

الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي و البحث العلمي
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
جامعة عمار ثلجي بالأغواط
UNIVERSITE AMAR TELIDJI LAGHOUAT



كلية العلوم
FACULTE DES SCIENCES
DEPARTEMENT D'INFORMATIQUE
Mémoire de Master

Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Système d'information décisionnel

Réalisé Par :

Amira Belhanafi

Nadjat Benabderrahmane

Thème

**La reconnaissance de l'écriture manuscrite dans les
prescriptions médicales basées sur des techniques
d'apprentissage profond**

Soutenue publiquement le 21-06-2023 devant les membres de jury composé de :

Mme. Fatima Zohra Bousbaa	M.C.(A)	Pésidente
Mr. Laradj Chellama	M.A.(A)	Examineur
Mr. Younes Guellouma	M.C.(A)	Examineur
Mr. Mustapha Bouakkaz	Professeur	Encadreur
Mlle. Saida Sarra Boudouh	Doctorante	Co-encadreur

Année universitaire 2022/2023

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dédicaces

Je dédie ce mémoire à ...

Mes chers parents et à celle qui m'a mis sur le chemin de la vie, la femme de ma vie, ma mère **Farida Belhanafi** qui m'a accompagné toujours et pour toujours et à chaque pas. Aussi pour mon hero père **Mohamed Belhanafi**, J'espère de tout mon cœur m'élever au niveau où je rêvais de me voir.

À mes frères, les étoiles de mon ciel **Rida, Rahma, AbdelNour, Rachid.**

Je le dédie spécialement à mon amie et compagne **Nadjat Benabderrahmane**, qui s'est toujours distinguée par sa positivité et son travail acharné. Je souhaite également que

Dieu vous bénisse avec succès et bonheur dans tous les domaines de la vie.

À moi-même, propriétaire d'un effort et d'un travail continu, et d'une position ferme malgré tout, je la remercie beaucoup pour sa continuation. À tous ceux qui ont contribué même un peu au développement de ce travail. ...

Amira Belhanafi

Dédicaces

Je dédie ce mémoire . . .

Pour l'âme de mon cher père, mon héro **Benharzallah** qui ma appris le sens du succès, pour ses encouragements constants qu'il ma donné, que la mort ne lui a laissé pas le temps pour voir la succès du sa fille, que Dieu lui fasse miséricorde.

Pour ma mère forte **Mamma Benabderrahmane** qui a joué le rôle de mère et de père dans ma vie, pour sa patience et sa confiance en moi pendant mes études, vous êtes la raison de ma réussite.

À mes chères soeurs **Samira, Souad, Soumia, Hadjer, Zahira** et **Djihad** pour leur amour et leur soutien pour moi.

À mon petit frère **Mohammed Yacine**, et bien sur à mon oncle **Hamid Benabderrahmane** qui était comme mon grand frère.

Mes dédicaces sont également adressées a mon binôme et ma chère amie **Amira**, c'était toujours un plaisir de travailler avec vous, Merci pour tous les souvenirs que nous avons passé pendant ce cycle universitaire, je vous souhaite un avenir plein de succès et que vous réalisez vos souhaits.

...

Nadjat Benabderrahmane

Remerciement

Nous tenons tout d'abord à remercier Allah le Tout-Puissant et Miséricordieux, qui nous a donné la force et la patience d'accomplir ce travail.

En second lieu, nous tenons à remercier notre encadreur *Pr* **Mustapha Bouakkaz** pour l'honneur qu'il nous a fait en acceptant de nous encadrer, ainsi que pour les conseils précieux, les orientations, et le temps qu'il nous a consacré afin de réaliser ce travail, sa disponibilité et ses orientations.

Nos vifs remerciements aussi pour les membres du jury qui ont accepté d'examiner et d'évaluer notre travail.

Nous exprimons aussi notre gratitude pour *Dr* **Fatima Zohra Bousbaa** pour tous ses encouragements et son aide dans la réalisation de ce travail.

Sans oublier *Dr* **Saida Sarra Boudouh** pour ces aides et ses conseils.

Nous tenons à exprimer nos sincères remerciements à tous les enseignants qui nous ont enseigné et qui par leurs compétences nous ont soutenu dans la poursuite de nos études. Enfin, on remercie tous ceux qui de près ou de loin, ont contribué à la réalisation de ce mémoire.

ملخص

يعانى الأشخاص بشكل عام من صعوبات كبيرة في قراءة خط الأطباء، حيث يعرفون غالبًا بكتابتهم غير المقروءة. يرجع ذلك إلى عدة أسباب، مما يجعل مهمة الصيدالفة صعبة عندما يتعلق الأمر بتوفير الدواء الصحيح للمرضى. في أسوأ الحالات، يمكن أن يؤدي ذلك حتى إلى فقدان حياة المريض وتعرض الطبيب لحقوبات قانونية. من خلال هذا العمل ، نحن مهتمون باستخدام تقنيات التعلم العميق للتعرف على النص المكتوب يدويًا في الوصفات الطبية. من خلال استغلال مجموعات بيانات مختلفة مثل مجموعة البيانات التي تم إنشاؤها بواسطة EMNIST ، ومجموعة بيانات IAM ، جنبًا إلى جنب مع تقنيات مختلفة مثل الشبكات العصبية التلافيفية (CNN) والدعم المتجه الآلي (SVM) ، قمنا ببناء نماذج متعددة أظهرت نتائج مختلفة، حيث تم تحقيق دقة قصوى تبلغ 91٪.

كلمات مفتاحية : DL ، EMNIST ، IAM ، CNN ، SVM.

Résumé

Les gens en général rencontrent de grandes difficultés pour lire l'écriture manuscrite des médecins, car ils sont généralement connus pour avoir une écriture illisible. Cela est dû à plusieurs raisons, ce qui rend la tâche des pharmaciens difficile lorsqu'il s'agit de fournir le bon médicament aux patients. Dans les pires cas, cela peut même entraîner la perte de vie du patient et exposer le médecin à des sanctions légales. À travers ce travail de master nous nous intéressons à l'utilisation des techniques de deep learning (DL) pour la reconnaissance de l'écriture manuscrite dans les prescriptions médicales, En exploitant différents dataset tels que une dataset réalisé par nous-même, EMNIST dataset et IAM dataset, ainsi que différentes techniques telles que CNN et SVM, en nous basant sur ce qui a été mentionné, construit plusieurs modèles qui ont montré des résultats différents, atteignant une accuracy maximale de 91 %.

Mots clés : DL, EMNIST dataset, IAM dataset, CNN, SVM.

Abstract

People in general encounter significant difficulties in reading doctors' handwriting as they are often known for having illegible writing. This is due to several reasons, which makes the task of pharmacists challenging when it comes to providing the correct medication to patients. In the worst cases, this can even result in the loss of a patient's life and expose the doctor to legal penalties. Through this master's work, we are interested in using deep learning (DL) techniques for recognizing handwritten text in medical prescriptions. By leveraging different datasets such as a dataset created by ourselves, the EMNIST dataset, and the IAM dataset, along with various techniques such as CNN and SVM, we have constructed multiple models that have shown different results, achieving a maximum accuracy of 91 %.

Key Words : DL, EMNIST dataset, IAM dataset, CNN, SVM.

TABLE DES MATIÈRES

Introduction générale	1
1 Écriture manuscrite	3
1.1 Introduction	3
1.2 Définition	3
1.3 Les caractéristiques de l'écriture manuscrite	4
1.4 Les types de l'écriture manuscrite :	4
1.5 Les facteurs de singularité de l'écriture manuscrite	5
1.6 Écriture manuscrite des médecins	5
1.6.1 Les conséquences nuisibles	6
1.6.2 Les raisons de la mauvaise écriture des médecins	6
1.7 Problématique	7
1.8 Solution proposée	7
1.9 Conclusion	7
2 Intelligence artificielle	9
2.1 Introduction	9
2.2 Définition	9
2.3 Test de turing	10
2.4 Les types d'intelligence artificielle	11
2.4.1 Intelligence artificielle faible (Weak AI)	11
2.4.2 Intelligence artificielle fort (Strong AI)	11
2.5 Les branches de l'intelligence artificielle	12
2.6 L'apprentissage automatique	13
2.6.1 Les données	13
2.6.2 Les types d'apprentissage automatique	13
2.7 Algorithmes d'apprentissage automatique	17

2.7.1	Machine à vecteur de support (SVM)	17
2.8	Réseaux de neurones et apprentissage profond	18
2.8.1	Réseaux de neurones	18
2.8.2	Apprentissage profond (deep learning)	26
2.9	Conclusion	27
3	État de l'art	28
3.1	Introduction	28
3.2	Solutions basées sur l'apprentissage profond	28
3.2.1	Étude 01 : Doctor's Handwritten Prescription Recognition System In Multi-Language Using Deep Learning [44]	28
3.2.2	Étude 02 : Doctor's Cursive Handwriting Recognition System Using Deep Learning [45]	30
3.2.3	Étude 03 : Medical Handwritten Prescription Recognition Using CRNN [46]	32
3.3	Solutions basées sur l'apprentissage automatique	32
3.3.1	Étude 01 : Medical Prescription Recognition using Machine Learning [47]	32
3.4	Solution pour l'écriture en général par le deep learning	34
3.5	Discussion et critique	34
3.6	Conclusion	35
4	Implémentation des algorithmes et présentation des résultats	36
4.1	Introduction	36
4.2	Logiciels et libraires utilisé dans l'implémentation	36
4.2.1	Python	36
4.2.2	Tensorflow	36
4.2.3	Keras	37
4.2.4	Spyder	37
4.2.5	Google Colaboratory	37
4.2.6	Pinterest	37
4.2.7	Adobe scan	38
4.3	Définition de datasets utilisées	38
4.3.1	Notre propre dataset	38
4.3.2	EMNIST ByClass	41
4.3.3	IAM Handwriting	41
4.4	Implementation et mise en œuvre	42
4.4.1	Modèle 1	43
4.4.2	Modèle 2	46
4.4.3	Modèle 3	47

4.4.4	Approche de modèle 1	49
4.4.5	Approche de modèle 2	49
4.5	Résultats obtenus	50
4.6	Discussion	51
4.7	Conclusion	53
	Conclusion générale	54
	Bibliography	55

TABLE DES FIGURES

1.1	La prescription a entraîné la mort du patient	6
2.1	Test de turing	11
2.2	Les branches de l'IA	12
2.3	Les types de ML	14
2.4	Exemple de l'apprentissage supervisé	14
2.5	Exemple de clustering	16
2.6	Le mécanisme de fonctionnement de L'apprentissage par renforcement	16
2.7	Classification des données à l'aide de la courbe SVM	17
2.8	Exemple de l'algorithme SVM	18
2.9	La structure d'un neurone biologique	19
2.10	une modèle de perceptron	19
2.11	Fonction d'activation sigmoïde	21
2.12	Fonction d'activation ReLU	21
2.13	Réseaux de neurones à propagation avant	22
2.14	Représentation d'une image en forme de matrice des pixels	22
2.15	Représentation d'une CNN	23
2.16	La convolution par un filter 3 x 3	24
2.17	Opération de convolution	25
2.18	illustration des deux types de pooling	25
2.19	Exemple d'un réseau de neurones profond.	26
3.1	Les phases de création	29
3.2	Pré-traitement des données	30
3.3	Les phases de l'approche	31
3.4	Illustration de CRNN	32
3.5	Aperçu du modèle proposé	33

4.1	Aperçu de l'ensemble de données pour chaque catégorie	39
4.2	Le nombre des images par partie.	40
4.3	Le pourcentage de chaque partie par rapport au total des données.	40
4.4	la distribution des données par classe présente dans EMNIST ByClass	41
4.5	La méthodologie de conception des modèles.	43
4.6	L'architecture complète du CNN de modèle 1.	43
4.7	Résultat du seuillage et de la dilatation.	45
4.8	Explication de la segmentation du nom du médicament.	45
4.9	L'architecture complète du CNN de modèle 2.	47
4.10	L'architecture complète du CNN de modèle 3.	48
4.11	Écriture manuscrite du nom du médicament bilastine.	51

LISTE DES TABLEAUX

2.1	Les définitions de l'intelligence artificielle, classées en quatre catégories . . .	10
3.1	Nombre d'images correctement identifiées par prescription	31
4.1	Le nombre des images dans chaque ensemble de données(dataset).	42
4.2	les techniques et les datasets utilisés dans chaque modèle.	42
4.3	La valeur d'accuracy pour chaque modèle.	51
4.4	Des mots prédits par chaque modèle.	51

LISTE DES ABRÉVIATION

- AI : Artificielle Intelligence .
- SAI : Super Artificielle Intelligence.
- GAI : General Artificial Intelligence.
- NLP : Natural Language Processing.
- ML : Machine Learning.
- SVM : Support Vector Machine.
- CNN : Convolutional Neural Network.
- RNN : Recurrent Neural Network .
- LSTM : Long Short-Term Memory.
- CTC : Connectionist Temporal Classification.
- RNN : Recurrent Neural Network.
- DL : Deep Learning.

INTRODUCTION GÉNÉRALE

Dans le passé et avant 3 200 ans Jc, l'homme a inventé l'écriture [1], qui était un moyen essentiel d'écrire l'histoire et de promouvoir la civilisation. Au fil du temps, elle est devenue un moyen pour les gens d'exprimer leurs idées, d'écrire diverses sciences et de communiquer avec eux. L'écriture manuscrite est également quelque chose qui distingue chaque individu des autres, en plus de cela, il est actuellement utilisé dans tous les domaines, y compris le domaine médical.

Il est courant et connu que la mauvaise écriture manuscrite des médecins est due à plusieurs raisons, notamment la pression du travail et la quantité d'informations à écrire, ce qui a entraîné de graves conséquences qui ont entraîné la mort de patients. 7 000 personnes meurent chaque année à cause de l'écriture incompréhensible des médecins [2].

Avec le développement de la technologie et l'émergence des techniques d'intelligence artificielle, de nombreux chercheurs ont utilisé certaines technologies comme solution à l'écriture erronée pour aider les agents de santé à comprendre l'écriture manuscrite des médecins. Ces techniques comprennent des algorithmes d'apprentissage automatique et des techniques d'apprentissage profond tel que les réseaux de neurones.

L'objectif premier de notre étude est d'élaborer un modèle capable de reconnaître l'écriture manuscrite des médecin dans les prescriptions médicales en utilisant des techniques d'apprentissage profond telles que les réseaux de neurones convolutifs, et nous avons organisé cette étude comme suit :

- Le premier chapitre se concentre sur l'écriture manuscrite et ses types, où nous expliquons les causes et les conséquences exactes de la mauvaise écriture manuscrite des médecins.
- Le deuxième chapitre est consacré à expliquer l'intelligence artificielle et ses différentes techniques, tout en clarifiant les points les plus importants.
- Le troisième chapitre se focalise sur des études des chercheurs spécialisées sur la reconnaissance de l'écriture manuscrite des médecins et les résultats de ces études

sont exposée.

- Le dernier chapitre représente notre contribution, où nous avons présenté les outils utilisés du début à la fin en présentant les données et en expliquant toutes les étapes du processus jusqu'à discuter des résultats obtenus.

Nous terminerons cette étude par une conclusion générale de nos travaux et indiquons les perspectives et les travaux futurs.

CHAPITRE 1

ÉCRITURE MANUSCRITE

1.1 Introduction

L'histoire de l'invention de l'écriture remonte à des milliers d'années, lorsque l'homme a réalisé qu'il avait besoin de communiquer avec les autres, de conserver son histoire, ses idées et sa culture. C'est le moyen par lequel les scientifiques, les écrivains, les historiens et les philosophes ont transmis leur savoir et leur connaissance à travers les générations, enregistrant ainsi l'histoire, les sciences sous toutes leurs formes et même les croyances. Jusqu'à aujourd'hui, l'écriture reste un axe important dans la vie quotidienne et est considérée comme une priorité pour les nations qui se concentrent, dans leurs différences, sur l'enseignement de l'écriture dès le plus jeune âge à leurs populations, Cependant, la pratique quasi quotidienne de l'écriture et la variété des personnalités a engendré une évolution des styles d'écriture et leur variété, de l'écriture lisible à l'écriture incompréhensible incluant l'écriture manuscrite des médecins, qui est célèbre pour être illisible en raison de divers facteurs.

1.2 Définition

L'écriture manuscrite est une pratique utilisée pour enregistrer et communiquer à l'aide d'instruments d'écriture, le mot manuscrite fait référence à un texte écrit à la main[3].

Pr Audrey van der Meer a mis en évidence la nécessité pour les enfants de recevoir une formation en écriture manuscrite car cela aide à renforcer la mémoire, car l'utilisation du papier et du stylo améliore la capacité du cerveau[4].

1.3 Les caractéristiques de l'écriture manuscrite

Il existe de nombreuses caractéristiques différentes qui entrent dans la création de l'écriture manuscrite et expriment la personnalité des individus, et nous les mentionnons :

— **Taille de l'écriture :**

La taille des lettres peut déterminer si une personne est timide ou extravertie. S'il écrit en grandes lettres au-delà de la première ligne, il est susceptible d'être extraverti et confiant. Une personne qui écrit en petites lettres et fermées est timide[5].

— **Pression du stylo :**

L'intensité de la pression exercée sur la page lors de l'écriture peut révéler des informations sur le niveau de stress d'une personne. Si une personne exerce trop de pression sur le stylo signifie qu'ils sont enclins à se sentir stressés et stressés par leur charge de travail. D'autre part, une pression légère indique la sensibilité et l'empathie[5].

1.4 Les types de l'écriture manuscrite :

Il existe trois styles différents dans lesquels l'écriture manuscrite peut être divisée :

— **Écriture cursive :**

L'écriture manuscrite cursive est une écriture "jointé", dans laquelle les lettres sont connectées, et vous pouvez écrire plus rapidement car vous n'avez pas à éloigner le stylo de la page[6].

L'écriture manuscrite cursive a été développée pour la première fois au 17ème siècle. Avant l'invention de la machine à écrire, l'écriture cursive était le style d'écriture préférée des professionnels comme les médecins.

— **Écriture manuscrite imprimée :**

L'écriture manuscrite imprimée, également appelée lettres majuscules, est faite lorsque le stylo quitte la page après chaque lettre.

Considéré comme plus clair et plus facile à lire et est souvent demandé lors du remplissage de formulaires[7].

— **Cursive modern :**

Il y a eu beaucoup de variations dans l'écriture cursive au fil des ans pour faciliter l'apprentissage des enfants. Il a été développé pour minimiser les traits inutiles afin qu'il soit moins fatigant d'écrire. De nos jours, il y a moins de boucles et de fioritures que dans les styles plus anciens, et le look est plus simple et épuré.

1.5 Les facteurs de singularité de l'écriture manuscrite

L'écriture manuscrite varie d'une personne à l'autre pour plusieurs raisons :

— Motricité et capacités physiques

L'écriture est une activité physique qui nécessite une coordination entre les muscles de la main, du bras et des yeux. Certaines personnes peuvent avoir de meilleures capacités motrices ou une meilleure coordination œil-main que d'autres, ce qui peut affecter leur écriture.

— personnage

L'écriture manuscrite peut également refléter les traits de personnalité d'une personne, tels que son niveau de confiance en soi, son état émotionnel et ses traits personnels. Par exemple, lorsqu'il y a des lettres incomplètes et une pression inégale, la personne est souvent sensible[8].

— Impact culturel et éducatif

La façon dont les gens apprennent à écrire à l'école et les normes culturelles entourant l'écriture peut également affecter leur style d'écriture.

Par exemple, quelqu'un qui a appris à écrire en cursive peut avoir un style différent de celui qui a appris à écrire en script.

— Pratiques et Habitudes

L'écriture est une compétence qui peut être améliorée avec la pratique, et quelqu'un qui écrit régulièrement peut développer un style d'écriture unique au fil du temps.

— Environnement physique

L'environnement physique dans lequel les gens écrivent affecter également leur écriture, comme le type de stylo ou de crayon qu'ils utilisent, l'angle et l'état de la surface d'écriture.

1.6 Écriture manuscrite des médecins

Dans le passé, les médecins prenaient des notes pour tenir un dossier personnel des antécédents médicaux d'un patient, qui étaient généralement consultées uniquement par le médecin. Cependant, aujourd'hui, la pratique médicale a évolué et les médecins ne travaillent plus en tant qu'individus isolés. Au lieu de cela, ils font partie d'une équipe de soins de santé multidisciplinaire composée de nombreux professionnels. Cependant, cette expansion a entraîné une conséquence inattendue : les médecins sont souvent pressés et écrivent rapidement des notes illisibles qui sont maintenant données à des collègues qui ont du mal à comprendre[9].

1.6.1 Les conséquences nuisibles

- Une écriture illisible peut retarder le traitement et entraîner des tests inutiles et des doses inappropriées qui, à leur tour, peuvent entraîner la mort. Chaque année, plus de 7 000 personnes sont victimes de la mauvaise écriture des médecins. En outre, d'après un rapport de juillet 2006 de l'Institut de médecine des Académies nationales des sciences, plus de 1,5 million d'Américains sont blessés à cause d'erreurs de médication évitables. Ces erreurs sont souvent liées à des abréviations et des indications de dosage peu claires, ainsi qu'à une écriture illisible sur certaines des 3,2 milliards d'ordonnances rédigées aux États-Unis chaque année[10].
- Une mauvaise écriture est un facteur contribuant au taux élevé d'erreurs médicales en Grande-Bretagne, qui entraînerait jusqu'à 30 000 décès par an[11].
- En 1999, un cardiologue américain a causé la mort d'un patient de 42 ans lorsque son pharmacien a mal interprété sa prescription contenant 20 mg d'Isordil, un antiangineux, comme 20 mg de Plendil, un antihypertenseur[12].



FIGURE 1.1 – La prescription a entraîné la mort du patient [12].

1.6.2 Les raisons de la mauvaise écriture des médecins

Les raisons pour lesquelles les médecins ont une écriture illisible sont multiples, notamment :

— La quantité

La grande quantité d'informations que les médecins écrivent au cours de leurs études dans des conférences et des cours afin d'écrire la plus grande quantité d'informations et de les mémoriser en peu de temps, leur a fait avoir une mauvaise écriture, car ils se concentrent sur la collecte d'informations et non sur la méthode pour l'écrire, et en

plus de cela, des ordonnances et des rapports qui doivent être remplis quotidiennement pour un grand nombre de patients en peu de temps.

— **La pression de travail**

La pression au travail, combinée avec la quantité d'informations, affecte les muscles de la main, car après un certain nombre d'heures, la main n'est plus en mesure de rattraper tous les détails de l'écriture, en plus d'autres facteurs liés à leur santé.

1.7 Problématique

Après avoir obtenu une vue détaillée sur l'écriture manuscrite et la mauvaise écriture des médecins et ces terribles conséquences, nous pouvons conclure que ce problème constitue une menace pour les patients et les médecins pour les raisons suivantes :

1. Mauvais diagnostic des patients.
2. L'émergence de nouvelles maladies symptomatiques.
3. Cause de décès des patients.
4. Perdre le temps.
5. Les médecins sont soumis à des sanctions légales s'ils causent des préjudices à leurs patients.
6. Le médecin subit des pertes financières et morales, financièrement, il perd ses patients, et il souffre d'une mauvaise réputation dans la société.
7. En cas d'accidents répétés, la communauté perd confiance dans les médecins, ainsi que la valeur que les médecins acquièrent.

1.8 Solution proposée

L'objectif de ce travail est de proposer une solution basée sur des techniques d'apprentissage profond qui vont permettre de comprendre l'écriture manuscrite des médecins dans les ordonnances, y compris les doses de médicaments, les noms et les heures d'utilisation afin de réduire les erreurs.

1.9 Conclusion

Ce chapitre examine en détail l'écriture manuscrite, en mettant l'accent sur l'écriture manuscrite des médecins et les raisons de son illisibilité, qui peut avoir des conséquences mortelles. Dans ce contexte, nous proposons une solution pour résoudre ce problème

dans le cadre de notre étude. Dans le prochain chapitre, nous présenterons l'intelligence artificielle et ses différentes techniques qui seront utilisées pour répondre à notre objectif.

2.1 Introduction

Dans ce chapitre, nous parlerons sur les techniques de l'intelligence artificielle, en mettant particulièrement l'accent sur les techniques qui nous avons utilisé dans notre travail qui propose une solution basée sur l'apprentissage profond pour reconnaître l'écriture manuscrite des médecins dans les prescriptions médicales.

2.2 Définition

La définition de l'intelligence artificielle diffère, car il y a ceux qui la définissent en termes de méthode et certains en termes de raisonnement, donc dans le tableau 2.1 ci-dessous il y a différentes définitions à travers deux dimensions différentes. La première dimension est divisée en penser et agir, et la deuxième est divisé en intelligence humaine et rationnelle, où ce dernier est cela signifie qu'il fait la bonne chose en fonction de ce qu'il sait[13].

<p><u>Penser humainement</u> ”L’automatisation d’activités que nous associons à la pensée humaine, des activités telles que la prise de décision, la résolution de problèmes, l’apprentissage... ”(Bellman, 1978) ”La tentative nouvelle et passionnante d’amener les ordinateurs à penser. . . des machines avec des esprits, au sens littéral et complet.” (Haugeland, 1985)</p>	<p><u>Penser rationnellement</u> ”L’étude des facultés mentales grâce à des modèles informatiques.” (Charniak et McDermott 1985) ”L’étude des moyens informatiques qui rendent possibles la perception, le raisonnement et l’action.” (Winston, 1992)</p>
<p><u>Agir humainement</u> ”L’art de créer des machines capables de prendre en charge des fonctions exigeant de l’intelligence quand elles sont réalisées par des gens.” (Kurzweil, 1990) ”L’étude des moyens à mettre en œuvre 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 Knight, 1991)</p>	<p><u>Agir rationnellement</u> ”L’intelligence artificielle (computational intelligence) est l’étude de la conception d’agents intelligents.”(Poole et al., 1998) ”L’IA. . . étudie le comportement intelligent dans des artefacts.” (Nilsson, 1998)</p>

TABLE 2.1 – Les définitions de l’intelligence artificielle, classées en quatre catégories [13].

2.3 Test de turing

Ce test a été proposé pour la première fois par le célèbre informaticien Alan Turing en 1950, et il est considéré comme l’un des tests les plus célèbres dans le domaine de l’intelligence artificielle, malgré les opinions divergentes à son sujet. Le test consiste à placer un être humain et un ordinateur dans deux salles isolées et à engager une conversation écrite avec un humain tiers. Si le tiers ne peut pas déterminer si l’interlocuteur est un humain ou un ordinateur, alors l’ordinateur est considéré comme intelligent, car il est capable de simuler la pensée humaine [14]. La figure 2.1 ci-dessous montre ce qui a été mentionné.

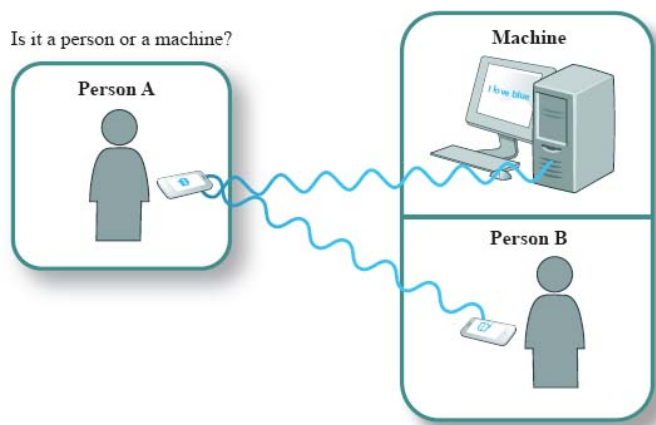


FIGURE 2.1 – Test de turing [15].

2.4 Les types d'intelligence artificielle

Il y a deux types principaux, qui sont les suivants :

2.4.1 Intelligence artificielle faible (Weak AI)

Également appelé IA étroite (Narrow AI), elle est formée et concentrée pour effectuer des tâches spécifiques. L'IA faible se concentre sur la résolution de problèmes spécifiques en utilisant des techniques d'apprentissage automatique et de traitement du langage naturel. Ce dernier couramment utilisé dans de nombreux systèmes d'IA que nous utilisons aujourd'hui, tels que les chatbots ou les assistants virtuels [16].

2.4.2 Intelligence artificielle fort (Strong AI)

L'Intelligence Artificielle forte (IA forte) se compose de l'Intelligence Artificielle Générale (IAG) et de la Super Intelligence Artificielle (SIA). L'Intelligence Artificielle Générale (AIG) est une forme théorique de l'IA dans laquelle la machine est supposée avoir une intelligence égale à celle de l'homme, possédant une conscience de soi, une capacité à résoudre des problèmes, à apprendre et à planifier pour l'avenir. L'Intelligence Artificielle Supérieure (SIA) est définie comme dépassant l'intelligence et les capacités du cerveau humain. Bien que l'IA forte soit encore entièrement théorique et qu'il n'y ait pas d'exemples pratiques utilisés jusqu'à présent [17].

2.5 Les branches de l'intelligence artificielle

Il y a plusieurs branches de l'intelligence artificielle, et certaines d'entre elles sont représentées dans la figure suivante :

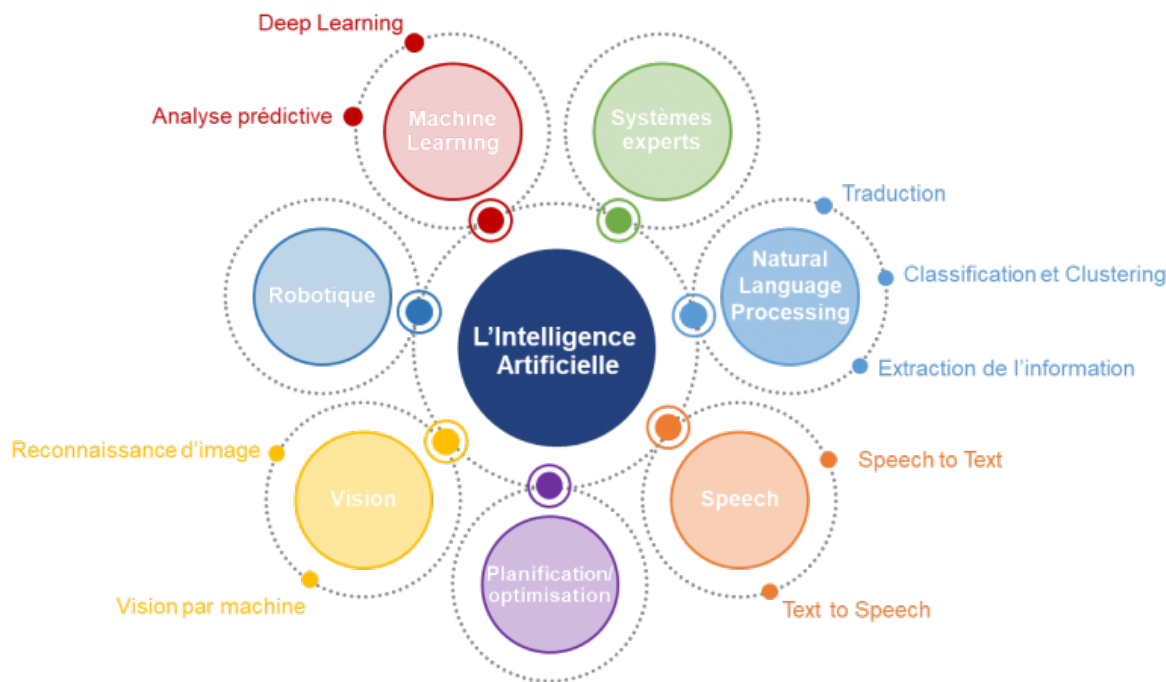


FIGURE 2.2 – Les branches de l'IA [18].

— Le traitement du langage naturel (Natural Language Processing)

Le traitement automatique des langues (TALN), également connu sous le nom de Natural Language Processing (NLP), est une branche de l'intelligence artificielle qui vise à doter les machines de la capacité de comprendre, de générer et de traduire le langage humain tel qu'il est parlé ou écrit. Les chatbots sont l'une des applications de NLP les plus couramment utilisés, et il existe de nombreuses autres applications célèbres telles que les assistants vocaux Alexa, Google Home et Siri [19].

— La robotique

Les robots sont des machines qui sont programmées par les êtres humains pour effectuer des tâches précises. Ils sont équipés de capteurs qui leur permettent de collecter des informations sur leur environnement extérieur, comme la température, les mouvements et les sons, qu'ils peuvent ensuite interpréter pour effectuer des actions appropriées. De plus, ils peuvent apprendre de leurs erreurs et s'adapter à de nouveaux environnements [13].

— La vision par ordinateur

La vision par ordinateur aide les ordinateurs et les systèmes d'extraire des informations utiles à partir d'images numériques, de vidéos et d'autres entrées visuelles, afin de prendre des mesures ou de formuler des recommandations en fonction de ces informations. Elle est utilisée dans la surveillance vidéo, les voitures autonomes, etc [20].

2.6 L'apprentissage automatique

L'apprentissage automatique ou Machine Learning (ML) en anglais, est un type de technologie qui relève de l'intelligence artificielle (IA). Son objectif principal est de permettre aux machines d'apprendre à partir des données et d'améliorer leurs performances par l'expérience, plutôt que de s'appuyer sur une programmation explicite. Les algorithmes d'apprentissage automatique sont conçus pour analyser des ensembles de données volumineux, identifier des patterns et des corrélations, et prendre des décisions et à émettre les meilleures prévisions. Comme ces algorithmes sont continuellement affinés par la pratique, ils deviennent de plus en plus précis, en particulier lorsqu'ils ont accès à de plus grandes quantités de données[21].

2.6.1 Les données

Comme la définition proposée précédemment, les algorithmes d'apprentissage automatique ont besoin de données pour fonctionner. Ces données sont également appelées des échantillons, des observations ou des exemples. Il existe deux grandes catégories de jeux de données (dataset) pouvant être utilisées :

1. Les données étiquetées : chaque observation x est associée à une étiquette y .
2. Les données non-étiquetées : comme leur nom l'indique, ces données ne sont pas accompagnées d'étiquettes [22].

2.6.2 Les types d'apprentissage automatique

Quatre types d'apprentissage automatique peuvent être distingués, qui sont présentés dans la figure 2.3, en fonction de la nature de l'apprentissage et du système d'apprentissage.

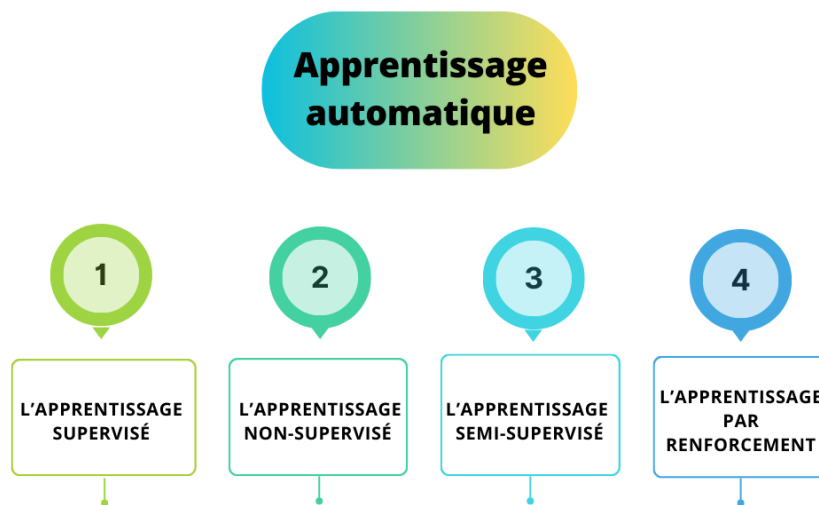


FIGURE 2.3 – Les types de ML

L'apprentissage supervisé (Supervised Learning)

Est une méthode qui se concentre sur les données étiquetées. L'objectif est de prédire l'étiquette y (qui est inconnue) associée à une nouvelle observation x en utilisant les informations fournies par les observations étiquetées du jeu de données [22]. L'exemple présenté dans la figure 2.4 illustre ce qui a été expliqué précédemment.

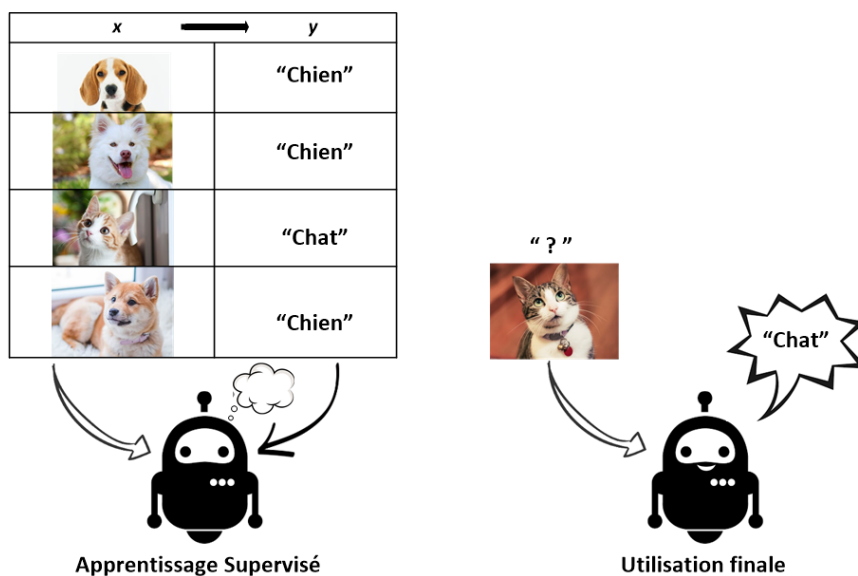


FIGURE 2.4 – Exemple de l'apprentissage supervisé [23].

Il existe deux types de modèle d'apprentissage supervisé :

— **Classification**

Un modèle de classification est utilisé pour prédire une valeur discrète (qualitative). Cela implique que l'ensemble des valeurs de sortie Y , que l'on cherche à estimer avec la fonction f , est un ensemble fini : $Y = \{0, 1, \dots, n\}$

Exemple : supposons que nous voulons créer un modèle $f : X \rightarrow Y$ qui prédit si le patient est atteint de la COVID-19 ou non. Le X représente l'ensemble des tests COVID-19 et $Y = 0, 1$; 1 s'il atteint la COVID-19 et 0 sinon. Alors si on veut analyser un test de COVID-19 m , on calcule $f(m)$ [24].

— **Régression**

Un modèle de Régression est utilisé pour prédire une valeur continue (quantitative). Cela implique que l'ensemble des valeurs de sortie Y , que l'on cherche à estimer avec la fonction f , est un ensemble de réel : $Y \subset \mathbb{R}$.

Exemple : supposons que nous voulons créer un modèle $f : X \rightarrow Y$ qui prédit le prix d'un ordinateur en fonction de sa caractéristique. Le X représente l'ensemble des caractéristiques des ordinateurs et Y l'ensemble des prix. Alors si l'on souhaite estimer le prix d'un ordinateur de caractéristique s , on calcule $f(s)$ [24].

L'apprentissage non-supervisé (Unsupervised Learning)

L'apprentissage non supervisé est une technique centrée sur des données sans aucune étiquette, et son objectif est de reconnaître des caractéristiques partagées entre différentes observations de manière automatique [22].

— **Clustering**

Le partitionnement, également connu sous le nom de clustering, consiste à regrouper des observations en clusters homogènes en fonction d'un critère de similarité mesurée par une distance D . Les méthodes de partitionnement est utilisée pour maximiser la similarité au sein de chaque groupe et minimiser la similarité entre les groupes[22].

Un exemple de clustering est présenté dans la figure ci-dessous.

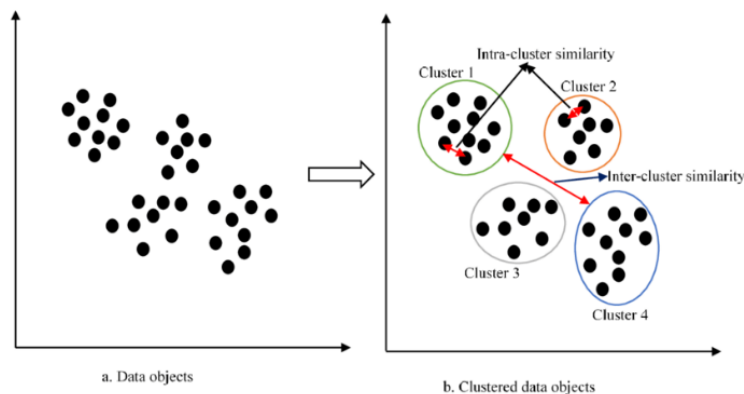


FIGURE 2.5 – Exemple de clustering [25].

L'apprentissage semi-supervisé (Semi-supervised Learning)

Cette technique consiste à apprendre des étiquettes à partir d'un jeu de données qui contient à la fois des données étiquetées et non étiquetées. Cette technique présente l'avantage de ne pas avoir à étiqueter manuellement les données non étiquetées, mais plutôt à utiliser des étiquettes approximatives basées sur des données étiquetées[26].

L'apprentissage par renforcement (Reinforcement Learning)

L'apprentissage par renforcement est une approche d'apprentissage automatique qui vise à enseigner à un agent comment prendre les bonnes décisions dans un environnement donné, en renforçant les comportements appropriés et en évitant les erreurs et les comportements inappropriés. Cette méthode est adaptée pour résoudre des problèmes où les actions appropriées ne peuvent pas être prédéterminées. L'une des particularités de cette approche est que l'agent apprend de ses expériences et améliore sa performance au fil du temps[27].

Voici le mécanisme de fonctionnement de cette approche :

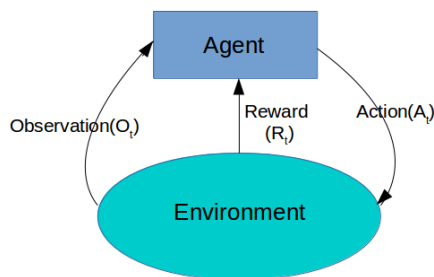


FIGURE 2.6 – Le mécanisme de fonctionnement de L'apprentissage par renforcement [28].

2.7 Algorithmes d'apprentissage automatique

Il existe de nombreux algorithmes utilisés dans l'apprentissage automatique, y compris l'algorithme mentionné ci-dessous :

2.7.1 Machine à vecteur de support (SVM)

L'algorithme SVM (Support Vector Machine) est un algorithme d'apprentissage automatique supervisé qui peut être utilisé pour résoudre des problèmes de classification ou de régression. Dans cet algorithme, chaque donnée est représentée par un point dans un espace à n dimensions (où n est le nombre d'attributs dont on dispose), la valeur de chaque attribut étant la valeur d'une coordonnée particulière. Ensuite, l'algorithme effectue une classification en trouvant l'hyperplan optimal qui permet de bien différencier les deux classes. Il existe deux types de SVM, le SVM linéaire et le SVM non linéaire[29].

voici un exemple d'un courbe SVM :

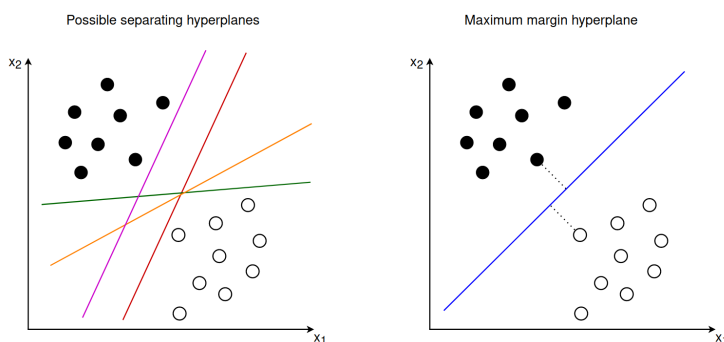


FIGURE 2.7 – Classification des données à l'aide de la courbe SVM [30].

- La fonctionnement de l'algorithme SVM

Dans l'exemple d'application décrit ci-dessous, nous pouvons observer que il contient deux classes distinctes : la classe des points de données carré et la classe des points de données cercle.

La tâche fondamentale de l'algorithme SVM (Support Vector Machine) est de trouver un hyperplan optimal. Cela signifie rechercher un hyperplan qui sépare parfaitement les deux classes, et qui il est également le plus éloigné possible de tous les points de données. En effet, pour obtenir un hyperplan optimal, il est nécessaire de conserver une marge maximale afin de pouvoir placer les nouveaux points de données dans les classes qui leur correspondent le mieux.

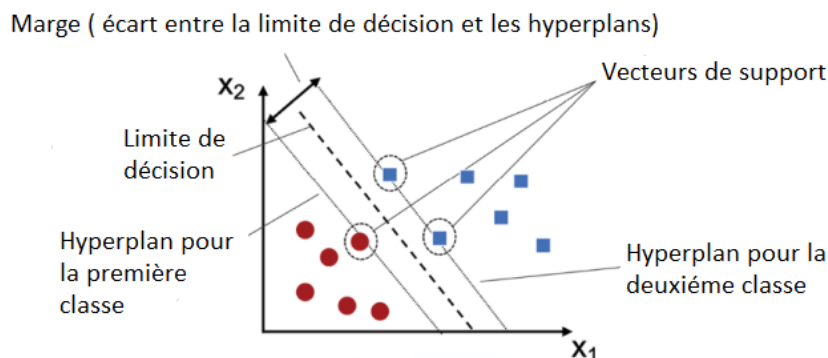


FIGURE 2.8 – Exemple de l’algorithme SVM [31].

- Hyperplan et les vecteurs de support dans l’algorithme SVM :

1. Hyperplan

Les hyperplans sont des frontières de décision qui aident à classer les points de données. Les points de données se trouvant de part et d’autre de l’hyperplan peuvent être attribués à différentes classes. De plus, la dimension de l’hyperplan dépend du nombre de caractéristiques. Si le nombre de caractéristiques d’entrée est de 2, alors l’hyperplan est simplement une ligne. Si le nombre de caractéristiques d’entrée est de 3, alors l’hyperplan devient un plan bidimensionnel [32].

2. Vecteurs de support

Les vecteurs de support sont les points de données les plus proches de l’hyperplan. C’est des points de données qui influent sur la position et l’orientation de l’hyperplan et qui aident à maximiser la marge [32].

2.8 Réseaux de neurones et apprentissage profond

2.8.1 Réseaux de neurones

- Définition :

Un réseau de neurones est composé de nœuds connectés entre eux par des connexions. Sa capacité de traitement dépend des poids et de puissance, qui sont déterminées par un processus d’adaptation et d’apprentissage à partir d’un ensemble de données d’entraînement.

Pour bien comprendre les réseaux de neurones, il est important d’avoir une brève

compréhension sur la neurobiologie. car les réseaux de neurones sont principalement inspirés de la cellule nerveuse [33].

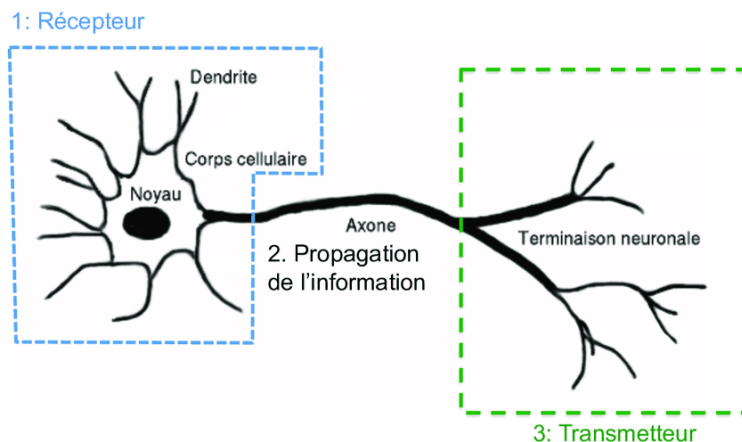


FIGURE 2.9 – La structure d’un neurone biologique [34].

Le cerveau humain est composé d’environ 100 milliards de neurones. Les neurones communiquent via des signaux électriques qui sont des impulsions de courte durée ou des ”pics”. Chaque neurone reçoit généralement des milliers de connexions provenant d’autres neurones et est donc constamment soumis à une multitude de signaux entrants, qui finissent par atteindre le corps cellulaire. Ensuite ils vont être additionnés ensemble d’une certaine manière puis décident s’il va générer une impulsion électrique ou pas qui va être transmise via les axons aux autres cellules[33].

- Structure d’un réseau de neurones artificielle :

”Perceptron” est la structure la plus simple et la première structure proposée pour un réseau de neurones artificielle, inspirée des cellules nerveuses biologiques[35].

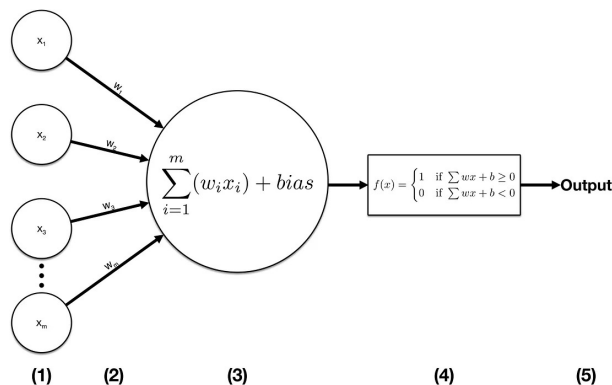


FIGURE 2.10 – une modèle de perceptron [35].

1. Sur le côté gauche, vous pouvez observer des neurones (représentés par de petits

cercles) de x avec des indices 1, 2, ..., m portant des données d'entrée.

2. Ici, chaque entrée est multipliée par le poids correspondant dans la branche, où le poids représente la force de la connexion.
3. Une fois que toutes les entrées sont multipliées par leur poids respectif, elles sont ensuite toutes additionnées, et un nombre prédéfini appelé le biais est ajouté.
4. Le résultat de l'opération précédente, est utilisé comme entrée pour la fonction d'activation. Dans cet exemple, la fonction d'activation renvoie 1 si l'entrée est supérieure à 0, sinon elle renvoie 0.
5. La sortie est soit 1 soit 0 [35].

- L'apprentissage :

Le processus d'apprentissage se déroule ici par la modification des paramètres de poids, il existe deux méthodes principales de l'apprentissage, qui sont les suivantes :

1. Supervisé

Dans ce type d'apprentissage, les entrées et les sorties sont préalablement connues, ce qui signifie que la base de données d'apprentissage est étiquetée. Le réseau traite ensuite les entrées et compare les résultats obtenus aux sorties désirées. Les poids sont ajustés en fonction des erreurs propagées à travers le système. Ce processus est répété tant que les poids continuent à s'améliorer. L'ensemble de données utilisé pour l'apprentissage est appelé ensemble d'apprentissage [33][36].

2. Non supervisé

Dans ce type d'apprentissage, les résultats ne sont pas donnés à l'avance. Le réseau ne reçoit que l'entrée, puis une fonction est utilisée pour regrouper les données d'entrée. Cette méthode s'appelle l'adaptation[33][36].

- Fonctions d'activation :

Il existe plusieurs fonctions telles que :

1. Fonction sigmoïde (Sigmoid)

Il permet de transformer les valeurs d'entrée, qui peuvent varier de moins l'infini à plus l'infini, en une valeur appropriée comprise entre 0 et 1.

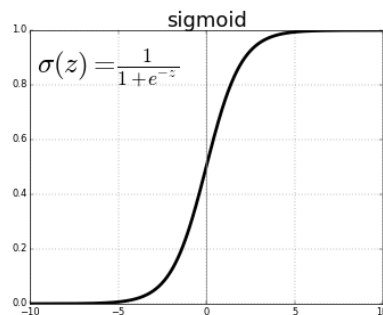


FIGURE 2.11 – Fonction d’activation sigmoïde [37].

2. Fonction ReLU (Rectified Linear Unit)

Cette fonction conserve toutes les valeurs positives et attribue simplement 0 aux valeurs négatives.

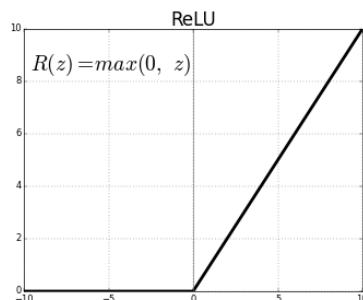


FIGURE 2.12 – Fonction d’activation ReLU [37].

- Les différents types d’architecture des réseaux de neurones :

L’architecture des réseaux de neurones diffère en fonction de la façon dont les neurones sont liés entre eux, en plus de la différence dans la manière dont ils seront traités. Voici des exemples de certaines structures :

1. Réseaux de neurones à propagation avant (Feedforward Neural Networks)

Il est également connu ce le nom ”Multilayer Perceptrons” ce type d’architecture est composé d’une couche d’entrée, une ou plusieurs couches cachées et une couche de sortie, Où chaque couche a sa propre fonction. La couche d’entrée reçoit les signaux d’entrée du monde extérieur et redistribue ces signaux à tous les neurones de la première couche cachée. Jusqu’à attendre la couche de sortie et obtenir le modèle de sortie. Ce type de réseau est appelé ”Feedforward Neural Networks” car le flux d’informations se déplace uniquement dans une direction, de la couche d’entrée à la couche de sortie, sans aucune connexion de rétroaction vers les couches précédentes. Cela signifie que les signaux se déplacent dans une seule direction [38].

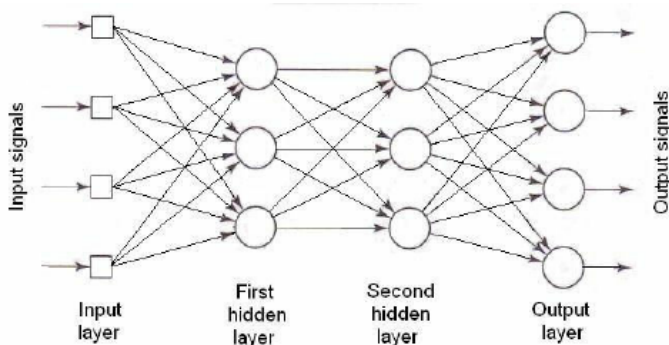


FIGURE 2.13 – Réseaux de neurones à propagation avant [38].

2. Réseau récurrent (Feed back)

Aussi appelé en anglais "Recurrent Neural Network (RNN)". Il s'agit d'un type de réseau neurones qui contient une mémoire et convient le mieux aux données séquentielles, où les sorties de la phase précédente sont mémorisées, puis transmises en tant qu'entrée lors de la phase suivante. Les RNN sont couramment utilisés dans la traduction automatique, la génération de texte, l'analyse du sentiment [39].

LSTM (Long Short-Term Memory)

Le problème des RNN (Réseaux Neuronaux Récurrents) est leur incapacité à conserver des données sur une longue période et à oublier progressivement leurs entrées précédentes. C'est pourquoi les LSTM ont été proposés, car ils traitent le problème des dépendances à long terme. Autrement dit, si l'état précédent qui influence les prédictions actuelles est dans le passé récent, les LSTM sont capables de le prendre en compte et de l'utiliser efficacement[39].

3. Réseaux de neurones convolutifs (Convolutional Neural Networks ou CNN)

Est un type spécifique de réseau neurones conçus pour traiter des données avec une structure en forme de grille, comme des images.

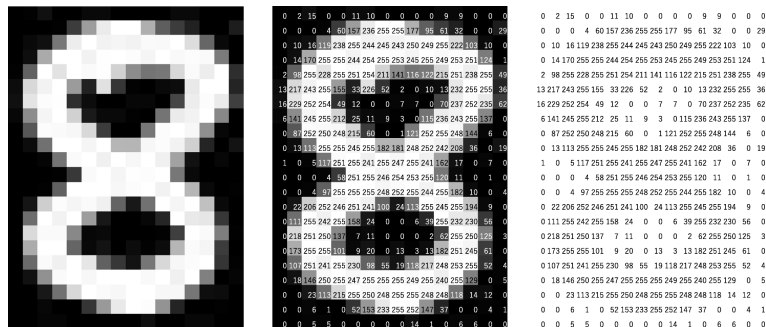


FIGURE 2.14 – Représentation d'une image en forme de matrice des pixels [40].

Un ordinateur voit une image comme une matrice de nombres. La matrice de droite représente les valeurs de luminosité des pixels dans l'image de gauche, avec des nombres allant de 0 à 255. Les deux images sont ensuite superposées dans l'image du milieu [40].

CNN est une construction mathématique qui est généralement composée de trois types de couches : les couches de convolution, de pooling (regroupement) et fully connected (entièrement connectées). Les deux premières couches, de convolution et de pooling, effectuent l'extraction des caractéristiques, tandis que la troisième couche, mappent les caractéristiques extraites vers la sortie finale, telle que la classification[40].

Une couche de convolution joue un rôle essentiel dans les CNN, qui est composée d'une séquence d'opérations mathématiques, dont la convolution, une opération linéaire spécialisée. Dans les images numériques, les valeurs des pixels sont stockées dans une grille bidimensionnelle (2D), c'est-à-dire un tableau de nombres, et un petit ensemble de paramètres appelé Kernel (noyau), qui est un extracteur de caractéristiques optimisable, est appliqué à chaque position de l'image. Cela rend les CNN très efficaces pour le traitement d'images, car une caractéristique peut apparaître n'importe où dans l'image. À mesure qu'une couche transmet sa sortie à la couche suivante, le processus d'optimisation des paramètres tels que les noyaux est appelé entraînement. Il vise à réduire la différence entre les sorties et les étiquettes de vérité terrain grâce à des algorithmes d'optimisation tels que la backpropagation[40].

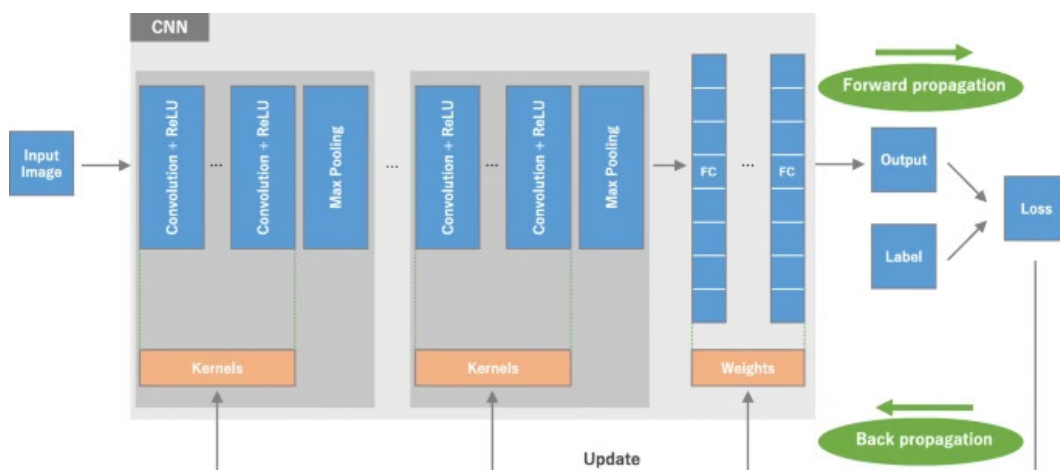


FIGURE 2.15 – Représentation d'une CNN [40].

Couche de convolution

Dans l'architecture CNN, une couche de convolution fait l'extraction des caractéristiques. Elle est composée principalement d'une opération de convolution A travers

un filtre ou un noyau dont la taille est généralement de 3 x 3 ou 2 x 2 et d'une fonction d'activation. La figure ci-dessous montre ce qui a été mentionné [40].

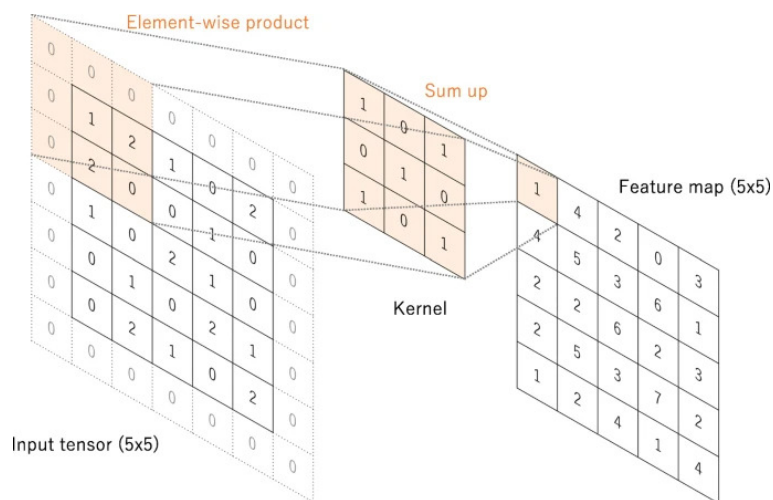


FIGURE 2.16 – La convolution par un filter 3 x 3 [40].

La matrice générée par convolution est appelée carte de caractéristiques (feature map). La taille de la carte est donnée par l'équation suivante :

$$\text{La taille de la carte} = 1 + \frac{W - F + 2P}{S}$$

- W représente la taille de la matrice source,
- F représente la taille du filtre,
- P a la valeur 1 si le padding a été ajouté et 0 sinon,
- S représente la valeur de déplacement du filtre dans la matrice.

La méthode de calcul des valeurs dans la carte des caractéristiques est présentée dans la figure 2.17 .

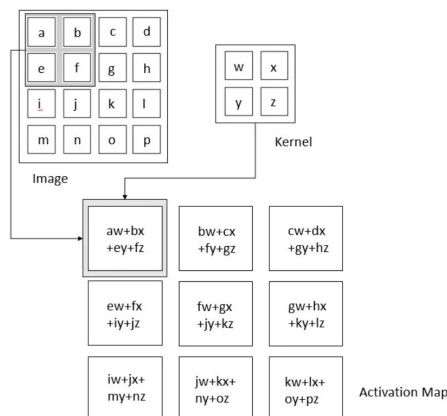


FIGURE 2.17 – Opération de convolution [41].

Après cette étape une fonction d'activation est appliquée sur la carte, certains d'entre eux ont déjà été expliqués.

Couche de pooling

Une couche de pooling fournit une opération de sous-échantillonnage qui réduit la dimensionnalité dans la carte de caractéristiques. Il existe deux types de pooling commun qui sont les plus couramment utilisés, et ils sont les suivants :

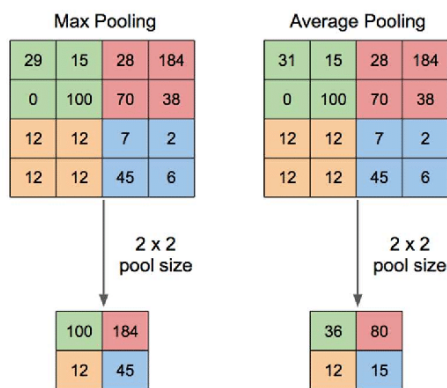


FIGURE 2.18 – illustration des deux types de pooling [42].

— Max pooling

C'est la forme la plus populaire, cela fonctionne en sélectionnant la valeur maximale de chaque groupe ou "pool" et ignorant les autres valeurs, le max pooling utilise un filtre de 2 x 2 avec un pas (stride) de 2 en pratique [40].

— Average pooling

Une autre opération qui calcule la valeur moyenne des pool d'une carte des caractéristiques et l'utilise pour créer une carte sous-échantillonnée [40].

Couche entièrement connectée (Fully connected layer)

La carte des caractéristiques de sortie de la dernière couche de convolution est généralement aplatie, c'est-à-dire transformée en un tableau unidimensionnel (1D) de nombres (ou vecteur), ensuite elle est passée à un ensemble de couches entièrement connectées. Ces couches utilisent les informations de chaque neurone pour calculer les sorties finales du réseau, telles que les probabilités pour chaque classe dans les tâches de classification [40].

Chaque couche entièrement connectée est généralement suivie d'une fonction non linéaire, comme la fonction ReLU [40].

2.8.2 Apprentissage profond (deep learning)

- Définition :

Est également appelé en anglais le deep learning, est un sous-domaine de l'apprentissage automatique qui a permis à l'intelligence artificielle de réaliser des progrès significatifs ces dernières années.

Le deep learning repose sur des réseaux de neurones artificiels, avec l'utilisation de techniques. Le modèle basé sur l'apprentissage profond se distingue par l'existence de différentes architectures qui effectuent différentes tâches, telles que la couche de convolution dans les couches cachées qui peut accomplir des tâches sans intervention externe comme l'extraction de caractéristiques.

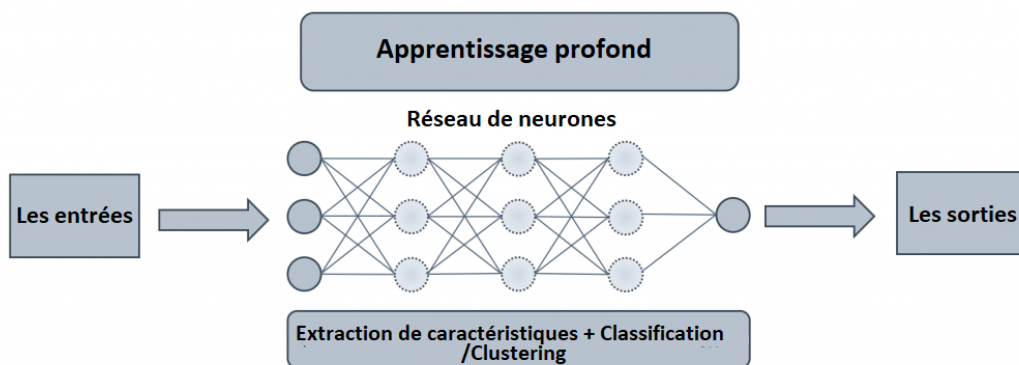


FIGURE 2.19 – Exemple d'un réseau de neurones profond.

- Principe de fonctionnement :

Les couches d'un modèle basé sur l'apprentissage profond sont regroupées en trois types :

1. La couche d'entrée (Input layer) :

Son rôle consiste à recevoir les données d'entrée et à les transmettre à la première couche cachée.

2. Les couches cachées (Hidden layers) :

Dans cette partie, chaque nœud effectue des calculs mathématiques sur les entrées et transmet le résultat obtenu en tant qu'entrée à la couche suivante. Une explication détaillée a été présentée précédemment dans la partie intitulée "Réseau de neurones".

L'entraînement d'un modèle basé sur l'apprentissage profond consiste à améliorer et ajuster les poids pendant le processus d'entraînement afin d'obtenir des valeurs précises.

3. La couche de sortie (Output layer) :

Elle est responsable de produire la sortie désirée ou les prédictions en fonction des informations traitées par les couches précédentes[43]. Telles que les probabilités des différentes classes dans une tâche de classification.

Les modèles de l'apprentissage profond à un besoin d'une grande puissance de calcul.

2.9 Conclusion

Dans ce chapitre, nous avons exploré les aspects de l'intelligence artificielle, notamment ses types, ses branches et ses techniques qui ont à le besoin dans notre étude comme les réseaux de neurones et l'apprentissage profond.

Le prochain chapitre, nous allons explorer diverses études de recherche sur l'application de l'intelligence artificielle pour la reconnaissance de l'écriture manuscrite des médecins et d'autres.

3.1 Introduction

Les études antérieures aident à guider le chercheur tout au long du processus de recherche scientifique, car les nouvelles études sont considérées comme des études complémentaires dans lesquelles de nouvelles techniques ou idées sont prouvées, modifiées ou ajoutées.

De nombreux chercheurs ont fait de grands efforts pour aider les pharmaciens ainsi que les travailleurs de la santé à comprendre l'écriture des médecins, et dans ce chapitre, nous allons citer quelques travaux de recherche qui ont utilisé l'IA pour reconnaître l'écriture des médecins.

3.2 Solutions basées sur l'apprentissage profond

Les chercheurs ont utilisé diverses techniques d'apprentissage profond pour reconnaître l'écriture manuscrite des médecins afin de réduire les erreurs catastrophiques. Ces études font partie d'un groupe d'études qui ont travaillé récemment sur ce sujet et ont également utilisé l'apprentissage profond comme solution.

3.2.1 Étude 01 : Doctor's Handwritten Prescription Recognition System In Multi-Language Using Deep Learning [44]

Dans l'étude [44] les chercheurs se reposent sur des techniques de l'apprentissage profond pour créer un outil capable de traduire l'écriture manuscrite des médecins dans

n'importe quelle langue. Ils utilisent des techniques d'apprentissage en profondeur, notamment CNN, RNN et LSTM.

Le système a été transformé en une application. Lorsqu'un utilisateur télécharge une image de prescription, le programme effectue un prétraitement de l'image et une segmentation des mots. Ces étapes sont effectuées pour chaque langue que le modèle doit détecter. De plus, la recherche floue et l'analyse du panier de marché sont utilisées pour améliorer les résultats à l'aide de la base de données pharmaceutiques. La figure ci-dessous représente les phases de design du système.

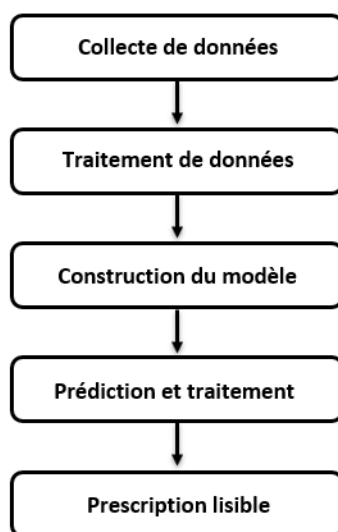


FIGURE 3.1 – Les phases de création [44].

Dans cette étude appliquée, les données sont divisées en deux parties : les données d'entraînement (training data) représentent 90 % des données, et les données de test (testing data), qui représentent 10 % de l'ensemble de données d'apprentissage.

Les données ont été prétraitées en utilisant plusieurs étapes mentionnées dans l'image ci-dessous.

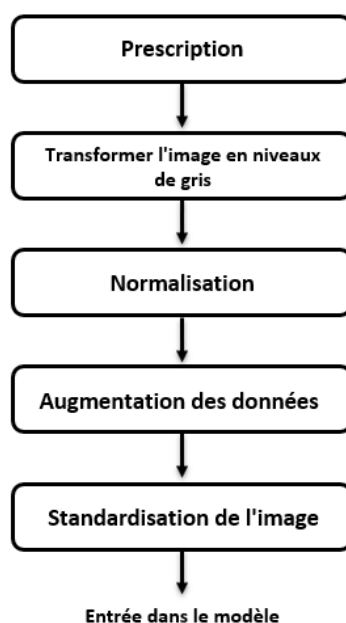


FIGURE 3.2 – Pré-traitement des données [44].

Les chercheurs ont essayé de maintenir la perte CTC (Connectionist Temporal Classification Loss) au niveau le plus bas possible pour atteindre les meilleurs résultats. Il a également été mentionné que le modèle a été testé dans une pharmacie dans le but de corriger les erreurs et d'améliorer tout en obtenant des résultats dans la langue que l'utilisateur recherché.

3.2.2 Étude 02 : Doctor's Cursive Handwriting Recognition System Using Deep Learning [45]

L'objectif général de cette étude est de développer un système neuronal récurrent convolutif profond pour la reconnaissance de l'écriture manuscrite cursive des médecins.

Les données ont été collectées auprès de 50 médecins différents, où chaque médecin a été invité à écrire une liste de 12 médicaments spécifiques prédéfinis, tels que des médicaments difficiles à lire. Chaque médecin a écrit la liste trois fois, ce qui a permis de collecter un total de 1800 images, qui ont ensuite été traitées pour être utilisées dans l'entraînement du modèle.

L'étude met en œuvre l'approche suivante qui comporte quatre étapes, comme le montre la figure.

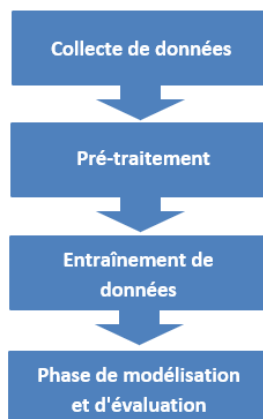


FIGURE 3.3 – Les phases de l'approche [44].

1260 images (70 % des données) ont été utilisées pour l'entraînement du modèle, atteignant une accuracy de 76 %. Ensuite, 540 images (les 30 % restants) ont été utilisées pour la validation, atteignant une accuracy de 72 %. Les 540 images ont été divisées en 12 groupes (les médicaments de la liste), chacun contenant 45 images et chaque groupe a atteint une précision différente, comme indiqué dans le tableau ci-dessous.

Prescription	Nombre d'images correctement identifiées
Prescription1	39
Prescription2	35
Prescription3	26
Prescription4	37
Prescription5	30
Prescription6	26
Prescription7	38
Prescription8	23
Prescription9	34
Prescription10	32
Prescription11	29
Prescription12	40

TABLE 3.1 – Nombre d'images correctement identifiées par prescription [45].

OpenCV, Python et Tensorflow ont été utilisés pour l'entraînement.

3.2.3 Étude 03 : Medical Handwritten Prescription Recognition Using CRNN [46]

Ce projet est basé sur le Convolutional Recurrent Neural Network (CRNN), qui s'inspire de l'architecture VGG16 combinée avec le Recurrent Neural Network (RNN).

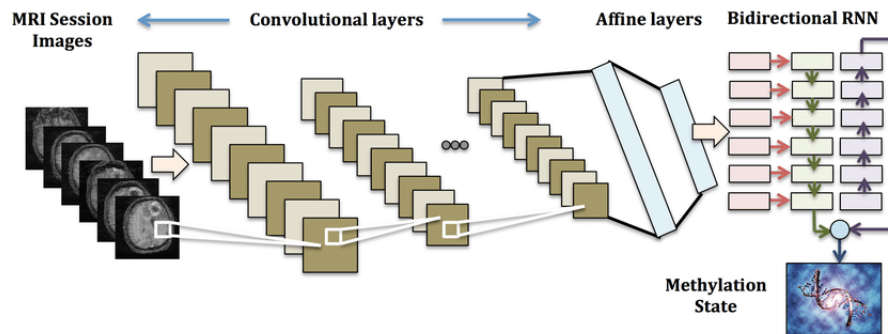


FIGURE 3.4 – Illustration de CRNN [46].

Cette étude se concentre sur les prescriptions médicales écrites en anglais et utilise 66 catégories distinctes, comprenant des caractères alphanumériques, des signes de ponctuation et des espaces. Avant leur utilisation dans la phase d'entraînement, les images ont subi un prétraitement qui a inclus le redimensionnement et la clarification des données.

L'accuracy de modèle a atteint 95 % et l'erreur a diminué progressivement jusqu'à atteindre finalement 0,003 et les résultats ont été testés en utilisant des prescriptions différentes de celles utilisées pour l'entraînement. Un total de 50 prescriptions ont été utilisées, avec 45 résultats corrects, soit une erreur de 10 %.

3.3 Solutions basées sur l'apprentissage automatique

Les chercheurs ont utilisé différentes méthodes pour résoudre le problème de l'écriture manuscrite des médecins, dont l'utilisation de l'apprentissage automatique et de ses algorithmes. En effet, plusieurs études ont été menées sur ce sujet ces dernières années, et nous pouvons présenter le suivant :

3.3.1 Étude 01 : Medical Prescription Recognition using Machine Learning [47]

Les chercheurs de ce travail ont proposé une application mobile basée sur l'apprentissage automatique qui reconnaît les noms de médicaments manuscrits et renvoie un texte numérique lisible du médicament et de sa dose.

L'ensemble de données est collecté auprès de plusieurs médecins et hôpitaux avec différentes spécialisations, car ces données ont subi de nombreuses modifications comme la conversion noir-blanc, la réduction du bruit et le redimensionnement de l'image. La figure ci-dessous montre une vue d'ensemble du système.

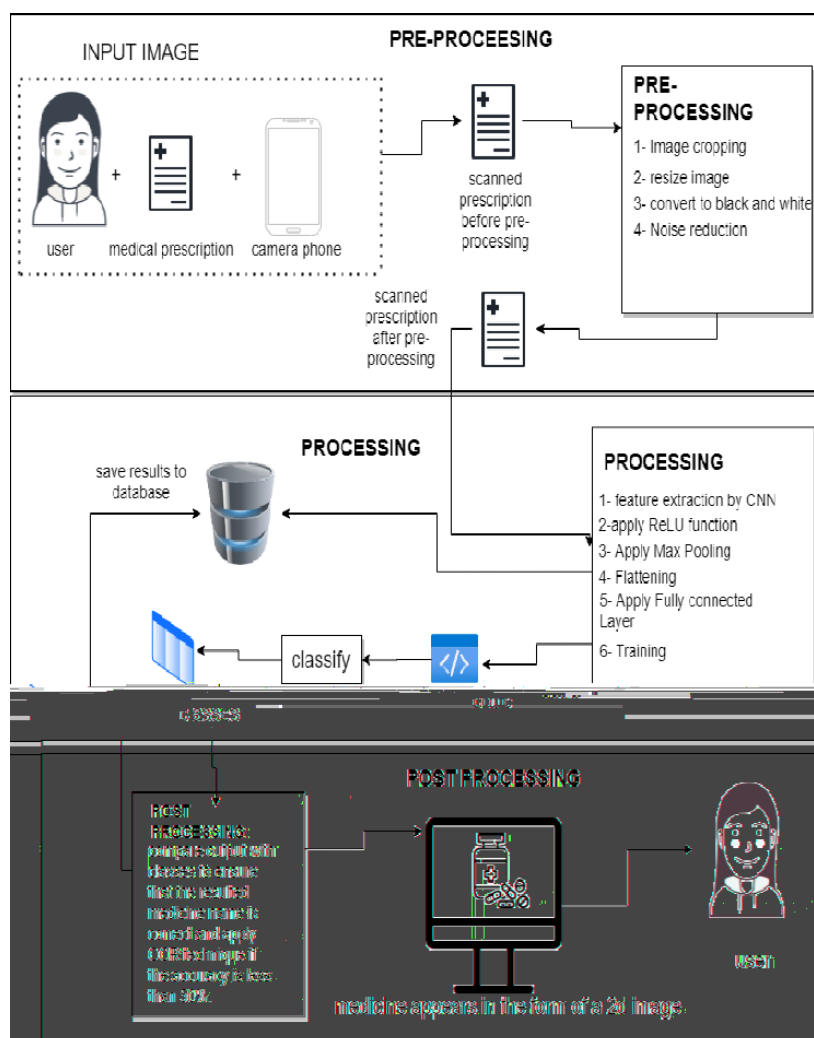


FIGURE 3.5 – Aperçu du modèle proposé [47].

L'ensemble de données a été divisé en 70 % pour l'entraînement et 30 % pour les tests afin de former le modèle proposé.

La prescription était divisée en trois parties, la partie supérieure qui contient le nom du médecin et sa spécialisation, la partie médiane qui contient les médicaments avec leurs doses, et c'est la partie la plus importante qui sera travaillée en la re-traitant par de nombreuses étapes, qui seront finalement classées. Quant à la partie inférieure, qui contient l'adresse avec le numéro du médecin, elle sera écartée.

Finalement, une accuracy de l'entraînement a donné 73 % et une accuracy du test a donné 50 %.

3.4 Solution pour l'écriture en général par le deep learning

Huruf : An application for arabic handwritten character recognition using deep learning [48]

Cette étude a développé un modèle pour reconnaître les lettres et les chiffres manuscrits arabes à l'aide de techniques d'apprentissage en profondeur. Où un réseau neuronal convolutif composé de 18 couches a été utilisé.

L'étude a utilisé deux ensembles de données publics. Le premier ensemble de données, appelé "MaDBase" (Modified Arabic handwritten digits Database), contenait des échantillons provenant de 700 rédacteurs. Chaque rédacteur a contribué avec 10 échantillons des chiffres arabes de 0 à 9, ce qui a donné un total de 70 000 images.

Le deuxième ensemble de données, connu sous le nom de "Arabic Handwritten Character Dataset (AHCD)", contenait 16 800 caractères écrits par 60 participants. Chaque participant a écrit les 28 lettres arabes dix fois.

Les données ont été prétraitées avant leur utilisation pour l'entraînement, où dans l'entraînement les chiffres et les lettres ont été entraînés séparément. L'ensemble global des données des caractères a été divisé en deux parties : 80 % (13440 images) pour l'entraînement et 20 % (3360 images) pour le test. Chaque partie, l'ensemble d'entraînement et l'ensemble de test, a été divisée en 28 classes.

Le modèle a obtenu une accuracy de 96,93 % lorsqu'il a été testé sur des données de lettres, et une accuracy de 99,35 % lorsqu'il a été testé sur des données de chiffres.

Plusieurs métriques ont été utilisées dans cette étude, telles que :

- Precision
- F1-score
- Recall

3.5 Discussion et critique

Pour l'étude de la référence [44], il a été mentionné dans l'article qu'il n'y avait pas des données disponibles, ce qui les a obligés à créer les données à partir de zéro, ce qui n'est pas correct, car il existe des données disponibles en ligne telles que le EMNIST. De plus, le modèle est censé reconnaître plusieurs langues, ce qui peut affecter sa performance, en plus du fait qu'il aurait besoin d'une grande quantité de données pour s'entraîner sur plusieurs langues, ce qui n'a pas été mentionné dans l'étude. En outre, il n'a pas été fait mention l'accuracy du modèle ou de toute autre métrique permettant d'évaluer le modèle.

Pour la deuxième étude, intitulée par "Doctor's Cursive Handwriting Recognition System Using Deep Learning", il a été mentionné que les 12 médicaments sur lesquels les

données ont été construites sont parmi les médicaments les plus souvent mal lus par les personnes, mais cela reste relatif, car toute erreur peut se produire pour n'importe quel médicament. Par conséquent, ces données ne sont pas suffisantes pour entraîner le modèle. De plus, le domaine médical étant très sensible, toute erreur peut coûter très cher. Il n'a pas non plus été fait mention de l'accuracy du test.

L'étude référencée par [47] a répété l'accuracy de l'entraînement à plusieurs reprises, mais ce qui est plus important à souligner est l'accuracy du test, qui a été mentionnée quelques fois avec peu d'explication sur les raisons de ses résultats.

3.6 Conclusion

Dans ce chapitre, nous avons discuté de certaines études qui ont traité de la reconnaissance de l'écriture manuscrite des médecins, où chaque étude l'a traitée avec son propre style et ses propres outils. Dans le prochain chapitre, nous présenterons en détail notre propre contribution, en expliquant toutes les étapes et les outils utilisés, avec une clarification des résultats et leurs discussions.

CHAPITRE 4

IMPLÉMENTATION DES ALGORITHMES ET PRÉSENTATION DES RÉSULTATS

4.1 Introduction

Dans ce chapitre, nous allons présenter en détail notre propre contribution à la résolution de ce problème. Nous expliquerons toutes les étapes et les outils que nous avons utilisés, en fournissant une clarification des résultats obtenus et leurs discussions.

4.2 Logiciels et libraires utilisé dans l'implémentation

De nombreux outils ont aidé à accomplir ce travail, que nous mentionnerons avec une brève explication de chaque outil

4.2.1 Python

Python est un langage de programmation orienté objet de haut niveau qui a la capacité de traiter différentes structures de données, en plus de la présence de packages et de bibliothèques prédéfini. Il est considéré comme le langage le plus utilisé dans les domaines de l'intelligence artificielle car il prend en charge des bibliothèques telles que TensorFlow et Keras.

4.2.2 Tensorflow

TensorFlow est une bibliothèque de machine learning open source créée par Google qui regroupe un grand nombre de modèles et d'algorithmes de machine learning et de

deep learning.

C'est une boîte à outils pour résoudre facilement des problèmes mathématiques extrêmement complexes. Il permet aux chercheurs de développer et de traduire des architectures d'apprentissage expérimentales en logiciels [49].

4.2.3 Keras

Keras est une API d'apprentissage profond écrite en Python, exécutée sur la plateforme d'apprentissage automatique Tensor Flow, Microsoft Cognitive Toolkit (CNTK) et Theano. Il a été développé dans le but de permettre une expérimentation rapide et de permettre la constitution rapide de réseaux neuronaux, de plus Keras est simple, flexible et puissant.

4.2.4 Spyder

Spyder, abréviation de "Scientific PYthon Development EnviRonment", est un environnement de développement intégré (IDE) spécialement conçu pour le calcul scientifique, l'analyse de données et la visualisation de données en Python. Il fournit un environnement puissant et intuitif pour écrire, exécuter et déboguer du code Python, en particulier pour les tâches liées à la recherche scientifique, à l'analyse de données et à l'apprentissage automatique.

4.2.5 Google Colaboratory

Google Colaboratory est un environnement de développement intégré (IDE) basé sur le cloud où tout le traitement a lieu sur les serveurs de Google.

Est construit sur Jupyter Notebook et prend en charge l'informatique interactive, l'analyse de données et l'apprentissage automatique, il fournit la puissance de calcul et les ressources nécessaires, y compris le CPU et le GPU.

Google Colaboratory est livré préinstallé avec plusieurs bibliothèques et packages Python populaires, notamment NumPy, Pandas, Matplotlib et TensorFlow. D'autres bibliothèques et packages peuvent également être installés si nécessaire.

4.2.6 Pinterest

Pinterest est un réseau social et un moteur de recherche où des photos et des vidéos sont publiées, ainsi que des idées sont partagées.

4.2.7 Adobe scan

C'est une application portable qui permet à ses utilisateurs de convertir des documents et des documents en fichiers PDF de haute qualité et permet également des outils d'amélioration de la photo en corrigeant les couleurs et de clarifier les textes [50].

4.3 Définition de datasets utilisées

Dans cette étude, de nombreuses sources de données ont été utilisées.

4.3.1 Notre propre dataset

Dans ce travail, nous construisons notre propre dataset composée d'images de lettres et de chiffres manuscrites (26 lettres de a à z, et 10 chiffres de 0 à 9).

On peut la diviser en trois parties :

1. Partie 1

Afin de rapprocher cette dataset du domaine médical (où nous travaillons), nous avons embauché 10 personnes (quatre médecins et six pharmaciens) et quatre autres pour augmenter le nombre de travaux. Ainsi, chacun écrire manuellement les chiffres (0 - 9) et les lettres majuscules et minuscules (a - z).

- Nous avons collecté les données manuscrites.
- Nous avons pris des photos de ces écrits.
- Nous avons utilisé l'application Adobe Scan pour clarifier les images et assurer une bonne qualité d'image.
- Le nombre total de photos collectées est de 303 au format JPG.

2. Partie 2

- Nous avons utilisé l'application Pinterest pour collecter différentes images de chiffres et de lettres manuscrites.
- La plupart du temps, nous avons l'habitude de trouver toutes les lettres majuscules et minuscules (de a à z) et les chiffres (de 0 à 9) dans une seule image, pour cela nous découpons ces images pour extraire chaque lettre et chaque chiffre séparément.
- Le nombre d'images que nous avons collectées à partir du programme Pinterest a atteint 1118 images au format JPG.

Remarque : dans les parties 1 et 2, les dossiers des lettres contiennent des images des lettres manuscrites majuscules et minuscules

3. Partie 3

Pour assurer un bon entraînement pour nos modèles, nous avons augmenté la taille des données en ajoutant un ensemble de données composé également d'images de chiffres et de lettres manuscrites.

- Le nom de cet dataset est CHoiCe, il contient totalement 2811 images d'un format PNG. Divisé en 62 dossiers : 10 dossiers de lettres (de 0 à 9), 26 dossiers de lettres majuscules (de A à Z) et 26 dossiers de lettres minuscules (de a à z) [51]. La graphique à barres ci-dessous représente le nombre de données dans chaque catégorie.

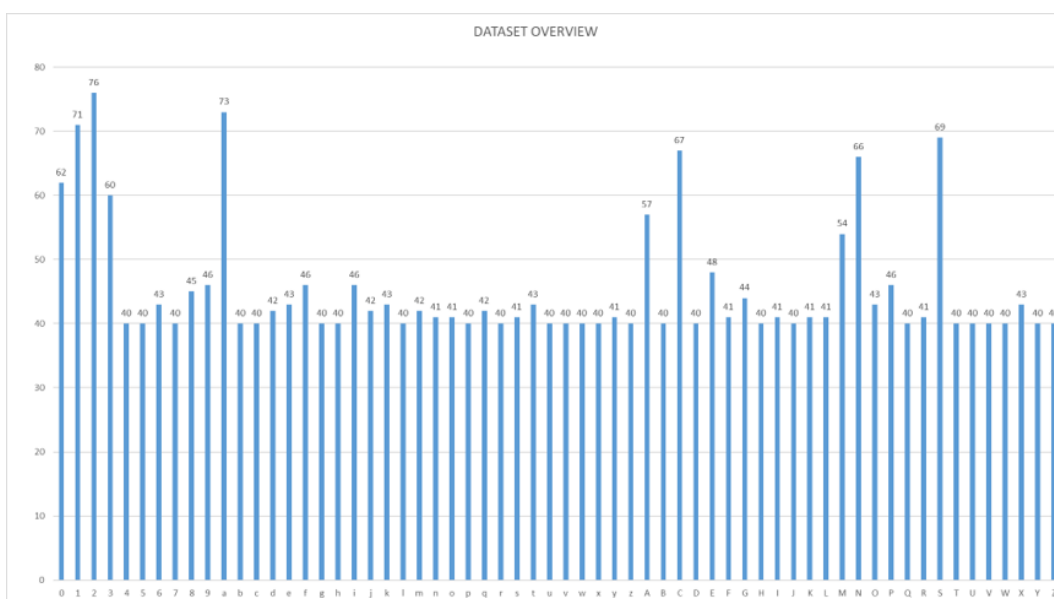


FIGURE 4.1 – Aperçu de l'ensemble de données pour chaque catégorie [51].

- On applique quelques modifications sur cette dataset telle que : les lettres majuscules et minuscules manuscrites dans le dataset "CHoiCe" se trouvent dans des dossiers séparés, nous avons donc mélangé tous les dossiers qui contiennent des lettres majuscules et minuscules pour une même lettre dans le même dossier.

Le diagramme à barres ci-dessous représente le nombre des images pour chaque partie de dataset mentionné précédemment.

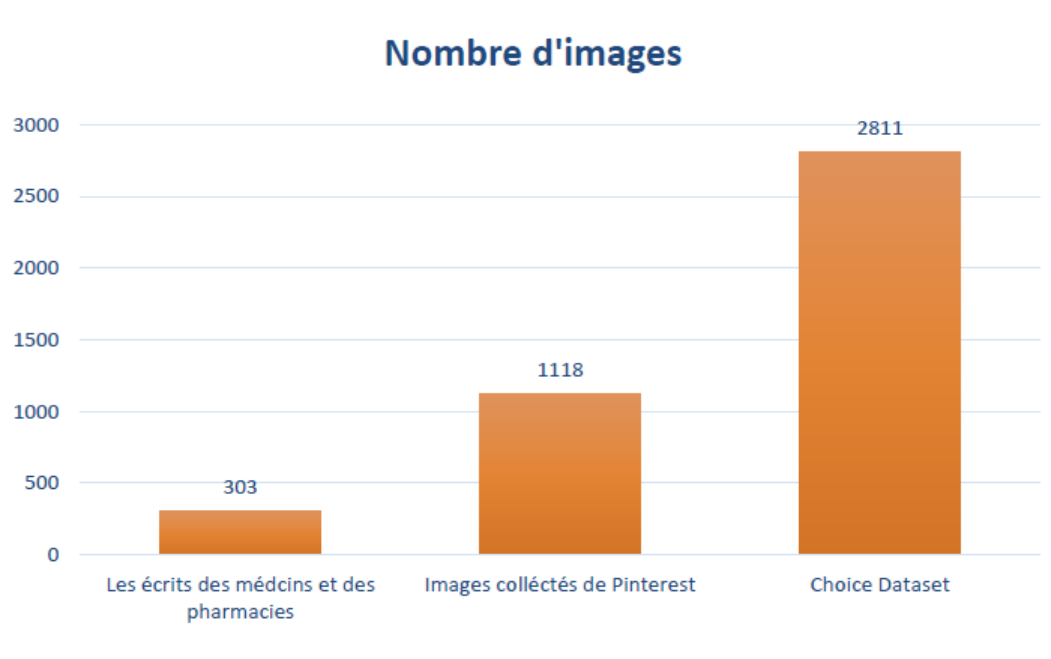


FIGURE 4.2 – Le nombre des images par partie.

Le diagramme circulaire ci-dessous montre le pourcentage de chacune des trois parties où chaque partie est représentée par une couleur particulière.

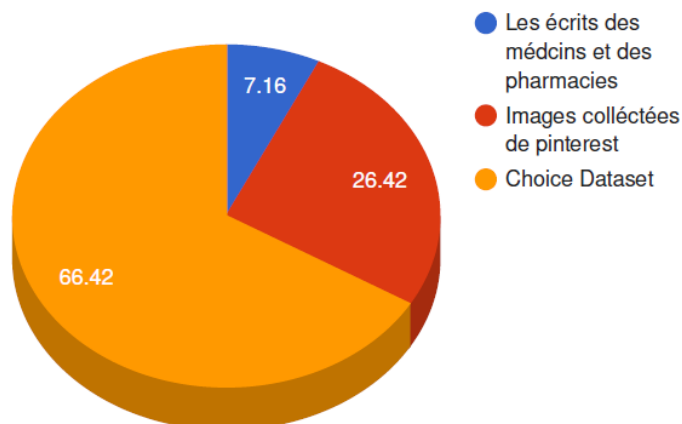


FIGURE 4.3 – Le pourcentage de chaque partie par rapport au total des données.

Nous avons également rassemblé toutes les parties ensemble, où chaque lettre représente la somme des trois parties respectives, et nous avons ainsi obtenu 36 dossiers qui diffèrent par le nombre d'images. Ensuite, nous avons effectué les étapes suivantes :

- Comme dernière étape dans la construction de notre ensemble de données, pour obtenir un dataset équilibré, nous avons choisi le plus petit dossier en termes de nombre d'images (dans notre cas 65 images).

- Nous avons modifié tous les dossiers pour avoir le même nombre d'images 65 images pour chaque dossier.
- Enfin, nous avons obtenu 36 dossiers équilibrés, où chaque dossier contient 65 images de letters et de chiffres manuscrits.

4.3.2 EMNIST ByClass

EMNIST est l'abréviation de "Extended Modified National Institute of Standards and Technology" et sous une forme plus abrégée "Expended MNIST". Ce dataset est développé et publié par le NIST en tant qu'extension de dataset MNIST original [52].

MNIST est un dataset populaire largement utilisé dans les domaines de l'apprentissage automatique, l'apprentissage profond et de la vision par ordinateur pour l'entraînement et l'évaluation des algorithmes de classification d'images. EMNIST a été créé pour répondre aux limites de dataset MNIST, qui ne contient que des images de chiffres manuscrits [52].

Le dataset EMNIST ByClass est l'un des sous-ensembles de dataset EMNIST qui vise à préserver la distribution originale des caractères dans le dataset spécial NIST 19, il contient des images des chiffres manuscrits et des lettres anglaises (majuscules et minuscules) avec un format d'image de 28x28 pixels divisés en 62 classes ou chaque classe représentant un caractère unique [52].

Le graphe ci-dessous clarifier la distribution des données pour chaque classe présente dans EMNIST ByClass.

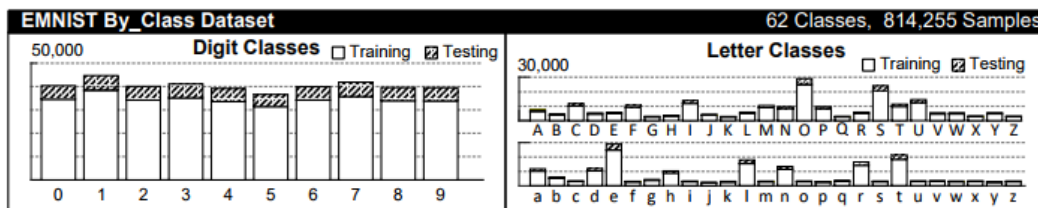


FIGURE 4.4 – la distribution des données par classe présente dans EMNIST ByClass [53].

4.3.3 IAM Handwriting

IAM c'est l'abréviation de "Institute for Applied Mathematics" est une dataset contient des images de formes de texte anglais manuscrit largement utilisé dans le domaine de la reconnaissance de l'écriture manuscrite et de la reconnaissance optique des caractères (OCR)[54]. Les échantillons de l'écriture manuscrite de IAM dataset sont fournis par 657 écrivains. IAM dataset est divisé en des datasets montré dans le tableau

suivant :

IAM dataset	le nombre des images
IAM forms	1539 pages de texte scanné
IAM sentenceces	5685 phrases isolées et étiquetées
IAM lines	13353 lignes de texte isolées et étiquetées
IAM words	115320 mots isolés et étiquetés

TABLE 4.1 – Le nombre des images dans chaque ensemble de données(dataset).

IAM Handwriting words est une partie de IAM Handwriting dataset qui contient seulement les images de mots (115320 mots). Avec un fichier texte (words.txt) qui sert de fichier de métadonnées contenant des informations sur chaque image de mot dans l'ensemble de données.

4.4 Implémentation et mise en œuvre

Dans ce travail, nous avons construit trois modèles différents en termes de techniques de réseaux de neurones et de différents dataset utilisées.

Le tableau suivant montre les techniques et les dataset utilisés dans chaque modèle.

	Modèle 1	Modèle 2	Modèle 3
dataset utilisé	Notre dataset	Notre dataset	EMNIST ByClass dataset
Les techniques utilisées	CNN	CNN et SVM	CNN

TABLE 4.2 – les techniques et les datasets utilisés dans chaque modèle.

Dans chacun des modèles ci-dessus, nous avons adopté la même méthodologie de conception, depuis la collecte des données jusqu'à l'obtention des résultats. Nous présentons cette méthodologie dans la figure suivante :

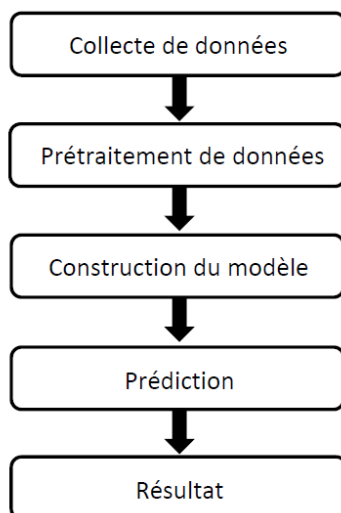


FIGURE 4.5 – La méthodologie de conception des modèles.

4.4.1 Modèle 1

1- Collecte et prétraitement de données

Comme le montre le tableau 4.2, dans notre premier modèle, nous avons utilisé notre propre ensemble de données, qui contient des images de chiffres et de lettres écrites à la main. Nous avons divisé cet ensemble de données en 70 % pour l'entraînement du modèle et 30 % pour les tests. Les images ont ensuite été converties à une taille de pixels 32 par 32. Quant aux couleurs, elles ont été converties en Grayscale.

2- Construction du modèle

Architecture de CNN

La figure ci-dessous illustre l'architecture complète du CNN utilisée dans ce modèle.

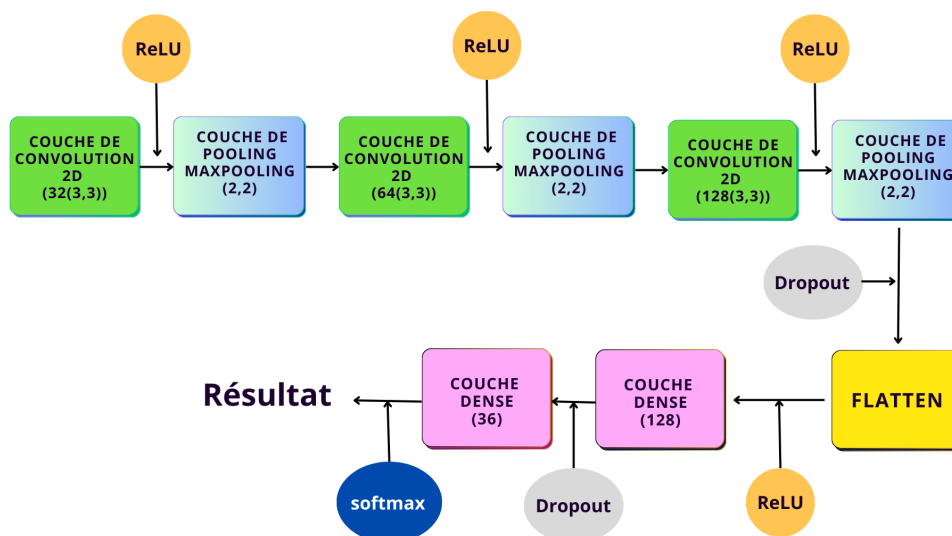


FIGURE 4.6 – L'architecture complète du CNN de modèle 1.

La première couche que nous avons ajoutée est une couche de convolution (deux dimensions) avec 32 filtres de taille (3,3). Cette couche est utilisée pour extraire les caractéristiques des images en appliquant des filtres à des zones de l'image. Nous avons utilisé la fonction d'activation "ReLU" (Rectified Linear Unit) pour introduire la non-linéarité dans le modèle. L'argument "padding" avec la valeur "same" pour extraire les caractéristiques extrêmes des bords et rendre les dimensions de sortie de la couche de convolution identiques à celles de l'entrée.

Ensuite, une couche de pooling (MaxPooling2D) est ajoutée pour réduire la dimension spatiale des caractéristiques extraites. Cette couche effectue un échantillonnage en sélectionnant la valeur maximale dans chaque région de pooling de taille (2,2).

Nous avons ajouté deux autres couches de convolution avec respectivement 64 et 128 filtres, suivies de couches de pooling pour extraire des caractéristiques plus abstraites à partir de l'image.

Puis, une couche de Dropout est ajoutée avec une valeur de 0,2, ce qui signifie que 20 % des neurones de la couche précédente seront aléatoirement désactivés pendant l'entraînement. Cela aide à prévenir le sur apprentissage en forçant le modèle à utiliser différentes combinaisons de neurones à chaque mise à jour des poids.

Ensuite, les caractéristiques extraites sont aplaties dans un vecteur unidimensionnel en utilisant la couche Flatten, afin de les passer à des couches entièrement connectées. Nous avons ajouté ensuite la couche Dense avec 128 neurones et utilisé la fonction d'activation "ReLU" pour introduire une non-linéarité. Enfin, la dernière couche que nous avons ajoutée est une couche Dense avec 36 neurones correspondant aux 36 classes de sortie possibles (10 chiffres et 26 lettres). La fonction d'activation softmax est utilisée pour obtenir une distribution de probabilités sur les différentes classes.

Segmentation

En raison de l'utilisation d'un dataset composé de chiffres et lettres, nous avons dû ajouter une partie dans le code qui divise l'image composée d'un mot entier en plusieurs lettres, puis la prédit. Cela se fait en construisant trois fonctions :

1. Première fonction

Cette fonction définit les contours des objets (lettres et chiffres) dans l'image et définit également les cases entourant chaque périmètre en utilisant la fonction `cv2.boundingRect()`, puis trie les contours et les rectangles englobants selon une méthode spécifiée, ce qui permet d'obtenir une liste de contours triés et de rectangles englobants correspondants.

2. Deuxième fonction

La deuxième fonction que nous avons utilisée est une fonction qui analyse une

image et extrait les caractères de l'image. Cela se fait en effectuant les étapes suivantes :

- Tout d'abord, elle lit l'image passée à la fonction en utilisant la bibliothèque OpenCV et la convertit en une image en niveaux de gris.
- L'application d'une technique appelée seuillage binaire inversé (Binary Inverse Thresholding) pour convertir l'image en niveaux de gris en une image binaire avec un contraste élevé.
- Ensuite, la fonction dilate les zones noires dans l'image binaire à l'aide de la technique de dilatation.
- Voici une illustration du résultat de chacune des deux fonctions de seuillage et dilatation.

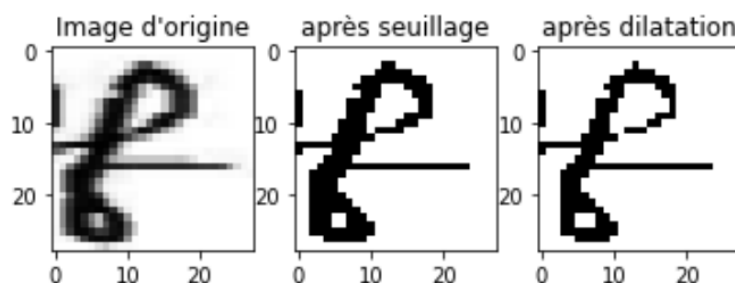


FIGURE 4.7 – Résultat du seuillage et de la dilatation.

- Elle utilise la fonction de détection de contours pour trouver les régions connectées dans l'image. Et après elle appelle à la première fonction qui trie les contours en fonction de leur longueur de gauche à droite et afin de déterminer le rectangle englobant de chaque région connectée et le dessine sur l'image d'origine.
- Elle modifie la forme de l'image pour correspondre à la forme des entrées attendues par le modèle d'apprentissage profond. Voici un exemple de la segmentation d'une image du nom d'un médicament.

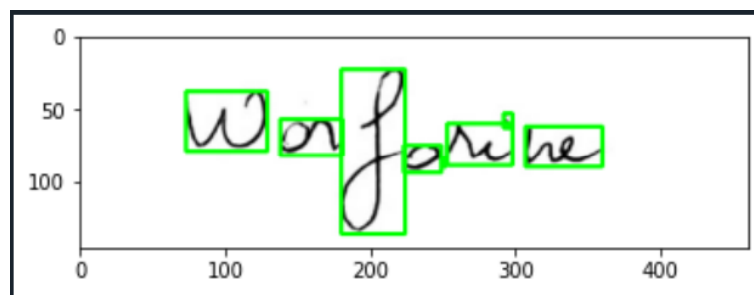


FIGURE 4.8 – Explication de la segmentation du nom du médicament.

- Enfin, en utilisant le modèle d'apprentissage profond préalablement chargé (modèle 1), elle prédit les caractères présents dans l'image. Les caractères prédits sont stockés dans une liste.

3. Troisième fonction

Cette fonction combine les caractères de la liste entrée comme paramètre dans l'ordre et donne le mot entier en sortie.

3- Prédiction

Dans cette étape, nous avons utilisé un ensemble de données composé des images de médicaments manuscrits que nous avons collectées auprès des mêmes médecins et pharmaciens que nous avons utilisés précédemment pour collecter des lettres et des chiffres manuscrits qui été mentionnés dans Partie 1 du titre 4.3.1.

4.4.2 Modèle 2

Dans ce modèle, nous avons suivi exactement les mêmes étapes que le premier modèle, depuis la collecte de données jusqu'au prétraitement des données et à la segmentation. La différence résidait dans l'architecture CNN, où nous avons ajouté une couche pour intégrer l'algorithme SVM à cette architecture. C'était comme suit :

- La première couche que nous avons ajoutée est une couche de convolution avec 32 filtres de taille (3,3), avec l'utilisation de la fonction d'activation "ReLU" et l'argument "padding" avec la valeur "same".
- Une couche de pooling est ajoutée pour réduire la dimension spatiale des caractéristiques extraites, avec une pooling de taille (2,2).
- Deux autres couches de convolution sont ajoutées avec 64 et 128 filtres respectivement, suivies de couches de pooling pour extraire des caractéristiques en plus abstraites à partir de l'image.
- Nous avons ajouté une couche Flatten. Puis, une couche de type Dense est ajoutée, comporte 128 unités, avec une activation "ReLU".
- Une couche de sortie de type Dense est ajouté avec 36 neurones correspondant aux 36 classes de sortie possibles, une activation softmax et un paramètre de régularisation (`tf.keras.regularizers.l2(0.01)`), ensuite le CNN est compilé avec une fonction de perte "squared hinge". Ce paramètre avec la fonction de perte "squared hinge" assurent que la couche finale qui est ajoutée est automatiquement une SVM couche, et alors la partie de classification multi classe dans ce modèle est fait par le SVM.

Et tout ce qui précède concernant l'architecture du réseau dans l'image ci-dessous peut être résumé.

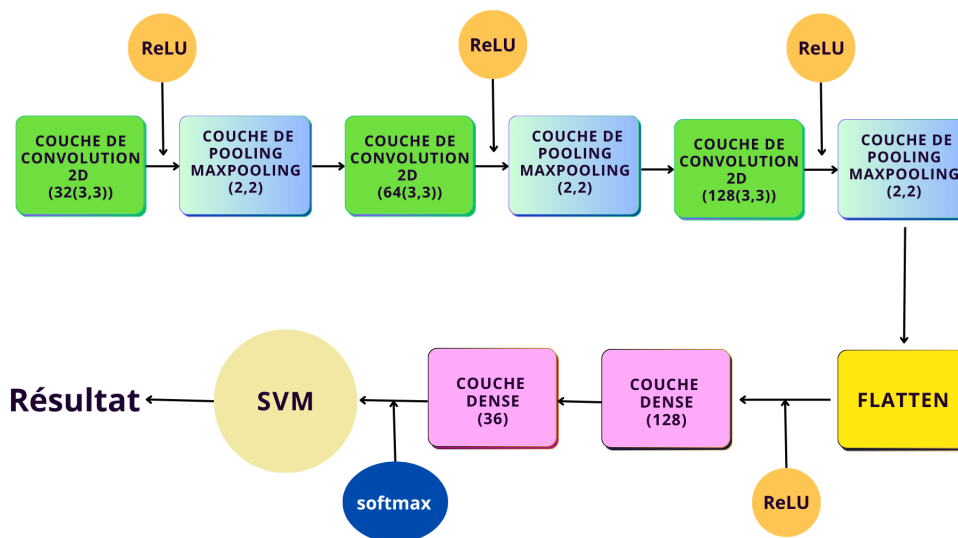


FIGURE 4.9 – L’architecture complète du CNN de modèle 2.

3- Prédiction

Pour faire la prédiction, nous avons utilisé le même ensemble composé d’images de médicaments manuscrites. Les deuxième et troisième fonctions sont appelées pour insérer l’image que nous voulons prédire afin d’obtenir le mot complet dans l’image.

4.4.3 Modèle 3

1- Collecte de données

Dans ce modèle, nous avons utilisé le dataset EMNIST ByClass, nous avons choisi l’environnement Python de Google Colaboratory pour charger ce dataset. Nous avons installé le package EMNIST. Afin d’utiliser ce package dans notre travail, nous avons importé les modules `extract-training-samples` pour l’entraînement et `extract-test-samples` pour le test. Ensuite, nous avons téléchargé le dataset EMNIST ByClass par les commandes `extract-training-samples("byclass")` qui extraire les échantillons d’entraînement de l’ensemble de toutes les classes de caractères disponibles dans le dataset EMNIST, et `extract-test-samples("byclass")` qui extraire les échantillons de test.

2- Prétraitement de données

Premièrement, nous avons normalisé les images téléchargées depuis EMNIST en les divisant par 255. La normalisation est effectuée en divisant chaque pixel de l’image par 255, ce qui ramène les valeurs de pixel dans la plage $[0,1]$, afin d’obtenir une représentation numérique des images qui facilite l’apprentissage du modèle. Puis, nous avons converti les étiquettes d’entraînement et les étiquettes de test en une représentation catégorique en utilisant la fonction `"to-categorical"`. Cette fonction convertit les valeurs qui représente

Les étiquettes d'origine (dans EMNIST ByClass dataset sont représentées de 0 à 61) en une représentation "one-hot encoded", où chaque étiquette est représentée par un vecteur binaire avec une seule valeur 1 et les autres valeurs 0.

3- Construction du modèle

Architecture CNN

La présente figure donne une représentation détaillée de l'architecture complète du CNN utilisée dans ce modèle.

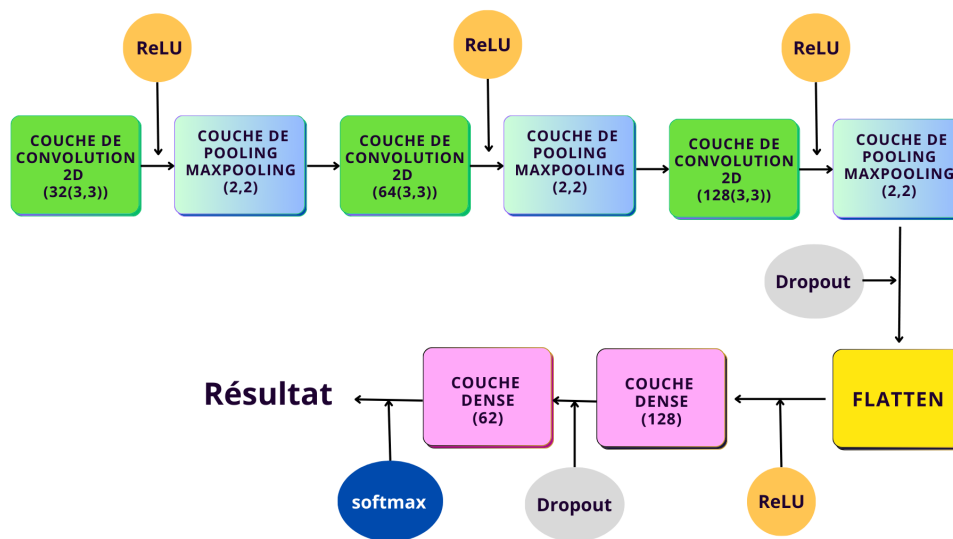


FIGURE 4.10 – L'architecture complète du CNN de modèle 3.

La première couche est une couche de convolution avec 32 filtres de taille (3,3), et les images d'entrée sont de taille 28x28 pour être compatibles avec les images du dataset utilisé.

La deuxième couche qui nous avons ajoutée est la couche de MaxPooling. Le filtre utilisé dans cette couche est de taille (2,2).

Nous avons ajouté puis, deux autres couches de convolution avec 64 et 128 filtres respectivement, suivies de couches de MaxPooling.

Ensuite, 25 % des neurones de la couche précédente seront aléatoirement désactivés pendant l'entraînement en ajoutant une couche de Dropout avec une valeur de 0,25.

Puis, une couche de Flatten. Et après, une couche Dense avec 128 neurones.

Une autre couche de Dropout avec la même valeur. Et Enfin, une couche Dense avec 62 neurones cette fois, correspondant aux 62 classes de sortie possibles existant dans le dataset EMNIST ByClass (10 nombres et 26 lettres en minuscule, et en majuscule).

Segmentation

Avec le même concept que nous avons expliqué dans la partie segmentation de modèle 1 4.4.1, nous prédisons les caractères présents dans l'image en utilisant ce modèle

d'apprentissage profond. Ensuite, nous combinons ces caractères prédits pour obtenir le mot complet dans l'image.

4- Prédiction

L'étape de prédiction se produit dans ce modèle exactement comme dans les premier et deuxième modèles, on insère l'image que l'on veut prédire, elle passe par les fonctions que l'on a définies, puis nous donne le mot complet comme résultat.

D'autre part, nous voulions essayer un dataset composé de mots, nous avons donc utilisé deux modèles, qui utilisent tous deux IAM words dataset 4.3.3, avec l'ajout des modifications sur ces deux modèles :

4.4.4 Approche de modèle 1

Tout d'abord, dans ce travail [55], le dataset IAM words est chargé, puis divisé en 90 % pour l'entraînement, 5 % pour la validation et 5 % pour les tests. Par la suite l'étape de prétraitement des données.

Ce modèle est basé sur les techniques CNN, RNN et CTC, de sorte que l'architecture de ce modèle est comme suit :

- Le modèle comprend deux couches de convolution. La première couche est composée de 32 filtres de taille (3,3), tandis que la deuxième couche est également constituée de 64 filtres de même taille. Chaque couche de convolution est suivie d'une couche de MaxPooling. Ensuite, une couche Dense avec 64 neurones est ajoutée.
- Deux couches LSTM avec 128, 64 unités de mémoire et un taux de dropout de 25% pour les deux couches, suivie d'une couche de sortie Dense avec une activation softmax.
- Enfin, une couche CTC (Connectionist Temporal Classification) pour calculer la perte CTC qui été conçue pour les tâches où nous avons besoin d'un alignement entre les séquences.

Notre modification

Nous avons ajouté des lignes de code pour charger l'image que nous souhaitons prédire. Ensuite, nous invoquons la fonction prédéfinie "preprocess-image" pour traiter cette image, suivie de l'affichage de celle-ci. Enfin, nous utilisons le modèle dédié à la reconnaissance des mots, afin de prédire le mot dans l'image.

4.4.5 Approche de modèle 2

Ce travail [56] définit une fonction pour charger les données du IAM Words dataset.

Pour l'entraînement, 95 % des échantillons du dataset IAM Words sont utilisés, tandis que les 5 % restants sont utilisés pour la validation. Tous les échantillons sont convertis en niveaux de gris (grayscale).

Il permet de lire les échantillons d'entraînement et de validation, de charger les images à partir de leurs chemins d'accès respectifs, et de fournir des lots d'images avec leurs textes de référence correspondants.

Ce modèle est basé sur les techniques de CNN, RNN et CTC. Son architecture est la suivante :

- Il comporte cinq couches de convolution. La première couche a 32 filtres de taille (5,5), puis une autre couche avec 64 filtres de taille (5,5), suivie de deux couches de convolution avec 128 filtres de taille (3,3). Ensuite, une couche avec 256 filtres de taille (3,3) est ajoutée. Dans toutes ces couches, la fonction d'activation ReLU est utilisée.
- Deux couches LSTM sont ensuite utilisées, chacune avec 256 unités de mémoire.
- Enfin, une fonction est mise en place pour calculer la perte CTC.

Notre modification

Le modèle a été entraîné sur l'ensemble des données d'entraînement pendant 15 époques. Étant donné que la taille du dataset utilisé est volumineuse, nous avons augmenté le nombre d'époques à 50. Ceci permet au modèle d'avoir davantage d'occasions d'apprendre et de s'ajuster aux données d'une part, et nous permet de faire une comparaison équitable avec les autres modèles (tous les autres modèles ont été entraînés sur 50 époques.) en termes de résultats d'autre part.

4.5 Résultats obtenus

Chaque modèle parmi les modèles que nous avons créés ou modifiés nous donne des résultats différents en termes de mots prédits, et donc en termes d'accuracy.

Accuracy est l'une des métriques permettant d'évaluer les modèles de classification. Elle a la définition suivante :

$$\text{Accuracy} = \frac{\text{Nombre de prédictions correctes}}{\text{Nombre total de prédictions}}$$

Nous avons modifié cette mesure pour servir notre travail afin que nous n'ayons pas besoin de calculer l'accuracy du modèle de classification, mais nous devons calculer l'accuracy des mots prédits, et donc nous avons défini l'accuracy comme suit :

$$\text{Accuracy} = \frac{\text{Nombre de caractères correctement prédits dans un mot}}{\text{Nombre total de caractères dans un mot}}$$

Nous donnons un exemple sur un mot présenté dans la figure suivante :



FIGURE 4.11 – Écriture manuscrite du nom du médicament bilastine.

Le mot prédit : bilaslime

La valeur d'accuracy obtenue pour ce mot est comme suit :

$$\text{Accuracy} = \frac{7}{9} \approx 0.78 = 78\%$$

Les résultats des modèles sont présentés dans le tableau ci-dessous :

	Modèle créé			Modèle modifié	
Les lodèles	Modèle 1	Modèle 2	Modèle 3	Modèle 1	Modèle 2
Accuracy	[0%, 22%]	[0%, 14%]	[0%, 21%]	[0%, 70%]	[0%, 91%]

TABLE 4.3 – La valeur d'accuracy pour chaque modèle.

Quelques mots prédits par chaque modèle sont présentés dans le tableau ci-dessous :

	Modèle créé			Modèle modifié	
Les mots	Modèle 1	Modèle 2	Modèle 3	Modèle 1	Modèle 2
bilastine	6b56a886	xh0aa8al	ktkesw	litartine	Bilaslime
evermictine	666x	8gwa	cvczzkw	Prbrmestin	Evernectine
diclofenac	8o888qr	dodam0x	dzekokz	iiarigran	diclefern

TABLE 4.4 – Des mots prédits par chaque modèle.

4.6 Discussion

Nous constatons que les résultats obtenus varient d'un modèle à l'autre, en fonction d'architecture utilisées dans chaque modèle et de dataset utilisés. Dans le premier modèle, construit avec l'architecture CNN et en utilisant nos propres dataset, nous avons remarqué que les chiffres étaient largement présents. Nous avons également observé des confusions

entre la lettre minuscule "b" et le chiffre "6" dans certains mots, en raison de l'écriture manuscrite de ces mots. L'accuracy la plus élevée que nous avons obtenue dans ce modèle était de 22 %.

En ce qui concerne le deuxième modèle, nous avons obtenu une accuracy maximale de 14 %. Malgré l'utilisation du même dataset, le changement dans l'architecture du deuxième modèle en combinant les techniques CNN et SVM a donné des résultats différents. Ainsi, nous pouvons conclure que la variation dans l'architecture des modèles a un impact significatif sur les résultats. Dans ce modèle, la lettre "w" était considérée comme la lettre "m" dans certains mots, et il y avait aussi une confusion entre les lettres "n" et "m" en d'autres termes.

En utilisant un dataset différent dans le troisième modèle, spécifiquement le EM-NIST byClass dataset, et en utilisant la même architecture que le premier modèle, nous avons obtenu une accuracy de 21 %. Les lettres les plus fréquentes dans ce modèle sont "k" et "z". Par exemple, les lettres minuscules "l" et "t" sont considérées dans ce modèle comme "K".

Dans ces trois modèles, notre ajout de la partie de segmentation qui divise le mot en lettres a impacté sur les résultats en considérant plusieurs lettres dans un mot comme une seule lettre, en raison des lettres superposées qui apparaissent dans l'écriture manuscrite des médecins.

Quant aux deux modèles que nous avons modifiés, qui sont basés sur les techniques CNN, RNN et CTC, en utilisant le même dataset dans les deux, qui est IAM Words Dataset mais avec une architecture différente (en termes de nombre de couches), il y avait des confusions entre les lettres "t" et "l" ainsi qu'entre les lettres "n" et "m". Cependant, malgré cela, nous avons obtenu de meilleurs résultats. L'accuracy la plus élevée que nous avons obtenue était de 70 % dans le premier modèle 4.4.4, et de 91 % dans le deuxième modèle 4.4.5.

Remarque : lors du calcul de la précision, un caractère est considéré comme correct uniquement s'il est à sa position correcte dans le mot. Par exemple, lorsque nous disons qu'un mot a obtenu 0 % de précision, cela signifie soit que le mot n'a pas du tout été reconnu, soit que tous les caractères ne sont pas à leur position correcte.

À partir des résultats de chaque modèle, nous concluons que le type et la taille du dataset et l'architecture de modèle ont un impact direct sur les performances du modèle et, par conséquent, sur les résultats obtenus.

4.7 Conclusion

Dans ce chapitre, nous avons présenté notre contribution par rapport à la reconnaissance de l'écriture manuscrite dans les prescriptions médicales, en utilisant des différents dataset, et différents modèles basés sur des techniques d'apprentissage profond.

Chaque modèle nous a fourni des prédictions de mots avec des valeurs d'accuracy fluctuantes, tels que la valeur la plus élevée que nous ayons obtenue est 91 %.

CONCLUSION GÉNÉRALE

La reconnaissance de l'écriture manuscrite des médecins est un problème courant auquel de nombreux chercheurs ont contribué. Dans ce mémoire de master, nous avons choisi le domaine de l'apprentissage profond pour trouver des solutions à ce problème. Nous avons utilisé différents dataset, dont certains que nous avons créés, contenant des images de caractères et de chiffres écrits par des médecins et des pharmaciens afin d'approximer les données au domaine médical, qui est le domaine dans lequel nous travaillons. Nous avons également collaboré avec les mêmes médecins et les mêmes pharmaciens pour créer un ensemble de données composé de noms de médicaments écrits à la main, destiné à tester ultérieurement les modèles. Nous avons également utilisé des ensembles de données préexistants tels que EMNIST byClass et IAM words dataset.

Sur la base de ces données, nous avons développé cinq modèles, dont trois ont été construits à l'aide de réseaux de neurones convolutifs (CNN) et de machines à vecteurs de support (SVM), et d'autre avec le CNN seulement, tandis que les deux autres ont été modifiés pour s'adapter au problème spécifique sur lequel nous travaillons. En conséquence, nous avons obtenu différents résultats en termes d'accuracy, pour que La meilleure accuracy que nous avons obtenue est de 91 %. Cette différence était due à la diversité et à la taille de dataset d'une part, et à l'architecture des modèles d'autre part.

Notre principale perspective est d'élargir l'ensemble de données composé de noms de médicaments écrits à la main par des médecins, afin de créer une plateforme et une application pour détecter et reconnaître les informations présentes dans les ordonnances et les dossiers médicaux. Par notre utilisation des datasets composé de caractères et de chiffres, ces modèles peuvent être généralisés à différents domaines, tels que la reconnaissance de l'écriture des étudiants dans des compétitions nationales telles que le BEM, BAC et concours de doctorat. De plus, à l'ère du numérisation, nos modèles peuvent être développés pour convertir des documents manuscrits en fichiers numériques.

BIBLIOGRAPHIE

- [1] t onec, “ Livre d’histoire et de géographie pour la troisième année du primaire.” <https://drive.google.com/file/d/13ULvTLXo-HZL2qTj2MA0Irl18AoPkrRgG/view>. consulté le 26-04-2023.
- [2] L-frii, “ La mauvaise écriture des médecins tuerait plus de 7 000 patients par an.” <https://l-frii.com/la-mauvaise-ecriture-des-medecins-tuerait-plus-de-7-000-patients-par-an/>. consulté le 26-04-2023.
- [3] LAROUSSE, “Définition : manuscrite.” <https://www.larousse.fr/dictionnaires/francais/manuscrit/49283>. consulté le 27-04-2021.
- [4] E. Ose Askvik, F. Van der Weel, and A. L. van der Meer, “The importance of cursive handwriting over typewriting for learning in the classroom : A high-density eeg study of 12-year-old children and young adults,” *Frontiers in Psychology*, 2020.
- [5] twinkl, “What is Handwriting?.” <https://www.twinkl.com/teaching-wiki/handwriting>. consulté le 28-04-2023.
- [6] M. Answer, “Quel style d’écriture est le meilleur pour les examens?.” <https://www.moreanswer.fr/quel-style-decriture-est-le-meilleur-pour-les-examens/>. consulté le 30-04-2023.
- [7] A. Jordanie, “Comment écrivez-vous D nealian?.” <https://celebrity.fm/fr/how-do-you-write-d-nealian/>. consulté le 29-04-2023.
- [8] MINIKAR, “Détermination des caractéristiques personnelles d’une personne à l’aide de l’analyse de l’écriture manuscrite. Graphologie.” <https://minikar.ru/fr/money/opredelenie-lichnostnyh-harakteristik-cheloveka-s-pomoshchyu-analiza>. consulté le 29-04-2023.
- [9] D. K. Sokol and S. Hettige, “Poor handwriting remains a significant problem in medicine,” *Journal of the Royal Society of Medicine*, 2006.

-
- [10] J. Caplan, “Cause of Death : Sloppy Doctors.” <https://content.time.com/time/health/article/0,8599,1578074,00.html>. consulté le 01-05-2023.
- [11] “Medical errors ’kill thousands’.” http://news.bbc.co.uk/2/hi/uk_news/682000.stm. consulté le 1-05-2023.
- [12] F. Charatan, “Family compensated for death after illegible prescription,” 1999.
- [13] S. R. et Peter Norvig, “Artificial intelligence a modern approach, third edition,” 2010.
- [14] A. M. Turing, “Computing machinery and intelligence (1950),” *The Essential Turing : the Ideas That Gave Birth to the Computer Age*, 2012.
- [15] “Turing Test.” https://science.jrank.org/programming/Turing_Test.html. consulté le 04-05-2023.
- [16] A. Crochet-Damais, “IA faible : définition et cas d’application.” <https://cutt.us/q5Xeg>. consulté le 04-05-2023.
- [17] I. Cloud, “What is artificial intelligence (AI)?.” <https://www.ibm.com/topics/artificial-intelligence>. consulté le 04-05-2023.
- [18] M. Romero, H. Aloui, L. Heiser, L. Galindo, and A. Lepage, *Un bref parcours sur les ressources, pratiques et acteurs en IA et éducation*. PhD thesis, Université Côte d’Azur, 2021.
- [19] A. Crochet-Damais, “Natural language processing (NLP) : définition et techniques.” <https://tinyurl.com/yc62wucu>. consulté le 05-05-2023.
- [20] IBM, “Qu’est-ce que la Computer Vision?.” <https://www.ibm.com/fr-fr/topics/computer-vision>. consulté le 05-05-2023.
- [21] SAP, “Qu’est-ce que le Machine Learning?.” <https://www.sap.com/france/products/artificial-intelligence/what-is-machine-learning.html>. consulté le 06-05-2023.
- [22] F. Sur., “Introduction à l’apprentissage automatique,” *École des Mines de Nancy*, 2021.
- [23] G. SAINT-CIRGUE, “Apprentissage Supervisé : Introduction.” <https://machinelearnia.com/apprentissage-supervise-4-etapes/>. consulté le 05-05-2023.
- [24] P. Artificielle, “Machine Learning pour débutant : Introduction au Machine Learning.” <https://penseeartificielle.fr/introduction-au-machine-learning/>. consulté le 06-05-2023.
- [25] A. Ezugwu, A. Shukla, M. Agbaje, O. Oyelade, A. José-García, and J. Agushaka, “Automatic clustering algorithms : a systematic review and bibliometric analysis of relevant literature,” *Neural Computing and Applications*, 2021.

-
- [26] A. Vidhya, “Introduction to Pseudo-Labeling : A Semi-Supervised learning technique.” <https://www.analyticsvidhya.com/blog/2017/09/pseudo-labeling-semi-supervised-learning-technique/>. consulté le 06-05-2023.
- [27] R. S. Sutton and A. G. Barto, “Reinforcement learning : An introduction,” *Robotica*, vol. 17, no. 2, pp. 229–235, 1999.
- [28] P. Baudin, “Machine learning et types d’apprentissage.” <https://blog.avanci.fr/comment-se-passe-lapprentissage-du-machine-learning>. consulté le 07-05-2023.
- [29] A. Vidhya, “Learn How to Use Support Vector Machines (SVM) for Data Science.” <https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>. consulté le 07-05-2023.
- [30] inovex, “Extensive Guide to Support Vector Machines.” <https://www.inovex.de/de/blog/support-vector-machines-guide/>. consulté le 07-05-2023.
- [31] D. Analytics, “Support Vector Machine (SVM) Python Example.” <https://vitalflux.com/classification-model-svm-classifier-python-example/>. consulté le 16-05-2023.
- [32] towardsdatascience, “Support Vector Machine — Introduction to Machine Learning Algorithms.” <https://shorturl.at/qEFT6>. consulté le 16-05-2023.
- [33] K. Gurney, *An introduction to neural networks*. 1997.
- [34] C. Hébert, “All diamond and diamond/carbon nanotube composite electrode for neurophysiological studies : form the material to the interfacial devices,” 11 2012.
- [35] medium, “Introducing Deep Learning and Neural Networks — Deep Learning for Rookies (1).” <https://shorturl.at/goN06>. consulté le 17-05-2023.
- [36] L. V. Fausett, *Fundamentals of neural networks : architectures, algorithms and applications*. 1993.
- [37] “Les réseaux de neurones.” https://ml4a.github.io/ml4a/fr/neural_networks/. consulté le 17-05-2023.
- [38] M. Negnevitsky, *Artificial intelligence : a guide to intelligent systems*. Pearson education, 2005.
- [39] medium, “Understanding recurrent neural network (rnn) and long short term memory(lstm).” <https://shorturl.at/dkDS6>. consulté le 19-05-2023.
- [40] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks : an overview and application in radiology,” *Insights into imaging*, vol. 9, 2018.

- [41] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [42] M. Yani, S. Irawan, and C. Setianingsih, “Application of transfer learning using convolutional neural network method for early detection of terry’s nail,” *Journal of Physics : Conference Series*, 05 2019.
- [43] freecodecamp, “Want to know how deep learning works? here’s a quick guide for everyone.” <https://shorturl.at/fpqrI>. consulté le 23-05-2023.
- [44] G. Pavithiran, S. Padmanabhan, N. Divya, V. R. Aswathy, P. IreneJerusha, and B. Chandar, “Doctors handwritten prescription recognition system in multi-language using deep learning,” *ArXiv*, 2022.
- [45] L. Fajardo, N. Sorillo, J. Garlit, C. Tomines, M. Abisado, J. M. Imperial, R. Rodriguez, and B. Fabito, “Doctor’s cursive handwriting recognition system using deep learning,” 11 2019.
- [46] R. Achkar, K. Ghayad, R. Haidar, S. Saleh, and R. A. Hajj, “Medical handwritten prescription recognition using crnn,” *2019 International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2019.
- [47] E. Hassan, H. Tarek, M. Hazem, S. Bahnacy, L. Shaheen, and W. H. Elashmwai, “Medical prescription recognition using machine learning,” *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0973–0979, 2021.
- [48] M. Kamal, F. Shaiara, C. M. Abdullah, S. Ahmed, T. Ahmed, and M. H. Kabir, “Huruf : An application for arabic handwritten character recognition using deep learning,” in *2022 25th International Conference on Computer and Information Technology (ICCIT)*, IEEE, 2022.
- [49] lebigdata, “Tensorflow : tout savoir sur la bibliothèque machine learning open source.” <https://www.lebigdata.fr/tensorflow-definition-tout-savoir>. consulté le 25-05-2023.
- [50] adobe. <https://www.adobe.com/acrobat/mobile/scanner-app.html>. consulté le 27-05-2023.
- [51] github, “A good choice : A complex handwritten character dataset.” <https://github.com/skyatmoon/CHoiCe-Dataset>. consulté le 26-05-2023.
- [52] G. Cohen, S. Afshar, J. Tapson, and A. van Schaik, “Emnist : an extension of mnist to handwritten letters,” 2017.
- [53] “Emnist (extended mnist).” <https://www.kaggle.com/datasets/crawford/emnist>. consulté le 27-05-2023.

- [54] C. Vision and A. Intelligence, “Iam handwriting database.” <https://fki.tic.heia-fr.ch/databases/iam-handwriting-database>. consulté le 28-05-2023.
- [55] keras, “Handwriting recognition.” https://keras.io/examples/vision/handwriting_recognition/. consulté le 6-06-2023.
- [56] Shanky, “Paper-evaluation-using-ai.” <https://github.com/shanky1947/Paper-Evaluation-using-AI>. consulté le 6-06-2023.