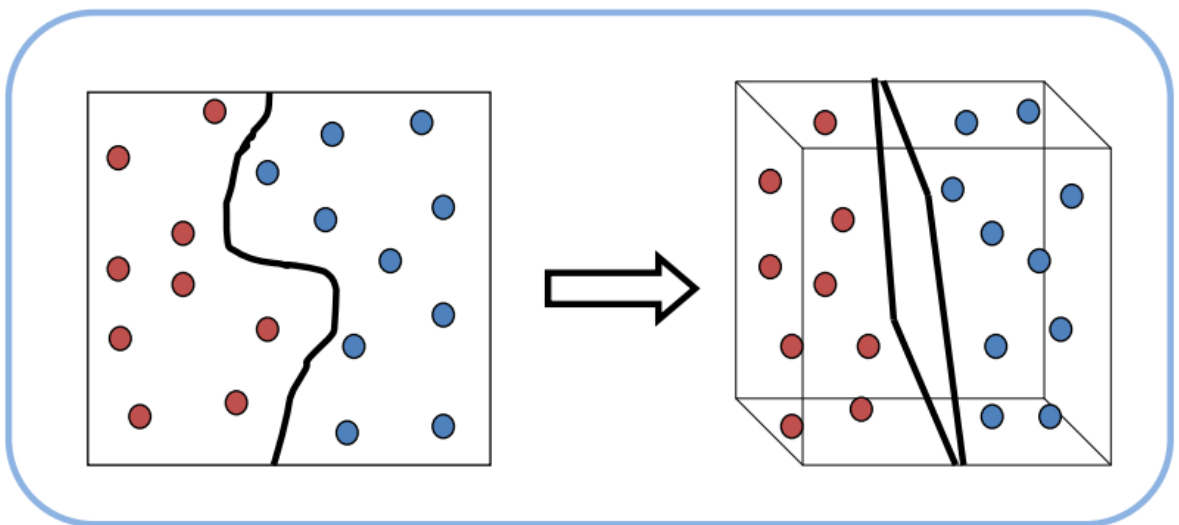


**Figure**

c) Cas non linéaire

Pour surmonter les inconvénients des cas non linéairement séparable, l'idée des SVM est de changer l'espace des données. La transformation non linéaire des données peut permettre une séparation linéaire des exemples dans un nouvel espace. On va donc avoir un changement de dimension. Ce nouvel espace est appelé « espace de re-description ».

En effet, intuitivement, plus la dimension de l'espace de re-description est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples est élevée. Ceci est illustré par le schéma suivant :



**Figure**

On a donc une transformation d'un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de plus grande dimension. Cette transformation non linéaire est réalisée via une fonction noyau. En pratique, quelques familles de fonctions noyau paramétrables sont connues et il revient à l'utilisateur de SVM d'effectuer des tests pour déterminer celle qui convient le mieux pour son application. On peut citer les exemples de noyaux suivants : polynomiale, gaussien, sigmoïde et Laplacien. [16] [17]

**II.3. Fondement mathématiques [18]**

a) Cas de données linéairement séparables

Considérons  $\ll l \gg$  points  $\{(x_1, y_1), (x_2, y_2), \dots, \dots, \dots, (x_l, y_l)\}, x_i \in R^N$

Avec  $i = 1, 2, \dots, L$  et  $y_i \in \{-1, 1\}$

Ces points sont classés en utilisant une famille de fonctions linéaires définis par :

$$\langle w, x \rangle + b = 0 \quad (1)$$

Avec  $w \in R^N$  et  $b \in R$  de telle sorte que la fonction de décision concernant l'appartenance d'un point à l'une des deux classes soit donnée par :

$$f(x) = \text{sgn}(\langle w, x \rangle + b) \quad (2)$$

La fonction (1) représente l'équation d'hyperplan H. La fonction de décision (2) va donc observer de quel côté de H se trouve l'élément de x.

On appelle la marge d'un élément la distance euclidienne prise perpendiculairement entre H et x. Si on prend un point quelconque t sur H, cette marge peut s'exprimer en :

$$M_x = \frac{w}{\|w\|} (x - t) \quad (3)$$

La marge de toutes les données est définie comme étant :

$$M = \min_{x \in E} M_x \quad (4)$$

L'approche de classification par SVM tend à maximiser cette marge pour séparer le plus clairement possible deux classes. Intuitivement, avoir une marge la plus large possible sécurise mieux le processus d'affectation d'un nouvel élément à l'une des classes. Un SVM fait donc partie des classificateurs à marge maximale.

Un classificateur à marge maximale est un classificateur dont l'hyperplan optimal séparant deux classes est une solution du problème d'optimisation mathématique suivant (forme primale) :

$$\text{minimiser } \frac{1}{2} \|w\|^2 (\langle w, x \rangle + b) \geq 1 \quad x \in E \quad (5)$$

La fonction objective de ce problème est le carré de l'inverse de la double marge qu'on veut maximiser. La contrainte unique correspond au fait que les éléments x doivent être bien placés.

La résolution de ce problème nécessite de fixer les paramètres w et b qui constituent les variables  $\omega_i$  de la machine d'apprentissage.

Les classificateurs à marge maximale donnent de bons résultats lorsque les données sont linéairement séparables.

La tâche de discrimination est de trouver un hyperplan qui sépare deux (ou plus) ensembles de vecteurs (Figure III 5).

Pour la détection et la reconnaissance des vertèbres, ces deux classes peuvent être région de vertèbre ou non.

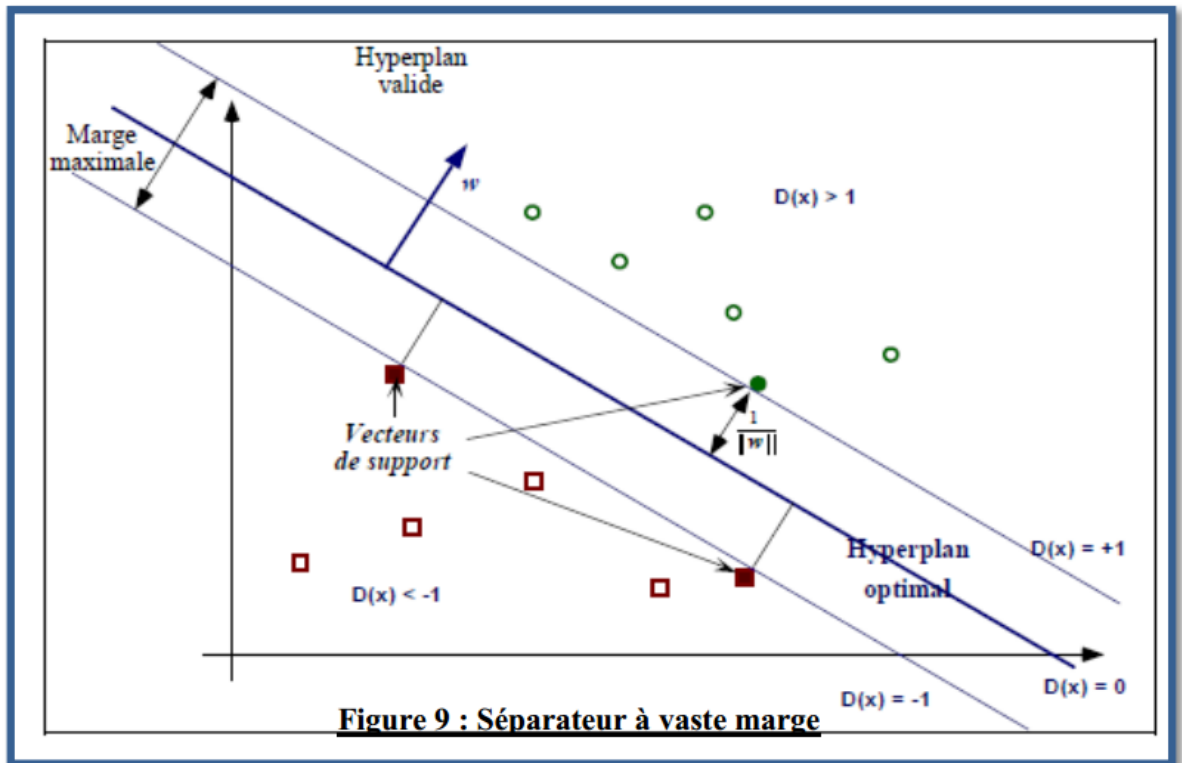


Figure III. 5 : Séparateur à vaste marge.

b) Forme duale

La formulation primale peut être transformée en formulation duale en utilisant les multiplicateurs de Lagrange. L'équation (5) s'écrit alors sous la forme suivante :

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l a_i (y_i (\langle w, x_i \rangle + b) - 1) \quad (6)$$

c) Cas des données non - linéairement séparables

En pratique, il est assez rare d'avoir des données linéairement séparables. Afin de traiter également des données bruitées ou non linéairement séparables, les SVM ont été généralisées grâce à deux outils : la marge souple (soft margin) et les fonctions noyau (kernel functions).

Le principe de la marge souple est d'autoriser des erreurs de classification. Le nouveau problème de séparation optimale est reformulé comme suit :

$$\text{MIN}_{w,b,\epsilon} \frac{1}{2} w^T W + C \sum_{i=1}^l \epsilon_i, C \geq 0 \quad \text{sous contraintes} \quad (7)$$

$$y_i (\langle w, x \rangle + b) \geq +1 - \epsilon_i$$



$$\epsilon_i \geq 0 \text{ pour } i = 0, \dots, l$$

Un terme de pénalité est introduit dans la formule (7), Le paramètre C est défini par l'utilisateur. Il peut être interprété comme une tolérance au bruit de classificateur.

**Remarque :**

L'idée de base pour les données non linéairement séparables, est de projeter l'espace d'entrée (espace des données) dans un espace de plus grande dimension appelé espace de caractéristiques (feature space) afin d'obtenir une configuration linéairement séparable (à l'approximation de la marge souple près) de nos données, et d'appliquer alors l'algorithme SVM.

### III. Architecture d'un classificateur SVM [18]

#### III.1. la fonction noyau

Afin de résoudre le problème de données non linéaires, la fonction noyau joue le rôle central de liaison des vecteurs d'entrées à l'espace de caractéristiques de grande dimension. (Voir figure III 6)

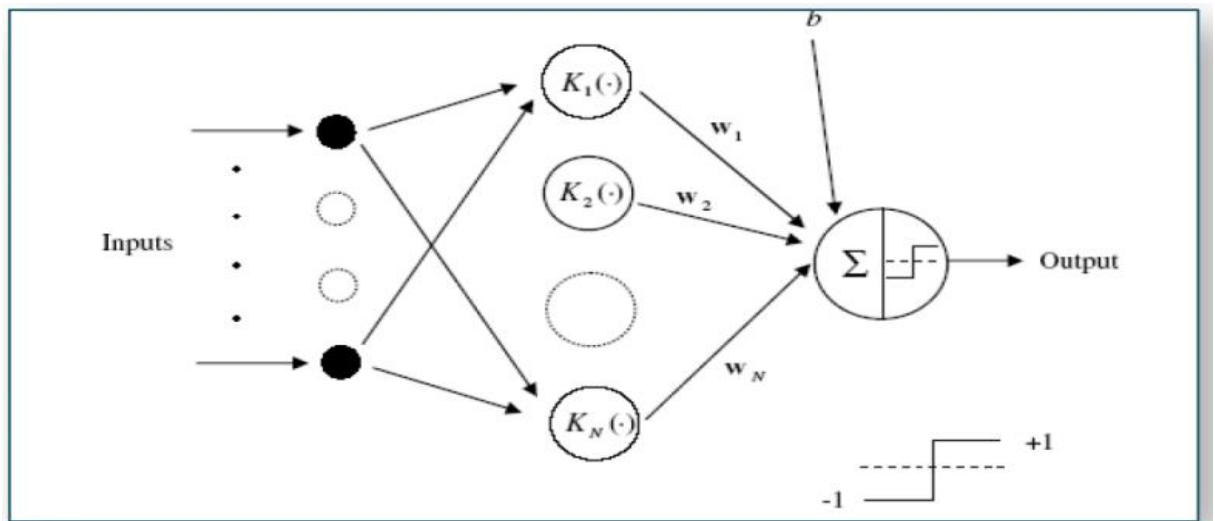


Figure III. 6 : Architecture d'une machine à vecteurs de support (d'un nombre N)

Les choix typiques pour la fonction noyau sont :

- 1- Noyau gaussien à base radiale (ou RBF : Radial Basis Fonction) :

$$K(x, xi) = \exp[-\gamma \|x - xi\|] \quad (8)$$

- 2- Noyau polynômiale :

$$K(xi - xj) = (x_i^T x_j + 1)^p \quad (9)$$

Avec p : une constante qui spécifie le degré du polynôme.

### III.2. Sélection de modèle SVM [18]

Une machine à vecteur de support binaire sépare les exemples positifs des exemples négatifs dans la phase d'apprentissage.

Les multiplicateurs de Lagrange  $\alpha_i$  pour chaque machine binaire sont déterminés par la minimisation de la fonction coût donnée par la formule (10) suivante :

$$P w = \frac{1}{2} \|w\|^2 \quad (10)$$

Ce problème est résolu à partir de la forme duale qui est exprimée par :

$$L_D = \sum_i \alpha_i - \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i x_j \quad (11)$$

Cela revient à chercher :  $\sum_i \alpha_i \alpha_j = 0$  et  $0 \leq \alpha \leq C$

Si la valeur du paramètre de régularisation « C » qui contrôle la tolérance aux erreurs de classification dans la phase d'apprentissage est élevée, donc plus de pénalité sera donnée à l'erreur.

Le vecteur d'apprentissage  $x_i$  qui a une valeur de  $\alpha_i$  non nul est appelé « vecteur de support ».

### III.3. Estimation de l'erreur de généralisation [18]

La technique la plus populaire pour l'estimation de l'erreur de généralisation est la validation croisée (en anglais cross validation ou Leave One Out ( LOO)) qui est utilisée indépendamment de la nature de la machine d'apprentissage utilisée. Le principe de cette méthode consiste à séparer en N sous-ensembles les échantillons de la base d'apprentissage, à apprendre sur N - 1 sous-ensembles, à valider sur le sous- ensemble restant, puis à faire tourner les sous-ensembles de façon à ce que chacun ait pu être testé. Au final, on note le nombre d'erreurs de classification de la procédure LOO par  $l(X_1, Y_1, \dots, X_n, Y_n)$

En effet, l'espérance  $E(\cdot)$  de la probabilité  $p_{\text{err}}^{n-1}$  de l'erreur de test par une machine entraînée à partir de n-1 exemples est donnée par la formule (25):

$$p_{\text{err}}^{n-1} = \frac{1}{n} E(l(X_1, Y_1, \dots, X_n, Y_n)) \quad (12)$$

Cependant la procédure est coûteuse en calcul, car nécessitant « n » apprentissages.

Une procédure simple d'estimation d'erreur de généralisation est la validation croisée dite K-fold. Elle consiste à diviser l'ensemble des données en k sous-ensembles

mutuellement exclusifs de taille approximativement égale. L'apprentissage de la machine est effectué en utilisant  $k-1$  sous ensemble et le test est effectué sur le sous-ensemble restant. Cette procédure est utilisée une fois pour le test.

La moyenne des  $k$  taux d'erreur obtenus estime l'erreur de généralisation. Nous avons montré dans notre travail que la procédure de la validation croisée donne de bonnes estimations de l'erreur de généralisation. Le nombre d'ensembles utilisé dans notre travail est égal à 5 (5fold).

### **III.4. Détection par SVM [18]**

Notre approche de détection est basée sur l'extraction automatique des fenêtres rectangulaires. Le contenu de ces fenêtres est transformé en vecteurs de caractéristiques en appliquant les méthodes d'extraction citées précédemment. Ces vecteurs sont ensuite injectés dans notre classificateur SVM afin de déterminer si une fenêtre correspond ou pas à une région de vertèbres.

Les fenêtres extraites sont de taille fixe, avec un chevauchement pour parcourir la totalité de l'image, respectant l'algorithme suivant (voir figure III 7):

1. Mettre une fenêtre de taille fixe  $W_{sub} * H_{sub}$  dans le coin supérieur gauche de l'image.
2. Extraire une région de l'image dans cette fenêtre.
3. Décaler la fenêtre à droite avec un pas de  $S_x$  ( $S_x < W_{sub}$ ).
4. Extraire une nouvelle région de l'image dans la fenêtre. Refaire l'étape 3 et 4 jusqu'à ce que la fenêtre soit au bord droit de l'image.
5. Décaler la fenêtre au bord gauche de l'image.
6. Décaler la fenêtre en bas avec un pas de  $S_y$  ( $S_y < H_{sub}$ ).
7. Extraire une nouvelle région de l'image dans la fenêtre.
8. Refaire l'étape 3 à 8 jusqu'à ce que la fenêtre soit au bord droit de l'image.

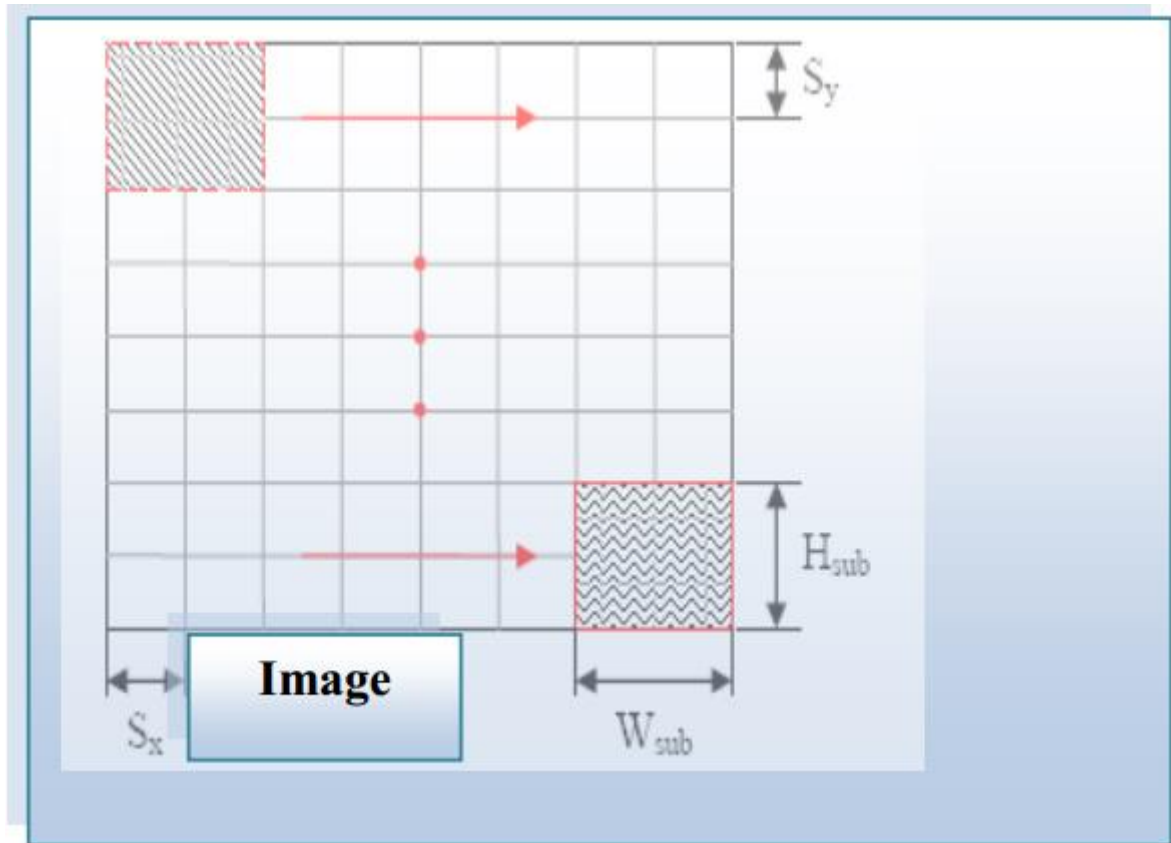


Figure III 7 : Processus d'extraction de fenêtres à la phase de détection.

**Post traitement :** Afin d'éliminer les fausses réponses du classifieur SVM une méthode d'analyse à base de coefficients de corrélation est utilisée.

#### IV. Le coefficient de corrélation [18]

Est un indice de mesure de l'intensité d'un lien qui peut exister entre deux variables. Le coefficient de corrélation peut prendre une valeur comprise entre -1 et +1. S'il est égal à 0, cela signifie qu'il n'existe aucun lien entre ces 2 variables. Il est très généralement utilisé dans le cadre de l'analyse de variables quantitatives.

Dans notre cas les deux variables représentent la fenêtre de détection qui a été mal classée et un modèle d'une région positive contenant la vertèbre.

Notons A la matrice représentant la fenêtre de balayage et B la matrice représentant notre modèle moyen. La formule utilisée pour calculer ce coefficient "r" est :

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \quad (13)$$

La valeur du Coefficient de Corrélation est toujours comprise entre -1 et +1 Plus il est proche de 1 ou de -1, plus les points sont proches d'une droite. S'il est égal à + 1, les points sont strictement alignés.

## **V. Conclusion**

Dans ce chapitre, on a tenté de présenter d'une manière simple et complète le concept de système d'apprentissage introduit par Vladimir Vapnik, les « Support Vecteur Machine bi-classes » on a donné une vision générale et une vision purement mathématiques des SVM. Cette méthode de classification est basée sur la recherche d'un hyperplan qui permet de séparer au mieux des ensembles de données. On a exposé les cas linéairement séparable et les cas non linéairement séparables qui nécessitent l'utilisation de fonction noyau (Kernel) pour changer d'espace. Cette méthode est applicable pour des taches de classification à deux classes, mais il existe des extensions pour la classification multi classes.

# **CHAPITRE IV**

## **Test & Résultat**

## I. Introduction

Dans ce chapitre, nous avons décrit la méthodologie que nous avons proposé pour réaliser la reconnaissance des chiffres des code postaux à partir de l'image binaire. Cette méthode utilise la corrélation entre les images de test et les images d'apprentissage.

Nous utilisons Matlab **2013a**, installé sous un PC portable de caractéristique ( $\mu$ Processeur Intel Core i5, RAM 4 GB), **Windows 10** .....

## II. Description notre travail

### II.1. Base de donné

Nous avons créé une base de données contient tous les codes postaux de la Wilaya de Ghardaïa, qui sont 35 codes. Ces codes sont décrits par 84 personnes, chaque personne répète chaque code postal 5 fois (Annexe). Finalement nous obtenons environ 15 000 images de taille 60\*250 pixel chaque images.

### II.2. 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).

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

$$TMBR(\%) = \frac{\sum TRC}{\text{nombre des classe}} \quad (4.2)$$

### II.3. Résultat & discussion

On a fait trois tests :

1. **le test 1:** dans le premier test, nous avons utilisé une (01) image pour chaque code pour chaque personne, c.-à-d au total 84 images d'apprentissage pour chaque code, et pour les tests nous utilisons trois (03) images pour chaque code pour chaque personne c.-à-d au total 252 images de test pour chaque code, on obtient les résultats suivantes :

Mis en forme : Police :13 pt, Police de script complexe :13 pt, Français (France)

Mis en forme : Police : (Par défaut) +Titres CS (Times New Roman), 13 pt, Police de script complexe :+Titres CS (Times New Roman), 13 pt, Français (France)

Mis en forme : Français (France)

Mis en forme : Français (France)

Mis en forme : Français (France)

Mis en forme : Français (France)

Mis en forme : Français (France)





47021	33%	
47022	0%	
47023	33%	
47024	100%	
47025	66%	
47026	0%	
47027	33%	
47028	33%	
47029	33%	
47030	33%	
47031	66%	
47800	0%	
47801	100%	

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme

Mis en forme

Mis en forme

Mis en forme

Mis en forme

Mis en forme

Mis en forme

Mis en forme

Mis en forme

Mis en forme

Mis en forme

On a trouvé pour le 1<sup>er</sup> test :

$$\text{TMBR1(\%)} = 44\%$$

2. **le test 2:** dans le deuxième test, nous avons utilisé deux (02) images pour chaque code pour chaque personne, c.-à-d au total 168 images d'apprentissage pour chaque code, et pour les tests nous utilisons trois (03) images pour chaque code pour chaque personne c.-à-d au total 252 images de test pour chaque code, on obtient les résultats suivantes :

Tableau 4.2 : résultat pour 2 images d'apprentissage et 3 images de test pour chaque code pour chaque personne.

Les codes postaux	TRC(%)	TMBR(%)
47000	66%	
47001	66%	

Mis en forme

Mis en forme

Mis en forme



[Test & résultat]

47025	0%	
47026	66%	
47027	66%	
47028	0%	
47029	33%	
47030	66%	
47031	100%	
47032	33%	
47800	66%	
47801	0%	

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme : Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

Mis en forme : Couleur de police : Accent 1

Mis en forme

Mis en forme : Couleur de police : Accent 1

Mis en forme

Mis en forme : Couleur de police : Accent 1

Mis en forme

Mis en forme : Couleur de police : Accent 1

Mis en forme

Mis en forme : Couleur de police : Accent 1

On a trouvé pour le 1<sup>ier</sup> test :

$$\text{TMBR2(\%)} = 51\%$$

3. **le test 03:** dans le deuxième test, nous avons utilisé deux (02) images pour chaque code pour chaque personne, c.-à-d au total 252 images d'apprentissage pour chaque code, et pour les tests nous utilisons trois (03) images pour chaque code pour chaque personne c.-à-d au total 252 images de test pour chaque code, on obtient les résultats suivantes :

Tableau 4.2 : résultat pour 3 images d'apprentissage et 3 images de test pour chaque code pour chaque personne.

Les codes postaux	TRC(%)	TMBR(%)
47000	100%	
47001	66%	
47002	66%	

Mis en forme

Mis en forme : Couleur de police : Accent 1

Mis en forme

Mis en forme : Couleur de police : Accent 1

Mis en forme

Mis en forme : Couleur de police : Accent 1



47026	66%
47027	33%
47028	33%
47029	66%
47030	100%
47031	66%
47032	66%
47800	66%
47801	33%

**TMBR3(%)= 65%**

Par comparaisons nous remarquons que le taux de reconnaissance de troisième test est mieux que les deux premiers, et c'est conséquent de la base d'apprentissage élevé et important dans le troisième test. Donc nous concluons que lorsque la base d'apprentissage augmente le taux de reconnaissance s'augmentera aussi.

### III. Conclusion

Dans ce chapitre, nous avons donné une idée sur notre base de données créée, puis une brève explication sur l'environnement de travail (Matériel et Logiciel). A la fin nous donnons nos résultats et le conséquent de la variation de ces résultats qu'il est la variation de la base d'apprentissage qui joue un rôle important pour obtenir un bon résultat.

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

**Mis en forme :** Police :Gras, Couleur de police : Accent 1, Police de script complexe :Gras

**Mis en forme :** Couleur de police : Accent 1

# **Conclusion Général**

## **Conclusion général**

Notre projet de fin d'étude s'est porté sur la reconnaissance des chiffres manuscrites, Afin d'atteindre notre objectif qui était de concevoir et de réaliser un Système de Corrélacion des images d'apprentissage et de test. Qui est une technique facilement à appliquer actuellement dans le domaine de traitement et de reconnaissance d'écriture.

A ce stade, il faut noter que les recherches dans ce domaine ne cessent de se développer, la plupart des systèmes proposés jusqu'ici, présentent de bonnes performances sur les même types de faux, celle-ci diminuent considérablement si toutes les catégories de faux sont présents.

# **BIBLIOGRAPHIE**



## Chapter I :

- [1] Mr Bendjeddou Toufik "Le choix de parametre pour la reconnaissance des chiffres manuscrit" these majustere. University BADJI MOUKHTAR ANNABA 2007-2008
- [2] M. Amrouch, 'Reconnaissance de caractères imprimés et manuscrits, textes et documents basée sur les modèles de Markov cachés', *Thèse de doctorat, Université Ibn Zohr Maroc*, 2014.
- [3] Mr MEGAR Ishak et GEUDDA Adel " ontribution à la reconnaissance d'écriture arabe manuscrite en utilisant la classification neuronale" these master proffessionnelle university kasdi merbah ouargla 2015-2016
- [4] F. Menasri, 'Contributions à la reconnaissance de l'écriture arabe manuscrite', *Thèse de doctorat, Université paris* , 2008.
- [5] K. Merdour , 'Reconnaissance de formes du chiffre arabe imprimé : Application au code à barre d'un produit', *Thèse de Magister , Université de Batna* , 2009.
- [6] IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 2, FEBRUARY 2002 Recognizing Handwritten Digits Using Hierarchical Products of Experts Guy Mayraz and Geoffrey E. Hinton
- [7] Ernest Kussul, Tatyana Baidyk : Improved Method of Handwritten Digit Recognition, Glushkov Institute of Cybernetics, Kiev, Ukraine, 2004.
- [8] Ming Wu Zhen Zhang, Handwritten Digit Classification using the MNIST Data Set, ResearchGate publications, 2010.
- [9] High density impulse denoising by a fuzzy filter Techniques:Survey TarunSrivastava(M.Tech-Vlsi) Suresh GyanVihar University
- [10] 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.
- [11] Lee et al, Generalizing Pooling Functions in Convolutional Neural Networks: Mixed, Gated, and Tree, NSF awards, 2015.
- [12] 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.*

## [BIBLIOGRAPHIE]

---

[13] Dundar et al, (2016).

[14]. G.Abdeldjalil, 2011, «Segmentation automatique pour la reconnaissance numérique des chèques bancaires Algériens » ; Thèse de MAGISTER, centre Universitaire de Khanchela

### **Chapitre II :**

[17] .M. Arif Mohamm, hybridation des données on/off-pour la vérification des signatures manuscrites, mémoire magister 2009.

[18]. L. Bréhélin et O. Gascual, “Modèles de Markov cachés et apprentissage de séquences”, Le temps, l’espace et l’évolutif en sciences du traitement de l’information, Edition Cépaduès, H. Prade, R. Jeansoulin et C. Garbay éditeurs, 2000, pp 407-421.

[19]. A. Belaïd, G. Saon, “Utilisation des processus markoviens en reconnaissance de l’écriture”, Traitement de signal, vol.14 n°2, pp 161-178, 1997.

[20]. N. Ben Amara, A. Belaïd et N. Ellouze, “Utilisation des modèles markoviens en reconnaissance de l’écriture arabe- Etat de l’art”, Colloque international francophone sur l’écrit et le document, Lyon , France 2000 , pp. 181-191

[21]. N. Ben amara, A. Belaid, “Une méthode stochastique pour la reconnaissance de l’écriture arabe imprimée”, Forum de la Recherche en Informatique, Tunisie, 16-18 juillet 1996.

[22]. L. R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”, Proc. of IEEE, VOL. 77, NO. 2. pp 257-286. Février 1989.

[23]. <http://www.gpds.ulpgc.es/download/index.htm>

### **Chapter III :**

[15] Nicolas Turenne , "Apprentissage automatique (Machine Learning) "INRA ,2006.

[16] Mohamadally Hasan et Fomani Boris , " SVM : Machines a Vecteurs de Support ou Separateurs a Vastes Marges"Versailles St Quentin, Francejanvier 2006.

## [BIBLIOGRAPHIE]

---

[17] Senoussaoui Mohammed,"application des modèles de markov cachés & les machines à vecteurs de support pour la reconnaissance des caractères isolés d'écriture en ligne" , pour l'obtention de diplôme de magister ,SIMPA,2007.

[18] Mme Benyahia Souad Née Belaidi << Application de la classe des méthodes d'apprentissage statistique SVM (support vector machine) pour la reconnaissance des formes dans les images>> These magister university abu baker belgaid telmcen année 2011/2012 page 27

### **Chapter IV:**

[19] <**Reconnaissance hors ligne des chiffres manuscrite isolé (Base de donnée M.N.I.S.T)**> présenté par MENASRIA Abdelaali & Z EMOULI Saddam, THESE MASTER UNIVERSITY TEBSSA 2015/2016

**ANNEXE :**

47000	47000	47000	47000	47000	47000
47001	47001	47001	47001	47001	47001.
47002	47002	47002	47002	47002	47002.
47003	47003	47003	47003	47003	47003
47004	47004	47004	47004	47004	47004.
47005	47005	47005	47005	47005	47005.
47006	47006	47006	47006	47006	47006.
47007	47007	47007	47007	47007	47007.
47008	47008	47008	47008	47008	47008.
47009	47009	47009	47009	47009	47009.
47010	47010	47010	47010	47010	47010.
47011	47011	47011	47011	47011	47011.
47012	47012	47012	47012	47012	47012.
47013	47013	47013	47013	47013	47013.
47014	47014	47014	47014	47014	47014.
47015	47015	47015	47015	47015	47015.
47016	47016	47016	47016	47016	47016.
47017	47017	47017	47017	47017	47017.
47018	47018	47018	47018	47018	47018.
47019	47019	47019	47019	47019	47019.
47020	47020	47020	47020	47020	47020.
47021	47021	47021	47021	47021	47021.
47022	47022	47022	47022	47022	47022.
47023	47023	47023	47023	47023	47023.
47024	47024	47024	47024	47024	47024.
47025	47025	47025	47025	47025	47025.
47026	47026	47026	47026	47026	47026.
47027	47027	47027	47027	47027	47027.
47028	47028	47028	47028	47028	47028.
47029	47029	47029	47029	47029	47029.
47030	47030	47030	47030	47030	47030.
47031	47031	47031	47031	47031	47031.
47032	47032	47032	47032	47032	47032.
47800	47800	47800	47800	47800	47800.
47801	47801	47801	47801	47801	47801.



47000	47000	47000	47000	47000	47000
47001	47001	47001	47001	47001	47001
47002	47002	47002	47002	47002	47002
47003	47003	47003	47003	47003	47003
47004	47004	47004	47004	47004	47004
47005	47005	47005	47005	47005	47005
47006	47006	47006	47006	47006	47006
47007	47007	47007	47007	47007	47007
47008	47008	47008	47008	47008	47008
47009	47009	47009	47009	47009	47009
47010	47010	47010	47010	47010	47010
47011	47011	47011	47011	47011	47011
47012	47012	47012	47012	47012	47012
47013	47013	47013	47013	47013	47013
47014	47014	47014	47014	47014	47014
47015	47015	47015	47015	47015	47015
47016	47016	47016	47016	47016	47016
47017	47017	47017	47017	47017	47017
47018	47018	47018	47018	47018	47018
47019	47019	47019	47019	47019	47019
47020	47020	47020	47020	47020	47020
47021	47021	47021	47021	47021	47021
47022	47022	47022	47022	47022	47022
47023	47023	47023	47023	47023	47023
47024	47024	47024	47024	47024	47024
47025	47025	47025	47025	47025	47025
47026	47026	47026	47026	47026	47026
47027	47027	47027	47027	47027	47027
47028	47028	47028	47028	47028	47028
47029	47029	47029	47029	47029	47029
47030	47030	47030	47030	47030	47030
47031	47031	47031	47031	47031	47031
47032	47032	47032	47032	47032	47032
47800	47800	47800	47800	47800	47800
47801	47801	47801	47801	47801	7801