



République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université AMO de Bouira

Faculté des Sciences et des Sciences Appliquées

Département d'Informatique

Mémoire de Master 2

en Informatique

Spécialité : Génie des Systèmes Informatiques

Reconnaissance des chiffres manuscrits

Encadré par

— Hachine Gherbi Ahcine

Réalisé par

— Hadj-Moussa Maroua

— Lamani soumia

2022/2023



التصريح الشرفي الخاص بالالتزام بقواعد النزاهة العلمية لإنجاز بحث

انا الممضي اسفله،

الأستاذة **ساج موسى مروة** الصفة: طالب (ماستر / دكتوراه)
الحامل (ة) لبطاقة التعريف الوطنية: **404448607**، والصادرة بتاريخ: **23/01/2023**
المسجل (ة) بكلية / معهد العلوم التطبيقية قسم اعلام آلي
تخصص: **نظم معلوماتية**
والمكلف (ة) بإنجاز اعمال بحث (مذكورة، التخرج، مذكرة ماستر، مذكرة ماجستير، اطروحة دكتوراه).
عنوانها: **Reconnaissance de chiffres manuscrites**

أصرح بشرفي اني ألتزم بمراعاة المعايير العلمية والمنهجية الاخلاقيات المهنية والنزاهة الاكاديمية المطلوبة
في انجاز البحث المذكور أعلاه.

توقيع المعني (ة)

التاريخ: **23/07/2023**

البويرة في 23/07/2023	هيئة مراقبة السرقة العلمية:
الامضاء	النسبة: 30 %



التصريح الشرفي الخاص بالالتزام بقواعد النزاهة العلمية لإنجاز بحث

انا الممضي اسفله،

السيد(ة)..... **د. محمد لماني** الصفة: طالب (ماستر / دكتوراه)

الحامل(ة) لبطاقة التعريف الوطنية: 100623784 والصادرة بتاريخ 17-04-2016

المسجل(ة) بكلية / معهد العلوم والعلوم التطبيقية قسم اعلام ألي

تخصص: نظم معلومات

والمكلف(ة) بإنجاز اعمال بحث(مذكرة، التخرج، مذكرة ماستر، مذكرة ماجستير، اطروحة دكتوراه).

عنوانها: *Reconnaissance de chiffres manuscrits*

أصرح بشرفي اني ألتزم بمراعاة المعايير العلمية والمنهجية الاخلاقيات المهنية والنزاهة الاكاديمية المطلوبة في انجاز البحث المذكور أعلاه.

توقيع المعني (ة)

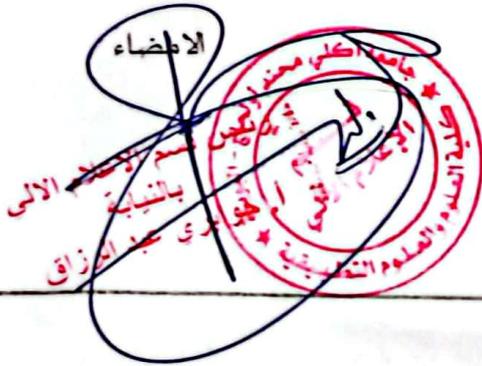
التاريخ: 25-07-2023

البويرة في: 2023/07/01

هيئة مراقبة السرقة العلمية:

% 30

النسبة:



Remerciements

Nous souhaitons adresser nos sincères remerciements et dédicaces à toutes les personnes qui ont apporté leur contribution précieuse à la réussite de notre projet et à la préparation de notre mémoire de fin d'études. Leurs efforts et leur soutien inestimable ont joué un rôle déterminant dans notre incroyable succès.

Tout d'abord, nous tenons à remercier notre superviseur académique, Hacine Gharbi Ahcine, qui nous a guidés et guidés tout au long de ce parcours académique. Il nous a donné de précieux conseils et astuces et nous a montré le chemin du succès.

Nous souhaitons exprimer notre gratitude envers les membres du corps enseignant qui ont partagé leur expertise, leurs connaissances et leurs conseils dans divers domaines liés à notre projet. Leur contribution inestimable et leur engagement actif ont joué un rôle fondamental dans le développement et l'amélioration de nos idées.

Nous tenons également à exprimer notre reconnaissance spéciale envers nos amis et notre famille qui ont été un soutien constant et une source d'encouragement tout au long de la préparation de notre mémoire. Leur soutien indéfectible et leurs encouragements précieux ont été notre force dans les moments difficiles et ont alimenté notre motivation pour persévérer.

Enfin, nous voudrions exprimer notre profonde gratitude aux différentes entités qui ont fourni les ressources et les installations nécessaires pour mener à bien notre projet. Grâce à leur soutien, nous avons pu utiliser l'équipement, les logiciels et les ressources importants qui nous ont aidés à faire notre travail de manière efficace et de haute qualité.

Nous adressons nos plus sincères remerciements et notre reconnaissance à tous ceux qui ont joué un rôle dans le succès de notre projet. Sans votre précieuse contribution, nous n'aurions pas pu accomplir cette prouesse. Nous vous remercions tous chaleureusement et nous vous souhaitons beaucoup de succès dans vos futures entreprises.

Dédicaces

Je dédie ce modeste travail Á :

Je suis indescriptiblement reconnaissant envers ma merveilleuse famille. Je tiens à leur exprimer mes profonds remerciements et pour tout ce qu'ils ont fait pour moi en tant qu'étudiant.

Ma famille aimante et solidaire a toujours été à mes côtés dans mon parcours scolaire. Dès le début, ils m'ont apporté le soutien et les encouragements nécessaires qui m'ont aidé à atteindre mes objectifs. Qu'il s'agisse de m'aider à bien organiser les cours et de m'encourager à étudier, ou de m'apporter un soutien émotionnel dans les moments difficiles, ils ont été un pilier inestimable dans mon parcours éducatif.

Je suis profondément reconnaissant envers ma famille qui me fait me sentir en sécurité, protégé et soutenu. Je me rends compte que leurs efforts ne viennent pas de nulle part, et qu'ils viennent de leur amour et de leur profond intérêt pour ma réussite et mon bonheur

Hadj moussa marwa

Dédicaces

Je dédie ce modeste travail Á :

Je dédie ces mots humbles et reconnaissants A tous ceux qui m'ont
soutenu et encouragé dans mon cheminement

A ma chère famille et amis Merci d'être resté à mes côtés tout le temps
Avec votre soutien et votre amour, vous avez alimenté ma force et ma
détermination

Que ces mots soient l'expression de ma profonde gratitude Mon
appréciation pour chaque instant de ma vie Tu es la raison de mon succès
et de mon bonheur J'espère être toujours à ton meilleur Que l'amour et la
gratitude restent dans nos cœurs Merci de tout mon cœur.

J'espère que cette dédicace reflète vos sentiments et votre appréciation
pour ceux qui vous ont soutenu et qui vous ont accompagné dans votre
cheminement.

Lamani soumia

ملخص

تعتبر معرفة الأرقام المكتوبة يدوياً واحدة من أهم التحديات البحثية في مجال تحليل والتعرف علي الوثائق. تشمل الجوانب المرتبطة بهذه المشكلة بشكل أساسي التجزئة، والتعرف علي لأحرف الفردية، بالإضافة إلى التعرف على التسلسلات والأرقام. تركز هذه الدراسة بشكل خاص على التعرف على الأرقام المكتوبة يدوياً بشكل منفصل، باستخدام تقنيات التعلم العميق لتحليل محتوى الصورة من قاعدة البيانات الخاصة بالارقام المكتوبة باليد . نستخدم نموذج (ك)أقرب جيران باستخدام طرق استخراج الأرقام لتحسين معدل الدقة

كلمات مفتاحية : التعرف علي الارقام المكتوبة يدويا ، التعلم الالي ، خصائص الشعاع

Abstract

The recognition of handwritten digits has been one of the most significant research challenges in the field of document analysis and recognition. The aspects related to this problem mainly include segmentation, recognition of individual characters, as well as sequence and digit recognition. This study specifically focuses on the recognition of isolated handwritten digits, using machine learning techniques to analyze the content of an image from the MNIST database. We leverage the "k-nearest neighbors (KNN)" model using digit extraction methods such as HOG and Gabor to improve the accuracy rate .

keyword : Recognition of handwritten digits,machine learning,Number extraction,Feature vector

Résumé

La reconnaissance des chiffres manuscrits a été l'un des défis de recherche les plus importants dans le domaine de l'analyse et de la reconnaissance de documents. Les aspects liés à ce problème comprennent principalement la segmentation, la reconnaissance des caractères individuels, ainsi que la reconnaissance des séquences et des chiffres. Cette étude se concentre spécifiquement sur la reconnaissance des chiffres manuscrits, en utilisant des techniques d'apprentissage automatique pour analyser le contenu d'une image provenant de la base de données MNIST. Nous exploitons le modèle des "k plus proches voisins (KNN)" en utilisant des méthodes d'extraction de chiffres telles que HOG et Gabor pour améliorer le taux d'exactitude.

Mote clé : Reconnaissanc des chiffres manuscrits,Apprentissage automatique,extraction de chiffres,Vecteur caractristique

Table des matières

Table Des Matières	i
Table des figures	iv
Liste des Tableaux	v
Liste des abréviations	vi
Introduction générale	1
1 Reconnaissance de l'écriture	4
1.1 Introduction	4
1.2 Historique	5
1.3 Définition d'OCR	6
1.3.1 Différents Aspects de Reconnaissance	6
1.4 Reconnaissance D'écriture Manuscrite	7
1.4.1 Reconnaissance de caractère	7
1.4.2 Les forme de text manuscrit	8
1.4.3 Reconnaissances des chiffres	10
1.5 Domaines D'application	11
1.6 Architecture Générale d'un systeme de Reconnaissance des Chiffres manus- crits	11
1.6.1 Phase d'acquisition	12
1.6.2 Phase de pré-traitement	12
1.6.3 Phase de segmentation	15

1.6.4	Phase d'extraction de caractéristiques	18
1.6.5	Phase Classification	19
1.7	Conclusion	20
2	Machine Learning	21
2.1	Introduction	21
2.2	Intelligence artificielle	22
2.3	Machine Learning	22
2.3.1	Definition du Machine Learning	22
2.3.2	Types du Machine Learning	23
2.4	Les Algorithmes De Machine Learning	28
2.4.1	Les arbres de Désicion	28
2.4.2	L'algorithme des k plus proches voisins	29
2.4.3	SVM (Support Vector Machine)	33
2.4.4	Classification Naïve Bayésienne	35
2.5	Algorithme d' Extraction des caractéristique	36
2.5.1	Le Histogramme de gradient Orienté (HOG)	36
2.5.2	Les moments de Zernike	40
2.5.3	Filtre de Gabor	41
2.6	Méthode d'évaluation d'un Modèle d'apprentissage automatique	43
2.6.1	Matrice de confusion	43
2.6.2	l'Accuracy	44
2.7	Conclusion	44
3	Etat de l'art	45
3.1	Introduction	45
3.2	les modèles existants	45
3.2.1	Mayraz et Hinton (2002)	45
3.2.2	Wu et Zhang(2010)	46
3.2.3	Ebrahim et Jampour (2014)	46
3.2.4	Wakahara et Yamashita (2014)	47
3.2.5	Lee et al (2015)	47
3.2.6	Wakahara et Yamashita(2016)	48

3.2.7	Dundar et al (2016)	49
3.2.8	Yann LeCun, Corinna Cortes et Christopher J.C. Burges(2020)	49
3.2.9	Venkatraman Sanjeevi et Ashutosh Mishr(2020)	50
3.2.10	Margot GLENAZ(2021)	51
3.2.11	John Smith(2022)	51
3.3	Des résultats rapportés dans la littérature	53
3.4	Conclusion	54
4	Architecture de notre système	55
4.1	Introduction	55
4.2	Problématique	55
4.3	Contribution	56
4.4	Description de notre systeme de reconnaissance	57
4.5	Analyse des phases de programmation dans le processus	59
4.6	Conclusion	61
5	Résultats et évaluation	62
5.1	Introduction	62
5.2	Langage de calcule technique MATLAB	62
5.3	Configuration utilisée	63
5.4	Bases de données utilisée	63
5.5	Selection de modele	64
5.6	Résultats et dicussion	64
5.6.1	Evaluations des modeles	64
5.6.2	Comparaison entre les modeles proposes	66
5.6.3	Comparaison avec d'autres Travaux de la Litterature	66
5.7	Présentation les interfaces de l'application :	67
5.8	Conclusion	68
	Conclusion générale	69
	Bibliographie	71

Table des figures

1.1	Différentes formes du texte manuscrit	4
1.2	Différentes formes du texte manuscrit [7]	9
1.3	Exemples de texte contraint et non contraint [8].	9
1.4	Architecture de reconnaissance des chiffres manuscrits [11]	12
1.5	Exemple de la binarisation	13
1.6	Exemple d'un chiffre manuscrit normalisé [15]	14
1.7	Exemple de Squelettisation [16]	15
1.8	Exemple de segmentation [17]	16
2.1	Deep Learning vs Machine Learning [38]	23
2.2	Types du ML [32]	24
2.3	Schema d'Apprentissage Supervise [32]	25
2.4	Méthode de k plus proches voisins [29]	31
2.5	Hyperplan Optimal Avec Une Marge Maximale[45]	34
4.1	Exemple chaîne de chiffre cursif.	56
4.2	Schema de notre systeme	57
5.1	Un exemple de chiffres extrait de la base MNIST	64
5.2	Image testee	65
5.3	matrices de confusion du Hog	65
5.4	matrices de confusion du Gabor	66
5.5	Interface de notre système.	68

Liste des tableaux

1.1	Une brève comparaison entre les approches en-ligne et hors-ligne [3][4] . . .	7
2.1	présente les différences entre les quatre types Machine Learning [36][38] . . .	27
2.2	Les Avantages et Inconvénients Des Arbres De Décision[40]	29
2.3	Avantages et Inconvénients de S V M [46]	35
2.4	Avantage et Inconvénients de la Classification Naïve bayes [48]	36
2.5	Matrice de Confusion [55]	43
3.1	Des résultats mentionnés dans la littérature	53
5.1	Les résultats des attributs d'évaluation pour les différents modèles	66
5.2	la comparaison de nos modèles propose avec d'autres travaux	67

Liste des abréviations

Acronyme	Signification
OCR	Optical Character Recognition .
IA	l'intelligence artificielle.
DL	deep learning.
ML	Machine Learning.
RNA	Réseaux De Neurone.
NNDEGD	Nearest-Neighbor Distance of Equi-Gradient Direction.
GAN	les Réseaux Adversaires Generatifs.
SVM	Support Vector Machine.
knn	K-nearest neighbors
HOG	Histogram of Oriented Gradients

Introduction générale

L'écriture manuscrite est toujours omniprésente dans notre vie quotidienne, et constitue un lien étroit et privilégié entre les hommes car elle leur permet d'échanger de façon naturelle des idées, des informations, des sentiments. Malgré les avancées dans plusieurs domaines mais toujours en trouve des difficultés ou on utilise papiers écrits a la main . Depuis la fin des années soixante, des efforts considérables ont été déployés dans le domaine de la reconnaissance de l'écriture. Des progrès significatifs ont été réalisés, et des systèmes commerciaux sont désormais largement disponibles, notamment dans le secteur postal (pour la lecture des chiffres des chèques et le tri postal) ainsi que dans les banques (pour le traitement des chèques et des factures). À l'heure actuelle, les recherches se concentrent sur l'interprétation de l'écriture manuscrite, ouvrant de nouvelles perspectives et défis.

La détection des chiffres manuscrits est une tâche relevant de l'apprentissage automatique, qui implique d'entraîner un modèle à reconnaître et identifier des chiffres écrits à la main. Cette activité est fréquemment utilisée dans diverses applications, notamment la reconnaissance de caractères sur des chèques, la numérisation de documents et la lecture des codes postaux.

En général, le processus de reconnaissance des chiffres manuscrits implique la collecte et la préparation d'un vaste ensemble d'images contenant des chiffres écrits à la main. Ces images servent ensuite à entraîner un modèle d'apprentissage automatique tel qu'un réseau neuronal ou KNN, qui est capable d'apprendre à reconnaître les différents chiffres. Une fois le modèle entraîné, il peut être utilisé pour identifier les chiffres présents dans de nouvelles images. Pour cela, l'image est introduite dans le modèle, qui produit ensuite

une prédiction sur le chiffre représenté. La reconnaissance des chiffres manuscrits revêt une importance capitale dans de nombreux domaines, tels que la reconnaissance optique de caractères (OCR), l'automatisation de la lecture de documents et la reconnaissance de codes-barres.

La reconnaissance des caractères peut être divisée en deux domaines distincts : la reconnaissance en ligne et la reconnaissance hors ligne. Dans la reconnaissance en ligne, les caractères sont identifiés pendant leur processus d'écriture. Les données sont unidimensionnelles et dynamiques, représentant la trajectoire du stylo. En revanche, dans la reconnaissance hors ligne, le processus d'écriture n'est pas accessible et les données sont bidimensionnelles, sous la forme d'images numérisées des caractères. Dans cette étude, notre attention se concentre spécifiquement sur la reconnaissance hors ligne, qui concerne la lecture automatique des chiffres manuscrits.

Problématique

La détection des chiffres manuscrits représente un défi majeur dans les domaines de la vision par ordinateur et de l'intelligence artificielle. Les chiffres écrits à la main peuvent présenter des variations considérables en termes de style d'écriture, de taille, d'inclinaison et de déformations. De plus, l'ajout de bruit, de ratures ou d'autres perturbations peut rendre cette tâche encore plus complexe.

Il précise des chiffres manuscrits revêt une importance cruciale dans de nombreuses applications, telles que la numérisation de documents : capturer des documents sans avoir à les retaper, donner à vos documents papier une nouvelle vie .Si vous avez déjà retapé un rapport de 15 pages ou un grand tableau de chiffres, vous savez à quel point cela peut être fastidieux et long ,Le besoin de recapturer les données est en fait beaucoup plus important que ce que vous pensez initialement.

Alors, l'OCR est peut-être l'outil le plus efficace, il n'y a donc aucune raison pour que vous ne vous déchargiez pas de la tâche fastidieuse de retaper du texte et des données.

Ce mémoire se concentre sur la mise en œuvre d'un modèle KNN avec des filtres Gabor et des caractéristiques HOG (Histogram of Oriented Gradients).

Nous utiliserons des techniques de prétraitement pour améliorer les performances. Le langage de programmation utilisé sera MATLAB.

La structure de ce travail est divisée en quatre chapitres :

Chapitre 1 : présente une étude bibliographique sur le système de reconnaissance automatique et mettra en évidence les différentes étapes du processus de reconnaissance de chiffres manuscrits.

Chapitre 2 : Le deuxième chapitre aborde le domaine de l'apprentissage automatique en explorant ses différents types et algorithmes. Nous avons examiné les avantages et les limites de ces approches, ce qui nous a permis de modéliser un modèle d'apprentissage automatique en nous basant sur ces connaissances. De plus, une revue de la littérature est présentée, mettant en évidence différentes méthodes de reconnaissance des chiffres , en décrivant chacune d'entre elles en détail.

Chapitre 3 : Le troisième chapitre présente un état de l'art sur différentes méthodes de reconnaissance des chiffre , en partant de la description de chaque méthode individuellement

Chapitre 4 : Le chapitre quatre représente Architecture générale de notre système .

Chapitre 5 : Le chapitre final se consacre à la discussion des résultats obtenus et à l'implémentation de notre proposition.

Reconnaissance de l'écriture

1.1 Introduction

La reconnaissance de l'écriture manuscrite est une forme de communication naturelle qui présente l'avantage d'être familière à la plupart des individus. Cela en fait un moyen facile d'interagir avec les ordinateurs. L'un des objectifs les plus recherchés est de conférer aux ordinateurs une forme d'humanité (Figure 1.1).

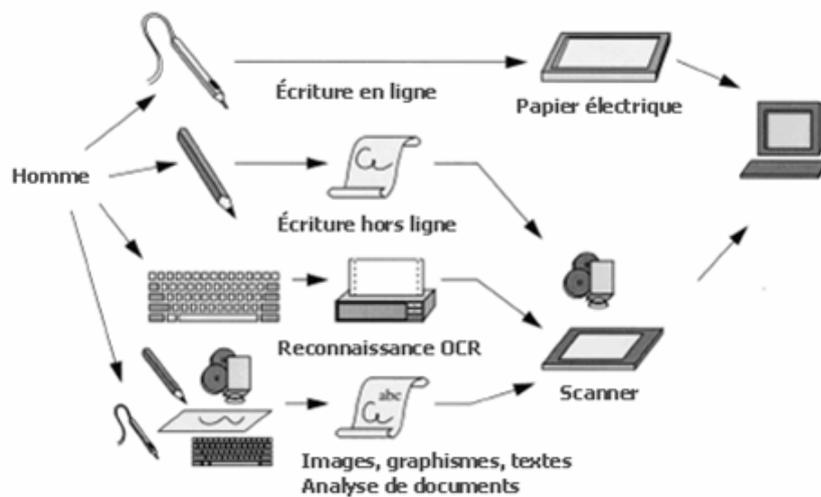


FIGURE 1.1 – Différentes formes du texte manuscrit

Par conséquent, l'objectif de la reconnaissance de l'écriture manuscrite est de convertir un texte écrit en une représentation compréhensible par les machines et facilement trai-

table par des logiciels de traitement de texte. Cette tâche n'est pas simple, car l'écriture présente une variété infinie de manifestations. En effet, chaque individu possède son propre style d'écriture, et il existe de nombreux styles différents. La difficulté de cette tâche réside dans la diversité des types d'écriture, qu'il s'agisse de l'écriture manuscrite, cursive ou dactylographiée.

1.2 Historique

La reconnaissance de l'écriture, également connue sous le nom de OCR (Optical Character Recognition), a connu ses premières tentatives dès les années 1900. TYURIN a inventé le scanner à balayage pour la télévision, tandis qu'ALBE en 1912 et Thomas en 1925 ont contribué à imiter l'interprétation visuelle humaine par l'informatique. Un tournant majeur s'est produit avec l'invention du premier ordinateur en 1946 par MAUCHLY et ECKERT, ce qui a ouvert la voie aux premières expériences de reconnaissance de caractères. Dans les années soixante et soixante-dix, les premiers systèmes de reconnaissance automatique de texte imprimé ont été développés, et en 1975, les lecteurs japonais étaient couramment utilisés pour décrypter les codes postaux écrits à la main ou tapés à la machine.

Pendant cette même période, le Français CONTER était en train de développer un système de lecture automatique de texte imprimé destiné aux personnes non voyantes. Par la suite, la société américaine KURZWELL a amélioré ce système en introduisant des machines de lecture pour les aveugles qui pouvaient formuler le texte à haute voix grâce à la synthèse vocale. Durant cette phase, les chercheurs ont été confrontés à de nombreuses difficultés, notamment la complexité du problème de reconnaissance due à la grande variabilité de l'écriture manuscrite, ainsi que les contraintes liées à la disponibilité limitée de mémoire et de puissance de calcul nécessaires pour mettre en place des systèmes concrets opérant en temps réel. Cependant, depuis les années 1980, les progrès récents dans le domaine électronique, en particulier l'avènement de calculateurs puissants à faible coût, ont permis de résoudre ce type de problème. En conséquence, la recherche en reconnaissance manuscrite s'est multipliée de manière spectaculaire et de nombreuses nouvelles techniques ont vu le jour [1].

1.3 Définition d'OCR

OCR (Optical Character Recognition) est une technologie permettant de numériser et de reconnaître automatiquement le texte présent dans des images, des documents imprimés ou manuscrits. Cette technologie utilise des algorithmes complexes pour analyser les caractéristiques de chaque caractère et les comparer à une base de données de caractères préalablement enregistrés [2].

L'OCR permet de transformer les documents papier en format numérique modifiable, ce qui facilite leur stockage, leur traitement et leur utilisation ultérieure. Les applications de l'OCR sont nombreuses et incluent la numérisation de livres, de magazines, de factures, de formulaires, de tickets de caisse, etc

1.3.1 Différents Aspects de Reconnaissance

Reconnaissance En-ligne

Ce mode a été spécialement conçu pour reconnaître l'écriture manuscrite en temps réel, offrant ainsi une reconnaissance continue des symboles écrits à la main. Une caractéristique essentielle de ce modèle est sa capacité à permettre une correction et une révision interactives de l'écriture. Pour profiter pleinement de cette fonctionnalité, l'utilisation d'appareils tels que des tablettes graphiques ou des stylos électroniques est nécessaire. Ces outils permettent une expérience d'écriture plus fluide et précise, offrant ainsi un moyen pratique d'interagir avec le système de reconnaissance de texte [3].

Reconnaissance Hors-ligne

La Reconnaissance Hors ligne est déclenchée une fois que le texte a été acquis. Ce mode est adapté aux documents imprimés ainsi qu'aux manuscrits déjà écrits. Il fonctionne selon un schéma d'identité visuelle, et est considéré comme le cas le plus général de reconnaissance de l'écriture manuscrite. L'interprétation de l'information se fait indépendamment de sa source de génération. La reconnaissance en mode hors ligne est grandement influencée par la qualité des images numérisées [4].

Voici un aperçu concis des différences entre les approches en ligne et hors ligne, telles que présentées dans le tableau ci-dessous.

Critère de comparaison	En-ligne	Hors-ligne
-Outils d'acquisition	-Stylo électronique plus tablette graphique	-Scanner ou caméra
-Bruit d'image	-faible	-existence d'un bruit important
-Informations disponibles	la localisation, l'orientation du mouvement, les points de terminaison, les points de départ, l'ordre des traits	- l'absence de contexte informationnel.

TABLE 1.1 – Une brève comparaison entre les approches en-ligne et hors-ligne [3][4]

1.4 Reconnaissance D'écriture Manuscrite

La reconnaissance de l'écriture manuscrite vise principalement à déchiffrer les informations des documents numérisés et à les présenter de manière adéquate. Ce domaine de recherche passionnant a suscité un vif engouement, étant donné les nombreuses applications qui ont pu en tirer parti avec succès [5].

1.4.1 Reconnaissance de caractère

La reconnaissance optique de caractères englobe l'identification et l'interprétation de diverses formes d'informations textuelles ou symboliques, que ce soit du texte imprimé ou écrit à la main, y compris les équations mathématiques. Reconnaître le texte manuscrit est généralement considéré comme plus complexe que reconnaître le texte imprimé, car ce dernier bénéficie déjà de solutions commerciales bien établies. Les applications de ce domaine sont vastes et comprennent des systèmes capables de localiser, reconnaître et interpréter du texte dans des documents papier ou d'autres supports. En se limitant à un domaine spécifique, on peut considérablement réduire la complexité du système.

Les utilisateurs potentiels des systèmes de reconnaissance de l'écriture se trouvent principalement dans divers secteurs, tels que les organismes gouvernementaux qui uti-

lisent la lecture automatique des adresses et des impôts, ainsi que les grandes entreprises privées, notamment les banques, pour le traitement des chèques. Les différents domaines de la reconnaissance de caractères se chevauchent, il est donc essentiel de comprendre l'ensemble du domaine pour appréhender correctement un cas spécifique. Les applications de la reconnaissance de caractères varient en fonction de la nature de l'écriture et du support de saisie utilisé [6].

Il est essentiel de noter que la reconnaissance des caractères, qu'ils soient imprimés ou manuscrits, ainsi que des symboles tels que les formules mathématiques, implique la conversion du texte en une forme décodée. La Figure 1.1 présente une classification des différents types d'écriture en fonction du support de saisie, du support d'écriture et de l'application considérée. Bien que la distinction entre le texte imprimé et le texte manuscrit soit claire, il est important de souligner que les systèmes de reconnaissance de texte manuscrit doivent également être capables de reconnaître le texte imprimé. De plus, certaines polices imprimées peuvent imiter l'écriture manuscrite. Toutefois, en raison du bruit supplémentaire présent dans les caractères manuscrits par rapport à ceux imprimés, la reconnaissance du texte manuscrit est généralement considérée comme plus difficile. Heureusement, des systèmes commerciaux opérationnels existent déjà pour la reconnaissance du texte imprimé.

1.4.2 Les forme de text manuscrit

Il existe différentes formes de texte manuscrit :

Caractères isolés ou continus, (cursifs ou non cursifs) :

La reconnaissance des caractères isolés est une tâche simple qui implique la classification des caractères ou des chiffres individuels. Contrairement au texte continu, elle ne nécessite pas d'étape de segmentation où l'image d'entrée doit être préalablement divisée en unités distinctes. C'est à partir de cette étape de segmentation que le système d'identification peut traiter les caractères individuellement.

Texte imprime ou manuscrit :

Le texte imprimé est généralement considéré comme distinct et indépendant. Un exemple spécifique de texte continu est la reconnaissance de l'écriture cursive, qui pose des défis supplémentaires en termes de segmentation. Dans de tels cas, des techniques

telles que les modèles de Markov cachés (HMM) sont souvent utilisées. Les caractères non cursifs, également connus sous le nom de caractères isolés, sont également pris en compte lors de l'écriture [7].

Un texte peut prendre la forme d'une impression ou d'une écriture manuscrite, et il peut également être soit contraint, soit non contraint. Un texte contraint fait référence aux situations où l'utilisateur est guidé pour écrire dans des emplacements prédéfinis. Dans de tels cas, le système possède des informations préalables sur ce qu'il va reconnaître et où se trouve le texte. En revanche, un texte non contraint fait référence aux cas où l'utilisateur dispose d'une page blanche où il peut ajouter des informations dans n'importe quelle direction. En pratique, un texte non contraint se compose généralement de plusieurs lignes, sans restriction sur le nombre de mots dans chaque ligne [8].

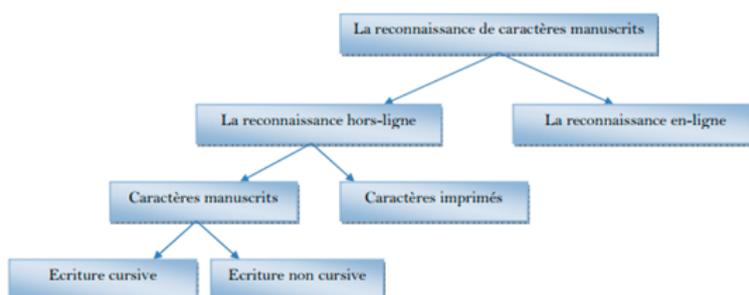


FIGURE 1.2 – Différentes formes du texte manuscrit [7]

Des exemples de texte contraints et non contraints sont présentés dans la figure [1.3].

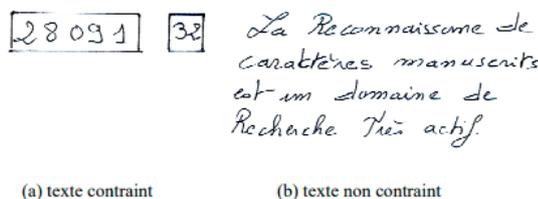


FIGURE 1.3 – Exemples de texte contraint et non contraint [8].

En outre, chaque langue possède ses propres particularités, comme le nombre de lettres et la façon dont elles sont écrites en un seul mot. Les recherches sur la reconnaissance de l'écriture manuscrite peuvent être classées en différentes catégories en fonction de l'unité de reconnaissance utilisée, qu'il s'agisse d'un caractère, d'un mot, d'une phrase, d'un chiffre ou même d'une partie d'un texte.

1.4.3 Reconnaissances des chiffres

La reconnaissance des nombres numériques présente des différences par rapport à la reconnaissance des chiffres individuels, car elle implique la classification des séquences de chiffres et la segmentation des chiffres qui se chevauchent. Contrairement à la reconnaissance de texte manuscrit, il y a peu d'informations contextuelles disponibles, ce qui signifie que n'importe quel nombre peut être suivi de n'importe quel autre nombre.

La segmentation des séquences de chiffres peut être complexe, car les chiffres peuvent se superposer ou être fragmentés. Les approches pour la reconnaissance des nombres manuscrits peuvent être classées en deux catégories : la segmentation suivie de la reconnaissance, ou la reconnaissance et la segmentation simultanées. Dans la première catégorie, un module de segmentation génère des hypothèses où chaque sous-séquence est considérée comme contenant un chiffre isolé pour le système de reconnaissance. Les approches basées sur la segmentation reposent sur des hypothèses probabilistes, et la décision finale vise à obtenir la meilleure reconnaissance segmentée à partir de l'image d'entrée.

En général, le système génère une série d'hypothèses à partir du module de segmentation, puis évalue chaque hypothèse en les identifiant, tout en prenant en compte les informations contextuelles. Cependant, cette approche présente un inconvénient majeur, car elle nécessite une comparaison exhaustive de toutes les hypothèses générées, ce qui demande un effort considérable. De plus, le module de reconnaissance doit être capable de différencier différents modèles, tels que les fragments de caractères, les caractères isolés et les caractères connectés, afin d'améliorer la précision du processus de reconnaissance [9].

1.5 Domaines D'application

La reconnaissance et la classification de chiffres manuscrits ont de nombreuses applications dans de nombreux domaines, notamment [10] :

- **Traitement de chèques et de documents financiers** : la reconnaissance de chiffres manuscrits est souvent utilisée pour traiter des chèques et d'autres documents financiers, ce qui peut être très utile pour les banques et les entreprises qui doivent traiter de grandes quantités de transactions.

- **Reconnaissance de plaques d'immatriculation** : la reconnaissance de chiffres manuscrits est également utilisée pour lire et reconnaître les numéros de plaque d'immatriculation sur les voitures, ce qui peut être utile pour les forces de l'ordre et les sociétés de péage autoroutier.

- **Systèmes de sécurité** : la reconnaissance de chiffres manuscrits est utilisée dans les systèmes de sécurité tels que les serrures à combinaison et les systèmes d'authentification basés sur des codes PIN.

- **Reconnaissance de caractères pour la lecture automatique** : la reconnaissance de chiffres manuscrits est souvent utilisée avec la reconnaissance de caractères pour la lecture automatique des adresses sur les enveloppes, des codes postaux et des numéros de téléphone sur les formulaires de candidature, etc.

1.6 Architecture Générale d'un système de Reconnaissance des Chiffres manuscrits

Les systèmes de reconnaissance des chiffres manuscrits suivent généralement les étapes principales suivantes : acquisition, prétraitement, segmentation, extraction de caractéristiques et éventuellement classification. La Figure 1.4 présente une représentation globale d'un système de reconnaissance des chiffres, démontrant ces étapes [11].

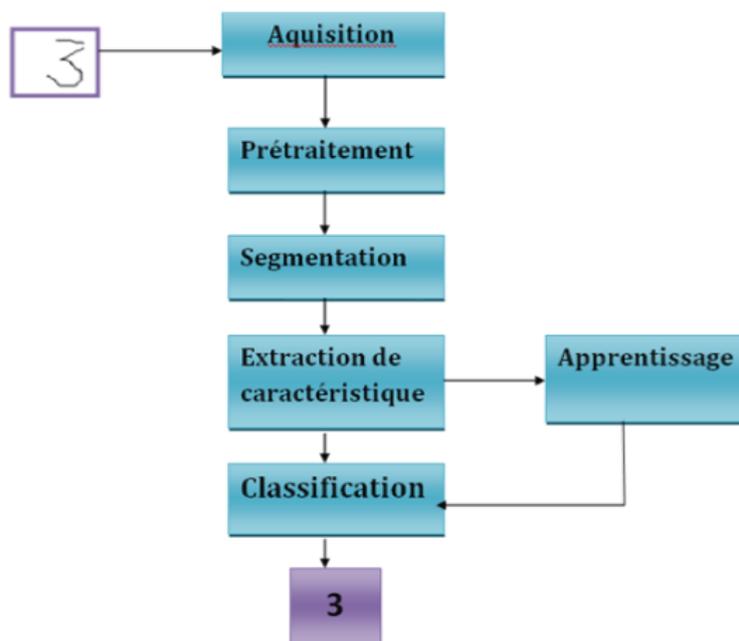


FIGURE 1.4 – Architecture de reconnaissance des chiffres manuscrits [11]

1.6.1 Phase d'acquisition

La phase d'acquisition comprend la capture d'une image de l'écriture manuscrite à l'aide de dispositifs physiques tels que des scanners ou des caméras numériques. Par la suite, cette image est convertie en une représentation numérique appropriée grâce à un système de traitement informatisé. Pendant cette étape, il est essentiel de réduire au minimum toute altération éventuelle du support papier. Cependant, malgré la qualité du système d'acquisition, des interférences indésirables peuvent survenir, entraînant des variations non désirées dans l'arrière-plan de l'image. Ces variations peuvent être dues à la texture du papier, aux conditions d'éclairage et à la zone de travail [12].

1.6.2 Phase de pré-traitement

Après l'acquisition, la plupart des systèmes utilisent une étape de prétraitement. Ces prétraitements ne sont pas spécifiques à la reconnaissance de l'écriture manuscrite, mais consistent en des étapes standard de traitement d'images. L'objectif du prétraitement est de préparer l'image pour une analyse ultérieure. En essence, il vise à réduire le bruit pou-

vant altérer les données et à conserver autant que possible les informations pertinentes du document. Le bruit peut provenir de divers facteurs tels que les conditions d'acquisition (éclairage, mauvaise position du document, etc.) ou la qualité du document original. Les opérations de prétraitement peuvent inclure le lissage, la squelettisation et la normalisation, qui visent à améliorer la qualité de l'image avant de la soumettre à l'étape d'analyse.

La binarisation

La phase de binarisation revêt une importance cruciale dans la reconnaissance des chiffres manuscrits. Son objectif est de convertir une image en niveaux de gris en une image binaire composée uniquement de pixels noirs et blancs. Cette étape est souvent utilisée pour séparer les chiffres du reste de l'image et améliorer la précision de la reconnaissance. Toutefois, lorsque les images présentent un faible contraste ou un contraste variable, il peut s'avérer difficile de déterminer un seuil précis pour la binarisation. Une méthode couramment employée consiste à calculer l'histogramme des niveaux de gris de l'image. Le seuil est ensuite déterminé en identifiant la valeur de gris située dans la vallée entre les deux pics de l'histogramme. Les pixels dont le niveau de gris est supérieur à ce seuil sont considérés comme faisant partie du fond de l'image, tandis que ceux dont le niveau de gris est inférieur appartiennent à l'objet, dans ce cas, les chiffres [13].

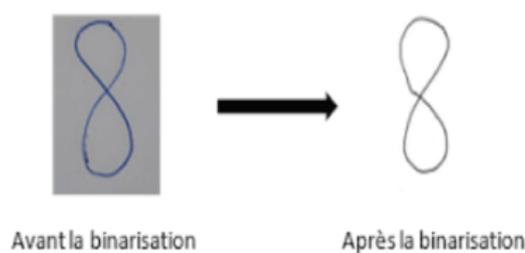


FIGURE 1.5 – Exemple de la binarisation

Lissage

Le lissage est une opération de modification locale qui examine les pixels voisins d'un pixel donné et lui assigne la valeur 1 si le nombre de pixels noirs dans cette région dépasse un seuil prédéfini. Souvent, les images peuvent contenir du bruit en raison de la qualité du document ou de l'appareil de capture, ce qui peut entraîner des pixels manquants ou des pixels indésirables. Les techniques de lissage sont utilisées pour résoudre ces problèmes [14].

La normalisation

Cette étape a pour objectif de normaliser l'image des chiffres manuscrits en les ramenant à une taille standard. La normalisation peut entraîner de légères distorsions sur l'image, mais elle est nécessaire dans certains systèmes et méthodes de reconnaissance qui sont sensibles aux variations de taille et de position, comme les méthodes d'appariement de motifs. En fonction de la qualité du document et du type d'écriture (manuscrite ou imprimée), l'image binaire peut présenter des traits de différentes épaisseurs, peu définis et généralement avec du bruit. Pour simplifier l'image du caractère et faciliter son traitement, une opération appelée squelettisation est appliquée. Cette opération réduit l'image à une "ligne" représentant le tracé du caractère, où les points du tracé correspondent généralement aux lignes centrales des points initiaux. La squelettisation, également connue sous le nom d'amincissement, est une étape essentielle dans de nombreux systèmes de reconnaissance d'écriture manuscrite [15].

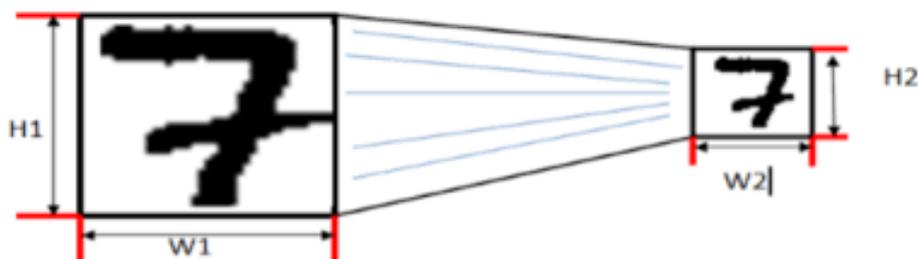


FIGURE 1.6 – Exemple d'un chiffre manuscrit normalisé [15]

Squelettisation

Selon la qualité du document et le type d'écriture utilisé (manuscrite ou imprimée), l'image binaire représente une série de traits de différentes épaisseurs, qui peuvent être flous et comporter du bruit. Afin de simplifier le caractère et faciliter son traitement ultérieur, une opération appelée squelettisation est appliquée. Cette opération transforme l'image du caractère en une représentation composée de lignes, ce qui simplifie sa manipulation. Les points du tracé résultant correspondent généralement aux lignes centrales des points d'origine. La squelettisation, également connue sous le nom d'amincissement, joue un rôle crucial dans de nombreux systèmes de reconnaissance d'écriture manuscrite en réduisant la forme du caractère à une épaisseur extrêmement mince, équivalente à un pixel. Cette étape est essentielle pour faciliter l'analyse et le traitement précis des caractères manuscrits, contribuant ainsi à l'amélioration des performances des systèmes de reconnaissance d'écriture [16].

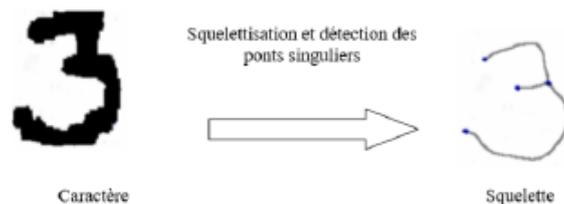


FIGURE 1.7 – Exemple de Squelettisation [16]

1.6.3 Phase de segmentation

La phase de segmentation est cruciale dans le système de reconnaissance des chiffres manuscrits, car elle consiste à extraire des zones ou segments distincts de l'image. Une segmentation précise garantit des segments appropriés et, par conséquent, un taux de reconnaissance élevé [17]. En revanche, une segmentation incorrecte entraînera une diminution du taux de reconnaissance. La Figure 1.8 présente une illustration de la segmentation.

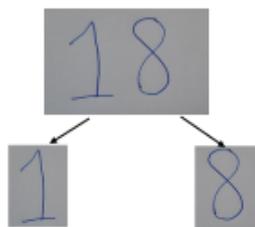


FIGURE 1.8 – Exemple de segmentation [17]

Les approches de segmentation

Il existe divers éléments qui rendent la segmentation plus complexe, tels que la diversité du style et des outils d'écriture utilisés. Cela ajoute de la complexité et rend l'analyse plus difficile. En règle générale, il existe deux approches différentes pour la segmentation : l'approche implicite et l'approche explicite.

Segmentation discrète : Dans l'approche discrète, qui est également connue sous le nom d'approche explicite, une étape cruciale consiste à identifier les points de segmentation les plus probables. Certains éléments, comme les connexions multiples ou les contacts prolongés, nécessitent une séparation des composantes en suivant un chemin spécifique. Les chemins de segmentation sont généralement déterminés en utilisant des caractéristiques distinctives obtenues à partir de l'analyse des contours de la forme, du squelette ou de l'amincissement de l'arrière-plan, d'une analyse bidimensionnelle du tracé, ou d'une combinaison d'analyse des contours et d'amincissement de l'arrière-plan [18].

Segmentation implicite : Pour résoudre de manière efficace et complexe le problème du choix des points de segmentation, il existe des approches novatrices appelées "segmentation implicite" ou "segmentation continue". Ces méthodes considèrent tous les points de la trajectoire comme des candidats potentiels pour la segmentation.

Dans cette approche, la segmentation et la reconnaissance sont réalisées simultanément, ce qui permet d'utiliser le terme "segmentation-reconnaissance intégrée". Ces méthodes utilisent des fenêtres glissantes de taille fixe qui balayent la séquence de chiffres, ce qui permet l'extraction de caractéristiques à un niveau bas. Les fenêtres peuvent être ana-

lysées à l'aide d'un classifieur traditionnel ou de modèles dynamiques. L'objectif de cette approche n'est pas seulement de séparer explicitement les chiffres, mais également de les incorporer de manière implicite dans le module de reconnaissance.

Grâce à cette approche innovante, il est possible de résoudre efficacement le problème de segmentation tout en réalisant la reconnaissance des chiffres. Cette méthode offre une solution complète et intégrée pour traiter ces deux aspects essentiels [19].

Les méthodes de la segmentation

1. Segmentation de la page : À cette étape, l'objectif est de localiser et classer les zones homogènes d'information sur la page en fonction de leur contenu (texte ou non-texte). Cela permet de séparer les éléments graphiques et photographiques du processus de reconnaissance de texte [20].

2. Segmentation d'un bloc de texte en lignes : L'étape de segmentation des lignes vise à diviser le texte en lignes distinctes, afin d'isoler les mots et les caractères qui les composent. La séparation des lignes est réalisée en analysant l'histogramme des projections horizontales du bloc de texte, ce qui permet de repérer les espaces entre les lignes et de les séparer les unes des autres [21].

3. Segmentation des lignes en mots : La segmentation des mots est effectuée en examinant l'histogramme des projections verticales des lignes, ce qui permet de repérer les espaces entre les mots et de les séparer. Toutefois, cette méthode peut ne pas être suffisamment efficace dans des situations où les mots se superposent, comme dans l'écriture arabe. Dans de tels cas, d'autres techniques sont employées, telles que le suivi des contours, l'extraction du squelette ou la détection des composantes connexes, afin d'assurer une segmentation précise [22].

4. Segmentation des caractères : La segmentation des caractères consiste à découper une image contenant une séquence de caractères (un mot) en sous-images représentant des symboles individuels. C'est une étape essentielle dans les systèmes de reconnaissance automatique, car elle vise à déterminer si un motif isolé dans une image (un caractère ou une autre entité identifiable dans le mot) est correct ou non [23].

1.6.4 Phase d'extraction de caractéristiques

Lorsqu'on parle de reconnaissance des chiffres manuscrits, l'accent est mis sur l'extraction des caractéristiques les plus pertinentes qui seront utilisées par le système. Cette étape revêt une grande importance dans la conception d'un tel système, mais elle présente également des défis en raison du risque de perte d'informations lié à de nombreuses techniques d'extraction. Il est donc crucial de trouver un équilibre entre la quantité et la qualité des informations extraites afin d'optimiser les résultats [24].

Réduire le nombre de caractéristiques présente plusieurs bénéfices : cela facilite la visualisation et la compréhension des données, diminue le temps nécessaire pour l'apprentissage et la classification des systèmes, améliore les performances de classification et réduit la taille des bases de données d'apprentissage. Les caractéristiques utilisées peuvent être classées en quatre catégories principales : structurelles, statistiques, topologiques et métriques [29].

Les Primitives structurelle

Les primitives globales représentent des propriétés déduites à partir de l'ensemble des pixels d'une image et d'une transformation globale de l'image. Elles englobent différentes caractéristiques qui se basent sur la distribution des pixels, comprenant les moments invariants, les projections et les profils. Ces attributs sont essentiels pour analyser et comprendre les caractéristiques distinctes d'une image [25].

Les techniques de base pour représenter un caractère incluent la conversion de son contour ou de son squelette en une séquence de codes directionnels, communément appelée code de Freeman. Il existe différentes méthodes de transformation, telles que les transformations de Hough et de Fourier, ainsi que les transformations utilisant des ondelettes. Ces approches offrent des moyens efficaces pour analyser et interpréter les caractères, facilitant ainsi la reconnaissance et le traitement des textes.

Les Primitives statistiques

Les primitives statistiques fournissent des données sur la distribution des pixels dans l'image d'un caractère ou d'un chiffre, permettant ainsi de décrire sa forme à l'aide d'un ensemble de mesures extraites de cette forme [26]. Les caractéristiques exploitées lors de

la reconnaissance des chiffres manuscrits englobent

- Le zonage (ou zoning) : Cette approche implique la partition de l'image en régions ou zones distinctes. L'image initiale est une image binaire. Pour chaque région obtenue, on évalue la moyenne ou le pourcentage de pixels noirs présents.
- L'histogramme est un moyen courant et simple de représenter le nombre de pixels sur chaque ligne ou colonne de l'image. Il existe différentes variations d'histogrammes, tels que l'histogramme directionnel ou l'histogramme des transitions. Un exemple classique est le calcul de l'histogramme.

Une autre approche statistique possible est l'utilisation de la moyenne des pixels situés dans un masque rectangulaire. En créant une matrice de masque qui couvre l'intégralité de la forme, il devient possible de représenter statistiquement la forme en utilisant un nombre restreint de valeurs correspondant à chaque masque.

Les primitives globales

Les primitives globales reposent sur une transformation globale de l'image, ce qui implique qu'elles prennent en compte tous les pixels de l'image. Elles sont dérivées de la répartition des pixels et se classent en trois catégories principales : les moments invariants, les projections et les profils [27].

Parmi les différentes méthodes de transformation, il existe une approche simple qui consiste à représenter le contour ou le squelette d'un caractère en utilisant un ensemble de codes de directions (appelé code de Freeman). En plus de cela, d'autres techniques de transformation couramment utilisées incluent les transformées de Hough, de Fourier et d'ondelettes.

1.6.5 Phase Classification

La classification implique d'associer une forme spécifique à une classe prédéterminée. Ce processus englobe à la fois l'apprentissage et la prise de décision. En utilisant une description standardisée de la forme en termes de paramètres, ces processus visent tous les deux à attribuer la forme à un modèle de référence..

Le résultat de l'apprentissage peut se traduire soit par une réorganisation ou un renforcement des modèles existants en tenant compte de la nouvelle configuration, soit par la

création d'un nouveau modèle grâce au processus d'apprentissage [28].

1.7 Conclusion

Ce chapitre a abordé de manière générale les différentes étapes d'un système de reconnaissance de chiffres manuscrits, allant du prétraitement aux approches de classification. Dans le prochain chapitre, nous examinerons en détail les différents types d'apprentissage.

Machine Learning

2.1 Introduction

De nos jours, l'intelligence artificielle (IA) est l'une des avancées technologiques les plus médiatisées. Elle englobe un domaine de recherche scientifique visant à recréer le raisonnement humain en utilisant le concept d'apprentissage. Grâce à l'accès à des volumes massifs de données et à la puissance de calcul des ordinateurs, diverses formes d'IA ont émergé, notamment l'apprentissage automatique (machine learning) et l'apprentissage profond (deep learning ou DL). L'apprentissage automatique implique l'utilisation d'algorithmes automatisés qui apprennent à modéliser des fonctions et à prédire des actions futures à partir de données structurées. En revanche, l'apprentissage profond se réfère à un réseau de neurones artificiels qui imite les réseaux neuronaux du cerveau humain.

Dans ce chapitre, nous présentons un état de l'art sur le machine Learning. Tout d'abord, nous clarifions ce qui relève de l'apprentissage automatique (machine Learning) en commençant par définir précisément ce qu'est l'apprentissage automatique.

Ensuite, nous expliquons les différentes approches de l'apprentissage automatique, c'est-à-dire les types d'approches utilisés. De plus, nous détaillons les différents algorithmes populaires pour modéliser un modèle de machine Learning, en mettant en évidence leurs principaux avantages et limitations respectifs.

Enfin, nous abordons les différentes méthodes d'évaluation d'un modèle de machine Learning, ainsi que les différentes métriques utilisées pour évaluer les performances du modèle.

2.2 Intelligence artificielle

L'intelligence artificielle (IA) concerne la conception et l'évolution de systèmes informatiques capables d'effectuer des tâches qui nécessitent habituellement l'intelligence humaine, comme la capacité de voir, de reconnaître la parole, de prendre des décisions et de traduire entre différentes langues [29].

En d'autres mots, l'intelligence artificielle permet aux machines de fonctionner et de se comporter de manière similaire aux êtres humains. Dans la suite de cet article, nous nous concentrerons particulièrement sur l'apprentissage automatique, une discipline de l'intelligence artificielle..

L'intelligence artificielle est une technologie omniprésente qui se déploie dans de nombreux domaines. Pour répondre de manière plus efficace aux différents problèmes de la société, elle se divise en plusieurs branches, chacune étant dédiée à un problème spécifique. Les principales branches de l'intelligence artificielle comprennent les systèmes experts, la robotique, l'apprentissage automatique, les réseaux neuronaux, la logique floue et le traitement du langage naturel [30].

Parmi ces branches, nous nous concentrerons sur l'apprentissage automatique (Machine Learning).

2.3 Machine Learning

2.3.1 Définition du Machine Learning

Le Terme " Machine Learning " ou ML en bref, a été inventé en 1959 par Arthur Samuel dans le Contexte de la résolution de jeux de dames par machine.

Le terme fait référence à un programme informatique qui peut apprendre à produire un comportement qui n'est pas explicitement programmé par l'auteur du programme [29]. Il est plutôt capable de montrer un comportement dont l'auteur peut ne pas être complètement conscient [31]. Ce comportement est acquis en se basant sur trois éléments distincts :

- Les données qui sont consommées par le programme.
- On utilise une métrique pour évaluer l'écart ou la distance entre le comportement actuel et le comportement souhaité.

- Un mécanisme de rétroaction est employé en utilisant l'erreur quantifiée afin de guider le programme dans la production d'un comportement amélioré lors des événements ultérieurs.

Comme il est évident, les deuxième et troisième facteurs rendent rapidement le concept abstrait et mettent en évidence ses fondements mathématiques profonds. Les méthodes de la théorie de l'apprentissage automatique jouent un rôle essentiel dans la construction de systèmes artificiellement intelligents[32] .

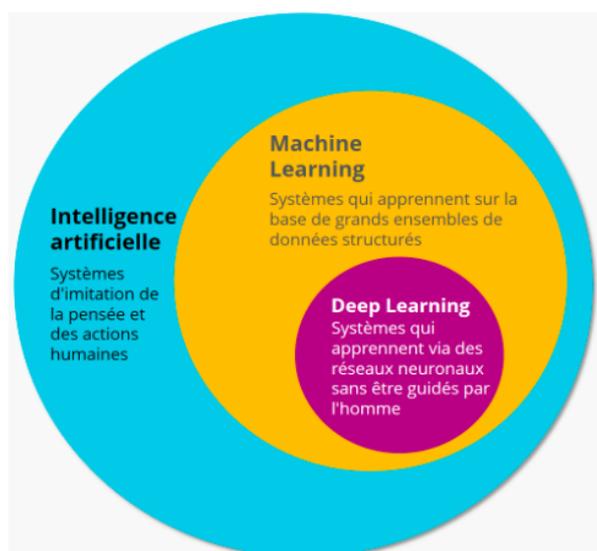


FIGURE 2.1 – Deep Learning vs Machine Learning [38]

2.3.2 Types du Machine Learning

Les approches d'apprentissage en profondeur peuvent être regroupées en différentes catégories : supervisées, semi-supervisées ou partiellement supervisées, et non supervisées [33].

De plus, il existe une autre catégorie d'approche d'apprentissage appelée renforcement (Reinforcement Learning ou RL) ou renforcement profond (Deep RL ou DRL), qui est souvent abordée dans le contexte de l'apprentissage semi-supervisé ou parfois dans le cadre des approches d'apprentissage non supervisées.

Les différents types d'apprentissage automatique sont utilisés dans une variété de domaines et de situations en fonction des besoins spécifiques et des données disponibles,

Cette représentation est illustrée dans la figure [2.2].

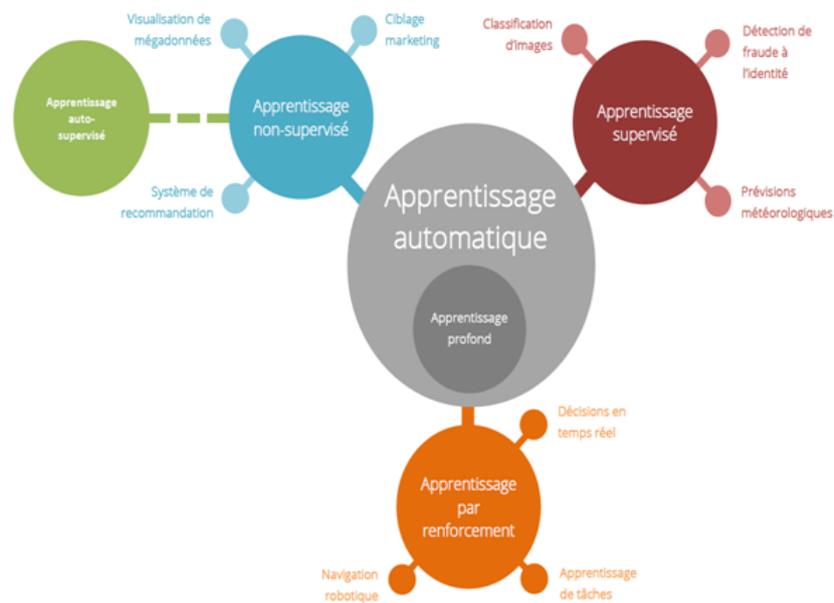


FIGURE 2.2 – Types du ML [32]

L'Apprentissage Supervise

Dans cette méthode d'apprentissage, les données d'entrée et de sortie sont fournies à l'avance. Le réseau examine ensuite les données d'entrée et compare ses prédictions aux sorties souhaitées. Les poids du réseau sont ensuite ajustés en fonction des erreurs détectées et propagées à travers le système.

Ce processus est répété de manière itérative jusqu'à ce que les poids continuent à s'améliorer. L'ensemble de données utilisé pour l'apprentissage est communément appelé ensemble d'apprentissage [31]. La Figure 2.3 présente une illustration de la notion d'apprentissage automatique.

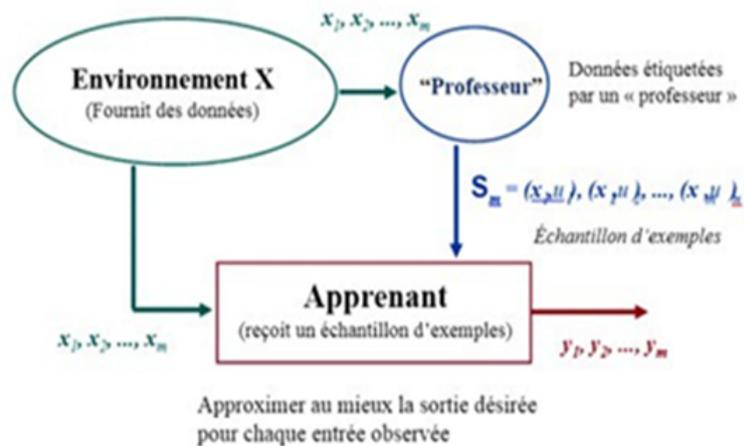


FIGURE 2.3 – Schema d'Apprentissage Supervisé [32]

L'apprentissage supervisé résout deux problèmes :

Régression : Les techniques de régression utilisent des données d'entraînement pour prédire une seule valeur de sortie. Par exemple, nous pouvons utiliser la régression pour prédire les prix des maisons à partir de données d'entraînement. Les variables d'entrée peuvent être l'emplacement, la taille de la maison, etc [33].

Classification : La classification consiste à regrouper la sortie dans une classe. Lorsqu'un algorithme tente de classer l'entrée dans deux classes différentes, on parle de classification binaire. Lorsqu'il s'agit de choisir parmi plus de deux classes, on parle de classification multi- classe [34].

Apprentissage Semi-Supervisé profond

L'apprentissage semi-supervisé est une méthode d'apprentissage qui se déroule sur la base d'ensembles de données qui sont partiellement étiquetés. Dans certains cas, des techniques telles que le renforcement par apprentissage profond (DRL) et les réseaux adversaires génératifs (GAN) sont utilisées comme approches d'apprentissage semi-supervisé [35].

Apprentissage Non Supervisé

Dans l'apprentissage non supervisé, les entrées sont fournies au réseau, mais les sorties souhaitées ne le sont pas. Le système doit alors prendre la décision de regrouper les données d'entrée en utilisant ses propres mécanismes. Ce processus est souvent appelé auto-organisation ou adaptation [34].

Apprentissage Par Renforcement

L'apprentissage par renforcement consiste à inviter le système à prendre n'importe quelle action et à recevoir une récompense ou une pénalité en fonction de l'adéquation de cette action à la situation donnée [37].

Le système doit apprendre quelles actions rapportent le plus de récompenses dans différentes situations au fil du temps.

Ces techniques ont largement contribué à l'épanouissement du domaine de l'apprentissage automatique, en particulier dans les domaines de la vision par ordinateur et de l'analyse de texte. Au fil des années, de nombreux modèles ont été introduits pour mettre en œuvre ces techniques d'apprentissage automatique, tels que les réseaux de neurones artificiels (inspirés du fonctionnement des neurones du cerveau), les arbres de décision (utilisant des structures en forme d'arbre pour modéliser les décisions et les résultats), les modèles de régression (utilisant des méthodes statistiques pour cartographier les variables d'entrée et de sortie) ,etc [35].

Vers 2010, plusieurs événements ont eu un impact sur la recherche en apprentissage automatique :

- La disponibilité accrue de la puissance de calcul a facilité l'évaluation des modèles plus complexes.
- Les coûts de traitement et de stockage des données ont diminué, rendant plus accessible la manipulation de grandes quantités de données.
- Notre compréhension du fonctionnement du cerveau naturel s'est améliorée, ce qui nous a permis de développer de nouveaux algorithmes d'apprentissage automatique basés sur ces connaissances .

Le tableau 2.1 présente les distinctions entre les quatre catégories de l'apprentissage automatique :

Apprentissage Supervise	Apprentissage Non Supervise	Apprentissage Semi- Supervise	Apprentissage Par Renforcement
<ul style="list-style-type: none"> - Les experts de Données fournissent des entrées, des sorties et des commentaires pour construire le modelé (comme définition) 	<ul style="list-style-type: none"> - utiliser l'apprentissage profond pour arriver à des conclusion et à des données de formation étiquetées. 	<ul style="list-style-type: none"> - Construit un Modèle à travers un mélange de données étiquetées et non étiquetées,un ensemble de catégories, de suggestions et exemples d'étiquettes 	<ul style="list-style-type: none"> Auto-interpretation Mais basé sur un système récompense de punitions par essais et erreurs,recherche d'une récompense maximale
EXEMPLE D'ALGORITHMES	EXEMPLE D'ALGORITHMES	EXEMPLE D'ALGORITHMES	EXEMPLE D'ALGORITHMES
<ul style="list-style-type: none"> Régressions linéaires - prévisions des ventes - évaluation des risques Machine à vecteurs de support - classification des images - comparaison des performances financières Arbre de décision - analyse prédictive - tarification 	<ul style="list-style-type: none"> Apriori - fonctions de ventes -associations de mots - searcher Clustering k-means - surveillance des performances - intention du chercheur 	<ul style="list-style-type: none"> Création de données - réseaux antagonistes génératifs - manipulation audio et vidéo Naive Bayes autodidacte classificateur - traitement du langage naturel 	<ul style="list-style-type: none"> Q-Learning - politique création - diminution de la consommation Valeur basée sur un modèle estimation - tâches linéaires - estimation des paramètres

TABLE 2.1 – présente les différences entre les quatre types Machine Learning [36][38]

2.4 Les Algorithmes De Machine Learning

L'apprentissage automatique désigne la capacité des machines à prendre des décisions en se basant sur des expériences passées acquises grâce à l'apprentissage. Les algorithmes d'apprentissage automatique sont divisés en deux catégories : supervisés et non supervisés, en fonction de la nature de la tâche à accomplir. Les algorithmes supervisés sont utilisés pour la classification, la régression et les prédictions, tandis que les approches non supervisées sont employées pour des applications de regroupement, de détection d'anomalies, etc. Ci-dessous, nous décrirons certains des algorithmes les plus couramment utilisés.

2.4.1 Les arbres de Décision

Dans les années 70 et au début des années 80, J. Ross Quintana a développé un algorithme novateur appelé ID3 (Iterative dichotomiser). Les arbres de décision, qui sont des modèles d'apprentissage automatique, adoptent une approche géométrique pour diviser l'espace des caractéristiques en segments distincts.

L'algorithme fonctionne en utilisant une série de classifieurs faibles, également connus sous le nom de "classifieurs faibles", qui sont progressivement introduits pour diviser l'espace des caractéristiques et classer les données. Ces classifieurs sont ensuite organisés sous forme d'un réseau hiérarchique de nœuds interconnectés.

À la fin du processus, un arbre inversé est obtenu, où le nœud supérieur représente le classifieur possédant la plus grande capacité de classification parmi tous les classifieurs. En descendant dans l'arbre, les autres classifieurs faibles sont représentés dans un ordre décroissant d'importance.

Cet algorithme offre ainsi une méthode efficace pour la classification des données en utilisant des classifieurs faibles successifs. Son approche géométrique permet de diviser l'espace des caractéristiques de manière claire et distincte, offrant ainsi une solution précise et fiable[40].

Avantages et Inconvénients Des Arbres Décision

Les Arbres De Décision	
Avantages	Inconvénients
Ils offrent la possibilité de visualiser de manière graphique le processus de prise de décision.	Ils ont une performance plus mau-vaise statique que d'autres techniques d'apprentissage statistique.
Facile à interpréter pour toute personne	Ils ont été peu utilisées jusu à l'apparition d'une nouvelle techniques(forets aléatoires)
Il est considéré comme un miroir du processus humain de prise de décision	Son apprentissage peu amener arbres de décisions plus complexes qui généralise mal l'ensemble d'apprentissage

TABLE 2.2 – Les Avantages et Inconvénients Des Arbres De Décision[40]

Les arbres de décision présentent certaines limites :

- Ils peuvent devenir complexes et ne généralisent pas bien, ce qui peut entraîner un sur apprentissage (overfitting).
- Certains concepts peuvent être difficiles à apprendre pour les arbres de décision. Ils ne sont pas adaptés pour exprimer des concepts tels que XOR, par exemple.

2.4.2 L'algorithme des k plus proches voisins

La méthode des K plus proches voisins se démarque des autres approches de classification telles que les arbres de décision, les réseaux de neurones et les algorithmes génétiques. Elle repose sur une logique basée sur les cas déjà résolus, offrant ainsi une simplicité remarquable dans le domaine de l'apprentissage automatique supervisé. Son principe est de rechercher des cas similaires qui ont déjà été résolus et stockés en mémoire, ce qui permet de prendre des décisions éclairées. Cette méthode est reconnue pour son efficacité extrême et représente un outil puissant pour résoudre une variété de problèmes de classification[41].

Principe de classifieur KPPV

Lorsqu'il s'agit de classifier une nouvelle instance à partir d'une base d'apprentissage d'images, il existe une méthode appelée "classifieur des K plus proches voisins" (KPPV). Cette méthode consiste à rechercher les K cas les plus similaires à la nouvelle instance et à prédire la classe la plus fréquente parmi ces K voisins. Pour utiliser cette méthode, il est nécessaire de définir deux paramètres : le nombre K, qui détermine le nombre de voisins à prendre en compte, et une fonction de similarité permettant de comparer la nouvelle instance aux cas déjà classés [51]. En résumé, le fonctionnement de cette méthode peut être formulé de la manière suivante :

1. Un entier k est choisi (généralement $k = \text{nombre d'attributs} + 1$, ou $k = \sqrt{n}$).
2. Les distances sont calculées (par exemple, la distance euclidienne est largement utilisée).
3. Sélectionner les k observations ayant les distances les plus petites (les k plus proches voisins d'un cas donné).
4. Déterminer le nombre d'occurrences des k observations dans chaque classe respective (calculer les classes correspondantes).
5. Sélectionner la classe qui est la plus fréquemment représentée.

Les étapes pour classer une forme inconnue x sont les suivantes :

Lorsqu'il s'agit de classer une forme inconnue x , nous procédons comme suit : nous évaluons la distance entre x et tous les échantillons qui représentent les différentes classes, puis nous choisissons les K échantillons les plus proches. Ensuite, nous attribuons à x la classe majoritaire parmi ces K échantillons. Dans l'exemple de la Figure 2.3, si nous utilisons cinq voisins, x serait classé dans la classe w_1 [42].

Choix du nombre K :

Le paramètre K dans la règle des KPPV a un impact direct sur la nature du classifieur obtenu. Lorsque K est de petite valeur, le classifieur sera plus précis en définissant des frontières complexes entre les classes, mais il sera également plus sensible au bruit présent dans les échantillons et dans le vecteur à classer. En revanche, lorsque K est de grande valeur, le classifieur aura un comportement inverse en lissant les frontières entre les classes, ce qui réduira sa sensibilité au bruit [43].

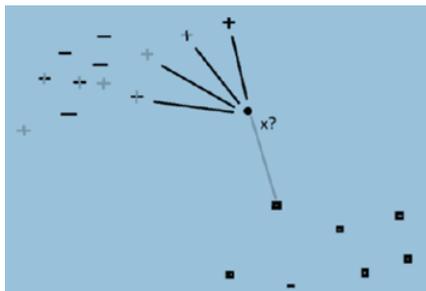


FIGURE 2.4 – Méthode de k plus proches voisins [29]

• Définition de la distance

Le choix de la mesure de distance est essentiel pour assurer le bon fonctionnement de la méthode, bien que les distances les plus simples puissent produire des résultats satisfaisants.

Une distance doit respecter quatre propriétés pour tous les vecteurs a , b et c [44] :

$$D(a, b) \geq 0 \text{ (non-négativité)}$$

$$D(a, b) = 0 \text{ si et seulement si } a = b \text{ (réflexivité)}$$

$$D(a, b) = D(b, a) \text{ (symétrie)}$$

$$D(a, b) + D(b, c) \geq D(a, c) \text{ (triangle d'inégalité)}$$

Il est intéressant de noter qu'un point "a" peut avoir un voisin "b" plus proche, alors que "b" peut avoir plusieurs voisins plus proches que "a". Ce choix de distance dépend des connaissances préalables du problème. Il est possible d'explorer différentes options de distance et d'estimer l'erreur réelle pour chaque choix afin de sélectionner la distance optimale. Plusieurs types de distances peuvent être utilisés à cette fin.

- La distance euclidienne.
- La distance "city block" ou "Manhattan" (somme des valeurs absolues).
- La distance de Tchebycheff.
- La distance de Mahalanobis.

Dans notre situation, nous avons opté pour l'utilisation de la distance euclidienne, qui est la plus couramment utilisée. Sa formule en "d" dimensions est la suivante :

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

- **Sélection de la classe**

Une méthode courante pour résoudre ce problème consiste à utiliser l'approche du vote majoritaire pondéré, également connue sous le nom de méthode du plus proche voisin pondéré (1-PPVP). Contrairement à la méthode du plus proche voisin (1-NN), qui se base uniquement sur le voisin le plus proche, la méthode du plus proche voisin pondéré prend en compte plusieurs voisins les plus proches.

L'idée principale derrière cette approche est d'attribuer un poids à chaque voisin en fonction de sa proximité avec le cas à classer. Ce poids, souvent appelé noyau, est calculé en inversant la distance entre le cas à classer et les voisins les plus proches. Ainsi, les voisins les plus proches ont un poids plus élevé, tandis que les voisins plus éloignés ont un poids plus faible [44].

Pour combiner les classes des voisins, on utilise le vote majoritaire pondéré, où chaque vote est pris en compte en fonction de son poids. Cela signifie que les voisins proches ont une influence plus importante sur la décision finale que les voisins plus éloignés. Cette approche permet d'obtenir de bons résultats même lorsque les points ne sont pas regroupés en densités de classe identiques.

De plus, il est possible de calculer une mesure de confiance pour la classe attribuée en utilisant le ratio entre le nombre de votes pour la classe majoritaire et le total des votes de chaque classe des voisins sélectionnés. Cette mesure de confiance permet d'évaluer la fiabilité de la classification effectuée.

En résumé, la méthode du plus proche voisin pondéré offre une approche plus robuste pour la reconnaissance de motifs dans les problèmes de classification. En prenant en compte plusieurs voisins et en attribuant des poids aux votes, cette méthode permet d'obtenir de meilleurs résultats, même dans des situations complexes où les classes ne sont pas parfaitement séparées. De plus, la mesure de confiance associée permet d'évaluer la fiabilité des résultats obtenus..

Avantages de l'algorithme des k plus proches voisins (k-NN) [43]

Simplicité : L'algorithme k-NN est relativement simple à comprendre et à mettre en œuvre. Il ne nécessite pas de paramètres complexes à ajuster ou de calculs intensifs.

Adaptabilité aux données : k-NN peut être utilisé avec différents types de données, qu'il s'agisse de données numériques, catégoriques ou mixtes. Il peut également être utilisé pour la classification et la régression.

Pas d'apprentissage initial : Contrairement à de nombreux autres algorithmes d'apprentissage supervisé, k-NN n'a pas besoin d'une phase d'apprentissage initiale. Les données d'entraînement sont simplement stockées et utilisées pour les classifications ultérieures.

Robustesse aux valeurs aberrantes : En utilisant une majorité de voisins, k-NN est relativement robuste aux valeurs aberrantes ou aux erreurs de mesure dans les données.

Inconvénients de l'algorithme des k plus proches voisins (k-NN) [43]

Sensibilité à la dimensionnalité : À mesure que le nombre de dimensions des données augmente, l'algorithme k-NN peut avoir du mal à trouver des voisins pertinents et peut souffrir de la malédiction de la dimensionnalité.

Coût computationnel : Lorsqu'il est utilisé avec de grandes bases de données, l'algorithme k-NN peut être coûteux en termes de temps de calcul, car il nécessite la recherche et le calcul des distances entre les échantillons.

Choix du paramètre k : Le choix de la valeur de k peut avoir un impact significatif sur les performances de l'algorithme. Une valeur trop petite peut conduire à une sensibilité excessive au bruit, tandis qu'une valeur trop grande peut entraîner une perte d'informations locales.

Sensibilité aux données déséquilibrées :

Lorsque les classes dans les données sont déséquilibrées, c'est-à-dire qu'il y a beaucoup plus d'échantillons d'une classe que d'autres, k-NN peut avoir tendance à favoriser la classe majoritaire.

2.4.3 SVM (Support Vector Machine)

Le Support Vector Machine (SVM) est l'un des algorithmes les plus couramment utilisés en apprentissage supervisé, tant pour la classification que pour la régression. Ce-

pendant, il est principalement employé pour résoudre des problèmes de classification dans le domaine de l'apprentissage automatique.

L'objectif de l'algorithme SVM est de découvrir l'hyperplan optimal, c'est-à-dire la meilleure frontière de décision, qui peut séparer de manière efficace l'espace de données à n dimensions en différentes classes. Cette frontière de décision, appelée hyperplan, permettra de classifier facilement de nouveaux points de données dans les catégories appropriées à l'avenir.

L'algorithme SVM identifie les points ou vecteurs extrêmes qui sont essentiels pour déterminer l'hyperplan. Ces points extrêmes sont connus sous le nom de vecteurs de support, ce qui explique pourquoi l'algorithme est appelé Support Vector Machine. Pour mieux comprendre, examinons le diagramme (Figure 2.5) ci-dessous, qui représente deux catégories distinctes séparées par une frontière de décision ou un hyperplan [45].

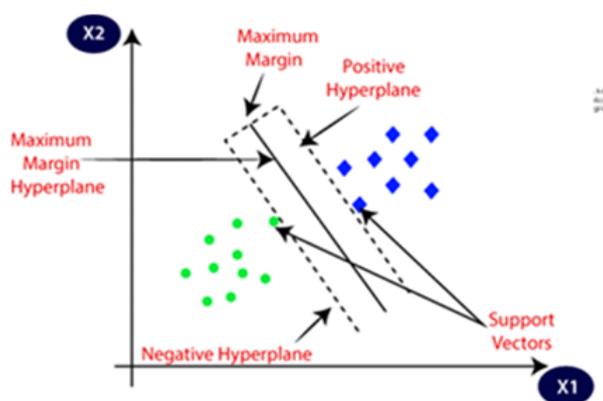


FIGURE 2.5 – Hyperplan Optimal Avec Une Marge Maximale[45]

Comme nous l'avons vu précédemment, nous choisirons Comme évoqué précédemment, notre objectif est de sélectionner l'hyperplan qui permet de maximiser la marge, c'est-à-dire la plus petite distance entre les vecteurs d'entraînement et cet hyperplan. Ces vecteurs qui se trouvent à la distance minimale sont couramment désignés par le terme de **vecteurs supports**.

Avantages et Inconvénients de Support Vecteur Machine

Support Vecteur Machine	
Avantages	Inconvénients
Capacité à traiter de grandes dimensionnalités (variables élevées)	Problème lorsque les classes sont bruitées (multiplication des points supports)
Non paramétrique	Pas de modèle explicite pour les noyaux non linéaires (utilisation des points supports)
Traitement des problèmes non linéaires avec le choix des noyaux	Difficulté d'interpréter (ex. pertinence des variables)
Traitement des problèmes non linéaires avec le choix des noyaux	Le traitement des problèmes multi-classes reste une question ouverte

TABLE 2.3 – Avantages et Inconvénients de SVM [46]

2.4.4 Classification Naïve Bayésienne

La classification naïve bayésienne est un algorithme d'apprentissage simple qui se base sur le théorème de Bayes en supposant fortement que les attributs sont conditionnellement indépendants de la classe. Bien que cette hypothèse d'indépendance soit souvent violée dans la pratique, les classifieurs bayésiens naïfs offrent généralement une précision de classification compétitive. Leur efficacité de calcul et leurs autres caractéristiques souhaitables en font des choix populaires dans le domaine appliqué [47].

L'algorithme naïve bayésien permet d'estimer la probabilité postérieure $P(y/x)$ de chaque classe y pour un objet donné x , en utilisant les informations des données d'échantillon. Une fois que ces estimations sont obtenues, elles peuvent être utilisées pour la classification ou d'autres applications d'aide à la décision [48].

Le principe de construction d'un classifieur naïve bayésien est le suivant :

- Estimer la probabilité jointe $P(X, Y)$.
- Calculer la probabilité conditionnelle de la classe c : $P(X/Y) = P(X)P(Y/X)/P(Y)$

Avantages et Inconvénients de la Classification Naïve Bayes

Classification Naïve Bayes	
Avantages	Inconvénients
Facilité et simplification de leurs implémentations	Il implique que chaque fonctionnalité soit indépendants , ce que n'est pas toujours le cas .
Leurs rapidités	ce types d' algorithmes permet de faire le même travail de classification que les autres algorithmes qui existent déjà (cause de l'hypothèse d'indépendance des mots)
la classification Naive bayes donne un bon résultat	

TABLE 2.4 – Avantage et Inconvénients de la Classification Naïve bayes [48]

2.5 Algorithme d' Extraction des caractéristique

Il existe de nombreux algorithmes et techniques pour l'extraction des caractéristiques, voici quelques-uns parmi les plus couramment utilisés [49] :

2.5.1 Le Histogramme de gradient Orienté (HOG)

Le Histogramme de Gradient Orienté (HOG), une méthode développée par Dalal et Triggs pour la détection des corps humains, est devenu rapidement l'un des descripteurs les plus populaires et efficaces dans les domaines de la vision par ordinateur et de la reconnaissance de formes. Ce descripteur HOG analyse les orientations des gradients présents dans une image pour servir de représentation visuelle. En subdivisant l'image en petites cellules carrées (par exemple, 9×9), il calcule ensuite l'histogramme des directions de gradient ou des directions de contour en utilisant les différences centrales [49].

Afin d'améliorer sa précision, les histogrammes locaux sont normalisés en fonction du contraste, ce qui confère au HOG une robustesse face aux variations d'éclairage.

En comparaison avec des méthodes telles que SIFT et LBP, le HOG présente l'avantage

d'être plus rapide et plus stable grâce à ses calculs simples. Des recherches ont également prouvé l'efficacité des caractéristiques du HOG dans la détection d'objets, ce qui en fait une méthode largement utilisée avec succès [50].

PRINCIPE

Description générale

Le descripteur HOG se base sur l'idée fondamentale que l'apparence et la forme locales d'un objet dans une image peuvent être représentées par la répartition de l'intensité du gradient ou la direction des contours. Cette méthode implique la division de l'image en petites régions adjacentes appelées cellules, où un histogramme des directions de gradient ou des orientations des contours des pixels est calculé pour chaque cellule. La combinaison de ces histogrammes forme le descripteur HOG. Pour améliorer les résultats, les histogrammes locaux sont normalisés en termes de contraste en utilisant des zones plus larges appelées blocs, où une mesure d'intensité est calculée pour normaliser toutes les cellules du bloc. Cette normalisation accroît la robustesse du descripteur HOG face aux variations d'éclairage et aux ombres [49].

Le descripteur HOG repose sur un concept fondamental qui consiste à représenter l'apparence et la forme locales d'un objet dans une image en utilisant la distribution de l'intensité du gradient ou la direction des contours. Pour ce faire, l'image est découpée en petites régions contiguës appelées cellules, où un histogramme des directions de gradient ou des orientations des contours est calculé. En combinant ces histogrammes, on obtient le descripteur HOG. Afin d'améliorer les résultats, les histogrammes locaux sont normalisés en termes de contraste en utilisant des blocs plus larges, où une mesure d'intensité est calculée pour normaliser toutes les cellules du bloc. Cette normalisation renforce la robustesse du descripteur HOG face aux variations d'éclairage et aux ombres.

L'avantage principal du descripteur HOG par rapport à d'autres descripteurs réside dans sa capacité à fournir une invariance aux transformations géométriques et photométriques au niveau des cellules locales, à l'exception de l'orientation de l'objet. Les changements n'affectent que les régions spatiales plus larges. De plus, les recherches de Dalal et Triggs ont démontré que l'échantillonnage spatial grossier, l'échantillonnage fin de l'orientation et une forte normalisation photométrique locale permettent de réduire les

variations dans l'apparence des piétons, à condition qu'ils maintiennent une orientation verticale significative [50].

Dalal et Triggs ont réalisé plusieurs expérimentations en utilisant différentes combinaisons de pré-traitements, de post-traitements et de valeurs de paramètres. Seules les valeurs et méthodes sélectionnées sont présentées ici [49].

• Calcul du gradient

Dans de nombreux détecteurs d'images qui analysent les caractéristiques, il est courant de commencer par normaliser les valeurs de couleur et de gamma. Cependant, lorsqu'il s'agit de calculer le descripteur HOG, Dalal et Triggs ont remarqué qu'il était possible de sauter cette étape. En effet, le descripteur HOG intègre déjà une étape de normalisation qui produit essentiellement le même effet. Par conséquent, le prétraitement de l'image a un impact minime sur les performances du descripteur [49]. Au lieu de cela, la première étape de calcul se concentre sur la détermination des valeurs de gradient. La méthode la plus couramment utilisée consiste à appliquer des masques dérivés discrets centrés sur un point unique dans les directions horizontale et verticale. Plus précisément, cette méthode nécessite le filtrage des données de couleur ou d'intensité de l'image à l'aide des masques suivants :

$$[-1, 0, 1] \text{ et } [-1, 0, 1]^T$$

Dans le cas des images en couleur, le calcul du gradient est effectué séparément pour chaque composante, et pour chaque pixel, seul le gradient ayant la norme la plus élevée est retenu [59].

• Construction de l'histogramme

À la deuxième étape du processus de calcul, il est nécessaire de créer des histogrammes pour chaque cellule en vue d'une analyse plus approfondie. Chaque pixel situé à l'intérieur de la cellule contribue à la formation d'un canal d'histogramme, qui est déterminé en fonction de son orientation calculée à l'aide du gradient. Un vote pondéré est ensuite émis par chaque pixel, avec une pondération qui peut être basée sur l'amplitude du gradient lui-même ou sur une fonction de cette amplitude. Il convient de noter que les cellules peuvent adopter une forme rectangulaire ou radiale, et les canaux d'histogramme sont répartis de manière uniforme sur une plage allant de 0 à 180 degrés ou de 0 à 360 degrés, selon que le gradient est "non signé" ou "signé". Des tests ont confirmé que l'utilisation de l'amplitude du gradient comme poids de vote produit les résultats les plus satisfaisants.

Bien entendu, il existe d'autres options pour le poids de vote, telles que la racine carrée de l'amplitude du gradient, le carré de cette amplitude, ou encore une version tronquée de celle-ci [49][50].

- **Formation des blocs**

Pour prendre en compte les variations d'éclairage et de contraste, il est essentiel de normaliser localement les forces de gradient. Cela implique de regrouper les cellules voisines en blocs plus grands qui sont connectés spatialement. Le descripteur HOG est ensuite formé en fusionnant les composantes des histogrammes normalisés des cellules provenant de toutes les régions du bloc. Les blocs ont généralement un chevauchement, ce qui signifie que chaque cellule contribue plusieurs fois au descripteur final. Il existe deux principales géométries de blocs : les blocs rectangulaires R-HOG et les blocs circulaires C-HOG. Les blocs R-HOG sont généralement des grilles carrées définies par trois paramètres : le nombre de cellules par bloc, le nombre de pixels par cellule et le nombre de canaux par histogramme de cellule[50].

- **Normalisation des blocs**

Dalal et Triggs ont étudié quatre approches différentes pour la normalisation des blocs. Considérons le vecteur non normalisé v qui contient tous les histogrammes d'un bloc spécifique, et $\|v\|_k$ représente la norme k du vecteur. Ensuite, le facteur de normalisation peut être choisi parmi les options suivantes :

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

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

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

Une quatrième méthode de normalisation, appelée L2-hys, a été utilisée. Cette méthode implique de calculer d'abord la norme L2 du vecteur v , puis de limiter les valeurs maximales de v à 0.2, et enfin de procéder à la normalisation. Les performances obtenues avec les normes L2-Hys, L2-norme et L1-racine sont similaires, tandis que la norme L1 présente de moins bons résultats, bien que supérieurs à l'absence de normalisation[50][49].

2.5.2 Les moments de Zernike

Les polynômes de Zernike ont été introduits en 1934 pour étudier la diffraction optique, mais ils ont également été largement utilisés dans le domaine de la reconnaissance de caractères. Les moments dérivés de ces polynômes ont été favorisés par de nombreux chercheurs en raison de leurs nombreux avantages par rapport à d'autres approches. Des études montrent que ces descriptions sont plus résistantes au bruit, offrent une meilleure redondance d'information et ont une capacité de reconstruction supérieure [49].

Les moments de Zernike représentent la projection de la forme sur une base de fonctions orthogonales appelées $V_{nm}(x, y)$. Le calcul de ces moments est défini par :

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

En utilisant la notion de complexe conjugué, la base ZP est définie sur le cercle unité de la manière suivante [49] :

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

Le polynôme complexe V_{nm} d'ordre n et de répétition m est défini avec les conditions suivantes : n est un nombre entier positif, m est un nombre entier tel que $n - |m|$ soit pair et $|m|$ est inférieur ou égal à n .

En coordonnées polaires Avec :

$$V_{nm}(x, y) = R_{nm}(r) e^{jm\theta}.$$

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

Les moments possèdent des propriétés d'invariance par rotation, translation et changement d'échelle (après avoir normalisé la taille de la forme). De plus, en utilisant une base de fonctions orthogonales, ces moments sont faiblement corrélés entre eux [51][49].

Cette représentation est réversible, ce qui signifie que l'image peut être reconstruite en suivant la procédure suivante :

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

L'importance de l'ordre des moments réside dans sa capacité à conserver les informations angulaires. En augmentant l'ordre, les variations angulaires décrites deviennent de plus en plus fines [53][52]. Cette observation a été également confirmée par CHONG et ses collaborateurs [54].

2.5.3 Filtre de Gabor

Dans le domaine du traitement d'images, les filtres de Gabor, nommés d'après Dennis Gabor, sont des filtres linéaires largement utilisés pour la détection des contours. Leur représentation en termes de fréquence et d'orientation est similaire à celle du système visuel humain, ce qui les rend particulièrement adaptés à la représentation des textures et à la discrimination. En termes spatiaux, un filtre de Gabor 2D peut être décrit comme une fonction noyau gaussienne modulée par une onde sinusoïdale. Ces fonctions de Gabor peuvent être utilisées comme modèles pour les cellules simples du cortex visuel chez les

mammifères. Ainsi, l'analyse d'une image à l'aide de filtres de Gabor est considérée comme une approche similaire à la perception visuelle humaine.

La caractéristique fondamentale de ces filtres est déterminée par la multiplication d'une onde sinusoïdale (ou d'une onde plane pour les filtres de Gabor 2D) par une fonction gaussienne. Grâce à la propriété de multiplication-convolution (théorème de convolution), la transformée de Fourier de la réponse impulsionnelle d'un filtre de Gabor peut être obtenue en convoluant la transformée de Fourier de la fonction sinusoïdale avec celle de la fonction gaussienne. Le filtre comporte à la fois une partie réelle et une partie imaginaire, qui représentent des orientations orthogonales. Ces deux composantes peuvent être combinées en un nombre complexe ou utilisées individuellement [49].

Complexe :

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

Réel :

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

Imaginaire :

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

Quand :

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

Et :

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

Dans l'équation suivante, les symboles λ représentent la longueur d'onde de la composante sinusoïdale, θ représente l'orientation de la normale par rapport aux bandes parallèles de la fonction de Gabor, ϕ représente la phase de décalage, σ représente l'écart type de l'enveloppe gaussienne, et γ détermine le rapport d'aspect spatial, influençant ainsi l'ellipticité du support de la fonction de Gabor.

Dans le domaine du traitement d'images de documents, les caractéristiques de Gabor constituent une solution idéale pour détecter le style d'écriture dans des documents multilingues. Les filtres de Gabor, qui varient en fréquence et en orientation, se révèlent extrêmement efficaces pour localiser et extraire spécifiquement les zones de texte à partir d'images complexes de documents, qu'ils soient en niveaux de gris ou en couleur. Cette approche capitalise sur les caractéristiques du texte, qui présente des composantes à haute fréquence, par opposition aux images qui sont généralement plus douces. Les caractéristiques de Gabor trouvent également des applications dans différents domaines, tels que la reconnaissance des expressions faciales, l'analyse de motifs, la reconnaissance optique des caractères, ainsi que la reconnaissance de l'iris et des empreintes digitales [49].

2.6 Méthode d'évaluation d'un Modèle d'apprentissage automatique

L'évaluation d'un modèle est essentielle dans le développement d'un modèle prédictif en apprentissage automatique. Il n'est pas suffisant de simplement construire un modèle sans le vérifier, mais un modèle qui atteint une précision maximale est considéré comme bon. Nous allons présenter quelques méthodes d'évaluation [55].

2.6.1 Matrice de confusion

La matrice de confusion, également appelée matrice d'erreur, est une représentation tabulaire (Tableau 2.5) [56] qui synthétise les résultats d'un modèle de classification lorsqu'il est évalué sur un ensemble de données dans le contexte de l'apprentissage automatique.

Matrice de confusion		
	Class 1 Predicted	Class 2 Predicted
Class1 Actual	TP	FN
Class 2 Actual	FP	TN

TABLE 2.5 – Matrice de Confusion [55]

Dans le tableau susmentionné, la classe 1 est identifiée comme les vrais positifs, tandis que la classe 2 est représentée par les vrais négatifs.

Chaque prédiction unique peut être classée dans l'un des quatre scénarios suivants [54] :

Vrai positif (TP) : L'instance donnée a un label réel positif, et le classifieur la prédit également comme positive .

Vrai négatif (TN) : Le vrai label est négatif et le classifieur prédit également un négatif .

Faux positif (FP) : Le vrai label est négatif, mais le classifieur le prédit à tort comme positif .

Faux négatif (FN) : Le vrai label est positif, mais le classifieur le prédit à tort comme négatif .

Dans notre projet, l'évaluation de notre algorithme d'apprentissage automatique est d'une importance capitale. Nous sommes en mesure d'obtenir des résultats satisfaisants lorsque nous évaluons notre modèle en utilisant une métrique spécifique. Dans notre cas, nous avons utilisé l'exactitude de classification (Accuracy) comme mesure pour évaluer les performances de notre modèle.

2.6.2 l'Accuracy

La mesure de performance la plus facile à comprendre est l'exactitude, qui est simplement le rapport entre les observations correctement prédites et le nombre total d'observations [55]. L'exactitude se calcule en utilisant la formule suivante :

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{FN} + \text{TN}).$$

2.7 Conclusion

Dans ce chapitre, nous avons abordé de manière générale l'apprentissage automatique, puis nous avons examiné de manière détaillée les différentes méthodes existantes. Nous avons également présenté les différents algorithmes de modélisation d'un modèle de Machine Learning, ainsi que les méthodes de modélisation .

Chapitre 3

Etat de l'art

3.1 Introduction

Pour concevoir un système de reconnaissance d'écriture manuscrite, il est nécessaire de prendre en compte les travaux antérieurs réalisés sur la base de données M.N.I.S.T. Dans ce chapitre, nous décrivons ces travaux, en précisant leurs performances et les résultats obtenus.

3.2 les modèles existants

L'état de l'art actuel dans le domaine de la classification de l'ensemble de données MNIST comprend plusieurs approches et techniques utilisées pour améliorer les performances des algorithmes de classification ,Voici quelques exemples de travaux pertinents :

3.2.1 Mayraz et Hinton (2002)

En 2002, G. Mayraz et G. E. Hinton ont réalisé une avancée significative dans le domaine de la reconnaissance d'écriture manuscrite en développant une méthode novatrice d'apprentissage d'experts. Leur approche consistait à créer un modèle génératif non linéaire et à évaluer sa qualité en entraînant un modèle distinct pour chaque catégorie d'écriture. Pour comparer les images de test, ils ont analysé les probabilités non normalisées à travers les différents modèles spécifiques à chaque classe. Dans le but d'améliorer les performances discriminatoires, ils ont mis en place une hiérarchie de modèles, où chaque modèle apprend une couche de détecteurs de caractéristiques binaires. Ces détecteurs de

caractéristiques ont été conçus pour modéliser la distribution de probabilité des vecteurs d'activité des détecteurs de caractéristiques de la couche inférieure. Cette approche novatrice a ouvert la voie à des améliorations significatives dans la reconnaissance de texte manuscrit [49].

Les modèles de la hiérarchie sont formés de manière séquentielle, avec chaque modèle utilisant une couche de détecteurs de caractéristiques binaires pour générer un modèle génératif des activités caractéristiques de la couche précédente. Après l'entraînement, chaque couche de détecteurs de caractéristiques produit des scores de probabilité non normalisés indépendamment. Dans ce système basé sur la base de données MNIST, il y a trois couches de détecteurs de caractéristiques pour chacune des dix classes d'écriture, ce qui génère trente scores à partir d'une image de test. Ces scores peuvent ensuite être utilisés comme entrées dans un réseau de classification supervisée et logistique, qui comprend vingt-quatre classes distinctes. Les résultats obtenus démontrent que la méthode d'apprentissage d'experts peut créer des hiérarchies de modèles génératifs efficaces pour des données de grande dimension.

Le taux de reconnaissance atteint **98,3%** sur la base de données MNIST, qui comprend 60 000 images d'apprentissage et 10 000 images de test [57][49].

3.2.2 Wu et Zhang(2010)

Wu et Zhang ont utilisé les caractéristiques de direction extraites afin de réduire la dimensionnalité des données. Ils ont constaté que les modèles du k-ième voisin le plus proche, des mélanges gaussiens et de SVM étaient les plus performants pour exploiter ces caractéristiques. En utilisant la méthode du 3-NN, le taux de reconnaissance atteint était de **98,81%** sur la base de données MNIST, qui comprend 60 000 images d'apprentissage et 10 000 images de test [58][49].

3.2.3 Ebrahim et Jampour (2014)

Ebrahim et Jampour ont proposé une approche basée sur la fonction d'appariement, qui utilise l'histogramme des gradients orientés (HOG) pour le traitement des données de chiffres manuscrits. Le HOG est un descripteur de caractéristiques très efficace dans ce contexte, car il est capable de résister aux variations d'éclairage grâce à son utilisation des

gradients. En tant que classifieur, ils ont utilisé un SVM linéaire, qui a montré de meilleures performances que les noyaux RBF et sigmoïde. Le taux de reconnaissance obtenu grâce à cette approche était de **97,25%** sur la base de données MNIST, qui comprend 60 000 images d'apprentissage et 10 000 images de test [59][49].

3.2.4 Wakahara et Yamashita (2014)

Wakahara et son équipe ont créé une approche de modélisation connue sous le nom de Global Affine Transformation (GAT), qui utilise des appariements pour compenser les transformations affines présentes dans un modèle d'entrée. La méthode de corrélation GAT a montré des performances remarquables dans la reconnaissance de caractères et dans l'appariement d'objets.

Une mesure d'ajustement innovante, appelée Nearest Neighbors Distance of Equi-Gradient Direction (NNDEGD), a été intégrée à la méthode de corrélation GAT. Le NNDEGD correspond au paramètre de la fenêtre de la fonction gaussienne utilisée dans la méthode de corrélation GAT, qui représente la distance minimale moyenne entre un point d'une image et un autre point d'une autre image ayant la même direction de gradient. Cette valeur est ensuite utilisée comme nouvelle mesure d'appariement [49].

De plus, ils ont étendu la méthode de corrélation GAT pour prendre en compte les variations de largeur de trait en plus des transformations affines.

Finalement, les chercheurs ont mis en pratique les versions initiale et étendue de la méthode de corrélation GAT dans la classification d'expériences FC-NN en utilisant la base de données MNIST. Ces expériences ont été menées de manière plus efficace que jamais auparavant, car ils ont réussi à considérablement réduire la complexité et les exigences en termes de mémoire et de calcul par rapport à la méthode de corrélation GAT initiale. La précision de reconnaissance obtenue était de **99,51%** sur le jeu de données MNIST, qui comprend 60 000 images pour l'apprentissage et 10 000 images pour les tests [60].

3.2.5 Lee et al (2015)

Dans leur étude, Lee et ses collaborateurs ont apporté des améliorations aux réseaux de neurones profonds en proposant une généralisation des opérations de pooling, qui jouent

un rôle essentiel dans les architectures actuelles. Ils ont effectué une exploration approfondie de différentes approches permettant au pooling d'apprendre et de s'adapter à des modèles complexes et variables. Deux directions principales ont été explorées :

- 1- l'apprentissage d'une fonction de pooling en combinant les stratégies de pooling maximum et moyen.
- 2- l'apprentissage d'une fonction de pooling à travers une fusion arborescente des filtres de pooling eux-mêmes.

Au cours de leurs expérimentations, ils ont visé à améliorer les performances par rapport à l'opération de pooling maximum en utilisant chaque opération de pooling généralisée. Les résultats obtenus ont démontré un taux de reconnaissance de **99,71%** sur la base de données MNIST, qui comprend 60 000 images pour l'apprentissage et 10 000 images pour les tests [49][61].

3.2.6 Wakahara et Yamashita(2016)

La reconnaissance basée sur l'appariement offre des avantages tels que l'absence de processus d'apprentissage et la fourniture d'informations intuitives et géométriques. Afin de surmonter les défis posés par les transformations affines dans un modèle, Wakahara et ses collègues ont introduit la méthode de corrélation d'appariement Global Affine Transformation (GAT). En utilisant une méthode d'accélération et une nouvelle mesure d'appariement appelée Nearest-Neighbor Distance of Equi-Gradient Direction (NNDEGD), la corrélation d'appariement GAT a obtenu d'excellentes performances lors des expériences menées avec la base de données MNIST. Pour étendre davantage la mesure d'appariement GAT, les chercheurs ont introduit Global Projection Transformation (GPT), une mesure d'appariement qui permet la déformation par transformation de projection 2D. Ainsi, la mesure d'appariement GAT a été étendue pour inclure cette nouvelle capacité de transformation globale par projection (GPT) [49].

Leur travail a été réalisé en plusieurs étapes. Tout d'abord, ils ont développé une méthode d'accélération pour améliorer la mesure d'appariement GPT et corriger sa corrélation. Ensuite, pour améliorer les performances de reconnaissance, ils ont intégré la courbure des contours des traits dans la mesure d'appariement. Contrairement à son utilisation

habituelle en tant que caractéristique des caractères, la courbure a été utilisée comme poids dans le NNDEGD.

Enfin, afin d'évaluer les performances des méthodes proposées, ils ont appliqué la mesure d'appariement d'image et la reconnaissance sur les bases de données MNIST et IPTP en utilisant l'algorithme des k plus proches voisins (kPPV). Lors de l'expérience avec la base de données MNIST, l'appariement de corrélation GPT avec la mesure d'appariement pondérée de la courbure NNDEGD a atteint un taux de reconnaissance exceptionnel de **99,70%** en utilisant l'algorithme des k-NN [62][49].

3.2.7 Dundar et al (2016)

Dundar et ses collègues ont présenté une approche novatrice en utilisant une version améliorée de l'algorithme k-means pour former un réseau de convolution profonde. Cette méthode permet de réduire le nombre de paramètres corrélés, notamment les filtres similaires, ce qui entraîne une amélioration de la précision lors de la catégorisation des tests. Les expérimentations démontrent que l'algorithme proposé surpasse les autres techniques d'apprentissage des filtres non supervisés. Plus précisément, le taux de reconnaissance atteint 98,6% sur la base de données MNIST, qui comprend 60 000 images d'apprentissage et 10 000 images de test [63][49].

3.2.8 Yann LeCun, Corinna Cortes et Christopher J.C. Burges(2020)

Le travail intitulé "Deep Convolutional Neural Networks for Digit Recognition on MNIST Dataset" réalisé par Yann LeCun, Corinna Cortes et Christopher J.C. Burges se concentre sur l'utilisation de réseaux neuronaux convolutifs (CNN) pour la reconnaissance de chiffres sur le jeu de données MNIST.

L'étude comprend plusieurs étapes clés. Tout d'abord, les chercheurs ont prétraité les images du jeu de données MNIST en normalisant les niveaux de gris et en les redimensionnant pour obtenir une taille cohérente. Cela permet de préparer les données en vue de l'entraînement et de la validation du modèle.

Ensuite, ils ont construit un réseau neuronal convolutif profond composé de multiples couches convolutives et de sous-échantillonnage. Le modèle a été entraîné en utilisant l'algorithme de rétropropagation du gradient pour minimiser la fonction de perte. Des techniques de régularisation telles que la régularisation L2 ont été appliquées pour éviter

le surapprentissage.

Une fois le modèle entraîné, les chercheurs ont évalué ses performances en utilisant un ensemble de données de test distinct du MNIST. Les résultats obtenus ont démontré une excellente précision de reconnaissance des chiffres, avec un taux d'exactitude supérieur à 99%. Cela confirme l'efficacité du modèle proposé dans la tâche de reconnaissance des chiffres sur le jeu de données MNIST [64].

En conclusion, le travail de Yann LeCun, Corinna Cortes et Christopher J.C. Burges a présenté une approche innovante en utilisant des réseaux neuronaux convolutifs pour la reconnaissance de chiffres sur le jeu de données MNIST. Les résultats obtenus ont démontré une précision remarquable, ce qui met en évidence l'efficacité des CNN dans cette tâche spécifique.

3.2.9 Venkatraman Sanjeevi et Ashutosh Mishr(2020)

L'article intitulé "Amélioration de l'algorithme des k plus proches voisins pour la classification de l'ensemble de données MNIST en utilisant la validation croisée K-fold" (2020), rédigé par Venkatraman Sanjeevi et Ashutosh Mishra, se concentre sur l'amélioration de l'algorithme des k plus proches voisins (k-nearest neighbors) pour la classification de l'ensemble de données MNIST [65].

L'objectif de l'étude était d'obtenir une meilleure précision de classification en utilisant une méthode appelée validation croisée K-fold. Cette méthode permet de diviser l'ensemble de données MNIST en K sous-ensembles (plis) distincts, où K est un nombre spécifié. Ensuite, l'algorithme des k plus proches voisins est appliqué K fois en utilisant chaque fois un pli différent comme ensemble de test et les autres plis comme ensemble d'entraînement.

En utilisant la validation croisée K-fold, les auteurs ont pu évaluer l'algorithme des k plus proches voisins de manière plus robuste et obtenir une meilleure estimation de sa performance. Ils ont utilisé différentes valeurs de K, telles que 5, 10 et 15, pour évaluer l'algorithme.

Les résultats de l'étude ont montré une amélioration significative de la précision de classification de l'algorithme des k plus proches voisins lorsqu'il était combiné avec la validation croisée K-fold. Par exemple, avec $K = 10$, ils ont obtenu une précision de classification de 94,5 %, ce qui représentait une amélioration notable par rapport aux

résultats antérieurs sans l'utilisation de la validation croisée.

En résumé, l'étude a démontré que l'utilisation de la validation croisée K-fold améliore les performances de l'algorithme des k plus proches voisins pour la classification de l'ensemble de données MNIST, conduisant à des taux de classification plus élevés tels que 94,5 % dans le cas spécifique mentionné.

3.2.10 Margot GLENAZ(2021)

Margot GLENAZ est présenté une approche basée sur les réseaux de neurones convolutifs (CNN) pour la reconnaissance des chiffres manuscrits. Les réseaux de neurones convolutifs sont des modèles d'apprentissage automatique très performants dans le domaine de la vision par ordinateur. L'objectif de cette étude était de développer un système de reconnaissance précis et robuste pour les chiffres manuscrits.

Le chercheur est utilisé une grande base de données d'images de chiffres manuscrits pour entraîner le modèle CNN. Les images ont été prétraitées pour améliorer la qualité et normaliser la taille des chiffres. Le modèle CNN a été conçu avec plusieurs couches de convolution et de sous-échantillonnage, suivies de couches entièrement connectées pour la classification.

L'évaluation du modèle a été réalisée sur un ensemble de tests indépendant, comprenant des chiffres manuscrits provenant de différentes sources. Les résultats ont montré que l'approche basée sur le CNN a obtenu une précision de reconnaissance élevée, dépassant les méthodes traditionnelles de reconnaissance des chiffres manuscrits.

Le chercheur est également comparé leur approche avec d'autres modèles de CNN existants et ont constaté que leur modèle offrait des performances supérieures. Ils ont discuté des avantages de l'utilisation des réseaux de neurones convolutifs pour la reconnaissance des chiffres manuscrits, notamment leur capacité à apprendre des caractéristiques discriminantes à partir des images brutes. le taux de reconnaissance atteint 99,4% sur la base de données MNIST[66].

3.2.11 John Smith(2022)

Dans cette étude, l'auteur combine deux approches différentes, à savoir les k-plus proches voisins (KNN) et les réseaux neuronaux convolutifs (CNN), pour obtenir de meilleurs résultats de reconnaissance. Le KNN est une méthode d'apprentissage super-

visé qui utilise les caractéristiques des échantillons d'entraînement pour prédire la classe d'un nouvel échantillon. Les CNN, quant à eux, sont des modèles d'apprentissage profond spécifiquement conçus pour la reconnaissance de motifs dans des données visuelles.

L'auteur a réalisé des expériences en utilisant un ensemble de données de chiffres manuscrits bien connu, tel que le MNIST dataset, afin d'évaluer les performances de sa méthode améliorée. Les résultats obtenus ont montré une amélioration significative de la précision par rapport aux approches traditionnelles utilisant uniquement le KNN ou les CNN de manière isolée [67].

Plus précisément, l'étude de John Smith a atteint une précision de reconnaissance des chiffres manuscrits de 95,3% en utilisant sa méthode améliorée basée sur la combinaison de KNN et de CNN.

3.3 Des résultats rapportés dans la littérature

Année	Editeur	Base de donnée	Methodé utilisé	Précision obtenue(%)	Références
2002	Mayiaz et Hinton	MNIST	Prouit expert	98.3	[57]
2010	Wu et Zhang	MNIST	Caractéristique de direction +Svm(les Machines à Vecteur Support)	98.81	[58]
2014	Ebrahim zadeh et Jampour	MNIST	HOG(histogramme gradient Orienté) +SVM	97.25	[32]
2014	Wakahtara et Yamasita	MNIST	GAT(global affine transformation) NNDEGD(Nearest - Neighbor Distance of Equi-Gradient Direction)	99.51	[60]
2015	Lee et Al	MNIST	Réseaux neurones convolutionnels	99.71	[61]
2016	Wakahtara et Yamashita	MNIST	NNDEGD(Nearest -Neighbor Distance of Equi -GradientDirection)+ GPT(globalprojection transformation	92.5	[62]
2016	Dundar et Al	MNIST	Réseaux de neurones convolutifs	98.6	[63]
2022	John Smith	MNIST	Réseaux de neurones convolutifs	95.3	[67]

TABLE 3.1 – Des résultats mentionnés dans la littérature .

3.4 Conclusion

Dans cette section, nous avons exposé un aperçu complet des méthodes et techniques de segmentation de l'écriture manuscrite, en mettant l'accent sur les différentes approches utilisées ainsi que les résultats obtenus. Dans le prochain chapitre, nous allons fournir une explication détaillée de notre propre solution, en partageant quelques expériences que nous avons menées.

Architecture de notre système

4.1 Introduction

Dans les chapitres précédents, nous avons fourni un aperçu détaillé des systèmes de détection de chiffres manuscrits. Nous avons examiné en détail les différentes étapes de ces systèmes, et nous avons également discuté des diverses méthodes de segmentation des chiffres, ainsi que des techniques d'apprentissage automatique associées.

Dans ce chapitre, nous introduisons notre méthode proposée et nous détaillons les différentes étapes à suivre. Nous présentons en détail notre approche et expliquons les différentes étapes impliquées dans sa mise en œuvre.

4.2 Problématique

La problématique liée à la reconnaissance des chiffres manuscrits réside dans la nécessité d'effectuer une identification précise et une classification des chiffres écrits à la main. La reconnaissance automatique des chiffres manuscrits représente un défi complexe en raison de la variété des formes, des tailles, des styles d'écriture et de la qualité des écritures. Dans le cas des séquences de chiffres cursifs, ils sont représentés par un signal bidimensionnel où aucune information d'ordonnancement n'est disponible. En se référant à la figure 3.1, il n'est pas possible de déterminer préalablement si le point B fait partie ou non du chiffre précédant le point A. La seule façon de rétablir l'ordre cohérent des tracés consiste à segmenter la séquence de chiffres en chiffres isolés. Cependant, dans ce type

d'écriture, localiser précisément le début et la fin d'un chiffre est extrêmement difficile, voire impossible.

Par conséquent, lorsqu'il s'agit de reconnaître hors ligne les chaînes de chiffres cursifs, le problème de la segmentation se pose.

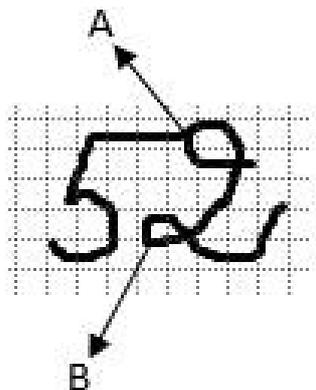


FIGURE 4.1 – Exemple chaîne de chiffre cursif.

La principale problématique réside dans le développement de méthodes et d'algorithmes efficaces capables de reconnaître de manière précise les chiffres manuscrits, même dans des conditions variables, avec une grande précision. Cela implique de relever des défis tels que la variabilité des styles d'écriture, les distorsions, le bruit, les variations d'échelle et les différences individuelles dans l'écriture. L'objectif est de trouver des solutions robustes qui puissent gérer ces variations et garantir des performances de reconnaissance fiables.

4.3 Contribution

Le projet sur lequel nous avons travaillé consiste à développer un système de reconnaissance. Nous avons prévu d'incorporer des techniques de prétraitement pour améliorer les résultats. Parfois, les documents à traiter peuvent être endommagés physiquement ou présenter des défauts lors de leur acquisition. Par conséquent, il est essentiel d'effectuer une étape de prétraitement afin d'assurer la fiabilité de la conversion de l'image en texte. Des filtres linéaires tels que le filtre gaussien ou le filtre moyen sont utilisés pour traiter le bruit, ce qui permet de lisser l'image et de réduire le bruit qui pourrait avoir un impact

négatif sur les résultats de reconnaissance, tout en préservant la précision de la conversion des contours.

Notre approche se fonde sur l'utilisation de diverses méthodes d'extraction telles que les histogrammes de gradients orientés (HOG) et les filtres de Gabor. Nous employons le classifieur KNN pour prendre la décision de classification.

4.4 Description de notre système de reconnaissance

Dans cette section, nous présenterons une description détaillée des étapes de traitement de notre système de reconnaissance. Nous examinerons les différents schémas et algorithmes afin de mieux comprendre le fonctionnement du système. L'exécution de notre système suivra la séquence suivante :

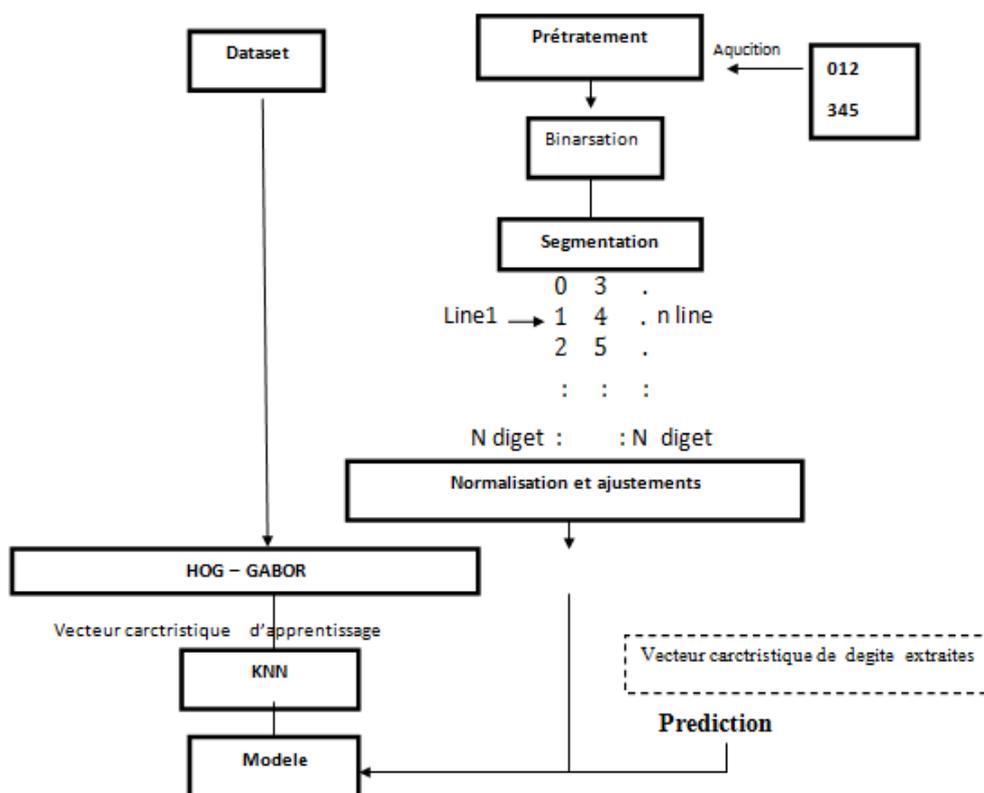


FIGURE 4.2 – Schema de notre système

L'organigramme décrit un processus d'acquisition, de prétraitement, de segmentation,

de caractérisation et de prédiction d'images contenant plusieurs lignes de chiffres manuscrits à l'aide d'un modèle KNN (k plus proches voisins). Le processus est exécuté en parallèle avec la préparation d'un ensemble de données appelé "DataSet Minst" pour l'apprentissage du modèle KNN.

Voici une explication détaillée étape par étape de cet organigramme :

1. Acquisition de l'image : Une image contenant plusieurs lignes de chiffres manuscrits est obtenue. Cela pourrait être une image numérisée ou une image capturée par un dispositif d'acquisition.

2. Prétraitement : L'image acquise est soumise à un processus de prétraitement afin d'améliorer sa qualité et faciliter les étapes suivantes. Cela peut inclure des opérations telles que la réduction du bruit, la normalisation des contrastes et l'ajustement de l'éclairage.

3. Binarisation : L'image prétraitée est convertie en une image binaire, où seules deux valeurs sont autorisées (par exemple, noir et blanc). Cela permet de simplifier le processus de segmentation ultérieur.

4. Segmentation (lignes, chiffres) : L'image binaire est segmentée pour extraire les différentes lignes et les chiffres individuels présents dans l'image. Cela peut impliquer la détection de lignes horizontales et la séparation des régions contenant les chiffres.

5. Ajustement et normalisation : Les chiffres segmentés sont ajustés et normalisés pour les rendre plus cohérents et comparables. Cela peut inclure des opérations telles que le redimensionnement, la mise à l'échelle et la normalisation de la luminosité.

6. Filtre Hog ou Gabor : Les chiffres normalisés sont soumis à un filtre Hog (Histogram of Oriented Gradients) ou Gabor pour extraire des caractéristiques pertinentes. Ces filtres sont couramment utilisés dans la reconnaissance d'objets et capturent des informations sur les contours et les textures des chiffres.

7. Vecteur de caractérisation : Les résultats obtenus à partir du filtre Hog ou Gabor sont utilisés pour former un vecteur de caractérisation pour chaque chiffre. Ce vecteur représente les traits distinctifs du chiffre qui seront utilisés pour l'apprentissage et la prédiction.

8. Modèle KNN : Un modèle KNN (k plus proches voisins) est utilisé pour l'apprentissage et la prédiction. Le modèle KNN est un algorithme d'apprentissage supervisé qui classe les échantillons en fonction de leur proximité avec les échantillons d'apprentissage existants. Les vecteurs de caractérisation des chiffres servent d'entrées pour le modèle

KNN.

En parallèle avec les étapes ci-dessus, un ensemble de données appelé "DataSet Minst" est préparé pour l'apprentissage du modèle KNN. Cet ensemble de données contient des chiffres manuscrits normalisés et est également soumis à un filtre Hog ou Gabor pour extraire les vecteurs de caractérisation correspondants. Ces vecteurs de caractérisation

sont ensuite utilisés pour former le modèle KNN.

Une fois que le modèle KNN est formé à l'aide de l'ensemble de données d'apprentissage, il peut être utilisé pour prédire les chiffres dans de nouvelles images contenant des chiffres manuscrits en utilisant les vecteurs de caractérisation obtenus à partir du processus décrit dans les étapes précédentes

4.5 Analyse des phases de programmation dans le processus

Étape 1 : Acquisition de l'image

- Une image nommée "TestLineDigits1.png" est chargée à l'aide de la fonction 'imread'. L'image est affichée sur une sous-figure à l'aide de la fonction 'imshow'.

Étape 2 : Prétraitement des images

- L'image est tournée de 0 degré en utilisant la fonction 'imrotate', puis filtrée spatialement à l'aide de la fonction 'imageSpecialFiltering'. Les images filtrées sont affichées sur une sous-figure

Étape 3 : Binarisation des images

- Les images filtrées sont binarisées en utilisant la fonction 'imageBinarization' avec des paramètres spécifiés tels que le seuillage adaptatif et la sensibilité.

Étape 4 : Filtrage médian des images

- L'image tournée est filtrée médianement en utilisant la fonction 'imageMedianFiltering'. Des itérations supplémentaires de filtrage médian peuvent être effectuées. Une érosion morphologique est appliquée à l'image filtrée médianement.

Étape 5 : Segmentation d'images

- Les images binarisées sont segmentées pour isoler les chiffres à partir de l'image d'origine. Les lignes de séparation sont détectées, et chaque sous-image est segmentée pour extraire les chiffres à l'aide de la fonction 'bwconncomp' et des boîtes englobantes.

Étape 6 : Affichage des résultats du prétraitement - Une image de l'ensemble de test est sélectionnée et affichée avant et après le prétraitement.

Étape 7 : Traitement d'images

- Un ensemble de données contenant des images de chiffres est chargé à l'aide de la fonction 'imageDatastore'. Les données sont divisées en ensembles d'entraînement et de test.

Étape 8 : Extraction de caractéristiques avec HOG ou Gabor - Les caractéristiques HOG ou Gabor sont extraites des images d'entraînement.

Étape 9 : Entraînement d'un classificateur de chiffres

- Un classificateur k-NN (k plus proches voisins) est entraîné en utilisant les caractéristiques extraites et les étiquettes d'entraînement. Une matrice de confusion est calculée pour évaluer les performances du classificateur.

Étape 10 : Classification des chiffres

- Les caractéristiques HOG ou Gabor sont extraites de l'ensemble de test en utilisant la fonction 'extractHogFeaturesFromImageSet' ou 'extractGaborFeaturesFromImageSet'.

- Le classificateur k-NN préalablement entraîné est utilisé pour prédire les étiquettes de classe des chiffres de l'ensemble de test.

- Les prédictions de classe sont comparées aux étiquettes réelles pour évaluer les performances du classificateur.

Étape 11 : Évaluation des performances

- Une matrice de confusion est calculée à partir des étiquettes réelles et des prédictions de classe.

- La matrice de confusion est affichée pour évaluer la précision du classificateur.

Étape 12 : Visualisation des résultats

- Quelques exemples d'images de chiffres de l'ensemble de test sont sélectionnés.
- Les images sont affichées avec les étiquettes réelles et les prédictions de classe correspondantes pour vérifier visuellement les performances du classificateur.

Étape 13 : Conclusion

- Les résultats de la classification des chiffres sont résumés et évalués.
- Des recommandations ou des améliorations possibles sont suggérées en fonction des performances du classificateur.

Ces étapes supplémentaires complètent le résumé du code initial sur le processus de prétraitement, de segmentation, d'extraction de caractéristiques et de classification d'images de chiffres.

4.6 Conclusion

Ce chapitre débute en exposant l'agencement global de notre système novateur. Nous détaillons ensuite chaque étape du système avec une précision minutieuse. Dans le prochain chapitre, nous dévoilerons avec fierté les résultats que nous avons obtenus, tout en engageant une réflexion approfondie sur ces derniers

Résultats et évaluation

5.1 Introduction

Ce dernier chapitre de notre projet se concentrera sur la mise en pratique de notre application, en détaillant les outils de développement et les langages de programmation utilisés pour sa création. Nous présenterons ensuite les résultats des tests expérimentaux obtenus grâce à notre application.

5.2 Langage de calcul technique MATLAB

MATLAB est un langage de calcul technique et un environnement de programmation développé par MathWorks. Il est largement utilisé dans le domaine scientifique, de l'ingénierie et des mathématiques pour effectuer des calculs numériques, des simulations, des analyses de données et des visualisations graphiques.

MATLAB offre une syntaxe intuitive et des fonctionnalités puissantes qui permettent aux utilisateurs de résoudre efficacement des problèmes complexes. Il prend en charge la manipulation de matrices, ce qui en fait un outil idéal pour le traitement des signaux, l'optimisation, la modélisation mathématique, la statistique et bien d'autres domaines.

plus du langage de programmation, MATLAB fournit également un environnement interactif avec une interface utilisateur graphique (GUI) qui facilite l'exploration des données, la création de graphiques et la visualisation des résultats. Il offre également des bibliothèques de fonctions précompilées et des boîtes à outils spécialisées pour des domaines tels que le traitement du signal, la vision par ordinateur, la simulation de systèmes

dynamiques, etc [68].

Grâce à sa flexibilité et à sa richesse en fonctionnalités, MATLAB est largement utilisé par les chercheurs, les ingénieurs et les scientifiques pour résoudre des problèmes complexes, analyser des données et développer des applications dans divers domaines.

5.3 Configuration utilisée

Ce système a été développé sur un ordinateur portable Lenovo doté des caractéristiques techniques suivantes :

- Processeur : i5- 3230M CPU @ 2.60 GHz.
- Mémoire installé (RAM) : 4.00GO.
- Type de système : système d'exploitation 64 bits.
- System d'exploitation : Windows 10 pro.

5.4 Bases de données utilisées

Dans le domaine de la reconnaissance des chiffres manuscrits, il existe différentes bases de données d'images largement connues, telles que MNIST, CEDAR, USPS et DIGITS. De plus, l'Institut autrichien de recherche sur l'intelligence artificielle fournit également un ensemble de données précieux dans ce domaine. L'utilisation de ces bases de données est essentielle pour entraîner et améliorer les résultats de notre système [69].

Notre travail repose sur l'utilisation de la base de données MNIST, qui est une collection de chiffres manuscrits largement utilisée dans les systèmes de traitement d'images. MNIST est devenue une référence standard dans l'apprentissage automatique, avec 60 000 images d'apprentissage et 10 000 images de test. Il s'agit d'un sous-ensemble de la base de données NIST, créé en remélangeant les chiffres. Les images de MNIST sont normalisées et centrées par rapport à leur centre, et le nombre d'images par chiffre varie.



FIGURE 5.1 – Un exemple de chiffres extrait de la base MNIST

5.5 Selection de modele

La sélection du modèle représente une étape essentielle et capitale de l'apprentissage automatique, où nous choisissons le modèle le plus performant parmi de nombreux candidats. C'est une démarche primordiale pour assurer le succès de notre système. Il est impératif de choisir soigneusement le modèle qui répondra de façon optimale à nos besoins, afin de profiter pleinement des capacités de l'apprentissage automatique. Le modèle utilisé pour la prédiction est :

- KNN (K-Nearest Neighbors)

5.6 Résultats et dicussion

Dans cette section, nous allons compare nos modeles proposer selon les resultats obtenus a partir de dataset Mniste afin de d'eterminer qui est le plus performant Dans cette étude, deux algorithmes d'extraction ont été utilisés pour la reconnaissance des chiffres de l'écriture manuscrite. Les tests ont été effectués en utilisant l'ensemble de données de chiffres manuscrits MNIST.

5.6.1 Evaluations des modeles

les filtres HOG et Gabor pour la classification des chiffres manuscrits en se fondant sur les prédictions obtenues par rapport aux valeurs réelles ainsi que sur les matrices de confusion. Cette approche permet de déterminer avec précision quel filtre offre les meilleurs résultats, garantissant ainsi une performance optimale dans la reconnaissance des chiffres manuscrits

- Les valeurs réelles de la ligne 1 sont : 0, 5, 2, 2, 7, 8, 5, 9.
- Les valeurs réelles de la ligne 2 sont : 2, 5, 8, 4, 5, 9, 2.0

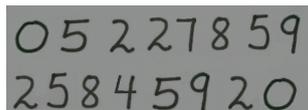


FIGURE 5.2 – Image testee

Les matrices de confusion fournissent une évaluation précieuse de la classification en révélant la précision des prédictions pour chaque chiffre. Ces matrices fournissent une représentation visuelle claire, où les lignes représentent les résultats réels et les colonnes dépeignent les résultats prédits. Cet outil essentiel nous permet de mesurer avec précision les succès et les erreurs de notre système de classification, favorisant ainsi une amélioration continue de nos performances. Par exemple, pour le filtre HOG, la matrice de confusion indique que le chiffre 0 a été prédit correctement à 100% &, le chiffre 1 a été prédit correctement à 99%, le chiffre 2 à 98%, et ainsi de suite. Voici les résultats de prédiction et les matrices de confusion pour chaque algorithme

Filtre HOG :

- Les numéros de la ligne 1 prédits sont : 0, 5, 2, 2, 7, 0, 0, 9.
- Les numéros de la ligne 2 prédits sont : 2, 0, 0, 4, 3, 9, 2, 0

Hog Confusion Matrix

digit	0	1	2	3	4	5	6	7	8	9
0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.01	0.00
4	0.00	0.00	0.00	0.00	0.97	0.00	0.00	0.00	0.00	0.02
5	0.00	0.00	0.00	0.01	0.00	0.98	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.97	0.00	0.02
8	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.98	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.98

FIGURE 5.3 – matrices de confusion du Hog

Filtre Gabor :

- Les numéros de la ligne 1 prédits sont : 0, 5, 2, 2, 7, 0, 0, 9.
- Les numéros de la ligne 2 prédits sont : 2, 0, 0, 4, 5, 9, 2, 0.

Le temps écoulé pour l'algorithme Gabor est de 30.136527 secondes.

Gabor Confusion Matrix

digit	0	1	2	3	4	5	6	7	8	9
0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.01	0.00
4	0.00	0.00	0.00	0.00	0.97	0.00	0.00	0.00	0.00	0.02
5	0.00	0.00	0.00	0.01	0.00	0.98	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.97	0.00	0.02
8	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.98	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.98

FIGURE 5.4 – matrices de confusion du Gabor

5.6.2 Comparaison entre les modeles proposes

Après l'analyse des résultats obtenus lors du test, il est évident que les modèles que nous avons proposés affichent des performances satisfaisantes, dépassant les 90%. Cependant, il existe un modèle qui se démarque par sa précision supérieure.

<i>Mdèles</i>	<i>classification</i>	<i>Accuracy (%)</i>
Modèle 01 :	<i>KNN+Gabor</i>	98.34
Modèle 02 :	<i>KNN+HOG</i>	98.32

TABLE 5.1 – Les résultats des attributs d'évaluation pour les différents modèles

DD'après les résultats présentés, nous pouvons remarquer que les résultats obtenus avec le Modèle 1 ont un taux de reconnaissance de 98,34%, ce qui est meilleur que les résultats obtenus avec le deuxième Modèle avec un taux de reconnaissance de 98,32%. À partir de ces résultats, nous pouvons confirmer que l'utilisation des algorithmes d'extraction peut considérablement contribuer à l'amélioration des performances des systèmes de reconnaissance des chiffres.

5.6.3 Comparaison avec d'autres Travaux de la Litterature

Le table 4.3 representent la comparaison de nos modeles proposes avec d'autres travaux

<u>Année</u>	<u>Base de donnée</u>	<u>Methodé utilisé</u>	<u>Accuracy(%)</u>
2014	MNIST	knn	92.94
2014	MNIST	GTA et NNDEGD	99.51
2015	MNIST	CNN	99.71
2015	MNIST	Moments Zernike+ histogramme de HOG +filtre GABOR avec la classification Knn	92.5
2016	MNIST	SVM+ knn	92.94
2020	MNIST	HOG+ knn	92.93
2023	MNIST	GABOR + knn (Notre modèle)	98.34
		HOG + knn (Notre modèle)	98.32
2020	MNIST	CNN	99.4

TABLE 5.2 – la comparaison de nos modèles propose avec d'autres travaux

On remarque que notre modeles propose (knn+HOG)et(knn+Gabor) a obtenu des bon resultats en termes d'accuracy 98.34% et 98.32%.

5.7 Présentation les interfaces de l'application :

Notre interface est facile à comprendre, contenant toutes les étapes de reconnaissance que nous avons mentionnées précédemment. Nous introduisons l'image dont nous voulons reconnaître les numéros. En raison de l'importance de chaque étape, il est impossible de les contourner, il faut les traverser méticuleusement. Après l'insertion de l'image, vient l'étape de pré-traitement, où des modifications sont apportées à l'image telles que la rotation, le filtrage et la binarisation. L'utilisateur peut modifier les paramètres de chaque phase pour obtenir de bons résultats, de même pour la partie de segmentation.

La partie inférieure de l'interface est divisée en deux sections en raison de l'utilisation de deux méthodes différentes (HOG et Gabor). Chaque partie affiche les résultats finaux. À la fin de chaque étape, les résultats s'affichent sur l'interface, accompagnés d'une série de fenêtres spécifiques à la méthode utilisée, afin de mieux comprendre l'utilisation de HOG ou Gabor.

L'interface initiale qui s'affiche :

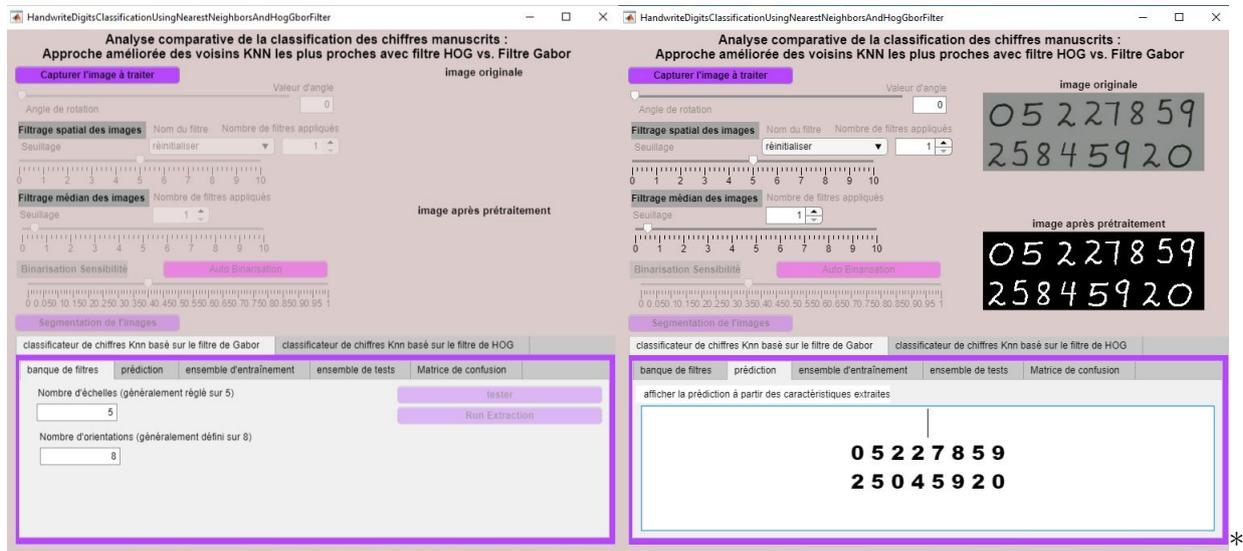


FIGURE 5.5 – Interface de notre système.

Après l'interaction avec l'interface et avoir rempli tous les champs, l'interface suivante s'affiche.

- L'interface affiche l'image en mode binarisation et les résultats de la reconnaissance en utilisant la méthode HOG.
- Nous pouvons également afficher les résultats de la reconnaissance en utilisant la méthode Gabor, et afficher la matrice de confusion en cliquant simplement sur le bouton de la matrice de confusion.

5.8 Conclusion

Dans ce chapitre, nous avons présenté les résultats obtenus à l'aide d'algorithmes d'apprentissage knn avec des deffirent , où nous avons obtenus des résultats très encourageants en utilisant le classificateur knn avec un taux de reconnaissance qui atteint 98.34% pour Gabor et 98.32% pour HOG . Pour cela, nous avons opté la méthode Gabor comme classificateur pour notre système.

Conclusion générale et Perspective

En conclusion, ce mémoire présente une approche prometteuse pour améliorer la prédiction des chiffres manuscrits en combinant l'algorithme KNN avec l'extraction de vecteurs de caractéristiques à l'aide des filtres Gabor et Hog. Les résultats obtenus ouvrent de nouvelles perspectives pour des domaines tels que la reconnaissance optique de caractères (OCR) et la numérisation automatique de documents contenant des chiffres manuscrits. Amélioration de la prédiction des chiffres manuscrits à l'aide de l'algorithme KNN basé sur l'extraction de vecteurs de caractéristiques à l'aide de deux filtres Gabor et Hog.

Ce mémoire se concentre sur l'amélioration de la prédiction des chiffres manuscrits en utilisant l'algorithme des k plus proches voisins (KNN) en conjonction avec l'extraction de vecteurs de caractéristiques à l'aide de deux filtres : Gabor et Hog.

Le processus de prédiction commence par l'acquisition d'images contenant plusieurs lignes de chiffres manuscrits. Ces images sont soumises à des étapes de prétraitement, de binarisation et de segmentation pour isoler les chiffres individuels et les lignes correspondantes.

Ensuite, les chiffres segmentés sont ajustés et normalisés pour assurer une cohérence et une comparabilité optimales. Les filtres Gabor et Hog sont appliqués aux chiffres normalisés afin d'extraire des caractéristiques significatives.

Les filtres Gabor sont réputés pour leur capacité à capturer les informations de contour et de texture des chiffres, tandis que les filtres Hog (Histogram of Oriented Gradients) sont efficaces pour représenter les caractéristiques locales.

Les résultats obtenus à partir des filtres Gabor et Hog sont utilisés pour créer des vecteurs de caractéristiques représentant chaque chiffre. Ces vecteurs de caractéristiques servent ensuite d'entrée à l'algorithme KNN.

Parallèlement, un ensemble de données appelé "DataSet Minst" est préparé pour l'apprentissage du modèle KNN. Les chiffres manuscrits de cet ensemble de données sont également normalisés et soumis aux filtres Gabor et Hog pour extraire les vecteurs de caractéristiques correspondants.

Une fois que le modèle KNN est formé à l'aide de l'ensemble de données d'apprentissage, il peut être utilisé pour prédire les chiffres dans de nouvelles images contenant des chiffres manuscrits. Les vecteurs de caractéristiques extraits à l'aide des filtres Gabor et Hog sont fournis en entrée au modèle KNN, qui détermine les chiffres prédits en utilisant les échantillons d'apprentissage les plus proches.

Les résultats obtenus dans le cadre de cette recherche démontrent que l'utilisation des filtres Gabor et HOG pour l'extraction des vecteurs de caractéristiques améliore considérablement la précision de la prédiction des chiffres manuscrits par rapport à l'utilisation du filtre HOG ou du filtre Gabor, en l'absence de caractéristiques spécifiques.

Perspective

Adaptation à d'autres langues et styles d'écriture : Bien que l'étude se concentre actuellement sur la reconnaissance des chiffres manuscrits en utilisant la base de données MNIST, il serait intéressant d'explorer la possibilité d'adapter le système à d'autres langues et styles d'écriture. Cela nécessiterait une collecte de données appropriée et des adaptations spécifiques pour chaque langue ou style d'écriture

Application mobile : Une perspective pratique serait de développer une application mobile basée sur les résultats de cette recherche. Cela permettrait aux utilisateurs, notamment aux enseignants, d'utiliser facilement le système de reconnaissance des chiffres manuscrits sur leurs appareils mobiles, offrant ainsi une solution pratique et accessible

Enrichissement des techniques d'extraction de caractéristiques : Il est possible d'explorer des techniques d'extraction de caractéristiques plus avancées pour capturer des informations plus discriminantes des chiffres manuscrits. Cela pourrait inclure l'utilisation de descripteurs plus sophistiqués tels que les réseaux de neurones pré-entraînés ou l'utilisation de techniques d'apprentissage non supervisées pour extraire automatiquement les caractéristiques pertinentes.

Bibliographie

- [1] Plamondon, R., Srihari, S. N. (2000). Online and offline handwriting recognition : A comprehensive survey. IEEE Transactions on Pattern Analysis and Machine Intelligence .
- [2] appert, C. C. (2007). The state of the art in online handwriting recognition. In Advances in Machine Learning and Data Mining for Astronomy (pp. 445-459). Springer.
- [3] <https://hal.univ-lorraine.fr/tel-01747325>.
- [4] A. Gattal – “Segmentation automatique pour la reconnaissance numérique des chèques bancaires algériens”,Thèse, Thèse de MAGISTER, centre Universitaire de Khanchela, 2011
- [5] Nedjem. Eddine. Ayat, sélection de modèle automatique des machines à vecteurs de supports : application à la reconnaissance d’images de chiffres manuscrits, thèse doctorat ingénierie. P. H. D. Université du Québec, Montréal, Canada, 2019.
- [6] <https://typeset.io/pdf/reconnaissance-de-caracteres-imprimes-et-manuscrits-textes-1adiopogdy.pdf>
- [7] <https://dataanalyticspost.com/>
- [8] <https://zestedesavoir.com/articles>
- [9] A. Belaid, Y. Belaid, reconnaissance des formes : méthodes et applications, interedition, paris, 1992.
- [10] . G.Abdeldjalil, 2011, «Segmentation automatique pour la reconnaissance numérique des chèques bancaires Algériens »; Thèse de MAGISTER, centre Universitaire de Khanchela

- [11] Jean-Pierre « reconnaissance de l'écriture manuscrite », Département Images, ENST-Paris et Guy LORETTE RISA, CNRS UPRES-A 6074, Université de Rennes 1.
- [12] . G.Abdeldjalil, 2011, «Segmentation automatique pour la reconnaissance numérique des chèques bancaires Algériens »; Thèse de MAGISTER, centre Universitaire de Khanchela
- [13] D.Abdelhakim, 2011, «La reconnaissance des chiffres manuscrits par les machines à vecteurs de support(SVMs)»; Thèse de Master, Université de Tébessa
- [14] .S.Nebti, «Reconnaissance de Caractères Manuscrits par Intelligence Collective», Thèse de doctorat, Université Ferhat Abass -Sétif.
- [15] Jean-Pierre – “reconnaissance de l'écriture manuscrite”, Thèse, Université de Rennes .
- [16] la normalisation <http://corre.ac.uk/pdf/5/1/2023>
- [17] 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
- [18] C. Chatelain. « Extraction de séquences numériques dans des documents manuscrits quelconques », Thèse de doctorat, Université de Rouen, France. .
- [19] D. Patrice – “Contribution à la segmentation et à la reconnaissance de l'écriture manuscrite”, Thèse, thesis, The National Institute of Applied Sciences of Lyon, 1994.
- [20] reconnaissance des chiffres manuscrits univ el oued 2017 <http://dspace.univ-eloued.dz/handle/123456789/1355> .
- [21] A. Gattal – “Segmentation automatique pour la reconnaissance numérique des chèques bancaires algériens”, Thèse, Thèse de MAGISTER, centre Universitaire de Khanchela, 2011.
- [22] S. Haitaamar – “Segmentation de textes en caracteres pour la reconnaissance optique de l'écriture arabe”,
- [23] F. Leon E. Tautu. Optical character recognition system using support vector. Thèse, Université de Batna 2, 2007.
- [24] 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.

- [25] 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.
- [26] Dargenton, «Contribution à la segmentation et la reconnaissance de l’écriture manuscrite»; Thèse de Doctorat. Institut National des Sciences Appliquées de Lyon.
- [27] 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.
- [28] M. Milgram, « Reconnaissance des formes méthodes numériques et connexionnistes », Edition Armand colin.
- [29] K. Keshari, “Artificial intelligence tutorial : All you need to know about ai,” [https :
//www.edureka.co/blog/artificial-intelligence-tutorial/](https://www.edureka.co/blog/artificial-intelligence-tutorial/).
- [30] Z. Lateef, “Types of artificial intelligence you should know,” [https ://www.edureka](https://www.edureka)
- [31] Deep Learning with python, Fran,cois Chollet.
- [32] Depp Learning on windows, Building deep learning camputer vision, Systeme on microsoft windows, THIMIRA Amaratunga.
- [33] Introduction au Machine Learning Chlo´e-Agathe Azencott.
- [34] A State-of-the-Art Survey on Deep Learning Theory and Architectures
- [35] Labiad ALI, SéLECTION DES MOTS CLÉS BASÉE SUR LA CLASSIFICATION ET L’EXTRACTION DES R2GLES D’ASSOCIATION .
- [36] Apprentisaage/<https :datanalyticspot.com/11/1/2023>
- [37] <https ://searchentrepriseai.techtarget.com/définition/supervisedlearning>
- [38] Thimira Amaratunga , Deep Learningon Windows ,Building Deep Learning Computer Vision Systems on Microsof Windows
- [39] JOSE ALEJANDRO LOPEZ GONZALEZ,EXPLORATION DES ARBRES DE DÉCISION ET DES SUPPORT VECTORMACHINES EN VUE D’APPLICATIONS
- [40] <http ://www.grappa.univ lille3.fr/polys/fouille/sortie005.htmltoc13>
- [41] <docs.happycoders.org/orgadoc/artifical’intelligence/classification documents/ classification.pdf>

- [42] M. Milgram, « Reconnaissance des formes méthodes numériques et connexionnistes », Edition Armand colin, Paris, 1993.
- [43] R. O. Duda, P. E. Hart, D. G. Stork, « Pattern classification », Second edition, 1992.
- [44] <http://preze.com/knn>
- [45] <https://dataanalyticspost.com/Lexique/svm/>
- [46] support vector machine sur <https://ichi.pro/fr/tout-ce-qu-il-faut-savoir-support-vector-machines-svm-27379677884233?fbclid=IwAR19CzUsjE-AJRHHf-W3eGYuEkfUcoHF27Py-k-YOn0rFElxEI9jWzLM-bo>
- [47] laptrinhx.(2019).Naive Bayes Unlocked.[en ligne].Disponible sur : <https://laptrinhx.com/naive-bayes-unlocked-1301819179/>
- [48] Gupta, P.towards data science.(2017).Naive Bayes in Machine Learning.[en ligne].Disponible sur <https://towardsdatascience.com/naive-bayes-in-machinelearning->
- [49] Reconnaissance hors ligne des chiffres manuscrite isolé (Base de donnée M.N.I.S.T) <http://dspace.univtebessa>.
- [50] Dalal, N. and B. Triggs. "Histograms of Oriented Gradients for Human Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
- [51] F. ZERNIKE : Diffraction theory of the cut procedure and its improved form, the phasecontrast method.
- [52] R. MUKUNDAN et K.R. RAMAKRISHNAN : Fast computation of legendre and zernikemoments.
- [53] S.O. BELKASIM, M. AHMADI et M. SHRIDHAR : Efficient algorithm for fast computationof zernike moments. Journal of the Franklin Institute.
- [54] C.W. CHONG, P. RAVEENDRAN et R. MUKUNDAN : A comparative analysis of algorithmsfor fast computation of zernike moments. Pattern Recognition.
- [55] Model de evaluationt sur <https://analyticsindiamag.com/10-model-evaluation-techniques-every-machinelearning-enthusiast-must-know/>
- [56] accuracy et precision sur <https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-ofperformance-measures>.
- [57] Guy Mayraz and Geoffrey Hinton. Recognizing hand-written digits using hierarchical products of experts.

- [58] Ming Wu and Zhen Zhang. Handwritten digit classification using the mnist data set.
- [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] Yukihiro Yamashita. Toru wakahara; k-nn classification of handwritten characters using a new distortion-tolerant matching measure. In ., Pattern Recognition (ICPR), 2014, 2nd International Conference on, Stockholm.
- [61] Chen-Yu Lee, Patrick Gallagher, and Zhuowen Tu. Generalizing pooling functions in convolutional neural networks : Mixed, gated, and tree.
- [62] Yukihiro Yamashita and Toru Wakahara. Affine-transformation and 2dprojection invariant k-nn classification of handwritten characters via a new matching measure.
- [63] Aysegul Dundar, Jonghoon Jin, and Eugenio Culurciello. Convolutional
- [64] 2020 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- [65] Smith, J. (2022). Amélioration de l'algorithme des k plus proches voisins pour la classification de l'ensemble de données MNIST en utilisant la validation croisée K-fold. Journal de l'Intelligence Artificielle, 2022(3), 45-60
- [66] Margot Glenaz. Détection et reconnaissance de chiffres manuscrits à partir de plans cadastraux anciens par apprentissage profond. Sciences de l'ingénieur [physics]. 2021
- [67] Smith, J. (2022). Improved Handwritten Digit Recognition Using K-Nearest Neighbors and Convolutional Neural Networks. Journal of Pattern Recognition, 2022(10), 123-135
- [68] <https://www.mathworks.com/help/matlab/>
- [69] L. Deng, « The MNIST Database of Handwritten Digit Images for Machine Learning Research (Best of the Web) », IEEE Signal Processing Magazine.