

**République Algérienne Démocratique Et Populaire
Ministère De L'enseignement Supérieur Et De La**

Recherche Scientifique

Université Larbi Ben M'hidi -Oum El Bouaghi-

Faculté Des Sciences Exactes Et Des Sciences De La Nature Et De La Vie

Département Mathématique et Informatique

N° D'ordre :

Série :

***MEMOIRE
EN VUE DE L'OBTENTION DU DIPLOME DE MASTER EN
INFORMATIQUE
OPTION : VISION ARTIFICIELLE***

THÈME

**Développement d'une application intelligente
pour la reconnaissance du Syndrome Main-Pied**

Présenté par

**Boushaba Imane
Riah Sara**

Dirigé par

Dr. Chebbout Samira

Devant le jury composé de

Présidente : Dr Maaziz Asma

Université Larbi Ben M'hidi Oum El Bouaghi

Directrice : Dr Chebbout Samira

Université Larbi Ben M'hidi Oum El Bouaghi

Examineur : Dr Ghoul Khaled

Université Larbi Ben M'hidi Oum El Bouaghi

Année universitaire 2022/2023

Dédicace

Nous tenons à saisir cette occasion et adresser nos profonds remerciements et nos profondes reconnaissances aux responsables et au personnel de l'Université L'arbi Ben M'hidi OEB et spécialement le département MI, nous désirons aussi remercier les professeurs de département MI de l'Université L'arbi Ben M'hidi OEB, qui nous a fourni les outils nécessaires à la réussite de nos études universitaires. Nous tiens à remercier spécialement notre encadreur Dr **Chebbout Samira**, qui a beaucoup contribué à ce travail par ses conseils et son suivi périodique.

Nous dédions ce mémoire

A nos parents pour leur amour inestimable, leur confiance, leur soutien,

Leurs sacrifices et toutes les valeurs qu'ils ont su m'inculquer.

A nos sœurs ainsi qu'à nos beaux-frères pour leur tendresse, leur complicité.

Boushaba Imane, Riah Sara

Résumé

Notre travail consiste à reconnaître le syndrome main-pied au niveau des mains/pieds des personnes atteintes d'un cancer et sous chimiothérapie. Dans un premier temps, nous avons procédé à la collection des images représentant ce syndrome. Le nombre des images collectées étant insuffisant, une augmentation hors ligne de ces images se fut nécessaire. Dans un deuxième temps, nous avons implémenté un réseau de neurones convolutif (CNN) capable de prédire si une personne est atteinte ou non du syndrome main pied. Nous avons développé une application bureau de santé intelligente qui offre des résultats de prédiction très encourageants.

Mots clés : Syndrome main pied, intelligence artificielle, apprentissage profond, réseau de neurone convolutif (CNN), augmentation de données.

Summary

Our work consists of recognizing hand-foot syndrome in the hands/feet of persons with cancer and under chemotherapy. First, we collected images of this syndrome. As the number of images collected was insufficient, an increase offline of these images was necessary. In a second phase, we set up a convolutive neural network (CNN) capable of predicting whether a person suffers from hand foot syndrome. We have developed a smart desktop health app that offers very encouraging prediction results.

Keywords: hand foot syndrome, artificial intelligence, deep learning, convolution neuran network (CNN), data augmentation.

المخلص

يتناسب عملنا مع سياق التعرف على متلازمة اليد والقدم في أيدي/أقدام الأشخاص المصابين بالسرطان وتحدث العلاج الكيميائي. أولاً، جمعنا صوراً لهذه المتلازمة. ولما كان عدد الصور التي تم جمعها غير كافٍ، كان من الضروري زيادة هذه الصور دون اتصال بالإنترنت. في المرحلة الثانية، أنشأنا شبكة عصبية ملتفة (CNN) قادرة على التنبؤ بما إذا كان الشخص يعاني من متلازمة اليد و القدم. لقد طورنا تطبيقاً صديقاً ذكياً لسطح المكتب يقدم نتائج تنبؤ مشجعة للغاية.

الكلمات المفتاحية: متلازمة اليد و القدم، التنبؤ ، الذكاء الاصطناعي ، التعلم الآلي ،

التعلم العميق ، زيادة البيانات ، الشبكة العصبية التلافيفية (CNN)

Table des figures.

Figure 1.1 cellule normale et cellule cancéreuse (Généralités sur le cancer, 2008).	4
Figure 1.2. Grade 1 du syndrome main-pied (Agnès TAUGOURDEAU Podologue, Docteur David SOLUB)	8
Figure 1.3. Grade 2 du syndrome main-pied (Agnès taugourdeau Podologue, Docteur David SOLUB)	8
Figure 1.4. Grade 3 du syndrome main-pied (Agnès TAUGOURDEAU Podologue, Docteur David SOLUB)	9
Figure 2.1. Les sous domaines de l'intelligence artificielle	13
Figure 2.2. Apprentissage supervisé (techno-science, 2021).....	14
Figure 2.3. Classification et regression.....	14
Figure 2.4. Apprentissage non supervisé (techno-science, 2021).....	16
Figure 2.5. Fonctionnement du deep Learning (futura-sciences , 2021).....	17
Figure 2.6. Mise en correspondance neurone biologique/neurone artificiel (Claude Touzet, 2016)...	19
Figure 2.7. Fonctionnement d'un neurone artificiel (Claude Touzet, 2016).....	20
Figure 2.8. Réseau de neurone monocouche (Perceptron) (Claude Touzet, 2016)	20
Figure 2.9 Structure d'un PMC composé de deux couches cachées (Claude Touzet, 2016)	21
Figure 2.10. Architecture d'un réseau convolutionnel.....	22
Figure 2.11. L'opération de convolution.....	22
Figure 2.12. L'opération de pooling.....	23
Figure 2.13. Fully Connected Layer (FC Layer) or classification.....	23
Figure 2.14. Processus d'apprentissage du CNN.....	23
Figure 2.15. Architecture VGGNet (Data Analytics, 2023).....	26
Figure 2.16. Types d'augmentation de données (Analytics Vidhya, 2021).....	29
Figure 3.1. Processus d'augmentation hors ligne des images collectées.....	35
Figure 3.2. Effet de rotation sur une image présentant le syndrome main pied.....	35
Figure 3.3. Effet du cisaillement sur une image présentant le syndrome main pied.....	36
Figure 3.4. Effet du Zoom sur une image présentant le syndrome main pied.....	36

Figure 3.5. Effet de réflexion horizontale sur une image présentant le syndrome main pied.....	37
Figure 3.6. Architecture du système de reconnaissance du syndrome HFS proposé.....	38
Figure 3.7. Architecture de modèle de réseau neuronal onvolutif.....	40
Figure 3.8. Courbe d'apprentissage (Accuracy/ Epoch).....	41
Figure 3.9. Courbe d'apprentissage (Loss/ Epoch).....	42
Figure 3.10. Matrice de confusion.....	43
Figure 3.11. Fenêtre d'accueil.....	44
Figure 3.12. Nouvelles prédictions.....	45
Figure 3.13. Prédiction positive image importée (main syndrome).....	45
Figure 3.14. Prédiction négative image importée (main non syndrome).....	46
Figure 3.15. Prédiction positive image prise (main syndrome).....	46
Figure 3.16. Prédiction négative image prise (main non syndrome).....	46

Table des Tableaux

Tableau 1. *Les types de réseaux de neurone convolutionnel CNN*.....25

Table des matières

Résumé	
Table des figures	
Table des Tableaux.....	
Introduction Général	

Chapitre 01 : Intelligence artificielle dans la santé

1. Introduction	1
2. Les applications de l'intelligence artificielle dans le domaine de la santé	2
2.1. Diagnostic médical.....	2
2.2. Détection et prédiction des maladies.....	2
2.3. Médecine personnalisée.....	2
2.4. Robotique médicale.....	2
2.5. Analyse de données de santé.....	2
3. L'IA et la prédiction des maladies	3
3.1. La prédiction des maladies cardiovasculaires	3
3.2. La prédiction du cancer du sein.....	3
3.3. La prédiction du cancer du poumon	3
4. L'IA et la détection des cancers	4
4.1. Les causes connues	4
4.2. Les traitements du cancer	6
4.2.1. La chirurgie	6
4.2.2. La radiothérapie.....	6
4.2.3. La chimiothérapie.....	6
4.2.4. Les autres traitements.....	7
4.2.5. Les thérapeutiques ciblées	7
5. Syndrome main-pied.....	7

5.1. Qu'est-ce que le syndrome main-pied ?.....	7
5.2. Les stades d'évolution du syndrome main-pied	8
5.3. Traitement symptomatique.....	9
5.3.1. Prévention.....	9
5.3.2. Traitement.....	10
6. Conclusion.....	11

Chapitre 02 : l'apprentissage automatique

1. Introduction	12
2. L'apprentissage automatique.....	12
2.1. Définition de l'apprentissage automatique.....	12
2.2. Types d'apprentissage automatique	13
2.2.1. L'apprentissage supervisé.....	13
2.2.2. L'apprentissage non-supervisé	15
2.2.3. L'apprentissage semi-supervisé	16
3. L'apprentissage profond « deep learning »	17
3.1. Applications du deep Learning.....	18
3.2. Les différents algorithmes de Deep Learning	18
3.3. Les réseaux neuronaux convolutifs (CNN).....	18
3.3.1. Algorithme de réseau de neurone.....	18
3.3.2. Structure d'un réseau de neurone artificiel.....	19
3.3.3. Fonctionnement d'un neurone artificiel.....	19
3.3.4. Type de Réseau de neurone.....	20
3.3.4.1. Réseau de neurone monocouche (Perceptron).....	20
3.3.4.2. Réseau de neurone multicouche.....	21
3.3.5. Les réseaux de neurones convolutionnels / Convolution neural network « CNN ».....	21
3.3.5.1. Définition d'un réseau de neurone convolutionnel.....	21
3.3.5.2. Les couches d'un réseau de neurone convolutionnel	22
3.3.6. Les types de réseau de neurone convolutionnel.....	25

3.4. Apprentissage par transfert.....	26
3.4.1. Utilité et fonctionnalités de l'apprentissage par Transfert.....	26
3.4.2. VGGNet Modèle de Transfert Learning	26
3.5. Augmentation de données	27
3.5.1. Définition de l'augmentation de données	27
3.5.2. Types d'augmentation de données	28
4. Conclusion.....	30

Chapitre 03 : Conception et implémentation de l'application SyndAP

1. Introduction	31
2. Les outils et les détails d'implémentation	32
2.1. Environnement de développement.....	32
2.1.1. Spyder	32
2.1.2. PySimpleGUI.....	32
2.2. Langage de programmation	33
2.2.1. Python.....	33
2.2.2. Pourquoi Python ?.....	33
3. Méthodologie de travail.....	34
3.1. Collection des images	34
3.2. Augmentation hors ligne des images collectée et création de la base d'image.....	35
3.3. Architecture du système de reconnaissance du syndrome HFS proposé	37
3.4. L'application SyndAp	44
4. Conclusion.....	47
Conclusion générale	48
Bibliography	49

Liste des abréviations

IA Intelligence Artificielle

CNN Convolutional Neural Network

FC Fully-Connected

RNN Recurrent Neural Network

RELU REctified Linear Unit

ANN Artificial Neural Network

Introduction générale

Le cancer est l'une des maladies les plus graves et répandues dans le monde, et il représente un défi majeur pour la médecine moderne. L'intelligence artificielle (IA) et ses sous-domaines, tels que l'apprentissage profond et les réseaux de neurones convolutifs (CNN), ont émergé comme des outils prometteurs dans la lutte contre le cancer. (LE CORGNE Aude ,2016)

L'IA utilise des algorithmes informatiques sophistiqués pour analyser de grandes quantités de données médicales, telles que des images médicales, des données génétiques et des dossiers médicaux, afin de détecter, diagnostiquer et prédire le cancer. La classification, qui consiste à catégoriser les données en différentes classes ou catégories, est un domaine clé de l'IA appliquée au cancer.

Les réseaux de neurones convolutifs (CNN) sont une architecture d'apprentissage profond spécialement conçue pour l'analyse d'images. Ils sont largement utilisés dans la classification des images médicales, y compris les scans, les radiographies et les images histologiques. Les CNN peuvent extraire des caractéristiques complexes et identifier des schémas dans les images qui sont souvent difficiles à détecter par les médecins humains. (Journal of Big Data ,2021)

L'apprentissage profond, qui fait partie de l'intelligence artificielle, permet aux modèles informatiques d'apprendre à partir de grandes quantités de données pour améliorer leurs performances au fil du temps. Les réseaux de neurones profonds peuvent être formés pour effectuer des tâches complexes de classification, de prédiction et de détection des cancers.

Lorsque les patients sont traités contre le cancer, ils peuvent éprouver certains effets secondaires liés aux traitements eux-mêmes. L'un de ces effets secondaires courants est le syndrome main-pied, également connu sous le nom de syndrome main-pied palmaire-plantaire.

Le syndrome main-pied se caractérise par des rougeurs, des gonflements, des douleurs et une sensibilité accrue au niveau des paumes des mains et de la plante des pieds. Ces symptômes peuvent rendre les activités quotidiennes, telles que la marche, la manipulation d'objets ou l'écriture, difficiles et inconfortables pour les patients. (Journal of Big Data ,2021)

Le cancer peut entraîner des effets secondaires tels que le syndrome main-pied, qui est une réaction cutanée spécifique à certains traitements contre le cancer. La prise en charge de ce syndrome fait partie intégrante de la prise en charge des patients atteints de cancer, dans le but de soulager les symptômes et de préserver leur qualité de vie pendant le traitement. De ce fait, l'objectif de ce travail consiste à développer un système intelligent qui puisse prédire la présence ou non du syndrome main-pied chez les personnes cancéreuses et sous chimiothérapie..

Ce mémoire est structuré selon trois chapitres distincts qui abordent différents aspects de la problématique étudiée. Voici un aperçu de chacun de ces chapitres :

Chapitre 1 : L'Intelligence artificielle dans la santé

Ce premier chapitre du mémoire porte sur le contexte médical dans lequel le syndrome main-pied est étudié. Le syndrome main-pied est un effet secondaire courant qui survient pendant le traitement du cancer. Afin de fournir un cadre complet, le chapitre commence par un aperçu du cancer, en soulignant ses causes possibles et en fournissant des informations sur les principaux signes avant-coureurs du cancer.

Ensuite, les différentes thérapies utilisées dans la lutte contre le cancer sont abordées, mettant en évidence les effets secondaires associés à ces thérapies. Parmi ces effets secondaires, une attention particulière est portée au syndrome main-pied.

Le chapitre se poursuit en expliquant en détail le syndrome main-pied, en le définissant et en décrivant ses caractéristiques. L'importance de comprendre le degré de développement du syndrome main-pied chez les patients cancéreux est soulignée, car il peut différer d'un individu à l'autre. Enfin, ce chapitre explore les développements de l'intelligence artificielle dans ce domaine.

Chapitre 2 : L'Apprentissage automatique

Le deuxième chapitre se concentre sur l'utilisation de l'apprentissage automatique pour classer les images. Il présente les concepts de base de l'apprentissage automatique et explore les différentes techniques et algorithmes utilisés pour la classification. Ce chapitre met également en évidence l'utilisation de l'apprentissage en profondeur, des réseaux de neurones convolutifs (CNN) et des techniques d'apprentissage par transfert pour améliorer la précision de la classification, étudier les bases de données et démontrer les techniques d'augmentation des données.

Chapitre 3 : Conception et implémentation de l'application SyndAP

Le troisième chapitre de ce mémoire se concentre sur l'application pratique du modèle de classification du syndrome main-pied en utilisant les techniques et les algorithmes discutés précédemment. Le processus de collecte et de traitement des images du syndrome main-pied est décrit en détail mettant en évidence la méthodologie utilisée. Après l'entraînement du modèle, ses performances sont évaluées à l'aide de mesure d'évaluation appropriées. Ces mesures aident à déterminer la précision et l'efficacité du modèle dans la classification du syndrome main-pied. Enfin, les résultats finaux obtenus par le modèle sont présentés et discutés.

Chapitre 1 :
Intelligence
artificielle dans la
santé

1. Introduction

L'intelligence artificielle (IA) a ouvert de nouvelles perspectives passionnantes dans le domaine de la santé. Grâce à ses capacités d'apprentissage automatique et d'analyse des données, l'IA peut être utilisée pour améliorer les diagnostics, optimiser les traitements, accélérer la découverte de médicaments et transformer la prestation des soins de santé.

L'IA dans la santé repose sur l'exploitation de vastes ensembles de données médicales, comprenant des dossiers de patients, des images médicales, des données génétiques et des données de recherche. Ces données sont analysées par des algorithmes d'apprentissage automatique et d'apprentissage profond, qui peuvent détecter des modèles, des tendances et des associations qui échappent souvent à l'œil humain.

L'un des domaines où l'IA a montré son potentiel est le diagnostic médical. En utilisant des modèles d'apprentissage automatique, l'IA peut analyser les symptômes, les antécédents médicaux et les résultats d'examens pour aider les médecins à poser des diagnostics plus précis et rapides. Elle peut également contribuer à la détection précoce de maladies, en identifiant les signes précurseurs et en alertant les professionnels de la santé.

L'IA peut également jouer un rôle clé dans la découverte de médicaments. Elle peut analyser de grandes quantités de données sur les composés chimiques, les interactions médicamenteuses et les essais cliniques pour accélérer le processus de recherche et de développement de nouveaux médicaments. Cela permet de réduire les délais et les coûts associés à la mise sur le marché de nouveaux traitements. L'IA peut aussi être utilisée pour détecter les effets secondaires de certains traitements pour certaines maladies.

Le cancer fait partie des maladies dans lesquelles l'intelligence artificielle a montré son potentiel en améliorant le diagnostic et la détection précoce de la maladie. Dans le chapitre, nous évoquerons cette maladie, ses caractéristiques, ses symptômes et ses divers effets secondaires du à la prise de médicaments pendant la période de traitement des personnes cancéreuses. Parmi ces symptômes, un accent particulier sera mis sur le syndrome main-pied. (WHO, 2021).

2. Les applications de l'IA dans le domaine de la santé

L'intelligence artificielle (IA) joue un rôle croissant dans le domaine de la santé et apporte des avancées significatives. Nous présentons quelques applications de l'intelligence artificielle dans le domaine de la santé

2.1. Diagnostic médical

L'IA peut aider les médecins à poser des diagnostics plus précis en analysant de grandes quantités de données médicales, y compris des images médicales, des dossiers médicaux électroniques et des informations génétiques. Les algorithmes d'apprentissage automatique/profond peuvent repérer des patterns et des anomalies qui échappent parfois à l'œil humain, permettant ainsi une détection précoce des maladies et une prise de décision éclairée.

2.2. Détection et prédiction des maladies

L'IA peut être utilisée pour détecter et prévoir certaines maladies. Par exemple, des modèles d'apprentissage automatique peuvent être développés pour prédire le risque de développer des maladies telles que le diabète, les maladies cardiaques, le cancer et les maladies neurologiques. Cela permet une intervention précoce et des mesures préventives pour réduire le risque de maladie.

2.3. Médecine personnalisée

L'IA peut être utilisée pour développer des traitements personnalisés en fonction des caractéristiques génétiques, des antécédents médicaux et des données cliniques d'un patient spécifique. Les algorithmes d'apprentissage automatique peuvent aider à identifier les traitements les plus efficaces et les plus adaptés à chaque individu, améliorant ainsi les résultats cliniques.

2.4. Robotique médicale

L'IA est utilisée dans la robotique médicale pour assister les chirurgiens lors d'interventions chirurgicales complexes. Les robots chirurgicaux peuvent être contrôlés par des algorithmes d'intelligence artificielle pour effectuer des mouvements précis et réduire les erreurs humaines. Cela permet des interventions plus précises, moins invasives et une récupération plus rapide pour les patients.

2.5. Analyse de données de santé

L'IA peut aider à analyser de grandes quantités de données de santé pour identifier des tendances, des schémas et des facteurs de risque. Cela peut contribuer à la recherche médicale, à l'identification de nouvelles découvertes et à l'amélioration des politiques de santé publique.

3. L'IA et la prédiction des maladies

L'IA intervient dans la détection et la prédiction de plusieurs maladies. A titre d'exemple, nous pouvons citer :

3.1 La prédiction des maladies cardiovasculaires

Les travaux de recherche de Mohammad Moshawrab sur la variabilité de la fréquence cardiaque ont récemment obtenu une attention internationale. Le candidat au doctorat en ingénierie et ses collègues ont décroché le prix du meilleur article lors d'une conférence internationale. Les travaux de recherche menés par l'étudiant de l'UQAR sont très prometteurs. « En se basant sur des modèles d'apprentissage automatique, nous avons analysé les paramètres de variabilité de la fréquence cardiaque. L'un de ces modèles a réussi à prédire l'apparition de futures maladies cardiovasculaire jusqu'à 12 mois en avance avec une exactitude de 91.8 % ». (UQAR-INFO, 2022)

3.2 La prédiction du cancer du sein

Les progrès technologiques ne cessent de faire des prouesses dans le domaine de la santé. Le Massachusetts Institute of Technology (MIT) aux Etats-Unis, le prouve en mettant au point une nouvelle intelligence artificielle, capable de détecter le cancer du sein, 5 années avant son apparition réelle. L'outil du MIT se base sur les résultats de plus de 60 000 patientes et près de 90 000 mammographies. Cette technique appelée deep learning permet d'analyser de grandes bases de données et de prédire un cancer du sein qui ne serait pas encore visible par l'œil humain. En effet, le logiciel compare votre image avec les images d'autres patientes dont on connaît le devenir sur une période de 5 ans. Ainsi, il voit si votre mammographie ressemble à telle ou telle autre patiente. (siecledigital-fr, 2019)

3.3 La prédiction du cancer du poumon

Des scientifiques de la Jameel Clinic du MIT ont développé un nouveau modèle d'intelligence artificielle capable de prédire le risque de cancer du poumon jusqu'à six ans à l'avance, grâce à un seul scanner à faible dose. Les chercheurs se sont dit surpris par le score élevé de prédiction du modèle, appelé Sybil, malgré l'absence de cancer visible. Cette innovation est cruciale car, aujourd'hui, il est extrêmement difficile de détecter le cancer du poumon à un stade précoce avec les scanners. (pourquoidocteur.fr, 2023)

4. L'IA et la détection des cancers

Causes, signes et traitement du cancer

Un cancer se définit par la perte de contrôle accidentelle de la régulation des cellules qui aboutit à leur prolifération anarchique (Généralités sur le cancer, 2008). Une tumeur est le résultat de la multiplication désordonnée des cellules d'un tissu ou d'un organe qui envahissent les tissus voisins en détruisant les capsules de séparation provoquant ainsi des métastases (Généralités sur le cancer, 2008).

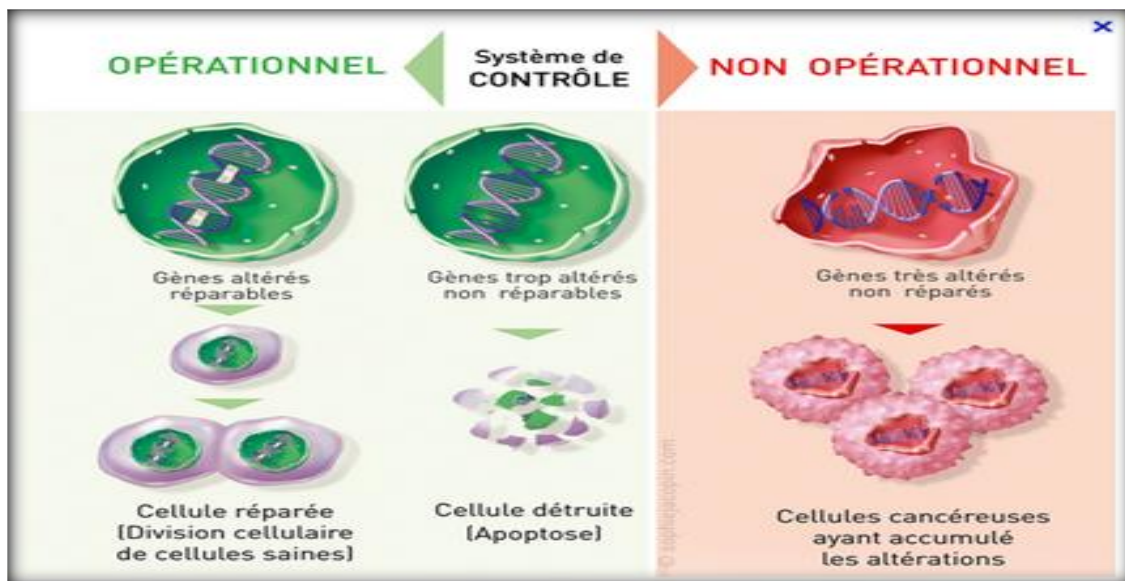


Figure 1.1. cellule normale et cellule cancéreuse (Généralités sur le cancer, 2008).

4.1. Les causes connues

Il est généralement difficile de savoir pourquoi une personne a développé un cancer. Les cancers sont en effet des maladies multifactorielles et l'ensemble des facteurs pouvant conduire à leur apparition ne sont pas connus. Toutefois, l'étude des mécanismes biologiques qui sont à l'origine des cancers (ou cancérogenèse) a permis d'identifier un certain nombre de facteurs qui augmentent le risque de cancer (Généralités sur le cancer, 2008).

Les Facteurs de risque professionnel

- ✓ Exposition à certaines substances (amiante, benzène, chlorure de vinyle, goudron...)
- ✓ Exposition aux radiations ionisantes (radiologues, militaires, centrales)
- ✓ Utilisation d'antimitotiques

Les facteurs de risque social

- ✓ Tabac, alcool, alimentation déséquilibrée
- ✓ Exposition solaire
- ✓ Sexualité à risque
- ✓ Stress

Les facteurs de risque viral

- ✓ Hépatite (foie)
- ✓ Papilloma-virus (col utérin)
- ✓ HIV (sarcome)

Les facteurs de risque génétique

- ✓ Maladies génétiques, certains cancers familiaux

Les autres facteurs

- ✓ Certains médicaments (distilbène)
- ✓ Pollution atmosphérique (centrales défectueuses, gaz des voitures ...)

Signes d'alarme des principaux cancers

➤ **Peau (mélanome)**

- ✓ Modification taille et couleur d'un grain de beauté, desquamation, suintement démangeaisons, douleur

➤ **Colon/rectum**

- ✓ saignement rectal ; selles sanglantes, modification du transit

➤ **Vessie / prostate**

- ✓ Mictions fréquentes, urines sanglantes, jet faible ou interrompu, douleurs dans la région pelvienne, aine ou en bas du dos

➤ **Sein**

- ✓ Épaississement œdème ou modification de la texture du tissu mammaire

➤ **Pancréas**

- ✓ Peu de signes précoces sauf douleur abdominale

- ✓ Prolongée et problèmes digestifs
- **Poumon**
- ✓ Toux lancinante, crachats sanglants, douleur thoracique, pneumonie ou bronchite récidivante
- **Gorge / larynx**
- ✓ Plaie buccale qui ne guérit pas, boule dans la gorge, voix rauque, difficulté à déglutir ou à mâcher
- **Utérus (et col)**
- ✓ Saignements vaginaux non liés au cycle menstruel et sécrétions inhabituelles
- **Leucémies et lymphomes**
- ✓ Fatigue, pâleur, perte de poids, infections à répétition, hématomes, ganglions enflés, hématomes, sueurs nocturnes, fièvre

4.2. Les traitements du cancer

La décision du choix du traitement est prise d'un commun accord entre le médecin et le patient (Le cancer, 2008).

4.2.1. La chirurgie

Pour les tumeurs solides, la chirurgie permet d'enlever le foyer principal, mais il persiste souvent une possibilité de maladie résiduelle c'est pourquoi on l'associe souvent à la chimiothérapie et/ou radiothérapie qui quelquefois sont même réalisées avant l'acte chirurgical Si la chirurgie doit être mutilante le patient doit être préparé psychologiquement et techniquement (stomathérapie).

4.2.2. La radiothérapie

L'utilisation actuelle de la radiothérapie haute énergie avec les accélérateurs de particules ainsi que le temps d'exposition et le nombre de séances dépendent du type de tumeur.

Effets secondaires de la radiothérapie

Fatigue, sclérose de la peau, des muscles, des lymphatiques avec œdèmes persistants, inflammation des muqueuses digestives, baisse de l'immunité en général.

4.2.3. La chimiothérapie

Elle représente le seul moyen théorique de contrôler la dissémination cancéreuse car véhiculée par le sang elle circulera dans tout le corps alors que la chirurgie et la radiothérapie ne s'adresse qu'à

des zones localisées Le principe consiste à introduire des substances chimiques capables d'interférer dans le métabolisme des cellules (cytostatiques).

Effets secondaires de la chimiothérapie

- ✓ Elle entraîne une fatigue générale
- ✓ Toxique également pour les cellules saines
 - alopecie
 - nausées, vomissements
 - toxicité du tube digestif (muqueuses)
 - toxicité de certains organes vitaux (cœur, foie, reins.)
 - toxicité médullaire (aplasie médullaire donc une baisse de l'immunité en général avec risque infectieux augmenté et besoins transfusionnels).

4.2.4. Les autres traitements

- ✓ Corticothérapie
- ✓ Hormonothérapie

4.2.5. Les thérapeutiques ciblées

- ✓ Les anti-angiogénèse, qui, par action angiostatique ou action angio-toxique, inhibent ou détruisent les vaisseaux tumoraux.
- ✓ Les anticorps monoclonaux dirigés contre les antigènes marqueurs spécifiques des cancers.

5. Le syndrome main-pied

5.1 Qu'est-ce que le syndrome main-pied ?

Le syndrome main-pied (Hand Foot Syndrom, en anglais) est un effet indésirable de certains traitements de chimiothérapie ou de certaines thérapies ciblées. Il se manifeste dans la paume des mains et sur la plante des pieds. Les symptômes apparaissent quelques semaines après le début du traitement. Il existe différents moyens de les soulager et dans tous les cas ils disparaissent après la fin du traitement. (roche.fr, 2021)

Le syndrome main-pied est une réaction inflammatoire à certains traitements de chimiothérapie qui fragilisent les micro-vaisseaux des mains et des pieds. Les premiers signes sont des rougeurs dans la paume des mains et sur la plante des pieds. Des fourmillements peuvent également être

ressentis. L'ampleur du syndrome main-pied dépend de chaque patient mais aussi de la posologie du traitement de chimiothérapie. (roche.fr, 2021)

5.2 Les stades d'évolution du syndrome main-pied

Selon son stade d'avancement le syndrome main-pied peut représenter un inconfort ou affecter l'aisance des manipulations quotidiennes et la marche. Trois grades peuvent être observés :

- **Le grade 1.** Il débute par des rougeurs, des gonflements, une sensibilité accrue de la peau, des fourmillements ou une sensation de brûlures. Ce sont des sensations désagréables mais elles n'affectent pas vos gestes quotidiens.



Figure 1.2. Grade 1 du syndrome main-pied (Agnès TAUGOURDEAU Podologue, Docteur David SOLUB)

- **Le grade 2.** La peau est extrêmement sèche et le gonflement s'accroît. Ces symptômes n'affectent pas vos activités quotidiennes.



Figure 1.3. Grade 2 du syndrome main-pied (Agnès taugourdeau Podologue, Docteur David SOLUB)

- **Le grade 3.** Des cloques, des crevasses et des œdèmes peuvent apparaître, associés à des douleurs s'apparentant à celles de piqûres ou de brûlures. La peau, à un stade plus avancé s'épaissit et présente une desquamation. À ce stade sévère, les gestes quotidiens sont difficiles, notamment la marche, qui peut devenir douloureuse, et les gestes fins de préhension. (roche.fr, 2021)



Figure 1.4. Grade 3 du syndrome main-pied (Agnès TAUGOURDEAU Podologue, Docteur David SOLUB)

5.3 Traitement symptomatique

5.3.1 Prévention

➤ Avant le début du traitement oncologique

- Une dizaine de jours avant le début du traitement, visite chez un podologue. Il peut être utile d'enlever en douceur et à titre préventif les callosités qui risquent de poser problème en cas de syndrome main-pied.
- Hydratation quotidienne et généreuse de la peau. Demandez conseil à votre cancérologue pour choisir un type de crème adapté. (roche.fr, 2021)

➤ Pendant l'administration des cures de chimiothérapie

- Port de gants et de pantoufles réfrigérés pendant chaque cure de chimio. Le froid contracte les petits vaisseaux sanguins et diminue ainsi l'exposition locale au médicament. Parlez-en à votre cancérologue pour savoir si ce type de mesure préventive est envisageable dans votre cas. (roche.fr, 2021)

➤ **Pendant toute la période du traitement oncologique**

- Eviter les chaussures serrantes et les hauts talons.
- Eviter les vêtements, les bandages ou les bijoux serrant.
- Pas d'exposition des mains et des pieds aux ultraviolets (soleil et banc solaire).
- Eviter l'exposition des mains et des pieds à la chaleur et au froid intenses.
- Utiliser des gants et des pantoufles pour protéger les extrémités.
- Utiliser un savon surgras pour la toilette quotidienne.
- Utiliser généreusement une crème hydratante adaptée. (roche.fr, 2021)

5.3.2 Traitement

Si un syndrome main-pied se déclare, un certain nombre de mesures permettent de soulager les symptômes :

- Eviter de mettre la peau en contact avec des sources de chaleur. Des bains de pieds et de mains à l'eau fraîche peuvent soulager, tout comme un sac de glaçons ou le port de semelles spéciales réfrigérées.
- Eviter toutes les irritations de la peau par frottement (vêtements, chaussures, bandages pansements ou bijoux serrant).
- Eviter de solliciter trop violemment vos mains et vos pieds (éviter la marche pieds nus, le jogging, ou les activités qui supposent d'agripper fermement un objet).
- Au moment du coucher, mettre une couche épaisse de crème hydratante sur la paume des mains et la plante des pieds puis couvrir les mains et les pieds avec des gants et des chaussettes lâches en coton.

Si ces mesures se révèlent insuffisantes, l'ultime recours consiste à arrêter ou à modifier le traitement cancérologique en cas de syndrome main-pied particulièrement sévère. On évitera cependant dans toute la mesure du possible d'en arriver là et le malade ne doit en aucun cas modifier de sa propre initiative une chimiothérapie ou un traitement ciblé. N'attendez pas que le problème s'aggrave. Dès l'apparition des premiers symptômes (rougeurs, gonflements, picotements, douleurs...) parlez-en sans hésiter à votre cancérologue. Il pourra envisager avec vous les différentes possibilités pour vous soulager avant de décider, si nécessaire, une adaptation de votre traitement. (roche.fr, 2021)

6. Conclusion

Le cancer demeure une maladie grave qui affecte divers segments de la société, y compris les enfants, les jeunes et les adultes, quel que soit leur sexe. La prévalence croissante du cancer a motivé des efforts pour intégrer différentes avancées technologiques dans le domaine médical, afin de trouver des solutions pour prévenir, diagnostiquer et traiter cette maladie dévastatrice.

L'intelligence artificielle (IA) joue un rôle de plus en plus crucial dans la médecine moderne. Elle offre des opportunités prometteuses pour la prédiction de diverses maladies, y compris le cancer, en utilisant des algorithmes sophistiqués pour analyser de grandes quantités de données médicales. Grâce à l'apprentissage automatique et à l'analyse de données, l'IA peut identifier des schémas et des caractéristiques subtiles dans les données médicales, permettant ainsi de détecter des signes précoces de cancer et de prédire les risques individuels.

L'IA est également utilisée pour accélérer les méthodes de diagnostic du cancer. Les techniques d'imagerie médicale, telles que la tomodensitométrie (CT), l'imagerie par résonance magnétique (IRM) et la mammographie, peuvent être améliorées grâce à l'IA. Les algorithmes d'IA peuvent aider à interpréter les images médicales, à détecter les anomalies et à assister les radiologues dans leurs diagnostics.

En combinant l'IA avec d'autres technologies, telles que la génomique et la médecine de précision, des progrès significatifs sont réalisés dans la lutte contre le cancer. Ces avancées permettent une détection précoce plus précise, des traitements personnalisés basés sur les caractéristiques génétiques des patients, et une meilleure compréhension des mécanismes sous-jacents du cancer.

Il convient de souligner que l'IA ne remplace pas les professionnels de la santé, mais elle constitue un outil puissant pour les soutenir dans leur pratique médicale. L'IA peut aider à améliorer l'efficacité, la précision et l'accessibilité des soins de santé, contribuant ainsi à la lutte contre le cancer et à l'amélioration des résultats pour les patients.

Chapitre 2

L' apprentissage automatique

1. Introduction

Dans certains cas, il peut être difficile de trouver une solution exacte à un problème donné en écrivant un programme informatique classique. C'est là que l'apprentissage automatique entre en jeu. L'apprentissage automatique, ou machine learning, est une approche qui permet à un programme informatique d'apprendre à partir de données et d'expériences passées, plutôt que d'être explicitement programmé pour résoudre un problème.

L'idée centrale de l'apprentissage automatique est de développer des algorithmes capables d'identifier des modèles et des relations dans les données, afin de prendre des décisions ou de faire des prédictions. Pour cela, on utilise souvent un ensemble de données appelé "dataset" qui contient des exemples du problème considéré, accompagnés d'annotations ou de résultats attendus.

L'algorithme d'apprentissage automatique apprend à partir de ces données en ajustant ses paramètres internes, de manière à minimiser une fonction de coût ou d'erreur. Plus l'algorithme est exposé à des exemples diversifiés et représentatifs, plus il a de chances d'apprendre des modèles précis et de généraliser ces connaissances à de nouvelles données.

L'apprentissage automatique peut être utilisé dans une grande variété de domaines et de problèmes, tels que la reconnaissance d'images, la traduction automatique, la prédiction de maladies, l'analyse de données, la recommandation de produits, et bien d'autres encore. Il permet de résoudre des problèmes complexes pour lesquels les solutions traditionnelles sont difficiles à programmer.

Il est important de noter que l'apprentissage automatique ne se limite pas à un seul type d'algorithme. On distingue notamment l'apprentissage supervisé, où l'algorithme apprend à partir de données étiquetées, et l'apprentissage non supervisé, où l'algorithme cherche à découvrir des structures et des regroupements dans les données sans avoir de résultats attendus. (Ah-Pine, 2019/2020).

2. L'apprentissage automatique

2.1. Définition de l'apprentissage automatique

L'apprentissage automatique est une application d'intelligence artificielle (IA) (Voir figure 1), qui permet aux systèmes d'apprendre et de s'améliorer automatiquement à partir de l'expérience elle-même sans être explicitement programmée. L'apprentissage automatique se concentre sur le

développement de programmes informatiques qui peuvent accéder à des données et les utiliser pour apprendre par eux-mêmes (Data Science Team, 2020).

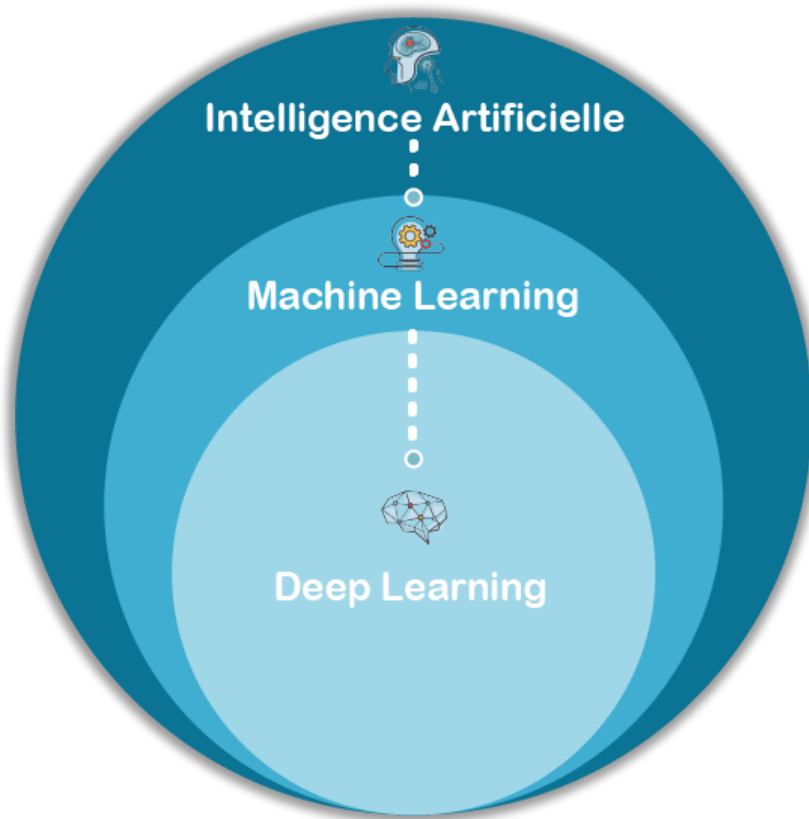


Figure 2.1. Les sous domaines de l'intelligence artificielle (Data Scientist ,2022)

2.2. Types d'apprentissage automatique

On peut citer trois types d'algorithmes d'apprentissage automatique

2.2.1. L'apprentissage supervisé

On parle d'apprentissage supervisé lorsque l'on dispose de données d'entraînement étiquetées, c'est à dire dont on connaît la sortie voulue. En notant les N entrées x_i et les sorties cibles associées, on dispose de l'ensemble de données $D = \{x_i, y_i\} \in [1, N]$. L'objectif est d'entraîner le modèle choisi pour qu'il puisse prédire correctement la sortie pour des entrées non étiquetées. (Ludovic DE MATTEIS Steeven JANNY, 2022)

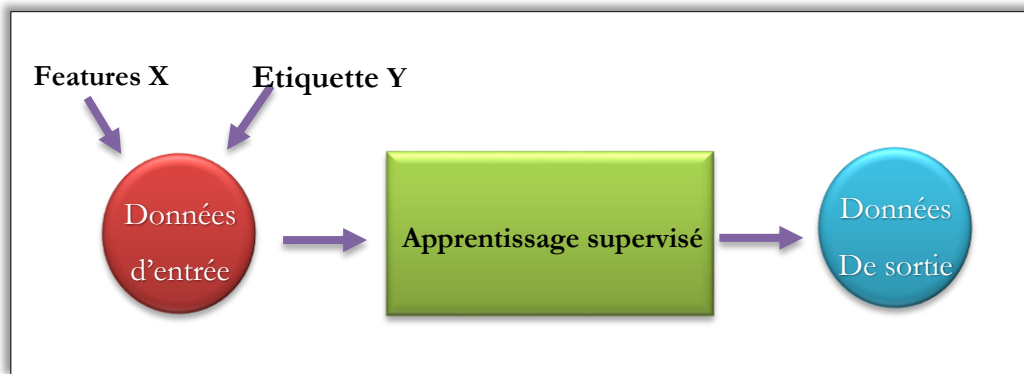


Figure 2.2. Apprentissage supervisé (techno-science, 2021).

L'apprentissage supervisé est généralement utilisé pour de la régression ou de la classification

- ✓ **La classification :** La classification est une tâche consistant à choisir une classe (valeur) parmi toutes celles possibles Exemple : Un algorithme prédisant le chiffre manuscrit sur l'image d'entrée ou un algorithme classifiant une tumeur comme « bénigne » ou « maligne ». (Ludovic DE MATTEIS Steeven JANNY, 2022).
- ✓ **La régression :** La régression est utilisée lorsque la sortie à prédire peut prendre des valeurs continues, il s'agit d'une variable réelle. Exemple : un algorithme prédisant la consommation électrique d'une installation ou un algorithme prédisant le cours d'actions en bourse. (Ludovic DE MATTEIS Steeven JANNY, 2022).

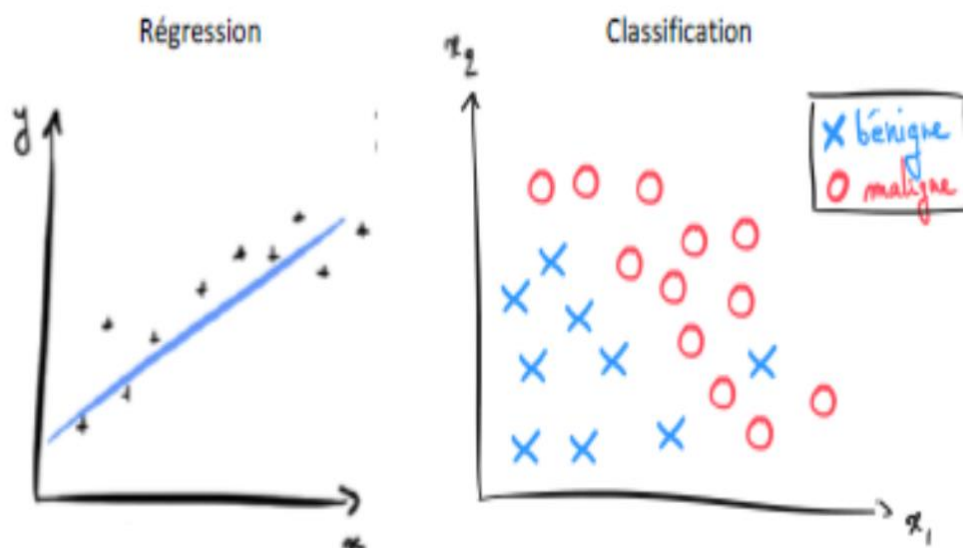


Figure 2.3. Classification et régressions (openclassrooms.com,2023)

Les algorithmes les plus célèbres utilisés dans cette approche sont les suivants (GAËL, 2019) :

- **La méthode des k plus proches voisins** : cet algorithme consiste à essayer différentes valeurs de K pour obtenir la séparation la plus satisfaisante.
- **Naïve Bayes** : est un classifieur assez intuitif à comprendre. Il se base sur le théorème de Bayes des probabilités conditionnelles, et il suppose que les variables sont indépendantes entre elles. Cela permet de simplifier le calcul des probabilités.
- **Les arbres de décision** : un arbre de décision sert à classer les futures observations, sachant qu'un corpus d'observations est déjà étiqueté.
- **Les Forêts Aléatoires** : cet algorithme fonde sur les arbres de décision, un modèle construit par de multiples arbres de décisions.
- **Régression Logistique** : l'algorithme de régression logistique consiste à trouver les meilleurs coefficients pour minimiser l'erreur entre la prédiction faite pour des destinations visitées et la vraie étiquette donnée (e.g. bon, mauvais etc.).
- **Les réseaux de neurones** : Les réseaux neuronaux sont un modèle informatique qui partage certaines propriétés avec le cerveau humain, dans lequel de nombreuses unités simples travaillent en parallèle sans centralisation, ils permettent de trouver des patterns complexes dans les données, il se compose de valeurs d'entrées, poids, fonction de transfert et une valeur de sortie.

2.2.2. L'apprentissage non-supervisé

On parle d'apprentissage non supervisé si les données ne sont pas étiquetées. On dispose donc de données d'entrée dont on ne connaît pas la sortie associée. L'ensemble de données est donc $D = \{ \} \in [1, N]$ et l'objectif du système est d'identifier des caractéristiques communes aux données d'entraînement.

L'apprentissage non-supervisé est principalement composé des algorithmes de regroupement (clustering). Ces algorithmes cherchent à séparer les données d'entrée en un nombre donné de groupes. Chaque élément du groupe doit avoir des caractéristiques proches de celles des éléments

du même groupe mais des caractéristiques relativement éloignées de celles des autres groupes. Ces algorithmes regroupent donc les entrées en familles afin de les étiqueter automatiquement.

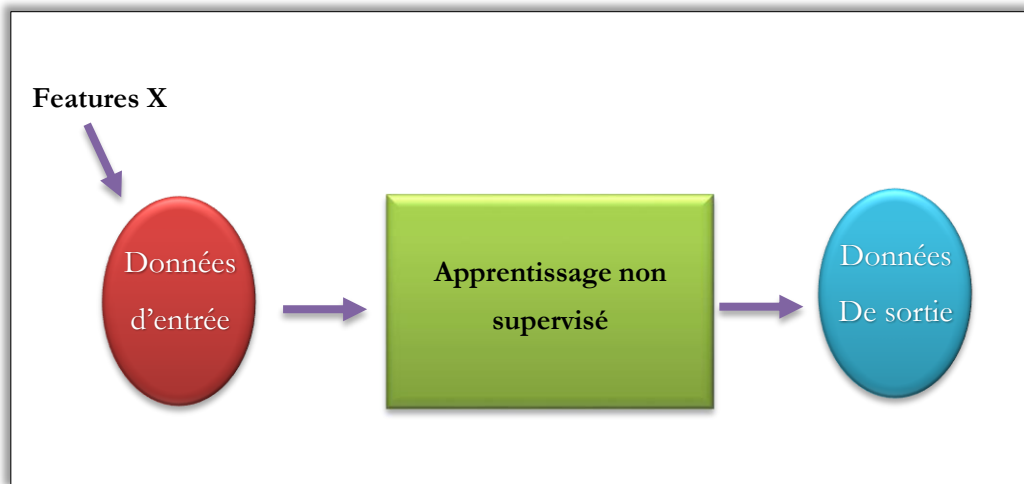


Figure 2.4. *Apprentissage non supervisé (techno-science, 2021).*

Les algorithmes les plus célèbres utilisés dans cette approche sont (Issarane, 2019) :

- **K-Moyenne** : c'est un algorithme de regroupement (clustering) qui regroupe dans les même cluster (groupes) les données similaires (qui se ressemblent). Il utilise un raffinement itératif pour produire un résultat final.
- **Analyse de classification hiérarchique (HCA)** : la mise dans un cluster hiérarchique est similaire à la mise dans un cluster normal, sauf que dans ce cas nous souhaitons mettre en place une hiérarchie des clusters. Cela peut s'avérer très important surtout quand nous désirons une flexibilité par rapport au nombre de clusters voulu.

2.2.3. L'apprentissage semi-supervisé

Effectué de manière probabiliste ou non, il vise à faire apparaître la distribution sous-jacente des « exemples » dans leur espace de description. Il est mis en œuvre quand des données (ou « étiquettes ») manquent. Le modèle doit utiliser des exemples non-étiquetés pouvant néanmoins renseigner. Exemple : En médecine, il peut constituer une aide au diagnostic ou au choix des moyens les moins onéreux de tests de diagnostics.

3. L'apprentissage profond « deep learning »

L'apprentissage profond est un ensemble de techniques d'apprentissage automatique qui a permis des avancées importantes en intelligence artificielle dans les dernières années. Dans l'apprentissage automatique, un programme analyse un ensemble de données afin de tirer des règles qui permettront de tirer des conclusions sur de nouvelles données. L'apprentissage profond est basé sur ce qui a été appelé, par analogie, des « réseaux de neurones artificiels », composés de milliers d'unités les « neurones » qui effectuent chacune de petites opérations simples. Les résultats d'une première couche de « neurones » servent d'entrée aux calculs d'une deuxième couche et ainsi de suite. Par exemple, pour la reconnaissance visuelle, des premières couches d'unités identifient des lignes, des courbes, des angles... des couches supérieures identifient des formes, des combinaisons de formes, des objets, des contextes... Les progrès de l'apprentissage profond ont été possibles notamment grâce à l'augmentation de la puissance des ordinateurs et au développement de grandes bases de données « big data ».

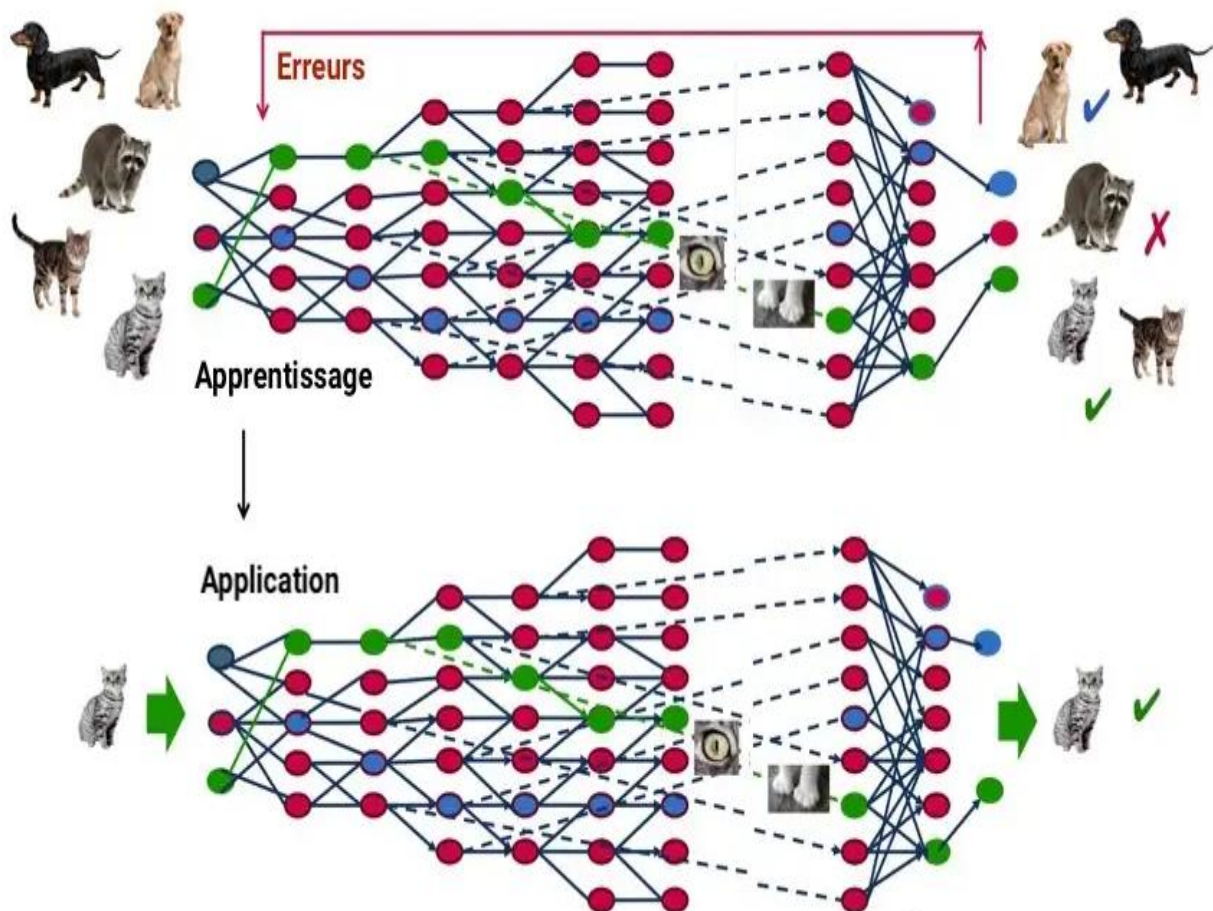


Figure 2.5. Fonctionnement du deep Learning (futura-sciences , 2021)

3.1. Applications du deep Learning

Le deep Learning est utilisé dans de nombreux domaines : (*futura-sciences, 2021*)

- ✓ Reconnaissance d'image
- ✓ Traduction automatique
- ✓ Voiture autonome
- ✓ Diagnostic médical
- ✓ Recommandations personnalisées
- ✓ Modération automatique des réseaux sociaux
- ✓ Prédiction financière et trading automatisé
- ✓ Identification de pièces défectueuses
- ✓ Détection de malwares ou de fraudes
- ✓ chabots (agents conversationnels)
- ✓ Exploration spatiale
- ✓ Robots intelligents.

3.2. Les différents algorithmes de Deep Learning

Il existe de nombreux algorithmes de ce type et parmi eux

- Réseaux neuronaux convolutifs (CNN)
- Réseaux neuronaux récurrents (RNN)
- Réseaux de fonction de base radiale (RBFN)
- Réseaux de mémoire à long et court terme (LSTM)
- Réseaux adversariaux génératifs (GAN)
- Machines de Boltzmann restreintes (RBM)...

3.3. Les réseaux neuronaux convolutifs (CNN)

3.3.1. Algorithme de réseau de neurone

Dans le domaine des technologies de l'information, un réseau de neurones est un système logiciel et / ou matériel qui imite le fonctionnement des neurones biologiques. Les réseaux neuronaux, aussi appelés réseaux de neurones artificiels (RNA ou ANN en anglais), font partie des technologies d'apprentissage profond ou « Deep Learning », couvertes également par l'intelligence artificielle (IA). (Tech Target, 2022)

3.3.2. Structure d'un réseau de neurone artificiel

La figure 7 montre la structure d'un neurone artificiel. Chaque neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones amonts. A chacune de ces entrées est associée un poids w abréviation de weight (poids en anglais) représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals. A chaque connexion est associée un poids.

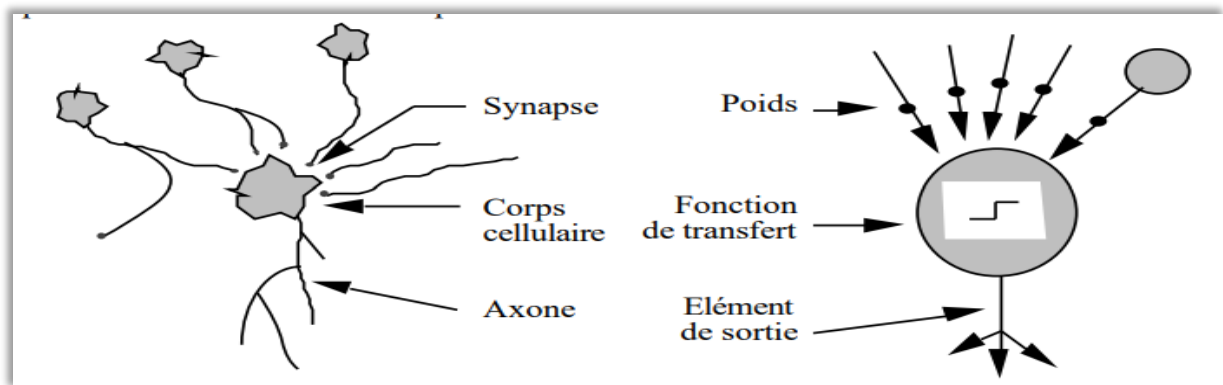


Figure 2.6. Mise en correspondance neurone biologique / neurone artificiel (Claude Touzet, 2016)

3.3.3. Fonctionnement d'un neurone artificiel

Un neurone seul est l'unité élémentaire de traitement d'un réseau de neurones.

Il est connecté à des sources d'information en entrée (d'autres neurones par exemple) et renvoie une information en sortie. Un neurone artificiel reçoit en entrée des entrées numériques x_i ($1 < i < N$ où N est le nombre de données qui va recevoir le neurone) valorisée chacune par un coefficient w_i . Le neurone artificiel (qui est une modélisation des neurones du cerveau) effectue alors une somme pondérée de ses entrées et lui ajoute un coefficient w_0 dit de biais supposé lié à une donnée $x_0 = -1$. $S = w_1 \rightarrow x_1 + \dots + w_N \rightarrow x_N + w_0 = \sum_{i=1}^N w_i \rightarrow x_i + w_0$

Cette donnée est passée à une fonction f dite d'activation qui représente un filtre permettant d'adapter la valeur de la somme précédente aux caractéristiques de la sortie désirée. C'est une fonction qui, généralement, doit renvoyer un réel proche de 1 quand les "bonnes" informations d'entrée sont données et un réel proche de 0 quand elles sont "mauvaises". La valeur de la fonction d'activation est la sortie y du neurone.

$$y = f(\sum_{i=1}^N w_i \rightarrow x_i + w_0)$$

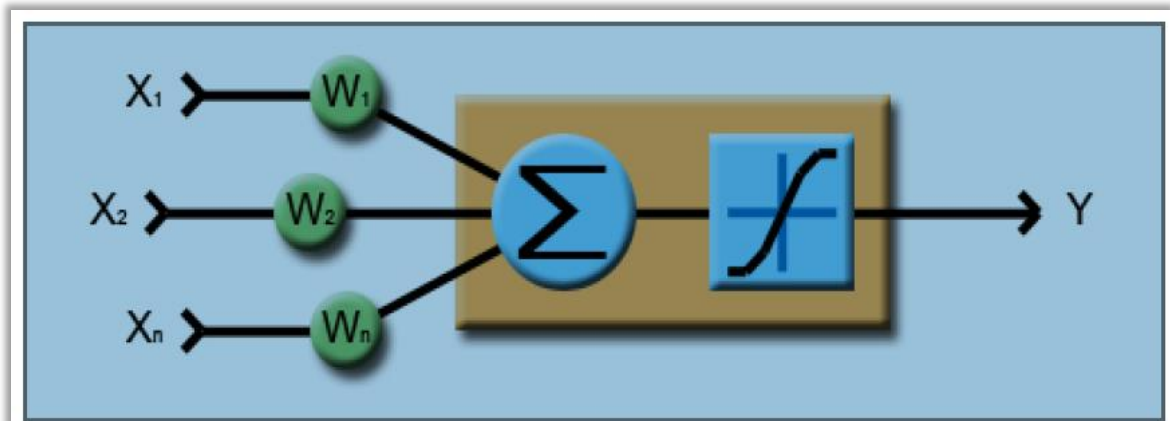


Figure 2.7. Fonctionnement d'un neurone artificiel (Claude Touzet, 2016)

3.3.4. Type de Réseau de neurone

L'architecture d'un réseau de neurones artificiel est définie par la structure de ses neurones et leur connectivité. Elle est spécifiée par le nombre d'entrées, de sorties, de nœuds et la façon selon laquelle sont interconnectés et organisés les nœuds.

3.3.4.1. Réseau de neurone monocouche (Perceptron)

Le réseau le plus simple est celui monocouche appelé le perceptron et caractérisé de la manière suivante :

- ✓ Il possède **N** informations en entrée
- ✓ Il est composé de **P** neurones, que l'on représente généralement alignés verticalement. Chacun peut en théorie avoir une fonction d'activation différente. En pratique, ce n'est généralement pas le cas
- ✓ Chacun des **p** neurones est connecté aux **N** informations d'entrée

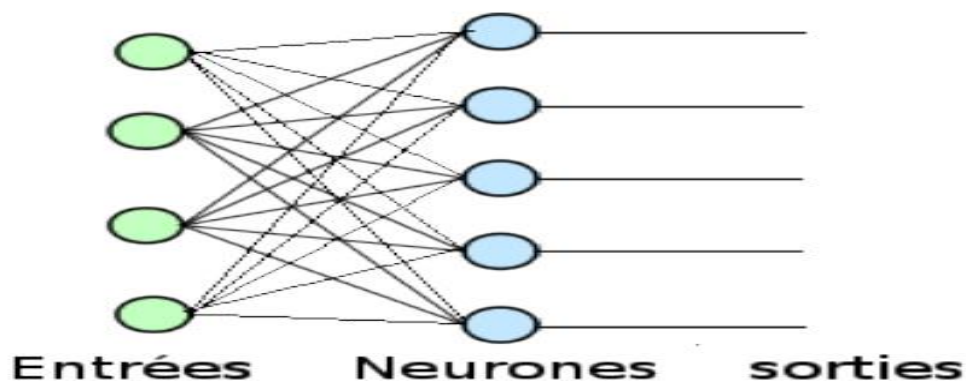


Figure 2.8. Réseau de neurone monocouche (Perceptron) (Claude Touzet, 2016)

3.3.4.2. Réseau de neurone multicouche

Dans ce réseau, les neurones de la première couche reçoivent toutes les informations entrées, ceux de la deuxième reçoivent toutes les sorties des neurones de la première couche, et ainsi de suite jusqu'au neurone de sortie qui reçoit celles de la dernière couche.

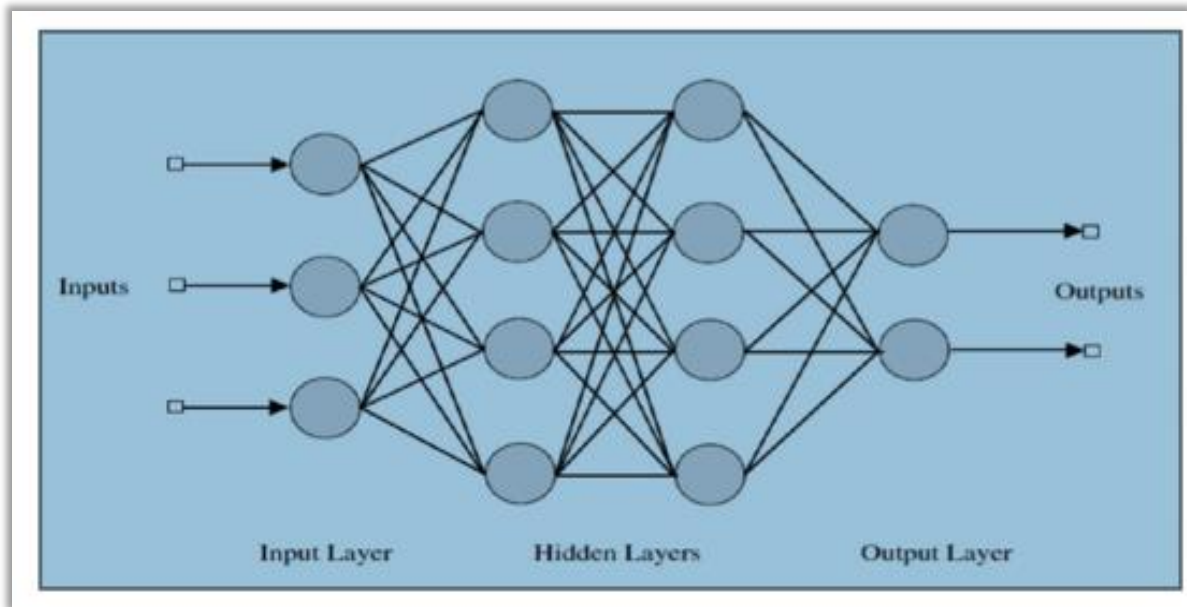


Figure 2.9. Structure d'un PMN composé de deux couches cachées (Claude Touzet, 2016)

3.3.5. Les réseaux de neurones convolutionnels / Convolution neural network « CNN »

Le réseau de neurone convolutionnel (CNN) est une forme spéciale du réseau neuronal artificiel. Il comporte plusieurs couches de convolution et est très bien adapté à l'apprentissage automatique et aux applications avec Intelligence artificielle (IA) dans le domaine de la reconnaissance d'images et de la parole, de la vente et du marketing ciblé et bien plus encore.

3.3.5.1. Définition d'un réseau de neurone convolutionnel

Un réseau de neurone convolutionnel (CNN) suppose une structure spatiale particulière dans son entrée. En particulier, il suppose que les entrées qui sont proches les unes des autres dans l'entrée originale sont sémantiquement liées (figure3-5). CNN est une séquence de couches, et chaque couche transforme un volume d'activations en un autre par une fonction différentiable. Les trois principaux types de couches pour construire ce type de réseau sont : couche convolutive, couche de pooling et couche entièrement connectée .

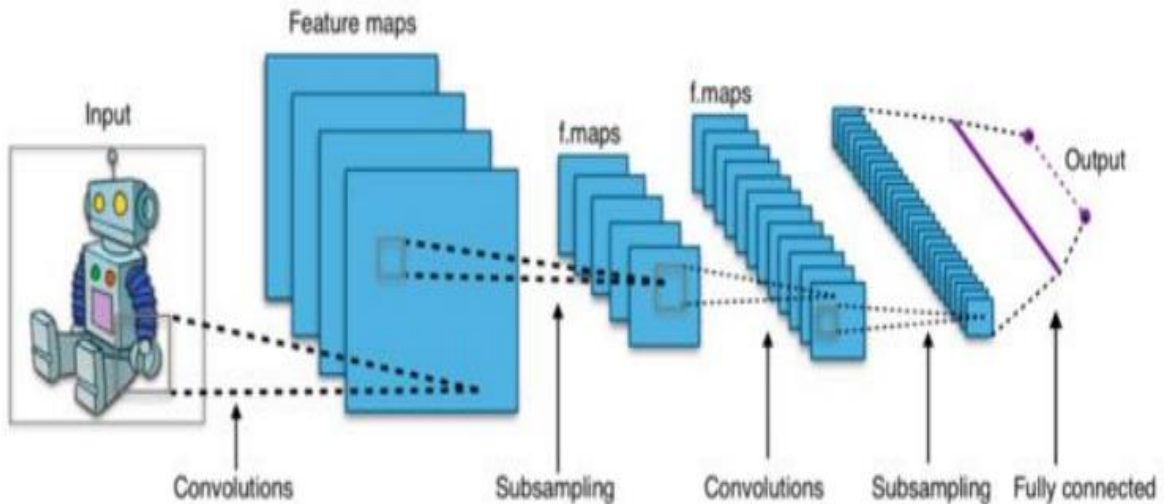


Figure 2.10. Architecture d'un réseau convolutionnel (ResearchGate,2023)

3.3.5.2. Les couches d'un réseau de neurone convolutionnel

➤ La couche convolutive

C'est la couche la plus importante et le cœur des éléments constitutifs du réseau convolutif, et c'est aussi elle qui effectue le plus de calculs lourds.

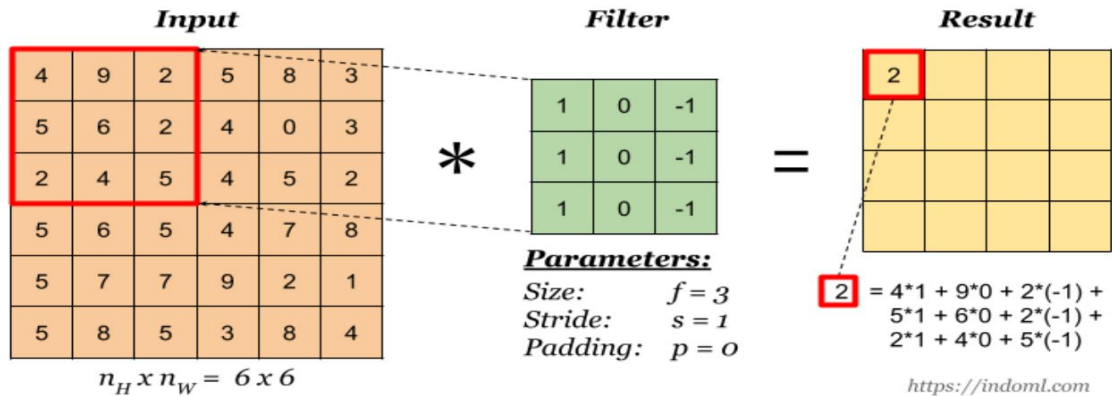


Figure 2.11. L'opération de convolution (ResearchGate,2023)

➤ La couche de pooling

Il est courant d'insérer périodiquement une couche Pooling dans ce type d'architecture. Sa fonction est de réduire progressivement la taille spatiale de la représentation pour réduire le nombre de paramètres et de calculs dans le réseau, et donc de contrôler également le surapprentissage.

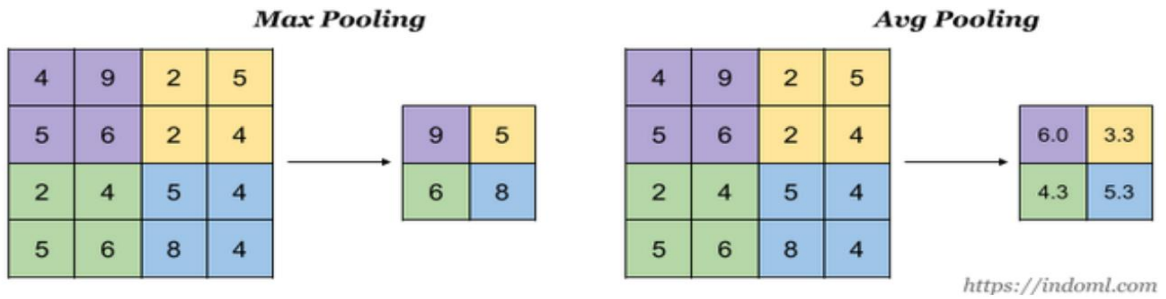


Figure 2.12. L'opération de pooling (ResearchGate,2023)

➤ **La couche entièrement connectée (Fully Connected Layer)**

Comme nous l'avons mentionné précédemment, les neurones d'une couche entièrement connectée ont des connexions complètes à toutes les activations de la couche précédente.

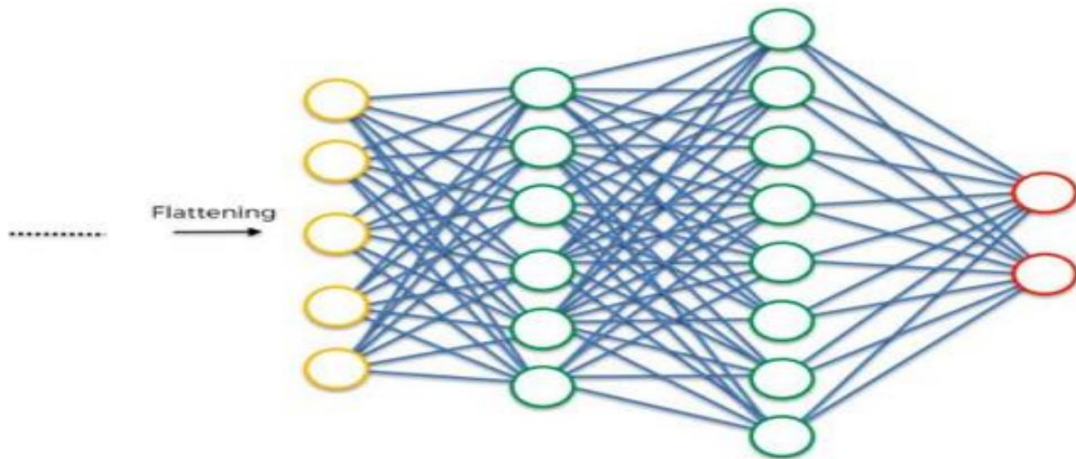


Figure 2.13. Fully Connected Layer (FC Layer) or classification (ResearchGate,2023)

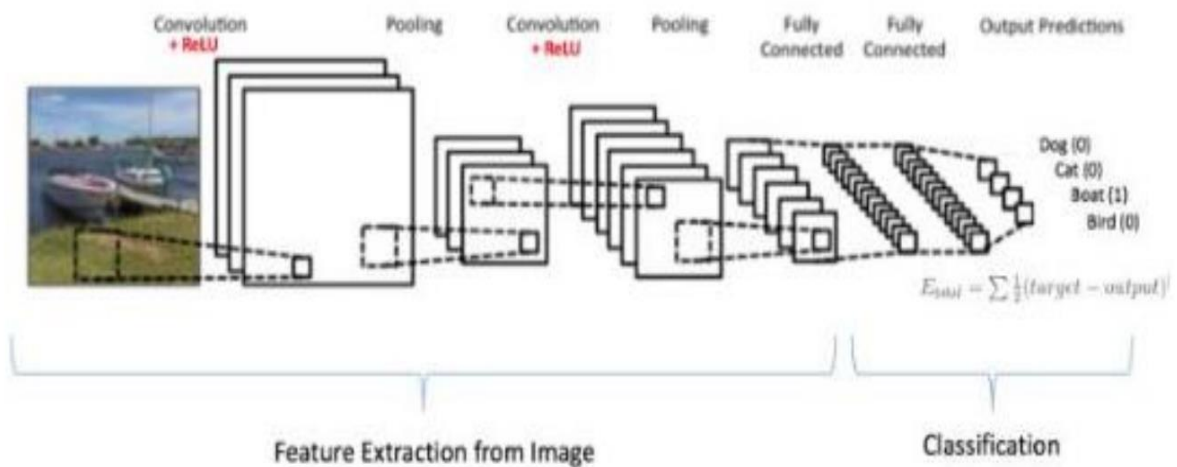


Figure 2.14. Processus d'apprentissage du CNN (ResearchGate,2023)

➤ Les fonctions d'activation

Les fonctions d'activations sont des fonctions mathématiques appliquées à la sortie d'un neurone dans un réseau neuronal. Elles sont utilisées pour introduire de la non-linéarité dans le modèle, ce qui permet au réseau de capturer des relations complexes et de modéliser des problèmes non linéaires. Voici quelques-unes des fonctions d'activation couramment utilisées:

- **Fonction d'activation linéaire**

Elle est définie par $f(x) = x$, où x est l'entrée du neurone. Cette fonction est souvent utilisée dans les couches de sortie lorsque la tâche consiste à prédire une valeur continue.

- **Fonction d'activation sigmoïde**

Elle est définie par $f(x) = 1 / (1 + \exp(-x))$. Cette fonction limite la sortie du neurone entre 0 et 1, ce qui est utile pour modéliser des probabilités ou des valeurs binaires.

- **Fonction d'activation tangente hyperbolique (tanh)**

Elle est définie par $f(x) = (\exp(x) - \exp(-x)) / (\exp(x) + \exp(-x))$. Cette fonction est similaire à la sigmoïde, mais elle permet des valeurs négatives en plus des valeurs positives.

- **Fonction d'activation rectifiée linéaire**

Elle est définie par $f(x) = \max(0, x)$. Cette fonction remplace les valeurs négatives par zéro, ce qui introduit de la non-linéarité tout en étant facile à calculer.

- **Fonction d'activation Leaky ReLU**

Elle est similaire à ReLU, mais elle permet une petite pente pour les valeurs négatives, ce qui peut aider à éviter le problème de "neurones morts" où un neurone a une sortie toujours nulle.

- **Fonction d'activation softmax**

Elle est utilisée dans les problèmes de classification multi classe. Elle normalise les sorties du neurone de manière qu'elles représentent des probabilités pour chaque classe.

Ces fonctions d'activation ne sont qu'une sélection parmi de nombreuses autres qui ont été proposées dans la littérature. Le choix de la fonction d'activation dépend du problème spécifique, de la nature des données et de l'architecture du réseau. Il est souvent recommandé d'expérimenter différentes fonctions d'activation pour trouver celle qui fonctionne le mieux pour une tâche donnée.

3.3.6. Les types de réseau de neurone Convolutionnel

Le tableau suivant représente les différents types d'architectures CNN :

Architecture	Année	Principales caractéristiques	Cas d'utilisation
LeNet	1998	Premières applications réussies des CNN, 5 couches (alternant entre convolutionnel et pooling), utilisé tanh/sigmoïde fonctions d'activation	Reconnaître les caractères manuscrits et imprimés à la machine
AlexNet	2012	Plus profond et plus large que LeNet, fonction d'activation ReLU utilisée, couches d'abandon implémentées, GPU utilisés pour la formation	Tâches de reconnaissance d'images à grande échelle
ZFNet	2013	Architecture similaire à AlexNet, mais avec des tailles et des nombres de filtres différents, Techniques de visualisation pour comprendre le réseau	Classement ImageNet
VGGNet	2014	Réseaux plus profonds avec des filtres plus petits (3×3), Toutes les couches convolutives ont la même profondeur, Configurations multiples (VGG16, VGG19)	Reconnaissance d'images à grande échelle
ResNet	2015	Introduction de « connexions ignorées » ou de « raccourcis » pour permettre la formation de réseaux plus profonds, configurations multiples (ResNet-50, ResNet-101, ResNet-152)	Reconnaissance d'image à grande échelle, a remporté la 1ère place à l'ILSVRC 2015
GoogleLeNet	2014	Introduction du module Inception, qui permet un calcul plus efficace et des réseaux plus profonds, plusieurs versions (Inception v1, v2, v3, v4)	Reconnaissance d'image à grande échelle, a remporté la 1ère place à l'ILSVRC 2014
MobileNets	2017	Conçu pour les applications de vision mobiles et embarquées, utilise des convolutions séparables en profondeur pour réduire la taille et la complexité du modèle	Applications de vision mobiles et embarquées, détection d'objets en temps réel

Tableau 01 : Les types de réseau de neurone convolutionnel CNN (Data Analytics, 2023)

3.4 Apprentissage par transfert

L'apprentissage par transfert, transfer learning en anglais, désigne l'ensemble des méthodes qui permettent de transférer les connaissances acquises à partir de la résolution de problèmes donnés pour traiter un autre problème. (DataScientest)

3.4.1 Utilité et fonctionnalités de l'apprentissage par Transfert

On distingue principalement deux catégories d'apprentissage par transfert : l'apprentissage par transfert inductif et l'apprentissage par transfert non supervisé. Avec l'apprentissage par transfert inductif, les problèmes à résoudre sont différents, mais les sources et cibles relèvent du même domaine. Pour l'apprentissage par transfert non supervisé, les tâches sont également différentes, mais les sources et cibles sont similaires.

3.4.2 VGGNet modèle de transfert learning

VGGNet est l'architecture CNN développée par Karen Simonyan, Andrew Zisserman et al. à l'Université d'Oxford. VGGNet est un CNN à 16 couches avec jusqu'à 95 millions de paramètres et formé sur plus d'un milliard d'images (1000 classes). Il peut prendre de grandes images d'entrée de 224×224 pixels pour lesquelles il dispose de 4096 caractéristiques convolutives. Les CNN avec des filtres aussi grands coûtent cher à former et nécessitent beaucoup de données, ce qui est la principale raison pour laquelle les architectures CNN aiment GoogLeNet.

Le modèle VGG CNN est efficace sur le plan informatique et sert de base solide pour de nombreuses applications de vision par ordinateur en raison de son applicabilité à de nombreuses tâches, y compris la détection d'objets. Le schéma ci-dessous représente le schéma d'architecture réseau standard VGG16 :

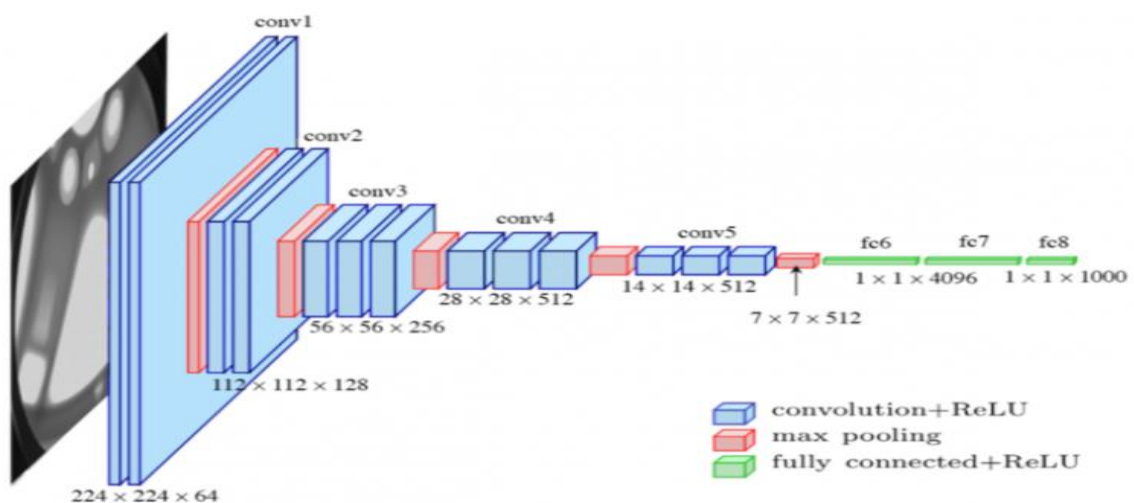


Figure 2.15. Architecture VGGNet (Data Analytics, 2023)

3.5 Augmentation de données

Un ensemble de données (data set) est une collection structurée de données utilisée pour l'apprentissage automatique, l'analyse statistique et d'autres tâches de traitement de données. Il est composé d'un ensemble d'exemples ou d'observations, où chaque exemple est constitué de plusieurs attributs ou caractéristiques.

Un ensemble de données peut prendre différentes formes selon le domaine d'application. Il peut s'agir d'un ensemble de textes, d'images, de séquences temporelles, de tableaux de valeurs numériques, etc. Les données peuvent provenir de diverses sources telles que des enquêtes, des expériences, des enregistrements de capteurs, des bases de données, des documents en ligne, etc.

L'ensemble de données est généralement divisé en deux parties principales : l'ensemble d'apprentissage (training set) et l'ensemble de test (test set). L'ensemble d'apprentissage est utilisé pour entraîner le modèle d'apprentissage automatique, tandis que l'ensemble de test est utilisé pour évaluer les performances du modèle sur des données non vues auparavant. Cette séparation permet de mesurer l'aptitude du modèle à généraliser à de nouvelles données.

Il est important de disposer d'un ensemble de données de qualité pour obtenir des résultats fiables et précis. Cela comprend des données représentatives, sans biais, complètes et étiquetées correctement (si l'apprentissage supervisé est utilisé). Dans certains cas, il peut être nécessaire de prétraiter les données en les nettoyant, en les normalisant ou en les transformant afin de les rendre appropriées pour l'apprentissage automatique.

3.5.1 Définition de l'augmentation de données

Les algorithmes, même les plus complexes et puissants, ont besoin de données pour entraîner le modèle. L'augmentation des données, data augmentation en anglais, est l'une des solutions les plus adaptées pour ce genre de problématique. Au lieu d'essayer de trouver et d'étiqueter plus de données, nous en construisons de nouveaux en fonction de ce que nous avons.

L'augmentation des données d'image est peut-être la technique la plus connue d'augmentation des données. Elle regroupe les techniques utilisées pour augmenter artificiellement la taille d'un groupe de données d'apprentissage en créant des versions modifiées d'images à partir des images d'apprentissage disponibles. Nous pouvons alors améliorer efficacement le processus d'apprentissage puisqu'il en résulte en plus d'échantillons d'entraînement pour le modèle de réseau de neurones. Les techniques d'augmentation peuvent créer des variations d'images qui peuvent

améliorer la capacité des modèles d'entraînement pour généraliser ce qu'ils ont appris à de nouvelles images, ce qui améliore fortement la performance du modèle. (Beyond Data, 2021)

3.5.2 Types augmentation de données

- **L'augmentation des données en ligne**

L'augmentation de données par défaut utilisée dans les modèles d'apprentissage automatique est l'augmentation de données en ligne, où les images des données de formation sont prises au hasard et les techniques d'augmentation de données sont appliquées. Le modèle est ensuite formé sur la base des données d'origine où les images sont augmentées de manière aléatoire. Dans ce cas, les images augmentées ne sont jamais enregistrées nulle part et il est impossible de savoir quelle image est augmentée.

Les modèles d'apprentissage en profondeur utilisent déjà l'augmentation de données en ligne comme la rotation, le retournement, le cisaillement, la plage de luminosité, etc., ceux-ci sont appliqués de manière aléatoire aux données d'apprentissage avant d'ajuster le modèle. Cette technique ne sauvegarde aucune donnée sur le disque et prétraite les images en temps réel. (Analytics Vidhya, 2021)

- **L'augmentation des données hors ligne**

La génération d'ensembles de données ou l'augmentation de données hors ligne permet au programmeur d'enregistrer les images augmentées sur le disque. Les techniques qui peuvent être utilisées pour effectuer une augmentation de données en ligne peuvent être utilisées pour effectuer une augmentation de données hors ligne, où les images sont stockées sur le disque, c'est ce qu'on appelle une augmentation de données hors ligne. Les images augmentées sont obtenues après l'application des techniques d'augmentation de données sur chaque image d'entraînement. Cela rend l'ensemble de données plus diversifié et le modèle plus robuste. Ce processus peut être utilisé pour améliorer le nombre d'images dans l'ensemble de données. (Analytics Vidhya, 2021)

- **Combinaison d'augmentation de données en ligne et hors ligne (Hybride)**

Comme son nom l'indique, dans ce cas, les deux méthodes sont utilisées pour analyser les données. Les images augmentées obtenues sont mélangées avec l'ensemble de données d'origine, puis l'augmentation des données en temps réel est effectuée juste avant la formation du modèle d'apprentissage automatique. Cette méthode est généralement utilisée à des fins de jeu ou lorsque les données disponibles ne sont pas claires. Dans le cas du troisième type d'augmentation de données, les techniques sont appliquées en modes hors ligne et en ligne. Les images sont stockées sur le disque pendant l'augmentation des données hors ligne, puis elles sont mélangées avec l'ensemble de données d'origine et l'augmentation des données en ligne est appliquée. (Analytics Vidhya, 2021)

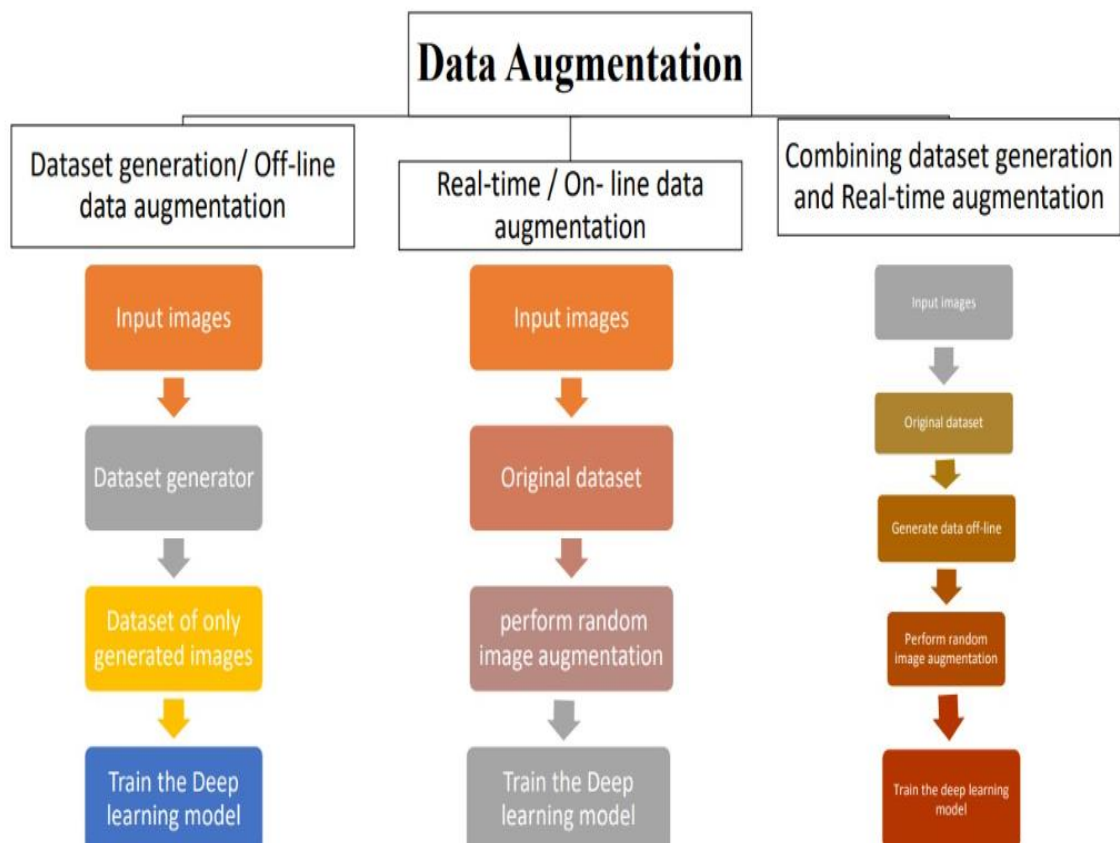


Figure 2.16. Types d'augmentation de données (Analytics Vidhya, 2021)

4. Conclusion

La technologie de l'intelligence artificielle ainsi que ses différentes branches à savoir l'apprentissage automatique et l'apprentissage profond a permis l'ouverture de nouveaux horizons de recherche et des découvertes importantes dans divers domaines tels que la santé et la médecine. Ces découvertes ont été un facteur important pour trouver des solutions à des problèmes qui semblaient difficiles à résoudre.

Chapitre 3

Conception et

Implémentation de

L' Application

SyndAP

1. Introduction

Dans ce chapitre, nous présentons la conception et l'implémentation de l'application bureau intelligente que nous avons développée, et qui sert à la prédiction du syndrome main-pied. Ce syndrome est reconnu comme un effet secondaire qui affecte les personnes atteintes d'un cancer et sous chimiothérapie. Pour commencer, nous expliquons deux techniques clés adoptées dans notre travail

➤ **Augmentation des données**

Nous utilisons la technique d'augmentation des données pour augmenter la taille de notre base d'images. Cela nous permet de générer de nouvelles images en appliquant des transformations géométriques telles que des rotations, des translations, des zooms, des changements d'éclairage, etc. Cela aide à améliorer la robustesse du modèle de reconnaissance développé en exposant davantage de variations possibles dans les images.

➤ **Apprentissage automatique supervisé**

Nous avons mis en œuvre un modèle d'apprentissage supervisé qui réalise la prédiction du syndrome main-pied. Nous définissons l'architecture du modèle, qui comprend des couches de convolution, des couches de pooling et des couches entièrement connectées. Le modèle apprend à extraire les caractéristiques pertinentes des images et à les utiliser pour prédire la présence ou l'absence du syndrome main-pied. Enfin, nous procédons à une évaluation des performances de notre modèle en utilisant des métriques telles que les courbes d'apprentissage et la matrice de confusion.

Ce chapitre sert de base pour comprendre les détails techniques de l'application développée, les outils utilisés et les résultats obtenus, tout en tirant des conclusions et discutant des implications de notre travail.

2. Les outils d'implémentation

2.1. Environnement de développement

Pour implémenter notre application, nous avons choisi *Spyder* comme un environnement de développement, avec l'utilisation des bibliothèques **PySimpleGUI**.

2.1.1. Spyder

Spyder est un environnement scientifique puissant écrit en Python, créé et développé par Pierre Raybaut en 2008. Il est conçu spécifiquement pour les scientifiques, les ingénieurs et les analystes de données.

L'objectif de Spyder est de fournir aux utilisateurs une combinaison unique de fonctionnalités avancées d'édition, d'analyse, de débogage et de profilage, similaires à celles d'un environnement de développement complet. En même temps, il offre des fonctionnalités de visualisation, d'exécution interactive, d'exploration de données et d'inspection approfondie, typiques d'un package scientifique.

Avec Spyder, les utilisateurs peuvent écrire, exécuter et déboguer leur code Python de manière interactive. L'environnement offre des fonctionnalités telles que la coloration syntaxique avancée, l'auto-complétion, l'intégration de la documentation, l'exploration des variables, le profiling de code, l'intégration de consoles IPython et bien plus encore. Cela permet aux scientifiques et aux analystes de travailler de manière efficace et productive dans leur flux de travail d'analyse et de développement.

En outre, Spyder propose des capacités de visualisation avancées, permettant aux utilisateurs de créer des graphiques et des représentations visuelles de leurs données. Cela aide à analyser et à comprendre les résultats plus facilement.

En résumé, Spyder est un environnement scientifique complet et polyvalent qui combine des fonctionnalités de développement avancées avec des capacités de visualisation et d'exploration des données. Il est spécialement conçu pour répondre aux besoins des scientifiques, des ingénieurs et des analystes de données travaillant avec Python. (Spyder Website Contributors, 2018).

2.1.2. PySimpleGUI

PySimpleGUI est une bibliothèque Python open source qui permet de créer rapidement et facilement des interfaces graphiques (GUI - Graphical User Interface). Elle vise à simplifier la création d'interfaces utilisateur en fournissant une syntaxe simple et intuitive.

PySimpleGUI offre une grande flexibilité et prend en charge plusieurs backends graphiques, notamment tkinter, Qt, wxPython, et Remi pour les interfaces web. Il fournit également une large

gamme d'éléments d'interface prêts à l'emploi tels que des boutons, des champs de saisie, des cases à cocher, des listes déroulantes, etc., ainsi que des mises en page flexibles pour organiser les éléments de l'interface.

L'une des caractéristiques les plus intéressantes de PySimpleGUI est sa syntaxe déclarative, ce qui signifie que vous pouvez décrire l'apparence et le comportement de votre interface en utilisant simplement une liste ou un tableau Python. Cela rend le processus de création d'interfaces graphiques plus rapide et plus intuitif, surtout pour les débutants.

(<https://pypi.org/project/PySimpleGUI>, 2023)

2.2. Langage de programmation

De nos jours, il existe plusieurs langages de programmation et chaque langage possède ses propres caractéristiques. Parmi ses langages, notre choix s'est focalisé sur Python.

2.2.1. Python

Il est l'un des langages de programmation les plus intéressants du moment, il est inventé par « Guido van Rossum », la première version de python est sortie en 1991. C'est un langage de programmation interprété, multiparadigme et multiplateformes. Il favorise la programmation impérative structurée, fonctionnelle et orientée objet

Python est un langage de programmation interprété, ce qui signifie que le code source est exécuté directement sans nécessiter de compilation préalable. Cela permet une approche de développement rapide et facilite la création de prototypes.

Python est un langage multi paradigme, ce qui signifie qu'il prend en charge plusieurs styles de programmation. Il favorise particulièrement la programmation impérative structurée, la programmation fonctionnelle et la programmation orientée objet.

Python est à la fois simple et puissant, il permet d'écrire des scripts très simples, et grâce à ses nombreuses bibliothèques, nous pouvons travailler sur des projets plus ambitieux (python.doctor, 2019).

2.2.2. Pourquoi Python ?

Python est un langage stable, flexible et il fournit divers outils pour les développeurs. Ce qui permet de le classer en premier choix pour l'apprentissage automatique, à partir du développement, la mise en oeuvre, et la maintenance. Python aide les développeurs de développer des produits confiants (Gupta, 2019).

Selon (Gupta, 2019), Python possède de nombreux avantages :

- **Simple et cohérent** : la simplicité de Python aide les développeurs à gérer l'algorithme complexe de l'apprentissage automatique.
- **La flexibilité** : Le facteur de flexibilité réduit la possibilité d'erreurs, il a laissé les programmeurs prendre le contrôle complètement, et de travailler sur elle confortablement.
- **Bibliothèques et Framework** : Les algorithmes de l'apprentissage automatique sont très complexes, mais Python est le secours avec une large gamme de bibliothèques comme *Scikit-learn*, *Keras* et de Framework, à titre d'exemple *TensorFlow*.
- **Lisibilité** : Python est facile à lire, de sorte que les développeurs peuvent facilement comprendre le code.
- **Indépendance de la plateforme** : Python est un langage indépendant de la plate-forme. Il est supporté par de nombreuses plateformes, y compris Windows, LINUX et macOS.

3. Méthodologie de travail

Dans notre travail, nous avons suivi les étapes suivantes :

3.1. Collection des images

Dans notre travail, nous avons utilisé un sous-ensemble de la base d'images « 11k Hands ». Cette base d'images est une collection de 11 076 images de mains (1600 x 1200 pixels). Chaque main a été photographiée de côté palmaire avec un fond blanc uniforme et placée approximativement à la même distance de l'appareil photo. L'ensemble de données proposé contient un grand nombre d'images de mains avec des métadonnées plus détaillées. L'ensemble de données est gratuit pour un usage académique raisonnable.

D'un autre côté, nous avons procédé à la collecte de **90** images depuis Internet qui représentent le syndrome main-pied. Nous nous sommes focalisés seulement sur des images de mains et non de pieds, dans le but de créer notre propre base d'images représentant le syndrome mains-pieds

Cette base d'images a été augmentée à l'aide de la technologie d'augmentation des données

3.2 Augmentation hors ligne des images collectées et création de la base d'image.

Après avoir procédé à la collection des images comme expliqué dans la figure ci-dessus, nous avons procédé à l'augmentation des images collectées comme le montre la figure ci-dessous. Nous avons procédé en premier lieu à une étape de pré-traitement qui consiste à un re-dimensionnement des images vue qu'elles ne sont pas de même dimension. Toutes les images ont été redimensionnées à la dimension **150 x 150** et convertit au format jpeg.

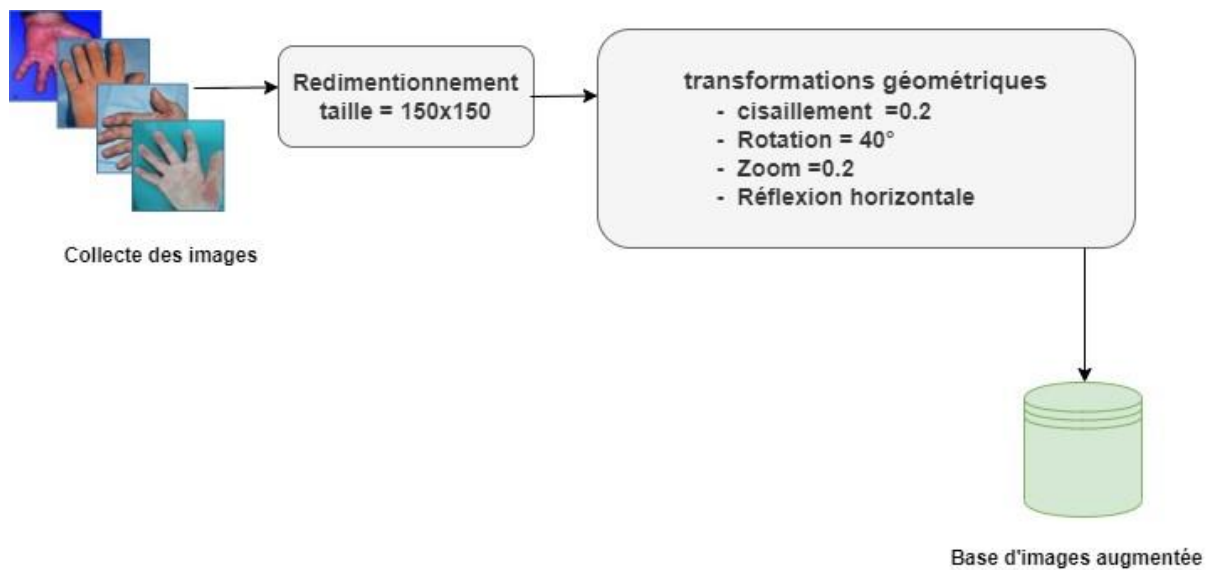


Figure 3.1. Processus d'augmentation hors ligne des images collectées.

➤ Technique de rotation

La technique d'augmentation « Rotation » comme son nom l'indique, consiste à modifier l'angle de l'image, à la faire tourner sur elle-même. Le modèle sera plus robuste aux changements d'angles et prises de vue. Nous avons utilisé un angle de rotation de 40°.

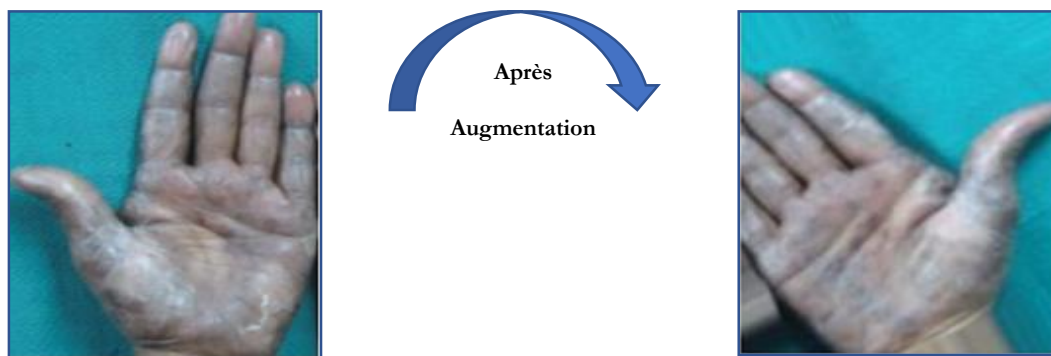


Figure 3.2. Effet de rotation sur une image présentant le syndrome main pied.

Technique de Cisaillement

L'inclinaison et la perspective fournissent une distorsion le long de l'axe des x et de l'axe des y. Il modifie une image de telle sorte qu'elle donne l'impression d'un angle de perception différent.

L'inclinaison et la perception sont utiles car le modèle en production recevra des images d'entrée sous toutes sortes d'angles. Cette augmentation rend un modèle robuste contre différents angles de perception visuelle.

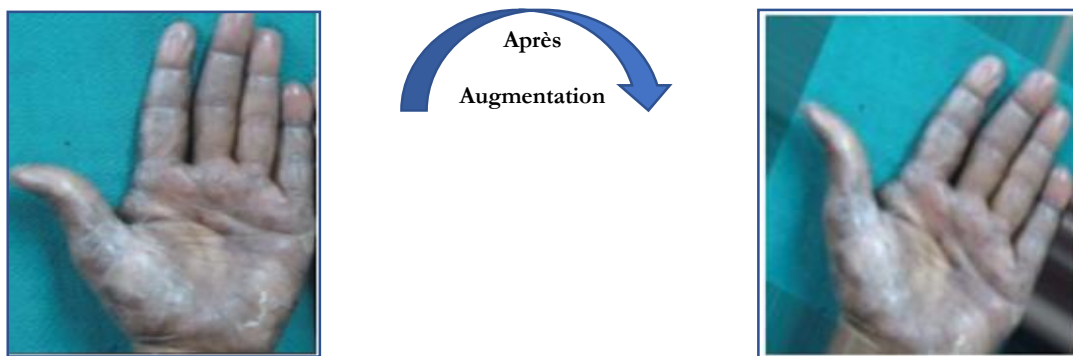


Figure 3.3. Effet du cisaillement sur une image présentant le syndrome main pied.

➤ Technique de zoom

C'est le changement d'échelle en x et/ou y. On trouve la réduction et l'augmentation de la taille de l'image. L'agrandissement est caractérisé par les paramètres S_x et S_y supérieur à 1 par contre la réduction est caractérisée par les paramètres S_x et S_y inférieur à 1.

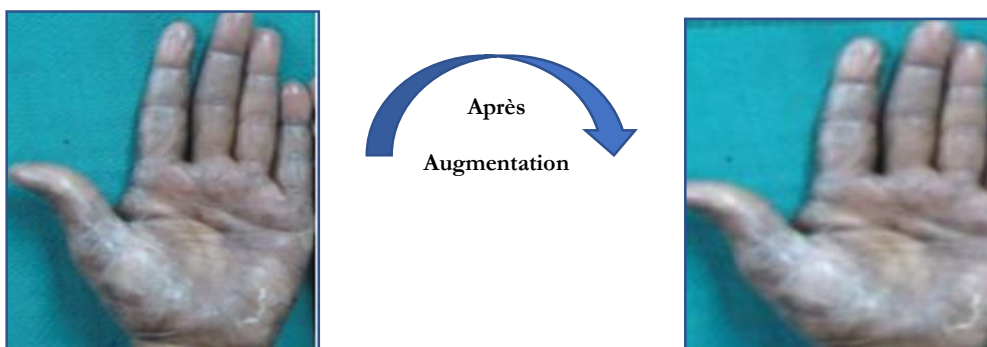


Figure 3.4. Effet du zoom sur une image présentant le syndrome main pied

➤ Technique de réflexion horizontale

Un retournement (Flip) d'image signifie l'inversion des lignes ou des colonnes de pixels dans le cas d'un flip vertical ou horizontal respectivement.

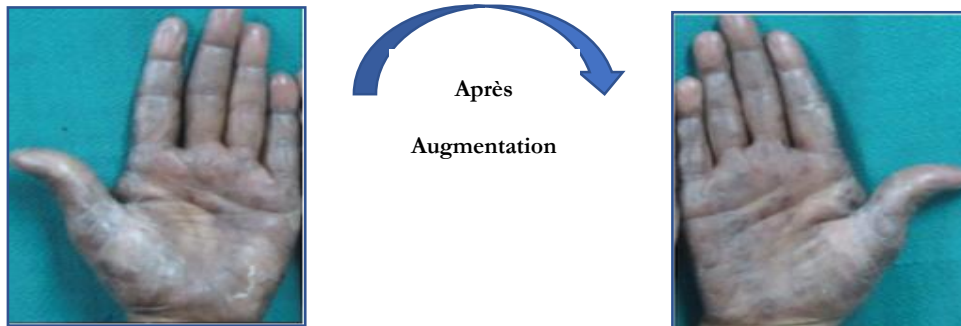


Figure 3.5. Effet réflexion horizontale sur une image présentant le syndrome main pied.

Après avoir procédé à l'augmentation des images, nous avons obtenu une base d'images composée de **21194** images représentant des images de mains normales et **21194** images représentant des images représentant le syndrome main pied. Nous avons sélectionné aléatoirement **2000** images réparti comme suit :

- **1600** images constituant l'ensemble d'apprentissage,
- **200** images constituant l'ensemble de validation,
- **200** images constituant l'ensemble de test.

3.3 Architecture du système de reconnaissance du syndrome HFS proposé

La figure ci-dessous illustre l'architecture détaillée du système de reconnaissance du syndrome HFS développé.

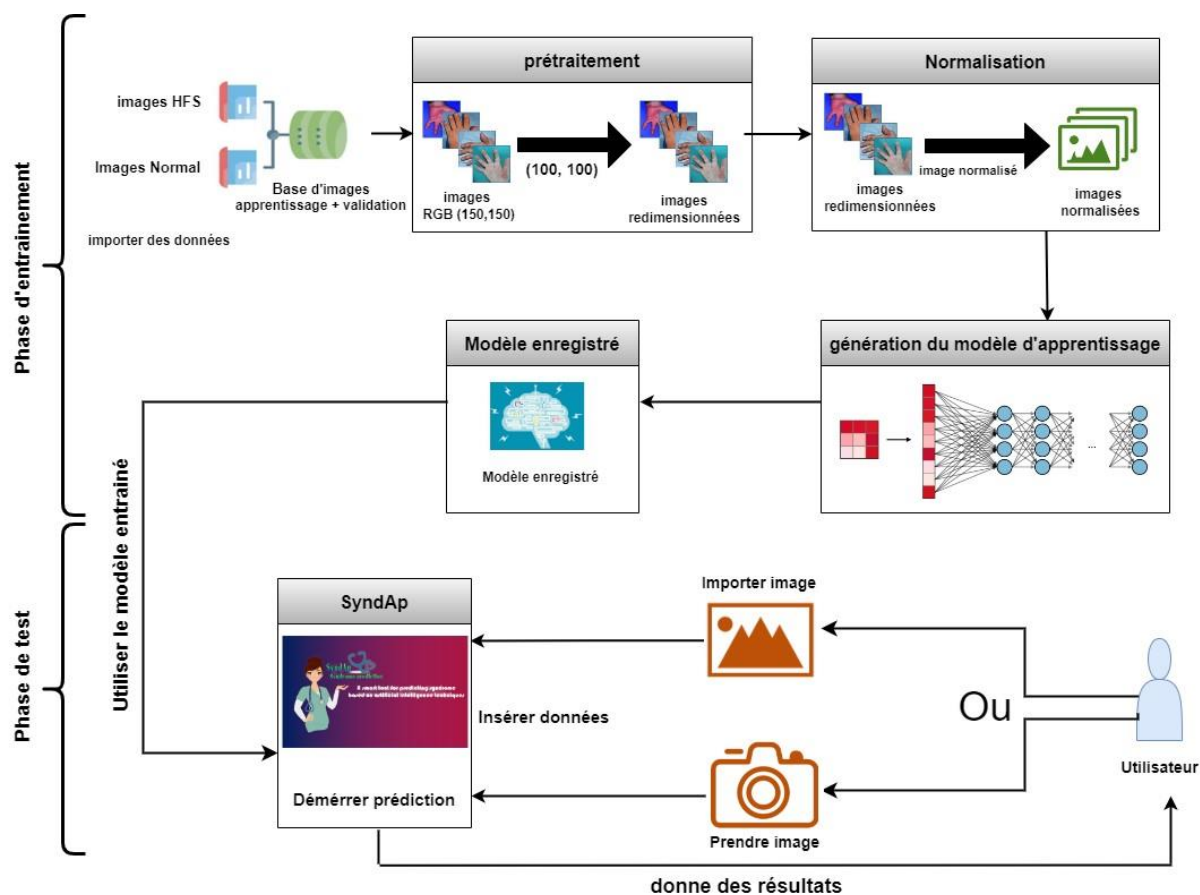


Figure 3.6. Architecture du système de reconnaissance du syndrome HFS proposé

Pré-traitements et normalisation des images

Nous procédons au redimensionnement des images augmentées de dimension 150 x150 à la dimension 100 x 100, suivi d'une opération de normalisation dans l'intervalle [0-1].

Génération du modèle d'apprentissage

Architecture du réseau CNN

Dans notre travail, nous avons implémenté un réseau de neurone convolutif composé de 4 couches de convolution, 4 couches de maxpooling et 2 couches entièrement connectées. L'image d'entrée a une taille de 100 x 100 pixels. Chaque couche de convolution utilise plusieurs filtres (32 pour les deux premières couches) de taille 2 x 2. La fonction d'activation ReLU est utilisée après chaque passage à travers une couche de convolution. La fonction ReLU force les neurones à renvoyer des valeurs positives, ce qui améliore l'apprentissage du modèle. Après chaque couche de convolution, une opération de maxpooling est appliquée pour réduire la taille de l'image et extraire les caractéristiques les plus importantes. Cela contribue à la réduction de la dimensionnalité et à l'augmentation de la vitesse de calcul.

Pour éviter le sur-apprentissage (overfitting en anglais), nous avons utilisé la technique de dropout. Le dropout désactive aléatoirement un certain pourcentage de neurones pendant l'entraînement, ce qui oblige le modèle à apprendre des caractéristiques plus robustes et généralisables.

Dans la deuxième couche de convolution, le nombre de filtres est augmenté à 64 au lieu de 32, tout en conservant la fonction d'activation ReLU et la même opération de maxpooling.

Dans la troisième couche de convolution, le nombre de filtres est à nouveau augmenté à 128, avec la même fonction d'activation ReLU et maxpooling.

Dans la quatrième couche de convolution, le nombre de filtres est encore augmenté à 256, avec la même fonction d'activation ReLU et maxpooling.

Après les couches de convolution et de maxpooling, nous utilisons la couche "flatten" pour convertir les données 3D en un vecteur 1D. Cela permet de passer des volumes 3D à une représentation linéaire afin d'être utilisée par les couches entièrement connectées.

Les dernières couches du modèle sont des couches entièrement connectées, qui ont des connexions complètes avec toutes les opérations de classification précédentes. Dans notre cas, nous utilisons la fonction d'activation sigmoïde, qui est couramment utilisée pour les tâches de classification binaire.

Ce modèle avec ses différentes couches de convolution et ses couches entièrement connectées permet de capturer et d'apprendre les caractéristiques pertinentes des images d'entrée, puis de les utiliser pour effectuer une classification précise. La figure suivante montre tout cela :

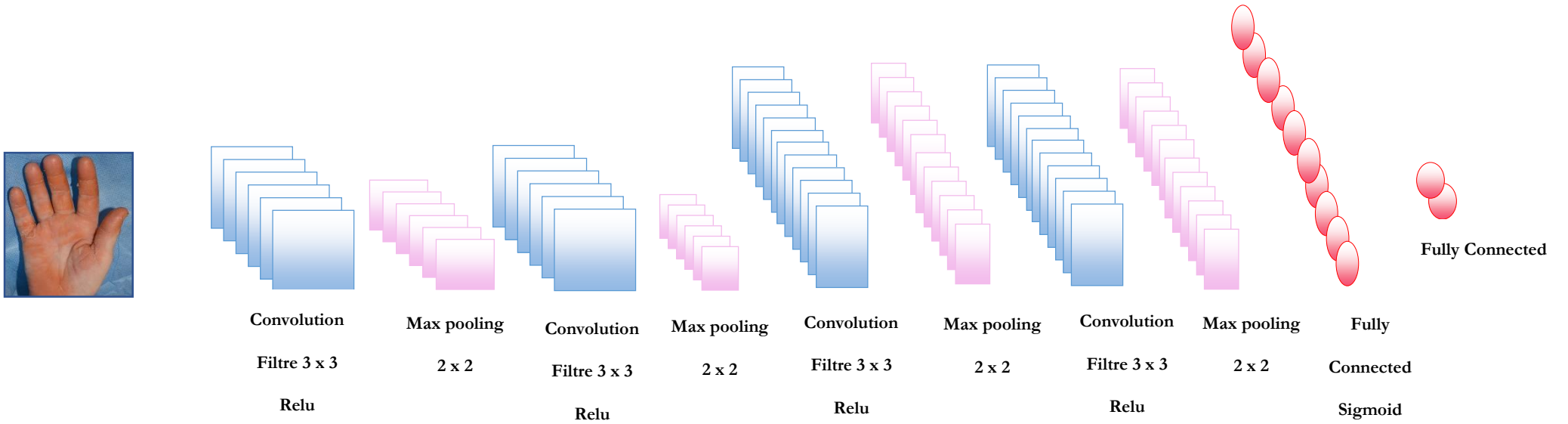


Figure 3.7. Architecture de modèle de réseaux neuronal convolutif

Courbes d'apprentissages

Afin de valider le modèle d'apprentissage, nous avons généré les courbes d'apprentissage comme le montre les figures suivantes. D'après la figure 28, le taux de reconnaissance (accuracy) de l'apprentissage et le taux de reconnaissance (accuracy) de la validation augmente à mesure que le nombre d'époques augmente jusqu'à ce que la validation commence à diminuer, de sorte à provoquer sur-apprentissage. Afin d'éviter que le modèle provoque un sur-apprentissage, nous avons utilisé un early stopping empêchant le processus d'apprentissage de continuer.

Nombre des époques = 20 batch_size = 32

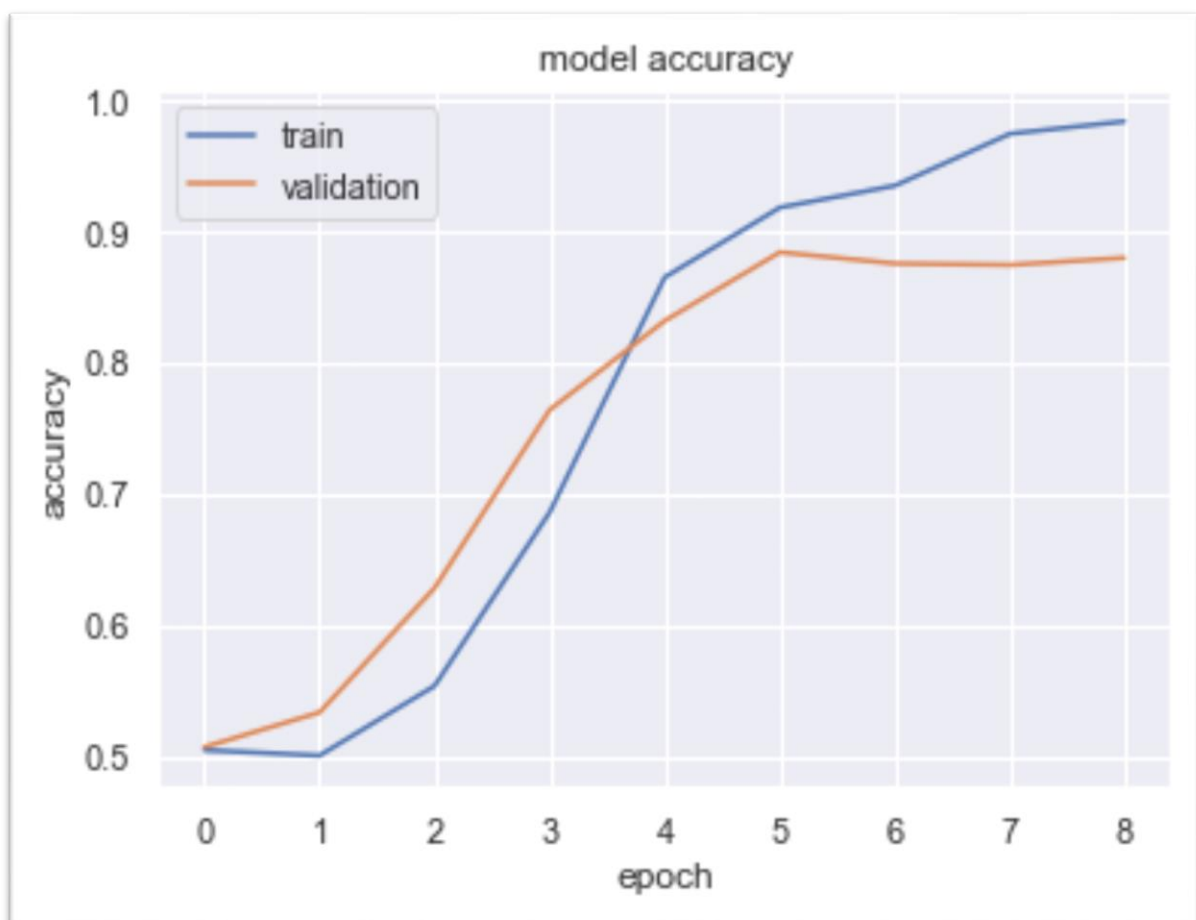


Figure 3.8. Courbe d'apprentissage (Accuracy/ Epoch)

D'après la figure 3.9, nous constatons que les erreurs d'apprentissage et de validation diminuent avec le nombre d'époques.

Nombre des époques = 20

batch_size=32

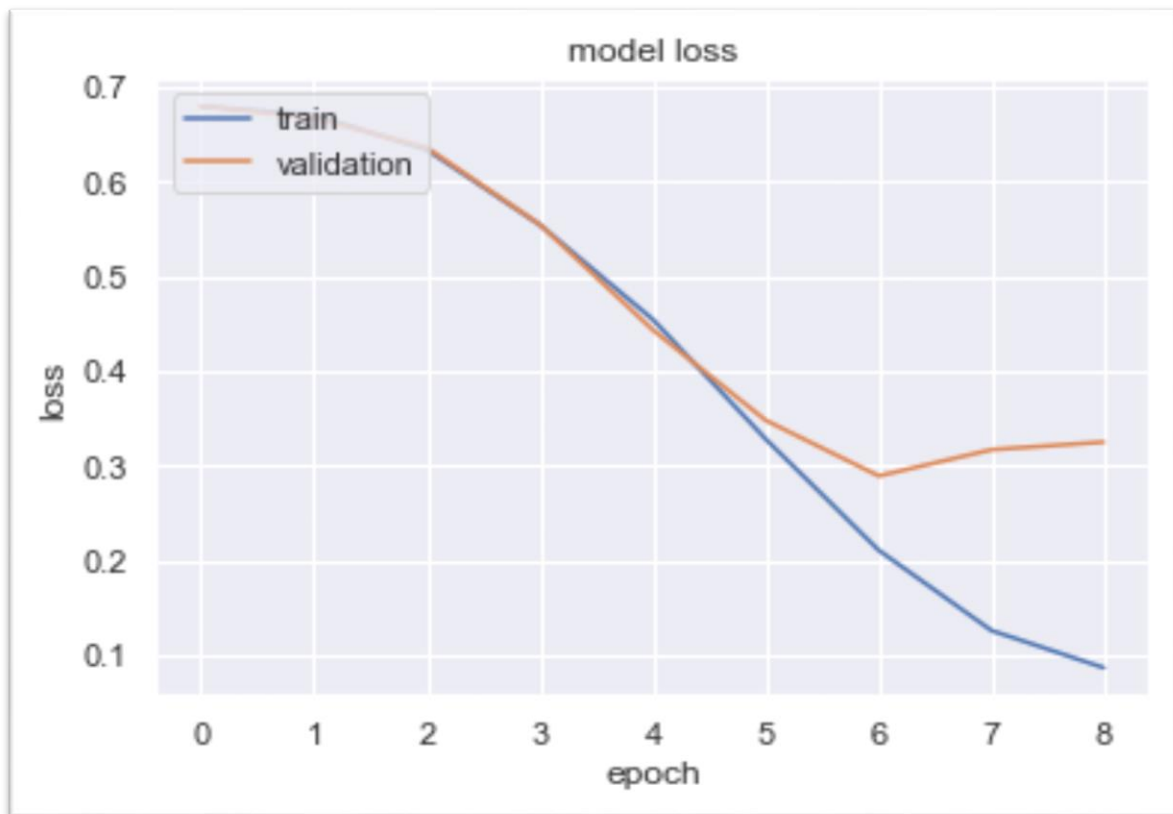


Figure 3.9. Courbe d'apprentissage (Loss/ Epoch)

Evaluation des performances

Afin d'évaluer les performances du système proposé sur la base de test, nous avons calculé les métriques évoquées ci-dessous. Ces indicateurs fournissent une évaluation globale des performances des modèles. Le taux de reconnaissance mesure la proportion d'instances correctement prédites, tandis que l'erreur représente le taux d'instances mal classées. La matrice de confusion fournit des informations détaillées sur les performances de classification en montrant le nombre d'instances prédites pour chaque classe réelle. Ces résultats permettent d'évaluer l'efficacité des modèles dans la tâche de prédiction.

Matrice de confusion

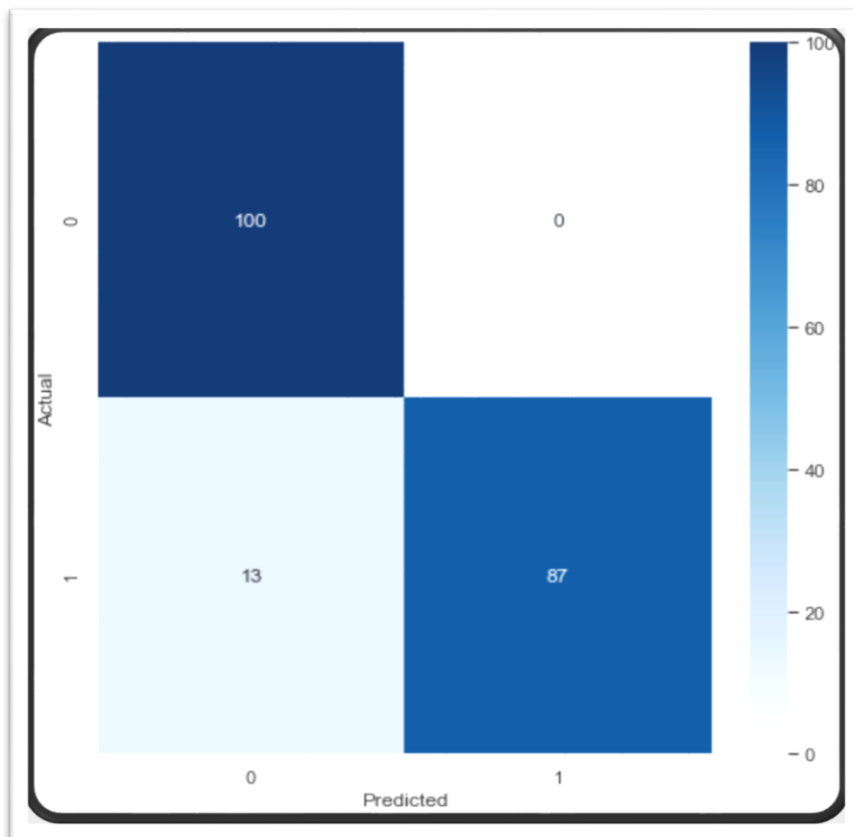


Figure 3.10. Matrice de confusion

- ✓ Pour mesurer les performances de ce classifieur, il est d'usage de distinguer 4 types d'éléments classés pour la classe voulue :
- **Positive (P)**: Observation est positive.
 - **Negative (N)**: Observation n'est pas positive.
 - ↪ **True Positive (TP)**: Observation est positive, et est prédite positive.
 - ↪ **True Negative (TN)**: Observation est négative, et est prédite négative.
 - ↪ **False Positive (FP)**: Observation est négative, mais est prédite positive.
 - ↪ **False Negative (FN)**: Observation est positive, mais est prédite négative.

Taux de classification /Accuracy

$$\text{Accuracy} = \frac{\text{True_Positive} + \text{True_Négative}}{\text{Nombre total d'exemples}}$$

Dans notre cas, nous avons obtenu un taux de classification estimé à **0.93%**

Taux d'erreur

$$\text{Erreur} = \frac{\text{False_Positive} + \text{False_Négative}}{\text{Nombre total d'exemples}}$$

Dans notre cas, nous avons obtenu un taux d'erreur estimé à **0.07%**

3.4 L'application SyndAP

Dans ce travail, nous avons développé est une application intelligente, SyndAP , qui permet de prédire si un patient cancéreux est atteint ou pas du syndrome main-pied. La figure ci-dessous illustre la fenêtre d'accueil de l'application SyndAP.

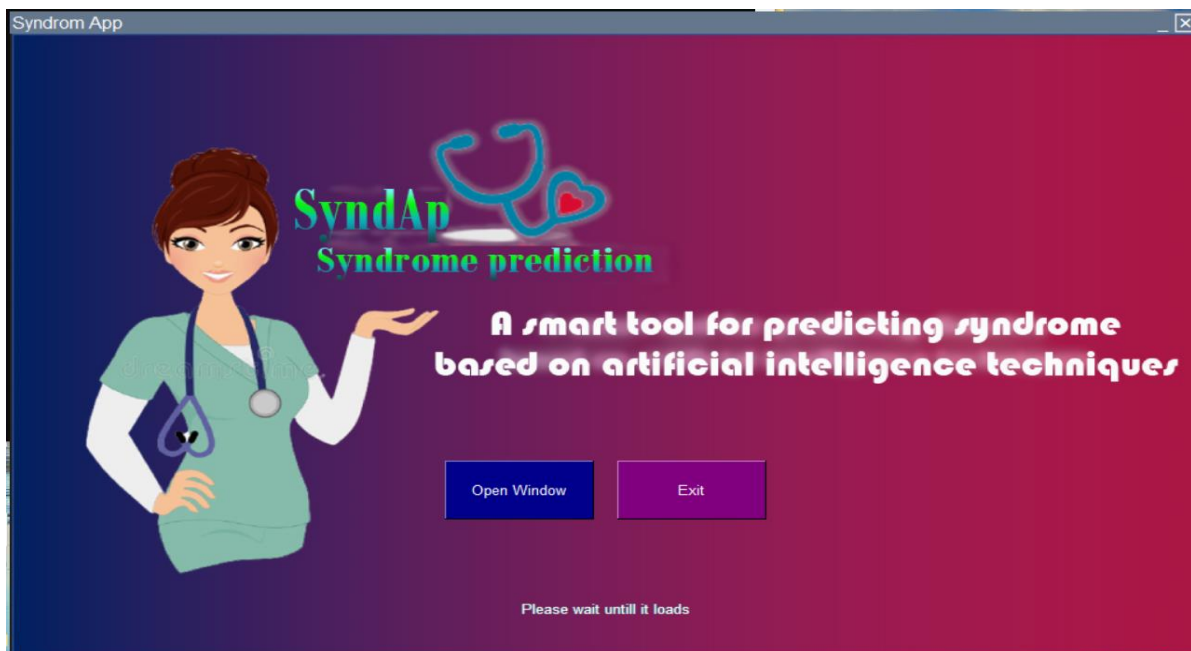


Figure 3.11. Fenêtre d'accueil.

La fenêtre principale de notre application contient deux boutons :

Exit : ce bouton est conçu pour quitter l'application.

Open Windows : ce bouton mène un utilisateur à la fenêtre principale de l'application comme le montre la figure ci-dessous.

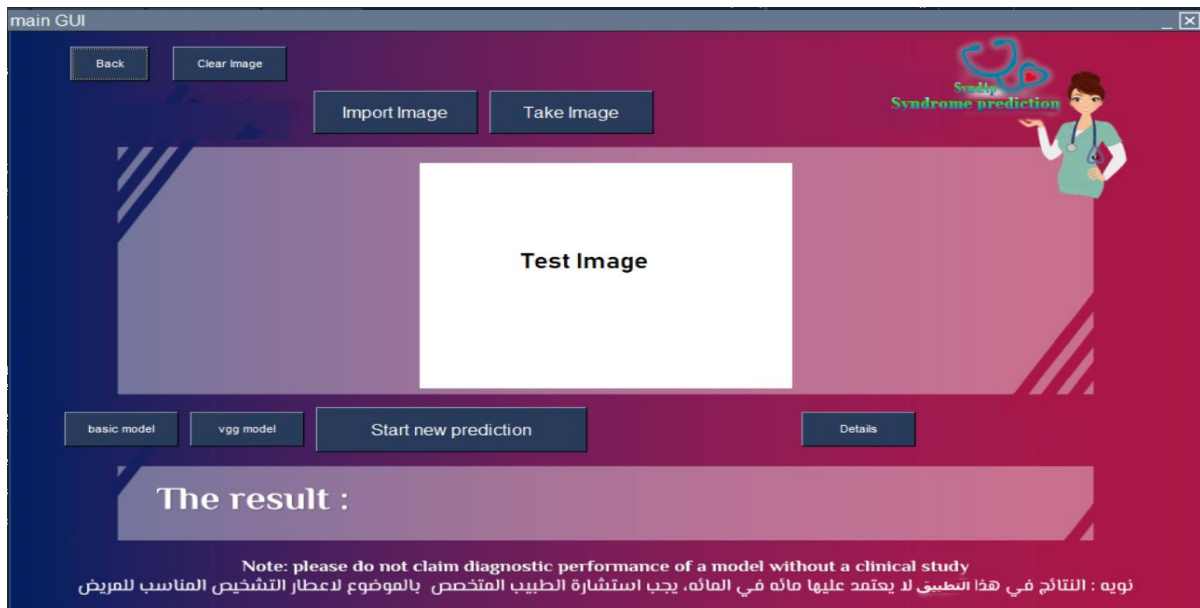


Figure 3.12. Nouvelles prédictions

Au niveau de cette fenêtre, l'utilisateur est invité à prendre ou importer une image photographiée de sa main afin de subir un diagnostic et savoir s'il a éventuellement développé le syndrome main pied comme l'illustre la figure ci-dessous.



Figure 3.13. Prédiction positive image importée (Hand syndrome)



Figure 3.14. Prédiction négative image importée (Hand No Syndrome)

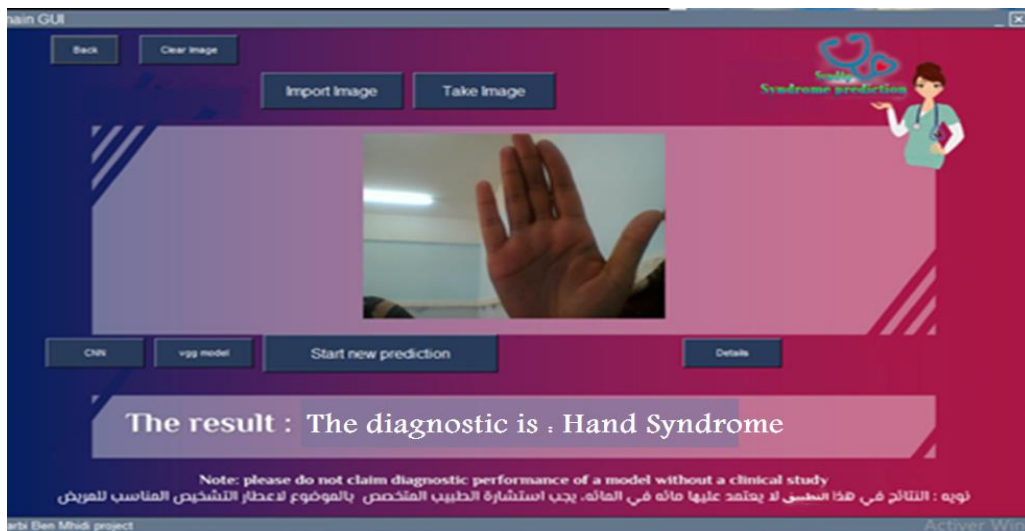


Figure 3.15. Prédiction positive image prise (Hand Syndrome)

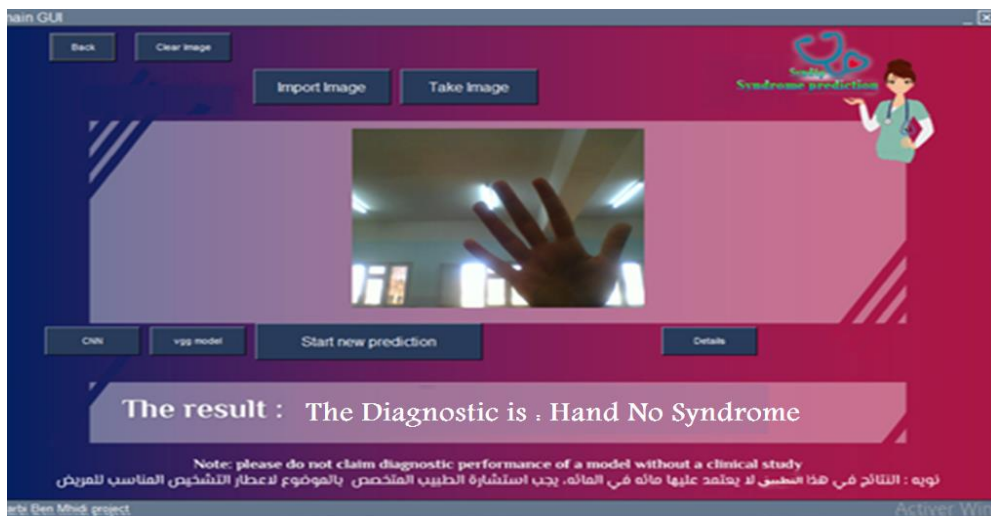


Figure 3.16. Prédiction négative image prise (Hand No Syndrome)

4. Conclusion

Au cours de ce chapitre, nous avons utilisé deux techniques essentielles : l'augmentation de données (data augmentation) et les modèles d'apprentissage de réseaux de neurones convolutifs (CNN) pour développer des modèles de prédiction du syndrome main-pied. Nous avons réalisé ces développements en utilisant le langage de programmation Python et l'environnement de développement Spyder.

L'augmentation de données a été utilisée pour augmenter la taille et la diversité de notre ensemble de données en appliquant des transformations telles que la rotation, le retournement, le redimensionnement, etc. Cela nous permet d'améliorer la capacité de notre modèle à généraliser et à reconnaître différents motifs.

Ensuite, nous avons utilisé les modèles d'apprentissage CNN, qui sont des réseaux neuronaux spécialisés dans le traitement de données avec une structure spatiale, telle que des images. Ces modèles sont particulièrement adaptés pour capturer les caractéristiques visuelles et les relations spatiales importantes dans les images du syndrome main-pied. Nous avons entraîné ces modèles en utilisant notre ensemble de données augmenté et avons ajusté les hyperparamètres pour obtenir de bonnes performances de prédiction.

Enfin, nous avons utilisé l'un de ces modèles entraînés pour développer une application de prédiction du syndrome main-pieds, spécifiquement une application de bureau. Cette application permet aux utilisateurs de charger une image et d'obtenir une prédiction sur la présence ou l'absence du syndrome main-pieds.

En combinant ces différentes étapes, nous avons créé un système complet qui utilise l'augmentation de données, les modèles CNN et une application de bureau pour prédire le syndrome main-pieds de manière efficace. Cela démontre l'utilisation pratique des techniques d'apprentissage automatique et des outils de développement Python pour résoudre des problèmes médicaux et améliorer les soins de santé.

Conclusion générale

La classification d'images est une tâche fondamentale dans le domaine de la vision par ordinateur et de l'apprentissage automatique. La notation automatique des images consiste à attribuer automatiquement une catégorie à une image à l'aide d'un système de notation. On retrouve ainsi la classification des objets, des scènes, la reconnaissance des visages et des maladies et des empreintes digitales.

Notre travail consiste à développer une application intelligente qui puisse prédire si une personne cancéreuse est atteinte ou non du syndrome main pied. Ce dernier est reconnu comme un effet indésirable dû à la prise des médicaments anti-cancéreux oraux.

Nous considérons que cet axe de recherche est plutôt nouveau, vu que nous ne prédisons pas la maladie, mais nous prédisons un effet secondaire résultant du traitement de la maladie, qui dans ce cas est le cancer.

Dans notre travail, nous avons utilisé l'apprentissage profond, une méthode d'apprentissage qui a montré ses performances ces dernières années. Nous avons choisi comme méthode de classification le réseau de neurones convolutifs (CNN). Ce choix se justifie par son efficacité.

Nous n'avons trouvé aucune base d'images dédiée à la reconnaissance de ce syndrome, ce qui nous a amené à collecter des images représentant ce syndrome. Le nombre des images collectées étant insuffisant pour pouvoir développer un modèle d'apprentissage suffisamment efficace, nous avons décidé d'utiliser la technique d'augmentation des données hors ligne dans le but d'augmenter le nombre des images utilisées dans le processus de développement d'un modèle d'apprentissage. En ce qui concerne les images représentant des mains saines, en d'autres termes des images ne représentant pas le syndrome main pied, la base d'images Hand11k a été utilisée.

Le système de reconnaissance du syndrome main pied développé a été implémenté en utilisant le langage de programmation Python. Les résultats expérimentaux obtenus sont très encourageants.

Notre travail n'en est qu'à sa version initiale, on peut dire que ce travail est encore ouvert à la comparaison et/ou à l'hybridation avec d'autres méthodes de classification et obtenir ainsi de meilleurs résultats.

Bibliographie

Ah-Pine, J. (2019/2020). *Apprentissage automatique*. Université Lyon 2, Lyon.

Data Science Team. (2020, mai). *Qu'est-ce que l'apprentissage automatique ? Une définition*. Retrieved 04 17, 2021, from datascience: <https://datascience.eu/fr/apprentissage-automatique/quest-ce-que-lapprentissage-automatique-une-definition/>

(ENI. 2020). *Intelligence artificielle vulgarisée - Le Machine Learning et le Deep Learning par la pratique*. Retrieved 04 19, 2021, from .editions-eni: <https://www.editions-eni.fr/open/mediabook.aspx?idR=4a1d225f6d8933d9887444217128e604>

(Nadia, M. 2013). *Apprentissage Incrémental*. Université HADJ LAKHDAR – BATNA, Département d'Informatique, Batna .

(techno-science. 2021, 4 16). *Apprentissage automatique - Définition et Explications*. Retrieved 4 18, 2021, from techno-science: <https://www.techno-science.net/glossaire-definition/Apprentissage-automatique.html>

(WHO. 2021). *Les maladies chroniques*. Retrieved from Organisation mondiale de la santé: https://www.who.int/topics/chronic_diseases/ar/#

(play.google.com,2021).<https://play.google.com/store/apps/details?id=co.medgic.medgic&hl=fr&gl=US&pli=1>

Mainenti, G., Campanile, L., Marulli, F., Ricciardi, C., & Valente, A. (2020, 1 20). Machine Learning Approaches for Classification: *reaserchgate*.

Text Data Augmentation for Deep Learning ink.springer.com/article/10.1186/s40537-021-00492-0 Journal of Big Data volume 8, Article number: 101 (2021)

(UQAR-INFO, 2022) Prédire les maladies cardiovasculaires avec l'intelligence artificielle Mardi 27 septembre 2022 Par Jean-François Bouchard

(<https://pypi.org/project/PySimpleGUI>, 2023)

[www.https://pypi.org/project/PySimpleGUI/](https://pypi.org/project/PySimpleGUI/) 22 mai 2023

(pourquoidocteur.fr, 2023) <https://www.pourquoidocteur.fr/Articles/Question-d-actu/42471-Cancer-poumon-IA-capable-predire-risque-jusqu-a-6-ans-l-avance>

<https://siecledigital-fr.cdn.ampproject.org/c/s/siecledigital.fr/2019/06/27/deep-learning-un-algorithme-du-mit-peut-predire-le-cancer-5-ans-a-lavance/?amp>

<https://www.roche.fr/fr/patients/info-patients-cancer/effets-secondaires-traitement-cancer/syndrome-main-pied-cancer.html>

Claw, S. (2020, 06 23). Python. Retrieved 07 19, 2020, from TkInter:

(cancer.be) <https://www.cancer.be › le-cancer › effets-secondaires>

(LE CORGNE Aude ,2016) THESE Présentée à la Faculté de Pharmacie de Dijon pour l'obtention du Diplôme d'Etat de Docteur en Pharmacie soutenue publiquement le 25/11/2016

(Journal of Big Data ,2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, Journal of Big Data, Published: 31 March 2021)

(docteur-eric-sebban.fr)<https://www.docteur-eric-sebban.fr> Traitement - août 24, 2020

(Data Analytics, 2023) AI, Data, Data Science, Machine Learning, Blockchain, Digital May 12, 2023 by Ajitesh Kumar

(Data Scientist , 2022) Michael Page - Cabinet de recrutement en France, lien vers la page d'accueil Fiche métier : Data Scientist Michael Page (2022)

M-FR-00003142-1.0 - Etabli en janvier 2021

(<https://openclassrooms.com>,2023) <https://openclassrooms.com/fr/courses/4011851-initiez-vous-au-machine-learning/4020611-identifiez-les-differents-types-dapprentissage-automatiques>

Panorama des cancers en France - Edition 2022