MINISTERE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE ABBES LAGHROUR KHENCHELA
FACULTE DES SCIENCES ET DE LA TECHNOLOGIE

Département de Math et Informatique

N° de série :2021\2022

Mémoire de fin d'études

Pour l'obtention du diplôme de Master (L.M.D)

Spécialité: Informatique
Option: Génie Logiciel et Système Distribué

La reconnaissance d'écriture manuscrite par des techniques des CNNs

Réalisé par :
Samra Benhanachi

Membres de jury :
***Dr. Abbes* Fayçal**
Dr.Hammam Mounir

Dirigé par :
Dr.Souad Saadi

Résumé

A ce jour, il existe de nombreux systèmes de la reconnaissance de l'écriture manuscrits, ces systèmes ont deux axes majeurs, qui sont Détection de texte et Reconnaissance de texte. Cependant, ces systèmes sont contrôlés par diverses conditions. Cette mémoire propose donc une nouvelle méthode de reconnaissance d'écriture manuscrite basée sur un nouveau machine Learning approche qui est l'apprentissage par transfert.

Notre méthode comprend une grande étape, qui est représentée dans keras-ocr, il est implémenté avec CRAFT comme détecteur de texte et CRNN comme reconnaissance de texte, le réseau neuronal récurrent convolutif en bref CRNN est un modèle très populaire pour la reconnaissance de texte.

Notre méthode nous donne un résultat satisfaisant.

Mots clés : La Détection, La Reconnaissance, l'Apprentissage Automatique, Le Transfert d'Apprentissage, keras-ocr,CRNN, CRAFT.

Abstract

To date, there are many handwriting recognition systems, these systems have two major axes, which are Text Detection and Text Recognition. However, these systems are controlled by various conditions. This thesis therefore proposes a new method of handwriting recognition based on a new machine learning approach which is transfer learning.

Our method includes a big step, which is represented in keras-ocr, it is implemented with CRAFT as text detector and CRNN as text recognition, convolutional recurrent neural network in short CRNN is a very popular model for text recognition.

Our method gives us a satisfactory result.

Key words: Détection, Recognition, Machine Learning, Transfer Learning, keras-ocr, CRNN, CRAFT.

ملخص

حتى الآن ، هناك العديد من أنظمة التعرف على خط اليد ، وهذه الأنظمة لها محورين رئيسيين ، وهما اكتشاف النص والتعرف على النص. ومع ذلك ، يتم التحكم في هذه الأنظمة من خلال ظروف مختلفة. تقترح هذه الأطروحة طريقة جديدة للتعرف على خط اليد بناءً على نهج جديد للتعلم الآلي وهو نقل التعلم.

تتضمن طريقتنا خطوة كبيرة ، وهي ممثلة في keras-ocr ، ويتم تنفيذها باستخدام CRAFT ككاشف للنص و CRNN مثل التعرف على النص ، والشبكة العصبية التلافيفية المتكررة باختصار CRNN هي نموذج شائع جدًا للتعرف على النص. طريقتنا تعطينا نتيجة مرضية.

الكلمات الرئيسية: الاكتشاف ، والاعتراف ، والتعلم الآلي ، ونقل التعلم ، و keras-ocr ، و CRNN ، و CRAFT.

Remerciement

Tout d'abord, nous remercions Dieu Tout-Puissant, le Très Miséricordieux, le Très Miséricordieux

Qui m'a donné force, courage et volonté

Pour mener à bien cet humble travail, je remercie mes parents pour leurs sacrifices et leurs efforts

Je ne saurai jamais où j'en suis aujourd'hui

Pour exprimer ma profonde gratitude à mon encadrant

Dr. Souad Saadi pour son aide précieuse et

Conseils avisés, contrôles et directions. Et autorisation de conduire

Ce travail. Son soutien, ses compétences et sa clairvoyance m'ont beaucoup aidée.

Inestimable.

Tous les enseignants du Département d'informatique de l'Université de

Khenchela, sur la qualité de leurs enseignements et conseils

Cela nous a permis de poursuivre notre séquence académique jusqu'à présent ;

Enfin, nous tenons également à remercier toutes les personnes qui ont participé de

Près ou de loin à la réalisation de ce travail.

Merci à tous et à toutes

Dédicace

D'un sentiment plain d'amour, de sincérité de fidélité, Je dédie ce modeste travail à :

A mes chers parents ma mère **Yasmina** et mon père **Nacer**

Pour leur patience, leur amour, leur soutient et leurs encouragement

A mes très chers frères

A mes très chères sœurs

À tout ma famille sans exception ;

A mes meilleurs amis : **Maissa, Romaissa, Nihed**

Tous ceux qui m'aiment et que j'aime

Samra

Table des matières

Liste des figures

Introduction générale

Malgré le développement accéléré des méthodes de communication, l'écriture manuscrite reste un lien étroit et supérieur entre les personnes, elle permet aux gens d'échanger des sentiments, des informations et des idées de manière naturelle. De manière plus prosaïque, il met en relation les citoyens avec leurs administrations, les clients avec leurs établissements bancaires, les consommateurs avec les sociétés de distribution. Son importance peut être appréciée en permanence par le nombre de messages manuscrits postés quotidiennement, de déclarations fiscales, etc.

La reconnaissance et l'interprétation de l'écriture manuscrite sont des technologies relativement nouvelles qui s'inscrivent dans le cadre général de la communication écrite homme-machine.

La lecture automatique de l'écriture manuscrite présente un intérêt indéniable dans l'accomplissement des tâches fastidieuses comme celles que l'on rencontre dans certains domaines : le tri postal, la lecture de chèques bancaires, la lecture des bordereaux, des bons de commande, des feuilles de déclaration... Elle offre aujourd'hui un surcroît d'intérêt avec le développement des nouvelles méthodes permettant de communiquer directement avec la machine de manière plus naturelle, grâce à l'essor des agendas, des blocs-notes électroniques et des ordinateurs sans clavier.

A priori, on pourrait concevoir la reconnaissance du manuscrit comme une émanation de la reconnaissance de l'écriture imprimée ou dactylographiée, pour laquelle les principaux obstacles semblent avoir été surmontés. Mais le passage de la reconnaissance de l'imprimé à celle du manuscrit n'est pas aussi évident que certains avaient pu le supposer initialement. Sa mise en œuvre est beaucoup plus complexe. Les difficultés rencontrées proviennent principalement de la nature même de l'écriture manuscrite, de son caractère cursif et de son extrême variabilité.

Cependant, malgré ces difficultés, la concrétisation du problème à résoudre et la concrétisation de la tâche à accomplir permettent aujourd'hui d'aboutir à des réalisations de plus en plus tangibles.

Revenons un peu en arrière avec un petit tour d'horizon du domaine très vaste de l'intelligence artificielle qui nous amène à parler de l'évolution de l'informatique dans le temps,

L'intelligence artificielle comprend de nombreux sous-domaines, tels que l'apprentissage en profondeur, qui est souvent appelé apprentissage en profondeur qui est une branche de l'apprentissage automatique ou de l'apprentissage automatique, après avoir passé beaucoup de temps à oublier, l'apprentissage en profondeur a refait surface grâce à l'amélioration de la la puissance de calcul des ordinateurs et l'émergence de nouvelles bases de données plus vastes et plus riches. L'objectif de notre travail est de développer un système de reconnaissance plus précise de L'écriture manuscrite à l'aide de techniques d'apprentissage en profondeur.

Notre mémoire se subdivise donc comme suit :

Chapitre1 : Dans le premier chapitre, nous avons invoqué quelques notions sur l'imagerie et des définitions sur l'intelligence artificielle ainsi que ses types, mais nous nous sommes surtout intéressés sur l'apprentissage automatique et l'apprentissage profond.

Chapitre2 : le deuxième chapitre expose les réseaux de neurone en terme général et les réseaux de neurones convolutif (CNN) en spécifique.

Chapitre3 : le troisième chapitre traite en global la reconnaissance de l'écriture manuscrite

Chapiter4 : le quatrième chapitre présent Les détails d'implémentation. Ce chapitre définit l'architecture de modèle qu'on a créé pour détecter la reconnaissance de l'écriture manuscrite.

Chapiter5 : conclut ce mémoire en présentant l'ensemble des résultats expérimentaux obtenus.

Une conclusion synthétiser les contributions et les résultats obtenus et proposera des Perspectives à nos travaux.

Chapitre 1

L'intelligence artificielle

1.1 Définition de l'intelligence artificielle

L'utilisation du terme intelligence artificielle prend note de l'émergence de l'IA dans les années 1950 et avec l'avancement de la technologie informatique et sa diffusion dans les activités humaines, l'IA est devenue une partie de notre culture dans une certaine mesure depuis lors. [9] S'il n'y a pas de définition générale, il y a deux axes principaux à suivre, le premier impliquant le processus de réflexion et de raisonnement, tandis que le second est basé sur le comportement, ainsi plusieurs définitions se dégagent :

1. **Des systèmes qui pensent comme les humains :**
   « La tentative nouvelle et passionnante d'amener les ordinateurs à penser ... d'en faire des machine dotées d'un esprit au sens plus le plus littéral »[Haugeland 1985]
   « L'automatisation d'activités que nous associons a la pensée humaines, des activité telle que la prise de décisions, la résolution de problème, l'apprentissage.. »[Bellman 1978
2. **Des systèmes qui pensent rationnellement :**
   « L'étude de facultés mentales grâce des modèles informatique»[Charniak et al 1985]
   « L'étude de moyens informatique qui rendent possible la perception le raisonnement et l'action »[Winston 1992]
3. **Des systèmes qui agissent comme des humains :**
   « L'art de créer des machines capables de prendre en charge des fonctions exigeant de l'intelligence quand elle sont réalisés par des gens » [Kurzweil 1990]
   « L'étude de moyens à mettre en ÷uvre pour faire en sorte que des ordinateurs accomplissent des choses pour lesquelles il est préférable de recourir à des personnes pour le moment »[Rich et al 1991]
4. **Des systèmes qui agissent rationnellement :**
   L'intelligence artificielle (computational intelligence) estl'étudede laconception d'agents intelligents [pool et al 1998]
   « L'IA ..... Étudie le comportements intelligent dans des artefacts » [Nillson 1998]

1.1.1 Type d'intelligence artificielle

Les scientifiques ne sont pas sur les mêmes longueurs d'onde concernant la notion d'intelligence articielle ils ont même redéni les deux grands en IA faible et IA forte :

- **Intelligence artificielle faible :**
  C'est des systèmes autonomes qui simulent l'intelligence mais qui ne sont pas du tout intelligents.
- **Intelligence artificielle forte :**
  On parle de machine intelligente qui pense en d'autres termes des machines qui éprouve une sensation de conscience, qui seront capables de raisonner de prendre des décisions et

de juger.
L'apprentissage est un aspect fondamental de l'IA, c'est grâce à lui qu'un système intelligent améliore ces performances avec de l'expérience, contrairement à l'approche manuelle ou on a connu quelques limites puisqu'on écrivait des programmes à la main pour la reconnaissance de caractères imprimés, jouez aux échecs etc. [31]

1.2 Brève histoire de l'IA

- **1943** McCulloch Pitts : modélisation des neurones
- **1950** "Computing Machinery and Intelligence" de A. Turing
- **1952-69** Période euphorique (une période euphorique de croissance économique)!
- **1950s** 1er programmes d'IA : programme de jeu de dame de A. Samuel, "Logic Theorist" de A. Newell H. Simon
- **1956** réunion de Dartmouth : le terme "Intelligence Artificielle" est adopté
- **1965** algorithme de J. A. Robinson pour le raisonnement logique 1966-74 IA découvre la complexité des calculs recherche en réseaux de neurones disparaît (presque) complètement
- **1969-79** 1er développements de système à base de connaissances
- **1980-88** industrie des systèmes experts en plein boum
- **1988-93** industrie des systèmes experts s'effondre : "Hiver de l'IA"
- **1985-95** réseaux de neurones redeviennent populaires
- **1988**- résurgence des méthodes de décision probabilistes "Nouvelle IA" : vie artificielle, algorithmes génétiques, ...
- **1995**- Agents, agents partout... [35]

1.3 Grands domaines de l'IA...

- **Reconnaissance et synthèse de la parole**
  ex : réservation d'hôtel, annuaire téléphone
- **Reconnaissance et synthèse d'images**
  ex : effets spéciaux au cinéma, vidéo-surveillance
- **Reconnaissance de l'écriture**
  ex : reconnaissance chèques, codes postaux
- **Langage naturel**
  ex : interfaces, text mining, Web Mining
- **Planification**
- **Aide à la décision**
  ex : contrôle de trajectoire du satellite Voyager
- **Aide à la programmation**
  ex : agents d'interface
- **Apprentissage / Adaptation**
  ex : construction de systèmes experts, classification automatique de galaxie, contrôleurs de robots ...
- **Jeux**
  ex : Echecs (DeeperBlue à 2600), Checkers (Champion), Othello (Champion), BackGammon (champion), GO (bon amateur).
- **Médecine**
  ex : Aide à la décision (Systèmes experts), prédiction de patients à risques, analyse automatique d 'images médicales [35]

1.4 Apprentissage automatique :

Le fameux apprentissage automatique (AA ou ML pour machine learning) est un sous-ensemble de l'intelligence artificielle (IA), à partir de là on peut dire que tout ML est IA, mais pas l'inverse, le machine learning ou apprentissage automatique nous permet de résoudre Le vieux problème d'être trop complexe pour les humains est qu'au lieu de dire aux ordinateurs comment gérer et résoudre les problèmes, l'apprentissage automatique permet aux ordinateurs d'apprendre à résoudre les problèmes par eux-mêmes. [5]

1.5 Pourquoi l'apprentissage automatique ?

L'apprentissage automatique (ML) utilise des ordinateurs pour simuler l'apprentissage humain et permet aux ordinateurs de reconnaître et d'acquérir des connaissances du monde réel et d'améliorer les performances de certaines tâches en fonction de ces nouvelles connaissances. [32]
Pour mieux comprendre l'utilité de l'apprentissage automatique, considérez les domaines d'application de l'apprentissage automatique : les voitures autonomes Google, la détection de fraude en ligne, les moteurs de recommandation en ligne - comme les recommandations d'amis Facebook, Netflix montrant vos films et émissions préférés, et "Plus à considérer". choses » et « procurez-vous un petit quelque chose » sur Amazon sont des exemples d'apprentissage automatique appliqué.
Tous ces exemples reflètent le rôle important que l'apprentissage automatique a commencé à jouer dans le monde d'aujourd'hui qui est riche en données. Les machines peuvent aider à filtrer les informations utiles qui pourraient conduire à des percées majeures, et nous avons vu comment cette technologie est mise en œuvre dans une variété d'industries.
Le flux présenté ci-dessous représente le fonctionnement de l'apprentissage automatique. [33]

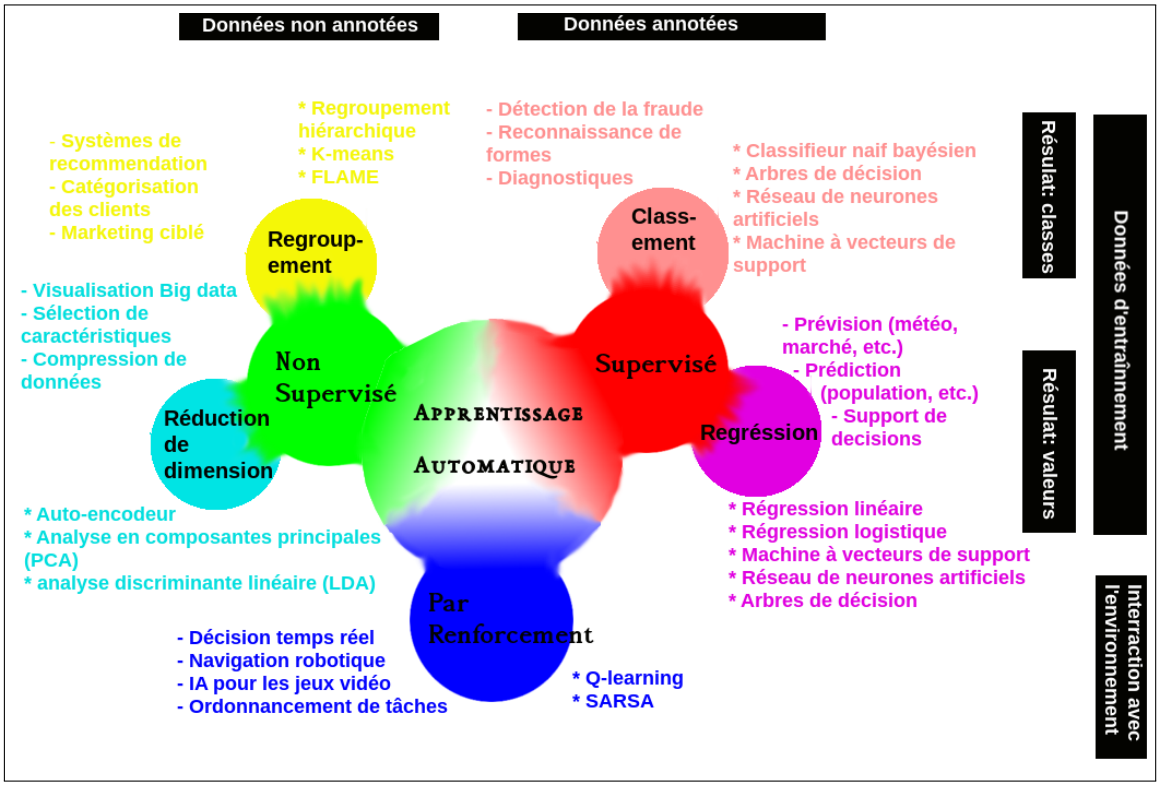


FIGURE 1.1 – Processus D'apprentissage Automatique.

1.6 Types de systèmes d'apprentissage automatique

Il existe différents types de systèmes d'apprentissage automatique, et il est utile de les répartir dans les grandes catégories suivantes :

- Qu'ils soient entraînés ou non avec une supervision humaine (supervisée, non supervisée, semi-supervisée et apprentissage par renforcement).
- Qu'ils puissent ou non apprendre de manière incrémentale à la volée (en ligne ou par apprentissage par lots).
- Fonctionne-t-il simplement en comparant des nouveaux points de données à des points de données connus, ou détecte-t-elle les modèles dans les données d'apprentissage et construit un modèle prédictif, comme le font les scientifiques (apprentissage basé sur l'instance ou modèle). [23]

1.7 Types d'apprentissage automatique (supervision humaine)

Les systèmes d'apprentissage automatique peuvent toujours être classés comme pertinents Il existe quatre catégories de critères de qualité et d'orientation appropriée qu'ils reçoivent au cours du processus de structuration Principales catégories Apprentissage supervisé Apprentissage supervisé Apprentissage semi-supervisé Observer et apprendre en consolidation

1.7.1 Apprentissage supervisé

Le principe de la méthode est qu'il existe une bibliothèque d'apprentissage, qui contient un ensemble de données ou d'exemples étiquetés, auxquels un expert (superviseur) attribue des catégories, le but des méthodes d'apprentissage supervisé est de générer automatiquement des règles de classification ou des bases fonctionnelles à partir des apprentissage, ces règles permettront de classer les objets. [20]

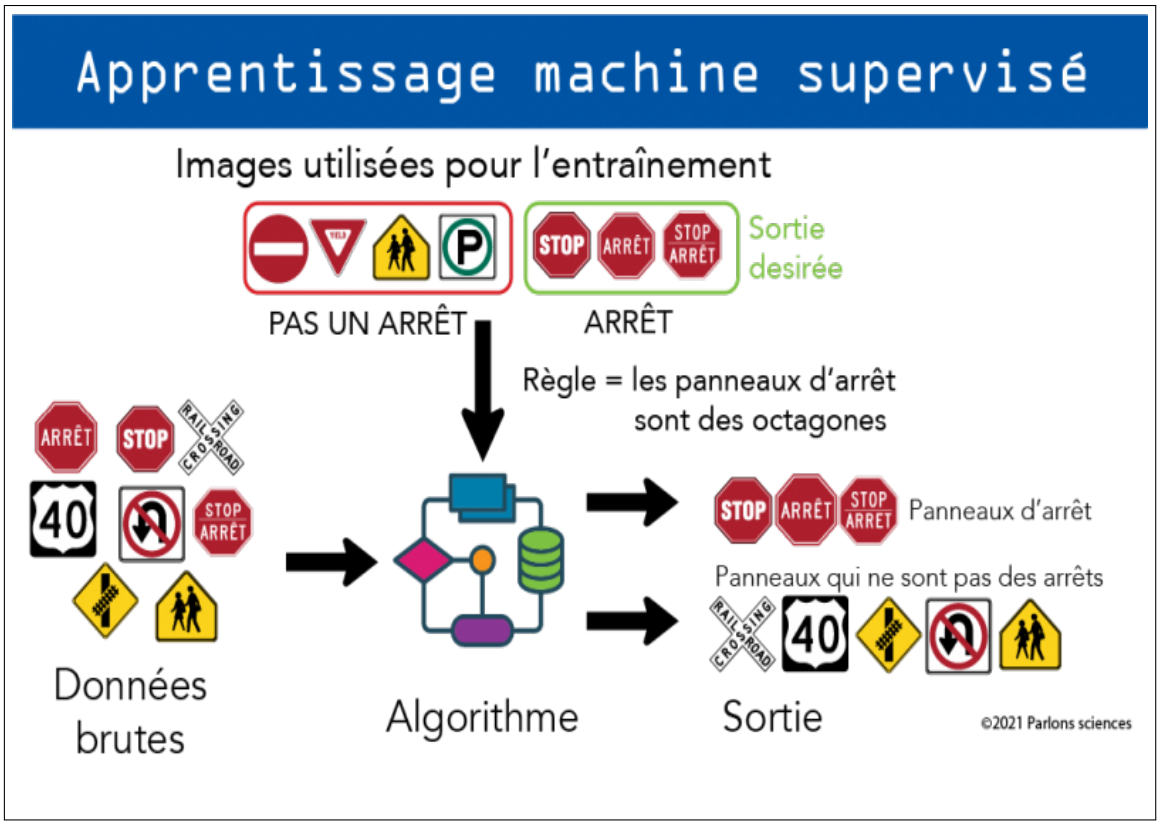


FIGURE 1.2 – Schématisation de l'apprentissage supervisé

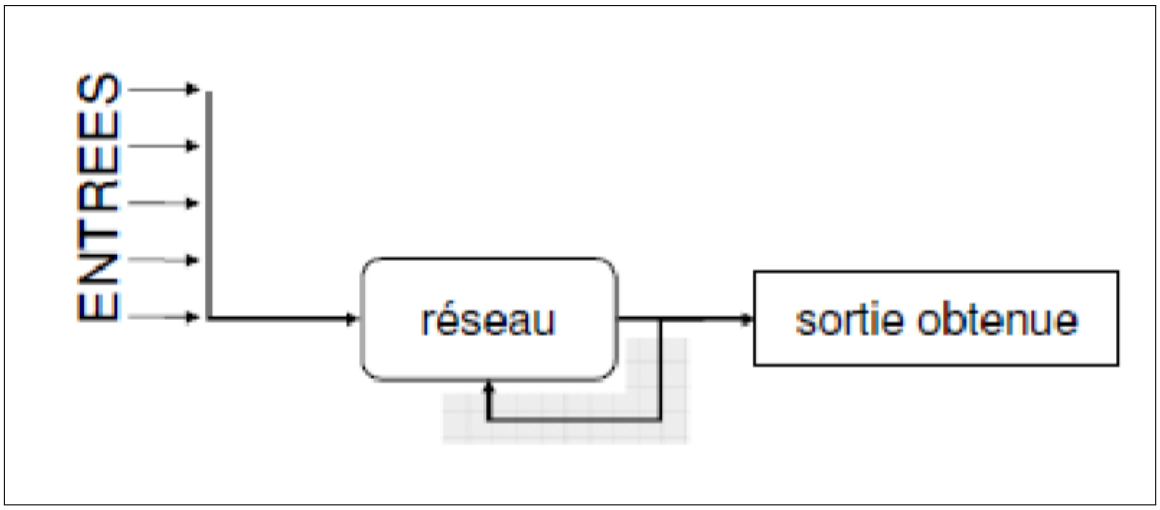


FIGURE 1.3 – Schématisation de l'apprentissage supervisé

les algorithmes d'apprentissage supervisé les plus importants :

- k-Plus proches voisins
- Régression linéaire
- Machines vectorielles de support (SVM).
- Arbres de décision et forêts aléatoires
- Les réseaux de neurones. [20]

Applications de la classification

- Reconnaissance des formes
  - Reconnaissance de visages : reconna^ıtre les personnes malgré les variations (pose, éclairage, lunettes, maquillage, coiffure)
  - Reconnaissance de caractéres manuscrits : reconnaitre malgré différents styles d'écriture
  - Reconnaissance de la parole : d´ependance temporelle de l'information, utiliser des dictionnaires de mots/structures valides
- Aide au diagnostic médical : d´eterminer les problèmes m´edicaux à partir des sympt^omes
- Extraction de connaissances : expliquer par des r'egles simples des masses de donn´ees
- Compression : repr´esenter par des mod'eles compacts les donn´ees
- D´etection d'irr´egularit´es : identifier des fraudes, des intrusions

FIGURE 1.4 – Reconnaissance de caractères

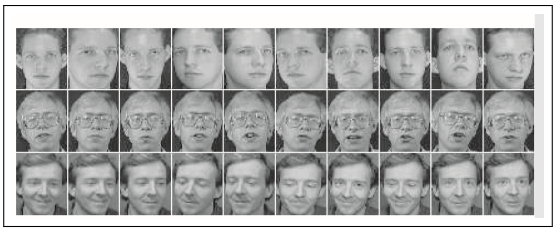


FIGURE 1.5 – Reconnaissance de visages

1.7.2 l'apprentissage non-supervisé

Appelé aussi apprentissage à partir d'observations, il s'agit d'une construction automatique des classes sans avoir recours a l'expert c'est à dire qu'il n'aurait pas de phase d'entrainement, on dispose dans ce cas d'une masse de données importante non étiqueté (sans indiquer leurs classes) qui sont introduites par l'utilisateur donc l'algorithme doit découvrir par lui-même la structure des donnés.[5][20]

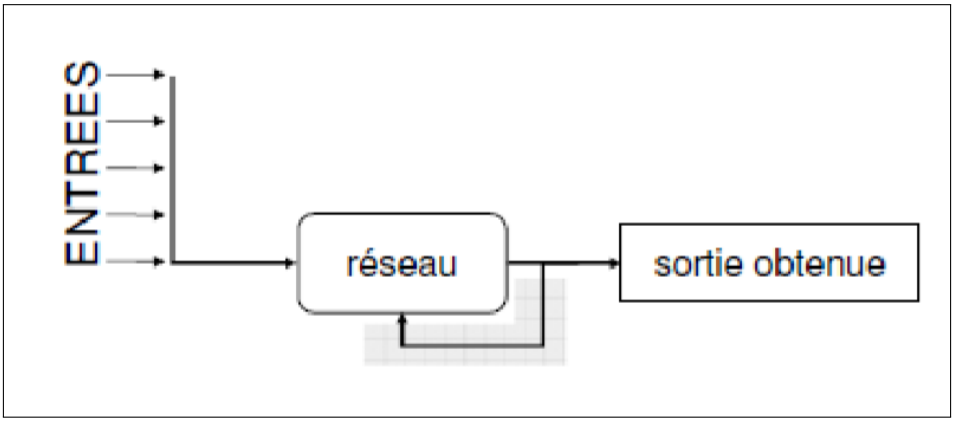


FIGURE 1.6 – Schéma d'un Modèle non-Supervisé.

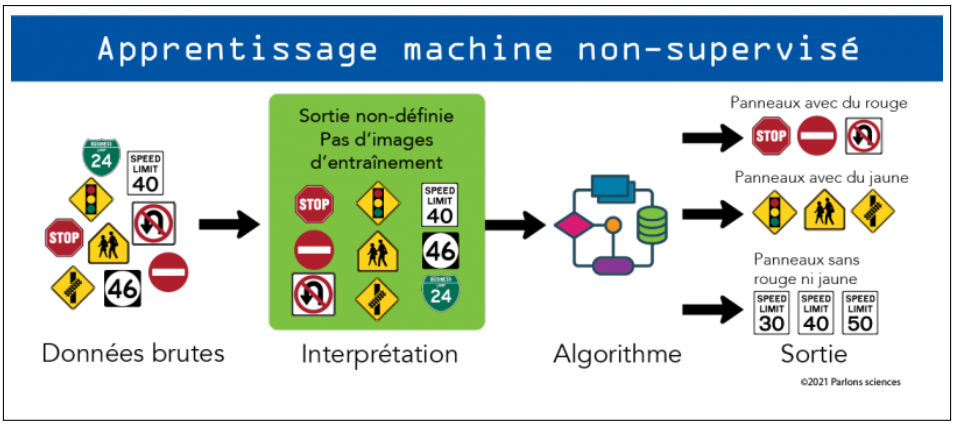


FIGURE 1.7 – Explique le Déroulement d'Apprentissage Non-Supervisé.

application :

- Segmenter les consommateurs dans les bases de données d'achats
- Compression d'images : réduire le nombre de couleurs utilisées
- Bio-informatique : découvrir des patrons dans l'ADN.

1.7.3 Apprentissage par renforcement

- Apprendre une politique : séquences d'actions
- L'apprentissage n'est pas supervisé, mais une récompense avec un délais est obtenue
- Problème d'assignation du crédit : quelles actions ont permis d'obtenir la récompense ?

application :

- Jeux, avec un ou plusieurs participants
- Robotique : navigation dans un environnement
- Agents : prises de décision

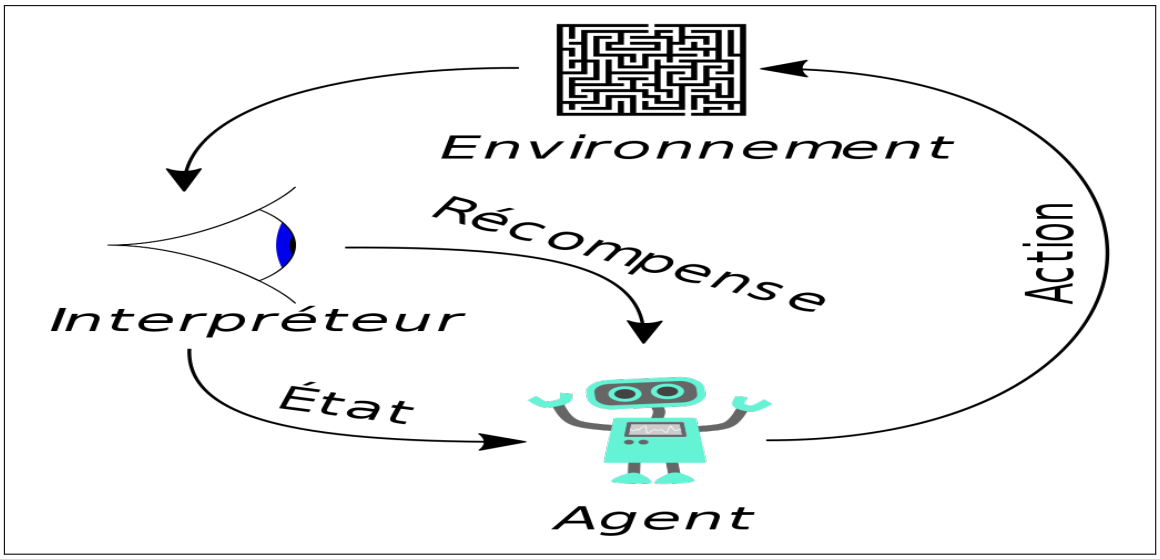


FIGURE 1.8 – Schéma d'un Modèle Par Renforcement.

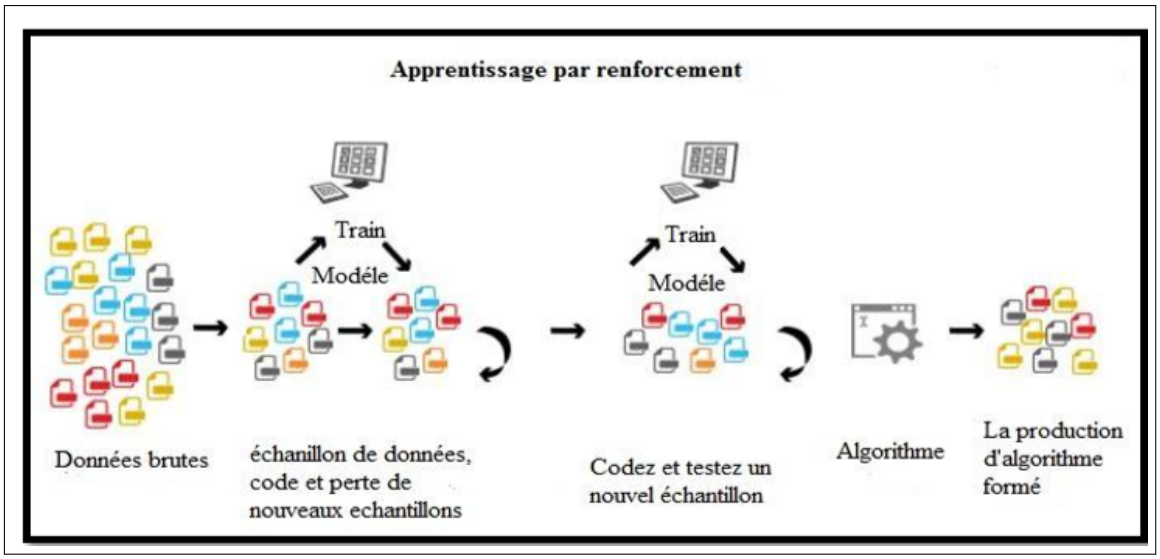


FIGURE 1.9 – Explique le Déroulement d'Apprentissage Par Renforcement.

1.7.4 Apprentissage Semi-Supervisé

Cette technique d'apprentissage est une alternative entre l'apprentissage supervisé et non supervisé. Dans certains cas, seulement une partie de données est signalées . Une grande quantité de données peuvent être divisées en deux groupes à l'aide d'un apprentissage non supervisé. Par conséquent, nous pouvons étiqueter ces données en fonction des caractéristiques de chaque cluster. Une fois les données étiquetées, l'apprentissage supervisé est utilisé pour la classification.

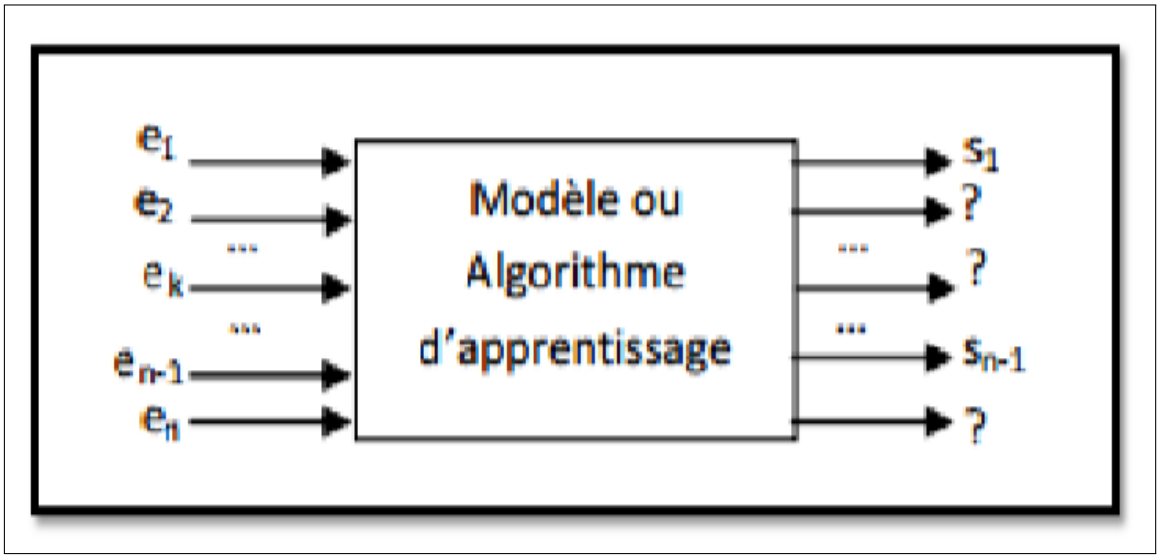


FIGURE 1.10 – Schéma d'un Modèle semi-Supervisé.

1.8 L'apprentissage profond

Appelé aussi Deep learning, c'est un sous-ensemble de l'apprentissage automatique (AA) donc on peut dire que tout Apprentissage profond est apprentissage automatique mais pas le contraire.

1.8.1 Définition :

L'apprentissage Profond limite la capacité du cerveau humain à traiter les données. Il fait référence aux algorithmes d'apprentissage automatique (ML) qui prennent en charge le traitement hiérarchique. Ce sont des techniques de traitement multi-étapes avec plusieurs couches de non-linéarité, d'où la notion de profondeur.

La première couche présente les données d'entrée et est appelée couche visible, quant aux couches suivantes, elles sont marquées comme masquées.

Une couche se compose de plusieurs neurones, où chaque neurone est connecté à d'autres neurones de la couche précédente par des connexions appelées poids, en utilisant une fonction non linéaire appelée fonction d'activation, le neurone traitera les poids d'entrée, chaque neurone caché Un neurone produit un sortie qui sera ensuite traitée par d'autres neurones.

Une couche se compose de plusieurs neurones, où chaque neurone est connecté à d'autres neurones de la couche précédente par des connexions appelées poids, en utilisant une fonction non linéaire appelée fonction d'activation, le neurone traitera les poids d'entrée, chaque neurone caché Un neurone produit un sortie qui sera ensuite traitée par d'autres neurones.
[10]

1.8.2 Apprentissages profonds supervisés

L'apprentissage profond supervisé consiste à réaliser des prédictions et des classifications sur les données d'entrée en fonction de la sortie souhaitée. Dans la phase de classification les valeurs cible fournie (sortie souhaitée) sont des valeurs discrètes qui représentent des étiquettent de données.

Le modèle d'apprentissage désire alors à classer les données d'entrer dans les catégories correspondantes.

En revanche dans la phase de prédiction les valeurs cibles sont continues et le modèle d'apprentissage consiste à modéliser la relation entre la cible et les variables d'entrées.

Les taches de classification et de prédictions sont exécutées de manière hiérarchique, en essayant de minimiser un objectif prédéfini, ou une fonction de perte, en utilisant l'algorithme de rétro propagation.[10]

1.8.3 Apprentissages profonds non supervisé

L'apprentissage profond non supervisé (unsuperviseddeeplearning) est la classe d'apprentissage la plus adaptée pour l'entrainement sur des bases de données à grandes échelles, pendant cet entrainement les informations sur les étiquettes ne sont pas utilisées en raison de leur indisponibilité et pour le cout de leur acquisition, la méthode la plus connue pour cet apprentissage est le deepbelief network.[10]

1.8.4 Quelques algorithmes de Deep Learning

Il existe différents algorithmes de Deep Learning. Nous pouvons ainsi citer :

- Les réseaux de neurones profonds (Deep Neural Networks).Ces réseaux sont similaires aux réseaux MLP mais avec plus de couches cachées. L'augmentation du nombre de couches, permet à un réseau de neurones de détecter de légères variations du modèle d'apprentissage, favorisant le sur-apprentissage ou sur-ajustement (« overfitting »).
- Les réseaux de neurones convolutionnels (CNN ou Convolutional Neural Networks). Le problème est divisé en sous parties, et pour chaque partie, un «cluster» de neurones sera créer afin d'étudier cette portion spécifique. Par exemple, pour une image en couleur, il est possible de diviser l'image sur la largeur, la hauteur et la profondeur (les couleurs).

- La machine de Boltzmann profonde (Deep Belief Network) : Ces algorithmes fonctionnent suivant une première phase non supervisée, suivi de l'entraînement classique supervisé. Cette étape d'apprentissage non-supervisée, permet, en outre, de faciliter l'apprentissage supervisé. [41]

1.8.5 La différence entre l'apprentissage profond et l'apprentissage automatique

L'apprentissage en profondeur est un sous-ensemble de l'apprentissage automatique et est l'une des 15 méthodes différentes. Tout apprentissage profond est un apprentissage automatique, mais tout apprentissage automatique n'est pas un apprentissage profond [27].

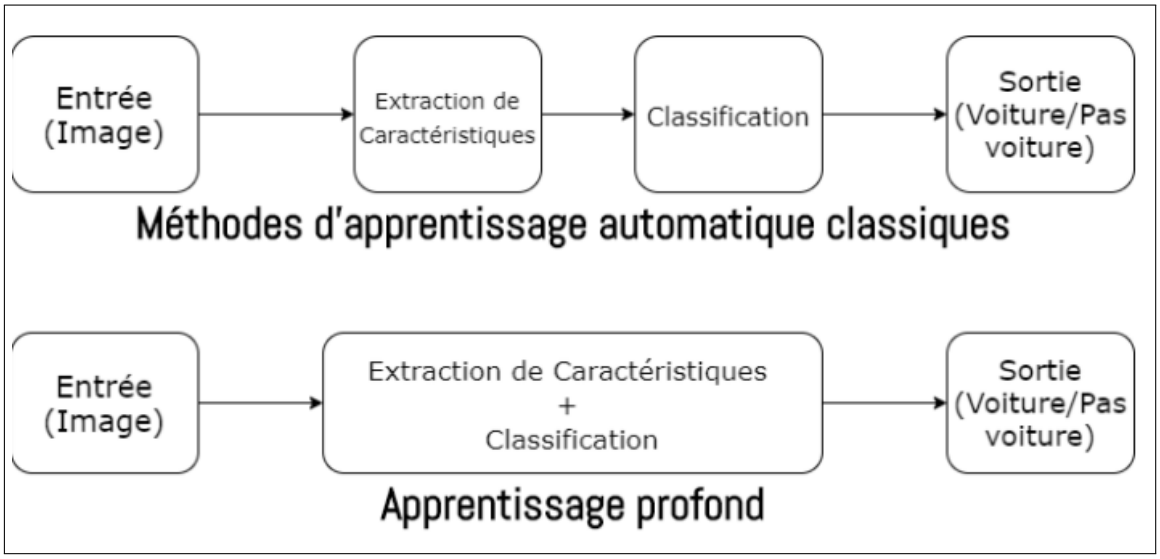


FIGURE 1.11 – La Différence Entre l'Algorithme de l'Apprentissage Profond et l'Apprentissage Automatique.

	Apprentissage automatique	L'apprentissage en profondeur
1	Le Machine Learning est un sur-ensemble du Deep Learning	Le Deep Learning est un sous-ensemble du Machine Learning
2.	Les données représentées dans le Machine Learning sont assez différentes de celles du Deep Learning car elles utilisent des données structurées	La représentation des données utilisée dans le Deep Learning est assez différente car elle utilise des réseaux de neurones (ANN).
3.	L'apprentissage automatique est une évolution de l'IA	Le Deep Learning est une évolution vers le Machine Learning. Fondamentalement, c'est la profondeur de l'apprentissage automatique.
4.	L'apprentissage automatique se compose de milliers de points de données.	Big Data : Des millions de points de données.
5.	Sorties : valeur numérique, comme la classification du score	Tout, des valeurs numériques aux éléments de forme libre, tels que le texte libre et le son.
6.	Utilise divers types d'algorithmes automatisés qui se tournent vers des fonctions de modèle et prédisent les actions futures à partir des données.	Utilise un réseau de neurones qui transmet les données à travers les couches de traitement pour interpréter les caractéristiques et les relations des données.
7.	Les algorithmes sont détectés par les analystes de données pour examiner des variables spécifiques dans des ensembles de données.	Les algorithmes sont largement autodécrits lors de l'analyse des données une fois qu'ils sont mis en production.
8.	Le Machine Learning est très utilisé pour rester dans la compétition et apprendre de nouvelles choses.	L'apprentissage profond résout des problèmes complexes d'apprentissage automatique.

FIGURE 1.12 – La Différence Entre l'Algorithme de l'Apprentissage Profond et l'Apprentissage Automatique. [3]

1.9 Conclusion

Dans ce chapitre,nous représentons trois segments : l'intelligence artificielle, l'apprentissage automatique et l'apprentissage en profond. En intelligence artificielle, nous proposons quelques définitions, des historiques, les Type d'intelligence artificielle et leurs grands domaines. Ensuite, en machine learning, nous exposons quelques définitions, pourquoi le machine learning , son type et son modèle. Enfin, dans la section sur l'apprentissage en profond, nous définissons l'apprentissage en profond, son fonctionnement ,son type et la différence entre l'apprentissage en profond et l'apprentissage automatique. Dans le chapitre suivant,on vas expliquer Les Réseaux de neurones.

Chapitre 2

Les Réseaux de neurones

2.1 Introduction

Les réseaux de neurones sont des outils puissants pour modéliser des relations complexes entre les données d'entrée et de sortie. La technologie a été développée dans l'ambition de construire un système artificiel capable d'effectuer des tâches "intelligentes" d'une manière similaire au cerveau humain. Un réseau de neurones est similaire au cerveau humain dans le sens qu'il apprend en apprenant et que ses connaissances sont stockées dans des connexions d'interneurones appelées "poids synaptiques". L'avantage des réseaux de neurones est leur capacité à apprendre à résoudre des problèmes complexes en modélisant des exemples d'apprentissage réels[30].

Les réseaux de neurones ont une capacité extraordinaire à tirer un sens de données incompréhensibles ou imprécises, et peuvent reconnaître et détecter des formes complexes difficiles à détecter par d'autres techniques. Un réseau de neurones bien formé peut être considéré comme un "expert" capable de classer les informations à analyser et d'apporter des réponses à de nouvelles informations ou situations.

Les réseaux de neurones sont largement utilisés pour la reconnaissance de caractères en raison de leur capacité de généralisation, selon leurs différents types d'architectures, comme les travaux suivants avec des taux de reconnaissance correspondants : Multilayer Perceptron Recognition of Handwritten Characters (Latin Alphabet) [14].

Ce chapitre présente quelques modèles architecturaux de réseaux de neurones artificiels et leur fonctionnement, en particulier les réseaux de neurones convolutionnels, les cartes auto- organisatrices, les réseaux de neurones probabilistes, les réseaux de neurones récurrents et les réseaux de neurones temporels.

2.2 Historique de Réseaux de neurones

- **1890 :** La loi de fonctionnement pour l'apprentissage est présentée par W. James.
- **1943 :** Warren Mc Culloch et Walter Pitts proposent le premier modèle du neurone formel.
- **1949 :** Règle de Donald Hebb qui décrit que si les neurones d'une synapse sont activés d'une façon synchrone et répétée, la force de connexion synaptique est croissante .
- **1958 :** Création du premier réseau de neurones par Rosenblatt, le "perceptron", ce dernier est inspiré du système visuel. Il permet d'apprendre et d'identifier des formes simples et aussi de Calculer certaines fonctions logiques. 1969 : Marvin Minsky et Seymour Papert montrent les limites du perceptron, surtout à l'incapacité de résoudre des problèmes non linéairement séparables (exemple du XOR). Période noire des Réseaux de neurones (15 ans) et beaucoup de déception chez les amateurs de l'intelligence artificielle.
- **1982 :** John Hopfield réactive l'intérêt de l'utilisation de ce domaine grâce à sa découverte sur l'utilisation des réseaux récurrents (feed-back) après la première classe du perceptron .

- **1975 :** Werbos propose l'idée d'une possible utilisation d'une rétropropagation du gradient.
- **1985-1986 :** Le Cun et Parker (1985) continuent leurs recherches du pricipe de Werbos (1975), mais les travaux de Rumelhart (1986) furent le vrai départ de l'apprentissage des réseaux de neurones multicouche avec la méthode de rétropropagation du gradient .[16]

2.3 Définition

Les réseaux de neurones sont l'un des algorithmes d'apprentissage automatique les plus populaires aujourd'hui. Au fil du temps, il a été prouvé sans équivoque que les réseaux de neurones surpassent les autres algorithmes en termes de précision et de vitesse. Avec l'avènement de diverses variantes telles que les CNN (Convolutional Neural Networks (CNN en abrégé)), les RNN (Recurrent Neural Networks), les AutoEncoders, etc., les réseaux de neurones deviennent progressivement automatisés pour les scientifiques ou les praticiens de l'apprentissage [25].

2.3.1 Réseaux de neurone biologique

Les neurones sont des cellules nerveuses. Il se compose d'un corps cellulaire appelé "Soma" qui contient un noyau étendu (où se déroulent d'importantes activités cellulaires) appelé "Neurite". Ces derniers sont de deux types, les dendrites comme canaux d'entrée et les axones comme seuls canaux de sortie. D'un point de vue fonctionnel, un neurone est considéré comme une entité polarisée, c'est-à-dire que l'information ne circule que dans une seule direction : des dendrites vers les axones. Par conséquent, un neurone recevra des informations d'autres neurones via ses dendrites. Toutes ces informations seront ensuite agrégées par des potentiels d'action (signaux électriques) au niveau du corps cellulaire. Les résultats de l'analyse suivront l'axone jusqu'au terminal synaptique.A cet endroit, lorsque le signal arrivera, la vésicule synaptique fusionnera avec la membrane cellulaire, ce qui permettra la libération de neurotransmetteurs (médiateurs chimiques) dans la fente synaptique. Puisque les signaux électriques ne peuvent pas traverser les synapses (dans le cas des synapses chimiques), les neurotransmetteurs permettent aux informations de passer d'un neurone à l'autre. Les synapses ont une sortie "mémoire" qui leur permet d'ajuster leur fonction. Selon leur "historique", c'est-à-dire s'ils se déclenchent à plusieurs reprises entre deux neurones, les connexions synaptiques changeront en conséquence. Par conséquent, les synapses vont faciliter ou non le passage de l'influx nerveux. Cette plasticité est à l'origine du mécanisme d'apprentissage [21].

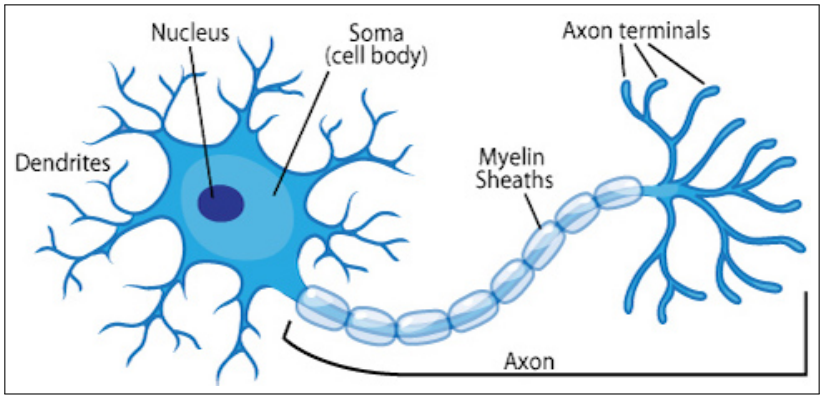


FIGURE 2.1 – RN biologique

2.4 Réseaux de neurones artificiels(formel) :

2.4.1 Définition

Un neurone artificiel est une unité de base qui reçoit de nombreuses entrées ou sorties d'autres neurones du réseau, avec des poids associés à chacune de ces entrées représentant la puissance de sa connexion au neurone. La figure 1.2 montre le modèle du neurone artificiel.

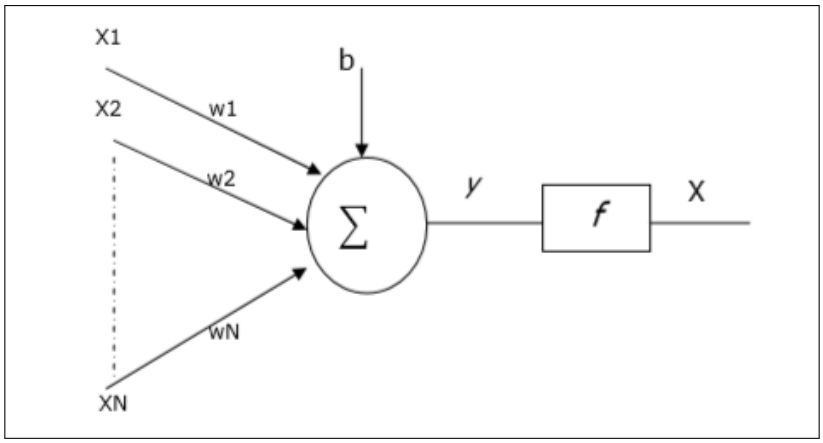


FIGURE 2.2 – Le neurone artificiel

La sortie d'un neurone est une somme pondérée de ses entrées plus le biais, propagée à travers une fonction d'activation, qui peut être (sigmoïde, tangente ou hyperbolique, etc.) [19]

$$\sum_{i=1}^{n} w_i x_j + b$$

$$x = f(y)$$

xj est la valeur de la jème entrée du neurone, Wj est le poids synaptique correspondant au neurone j, b est le biais et f(.) est la fonction d'activation. La fonction la plus couramment utilisée est de type sigmoïde [24]. Il est défini comme :

$$f(y) = \frac{1}{1 + e^{-y\delta}}$$

Où δ dénote le paramètre de la sigmoïde qui définit le degré de non-linéarité.

2.4.2 Topologies

L'unité de calcul de base dans un réseau de neurones artificiels est le neurone, défini en 1959. Un neurone artificiel reçoit une entrée d'un autre neurone ou d'une source externe avec des valeurs x1, x2, ..xn. Connectez-vous via les synapses et calculez la sortie y. Chaque entrée xi a un poids associé wi, qui est attribué en fonction de son importance relative par rapport aux autres entrées. La valeur d'entrée x d'un neurone est une somme pondérée de ses entrées, en ajoutant une autre entrée avec un poids b appelé biais. Le neurone applique alors une fonction f sur cette somme.
[11] Comme le montre la FIGURE 1.2, un réseau de neurones à propagation avant peut être constitué de trois types de neurones :

- Neurones d'entrée : les neurones d'entrée fournissent des informations provenant du monde extérieur au réseau et, ensemble, ils forment une "couche d'entrée". Aucun calcul n'est effectué dans le nœud d'entrée. Ils ne font que transmettre des informations aux nœuds cachés.
- Neurones cachés : les neurones cachés n'ont aucun lien direct avec le monde extérieur (d'où le nom "caché"). Ils effectuent des calculs et transmettent des informations des neurones d'entrée aux neurones de sortie. Un groupe de neurones cachés forme une "couche cachée". Bien que le réseau de propagation vers l'avant n'ait historiquement eu qu'une seule couche d'entrée et une seule couche de sortie, il peut avoir zéro ou plusieurs couches cachées selon le principe de généralisation.
- Neurones de sortie : les neurones de sortie forment ensemble une "couche de sortie". Ils sont responsables du calcul et de la transmission des informations du réseau vers le monde extérieur.[11]

Un réseau de neurones peut être défini de la manière suivante. Soient :

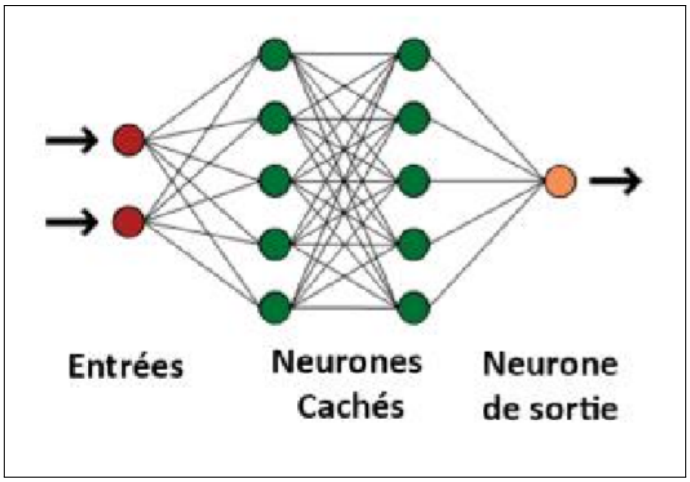


FIGURE 2.3 – les Couche d'un RN

- R le nombre des neurones sur la couche d'entrée
- S le nombre des neurones sur la couche de sortie
- L le nombre total des couches du réseau de neurones
- Nk le nombre des neurones sur la k ème couche avec 0 k l tel que n0 = r et nl = s
- les fonctions affines Ak i (x) = Pnk1 j=1 w k i j xj +b k i avec 1 k l, 1 i nk,x R nk1
- F la fonction d'activation du réseau de neurones
- Le vecteur de sortie Y = (y1,..., ys) (historique)

Fonction d'activation

Les fonctions d'activation (ou fonctions de seuil, voire fonctions de transfert) sont utilisées pour introduire de la non-linéarité dans le fonctionnement des neurones. Les fonctions de seuil ont généralement trois plages :

- En dessous du seuil, le neurone est non actif (souvent dans ce cas, sa sortie vaut 0 ou - 1).
- Aux alentours du seuil, une phase de transition
- Au-dessus du seuil, le neurone est actif (souvent dans ce cas, sa sortie vaut 1) .

Dans sa première version, le neurone formel était implémenté avec une fonction à seuil, mais de nombreuses versions existent. Ainsi, le neurone de McCulloch et Pitts a été généralisé de différentes manières, en choisissant d'autres fonctions d'activations, comme les fonctions énumérées dans le tableau suivant. Les trois fonctions les plus utilisées sont les fonctions « seuil » « linéaire » , « sigmoïdes ». [13]

2.5 Architecture des Réseaux de neurones

Après avoir vu que l'on peut classer les réseaux de neurones en fonction des modèles d'apprentissage, nous allons maintenant voir la classification par architecture, qui se décompose également en deux grandes catégories : les réseaux Feedforward et les réseaux Feedback.

2.5.1 Les réseaux Feed-forward

Ce sont des réseaux dans lesquels les informations se propagent successivement de couche en couche sans retour en arrière. On y trouve :
Le perceptron monocouche

En termes d'architecture, c'est le réseau le plus simple, et en termes de mise en œuvre, il se compose d'une couche d'entrée et d'une couche de sortie. Le réseau est capable de résoudre des

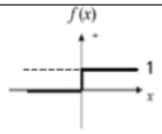
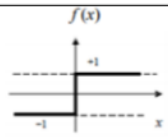
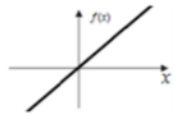
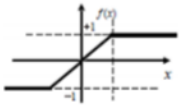
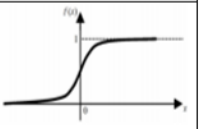
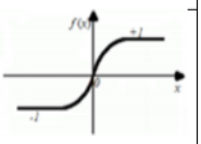
Catégories	Type	Equation	Allure
Seuil	Heaviside	$f(x) = \begin{cases} 0 \; si \; x < 0 \\ 1 \; si \; x \geq 0 \end{cases}$	
	Signe	$f(x) = \begin{cases} -1 \; si \; x \leq 0 \\ 1 \; si \; > 0 \end{cases}$	
Linéaire	Identité	$f(x) = \{x$	
	Saturé symétrique	$f(x) = \begin{cases} -1 \; si \; x \leq -1 \\ 1 \; si \; x \geq 1 \\ x \; sinon \end{cases}$	
Non linéaire	Sigmoïde	$f(x) = \frac{1}{1+e^{-x}}$	
	Tangente hyperbolique	$f(x) = \frac{2}{1+e^{-x}} - 1$	

FIGURE 2.4 – Quelque fonction d'activation [16]

problèmes linéairement séparables (ex : fonction logistique « OU » ou « ET »); notez que tous les neurones de la couche d'entrée sont reliés à tous les neurones de sortie, comme le montre la figure suivante :

- E1 :Entrée 1 de la couche d'entrée.
- Ei :Entrée i de la couche d'entrée.
- En :Entrée n de la couche d'entrée.
- S1 :Première sortie du réseau.
- Sk : Sortie k du réseau (k étant le nombre de sortie).

Le perceptron multicouche (PMC)

Contient une couche cachée entre les couches d'entrée et de sortie, le réseau résout un problème non-linéairement séparable (par exemple : fonction logistique 'XOR'); encore une fois, chaque neurone de la couche cachée est connecté à toutes les couches avant et après lui Neurones. La figure suivante représente un PMC :

- E1 : Entrée 1 de la couche d'entrée.
- Ei : Entrée i de la couche d'entrée.
- En : Dernière entrée de la couche d'entrée.
- C1j : Premier neurone de la couche cachée j.
- Cij :Neurone i de la couche cachée j.

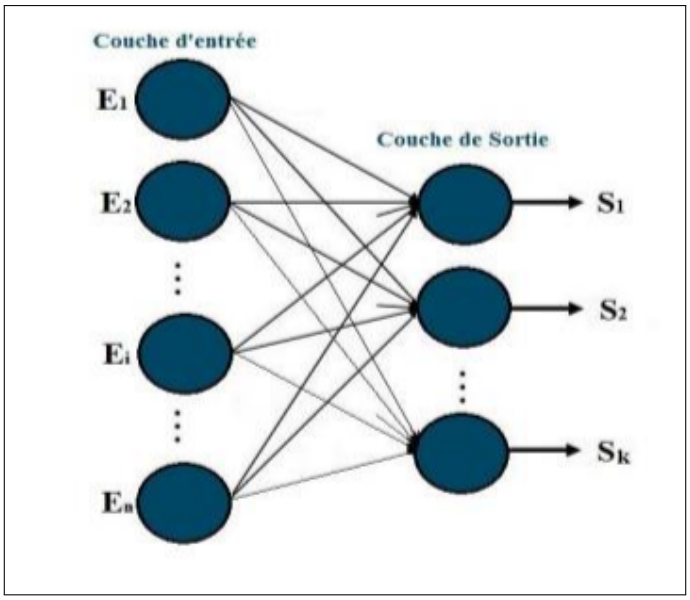


FIGURE 2.5 – Perceptron monocouche

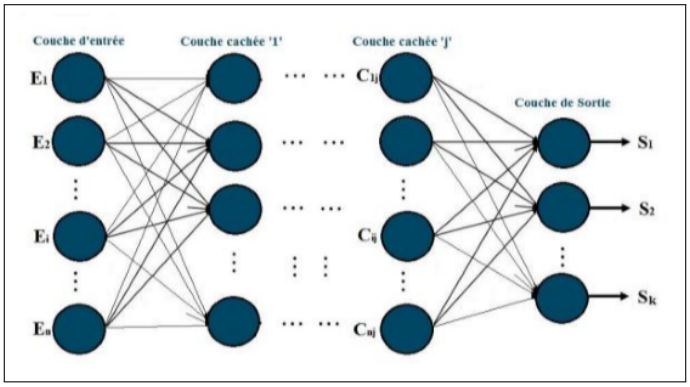


FIGURE 2.6 – Perceptron multicouche

- :Cnj : Dernier neurone de la couche cachée j (m : nombre de neurones dans la couche cachée j).
- S1 : Première sortie du réseau.
- Sk : Sortie k du réseau (k étant le nombre de sortie).

Le réseau à fonctions radiales (RBF)
A la même architecture que PMC, la différence est la fonction d'activation des neurones, puisque RBF utilise une fonction gaussienne.

2.5.2 Les réseaux Feed-back

Également connu sous le nom de « réseau récurrent », est un réseau qui peut revenir en arrière. On y trouve :
Cartes auto-organisatrices de Kohonen
C'est un réseau à apprentissage non supervisée, il est sous forme d'une grille dont chaque nœud représente un neurone associé à un vecteur de poids ; la figure ci-dessous explique mieux son architecture :

Une carte auto-organisatrice est un procédé qui convertit un signal d'entrée complexe (Plusieurs variables par exemple) en une nouvelle variable catégorielle : c'est donc un procédé de classification (modélisation non-supervisée). Les SOM sont une généralisation de l'analyse en composantes principales. Elle fonctionne comme un réseau de neurones sans variable cible et avec

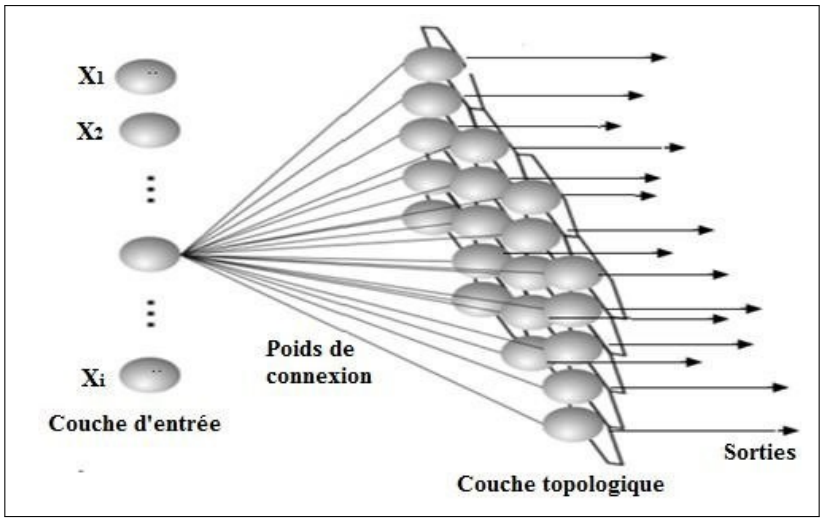


FIGURE 2.7 – Schéma d'une carte auto-organisatrice

plusieurs nœuds dans la couche de sortie. La carte structure les nœuds en sortie en classes de nœuds. Contrairement aux réseaux de neurones, SOM n'a pas de couches cachées. La couche de sortie contient plusieurs nœuds, représentés par un réseau rectangulaire

. Le nombre de nœuds dans la couche de sortie est défini arbitrairement par l'utilisateur. Il définit le nombre maximum de classes[2].

a) - Principes de fonctionnement

Fonction de score : Les valeurs des nœuds de la couche d'entrée (valeurs normalisées des variables prises en compte par le modèle) sont distribuées dans les nœuds de la couche de sortie après transformation en fonction des pondérations du réseau : on parle de « fonction de score ». Cette fonction est généralement une fonction de distance euclidienne. Le nœud de sortie qui a le meilleur résultat (on dit le «meilleur score ») est le « nœud gagnant » : il Reçoit l'individu en question. Le meilleur score c'est la plus petite distance entre les poids de connexion et les données d'entrée [2].

b) -Principe de fonctionnement : liaison de voisinage des nœuds de la couche de sorties

Comme tous les réseaux de neurones, les nœuds d'une même couche, notamment les nœuds de la couche de sortie, ne sont pas liés les uns aux autres. Cependant, dans le cas de données similaires, les poids des nœuds voisins du nœud gagnant sont ajustés pour favoriser leur victoire. C'est ce qu'on appelle la coopération et l'adaptation des nœuds de la couche de sortie. S'adapter, c'est apprendre. Comme tous les réseaux de neurones, les nœuds de la même couche, en particulier les nœuds de la couche de sortie, n'ont aucun lien entre eux
[16]

Réseaux de Hopfield

Ce sont des réseaux entièrement connectés. Dans ce type de réseau, chaque neurone est connecté à chaque autre neurone et il n'y a aucune différenciation entre les neurones d'entrée et de sortie. Ils sont capables de trouver un objet stocké en fonction de représentations partielles ou bruitées. L'application principale des réseaux de Hopfield est l'entrepôt de connaissances mais aussi la résolution de problèmes d'optimisation. Le mode d'apprentissage utilisé est là aussi le mode non-supervisé.
[16]

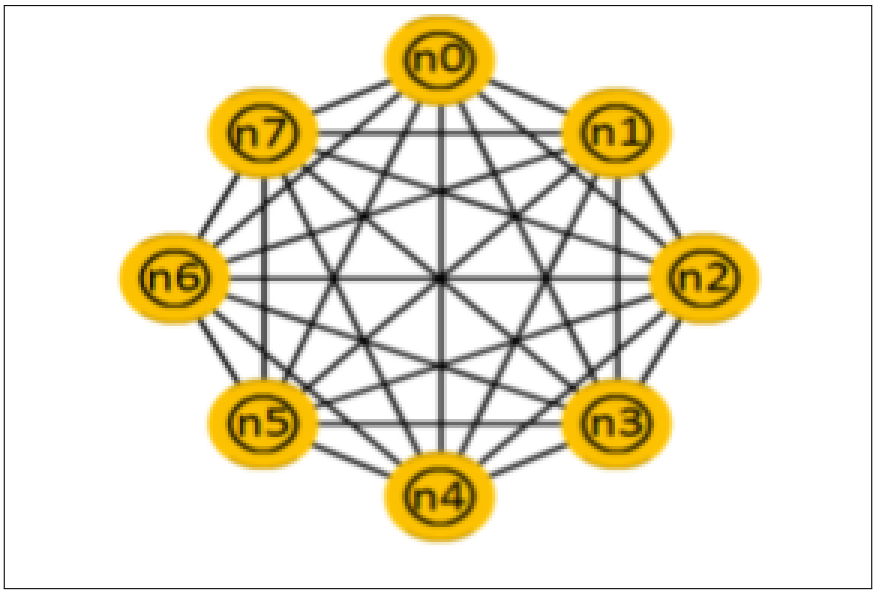


FIGURE 2.8 – Architecture d'un Réseau de Hopfield

2.6 les différents types de réseaux de neurones

2.6.1 Réseau de neurones récurrents(RNN)

Les humains ne commencent pas leur réflexion à partir de zéro chaque seconde. En lisant cet essai, vous comprenez chaque mot en fonction de votre compréhension des mots précédents. Vous ne jetez pas tout et recommencez à penser à partir de zéro. Vos pensées ont de la persévérance. Les réseaux de neurones traditionnels ne peuvent pas le faire, et cela semble être une lacune majeure. Par exemple, imaginez que vous souhaitiez classer le type d'événement qui se produit à chaque étape d'un film. On ne sait pas comment un réseau de neurones traditionnel pourrait utiliser son raisonnement sur des événements antérieurs dans le film pour les informer plus tard [14]

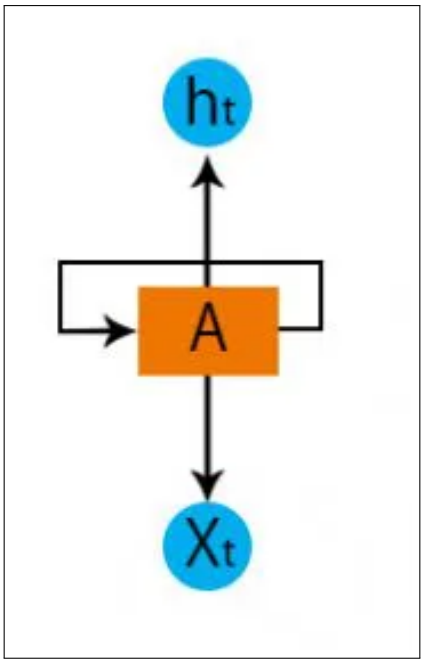


FIGURE 2.9 – Les réseaux neuronaux récurrents ont des boucles

Les réseaux de neurones récurrents résolvent ce problème. Ce sont des réseaux avec des boucles qui permettent à l'information de persister. Dans la figure cidessus, un segment de réseau neuronal ; "A" regarde une entrée Xt et fournit une valeur ht. Une boucle permet de passer des informations d'une étape du réseau à l'autre. Ces boucles rendent les réseaux neuronaux récurrents un peu mystérieux. Cependant, si vous pensez un peu plus, il s'avère qu'ils ne sont pas tous différents d'un réseau de neurones normal. Un réseau de neurones récurrent peut être considéré comme des copies multiples du même réseau, chacune transmettant un message à un successeur comme le montre la figure ci- dessus. Considérez ce qui se passe si nous déroulons la boucle : [5]

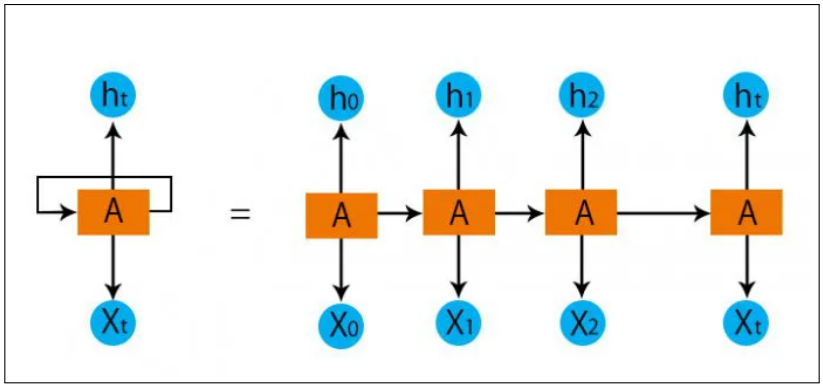


FIGURE 2.10 – Un réseau neuronal récurrent déroulé

Cette nature en chaîne révèle que les réseaux neuronaux récurrents sont intimement liés aux séquences et aux listes. Ils sont l'architecture naturelle du réseau de neurones à utiliser pour de telles données. Au cours des dernières années, il y a eu un succès incroyable en appliquant les RNN à une variété de problèmes : la reconnaissance de la parole, la modélisation du langage, la traduction, la reconnaissance de caractères, le sous-titrage des images ...etc [7].

2.6.2 Réseau de neurone convolutifs(CNN) :

En apprentissage automatique, un réseau de neurone convolutifs (ou réseau de neurones à convolution, ou CNN ou ConvNet) [36] est un type de réseau de neurones artificiels acycliques dans lequel le motif de connexion entre les neurones est inspiré par le cortex visuel des animaux.

Les neurones de cette région du cerveau sont arrangés de sorte à ce qu'ils correspondent à des régions qui se chevauchent lors du pavage du champ visuel. Leur fonctionnement est inspiré par les processus biologiques, ils consistent en un empilage multicouche de perceptrons, dont le but est de prétraiter de petites quantités d'informations. Les réseaux neuronaux convolutifs ont de larges applications dans la reconnaissance de caractère et d'image, les systèmes de recommandation et le traitement du langage naturel [15].

Il y a plusieurs couches différentes dans CNN comme le montre la FIGURE 1.10 :

- Couche d'entrée (Input layer).
- Couche de convolution (Convo layer : Convolution + ReLU).
- Couche de Pooling.
- Couche entièrement connectée (Couche Fully connected).
- Couche Softmax/logistique.
- Couche de sortie (Output layer).

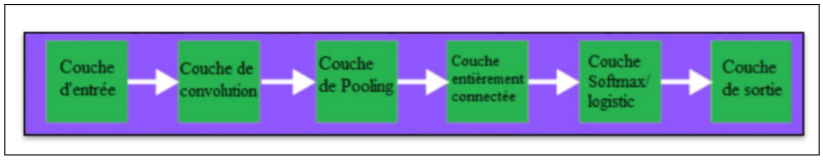


FIGURE 2.11 – Les couches de CNN

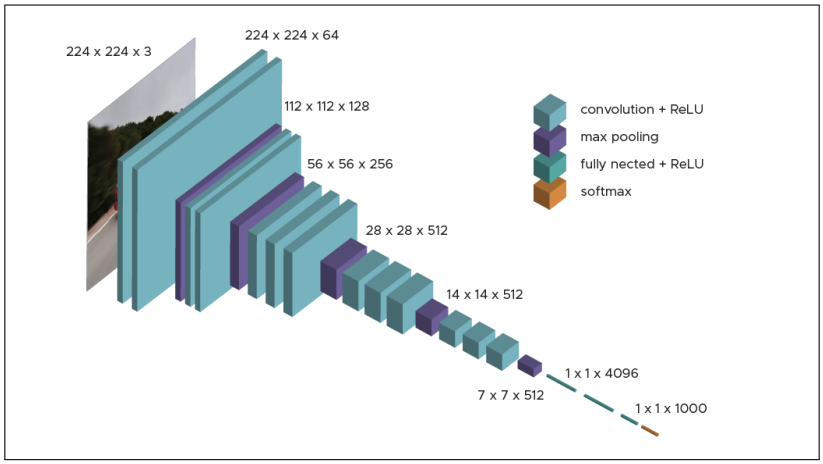


FIGURE 2.12 – Exemple d'architecture CNN

Couche d'entrée du CNN :

La couche d'entrée dans CNN doit contenir des données décrivant l'image. Les données d'image sont représentées par une matrice tridimensionnelle qui en général doit être remodelée en une seule colonne (représentation vectorielle) [18]

Couche de convolution

La couche de convolution est la composante clé des CNN, elle est présente au moins dans la première couche. Le but de cette couche est de trouver un ensemble de caractéristique dans les images reçues en entrée. C'est là qu'intervient l'opération de convolution : le principe est de faire "glisser" une fenêtre représentant les caractéristiques sur l'image. Le produit de convolution entre les caractéristiques et chaque portion de l'image balayée est alors calculé. Dans ce cas, la caractéristique est considérée comme un filtre. Les filtres correspondent exactement aux caractéristiques que l'on souhaite retrouver dans les images. En sortie, on obtient une carte d'activation ou caractéristique map, indiquant l'emplacement des caractéristiques dans 47 l'image. Dans un CNN, les

caractéristiques ne sont pas prédéfinies selon un formalisme particulier comme pour les méthodes traditionnelles (Viola Jones ou HOG par exemple), elles sont apprises lors de la phase d'entrainement du réseau. Les CNN sont donc capables de déterminer tout seul les éléments discriminants d'une image.
Le calcul de la dimension de la sortie associée à une couche de convolution nécessite la connaissance de trois hyper paramètres

- Le pas : dans le contexte d'une opération de convolution, le pas ou « stride » désigne le pas de déplacement du noyau après chaque opération à travers l'image.
- La profondeur : c'est le nombre de noyaux de convolution ou filtre.
- La marge à zéro ou zéro padding : il est commun d'utiliser une épaisseur de marge de zéros autour de l'entrée, le zero-padding est une technique consistant à ajouter P zeros à chaque côté des frontières de l'entrée[18].

Soit I le côté associé à l'entrant, F la dimension du noyau, P l'épaisseur de la marge à zéro et s l'amplitude du pas utilisée ; la dimension O de la sortie est alors donnée par :

$$O = \frac{1 - F + 2P}{S} + 1$$

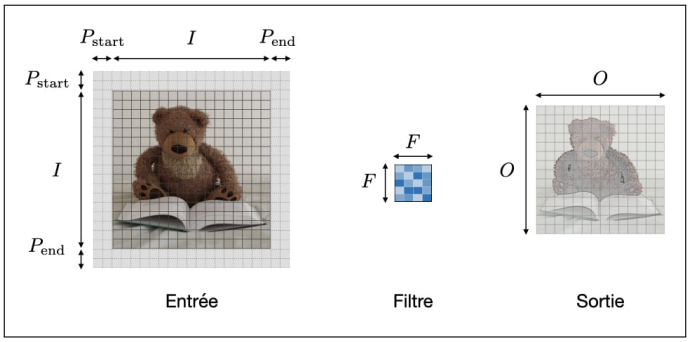


FIGURE 2.13 – Illustration de la taille d'une sortie de couche de convolution

$$O = \frac{1 - F + Pstart + Pend}{S} + 1$$

Couche de mise en commun (Pooling)

Après chaque couche convolutives, il peut y avoir une couche de Pooling. L'idée derrière l'opération Pooling est de retirer une certaine information dans un même voisinage pour l'entrant. L'opération est effectuée sur chacune des épaisseurs liées à la profondeur de l'entrant. Généralement, on utilise le max ou l'average Pooling, où les valeurs maximales et moyennes sont prises respectivement [5]

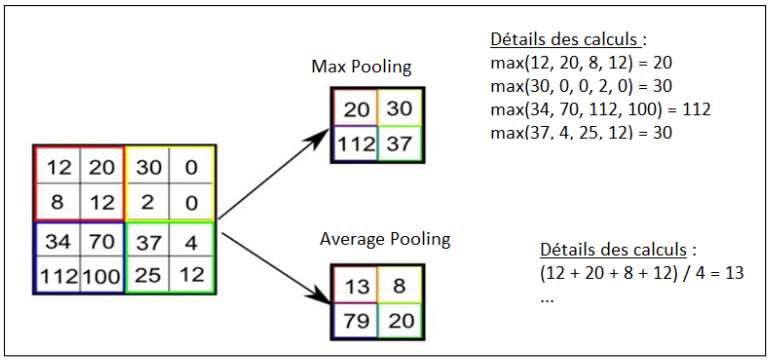


FIGURE 2.14 – exemple de calcule du max Pooling et average Pooling

Dans la figure II.15, on est en présence de cellules de taille 2x2. Pour l'average Pooling : chaque case correspond à la moyenne du carré d'entrée de la même couleur, exemple pour le cas de la case rouge, on a : $\frac{12+20+8+12}{4} = 13$
Pour le max Pooling : chaque case correspond à la valeur maximum du carré d'entrée de la même couleur, exemple pour la case bleu : :max(34,70,112,100) = 112

Couche entièrement connecté (Fully Connected)

La couche entièrement connectée ou Fully Connected layer en anglais constitue toujours la dernière couche d'un réseau de neurones. Cette couche n'est donc pas caractéristique d'un CNN. Elle s'applique sur une entrée préalablement aplatie où chaque entrée est connectée à tous les neurones et produit un nouveau vecteur en sortie. La dernière couche classifie l'image d'entrée en plusieurs classes et renvoie un vecteur dont la taille est égale au nombre de classe du problème. Chaque élément du vecteur indique la probabilité pour l'image en entrée d'appartenir à une certaine classe.[18]

Pour calculer les probabilités, la couche entièrement connectée multiplie donc chaque élément en entrée par un poids, fait la somme, puis applique une fonction d'activation. L'apprentissage des valeurs des poids se fait lors de phase d'entraînement, par rétropropagation du gradient. La couche entièrement connectée cherche les liens entre la position des features dans l'image et une classe. En effet, le tableau en entrée correspond à une carte d'activation pour une feature donnée : les valeurs élevées indiquent la localisation de cette feature dans l'image. Si la localisation d'une feature à un certain endroit de l'image est caractéristique d'une certaine classe, alors on accorde un poids important à la valeur correspondante dans le tableau [18].

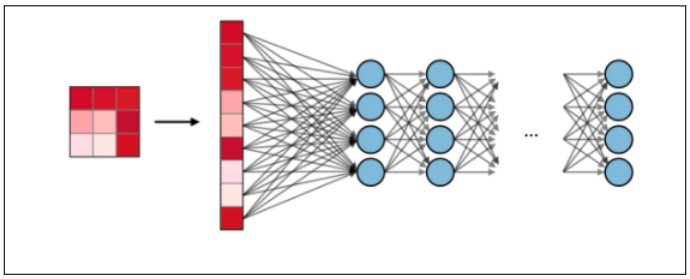


FIGURE 2.15 – Mis en évidence dune couche entièrement connectée

Couche Logistique ou Softmax :

Softmax ou couche logistique est la dernière couche de CNN. Elle réside à la fin de la couche FC. La logistique est utilisée pour la classification binaire et Softmax est pour la multi-classification.

Couche de sortie (output layer)

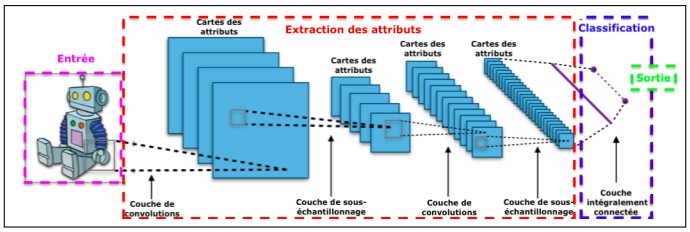


FIGURE 2.16 – Exemple montrant l'étiquette codée de la couche de sortie CNN

2.7 conclution

Nous avons présenté dans ce chapitre Introduction générale aux réseaux de neurones , sa définition et son histoire,Puis nous nous sommes abordés l'Architecture des Réseaux de neurones .Enfin, nous avons introduit les deux types de réseaux de neurones les plus populaires qui sont les Réseaux de neurones récurrents(RNN) et les Réseaux de neurones convolutifs(CNN) . Dans le chapitre suivant, on vas expliquer la reconnaissance de l'écriture manuscrite.

Chapitre 3

LA RECONNAISSANCE D'ÉCRITURE MANUSCRITE

3.1 Introduction

L'intérêt de la lecture automatisée de l'écriture manuscrite est indéniable pour l'accomplissement des tâches fastidieuses rencontrées dans certains domaines : tri postal, lecture de chèques bancaires, lecture de bordereaux, bons de commande, déclarations. A priori, on peut imaginer la reconnaissance des manuscrits comme une manifestation de la reconnaissance du texte imprimé ou dactylographié, et le principal obstacle semble avoir été surmonté. Mais la transition de la reconnaissance d'imprimés à la reconnaissance de manuscrits n'est pas aussi simple que certains le pensaient initialement. Sa mise en œuvre est beaucoup plus compliquée. Les difficultés rencontrées proviennent principalement de la nature de l'écriture manuscrite, des caractères cursifs et de l'extrême variabilité.

Cependant, malgré ces difficultés, la concrétisation du problème à résoudre et la concrétisation de la tâche à accomplir permettent aujourd'hui d'aboutir à des réalisations de plus en plus tangibles.[34]

3.2 Historique

Les premiers pas dans le domaine de la reconnaissance de l'écriture manuscrite sont anciens . Ils utilisent des systèmes mécaniques simples mais inefficaces et lents. Les avancées dans ce domaine ont été longues et complexes en raison de la nature de l'écriture : la variabilité inter-texte et intra-caractère est très élevée, comme le montre la figure 1.1. De l'écriture. Au fil du temps, cette variabilité a conduit de plus en plus de personnes à abandonner des techniques simples telles que les masques faciaux au profit de systèmes dits "intelligents". Ces systèmes s'adaptent au style de l'auteur, utilisent des contextes locaux et globaux pour prendre des décisions et utilisent des dictionnaires et des modèles de langage.

Ce bref historique combine des technologies de reconnaissance de l'écriture manuscrite et de la frappe identiques et souvent transférables d'un style à l'autre. La figure 1.2 montre un classement chronologique des différentes périodes, y compris les publications importantes et les différentes percées technologiques.[29]

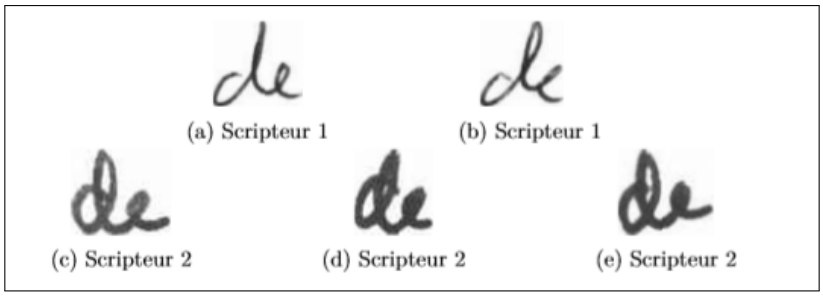


FIGURE 3.1 – Variabilité de l'écriture manuscrite

Sur la base de ce schéma, nous présentons l'histoire en quatre phases principales : les pionniers, les modèles statistiques, les modèles de Markov hybrides et l'essor des réseaux de neurones.

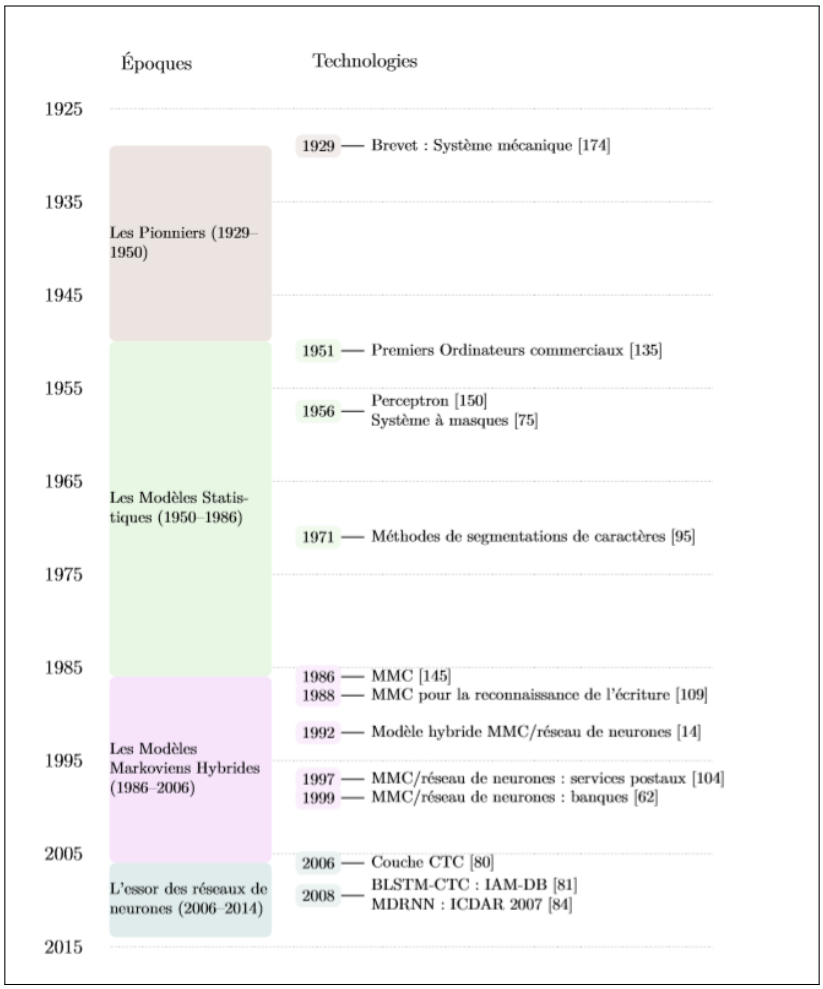


FIGURE 3.2 – Frise simplifiée de la reconnaissance de l'écriture

3.3 Différents aspects de reconnaissance

La première étape d'un système de reconnaissance consiste à convertir l'écriture manuscrite en une quantité numérique adaptée au système de traitement, avec le moins de dégradation possible. Tout dépend du type de dispositif d'acquisition et des données à traiter.

Nous allons présenter quelques aspects des systèmes de reconnaissance manuscrit tels que la différence entre la reconnaissance et la production et les différents types de systèmes de reconnaissance de l'écritures manuscrits selon le mode d'acquisition (Les systèmes de reconnaissance en ligne, Les systèmes de reconnaissance hors-ligne).[22]

3.3.1 Production et reconnaissance

La reconnaissance consiste à convertir une image d'un document en une représentation sous forme de fichier informatique. On peut aussi dire que la reconnaissance est le processus inverse de la production. La figure 1.3 illustre ces deux processus. Capturez des documents avec un scanner ou un appareil photo pour obtenir des images numériques. Cette image contient généralement du bruit. Cette image est ensuite prétraitée pour fournir une image propre avec une représentation plus nette. [40]

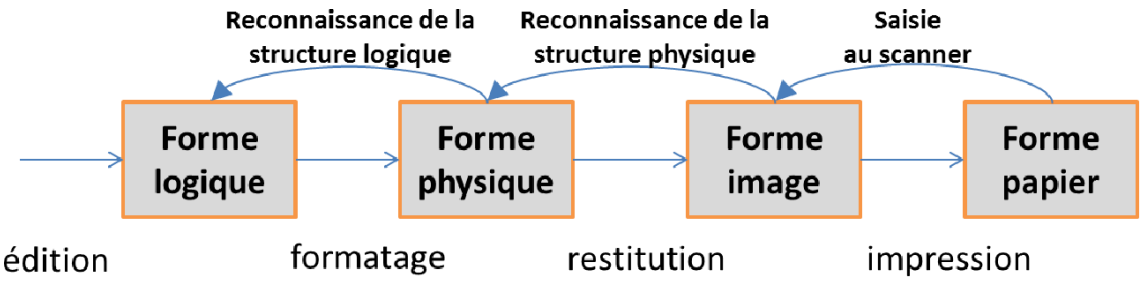


Figure 9 :Processus de production et de reconnaissance de documents

FIGURE 3.3 – Processus de production et de reconnaissance de documents

3.3.2 Reconnaissance en-ligne / hors-ligne

Dans la reconnaissance de l'écriture manuscrite en ligne, la collecte de l'écriture manuscrite est effectuée pendant le processus d'écriture, tandis que dans la reconnaissance de l'écriture manuscrite hors ligne, la collecte de l'écriture manuscrite est effectuée à partir de documents papier numérisés par des scanners ou des caméras.[12]

3.3.3 Reconnaissance en-ligne

Dans le cas en ligne, il s'agit de reconnaître l'écriture telle qu'elle est écrite. Le texte est saisi à l'aide d'un stylo et d'une tablette numérique, et les informations collectées consistent en une séquence ordonnée de points (définis par leurs coordonnées) échantillonnés à un taux fixe. L'écriture prend la forme d'une paire de signaux horaires numérisés.
Les avantages majeurs de reconnaissance d'écriture manuscrite en-ligne

- Moins de bruit car l'utilisateur écrit sur une table spéciale
- la possibilité de correction et de modification de l'écriture de manière interactive vu la réponse en continu du système [6][42]

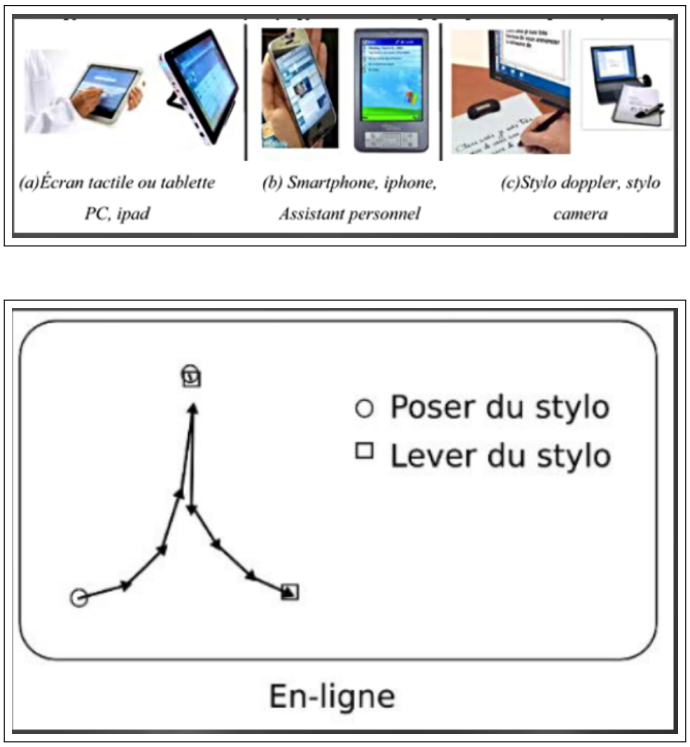


FIGURE 3.4 – Reconnaissance en-ligne

3.3.4 Reconnaissance hors-ligne

Dans les situations hors ligne, il s'agit d'un problème de reconnaissance de texte manuscrit à partir de documents précédemment écrits. Les images du texte écrit sont numérisées à l'aide de scanners et les informations collectées prennent la forme d'images discrètes constituées d'un ensemble de pixels. Le texte prend l'apparence d'un signal spatial bidimensionnel numérisé.

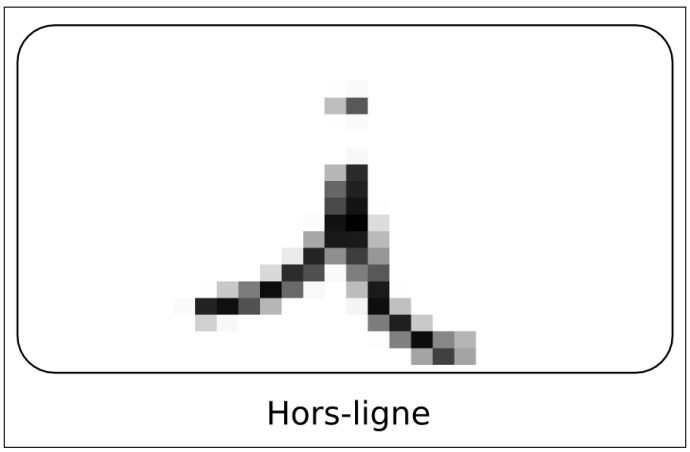


FIGURE 3.5 – Reconnaissance hors-ligne

3.4 Type de mode hors ligne

3.4.1 Reconnaissance de texte ou analyse de documents

Dans le premier cas, il s'agit d'un problème d'identification d'un texte structuré, qui se limite à quelques lignes ou mots. La recherche consiste en une simple identification des mots dans une ligne, puis en décomposant chaque mot en caractères. Dans le second cas (analyse documentaire), il s'agit de données bien structurées dont la lecture nécessite une compréhension de la typographie et de la mise en page du document. L'approche ici n'est plus un simple prétraitement, mais une approche experte de l'analyse de documents, y compris la localisation des régions, la séparation des régions graphiques et photo, l'étiquetage sémantique des régions et des modèles de texte, la détermination de l'ordre de lecture et la structure du document[42].

3.4.2 Reconnaissance de l'imprimé ou du manuscrit

Les approches diffèrent selon qu'il s'agisse de reconnaissance de caractères imprimés ou manuscrits. Les caractères imprimés sont dans le cas général alignés horizontalement et séparés verticalement, ce qui simplifie la phase de lecture. La forme des caractères est définie par un style calligraphique (fonte) qui constitue un modèle pour l'identification. Dans le cas du manuscrit, les caractères sont souvent ligaturés et leur graphisme est inégalement proportionné provenant de la variabilité intra et inter scripteurs. Cela nécessite généralement l'emploi de techniques de délimitation spécifiques et souvent des connaissances contextuelles pour guider la lecture [42].

3.5 La différence entre les Reconnaissances en ligne et hors ligne

La reconnaissance hors ligne ne peut pas se baser a priori sur l'information temporelle des tracés manquants, mais elle peut prendre en compte l'épaisseur des tracés (plein et ascendant). La reconnaissance en ligne peut avoir des informations temporelles (vitesse, accélération, levée du stylet, retour en arrière, barres en T, points d'inflexion), mais sans signal de pression de la pointe du stylet au support, il n'y a aucune information sur l'épaisseur du dessin. On a aussi le tableau 1 présente quelque différence entre les deux modes [21].

Caractère en ligne	Caractère hors ligne
Utilisation de stylos numériques	Utilisation de papier
Disponibilité de coup de stylo	Non disponible de stylo
Taux de reconnaissance élevé	Taux de reconnaissance bas
Grande précision	Faible précision

FIGURE 3.6 – La différence entre les Reconnaissances en ligne et hors ligne [4]

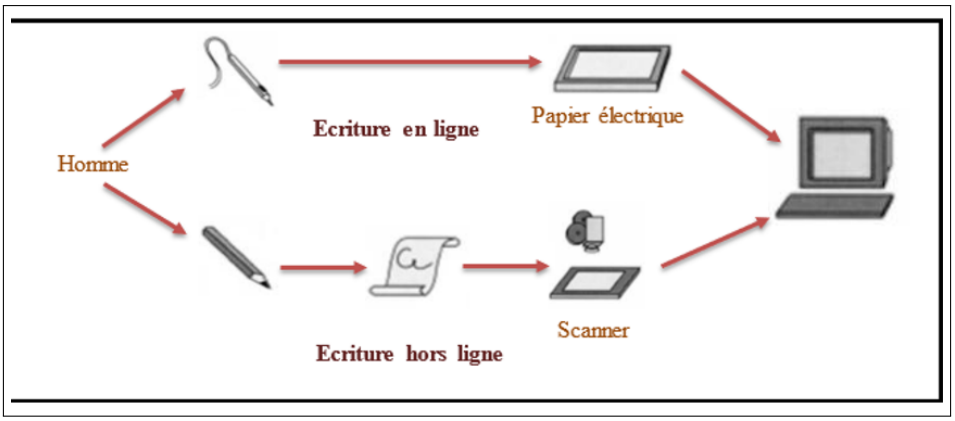


FIGURE 3.7 – Différence entre les deux types de reconnaissance (online , offline).

3.6 Organisation Générale D'un Système De Reconnaissance De L'écriture

À première vue, on peut considérer que la reconnaissance d'écriture manuscrite en ligne et hors ligne est généralement un problème de reconnaissance de formes. Elle s'articule autour des principes usuels : acquisition d'informations par des organes sensoriels (capteurs), extraction de représentations codées de la forme, classement d'objets à partir d'un ensemble de classes possibles ou de représentations codées en classes.[28]

Pour faciliter la compréhension, nous considérons que l'organisation d'un système classique de reconnaissance de l'écriture manuscrite se déroule en cinq opérations séquentielles (Figure 1.7) :

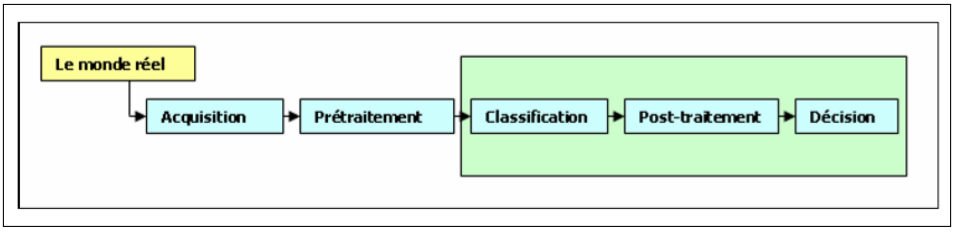


FIGURE 3.8 – Organisation d'un système de reconnaissance d'écriture manuscrite.

3.6.1 Le monde physique (réel)

C'est un espace de simulation de dimension infinie appelé espace de forme. Les objets de cet espace sont décrits de différentes manières, avec plusieurs propriétés. Les lois conduisant au monde discret nécessitent des choix et donc des simplifications.[38]

3.6.2 Acquisition de l'écriture

L'écriture est numérisée et elle est transformée en une image. C'est l'entrée de ce système. Cette étape est assez simple mais très importante car elle influence sérieusement les étapes suivantes.

Il y a deux paramètres importants :

- **Résolution :** la résolution normale est 300 dpi. Pourtant, quand la taille de l'écriture est petite, il faut augmenter la résolution.

- **Niveau d'éclairage :** si on ajuste le scanneur pour que l'image soit plus claire, le bruit est réduit mais des traits minces disparaissent aussi.

1. Acquisition de l'écriture hors ligne

Dans le cas de l'écriture hors ligne, les systèmes d'acquisition les plus courants sont essentiellement des scanners ou des caméras linéaires (matrices CCD). La résolution de l'image numérisée affecte les étapes ultérieures du système de lecture automatisé. Il est généralement admis que la résolution optimale d'une image dépend de l'épaisseur de la ligne d'écriture.

2. Acquisition de l'écriture en ligne

Dans le cas de l'acquisition « en ligne », les dispositifs d'acquisition les plus répandus sont des tablettes à numériser ou des « papiers électroniques ». L'échantillonnage du tracé délivre une série de coordonnées décrivant la trajectoire du stylet au cours du temps. Les tablettes les plus perfectionnées permettent également d'avoir accès aux informations de pression et d'inclinaison du stylo.[34]

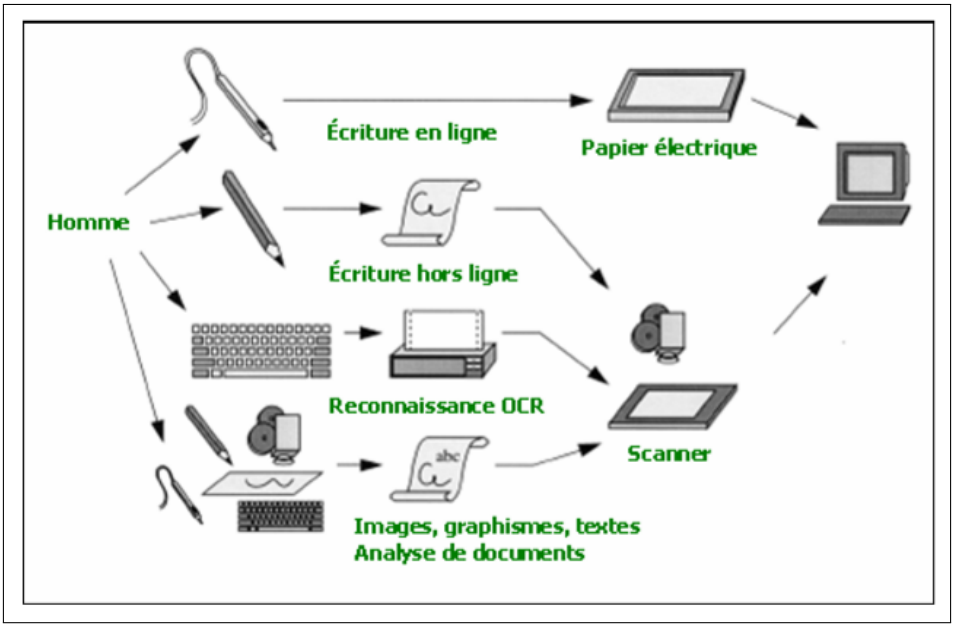


FIGURE 3.9 – Communication écriture Homme-machine.

3.6.3 Prétraitement

L'étape de prétraitement est l'une des premières tâches d'un système de reconnaissance après l'étape d'acquisition. Cette étape n'est pas spécifique à la reconnaissance du manuscrit mais fait partie de tout système de reconnaissance de forme visant à améliorer la qualité de l'information pour les étapes de traitements qui vont suivre. Dans le cadre du manuscrit, elle regroupe l'ensemble des processus visant au bon conditionnement du message écrit et qui sont indispensables à son identification. On peut catégoriser les pré-traitements en fonction du mode d'acquisition de l'écriture (en-ligne, hors-ligne). On peut subdiviser, dans les deux modes, les étapes de prétraitements en deux groupes d' une part pré-traitements qui concernent le rééchantillonnage et la quantification du signal provenant du capteur (scanner dans le cas horsligne, tablette à digitaliser dans le cas en-ligne), et d'autre part les pré-traitements qui concernent la localisation ainsi que la normalisation de l' information manuscrite afin d'éliminer dans une certaine mesure la variabilité due à la disposition spatiale de l'information (taille, inclinaison…). Si le premier groupe de pré-traitements reste propre à chaque mode d'acquisition, le second en revanche est commun aux deux.[34]

3.6.4 Reconnaissance [22]

La reconnaissance combine les deux tâches d'apprentissage et de prise de décision, qui jouent des rôles très similaires dans les systèmes de reconnaissance de formes. En fait, à partir d'une même description de la forme dans les paramètres, ils tentent tous d'attribuer cette forme au modèle de référence. Le résultat de l'apprentissage est de réorganiser ou d'améliorer les modèles existants pour prendre en compte les apports de nouvelles formes, ou de créer de nouveaux modèles qui représentent les formes d'entrée. Le résultat de la décision est une "opinion" sur l'appartenance du formulaire au modèle appris. Nous discuterons des capacités de ces tâches plus tard en développant quelques méthodes classiques pour les exécuter.

1. Apprentissage

Dans le cas d'apprentissage il s'agit en fait de fournir au système un ensemble de formes qui sont déjà connues (on connaît la classe de chacune d'elles). C'est cet ensemble d'apprentissage qui va permettre de « régler » le système de reconnaissance de façon à ce qu'il soit capable de reconnaître ultérieurement des formes de classes inconnues.
Il existe deux types d'apprentissage, supervisé et non supervisé.[34]

• **Apprentissage supervisé**

L'apprentissage est dit supervisé si les différentes familles des formes sont connues a priori et si la tâche d'apprentissage est guidée par un superviseur ou professeur, c'est-à-dire le concepteur, indique, pour chaque forme échantillon rentrée, le nom de la famille qui la contient. La tâche d'apprentissage tente de conserver ses liens de parenté en répartissant les familles dans des classes séparées entre elles.[34]

• **Apprentissage non supervisé**

On l'appelle aussi, suivant l'approche utilisée, classification automatique, inférence ou encore apprentissage sans professeur. Il s' agit, à partir d' échantillon de référence et de règles de regroupement ou de modélisation, de construire automatiquement les classes ou les modèles sans intervention de l'opérateur. Ce mode d'apprentissage nécessite un nombre élevé d'échantillons et des règles de construction précise et non contradictoires pour bien assurer la formation des classes. Il évite l'assistance d'un opérateur mais n'assure pas toujours une classification correspondant à la réalité (celle de l'utilisateur).[34]

2. La décision

La décision est la dernière étape de la reconnaissance. A partir de la description des paramètres, il recherche les modèles "les plus proches" parmi les modèles appris existants. La notion de proximité varie selon la nature de la représentation et le type de méthode. Si la réponse est unique (un seul modèle correspond à la description de la forme), la décision peut mener au succès. Si la réponse est multiple, cela peut prêter à confusion (plusieurs modèles correspondent à des descriptions, ce qui indique une sorte d'ambiguïté, intentionnelle ou non, dans le processus d'apprentissage). Enfin, si aucun modèle ne correspond à sa description, la décision peut entraîner le rejet du formulaire. Dans les deux premiers cas, la décision peut être accompagnée d'une mesure de vraisemblance, aussi appelée score ou taux de reconnaissance.

3.7 Méthodes de reconnaissances

Deux approches s'opposent en reconnaissance des mots : globale et analytique :

3.7.1 Méthodes globales

Les méthodes de reconnaissance globale de mots sont a priori attractives car elles ne souffrent pas des difficultés liées à l'ambiguïté induite par la segmentation. Cependant, ils ne peuvent être conçus qu'avec un vocabulaire limité, ou avec un vocabulaire plus large mais dynamiquement limité. Ils sont basés sur les caractéristiques globales des mots.

3.7.2 Méthodes analytiques

Contrairement aux méthodes de reconnaissance globales, les méthodes analytiques cherchent à identifier les graphèmes issus de la segmentation (fragments de lettres, des lettres ou des re-

groupements de lettres) pour reconstituer les mots. Elles présentent l' intérêt de pouvoir se généraliser à la reconnaissance d'un vocabulaire étendu ou limité dynamiquement, à partir d'une phase d'apprentissage des classes de graphèmes.

3.8 Différentes approches

Après avoir extrait les caractéristiques de la forme à reconnaître, le système OCR procède automatiquement à la classification. La classification est le stade principal de prise de décision. Celle-ci concerne l'identification de la forme, comme étant le membre d'une classe. Elle se fait en comparant les caractéristiques de la forme à celles d'un modèle de référence. Souvent, le résultat de cette étape est un ensemble de solutions au lieu de générer une solution unique. La classification cherche, à partir des caractéristiques de la forme, parmi les modèles d'apprentissages en présence, ceux qui sont les plus proches. La nature de la proximité a un sens différent en fonction de la nature de représentation et de type de méthode. La décision peut conduire à un succès si la réponse est unique, Elle peut conduire à une confusion si la réponse est multiple. Enfin, la décision peut conduire à un rejet de la forme si aucun modèle ne correspond à sa description.

Les principales méthodes utilisées sont les méthodes stochastiques, neuro-mimétiques, markoviennes, des méthodes issues de l'intelligence artificielle, de la théorie des sous-ensembles flous ou des méthodes mixtes combinant ces différentes approches.

3.8.1 Méthodes stochastiques

Dans ces méthodes de reconnaissance des formes ou il va s' agir de regrouper les formes inconnues dans des classes, on va considérer que chacune des classes est régie par un phénomène stochastique (par exemple on dira qu'on s' intéresse à une classe de comportement gaussien : chaque forme de la classe a une certaine probabilité de se produire et la loi de probabilité correspondante est gaussienne).

Décision bayésienne
La théorie de la décision Bayésienne est la théorie centrale des méthodes stochastiques où les problèmes de décision sont traités en termes de probabilités. Le point central de cette théorie est la règle de Bayes qui permet en fait de choisir l'hypothèse ayant la probabilité la plus élevée. La règle de Bayes permet de calculer ce qu'on appelle les probabilités a posteriori des classes, c'est à dire$P(w_i/x)$ à partir des probabilités a priori des classes$P(w_i)$ et des fonctions de densité des probabilités$p(x/w_i)$.

$$P(w_i/x) = \frac{p(x/w_i) * p(w_i)}{p(x)} \quad \forall i = 1,2 \quad avec \quad p(x) = \sum_{J=1}^{2} p(x/w_j) * p(w_j)$$

L'interprétation de $P(w_i/x)$ est la suivante : ayant obtenu une forme (décrite par une valeur x de) $P(w_i/x)$ donne la probabilité d'avoir alors la classe w_i

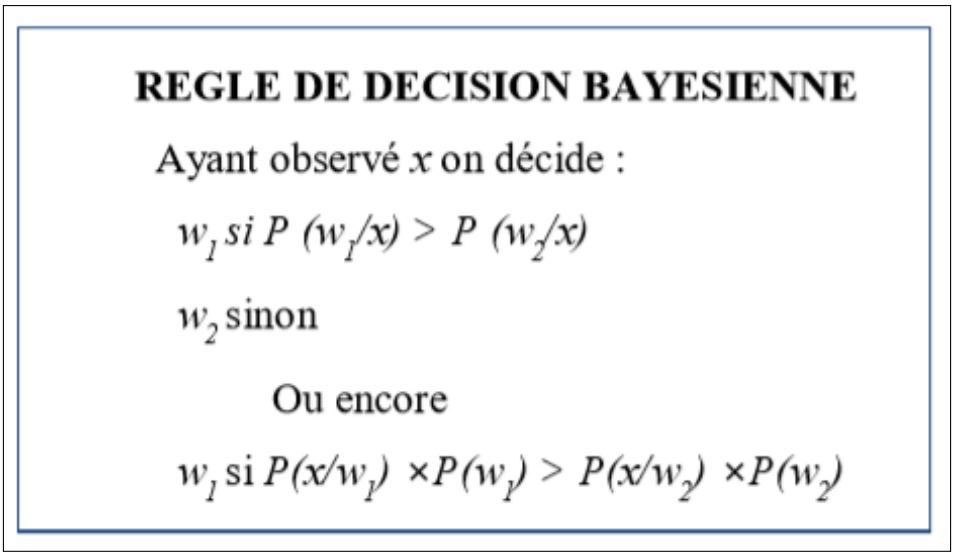
La règle de décision Bayésienne est alors pratiquement immédiate :

Par exemple, pour i=1,2, ayant remarqué une forme inconnue x (la forme est assimilée au paramètre qui la caractérise) on dira que la forme inconnue x est de la classe w_1 si :

$$P(w_1/x) > P(w_2/x)$$

Sinon elle est de la classe w2 Soit sous une forme simplifiée :

Dans cette règle on voit de façon particulièrement claire qu'il est nécessaire et suffisant de connaître pour toutes les classes cherchées la probabilité a priori et la loi de densité de probabilité de cette classe. [28]



3.8.2 Méthodes géométriques ou statiques

Elles consistent à extraire de la forme donnée un ensemble de mesures qui consistent les composants d'un vecteur d'un espace de représentation de dimension m. Ces mesures sont en nombre assez élevé pour la reconnaissance hors ligne ; elles comprennent principalement des informations topologiques et métriques : surfaces, régions, profils, concavités, boucles, intersections par des droites, rapports de longueurs et rapports de surfaces…

Par la reconnaissance en ligne, les mesures sont en nombre plus réduit ; elles consistent notamment en nombre de lettres dépassant le corps du mot (zone médiane dans laquelle s'inscrivent les minuscules sans extension) vers le haut ou le bas, en coefficients de transformées (ex : descripteurs de Fourier), en mesures sur des portions de traits (courbure, rapports, de longueurs, etc.). Le processus de classification s' effectue généralement par une partition de l'espace des représentations à l'aide d'une méthode de classement, notamment par séparation linéaire (ex : méthode des hyperplans).[34]

3.8.3 Méthodes neuro-mimétiques

Elles sont basées sur l ' utilisation de réseaux de neurones de tous types, perceptrons multicouches, carte de Kohonen, TDNN (Time Delay Neural Network), RBF (Radial Basis Function), neocognitron, etc. Les premières méthodes partaient directement des signaux ou des images normalisés, les méthodes plus récentes extraient d'abord un ensemble de caractéristiques qui servent ensuite à la reconnaissance d'écriture.

Les résultats obtenus par ces sont bons à condition de disposer de bases de données d'apprentissage de très grande taille. En revanche, les temps d'apprentissage peuvent être très longs. Comme l ' information est répartie, les causes d ' erreurs de ces systèmes peuvent difficilement être analysées.

3.8.4 Méthodes markoviennes

Les MMC ou modèles de Markov cachés (HMM : Hidden Markov Models) sont devenus une technique largement utilisée dans le domaine de la reconnaissance de l'écriture en ligne et hors ligne. Ils ont été créés pour modéliser la production par une source cachée de séquences de signaux variables au cours du temps. Ces derniers sont engendrés par un double processus aléatoire. Le premier processus peut se trouver, à un instant donné, dans un état appartenant à un ensemble de n états distincts. Le passage d'un état à un autre s'effectue selon une matrice de probabilités de transition. La séquence des états successifs n'est pas directement observable, d'où le

nom de caché. Chaque transition entre deux états successifs(ou chaque état) émet une observation Oi selon une loi de probabilité associée à chaque transition (où à chaque état).[34]

3.8.5 Classificateur euclidien

Il s'agit de l'un des plus simples classificateurs qui puissent être conçus. La classe dont le vecteur de caractéristiques moyen est le plus proche, au sens de la distance euclidienne, du vecteur de caractéristiques de l(objet à classifier est assignée à ce dernier. Les fonctions discriminantes utilisées sont donc de la forme suivante : [40]

$$\Phi_i(X) = -\frac{1}{2}(X - M_i)^T(X - M_i)$$

3.8.6 Le classificateur quadratique

Comme le nom l'indique, les frontières de décision fournies par ce modèle de classificateur sont quadratiques. Les fonctions discriminantes s'expriment par :

$$\Phi_i(X) = -\frac{1}{2}(X - M_i)^T(X - M_i)$$

Où $M_i = E(X/C_i)$ est le vecteur de caractéristiques moyen des éléments qui appartiennent à la classe C_i, $E\{.\}$ CE désignant l'opérateur d'opérateur d'espérance mathématique et $\{.\}^T$celui de transposition.

Le terme quadratique XX^t est indépendant de la classe de l'objet, et les fonctions discriminantes peuvent également s'écrire :

$$\Phi_i(X) = M_i^t X - \frac{1}{2}M_i^t M_i$$

Les frontières qui séparent les classes dans l'espace R^d sont ici linéaires. [34]

3.8.7 La méthode du plus proche voisin

L'algorithme KNN (K Nearest Neighbors) associe une forme inconnue à la classe. De son plus proche voisin en le comparant aux formes stockées dans une classe de références nommés prototypes. Il reste les K formes les plus proches de la forme à reconnaître suivant un critère de similarité. Une stratégie de décision permet d'affecter des valeurs de confiance à chacune des classes en compétition et d'attribuer, la plus vraisemblance (au sens de la métrique choisie) à la forme inconnue, le critère de similarité entre deux formes est basé sur la distance euclidienne.

3.8.8 Méthodes mixtes

Elles consistent à mettre en œuvre des systèmes à structure mixte comme, par exemple, un système combinant n règles floues et modèles de Markov ou encore carte kohonen et règles floues, etc.

Une autre approche consiste à combiner les résultats pondérés de plusieurs modules de classification qui sont de nature différente [28] et à prendre pour décision, la réponse fournie par le principe du "vote majoritaire" ou des techniques plus complexes comme les techniques de décision bayésienne ou la théorie de l'évidence de Dempster-schafer. Pour augmenter les performances, cette combinaison doit utiliser les points forts de chaque classifieur et écarter leur faiblesse

Post-traitements Lorsque la reconnaissance aboutit à la génération d'une suite de mots possibles éventuellement classés par ordre de vraisemblance, les post-traitements permettent d' améliorer les taux de reconnaissance par la prise en compte de connaissances pragmatiques et de connaissances linguistiques.[34]

- **Connaissances pragmatiques**
  Dans certaines applications comme celles des systèmes de lecture d'adresses, le nombre de chiffres du code postal est connu. D'autre part, la confrontation entre la reconnaissance de ce code postal et du bureau distributeur permet de lever certaines ambiguïtés.

- **Connaissances linguistiques**
  L'intégration, à un système de lecture, de connaissances linguistiques de différents niveaux (lexical, syntaxique ou sémantique) permet au système d' effectuer des vérifications et des corrections.

  L'utilisation d'un lexique ou d'un dictionnaire permet de valider a posteriori la reconnaissance effectuée. Mais pour éviter un temps de calcul prohibitif. L'utilisation d'un modèle du langage permet de moduler le taux de confiance des hypothèses de mots reconnus.

  La syntaxe confirme ou non, suivant les règles grammaticales prédéfinies, la séquence de mots proposés. Celle-ci est largement utilisée dans le cadre de la lecture des montants littéraux et numériques des chèques. Le contrôle sémantique est lié à l'aspect polysémique d'un mot ou d'une phrase. Il permet de réduire la liste des mots candidats, mais sa généralisation à de grands vocabulaires pose des difficultés. [34]

3.9 Mesure des performances

Après aborder les principes liés à la reconnaissance, il est bon de rappeler les principaux critères permettant de juger le degré d ' efficacité de ces systèmes et de mesurer leurs performances. Les évaluations sont généralement caractérisées comme dans le cas de l'imprimé, par les différents taux : [29]

- **le taux de reconnaissance** : pourcentage de mots bien reconnus ;
- **le taux de confusion** : pourcentage de mots pour lesquels le système fait une erreur ;
- **le taux de rejet** : pourcentage de mots pour lesquels le système refuse de se prononcer ;
- **le taux de confiance** : pourcentage de mots bien reconnus par rapport à la somme des mots bien reconnus et des mots mal reconnus.

Néanmoins, il ne faut pas considérer ces évaluations de manière absolue mais les replacer dans le contexte de l'application. Pour un type d'application donné, la comparaison des performances des différents systèmes de reconnaissance ne prend une signification que si ces systèmes sont testés sur des bases de données communes. Certains systèmes peuvent présenter des bonnes performances pour un type d'application et être très décevants pour un autre type. Tout dépend de la difficulté de l'application. Cette difficulté est déterminée par deux facteurs : l'ambiguïté des formes à reconnaître qui est difficilement mesurable, la dimension de l'espace des solutions ; la reconnaissance d'un vocabulaire réduit comprenant vingt mots n'est pas comparable à celle d'un vocabulaire de mille mots, encore moins à celle de 40 000 mots. D'autre part, la reconnaissance dépend de l'occurrence de chacun de ces mots, c'est pourquoi certains auteurs ont introduit pour l'écriture la notion de perplexité, déjà employée dans la reconnaissance de la parole. La perplexité est un indicateur basé sur l'entropie des solutions possibles. [29]

3.10 Conclusion

Dans ce chapitre nous avons présenté une étude générale sur la reconnaissance d'écriture manuscrite.

Nous abordons d'abord les aspects historiques et principaux liés à la reconnaissance automatique de l'écriture manuscrite. Deuxièmement, la différence entre la reconnaissance en ligne et hors ligne. Procédure d'identification générale. Ensuite, nous avons discuté de chaque étape du modèle (acquisition, prétraitement, extraction de caractéristiques, classification, post-traitement). En spécifiant la méthode d'identification principale. on vas montrer dans le chapiter suivant l'implémentation de notre approche.

Chapitre 4

Implémentation

4.1 Introduction

Dans ce chapitre, nous allons définir le schéma du modèle que nous avons créé pour la détection et la reconnaissance de l'écriture manuscrite. Pour aboutir à notre objectif, nous allons utiliser le langage de programmation python et ses bibliothèques keras-ocr, tensorflow et matplotlib.

4.2 Reconnaissance d'écritre manuscrite avec le transfert d'apprentissage

L'apprentissage par transfert est une technique d'apprentissage automatique où un modèle développé pour une tâche est réutilisé comme point de départ pour un modèle sur une deuxième tâche. Le Transfer Learning permet de faire du Deep Learning sans avoir à passer un mois de calculs. Le principe est d'utiliser les connaissances acquises par un réseau de neurones lors de la résolution d'un problème afin d'en résoudre un autre plus ou moins similaire. Cela permet un transfert de connaissances.

4.3 Architecture général de notre approche

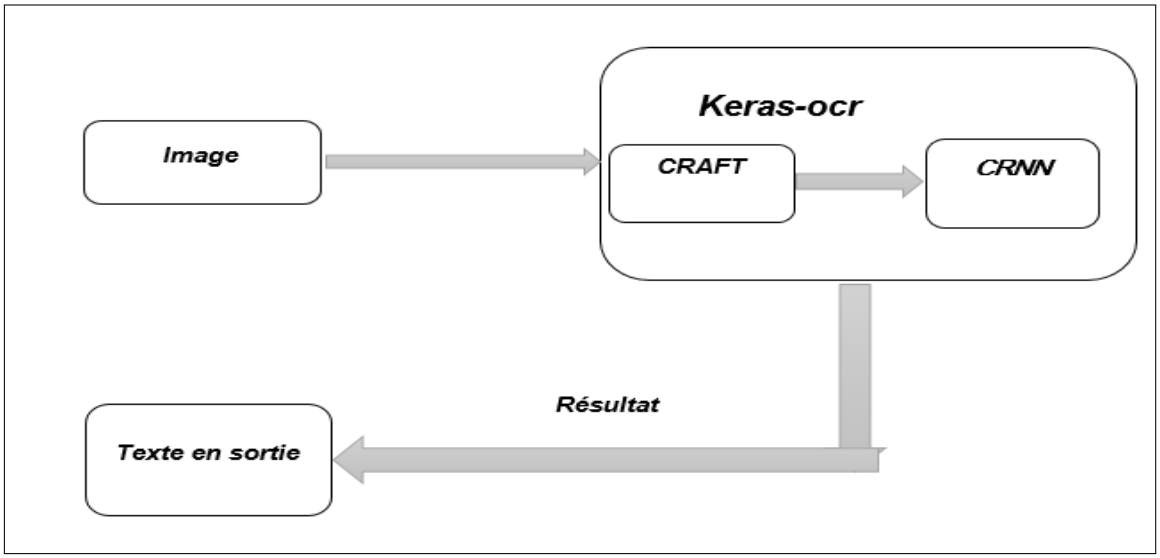


FIGURE 4.1 – Architecture Général

4.4 keras$_{o}cr$

keras-ocr fournit des modèles OCR prêts à l'emploi et un pipeline de formation de bout en bout pour créer de nouveaux modèles OCR Il s'agit d'une version légèrement peaufinée et packagée de l'implémentation Keras CRNN et du modèle de détection de texte CRAFT publié. Il fournit une API(Application Programming Interface) de haut niveau pour la formation d'un pipeline de détection de texte et d'OCR.

Dans nos expériences, nous avons utilisé keras ocr il se compose de deux étapes ; la première étape c'est détection du texte, la deuxième étape c'est la reconnaissance de texte.

4.4.1 CRAFT (Character-Region Awareness For Text detection)

Le modèle CRAFT est un détecteur du texte qui détecte efficacement la zone du texte en explorant chaque région de caractère et la distance entre les caractères pour distinguer les mots. La boîte englobante des textes est obtenue en trouvant simplement des rectangles de délimitation minimum sur une carte binaire après seuillage de la région des caractères et des scores d'affinité

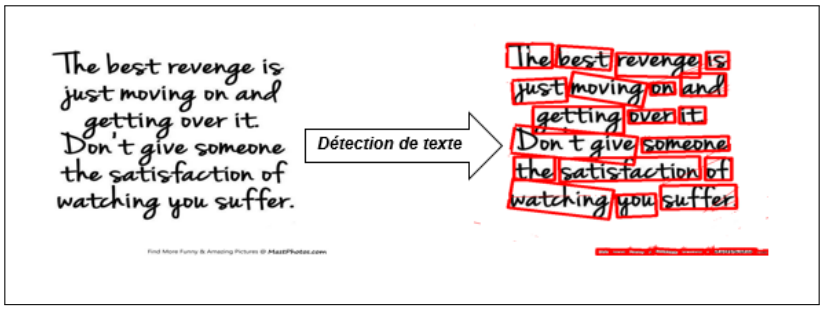


FIGURE 4.2 – Détection de texte de première étape.

4.4.2 CRNN (Convolutional Recurrent Neural Networks)

Il s'agit d'une implémentation d'un réseau de neurones profonds pour la reconnaissance de texte de scène. Le modèle consiste en une étape CNN extrayant des caractéristiques qui sont transmises à une étape RNN (Bi-LSTM) et une perte CTC(Classification Temporelle Connectioniste) pour les tâches de reconnaissance de séquences basées sur l'image, telles que la reconnaissance de texte de scène et l'OCR.

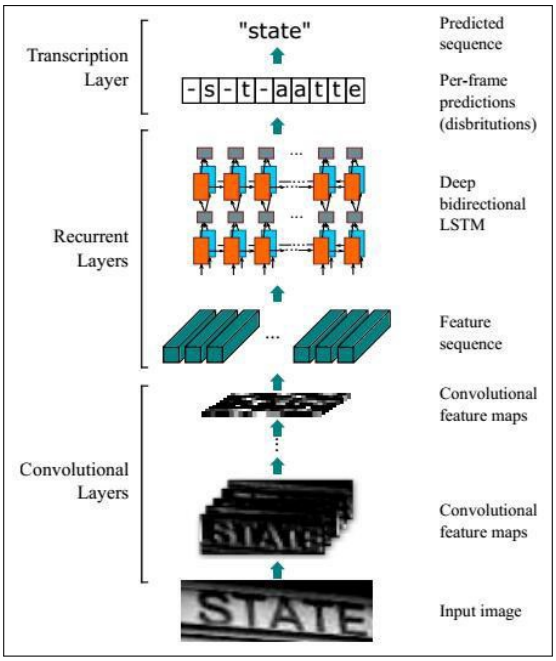


FIGURE 4.3 – La structure du réseau de neurones récurrent à convolution

Nous avons utilisé une fonction spéciale dans la bibliothèque keras-ocr pour Extraire et dessiner des annotations de texte sur l'image.

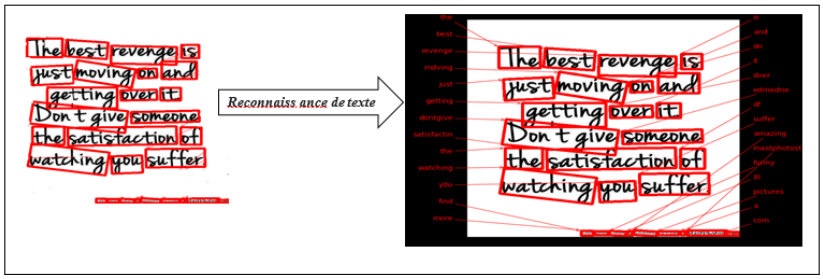


FIGURE 4.4 – Reconnaissance de texte de deuxième étape

4.5 Conclusion

Dans ce chapitre, nous avons introduit une méthode de reconnaissance de l'écriture manuscrite basée sur CRNN (Convolutional Recurrent Neural Network). Nous montrerons dans le chapitre suivant les résultats obtenus par notre méthode .

Chapitre 5

Résultat

5.1 Introduction

Dans ce chapitre, nous allons présenter la configuration du matériel, les bibliothèques et le langage utilisés dans notre travail, puis nous allons présenter les résultats obtenus par notre méthode.

5.2 La Configuration Du matirièle Utilisé dans l'implémentation

Nous avons utilisés un ordinateur portable DELL

- un processeur Intel Core i5-4210M 1.70 GHz
- Mémoire installée (RAM) : 4,00 Go.
- Système d'exploitation Windows 10

5.3 Logiciels et librairies Utilisés dans l'implémentation

- **python** 

  Python est un langage de programmation puissant de haut niveau, à la fois facile à apprendre, il prend en charge plusieurs modèles de programmation (procédural, fonctionnel et orienté objet). Les bibliothèques (packages) de python encouragent la modularité et la réutilisabilité des codes existants. Python et ses bibliothèques sont disponibles sans difficulté pour la majorité des plateformes et il peut être redistribué gratuitement. On estime que c'est l'un des langages de programmation les plus utilisés au monde. [38]

- **jupyter** 

  L'ordinateur portable étend l'approche basée sur la console à l'informatique interactive dans une direction qualitativement nouvelle, offrant un application Web adaptée à la capture de l'ensemble du processus de calcul : développement, documentation et exécution code, ainsi que la communication des résultats. Le notebook Jupyter combine deux composants :
  Une application Web : un outil basé sur un navigateur pour la création interactive de documents qui combinent un texte explicatif, les mathématiques, les calculs et leur production rich media.

Documents de bloc-notes : une représentation de tout le contenu visible dans l'application Web, y compris les entrées et les sorties de les calculs, le texte explicatif, les mathématiques, les images et les représentations multimédias enrichies d'objets.[4]

5.4 Résultat

Dans cette section, nous évaluerons notre méthode en termes de précision et de performance. Pour montrer son efficacité. La figure ci-dessous montre quelques résultats obtenus par notre méthode.

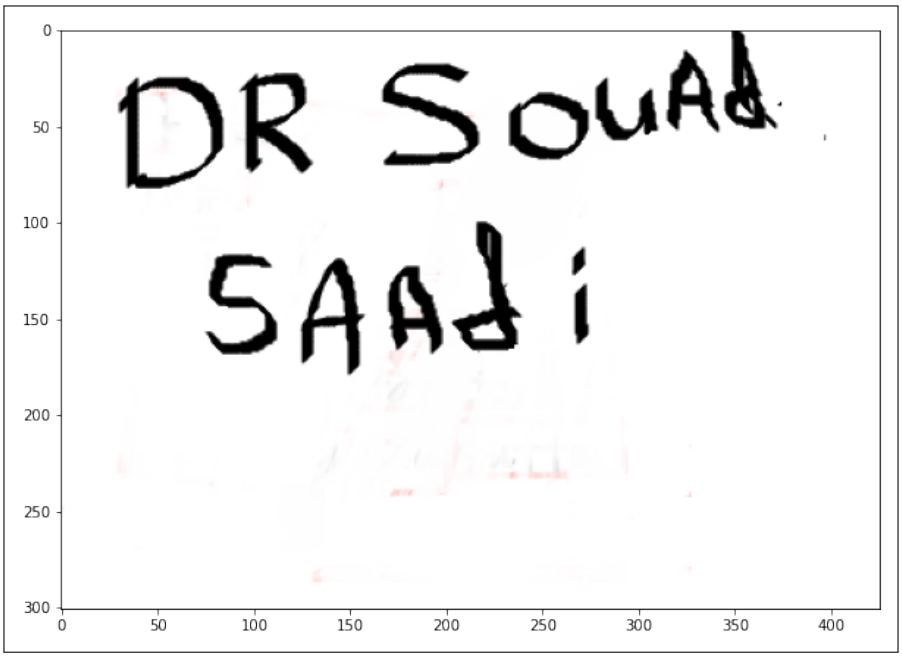


FIGURE 5.1 – L'image 01 avant reconnaissance d'écriture

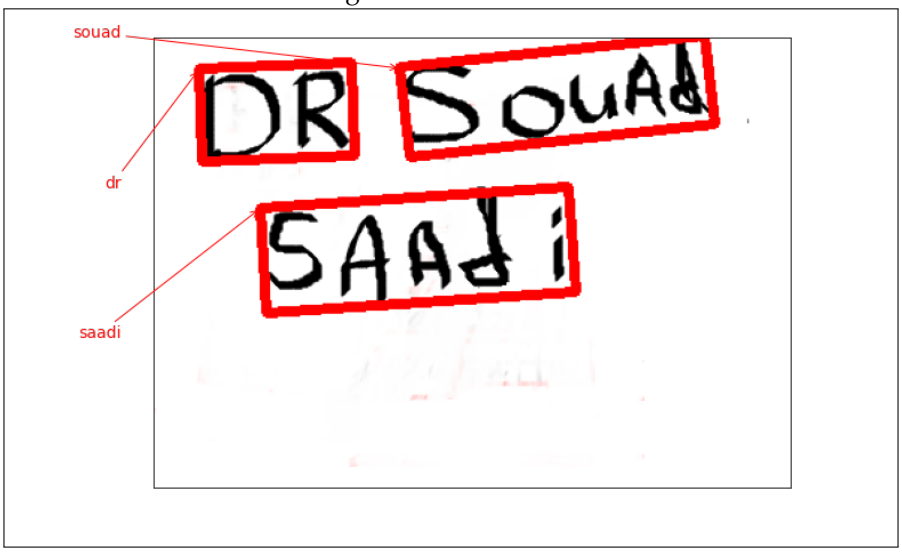


FIGURE 5.2 – L'image 01 aprés reconnaissance d'écriture

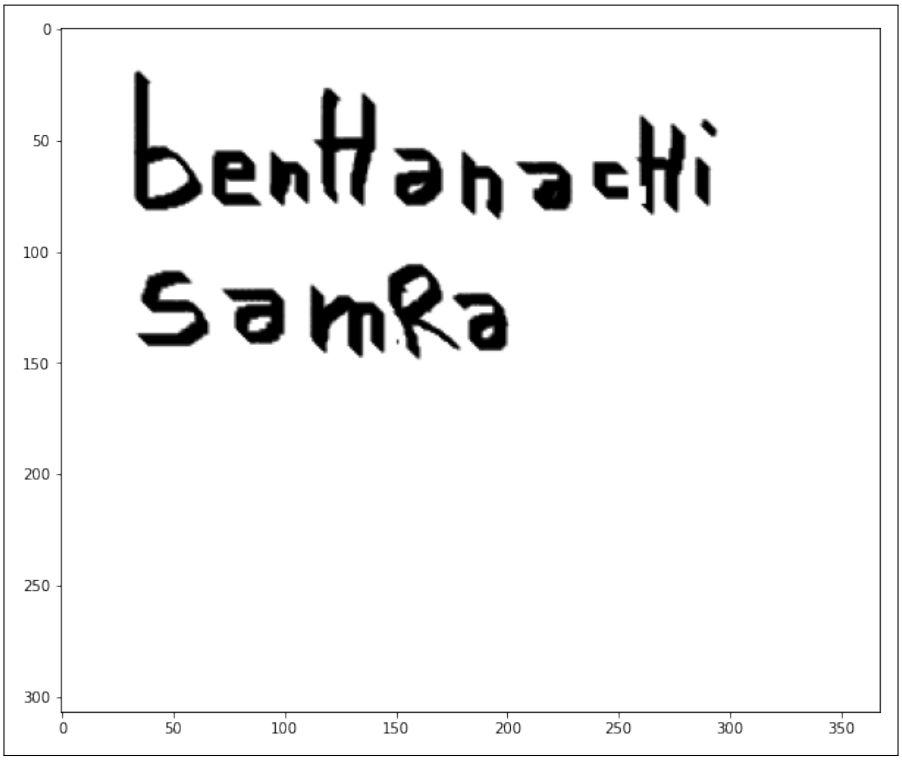


FIGURE 5.3 – L'image 02 avant reconnaissance d'écriture

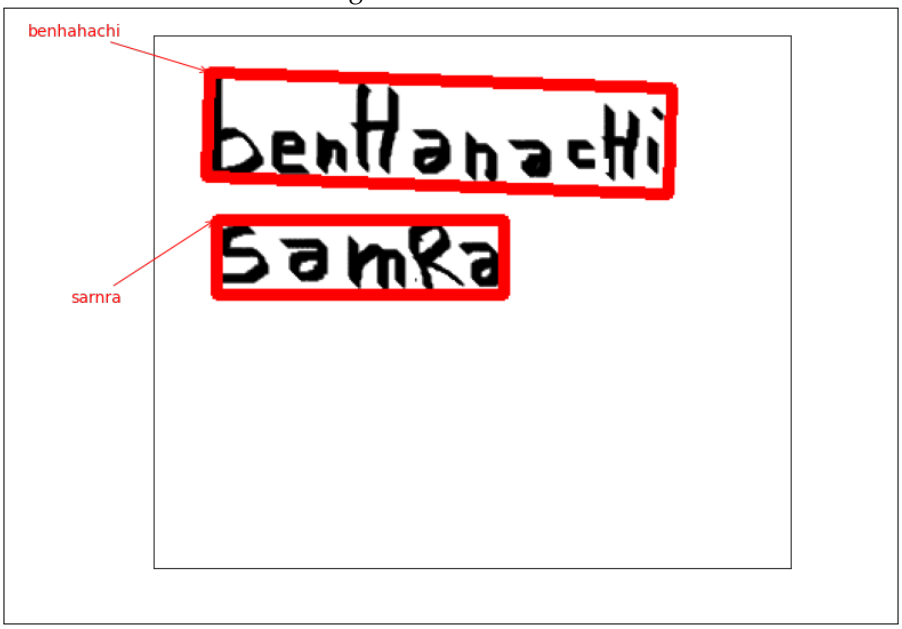


FIGURE 5.4 – L'image 02 aprés reconnaissance d'écriture

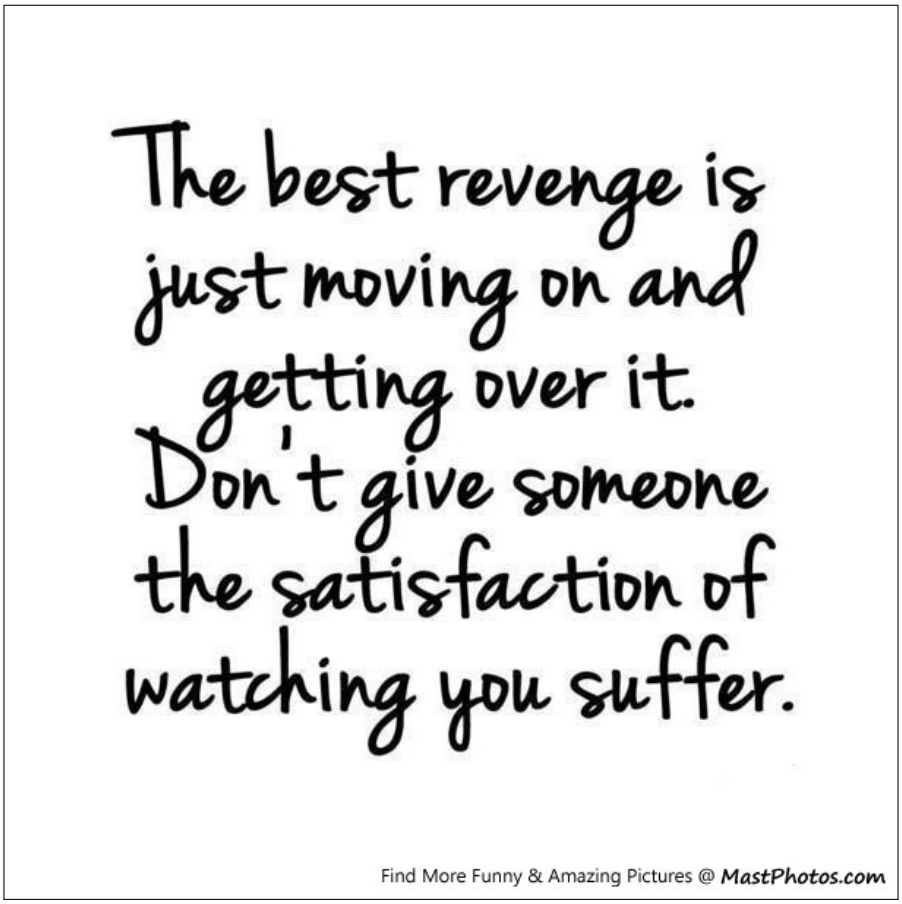


FIGURE 5.5 – L'image 03 avant reconnaissance d'écriture

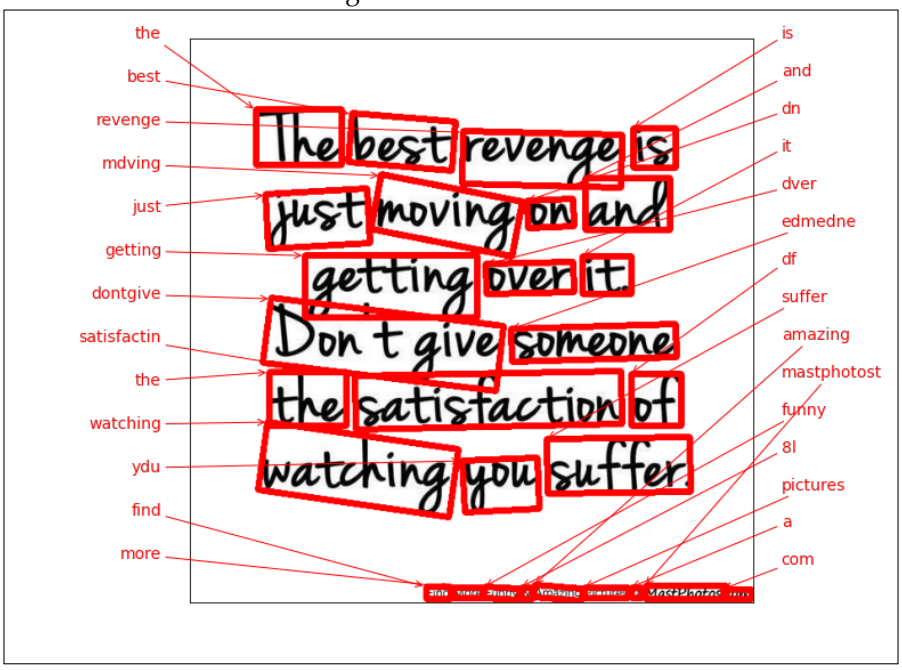


FIGURE 5.6 – L'image 03 aprés reconnaissance d'écriture

Les prochaines images proviennent également de notre environnement mais avec une différente, ces images contient des chiffres :

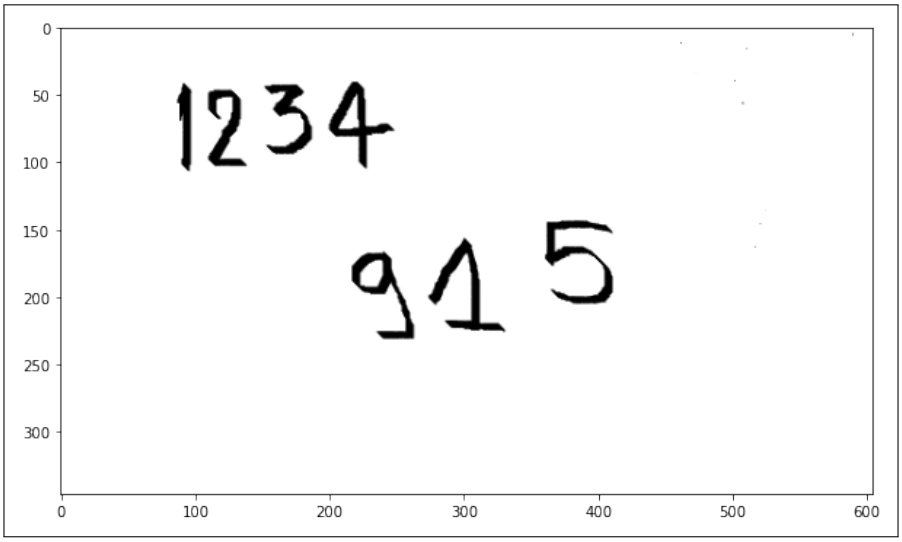


FIGURE 5.7 – L'image 01 avant reconnaissance d'écriture

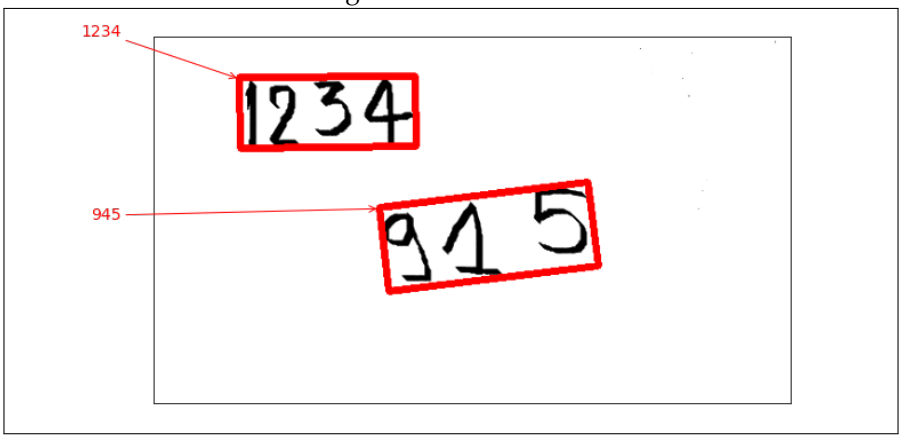


FIGURE 5.8 – L'image 01 aprés reconnaissance d'écriture

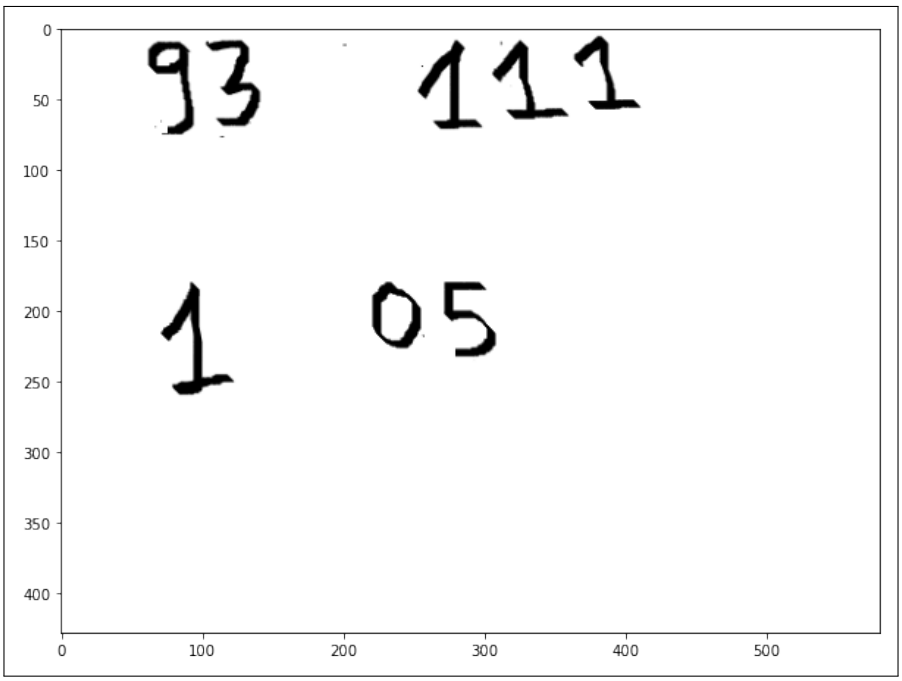


FIGURE 5.9 – L'image 02 avant reconnaissance d'écriture

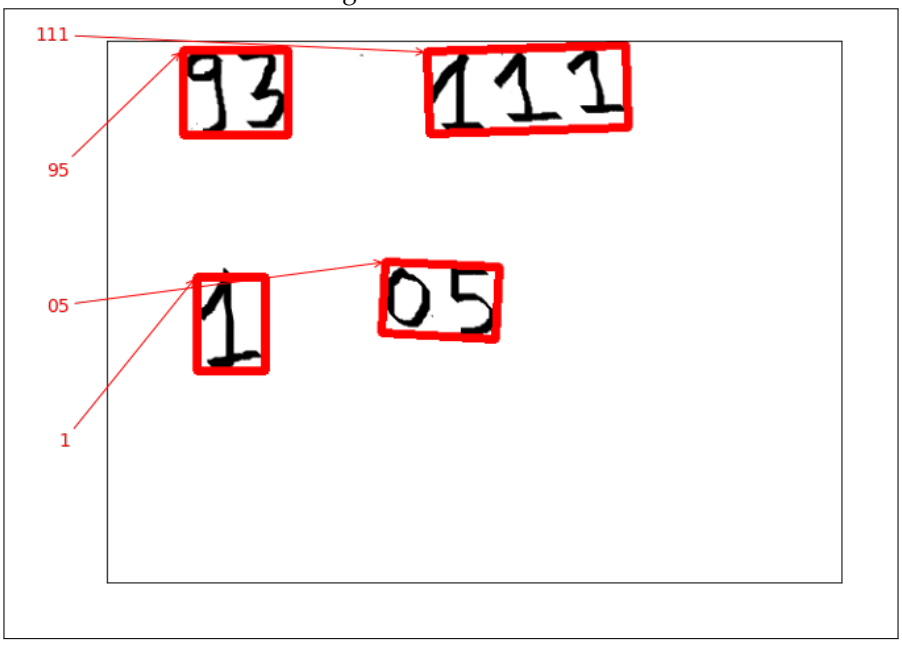


FIGURE 5.10 – L'image 02 aprés reconnaissance d'écriture

Enfin, dans la figure suivante, nous allons essayer une image qui contient des chiffres et des mots ensemble :

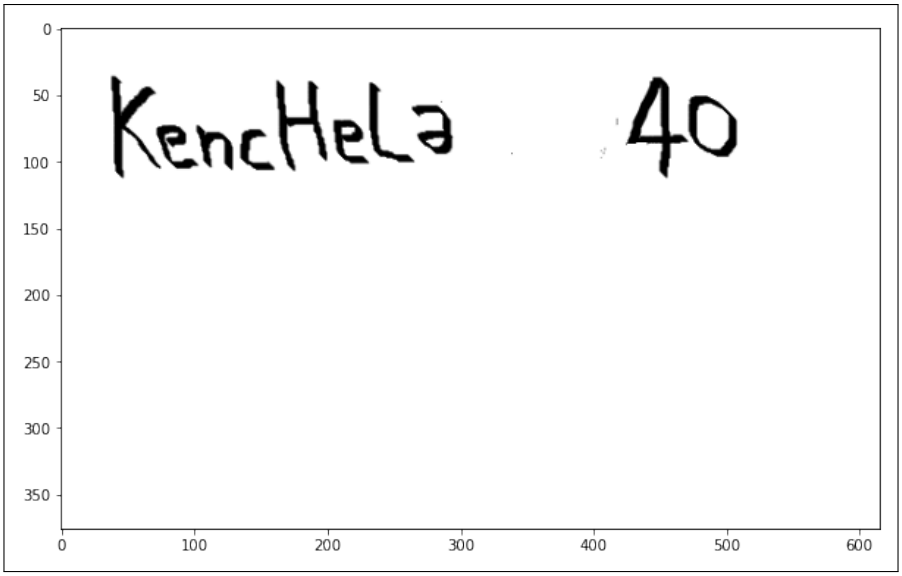


FIGURE 5.11 – L'image 01 avant reconnaissance d'écriture

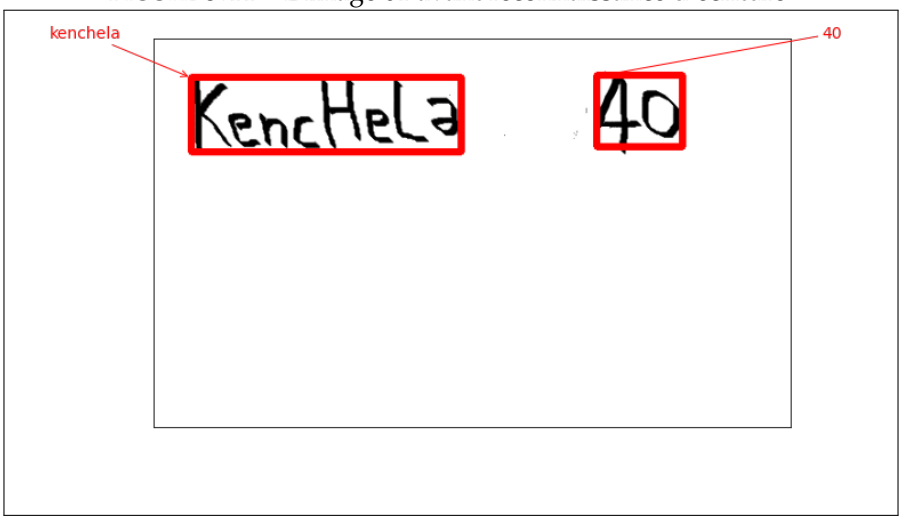


FIGURE 5.12 – L'image 01 aprés reconnaissance d'écriture

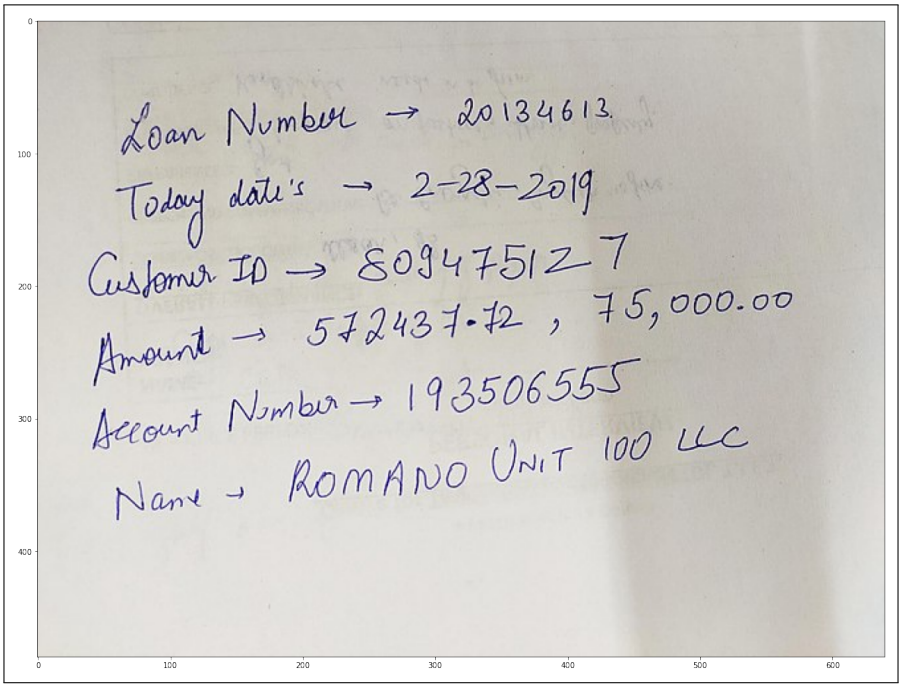


FIGURE 5.13 – L'image 02 avant reconnaissance d'écriture

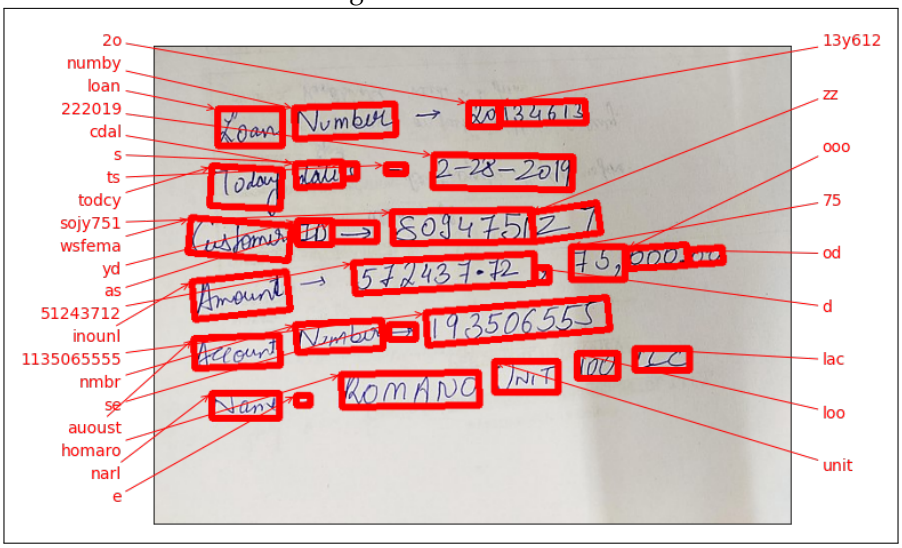


FIGURE 5.14 – L'image 02 aprés reconnaissance d'écriture

Remarquant dans les figures précédentes que la détection et la reconnaissance de texte a été bien fait par notre méthode. Mais on a remarqué qu'il y a des erreurs qui apparaissent lorsque certaines circonstances involontaires interfèrent surtout si les mots sont très proches les uns des autres (comme dans la dernière image).

5.5 Conclusion

Dans ce chapitre, nous avons montré les résultats obtenus par notre méthode . Ces résultats sont avérés satisfaisants en termes de détection et de reconnaissance d'écriture manuscrite apartire des images contenant Des textes manuscrits .Dans la partie suivante on vas présenter la conclusion générale de ce mémoire et les perspectives proposées .

Conclusion Générale

La reconnaissance automatique de l'écriture et des tracés manuscrits en général permet maintenant d'analyser, de comprendre, d'interpréter et de retranscrire « au propre » les actions réalisées à main levée par l'utilisateur. C'est alors l'ordinateur qui produit l'effort d'interprétation, libérant ainsi l'utilisateur des contraintes classiques de communication avec la machine : l'utilisateur pourra rester concentré sur le processus créatif associé à la réalisation de son document.

Le travail qui suit est le résultat de recherches dans le domaine de la détection et de la reconnaissance de l'écriture manuscrits. L'objectif de notre travail est de développer un outil de détection et de reconnaissance d'écriture méthode permettant d'extraire du texte et des mots.

Pour atteindre notre objectif, nous avons utilisé keras-ocr il consiste de deux étapes :

Il est implémenté avec CRAFT comme détecteur de texte et CRNN comme reconnaissance de texte.

Les résultats obtenus à partir de notre méthode montrent une amélioration remarquable et une grande efficacité.

Nous sommes sortis avec de nouvelles perspectives après avoir vu les résultats de ce travail, les

le plus important est que nous pourrons faire quelques ajustements sur notre méthode à l'avenir

. Nous pouvons également utiliser le flux vidéo à la place des images...

Bibliographie

[1] https://www.widoobiz.com/wp-content/uploads/2019/05/capture-dexxcran-2019-05-22-axx-18.37.20.png.

[2] Cours de data mining – 8 : Réseaux de neurone.

[3] Différence entre l'apprentissage automatique et l'apprentissage profond. http://fr.acervolima.com/difference-entre-l-apprentissage-automatique-et-l-apprentissage-profond/.

[4] Jupyter notebook documentation release 7.0.0a5.

[5] Safa ABA and Maroua CHIKH. Reconnaissance des mots arabes manuscrites. 2016.

[6] Zerdani Amina. " reconnaissance de caractères arabes manuscrits par réseau de hopfield » mastère en informatique université oum el bouaghi juin 2014 ,. 2011.

[7] ALLAL Mohammed Anes. " « utilisation-du-deep-learning » université tlemcen mastère en informatique université. 2017/2018.

[8] LOUAM ABDELHAK BILAL. "deep learning basé sur les méthodes de réduction pour la reconnaissance de visage, Le : Samedi 6 juillet 2019.

[9] Olivier Boisard. intelligence artificielle.

[10] Sanaa Chafik. Machine learning techniques for content-based information retrieval. diss. université paris-saclay. 2017.

[11] Khouloud Dahmane.

[12] Hala Djerouni. " développementd'unsystème dereconnaissance dechiffresmanuscrits,universitÉ echahid hamma lakhdar - d'el oued. 2020/2021.

[13] G. Dreyfur. " « réseaux de neurones : Méthodologie et application », édition eyrolles. 2004.

[14] P.Purohit D.Singh, N.Mehta. "text based image recognition using multilayer perceptron", special issues on ip multimedia communications, pp : 143-146. 2011.

[15] Boughaba Mohammed et Boukhris Brahim a. "« l'apprentissage profond (deep learning) pour la classification et la recherche d'images par le contenu » mastère en informatique université ouargla. 2016/2017.

[16] Chekilla Zanoubia et Guellab Salim. "vers un nouveau modèle organisationnel basé sur les réseaux de neurones pour les systèmes multi agents, universite « abbés laghrour » de khenchela ,. 2021.

[17] A. Belaïd et Y. Belaïd. " reconnaissance des formes méthodes et applications, interedition,. 1992.

[18] RAKOTONDRAINA Tahina Ezéchiel. " « systeme de detection d'objet base sur le deep learning dans un magasin intelligent ». mastère en informatique Université Biskra2018/2019.

[19] C. Y. Suen F. Lauer and G. Bloch. "a trainable feature extractor for handwritten digit recognition,.

[20] Najeh Naakhi Faiz, Rim and Khaled Mellouli. Apprentissage supervisé pour la classification des images basé sur la structure p-tree. EGC. 2005.

[21] Daoud fouad et louli farouk. " la reconnaissance des caractères arabes manuscrits par les reseaux convolutifs » thèse magistère en informatique université blida. 2020.

[22] A. Gattal. " segmentation automatique pour la reconnaissance numérique des chéques bancaires algériens", thèse, thèse de magister, centre universitaire de khanchela. 2011.

[23] A. Géron. Hands-on machine learning with scikit-learn and tensorflow : concepts, tools, and techniques to build intelligent systems. "o'reilly media, inc. 2017.

[24] C. J. Darken J. Moody. "fast learning in networks of locally-tuned processing units". neural computation, 1, pp : 281-294. 1989.

[25] Benmedjber kahina. "la reconnaissance des caractères arabes manuscrits par les transformateurs,universite larbi ben m'hidi -oum el bouaghi. 2007/2008.

[26] Benmedjber kahina. "la reconnaissance des caractères arabes manuscrits par les transformateurs ,universite larbi ben m'hidi -oum el bouaghi, 2020/2021.

[27] Ludovic L. : Machine learning et deep learning, comment ça marche ?, siecledigital.fr. 22 décembre 2016.

[28] A. Menasria. "reconnaissance d ' écriture manuscrite par technique à base de réseaux de neurones ", thèse de magister de l'université de constantine. MARCH 1999.

[29] Luc Mioulet. " reconnaissance de l'écriture manuscrite avec des réseaux récurrents. traitement du texte et du document. université de rouen,. 2015. Français. tel-01301728.

[30] Dong Xiao Ni. ""application of neural networks to character recognition", proceedings of students/faculty research day, csis, pace university.

[31] Nicolle. Anne. l'expérimentation et l'intelligence artificielle. intellectica. 1996.

[32] Alencar P. Cowan D. Portugal, I. The use of machine learning algorithms in recommender systems : a systematic review. expert systems with applications. 2017.

[33] Priyadharshini. Machine learning : What it is and why it matters, ''simplilearn,". 2018.

[34] Azizi Rebiai. "une approche hybride pour la reconnissance d'ecriture arabe manuscrite. ANNEE 2006/2007.

[35] reda adjoudj. l'intelligence artificielle,. *chapter de univeersity de sidi-bel-abbes*, 2006.

[36] Richard Kinh Gian Do Kaori Togashi Rikiya Yamashita, Mizuho Nishio. "« convolutional neural networks : an overview and application in radiology. ». Thèse du doctorat 2018.

[37] N. SALIMA. ""reconnaissance de caractères manuscrits par intelligence collective", thèse, université de sétif 1-ferhat abbas. 1992.

[38] G. (1990). The python language. See http ://www. python. org. Van Rossum. https://www.python.org.

[39] Rachid Zaghdoudi. "reconnaissance hors ligne des mots arabes manuscrits par les réseaux de neurones. 2008.

[40] F. Zaiz. " technique basée puzzle/svm pour l'amélioration de la reconnaissance du texte arabe manuscrit", thèse, université mohamed khider-biskra. 2017.

[41] Mr Mokri Mohammed Zakaria. Classification des images avec les réseaux de neurones convolutionnels. 2017.

[42] zermi narimane. " « reconnaissance de mots manuscrits arabe par les modèles de markov » thèse de doctorat université annaba. Juin 2014.