



République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique  
Université Ahmed Draïa d'Adrar  
Faculté des Sciences et de la Technologie  
Département de Mathématiques et Informatique

Mémoire de fin d'étude, en vue de l'obtention du diplôme de master en informatiq

**Option** : Systèmes Intelligents

**Thème:**

**Reconnaitances de chiffres manuscrits  
par réseau de neurones convolutifs**

**Préparé par :**

BELBALI Hlima

**Encadrés par :**

Mr. MAMOUNI El Mamoun

Année universitaire : 2021/2022

الجمهورية الجزائرية الديمقراطية الشعبية

People's Democratic Republic of Algeria

Ministry of Higher Education and  
Scientific Research  
University Ahmed Draia of Adrar  
The central library



وزارة التعليم العالي و البحث العلمي  
جامعة أحمد دراية- أدرار  
المكتبة المركزية  
مصلحة البحث البليوغرافي

## شهادة الترخيص بالإيداع

انا الأستاذ : ماموني المأمون

المشرف على مذكرة الماستر الموسومة بـ :

Reconnaisances de chiffres manuscrits par réseau de neurones convolutifs

من إنجاز الطالبة : بلبالي حليلة

كلية : العلوم والتكنولوجيا

القسم : الرياضيات والاعلام الآلي

التخصص: إعلام آلي " أنظمة ذكية "

تاريخ تقييم / مناقشة: 2022/06/21

أشهد أن الطالبة قد قامت بالتعديلات والتصحيحات المطلوبة من طرف لجنة التقييم / المناقشة، وان المطابقة بين  
النسخة الورقية والإلكترونية استوفت جميع شروطها.  
ويمكنها إيداع النسخ الورقية (02) والايكترونية (PDF).

- امضاء المشرف:

ادرار في : 2022/07/13

مساعد رئيس القسم:



و. قاضي  
مساعد رئيس قسم الرياضيات والإعلام الآلي  
مكلف بما بعد التدرج والبحث العلمي بكلية  
العلوم والتكنولوجيا

ملاحظة: لا تقبل أي شهادة بدون التوقيع والمصادقة.

# Remerciements

"Celui qui ne remercie pas les gens n'a pas remercié Dieu. Pour cela, je remercie d'abord Dieu Tout-Puissant, qui m'a accordé sa grâce et m'a permis de mener à bien ce travail.

Je remercie également mes chers parents qui ont tenu à m'encourager tout au long de ma parcours académique.

Ce projet a été mis en œuvre sous la supervision du **Mr. MAMOUNI EI Mamoun**, à qui je tiens à adresser mes sincères remerciements pour sa patience et sa disponibilité, ainsi que ses conseils, qui m'ont aidé à terminer ce travail.

Toute ma gratitude va également aux professeurs et au personnel du Département MI ainsi qu'à mes camarades de classe. À toute personne qui a travaillé directement ou indirectement à la réalisation de ce projet sous quelque forme de contribution que ce soit, retrouvez ici un témoignage de ma profonde gratitude

## Dédicaces

Je dédie cet humble travail à tous les honorables membres de ma famille, chacun en son propre nom, et à tous ceux qui portent mon nom de famille "**Belbali**", à mes camarades de classe, que Dieu leur accorde le succès, à mes distingués professeurs, que Dieu les récompense avec tout le meilleur, et à tous ceux que j'aime et que j'aime sans exception.

## ***Résumé***

Le travail présenté dans Cette thèse présente le développement du système de reconnaissance de chiffres manuscrits à l'aide d'un réseau de neurones convolutifs. Notre système comprend les étapes suivantes : charger une base de données, extraire des caractéristiques de l'image via des couches de réseau de neurones convolutionnels, puis les classer. Afin d'évaluer ce système et d'améliorer sa qualité, nous avons augmenté le nombre de couches au niveau des deux parties du modèle de base (la partie convolutive et la partie classification), et les résultats obtenus ont été très bons au service de notre objectif.

## ***Mots-clés***

Reconnaissance de chiffres manuscrits, Reconnaissance optique de caractères, convolutif, Maxpooling, Fully connected, base de données , CNN ,Kernel ,Filtre, Feature map, MNIST, Modèle de base, RuLu, Softmax.

## **ملخص**

يتمثل العمل المقدم في هذه الأطروحة في تطوير نظام التعرف على الأرقام المكتوبة بخط اليد باستخدام شبكة عصبية تلافيفية. يشتمل نظامنا على الخطوات التالية: تحميل قاعدة بيانات ، واستخراج الميزات من الصورة عبر طبقات الشبكة العصبية التلافيفية ، ثم تصنيفها. من أجل تقييم هذا النظام وتحسين جودته قمنا بزيادة عدد الطبقات على مستوى جزئي النموذج الاساسي ( الجزء التلافيفي وجزء التصنيف) وكانت النتائج المتحصل عليها جيدة للغاية تخدم هدفنا هذا.

## **كلمات البحث**

التعرف على الأرقام المكتوبة بخط اليد ، التعرف الضوئي على الأحرف ، التلافيف ، Maxpooling ، الاتصال بالكامل ، قاعدة البيانات ، CNN ، Kernel ، Filter ، Feature map ، MNIST ، Basic model ، RuLu ، Softmax.

## ***Abstract***

The work presented in this thesis presents the development of the handwritten number recognition system using a convolutional neural network. Our system includes the following steps: loading a database, extracting features from the image via convolutional neural network layers, then classify them. In order to evaluate this system and improve its quality, we have increased the number of layers at the two parts of the base model (the convolutional part and the classification part), and the results obtained have been very good in the service of our objective.

## ***Keywords***

Recognition of handwritten digits, optical character recognition, convolutional , Maxpooling, Fully connected, database, CNN ,Kernel ,Filtre, Feature map, MNIST, Basic Model, RuLu, Softmax.

## *Table des matières*

Remerciements.....	I
Dédicace.....	II
Résumé.....	IV
Abstract.....	IV
ملخص.....	V
Tables desmatières.....	VI
Liste des tableaux .....	IX
Liste des figures.....	X
Liste des abréviations.....	XI
Introduction générale.....	1
<b>Chapitre I : Reconnaissance d'écriture arabe manuscrite</b>	
1. Introduction.....	1
2. Définition.....	2
3. Histoire.....	3
4. Domaines d'application.....	3
5. Classification de l'ocr.....	3
5.1 Reconnaissance en-ligne et hors-ligne.....	4
5.2 Approche de reconnaissance.....	5
6. Structure d'un système de reconnaissance optique de caractères.....	6
.6.1. Acquisition d'images.....	6
6.2. Segmentation des caractères .....	7
6.3. Extraction de caractéristiques.....	7
6.4. Classification.....	7
7 Conclusion.....	7
<b>Chapitre II: Réseaux de neurones convolutifs</b>	
1. Introduction.....	10
2. Réseau de neurones artificiels (RNA).....	10
3. Types de réseaux de neurones artificiels.....	10
3.1. Réseaux de neurones récurrents.....	10
3.2. Réseaux de neurones à propagation avant ( feed-forward).....	11
3.2.1. Perceptron.....	11
3.2.2. Perceptron multi couches (Multi Layers Perceptron(MLP)).....	11

a. Rétropropagation du gradient.....	12
3. Réseaux de neurones convolutifs (ConvNet /CNN)	
4. Histoire.....	13
5. Construction du réseau de neurones.....	13
5.1. Partie convolutive.....	13
5.2. Partie classification.....	13
6. Couches de CNN.....	14
6.1 Couche de convolution.....	14
a. Convolution.....	14
b. Couche convolutif.....	14
6.2. Couche de <i>pooling</i> .....	14
6.2.1. Max pooling.....	14
6.2.2. Average pooling.....	14
6.3. Couche fully-connected.....	14
7 Fonctions d'activation.....	15
7.1.Fonction ReLU ( <i>Rectified Linear Units</i> ).....	15
8. Conceptions d'architecture CNN.....	18
9. Conclusion.....	19
<b>Chapitre III: Résultats expérimentaux et Discussions</b>	
1. Introduction.....	22
2. Ressources matérielles et logicielles.....	22
2.1. Ressources matérielles .....	22
2.2. Ressources logicielles.....	22
2.3. Tensorflow & Keras.....	22
a. Tensorflow .....	22
b. Keras.....	23
3. Base de données MNIST.....	23
4. Développer notre système de reconnaissance.....	24
4.1 Télécharger la base de données .....	24
4.2. Préparation des données de pixels.....	25
4.3. Définition du modèle.....	25
a. Partie convolutive.....	25
b. Partie de classifieur.....	25
4.4. Évaluation du modèle.....	25



5 Résultats.....	25
a. Diagnostiquer le comportement d'apprentissage du modèle pendant l'entraînement.....	26
b. Estimer les performances du modèle.....	26
5.1 Améliorations des performances du modèle.....	27
5.2 Résultats expérimentaux .....	27
a. Expérimentation 1 .....	27
b. Expérimentation 2.....	28
c. Expérimentation 3.....	29
5.3. Résumé des expérimentations .....	30
6 Conclusion.....	30

### *Liste des tableaux*

Tableau 1.1 : Fonctions d'activation.....	15
Tableau 3.2 : Résultat de comparaison .....	28
Tableau 3.3 : Résultat de comparaison .....	29
Tableau 3. 4: Résultat de comparaison .....	30

## *Liste des figures*

Figure 1.1 : technique d'OCR.....	3
Figure 1.2: Type de reconnaissance d'écriture.....	4
Figure 1.3 : Exemple de reconnaissance en ligne.....	5
Figure 1.4 : Exemple de reconnaissance hors- ligne.....	5
Figure 1.5 : Schéma général d'un système de reconnaissance d'écriture manuscrite...6	
Figure 2.1 : un neurone biologique VS réseau de neurones artificiel.....	10
Figure 2.2 : Réseaux de neurones à propagation avant (feed-forward) .....	11
Figure 2.3 : Schéma d'une LT.....	12
Figure 2.4 : Perceptron multicouche.....	13
Figure 2.5 : Schéma représentant l'architecture d'un CNN.....	14
Figure 2.6 : Schéma représentant les couches de CNN.....	15
Figure 2.7 : Schéma du parcours de la fenêtre de filtre sur l'image.....	16
Figure 2.8: Processus de Max-Pooling .....	16
Figure 2.9: Processus Average pooling.....	17
Figure 2.10 : Schéma représentant couche fully-connected .....	17
Figure 2.11 : Fonction ReLU.....	19
Figure 2.12 : Fonction Softmax.....	19.
Figure 3.1 : Exemple MNIST .....	24
Figure 3. 2 : Schéma représentant l'architecteur de CNN.....	24
Figure 3.3 : Résolution de classification pour chaque pli de la validation croisée.....	26
Figure 3.4 : graphique de diagnostic des performances du basic model.....	26
Figure3.5 : Résolution de classification pour chaque pli de la validation croisée .....	27
Figure 3.6 : graphique de diagnostic des performances de l'Improved model1.....	27
Figure3.7 : Résolution de classification pour chaque pli de la validation croisée.....	28
Figure 3.8 : graphique de diagnostic des performances du Improved model2.....	28
Figure 3.9 : Résolution de classification pour chaque pli de la validation croisée.....	29
Figure 3.10 : graphique de diagnostic des performances du Improved model3.....	29

## *Liste des abbreviations*

**OCR:** Optical Character Recognition

**CNN:** Convolutional Neural Networks

**LTU:** Linear Threshold Unit

**IA:** Intelligence Artificielle

**RNA :** Artificiel neural network

**MLP:** Multi Layers Perceptron

**MNIST:** Modified ou Mixed National Institute of Standards and Technology

## **Introduction générale**

L'intelligence artificielle connaît une croissance exponentielle pour combler le fossé entre les capacités des humains et des machines. Les chercheurs travaillent dans de nombreux aspects du domaine pour réaliser des choses incroyables. Le domaine de la vision par ordinateur est l'un des nombreux domaines similaires.

L'objectif de ce domaine est de permettre aux machines de voir le monde comme les humains, de le percevoir de manière similaire et même d'utiliser les connaissances pour de nombreuses tâches telles que la reconnaissance d'images et de vidéos, l'analyse et la classification d'images, la création des médias, les systèmes de recommandation, le langage naturel, traitement , etc...

Les progrès de la vision par ordinateur avec apprentissage en profondeur ont été générés et perfectionnés au fil du temps, principalement sur un algorithme spécifique " un réseau neuronal convolutif".

L'objectif de notre projet est de proposer un système de Reconnaissance de chiffres manuscrits par réseau de neurones convolutif. Ce système utilise un algorithme de réseau de neurones convolutif pour extraire les caractéristiques et leur calcification.

Notre mémoire est basée sur trois chapitres:

Le premier chapitre présente le concept général du système de reconnaissance de caractères et quel sont les processus de reconnaissance optique de caractères et mettre en évidence ses différents aspects, ainsi que les différentes étapes du processus de reconnaissance.

Le deuxième chapitre représente comment reconnaître des images par réseau de neurones convolutifs avec ces différents couches et quelque Conceptions d'architecture qui connaît encore de belles mises à jour dues à l'intérêt des chercheurs dans ce domaine.

Dans le dernier chapitre, nous avons fourni une description détaillée du système de reconnaissance des chiffres manuscrits, Comment développer ce système et extraire des résultats avec lesquels nous évaluons le système.

Le travail se terminera par une conclusion générale indiquant la force de ce domaine dans la reconnaissance des images en général et chiffres manuscrits en particulier.



# Chapitre I

Généralité sur le système de reconnaissance  
de caractères

## 1.1 Introduction

L'écriture sous ses diverses formes, imprimée et manuscrite, a toujours joué un rôle important dans la communication humaine, elle est utilisée pour conserver et archiver les connaissances. Certaines connaissances peuvent voir leurs pages endommagées avec le temps ou par l'eau ou le feu, ou l'écriture peut être obsolète ou mal écrite, Provoquant un problème pour les ordinateurs ou pour la personne elle-même Ce problème a donné naissance au domaine de la reconnaissance automatique de l'écriture manuscrite et un traitement d'image informatiquement qui a aidé les humains à lire des textes très difficiles.

Dans ce chapitre nous allons présenter ce qu'est un OCR ainsi que ses domaines d'application et Les différents aspects de son processus

## 1.2 Définition

La reconnaissance optique de caractères, ou (Optical Character Recognition OCR en anglais), est une technique basée sur la conversion électronique d'images textuelles dactylographiées, manuscrites ou imprimées. Toute traduction des images dans un fichier au format texte. Il analyse le fichier image (ex. PDF, jpg, .png, etc.) et convertit les caractères en un fichier texte (Microsoft Word, Excel, etc.) afin de pouvoir modifier et travailler sur ce texte.

Ce programme permet de pouvoir récupérer du texte sous forme de texte imprimé et de l'enregistrer dans un fichier exploitable dans un traitement de texte pour l'enrichissement et stocké dans une base de données ou tout autre support exploitable par le système informatique.

Avec l'OCR, un grand nombre de documents papier peuvent être convertis en texte lisible par machine, quels que soient la langue et le format dans lesquels ils ont été rédigés. Pour qu'il parte d'une image numérique produite par un scanner d'une page (document imprimé, papier imprimé, etc.), ou un appareil photo numérique, et produise sous forme de fichier texte une sortie dans divers formats (texte simple, traitement de texte, XML, etc. ).

Cette technologie facilite non seulement le stockage, mais rend également disponible des données auparavant difficiles d'accès.

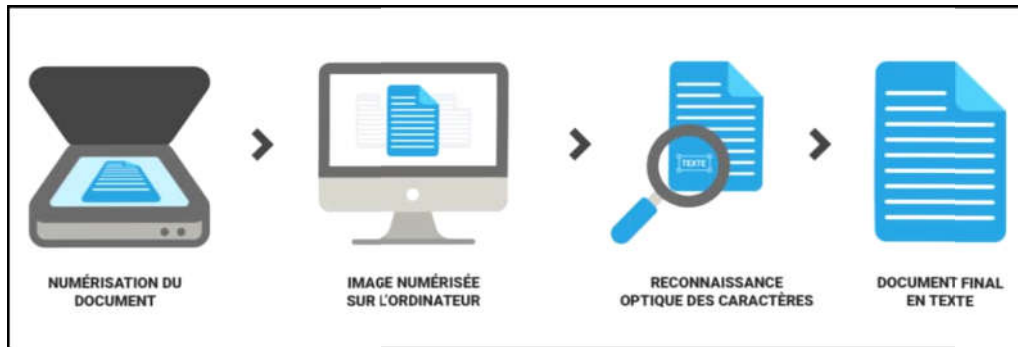


Figure 1.1 : technique d'OCR

### 1.3 Histoire

La première machine OCR a été créée par l'ingénieur allemand Gustav Tauschek en 1929, il contenait un détecteur sensible à la lumière qui dirigeait la lumière vers un mot lorsqu'il correspondait à un bloc de sa mémoire. Ensuite, David Shepard (un cryptanalyste à l'Agence de sécurité nationale des États-Unis (AFSA)) et fondateur de l'IMR a présenté le premier appareil ROC au monde en 1950, IBM et d'autres ont ensuite utilisé les brevets de Shepard.

Il semblait que ces nouvelles connaissances pourraient aider les aveugles à lire des documents enregistrés sur un support informatique, mais il fallut attendre 1974 lorsque Ray Kurzweil appliqua l'OCR et, à l'aide de techniques modernes, créa un synthétiseur audio pour "dire" le texte et perfectionner les processus de numérisation.

### 1.4 Domaines d'application

Les domaines d'utilisation de l'OCR sont très larges, y compris :

- **Domaine bancaire:** la gestion automatique des chèques bancaires ou postaux applications commerciales codes de paquets, dates d'expiration.
- **L'aide à la lecture pour les non-voyants** Les systèmes de reconnaissance associés à des synthétiseurs vocaux permettent la compréhension de documents et livres pour les aveugles.
- **Santé et applications cliniques.**
- ..... etc.

### 1.5 Classification de l'OCR

Il existe plusieurs modes de classification des systèmes OCR parmi lesquels on peut citer :

- Les systèmes " en-ligne " ou " hors-ligne " suivant le mode d'acquisition.
- Les approches globales ou analytiques selon que l'analyse s'opère sur la totalité du mot, ou par segmentation en caractères



### 1.5 .1 Reconnaissance en-ligne et hors-ligne

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

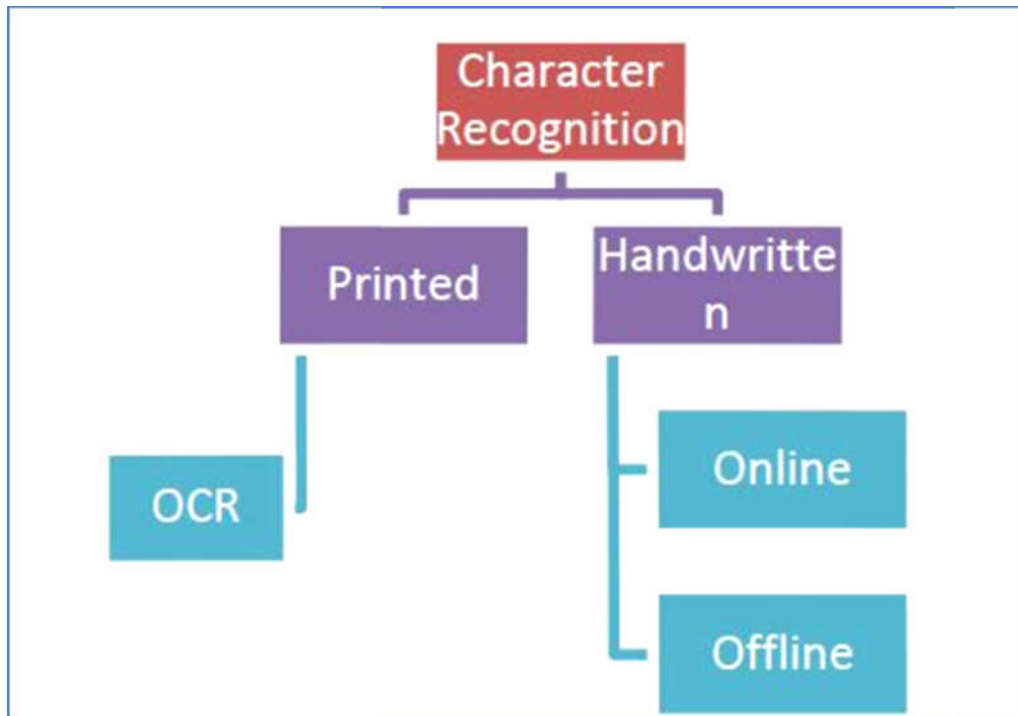


Figure 1.2: Type de reconnaissance d'écriture

#### Reconnaissance en-ligne (on-line)

Cette classe est liée à la connaissance dynamique, dans ce mode la reconnaissance se fait en temps réel (pendant la frappe), de sorte que le signal d'écriture est enregistré directement lors de sa production grâce à un dispositif spécifique, et les symboles sont reconnus comme manuscrits.

La reconnaissance en-ligne présente un avantage majeur c'est la possibilité de correction et de modification de l'écriture de manière interactive vu la réponse en continu du système. L'acquisition de l'écriture est généralement assurée par certains appareils comme une tablette graphique avec un stylet.

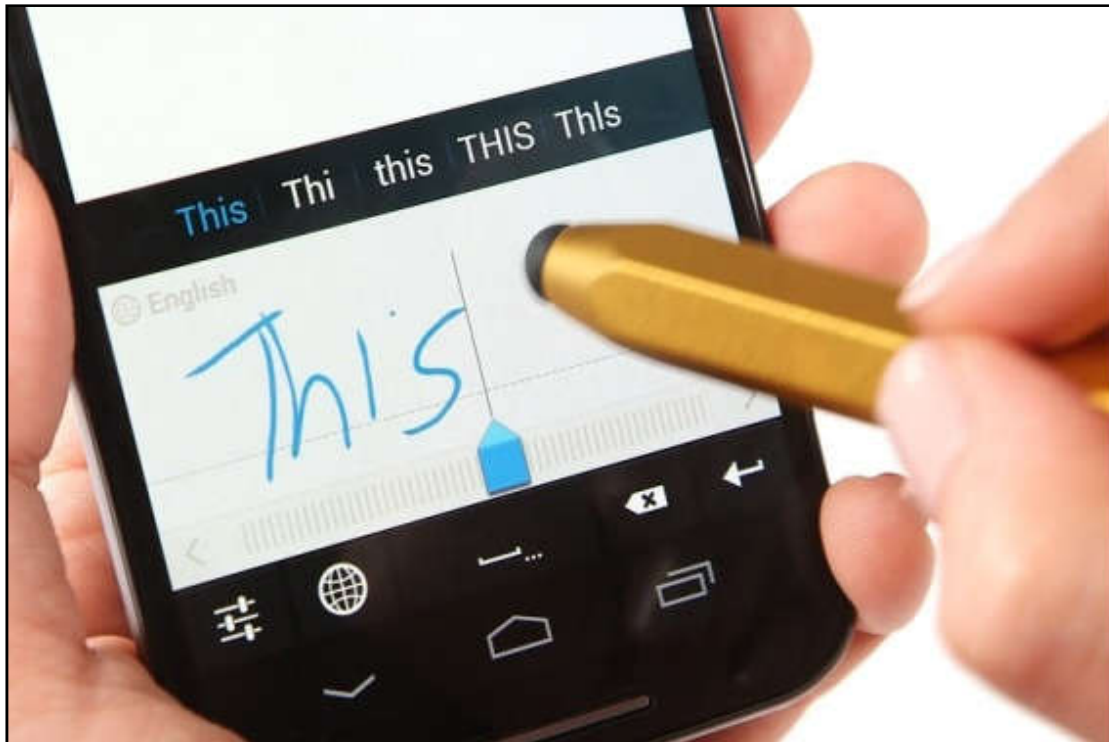


Figure 1.3 : Exemple de reconnaissance en ligne [1]

### Reconnaissance hors-ligne (off-line)

Cette classe est liée à la reconnaissance statique, elle démarre après l'acquisition afin que la donnée disponible soit le résultat de la numérisation d'un document papier contenant de l'écriture (une image numérique), mais avant d'aborder l'étape de la reconnaissance il faut identifier et extraire le type du signal du reste de l'image. Il convient aux documents réellement imprimés et manuscrits (chèques postaux ou bancaires). Ce mode peut être considéré comme le cas le plus général de reconnaissance d'écriture manuscrite.

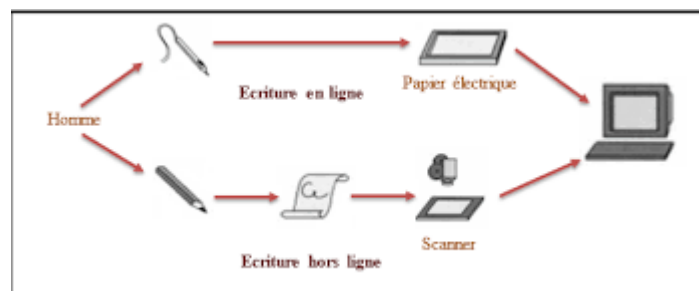


Figure 1.4 : Exemple de reconnaissance hors- ligne

## 1.5 .2 Approche de reconnaissance

### Approche globale

Elle considère un mot comme une entité unique et le décrit indépendamment des caractères qui le composent.

## Approche analytique

Contrairement à l'approche globale, le mot est segmentés en caractères ou en fragments puis les caractères sont reconnues et combinées pour former le mot

### 1.6. Structure d'un système de reconnaissance optique de caractères

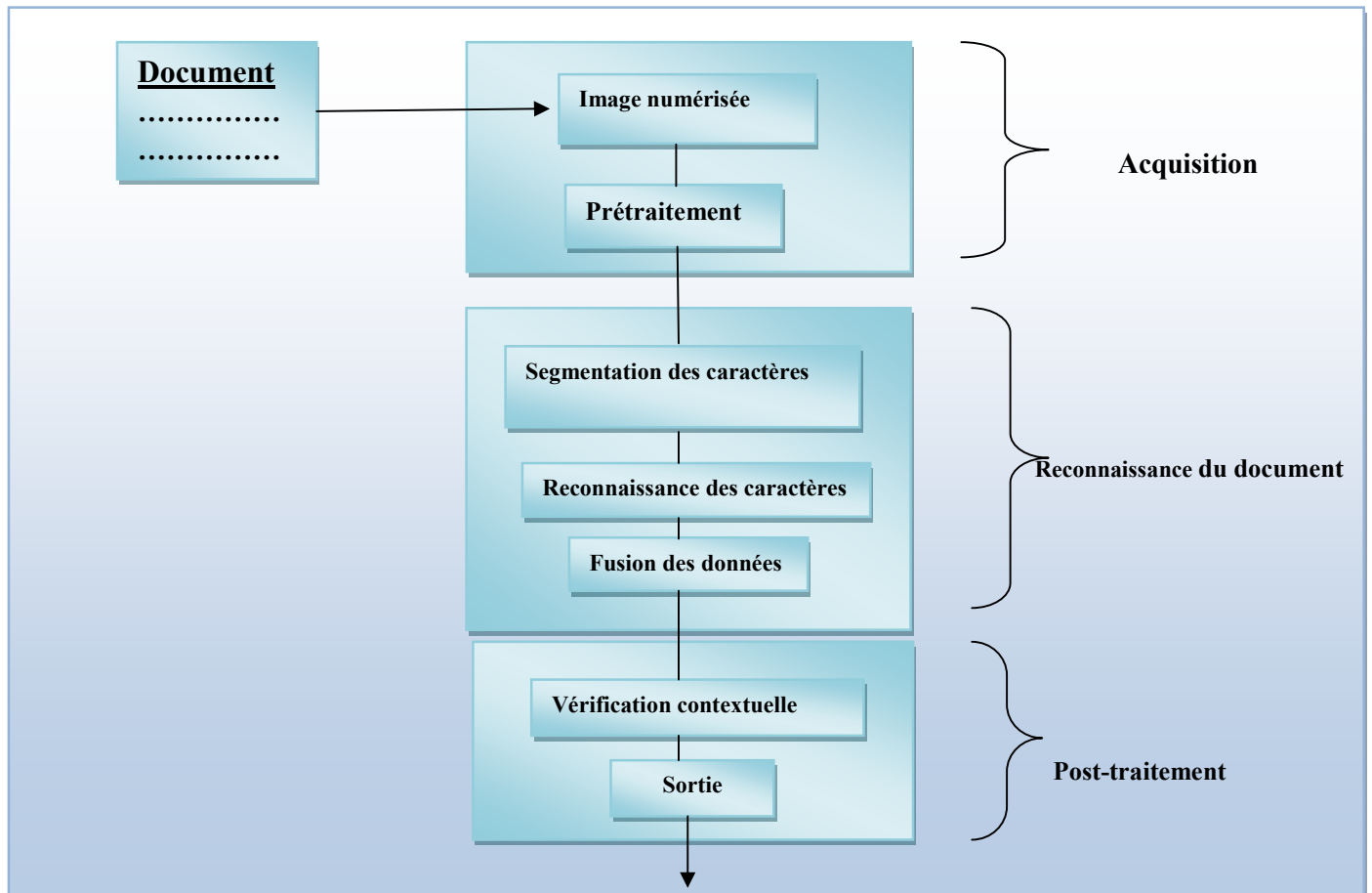


Figure 1.5 : Schéma général d'un système de reconnaissance d'écriture manuscrite [3]

Le processus OCR est une activité complexe composée de plusieurs étapes. Ces étapes sont les suivantes:

#### 1.6.1. Acquisition d'images

L'acquisition d'images est la première étape de l'OCR. Ce qui comprend l'obtention d'images numériques et leur conversion en formes appropriées faciles à manipuler par un ordinateur à partir de sources externes telles que des scanners ou des caméras.

Prétraitement : une fois l'acquisition de l'image terminée, diverses étapes de prétraitement peuvent être effectuées pour améliorer la qualité de l'image pour garantir la lisibilité et pour augmenter les chances de reconnaissance. Parmi les différentes techniques de prétraitement, il y a la réduction du bruit, le seuillage et l'extraction de la ligne de base de l'image.

### **1.6.2. Segmentation des caractères**

Dans cette étape, l'image est divisée en caractères avant d'entrer dans l'étape de classification. La segmentation peut être effectuée explicitement ou implicitement en tant que sous-produit de l'étape de classification. De plus, d'autres étapes de l'OCR peuvent aider à fournir des informations contextuelles qui facilitent la segmentation des images.

### **1.6.3. Extraction de caractéristiques**

Dans cette étape, diverses caractéristiques de personnalité sont extraites, qui définissent de manière unique la personnalité. Différents types de caractéristiques peuvent être utilisées, telles que l'image elle-même, des caractéristiques géométriques (cycles, lignes) et des caractéristiques statistiques (moments), diverses techniques telles que l'analyse en composantes principales peuvent être utilisées pour réduire les dimensions de l'image.

### **1.6.4. Classification**

Il est défini comme le processus de classification d'un personnage dans sa classe appropriée. La méthode de classification structurelle est basée sur les relations dans les composants de l'image. Les méthodes statistiques sont basées sur la classification des images à l'aide de fonctions discriminantes. Les méthodes de classification statistique comprennent la classification bayésienne, la classification par arbre de décision, la classification des réseaux neuronaux, la classification du voisin le plus proche, etc, il existe des classificateurs basés sur des méthodes syntaxiques, qui utilisent des méthodes syntaxiques pour regrouper des images à partir de sous-composants d'images.

Une fois les caractères classés, diverses méthodes peuvent être utilisées pour améliorer la précision des résultats de l'OCR. Une façon consiste à utiliser plusieurs classificateurs pour classer les images. Ce classificateur peut être utilisé pour des structures séquentielles, parallèles ou hiérarchiques. Diverses méthodes peuvent alors être utilisées pour combiner les résultats des classificateurs. Pour améliorer les résultats de l'OCR, vous pouvez également effectuer une analyse contextuelle. La géométrie de l'image et le contexte du document contribuent à réduire les risques d'erreurs. Le traitement lexical basé sur le modèle et le dictionnaire de Markov permet également d'améliorer les résultats de l'OCR.

## **1.7 Conclusion**

Dans ce chapitre, nous avons présenté le système de reconnaissance de caractères de manière générale. L'OCR n'est pas un processus atomique, il comprend l'acquisition, le prétraitement, la segmentation, l'extraction de caractéristiques, la classification et le post-traitement. Le système OCR peut également être utilisé dans diverses applications pratiques, telles que la

reconnaissance de plaques d'immatriculation, les bibliothèques intelligentes et de nombreuses autres applications.



## **Chapitre II**

### **Réseaux de neurones convolutifs**

## 2.1 Introduction

Dans cette chapitre, nous allons focaliser sur un des algorithmes les plus performants du Deep Learning, Réseaux de neurones convolutifs, ce sont des modèles de programmation puissants permettant notamment la reconnaissance d'images . Nous allons examiner ses couches et les processus qui s'y produisent afin de comprendre son principe de fonctionnement .

## 2. 2 Réseau de neurones artificiels (RNA)

Un réseau de neurones artificiels est une copie du cerveau humain basée sur un modèle biologique, elle est un système de neurones virtuels créés par des programmes informatiques pour imiter le fonctionnement des neurones biologiques. S'adapte aux ordinateurs équipés de fonctions d'intelligence artificielle.

Le réseau de neurones artificiels est basé sur plusieurs processeurs fonctionnant en parallèle. Ces processeurs sont organisés en couches. La première couche a pour fonction de recevoir les entrées de données brutes. Chacun des couches reçoit ensuite les sorties d'informations transmises par le couche précédent. La dernière couche est chargée de produire les résultats **du système**. Plus le problème est complexe, plus il faut de couches pour le traiter.

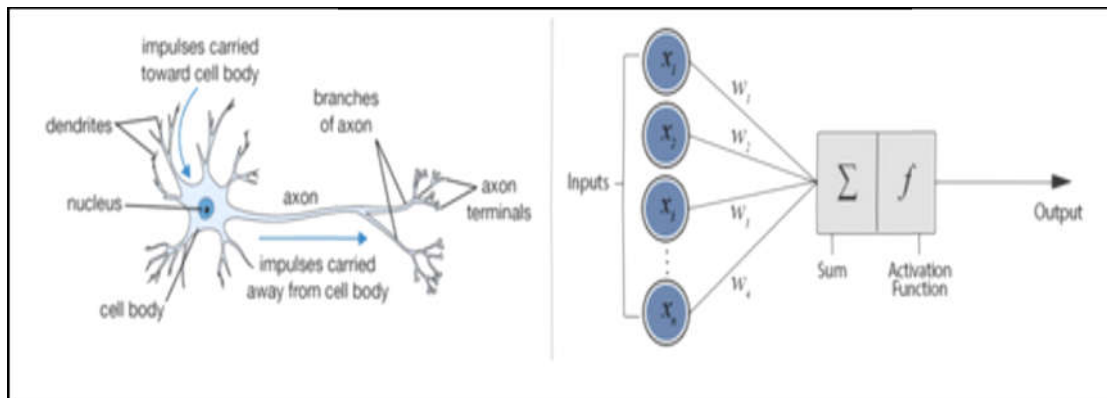


Figure 2.1 : un neurone biologique VS réseau de neurones artificiels

## 2.3 Types de réseaux de neurones artificiels

Les réseaux de neurones ont connu une croissance importante, due au développement conjoint de l'électronique numérique qui a fourni plus de puissance de calcul pour le développement des réseaux de neurones artificiels dans les années 1980. Ce développement a abouti à de nombreux types de réseaux de neurones artificiels Parmi eux:

### 2.3. 1. Réseaux de neurones récurrents

Est un réseau capable de sauvegarder les résultats obtenus après le passage des informations par les nœuds de traitement. Le modèle est progressivement alimenté et formé à partir des

résultats enregistrés. Les informations peuvent circuler dans la boucle de rétroaction et revenir à la couche précédente. De cette façon, une mémoire est constituée à l'intérieur du système.

### 2.3. 2. Réseaux de neurones à propagation avant (feed-forward)

Groupe de neurones disposés en couches composées principalement d'une couche d'entrée, d'une couche de sortie et d'une ou plusieurs couches cachées. Les informations sont transmises dans une seule direction passent directement de l'entrée aux nœuds de traitement, De là, ils sont acheminés directement aux sorties.

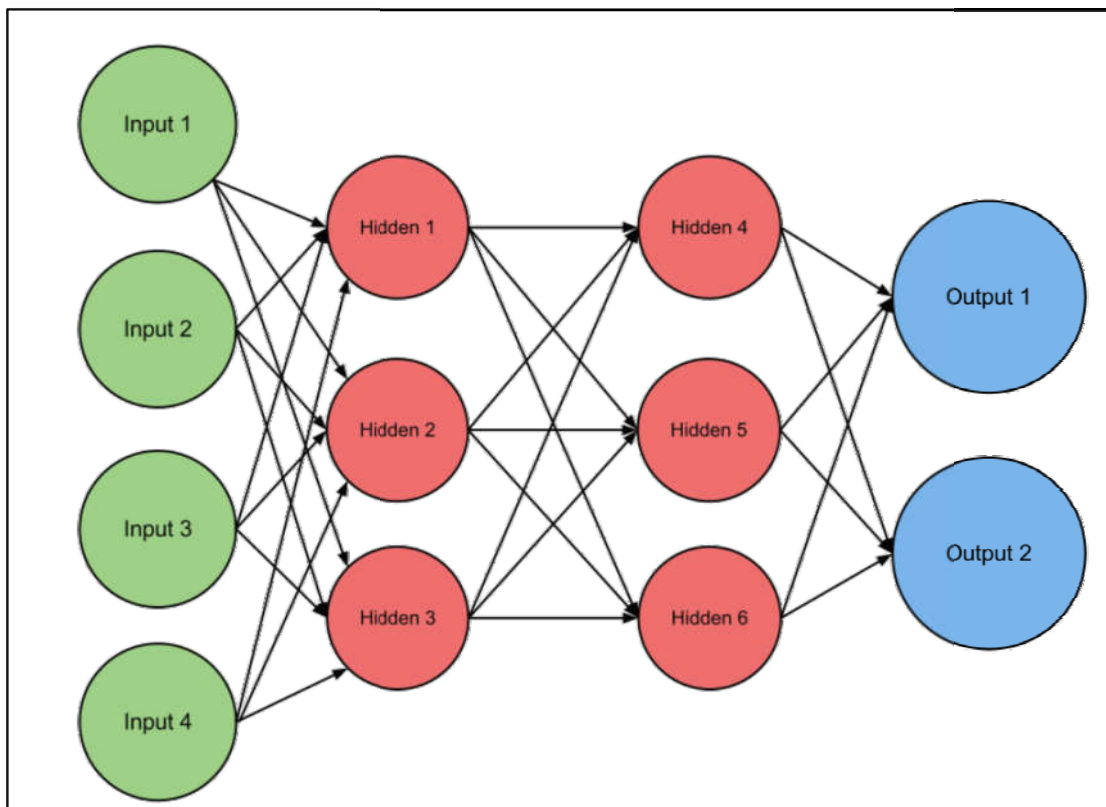


Figure 2.2 : Réseaux de neurones à propagation avant (feed-forward)

Il en existe certains types réseaux de neurones à propagation avant, notamment:

#### 2.3. 2. 1. Perceptron

Perceptron est un neurone formel muni d'une règle d'apprentissage qui permet de déterminer automatiquement les poids synaptiques; il est le plus petit réseau de neurones possible, dont la fonction d'activation est une fonction échelon également appelé Linear Threshold Function, ce qui fait du perceptron une unité linéaire à seuil (ou Linear Threshold Unit" LTU").

Frank Rosenblatt a inventé le Perceptron en 1957 au Cornell Flight Laboratory. Dans cette première version, le capteur était alors monocouche et ne disposait que d'une seule sortie à laquelle toutes les entrées étaient connectées.



Le perceptron a en entrées des nombres quelconques (contrairement aux neurones formels de McCulloch et Pitts qui ont des entrées binaires) et chaque entrée est pondérée par un poids ( $w$ ). Il possède un algorithme d'apprentissage supervisé de classifieurs binaires lui permettant sur un ensemble de données avec des étiquettes d'apprendre les poids des entrées.

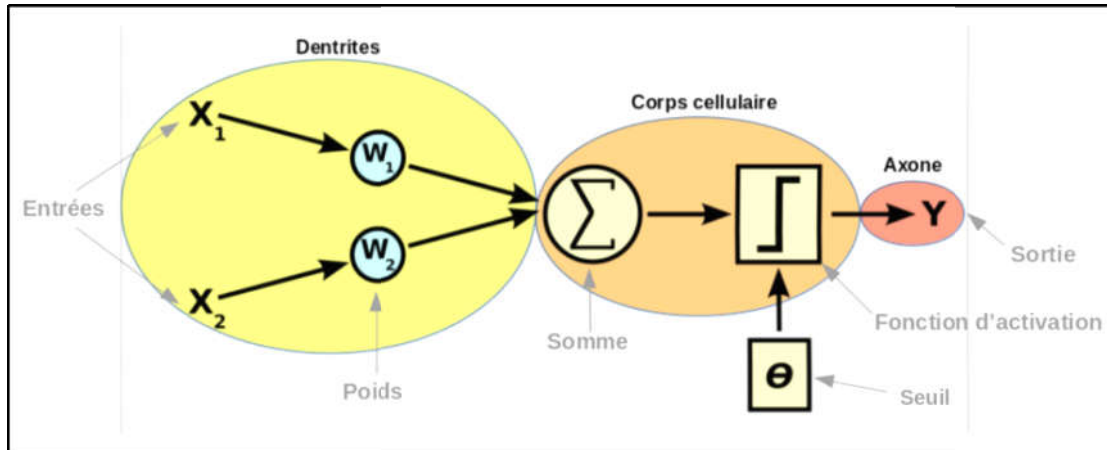


Figure 2.3 : Schéma d'une LTU

Le concept de perception est basé sur la classification des données qu'il reçoit en entrée en deux groupes (0 ou 1), de sorte que ces données soient linéairement séparables, c'est-à-dire qu'elles puissent être séparées en deux groupes.

Certaines modifications ont été ajoutées au perceptron pour séparer les données en plus de deux groupes afin de produire ce qu'on appelle un perceptron multi couches.

### 2.3. 2. 2. Perceptron multi couches (Multi Layers Perceptron(MLP))

MLP est un réseau neuronal artificiel type de feed-forward organisé en plusieurs couches, chaque couche est constituée d'un nombre variable de neurone, la première couche est reliée aux entrées, puis ensuite chaque couche est reliée à la couche précédente. C'est la dernière couche qui produit les sorties du MLP. Les sorties des autres couches ne sont pas visibles en dehors du réseau, et elles sont appelées pour cette raison couches cachées.

Ils visent à classer des données plus complexes que celles classifiées par perception Pour cela, MLP surveille chacune des données dont il dispose et met à jour chaque poids de chaque neurone dans chaque couche de son réseau afin de mieux classer cette base de données.il utilise un algorithme appelé la rétropropagation du gradient pour mettre à jour les poids.

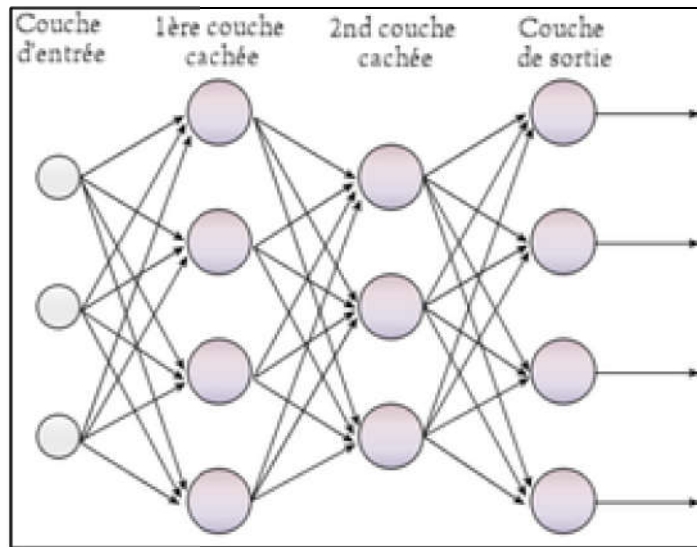


Figure 2.4 : Perceptron multicouche

### a. Rétropropagation du gradient

La rétropropagation du gradient La mise à l'échelle inverse est une méthode d'entraînement d'un réseau de neurones, basée sur la mise à jour des poids de chaque neurone de la dernière couche à la première, afin de corriger les erreurs en fonction de l'importance de la contribution de chaque élément à celle-ci. Les poids synaptiques qui contribuent plus à une erreur seront modifiés de manière plus importante que les poids qui provoquent une erreur marginale.

Bien que leur efficacité dans le traitement d'image, les MLP ont du mal à traiter de grandes images, en raison de la croissance exponentielle du nombre de connexions avec la taille de l'image, car chaque neurone est "entièrement connecté" à la fois au neurone précédent et au neurone suivant. Pour cela, il a fallu apporter quelques modifications à la structure MLP pour résoudre ce problème.

### 2.4. Réseaux de neurones convolutifs (ConvNet /CNN)

CNN est un autre type de feed-forward qui s'inspire des processus biologiques qui se produisent dans le lobe visuel, en particulier dans le cerveau des organismes vivants. Il s'agit d'un algorithme d'apprentissage en profondeur qui peut prendre une image d'entrée, attribuer une importance (poids et biais apprenables) à divers aspects ou objets d'une image et être capable de les distinguer les uns des autres.

### 2.5. Histoire

CNN a été développé et utilisé pour la première fois dans les années 1980. Le mieux que CNN pouvait faire à cette époque était de reconnaître les numéros manuscrits. Il était

principalement utilisé dans les secteurs postaux pour lire les codes postaux, les codes personnels, etc.

Mais tout modèle d'apprentissage en profondeur nécessite une grande quantité de données pour s'entraîner et nécessite également beaucoup de ressources de calcul. C'était un inconvénient majeur des CNN à l'époque, et ils étaient donc limités aux seuls secteurs postaux et ils n'ont pas réussi à entrer dans le domaine de l'apprentissage automatique.

En 2012, Alex Krizhevsky a réalisé qu'il était temps de ramener la branche de l'apprentissage en profondeur qui utilise des réseaux de neurones multicouches. La disponibilité de grands ensembles de données et d'ensembles de données ImageNet plus spécifiques avec des millions d'images classifiées et une abondance de ressources informatiques ont permis aux chercheurs de donner vie aux CNN.

## 2.6. Construction du réseau de neurones

Contrairement au classique modèle MLP, qui ne contient que la partie classification, CNN contient d'une partie convolutive, Il se compose donc de deux parties bien distinctes:

### 2.6.1. Partie convolutive

Sa tâche est d'extraire les caractéristiques de chaque image en la compressant pour réduire sa taille initiale, afin que l'image d'entrée passe à travers une série de filtres, pour créer de nouvelles images appelées cartes de convolution, qui est compilé dans un vecteur de caractéristiques appelé code CNN.

### 2.6.2. Partie classification

Le code CNN qui est généré par la partie convolutive est considéré comme une entrée dans cette partie. ce dernier est un MLP, son rôle est de combiner les caractéristiques du code CNN pour classer l'image.

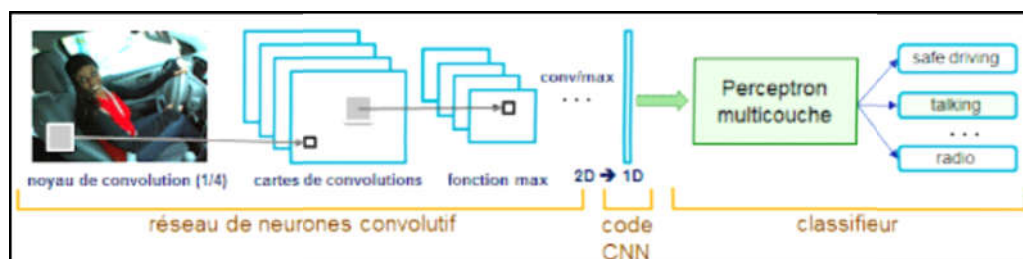


Figure 2.5 : Schéma représentant l'architecture d'un CNN

## 2.7. Couches de CNN

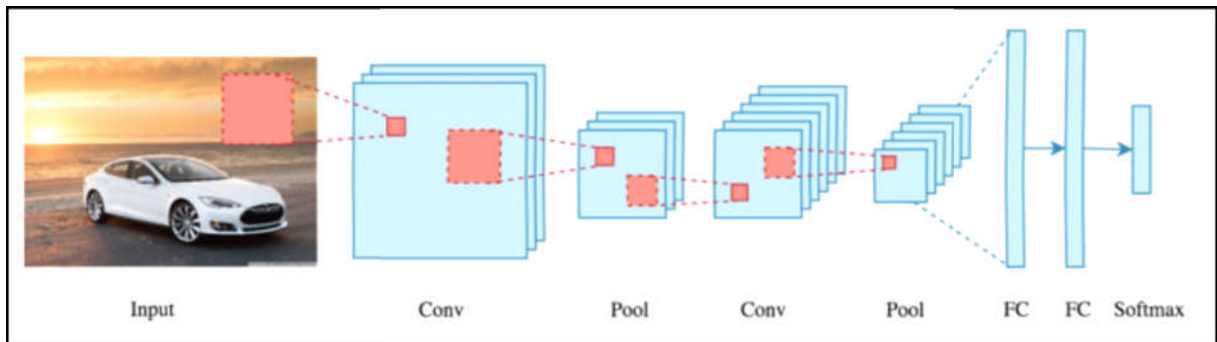


Figure 2.6 : Schéma représentant les couches de CNN

### 2.7.1. Couche de convolution

#### 2.7.1.1. Convolution

C'est une opération arithmétique simple généralement utilisée pour le traitement et la reconnaissance d'images. Ce qui explique pourquoi les réseaux de neurones convolutifs sont particulièrement adaptés à la reconnaissance d'images. Cela se fait par filtrage (La convolution agit comme un filtre).

#### 2.7.1.2. Couche convolutif

C'est la première couche de CNN et son composant principal. Le rôle de cette couche est d'analyser les images fournies en entrée et de découvrir un ensemble de caractéristiques (features). Pour ce faire, le filtrage est effectué par convolution et à l'aide de filtres. Le principe de fonctionnement des filtres est de "glisser" une fenêtre représentant la feature sur l'image, en calculant le produit de la convolution entre la feature et chaque partie de l'image numérisée. Une feature est alors vue comme un filtre. On obtient en sortie de cette couche un ensemble de features maps. [4]

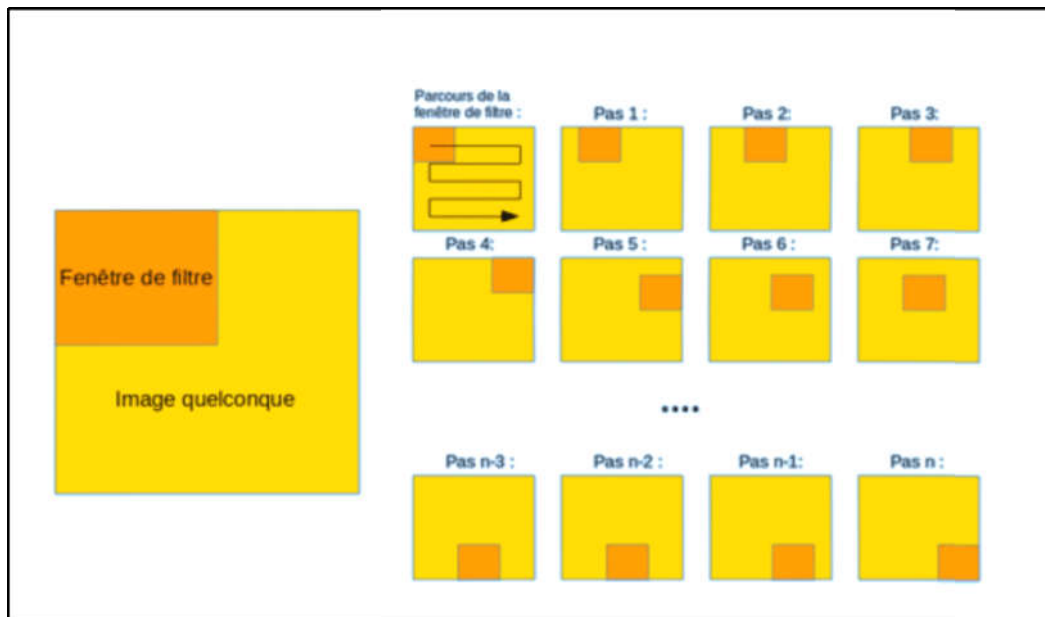


Figure 2.7 : Schéma du parcours de la fenêtre de filtre sur l'image [4]

### 2.7.2. Couche de pooling

Ce type de couche est souvent placé entre deux couches de convolution : elle reçoit en entrée plusieurs feature maps, et applique à chacune d'entre elles l'opération de pooling.

Son rôle est de réduire la taille des images, tout en préservant leurs propriétés importantes, afin de réduire la puissance de calcul nécessaire au traitement des données en réduisant les dimensions

Il existe deux types :

#### 2.7.2.1. Max pooling (subsampling)

Chaque opération de pooling sélectionne la valeur maximale du pixel

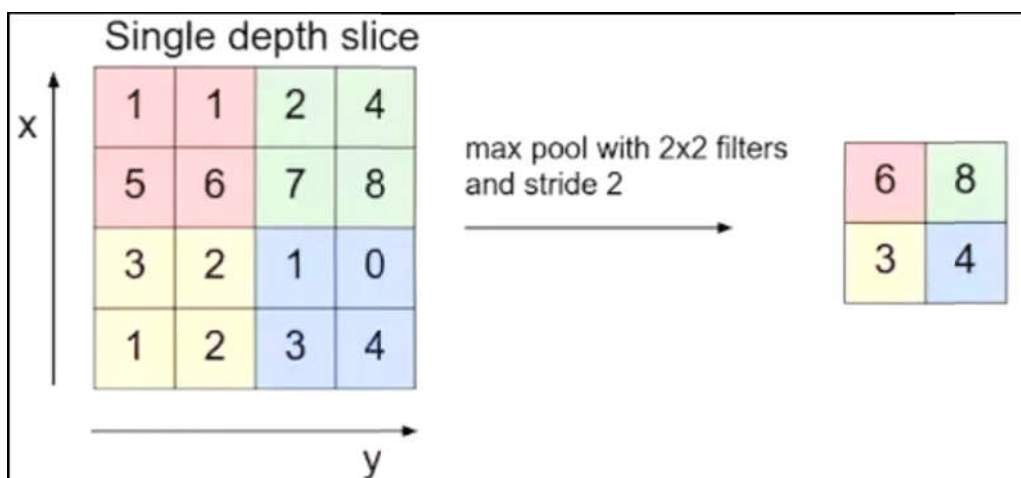


Figure 2.8: Processus de Max-Pooling [4]

### 2.7.2.2. Average pooling

Chaque opération de pooling sélectionne la valeur moyenne du pixel.

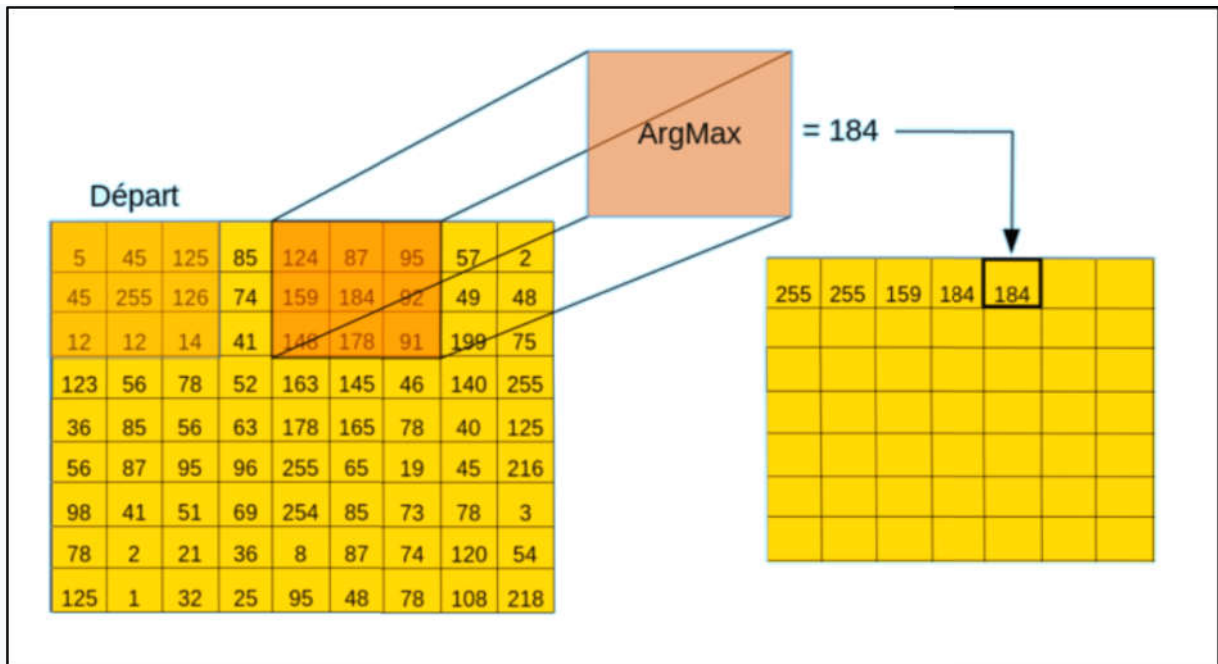


Figure 2.9: Processus Average pooling [4]

Mais la performance de max pooling est bien meilleures que average pooling, ce qui le rend plus utilisable.

### 2.7. 3. Couche fully-connected

C'est la dernière couche du réseau de neurones, elle est utilisée pour classer l'image d'entrée du CNN en déterminant le lien entre la position des features dans l'image et une classe.

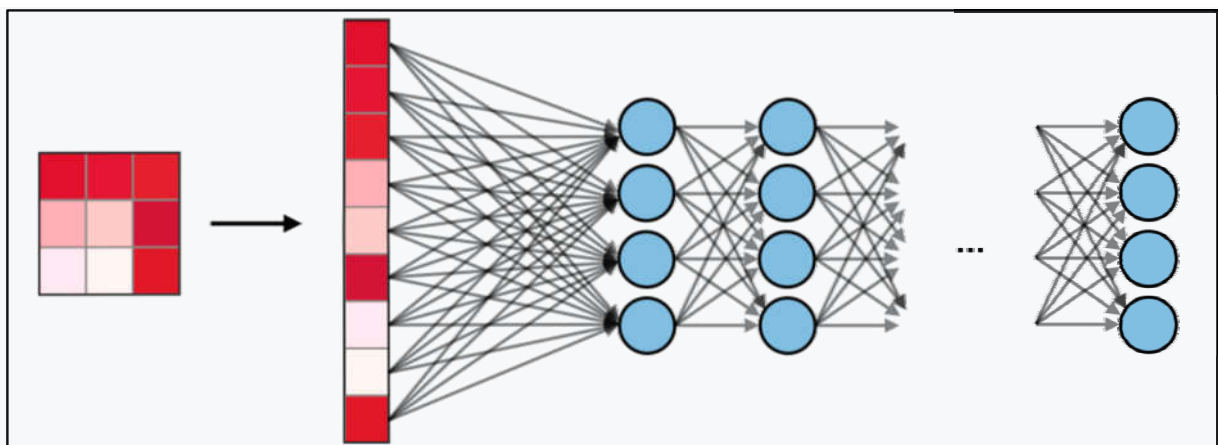


Figure 2.10 : Schéma représentant couche fully-connected

## 2.8 Fonctions d'activation

C'est une fonction mathématique qui s'applique à un signal en sortie d'un neurone artificiel, Son but est de déterminer comment la somme des signaux d'entrée est activée.

Il existe de nombreux types de fonctions d'activation.

Chaque couche de CNN crée plusieurs fonctions d'activation qui sont transmises à la couche suivante.










Nom de la fonction	Relation d'entrée/sortie	Icône
seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$	
seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$	
linéaire	$a = n$	
linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$	
linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$	
linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$	
sigmoïde	$a = \frac{1}{1+\exp^{-n}}$	
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	
compétitive	$a = 1$ si $n$ maximum $a = 0$ autrement	

Tableau 1.1 : Fonctions d'activation

### 2.7.1. Fonction ReLU (Rectified Linear Units)

Désigne la fonction réelle non-linéaire .Lorsque la valeur d'entrée est égale ou inférieure à 0, elle devient 0, et lorsqu'elle est supérieure à 1, l'entrée est sortie telle quelle.

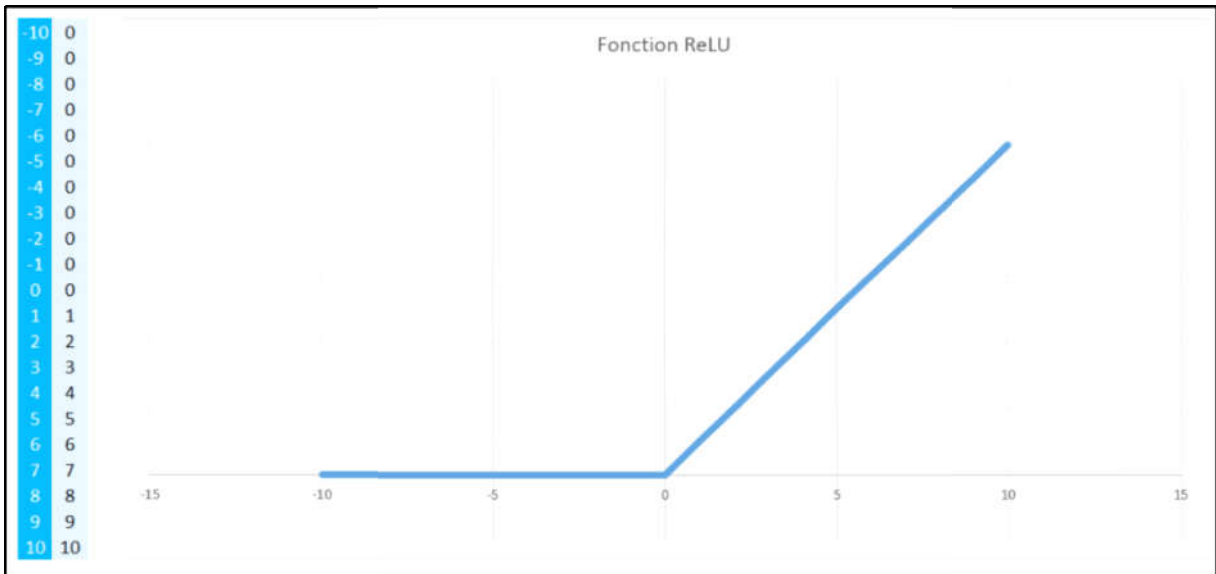


Figure 2.11 : Fonction ReLU

### 2.7.2. Fonction Softmax

Une fonction utilisée dans la couche de sortie. Couramment utilisé dans les problèmes de classification.

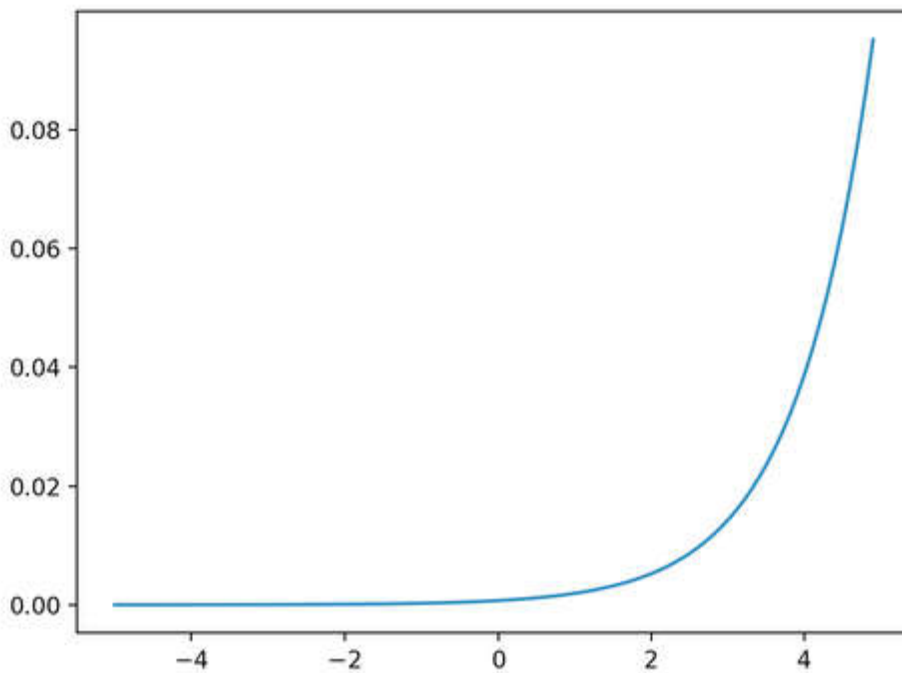


Figure 2.12 : Fonction Softmax

## 2.9. Conceptions d'architecture CNN

Il existe de nombreuses architectures disponibles pour les CNN qui ont joué un rôle déterminant dans la construction des algorithmes qui alimenteront l'IA dans son ensemble dans un avenir prévisible. Les plus courantes sont:



- LeNet
- AlexNet
- VGGNet
- GoogLeNet
- ResNet
- ZFNet
- ImageNet
- DenseNet
- SENet

Mais Le-Net 5 est le modèle le plus populaire et le plus pratique.

### **1.10. Conclusion**

Dans ce chapitre, nous avons abordé le concept de CNN après avoir introduit quelques concepts de réseaux neuronaux artificiels pour aider à comprendre le principe de fonctionnement de CNN, en mettant l'accent sur ses différentes couches. À ce jour, les réseaux de neurones convolutifs restent le modèle le plus efficace pour classer les images.

## **Chapitre III**

### **Résultats expérimentaux et Discussions**

### 3.1. Introduction

Dans ce chapitre nous présentons les résultats obtenus pour la validation de notre système de reconnaissance des chiffres manuscrits par réseau de neurones convolutifs mais tous d'abord nous présentons la base de données MNIST .Ensuit, nous fournissons une description détaillée de chaque couche de CNN, à la fin nous présentons les différents résultats obtenus.

## 3.2. Ressources matérielles et logicielles

### 3.2.1. Ressources matérielles

Notre système est développé dans un ordinateur avec les caractéristiques techniques suivantes:

- Processeur : Intel(R) Core(TM) i5-3210M CPU @ 2.50GHz 2.50 GHz
- RAM :4.00 Go
- Disque dur : 500 Go

### 3.2.2. Ressources logicielles

- Windows 10 Professionnel 64bits
- Python: version Python 3.8

Nous avons développé notre système en utilisant le langage Python, il est un langage de programmation multi-paradigme. La programmation orientée objet et la programmation structurée sont entièrement prises en charge, et bon nombre de ses fonctionnalités prennent en charge la programmation fonctionnelle et la programmation orientée aspect (y compris la métaprogrammation et les méta-objets [méthodes magiques]). De nombreux autres paradigmes sont pris en charge via des extensions, notamment la conception par nœuds et la programmation logique.

La plupart des implémentations Python (y compris CPython) incluent une boucle de lecture-évaluation-impression (REPL), leur permettant de fonctionner comme un interpréteur de ligne de commande pour lequel les utilisateurs saisissent des instructions de manière séquentielle et reçoivent les résultats immédiatement.

### 3.2.3. Tensorflow & Keras

Nous avons utilisé l'implémentation Tensorflow et Keras

#### a. Tensorflow

C'est une bibliothèque open source qui utilise Python pour le calcul numérique qui rend l'apprentissage automatique plus rapide et plus facile, qui facilite l'acquisition de données et les modèles de formation, sert les prédictions et aide à améliorer les résultats futurs. Il a été créé par l'équipe Google Brain. Il peut former et exécuter des réseaux de neurones profonds

pour la classification de l'écriture manuscrite, la reconnaissance d'images, l'intégration de mots, les réseaux de neurones récurrents et les modèles de chaînes pour la traduction automatique, le traitement du langage naturel et les PDE basés sur la simulation. Mieux encore, TensorFlow prend en charge les prévisions de production à grande échelle, avec les mêmes modèles que ceux utilisés pour la formation. [5]

### **b. Keras**

Une bibliothèque de réseaux de neurones open source écrite en Python. Elle peut s'exécuter sur Tensorflow, Microsoft Cognitive Tools, R, Theano ou PlaidML. Conçue pour permettre une expérimentation rapide des réseaux de neurones profonds, elle se concentre d'être facile à utiliser, flexible et évolutif. Il a été développé dans le cadre de l'effort de recherche du projet ONEIROS (Open Neuroelectronic Intelligent Robot Operating System) et son principal créateur et mainteneur est François Cholet, un ingénieur de Google. Cholet est le créateur d'un modèle de réseau neuronal profond appelé Xception. [5]

### **3.3. Base de données MNIST**

La base de données MNIST pour Modified ou Mixed National Institute of Standards and Technology, est une base de données de chiffres écrits à la main. Il s'agit d'un test standardisé qui est un ensemble de données largement utilisé dans l'apprentissage automatique. Il contient 60 000 images d'entraînement et 10 000 images de test. Ces images sont de petites images carrées de 28 x 28 pixels en noir et blanc (niveaux de gris), de nombres simples écrits à la main entre 0 et 9.

Il est un ensemble de données largement utilisé et profondément compris. Les réseaux de neurones convolutifs d'apprentissage en profondeur sont considérés comme les modèles les plus performants car ils atteignent une précision de classification de plus de 99 %, avec un taux d'erreur compris entre 0,4 % et 0,2 % dans l'ensemble de données de test.



Figure 3.1 : Exemple MNIST

### 3.4. Développer notre système de reconnaissance

Nous avons développé notre système en utilisant cinq composants principaux:

Nous avons téléchargé la base de données, préparé la base de données, défini le modèle, évalué le modèle et affiché les résultats.

#### 3.4.1 Télécharger la base de données

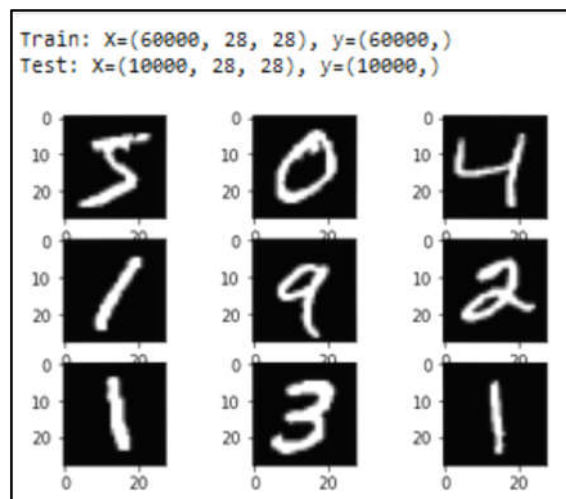


Figure 3.2 : Graphique des neuf premières images de l'ensemble de données

Nous pouvons voir qu'il y a 60 000 exemples dans l'ensemble de données d'apprentissage et 10 000 dans l'ensemble de données de test et que les images sont de taille  $28 \times 28$  pixels. Il montre également la nature manuscrite naturelle des images à classer.

### 3.4.2. Préparation des données de pixels

Les valeurs de pixel de chaque image de l'ensemble de données sont des entiers non signés dans la plage noir et blanc comprise entre 0 et 255.

Nous avons effectué une opération de mise à l'échelle afin de normaliser les valeurs de pixel des images en niveaux de gris en les redimensionnant dans la plage [0,1].

### 3.4.3. Définition du modèle

Le modèle a deux aspects principaux : le front-end pour extraire l'attribut compréhensible des couches convolutifs et pooling, et le back-end du classifieur qui fera la prédiction.

#### a. Partie convolutive

Pour le frontal convolutif, nous avons commencé avec une couche convolutif avec une petite taille de filtre (3.3) et un nombre modeste de filtres (32) suivie d'une couche de maxpooling.

Les cartes de filtre sont ensuite aplaties en appliquant un processus d'aplatissement pour fournir des caractéristiques permettant au classificateur de prédire les probabilités de sortie.

#### b. Partie de classifieur

Étant donné que le problème est une tâche de classification multi-classes, nous aurons besoin d'une couche de sortie avec 10 nœuds afin de prédire la distribution de probabilité d'une image appartenant à chacune des 10 classes. Cela nécessitera également l'utilisation de la fonction d'activation softmax. Entre l'extracteur de caractéristiques et la couche de sortie, nous pouvons ajouter une couche dense pour interpréter les caractéristiques à 100 nœuds.

Toutes les couches utilisent la fonction d'activation ReLU.

### 3.4.4. Évaluation du modèle

Le modèle est évalué à l'aide d'une validation croisée quintuple.

```
# evaluate a model using k-fold cross-validation
def evaluate_model(dataX, dataY, n_folds=5):
```

La valeur de  $k = 5$  a été choisie pour fournir une ligne de base pour une évaluation répétée et ne pas être trop grande pour nécessiter une longue durée d'exécution.

Chaque ensemble de test représentera 20 % de l'ensemble de données d'apprentissage, soit environ 12 000 exemples, ce qui est proche de la taille de l'ensemble de test réel pour ce problème.

## 3.5 Résultats

Il y a deux aspects principaux à présenter : Diagnostiquer le comportement d'apprentissage du modèle pendant l'entraînement et estimer les performances du modèle.

### a. Diagnostiquer le comportement d'apprentissage du modèle pendant l'entraînement

Les diagnostics impliquent la création d'un graphique linéaire montrant les performances du modèle sur l'ensemble d'apprentissage et l'ensemble de test lors de chaque pli de la validation croisée k-fold.

Nous avons créé une forme à l'aide de sous-parcelles, une pour la perte et une pour la précision. Les lignes bleues montrent les performances du modèle dans l'ensemble de données d'apprentissage tandis que les lignes oranges montrent les performances dans l'ensemble de données de test.

### b. Estimer les performances du modèle

Les scores de précision de la classification collectés lors de chaque pli peuvent être résumés en calculant la moyenne et l'écart type. Cela fournit une estimation de la performance moyenne attendue du modèle formé sur cet ensemble de données, avec la variance moyenne estimée en moyenne.

Les résultats imprimés après l'exécution du programme sont présentés dans les figures suivantes :

```
> 98.658
> 98.592
> 98.542
> 98.892
> 98.800
```

```
Accuracy: mean=98.697 std=0.131, n=5
```

Figure 3.3 : Résolution de classification pour chaque pli de la validation croisée

Nous notons que le modèle a atteint une compétence importante, avec une précision de 98,6 %, ce qui est raisonnable.

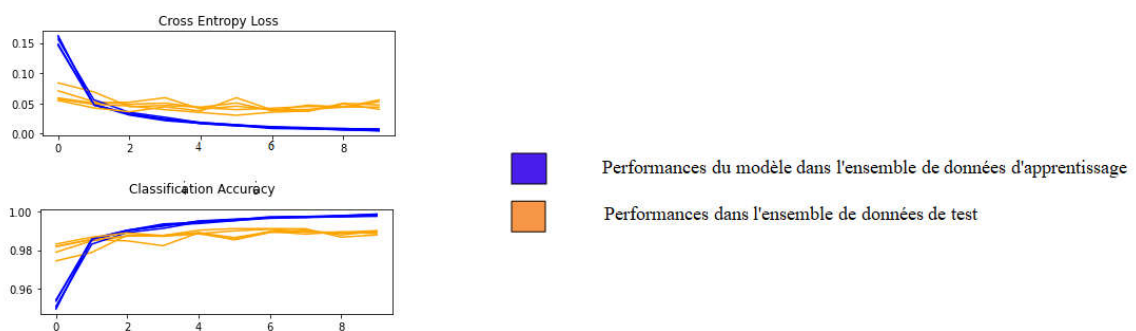


Figure 3.4 : graphique de diagnostic des performances du basic model

Nous pouvons voir que le modèle atteint généralement un bon ajustement, les courbes d'apprentissage et de test convergeant.

### 3.5.1 Améliorations des performances du modèle

Il existe de nombreuses façons d'explorer des améliorations au modèle de base. Nous examinerons les domaines de la formation de modèles qui conduisent souvent à l'amélioration,

### 3.5.2 Résultats expérimentaux

#### a. Expérimentation 1

Improved model1: modification de la capacité de la partie extraction de caractéristiques du modèle.

Nous avons augmenté dans la partie extracteur de caractéristiques de Basic modèle (Partie convolutive), en ajoutant plus de couches de convolution et max pooling de même taille, tout en augmentant le nombre de filtres.

Dans cette expérimentation, nous avons ajouté une double couche convolutive avec 64 filtres chacun, suivie d'une autre couche de max pooling. Et nous avons comparé Improved model1 avec Basic modèle.

Les résultats imprimés après l'exécution de programme:

```
> 98.958
> 99.017
> 98.883
> 98.917
> 99.183
```

```
Accuracy: mean=98.992 std=0.106, n=5
```

Figure 3.5 : Résolution de classification pour chaque pli de la validation croisée



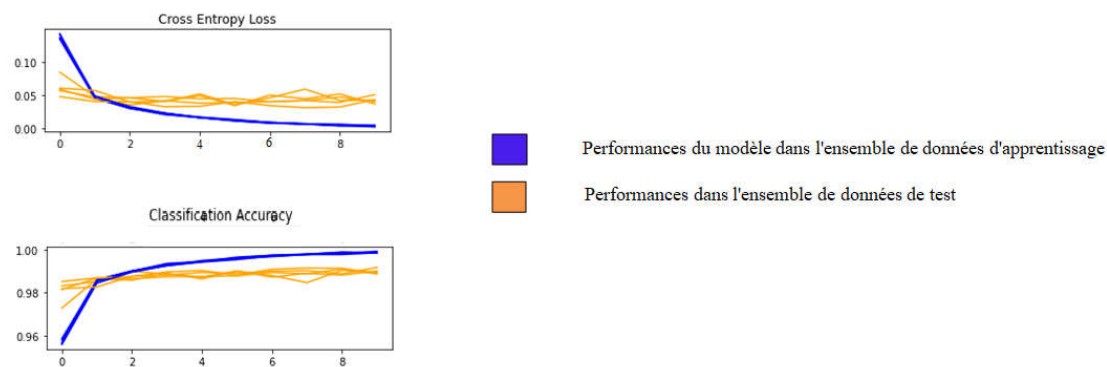


Figure 3.6 : graphique de diagnostic des performances du Improved model1

Les résultats de comparaison sont présentés dans le tableau ci-dessous :

Modèle	Acruacy	Moyenne	Std (standard deviation)
Basic model		98.697	0.131
Improved model1		98.992	0.106

Tableau 3.2 : Résultats de comparaison

Nous constatons à travers les résultats que le model atteint une compétence idéale estimée à 98,99%. Ce sont de bons résultants, et Improved model1 a bonne performance.

## b. Expérimentation 2

Improved model2: modification de la fonctionnalité de la partie classifieur du modèle.

Dans cette expérimentation, nous avons augmenté dans la partie classifieur de Basic modèle, nous avons ajouté une double couche dense avec 200 nœud chacun.

Les résultats imprimés après l'exécution de programme:

```
> 98.750
> 98.650
> 98.683
> 98.825
> 98.850
```

```
Accuracy: mean=98.752 std=0.078, n=5
```

Figure 3.7 : Résolution de classification pour chaque pli de la validation croisée

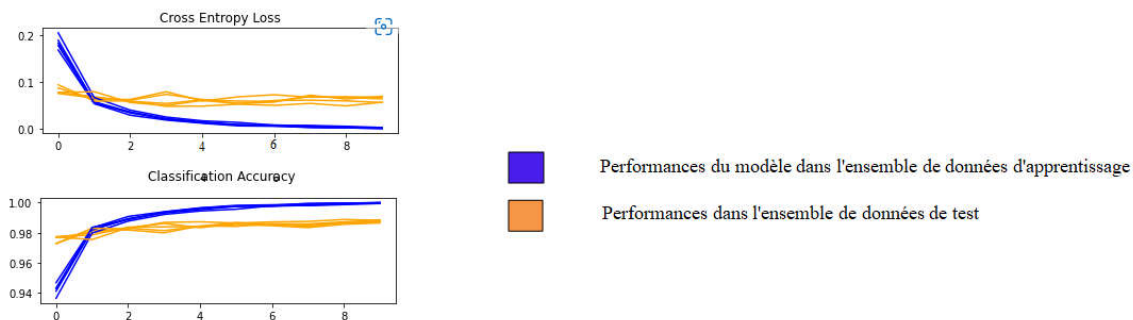


Figure 3.8 : graphique de diagnostic des performances du Improved model2

Les résultats de comparaison obtenus sont présentés dans le tableau ci-dessous

Modèle	Acuracy	Moyenne	Std (standard deviation)
Basic model		98.697	0.131
Improved model2		98.752	0.078

Tableau 3.3 : Résultat de comparaison

Nous notons dans le tableau ci-dessus que les performances du modèle se sont améliorées, quoique légèrement, mais de manière significative.

### c. Expérimentation 3

Improved model3: modification dans les deux parties en même temps.

Dans cette expérimentation, nous avons modifié les deux parties du basic model en même temps et nous avons comparé le basic model avec Improved model3.

Les résultats imprimés après l'exécution de programme:

```
> 99.008
> 98.983
> 98.850
> 98.792
> 98.925
```

```
Accuracy: mean=98.912 std=0.081, n=5
```

Figure 3.9 : Résolution de classification pour chaque pli de la validation croisée

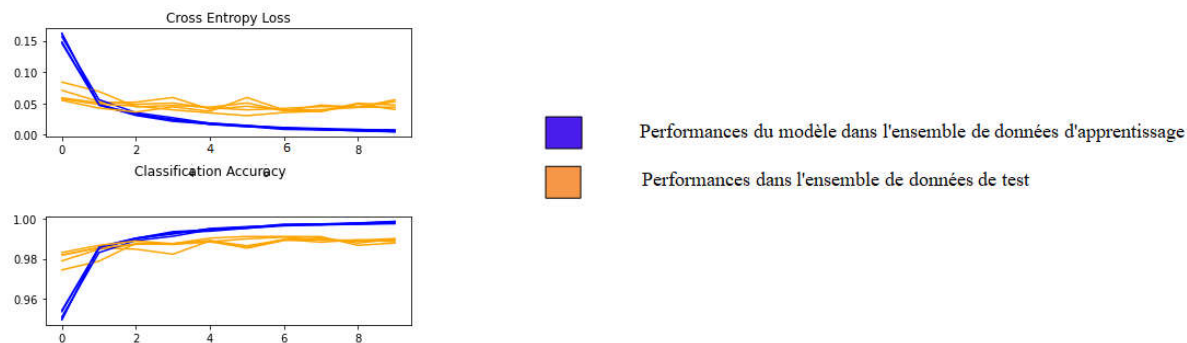


Figure 3.10 : graphique de diagnostic des performances du Improved model3

Les résultats de comparaison obtenus sont présentés dans le tableau ci-dessous

Modèle	Accuracy	Std (standard deviation)	
	Moyenne		
Basic model	98.697	0.131	
Improved model3	98.912	0.081	

Tableau 3.4 : Résultat de comparaison

Nous remarquons des résultats que le modèle a atteint une efficacité idéale de 98,91 %. Ce sont de bons résultats qui indiquent une amélioration des performances du modèle.

### 3.5.3. Résumé des expérimentations

A partir de trios expérimentations, nous concluons que la modification de la capacité de la partie extraction de caractéristiques du modèle (partie convolutive) est le point le plus influent dans l'amélioration des performances du modèle qui correspond à Improved model1 de notre programme.

### 3.6 Conclusion

Dans ce chapitre, nous fournissons une description détaillée du système de reconnaissance manuscrite des chiffres par CNN, en partant de la définition de ses différentes couches et de l'ensemble des opérations dans lesquelles elles se produisent, en passant par l'évaluation du système afin d'améliorer ses performances et de prédiction finale.

### **Conclusion générale**

Le réseau de neurones a plusieurs caractéristiques, y compris la taille de l'ensemble d'apprentissage fixe, les poids articulaires, la stabilité de la traduction, etc. Toutes sont considérées comme des qualifications qui l'ont aidé à être le premier candidat à la reconnaissance d'images en général, donc l'objectif de notre projet consistait à concevoir et développer un système de reconnaissance des chiffres manuscrits à l'aide d'un réseau de neurones convolutifs.

Afin de bien implémenter notre projet, nous avons mené une étude générale sur le système OCR puis mené une étude approfondie sur le CNN. Pour comprendre ce qui se passe au niveau de chacune de ses couches, où les filtres extraient les caractéristiques au niveau de la partie convolutive pour être combinés et catégorisés au niveau de la partie classifieur.

Les résultats de cette recherche et de cette étude pour la conception et le développement d'un système de reconnaissance des chiffres manuscrits ont été très satisfaisants, et nous en concluons que CNN est une proposition réussie pour le développement de ce type de système.

## **Bibliographie**

[1] Baca, A ;Fillali ,A «Traitment et Reconnaissance des caractères ».Mémoire DE MASTER, Univercité M'hamed Bougara,2016.

[2] Emilie, P ; Christian, G ; « Réseau de neurone a convolution : reconnaissance de l'écriture manuscrite non certaine », Ecole polytechnique de l'université de Nantes, France,2014.

[3] <https://www.natural-solutions.eu/blog/la-reconnaissance-dimage-avec-les-rseaux-de-neurones-convolutifs>

[4] A. Belaïd et H. Cecotti , « La numérisation de documents : Principe et évaluation des performances », Université Nancy 2 – LORIA

[5] <https://keras.io/about/>