

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université 20 Août -1955 - Skikda

Faculté des sciences - Département d'Informatique



Mémoire de fin d'études en vue de l'obtention du diplôme de

MASTER en INFORMATIQUE

Option : Génie Logiciel Avancé et Applications

Thème

Systeme pour la prédiction des
maladies de la peau.

Réalisé par :

Berrahal Aymen
Bahloul Moussab Seif el Islam

Encadreur.

Dr. MANSOUL A.

Juin 2023

Remerciements

Au terme de ce travail, nous devons remercier tout d'abord notre dieu qui nous a donné la force et le courage de suivre nos études et d'arriver à ce stade et à nos parents qui nous ont beaucoup soutenus pendant tout notre parcours.

Un grand merci à notre encadreur Mr MANSOUL A. qui nous a beaucoup aidé, soutenu et nous a permis d'arriver à ce niveau et pour ses conseils et pour son temps passé avec nous et sa patience, sans lui on n'aurait pas pu réaliser ce modeste travail et pour sa confiance en nous.

Nous tenons aussi à remercier les membres du jury qui nous ont fait l'honneur d'examiner ce travail.

Enfin, nos remerciements vont à ceux qui nous ont aidés de près ou de loin pour réaliser ce travail sans oublier les enseignants qui ont contribué à notre formation.

Dédicase

Je dédie ce modeste travail à ma mère qui m'a donné la vie et qui incarne la tendresse et le sacrifice pour mon bonheur et ma réussite. À mon père, en tant qu'école de mon enfance, a été mon ombre bienveillante tout au long de mes années d'études, toujours là pour m'encourager, m'aider et me protéger. À mes soeurs et à mon frère Moudou que j'aime. Je tiens à exprimer ma gratitude envers mon coéquipier Aymen Berrahal, avec qui j'ai partagé des moments à la fois joyeux et stressants. À toute ma famille, qu'elle soit petite ou grande, ainsi qu'à mes chers amis oussama tachi, Loulou, Seif, Ramy, Moncef, Hichem, Sohaib, Chouaib, Youcef.... je vous suis profondément reconnaissant pour votre soutien indéfectible et votre amour inconditionnel.

Moussaab Seif el islam Bahloul

Dédicase

Je dédie ce modeste travail à ma mère qui m'a donnée la vie et qui incarne la tendresse et le sacrifice pour mon bonheur et ma réussite. Mon père, en tant qu'école de mon enfance, a été mon ombre bienveillante tout au long de mes années d'études, toujours là pour m'encourager, m'aider et me protéger. A mes soeurs et mes frères. Je tiens à exprimer ma gratitude envers mon binôme Moussab Bahloul, avec qui j'ai partagé des moments à la fois joyeux et stressants. A toute ma famille, qu'elle soit petite ou grande, ainsi qu'à mon cher ami Nouh auquel je prie pour une miséricorde de Dieu. Que dieu bénisse son âme et lui pardonne in sha Allah.

Aymen Berrahal

Résumé

Les maladies de la peau sont considérées comme l’une des maladies les plus répandues dans le monde et les plus contagieuses et infectieuses. Le diagnostic de ces maladies dépend de tests médicaux et de certains indicateurs souvent apparents et même parfois cachés, ce qui rend le diagnostic un peu plus compliqué, en raison du grand nombre d’analyses et d’indicateurs dans ce diagnostic

Cela a conduit de nombreux chercheurs à explorer plusieurs techniques basées sur des données brutes pour prédire ces maladies au vu des caractéristiques apparentes sur la peau. Ceci nous a incités à nous pencher vers une approche basée sur les symptômes apparents sur la peau afin de prédire la maladie en question.

Notre approche se base sur un data set d’images qui présentent plusieurs lésions de la peau et qui touchent les maladies de la peau telles que : le psoriasis, l’eczéma, etc.

Nous avons utilisé une approche basée sur l’analyse de différentes images, et pour un cas donné (patient) nous utiliserons un réseau de neurones convolutif pour prédire la maladie de la peau en question.

Notre approche nous a permis d’atteindre un taux de prédiction de **87%** . .

Mots clés : maladie, peau, prédiction, CNN, réseau, neurone.

Abstract

Skin diseases are considered to be one of the most widespread and contagious/infectious diseases worldwide. The diagnosis of these diseases relies on medical tests and certain often apparent, and sometimes hidden, indicators, which make the diagnosis somewhat more complicated due to the large number of analyses and indicators involved.

This has led many researchers to explore several techniques based on raw data to predict these diseases based on visible skin characteristics. This has prompted us to adopt an approach based on the apparent symptoms on the skin to predict the specific disease.

Our approach is based on a dataset of images depicting various skin lesions associated with skin diseases such as psoriasis, eczema, etc. We have employed an approach that involves analyzing different images, and for a given case (patient), we will utilize a convolutional neural network (CNN) to predict the respective skin disease.

Our approach has achieved a prediction rate of **87%**.

keywords : : disease, skin, prediction, CNN, network, neural.

Table des matières

Introduction Générale	1
1 L'image	3
1.1 Introduction	3
1.2 Les types d'images	4
1.2.1 Les images à deux niveaux	4
1.2.2 les images à plusieurs niveaux de gris	5
1.2.3 Les images couleurs	6
1.3 Les caractéristiques d'une image	7
1.3.1 Le pixel	7
1.3.2 La définition d'une image	8
1.3.3 La résolution d'une image	8
1.3.4 La profondeur d'une image	9
1.3.5 La taille d'une image	9
1.3.6 Le format d'une image	9
1.3.7 La qualité d'une image	10
1.4 Conclusion	10
2 Etat de l'art en prédiction des maladies	12

2.1	Introduction	12
2.2	la prédiction des maladies	12
2.3	Les maladies de la peau	13
2.4	Prédiction des maladies de la peau	14
2.5	l'état de l'art	14
2.6	Conclusion	17
3	Les réseaux de neurones convolutionnels (CNN)	18
3.1	Introduction	18
3.2	Les réseaux de neurones convolutionnels (CNN)	18
3.2.1	Les couches du réseaux de neurones convolutionnels (CNN)	19
3.3	Utilisation du CNN en traitement d'images	21
3.4	Conclusion	21
4	Conception et Implémentation	23
4.1	Introduction	23
4.2	Architecture générale du système	23
4.3	Description générale	25
4.4	Sélection de données	25
4.5	Le modèle de classification	25
4.6	Prétraitement des données	26
4.6.1	Redimensionnement	27
4.6.2	Mis les images en niveau de gris	27
4.6.3	Encodage des données	28
4.6.4	Normalisation des données	29
4.6.5	Régularisation des données	29

4.6.6	Subdiviser les données pour l'entraînement et le test	30
4.7	L'entraînement de modèle	30
4.7.1	Le fonctionnement	31
4.7.2	La fonction d'activation	31
4.7.3	La fonction de perte	34
4.7.4	L'algorithme d'optimisation	34
4.8	Plateforme matérielle et logicielle utilisée	35
4.8.1	Plateforme matérielle	35
4.8.2	Plateforme logicielle	36
4.9	L'analyse du résultat	40
4.9.1	Résultat de l'entraînement	40
4.9.2	Évaluation des résultats de la perte (Loss)	41
4.9.3	Évaluation des résultats de la précision (Accuracy)	42
4.9.4	Évaluation du résultat de la matrices de confusion	43
4.10	Conclusion	44

Table des figures

1.1 image à deux niveaux.[2]	5
1.2 image à plusieurs niveaux de gris	6
1.3 image couleur.[3]	7
1.4 représentation du pixels.[6]	8
2.1 exemple pour une maladie de la peau.[16]	14
3.1 Schéma d'une convolution [25]	19
3.2 Schéma d'une convolution pooling [25]	20
3.3 Schéma d'une couche entièrement connecté avec 6 classes [25]	21
4.1 Architecture générale du système proposé	24
4.2 Architecture générale du système proposé	26
4.3 Subdiviser les données	30
4.4 La fonction RELU	33
4.5 La fonction Softmax	33
4.6 Python logo	37
4.7 Keras logo	37
4.8 Numpy logo	38
4.9 Pandas logo	38

4.10 Google colab logo	39
4.11 Tensorflow logo	39
4.12 Sckit logo	40
4.13 La résultat d'évaluation du modèle	41
4.14 La résultat d'évaluation du modèle	41
4.15 La résultat d'évaluation du modèle	42
4.16 La résultat d'évaluation du modèle	43

Liste des tableaux

4.1 L'architecture du modèle	31
4.2 La Configuration du Colab	36
4.3 Les caractéristiques des machines utilisées	36

Introduction Générale

Les maladies de la peau constituent un problème de santé majeur qui affecte des millions de personnes dans le monde. Ces affections cutanées peuvent engendrer des symptômes gênants, altérer la qualité de vie et, dans certains cas, entraîner des complications graves si elles ne sont pas détectées et traitées précocement. La détection précoce et le diagnostic précis des maladies de la peau revêtent donc une grande importance pour assurer des soins de santé efficaces et améliorer les résultats cliniques.

Avec les avancées technologiques récentes, notamment dans le domaine de l'intelligence artificielle et de la vision par ordinateur, de nouvelles possibilités ont émergé pour développer des systèmes de détection des maladies de la peau plus rapides, précis et fiables. Ces systèmes exploitent les capacités de l'intelligence artificielle pour analyser des images de la peau, identifier les caractéristiques spécifiques des différentes affections cutanées et fournir des diagnostics préliminaires.

Notre projet de fin d'étude vise à concevoir et à développer un système de détection des maladies de la peau basé sur l'intelligence artificielle. Ce système utilisera des techniques avancées de traitement d'images et d'apprentissage automatique pour analyser des images dermatologiques et fournir des évaluations préliminaires des affections cutanées présentes.

L'objectif principal de ce projet est de fournir aux dermatologues et aux professionnels de la santé un outil d'aide au diagnostic précis et efficace. Ce système de détection des maladies de la peau permettra d'accélérer le processus de diagnostic, d'améliorer la précision des résultats et de faciliter

la prise de décision clinique. Il peut également être utilisé comme un outil d'auto-évaluation pour les patients, leur permettant de surveiller l'évolution de leurs affections cutanées et de prendre des mesures appropriées en temps opportun.

Chapitre 1

L'image

1.1 Introduction

Les images sont des représentations visuelles qui ont toujours été au coeur de la communication humaine. Depuis les peintures rupestres et les fresques murales des temps anciens jusqu'aux images numériques sophistiquées d'aujourd'hui, les images ont été utilisées pour raconter des histoires, transmettre des informations, exprimer des émotions et embellir notre monde. Les images sont des outils puissants qui nous permettent de communiquer rapidement et efficacement des idées complexes, des concepts abstraits et des informations scientifiques et techniques.

Elles sont également utilisées dans les domaines de l'art, du divertissement et de la publicité pour captiver l'attention du public et stimuler la créativité. Dans cette section, nous allons explorer les différentes formes d'images, les techniques de création et les principes de conception qui permettent d'utiliser les images de manière optimale pour renforcer l'impact de votre message.

Nous allons également discuter des avantages et des limites de l'utilisation des images dans différents contextes de communication, ainsi que des défis et des opportunités offerts par les technologies émergentes, telles que l'intelligence artificielle et la réalité virtuelle.

Définition : Une image est une représentation visuelle d'un objet, d'une personne, d'un paysage ou d'une idée exprimée en deux dimensions. Elle peut être produite par divers moyens tels que la peinture, la photographie ou la numérisation. Les images sont utilisées pour communiquer des idées, évoquer des émotions ou fournir une représentation visuelle d'un objet ou d'un concept. Elles peuvent également être utilisées à des fins artistiques, éducatives, publicitaires ou informatives. [\[1\]](#)

1.2 Les types d'images

1.2.1 Les images à deux niveaux

Les images à deux niveaux, également appelées images binaires, sont un type particulier d'image où chaque pixel peut avoir seulement deux valeurs possibles : noir ou blanc (0 ou 1). Ces images sont souvent utilisées pour des applications où une distinction binaire est nécessaire, telles que la reconnaissance de caractères, la détection de contours, et la segmentation d'image.

Les images à deux niveaux peuvent être créées à partir de n'importe quelle image en convertissant les niveaux de gris en noir et blanc à l'aide d'un seuillage. Le seuillage est un processus qui compare chaque pixel à une valeur seuil prédéfinie et qui décide si le pixel doit être noir ou blanc en fonction de cette valeur seuil.

Bien que les images à deux niveaux soient plus simples que les images en niveaux de gris ou en couleurs, elles peuvent être très utiles pour des applications spécifiques. Par exemple, elles peuvent être utilisées pour extraire des informations sur la forme et la texture des objets dans une image. En outre, elles nécessitent moins de stockage et de traitement que les images en niveaux de gris ou en couleurs, ce qui les rend idéales pour les applications où la vitesse et l'efficacité sont des facteurs critiques.



FIGURE 1.1 – image à deux niveaux.[\[2\]](#)

1.2.2 les images à plusieurs niveaux de gris

Les images à plusieurs niveaux de gris sont un type d'image où chaque pixel peut prendre une valeur de gris entre 0 (noir) et un nombre défini de niveaux de gris (généralement 256 pour une image en niveaux de gris standard). Ces images sont couramment utilisées pour des applications telles que la photographie, la télédétection, la microscopie et l'imagerie médicale.

Dans les images à plusieurs niveaux de gris, chaque pixel peut avoir une valeur de niveau de gris différente. Cette valeur de niveau de gris représente l'intensité de la lumière réfléchie ou transmise par le pixel. Les images à plusieurs niveaux de gris peuvent être considérées comme des images en noir et blanc avec une gamme de nuances de gris. Elles sont souvent utilisées pour capturer des détails fins et des variations subtiles dans les images.

Les images à plusieurs niveaux de gris sont couramment utilisées dans des applications de traitement d'image telles que la segmentation, la détection

de bordures, la mesure de la texture, la compression d'image, et la reconnaissance d'objets. Les algorithmes de traitement d'image peuvent utiliser ces images pour extraire des informations sur la texture, la forme, la taille et l'orientation des objets dans une image .

Les images à plusieurs niveaux de gris sont également utilisées en imagerie médicale pour la visualisation de structures anatomiques et de pathologies. Par exemple, les images à plusieurs niveaux de gris peuvent être utilisées pour visualiser des tumeurs, des fractures osseuses, ou des lésions cérébrales.

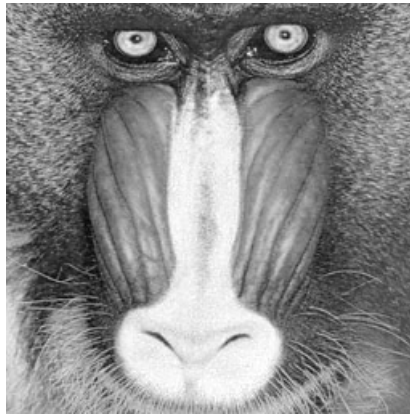


FIGURE 1.2 – image à plusieurs niveaux de gris [3].

1.2.3 Les images couleurs

Les images couleurs sont un type d'image qui utilise une combinaison de trois couleurs primaires (rouge, vert et bleu) pour créer une image en couleur. Chaque pixel d'une image couleur contient trois valeurs de couleur, chacune représentant l'intensité de la lumière réfléchie dans la gamme de couleurs rouge, verte et bleue [4].

Les images couleurs sont utilisées dans de nombreuses applications, telles que la photographie, la vidéo, l'imagerie médicale, l'imagerie satellite et la conception graphique [4].

Elles sont capables de représenter une gamme étendue de couleurs et

sont donc bien adaptées aux applications qui nécessitent des images avec une haute résolution de couleur [4].

Les images couleurs peuvent être traitées à l'aide d'algorithmes de traitement d'image pour extraire des informations sur la couleur, la texture, la forme et la taille des objets dans une image. Les applications de traitement d'image utilisant des images couleurs incluent la reconnaissance d'objets, la segmentation d'image, la classification d'image, et la détection de contours [4].



FIGURE 1.3 – image couleur.[3]

1.3 Les caractéristiques d'une image

1.3.1 Le pixel

Un pixel (ou "Picture element") est la plus petite unité d'une image numérique. C'est un point unique sur une grille régulière qui compose l'image. Chaque pixel est défini par une valeur numérique qui correspond à l'intensité de la couleur ou de la luminosité de ce point dans l'image. Dans les images en noir et blanc, chaque pixel est représenté par une seule valeur, généralement de 0 (noir) à 255 (blanc) [5].

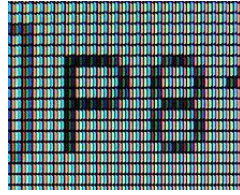


FIGURE 1.4 – représentation du pixels. [6]

1.3.2 La définition d'une image

La définition est le nombre de pixels composant une image. Pour calculer la définition d'une image numérique, il suffit de multiplier le nombre de pixels sur la hauteur par le nombre de pixels sur la largeur de l'image [7].

Par exemple : Une image de définition 800 X 600 est composée de 800 pixels en largeur et de 600 pixels en hauteur, soit 480 000 pixels ou 0.48 Méga Pixels [7].

1.3.3 La résolution d'une image

La résolution d'une image est une mesure de la densité de pixels dans une image numérique. Elle correspond au nombre total de pixels qui composent l'image. Plus précisément, la résolution est définie comme le nombre de pixels dans l'image horizontalement multiplié par le nombre de pixels verticalement. Elle est souvent exprimée en pixels par pouce (ppp) ou en pixels par centimètre (ppc) .

Une image de haute résolution aura une plus grande quantité de pixels par unité de surface, ce qui la rend plus nette, plus détaillée et plus adaptée aux impressions de grande qualité. En revanche, une image de faible résolution aura moins de pixels par unité de surface, ce qui la rend plus floue, moins détaillée et moins adaptée aux impressions de grande qualité [8].

1.3.4 La profondeur d'une image

La profondeur d'une image, également connue sous le nom de profondeur de couleur ou de bit depth, est une mesure de la quantité d'informations de couleur qui peut être stockée dans chaque pixel d'une image numérique [8].

La profondeur d'une image est exprimée en bits et représente le nombre de bits de données de couleur stockés pour chaque pixel de l'image. Par exemple, une image en noir et blanc de 8 bits a une profondeur de couleur de 8 bits, ce qui signifie que chaque pixel de l'image peut stocker l'une des 256 nuances de gris ($2^8 = 256$). [8].

Plus la profondeur d'une image est élevée, plus elle peut stocker d'informations de couleur et plus elle est capable de reproduire des dégradés de couleur subtils. Les images de haute qualité nécessitent souvent une profondeur de couleur élevée pour obtenir une reproduction précise des couleurs [8].

1.3.5 La taille d'une image

La taille d'une image est une mesure de l'espace de stockage occupé par l'image numérique. Elle est généralement exprimée en kilooctets (ko) ou en mégaoctets (Mo) [9].

La taille d'une image dépend de plusieurs facteurs, notamment la résolution de l'image, la profondeur de couleur, le format de fichier et la qualité de compression. Plus la résolution de l'image est élevée, plus elle contient de pixels et plus sa taille sera grande. De même, une profondeur de couleur plus élevée entraîne également une taille de fichier plus importante [9].

1.3.6 Le format d'une image

Le format de fichier joue également un rôle important dans la taille de l'image. Certains formats de fichier, comme JPEG, utilisent des techniques de compression pour réduire la taille du fichier sans perdre de qualité d'image

significative. D'autres formats, comme BMP ou TIFF, peuvent ne pas utiliser de compression, ce qui peut entraîner des tailles de fichier plus importantes .

En général, il est important de trouver un équilibre entre la qualité d'image et la taille de fichier pour garantir que l'image soit facilement accessible et facile à utiliser sans prendre trop de place sur le disque dur .

Le format de fichier et la compression peuvent également affecter la qualité de l'image. Certains formats de fichier, comme JPEG, utilisent des techniques de compression qui peuvent entraîner une perte de qualité d'image. D'autres formats, comme PNG ou TIFF, utilisent des méthodes de compression sans perte pour conserver une qualité d'image maximale **rasch2009interactive**.

1.3.7 La qualité d'une image

La qualité d'une image est une mesure de sa fidélité par rapport à l'original ou à l'objectif visé. Elle est généralement évaluée en termes de netteté, de précision des couleurs, de clarté et de détails [10].

La qualité d'une image dépend de plusieurs facteurs, notamment la résolution, la profondeur de couleur, le format de fichier, la compression et le traitement numérique. Plus la résolution de l'image est élevée, plus les détails sont clairs et plus l'image est nette. De même, une profondeur de couleur plus élevée permet d'obtenir une reproduction plus précise des couleurs [10].

Le traitement numérique, tel que la retouche d'image ou la réduction du bruit, peut également affecter la qualité de l'image. Une manipulation excessive de l'image peut entraîner une perte de détails ou une altération de la couleur [10].

1.4 Conclusion

Dans ce chapitre, nous avons abordé le sujet des images, en mettant l'accent sur leur importance, leurs types et leurs caractéristiques. Dans

le chapitre suivant nous allons expliqué la prédiction des maladies et voir quelque travaux.

Chapitre 2

Etat de l'art en prédiction des maladies

2.1 Introduction

Ce chapitre se concentre sur la prédiction des maladies et l'état de l'art. Les maladies dermatologiques ont un impact considérable sur la qualité de vie des individus, et il est essentiel d'effectuer un diagnostic précoce pour une prise en charge efficace. Dans ce contexte, le Deep Learning offre de nouvelles perspectives pour améliorer les capacités de prédiction et de diagnostic .

Nous explorons les avancées récentes, les approches adoptées et les défis rencontrés dans ce domaine.

2.2 la prédiction des maladies

La prédiction des maladies a retenu l'attention des chercheurs pour la prédiction et l'amélioration des soins. Plusieurs travaux ont été menés, parmi lesquels nous avons

Shankar et al., ont utilisé les réseaux de neurones convolutifs (CNN) pour prédire avec précision si un individu est à risque de maladie cardiaque ou

non. Ils ont implémenté un modèle de prédiction sur des données hospitalières réelles en utilisant des données structurées des patients [11].

Rustam et al., ont utilisé les dossiers de santé électroniques pour diagnostiquer les maladies cardiovasculaires. Cette étude utilise un réseau de neurones convolutifs (CNN). Un modèle CNN est utilisé afin de classier quatre ensembles de données différents et les résultats des test ont montré une performance importante [12].

Pant et al., ont utilisé Les réseaux de neurones convolutifs (CNN) pour créer un système de prédiction et de diagnostic médical à un stade précoce pour les maladies du cur. Le CNN reçoit 13 caractéristiques cliniques en entrée pour prédire si un patient aura ou non une maladie cardiovasculaire. Au cours des tests, il a été découvert que CNN prédit l'absence et la présence de maladies cardiaques avec une précision supérieure à 95% [13].

Farooq et al., ont proposé un système basé sur un CNN pour la prédiction de la maladie d'Alzheimer à l'aide d'imagerie par résonance magnétique (IRM). Un classificateur est implémenté pour classer la maladie d'Alzheimer (AD), les troubles cognitifs légers (MCI), les troubles cognitifs légers tardifs (LMCI) et les personnes en bonne santé. La technique proposée a aboutit à une précision de prédiction importante qui révèle l'efficacité des CNN [14].

2.3 Les maladies de la peau

Les maladies de la peau sont des affections qui affectent la peau et ses annexes, comme les ongles, les cheveux et les glandes sudoripares. Elles peuvent être causées par une variété de facteurs, tels que des infections, des allergies, des maladies auto-immunes, des troubles génétiques, des traumatismes, l'exposition à des substances nocives et des facteurs environnementaux. Les maladies de la peau peuvent avoir des manifestations diverses, notamment des rougeurs, des démangeaisons, des éruptions cutanées, des desquamations, des lésions, des bosses, des plaies ou des ulcérations. Certaines maladies de la peau sont bénignes et se résolvent spontanément, tandis que d'autres peuvent

être chroniques, invalidantes et potentiellement mortelles. Le diagnostic et le traitement des maladies de la peau peuvent être complexes et nécessitent souvent l'intervention d'un dermatologue ou d'un autre spécialiste de la santé [15].



FIGURE 2.1 – exemple pour une maladie de la peau. [16]

2.4 Prédiction des maladies de la peau

Définition : La prédiction des maladies de la peau est une méthode d'analyse et de détection précoce des maladies de la peau grâce à l'utilisation de techniques d'apprentissage automatique et de modèles statistiques. Cette approche consiste à analyser les caractéristiques de la peau, telles que la couleur, la texture, la forme et la taille des lésions, afin de prédire le type de maladie de la peau avec une précision accrue.

Les données peuvent être collectées à partir de photographies de la peau, d'examen physiques ou de l'historique médical du patient. La prédiction des maladies de la peau peut aider les médecins à poser un diagnostic plus précis et à traiter les maladies de la peau plus rapidement, ce qui peut améliorer les résultats pour les patients [17].

2.5 l'état de l'art

— Prédiction des mélanomes

Des chercheurs ont développé des modèles de deep learning capables de prédire avec précision la présence de mélanomes, une forme agressive de cancer de la peau. En utilisant des réseaux neuronaux convolutifs entraînés sur de vastes ensembles de données d'images dermatologiques, ces modèles peuvent détecter des caractéristiques spécifiques associées aux mélanomes et les distinguer des lésions cutanées bénignes .

— **Classification des types d'eczéma**

L'eczéma est une maladie inflammatoire de la peau qui peut présenter différents sous-types. Des chercheurs ont utilisé le deep learning pour développer des modèles de classification capables de distinguer les différents types d'eczéma en analysant des images dermatologiques. Ces modèles permettent une classification précise, ce qui peut aider les dermatologues à déterminer les traitements appropriés pour chaque patient [18].

— **Détection des infections fongiques de la peau**

Les infections fongiques cutanées sont courantes et peuvent être difficiles à diagnostiquer avec précision. Des modèles de deep learning ont été développés pour détecter automatiquement les signes d'infections fongiques à partir d'images dermatologiques. Ces modèles peuvent identifier les caractéristiques spécifiques des infections fongiques et aider les professionnels de la santé à prendre des décisions de diagnostic plus éclairées [19].

— **Prédiction de la progression du psoriasis**

Le psoriasis est une maladie chronique de la peau caractérisée par des plaques rouges et squameuses. Des modèles de deep learning ont été utilisés pour prédire la progression et la gravité du psoriasis en analysant des données cliniques, des images dermatologiques et des facteurs de risque génétiques. Ces modèles permettent d'identifier les patients à haut risque de développer une forme plus sévère de la maladie et de proposer des interventions précoces pour améliorer les résultats [20].

— **Projet "Skin Cancer MNIST" :**

Ce projet a utilisé un ensemble de données appelé "Skin Cancer MNIST" contenant 10 classes de lésions cutanées, comprenant des mélanomes, des carcinomes basocellulaires et des carcinomes épidermoïdes. En utilisant des réseaux de neurones convolutifs (CNN), les chercheurs ont développé des modèles de deep learning capables de prédire avec précision la présence de différentes formes de cancer de la peau à partir d'images dermatologiques .

- **Projet "Dermatologist-Level Classification of Skin Cancer" :**
Cette étude, menée par des chercheurs de Google, a utilisé un réseau neuronal convolutif pour réaliser des diagnostics dermatologiques. Le modèle a été entraîné sur un vaste ensemble de données d'images de lésions cutanées provenant de différentes classes, et il a été capable de rivaliser avec les dermatologues humains en termes de précision diagnostique dans la classification de différentes affections cutanées [21].
- **Projet "Automated Dermatological Diagnosis" :**
Ce projet a développé un système automatisé de diagnostic dermatologique en utilisant le deep learning. Les chercheurs ont utilisé des réseaux de neurones profonds pour analyser des images de lésions cutanées et prédire différentes affections, y compris des infections fongiques, des verrues et des tumeurs bénignes. Le modèle a démontré une précision élevée dans la prédiction des affections cutanées, offrant un outil d'aide au diagnostic pour les dermatologues [22].
- **Projet "Automated Eczema Classification" :**
Cette recherche a utilisé le deep learning pour classer différents types d'eczéma à partir d'images dermatologiques. Les chercheurs ont entraîné un réseau neuronal convolutif à reconnaître les caractéristiques distinctives des différents sous-types d'eczéma, tels que l'eczéma atopique, l'eczéma de contact et l'eczéma dyshidrosique. Le modèle a obtenu de bons résultats dans la classification précise des différents types d'eczéma, offrant un soutien aux dermatologues dans le diagnostic et le traitement [23].

2.6 Conclusion

Dans ce chapitre, nous avons démontrons l'importance croissante de la prédiction des maladies et de l'apprentissage automatique dans le domaine de la prédiction des maladies de la peau. Les avancées technologiques récentes ont permis le développement de systèmes sophistiqués. Le chapitre suivant explique en détail le fonctionnement du réseaux de neurones convolutionnels (CNN).

Chapitre 3

Les réseaux de neurones convolutionnels (CNN)

3.1 Introduction

Dans ce chapitre, nous explorerons en détail Les Réseaux de Neurones Convolutionnels (CNN) qu'ont révolutionné notre capacité à comprendre, analyser et interpréter des données complexes telles que des images et des séquences Inspirés par le fonctionnement du cerveau humain.

3.2 Les réseaux de neurones convolutionnels (CNN)

Définition : Les réseaux de neurones convolutifs, ou CNN (Convolutional Neural Networks) en anglais, sont une sous-catégorie de réseaux de neurones artificiels spécialement conçus pour traiter des données d'entrée structurées, telles que des images. Ils sont largement utilisés dans le domaine de la vision par ordinateur pour des tâches telles que la classification d'images, la détection d'objets,...etc

3.2.1 Les couches du réseaux de neurones convolutionnels (CNN)

Les Réseaux de Neurones Convolutifs (CNN) sont composés de plusieurs éléments qui travaillent ensemble pour effectuer des tâches de traitement des données. Voici les principaux couches d'un CNN

Couche de convolution (CONV)

La couche de convolution est le bloc de base d'un réseau de neurones convolutionnel. Elle est responsable de l'extraction des caractéristiques visuelles en appliquant des opérations de convolution sur l'image d'entrée à l'aide de filtres. Chaque filtre est utilisé pour détecter des motifs spécifiques dans l'image, tels que des bords, des textures ou des formes [24].

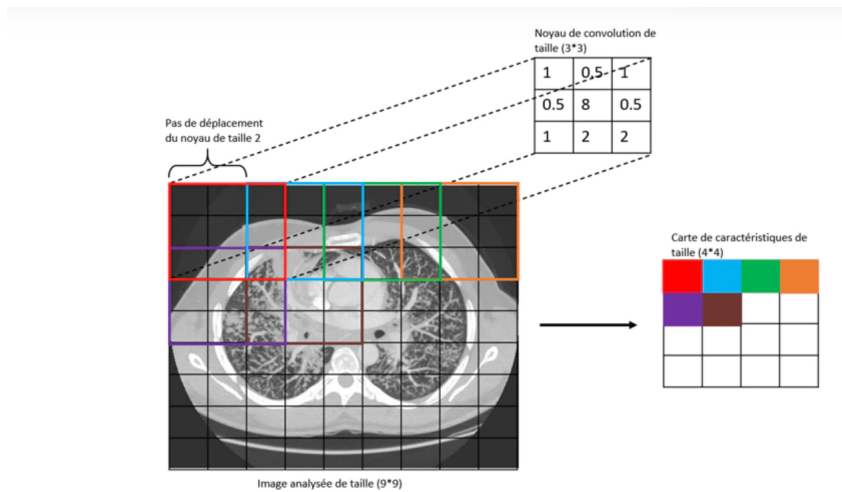


FIGURE 3.1 – Schéma d'une convolution [25]

Couche de pooling (POOL)

La couche de pooling est utilisée pour réduire la dimensionnalité des caractéristiques extraites par la couche de convolution. Elle regroupe les valeurs les plus significatives dans des régions spécifiques, ce qui permet de réduire

le nombre de paramètres et de rendre le modèle plus robuste aux variations mineures dans les données d'entrée [24].

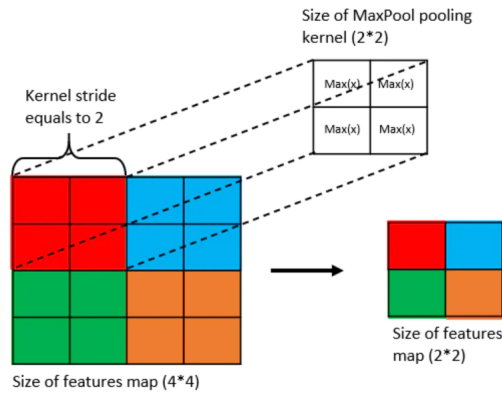


FIGURE 3.2 – Schéma d’une convolution pooling [25]

Couches de correction (ReLU)

Les couches de correction utilisent une fonction d’activation appelée Rectified Linear Unit (ReLU) pour introduire de la non-linéarité dans le réseau. Cette non-linéarité permet au modèle de capturer des relations complexes et des interactions entre les différentes caractéristiques extraites [26].

Couche entièrement connectée (FC)

La couche entièrement connectée est située à la fin du réseau et est responsable de la classification finale ou de la prédiction. Chaque neurone de cette couche est connecté à tous les neurones de la couche précédente, ce qui permet de combiner les caractéristiques apprises pour effectuer une classification précise [26].

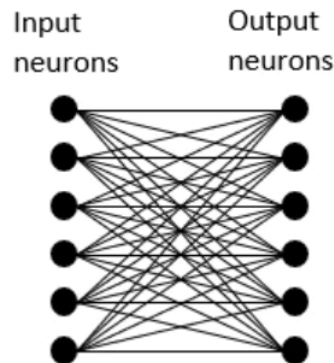


FIGURE 3.3 – Schéma d'une couche entièrement connectée avec 6 classes [25]

3.3 Utilisation du CNN en traitement d'images

L'utilisation des CNN en traitement d'images représente une avancée majeure dans de nombreux domaines, y compris la prédiction des maladies de la peau. Dans cette étude, nous avons exploré l'application des CNN pour analyser et interpréter les images dermatologiques. On n'oublie pas que cette utilisation offre des perspectives prometteuses. Les résultats obtenus jusqu'à présent ont montré des performances encourageantes, avec des modèles de CNN surpassant parfois les experts humains dans la classification et la détection de maladies dermatologiques.

3.4 Conclusion

En conclusion, l'utilisation du deep CNN traitement d'images représente une approche puissante et innovante pour prédire les maladies de la peau. Grâce aux CNN et à d'autres techniques de deep learning, nous sommes en mesure d'extraire des informations précieuses à partir d'images dermatologiques, ce qui peut contribuer à un diagnostic précoce et à une prise en charge efficace des patients ; Dans le chapitre suivant, nous allons tenter de mettre en œuvre la solution en utilisant plusieurs méthodes et discuterons des

différents avantages et inconvénients de chacune d'entre elles.

Chapitre 4

Conception et Implémentation

4.1 Introduction

La conception et l'implémentation d'un système de détection et de prédiction des maladies de la peau représentent une avancée majeure dans le domaine médical. Ce système vise à fournir un outil précieux pour un diagnostic précoce et une prise en charge efficace des patients. Dans ce chapitre, nous explorerons les aspects clés de la conception et de l'implémentation d'un tel système, en mettant l'accent sur les choix techniques et les considérations pratiques nécessaires pour développer une solution efficace.

4.2 Architecture générale du système

Dans notre application, nous avons choisi une architecture qui favorise une représentation abstraite du système ainsi qu'une communication entre ses divers composants, illustrée dans la figure 4.1.

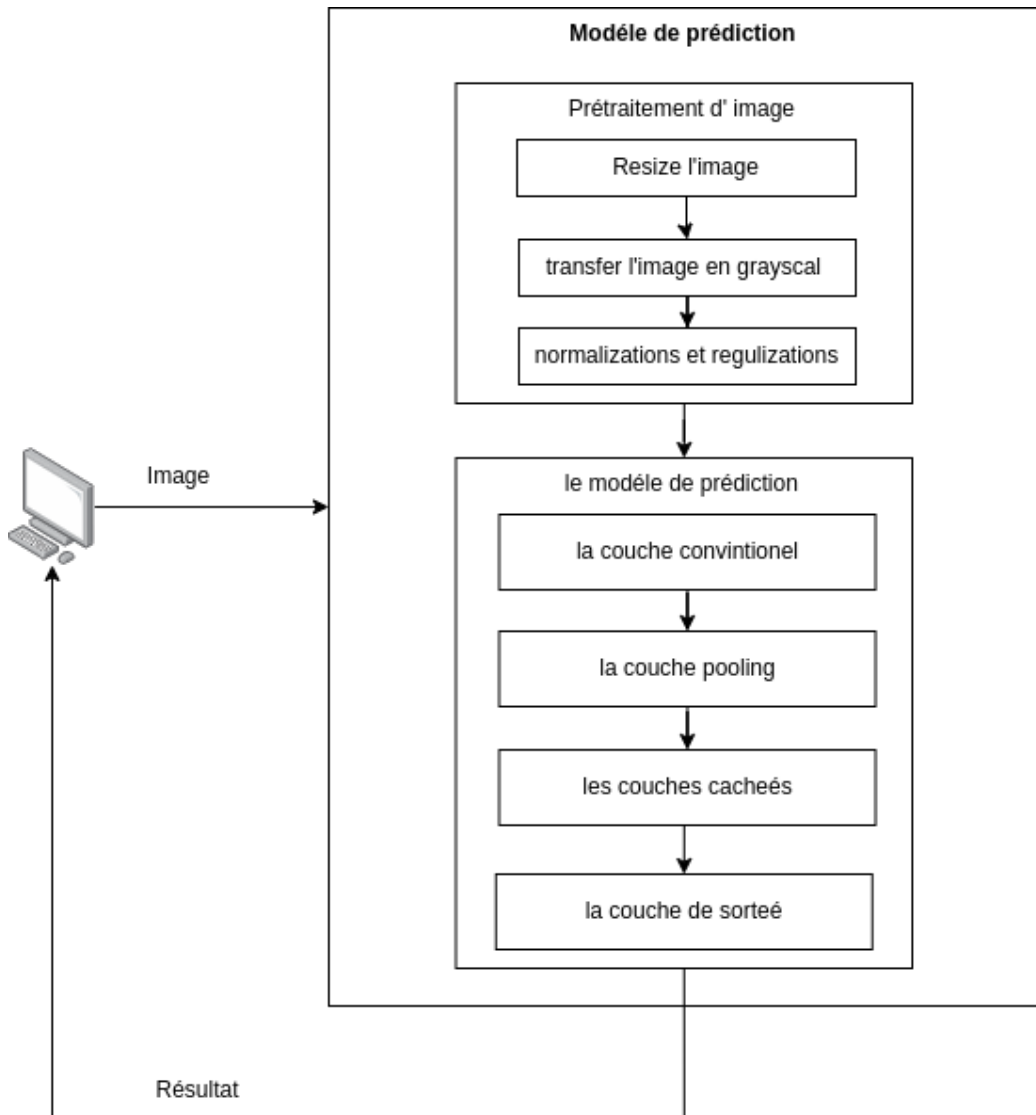


FIGURE 4.1 – Architecture générale du système proposé

4.3 Description générale

4.4 Sélection de données

Pour notre étude, nous avons utilisé le dataset appelé Dermnet. Ce dataset est largement utilisé dans le domaine de la prédiction des maladies de la peau. Le dataset Dermnet est composé de plus de 19 500 images et contient une variété de 23 types de maladies.

4.5 Le modèle de classification

Pour créer un modèle de classification, nous utilisons un réseau neuronal convolutif (CNN), en suivant les étapes indiquées dans la figure 4.2.

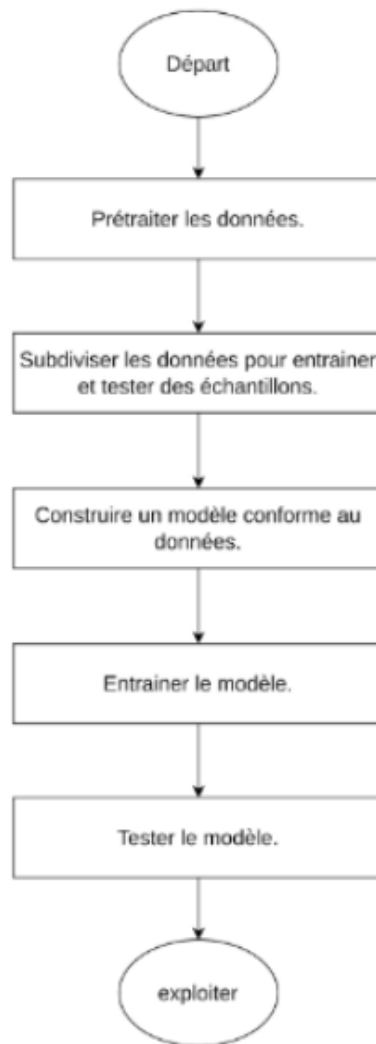


FIGURE 4.2 – Architecture générale du système proposé

4.6 Prétraitement des données

Lors de la création d'un modèle de classification, il est essentiel de pré-traiter les données, notamment par le biais de techniques telles que le redimensionnement et la normalisation. Ces étapes sont cruciales pour garantir des résultats précis et fiables lors de l'apprentissage du modèle.

4.6.1 Redimensionnement

Nous redimensionnons les images dans le cadre du prétraitement des données pour plusieurs raisons importantes

- **Cohérence de la taille** Les images peuvent avoir différentes tailles d'origine, ce qui peut entraîner des difficultés lors de l'entraînement du modèle. En redimensionnant les images, nous pouvons les rendre toutes de la même taille, ce qui facilite le traitement et l'analyse ultérieure.
- **Réduction de la complexité** Les images haute résolution peuvent contenir une quantité excessive de détails qui ne sont pas pertinents pour la tâche de classification. En les redimensionnant à une résolution appropriée, nous réduisons la complexité des données, ce qui peut améliorer l'efficacité et les performances du modèle.
- **Réduction des exigences en matière de mémoire et de calcul** Les images de grande taille occupent plus d'espace mémoire et nécessitent davantage de puissance de calcul lors de l'entraînement du modèle. En redimensionnant les images à une taille plus petite, nous réduisons les exigences en matière de mémoire et de calcul, ce qui permet de traiter un plus grand nombre d'images dans un temps raisonnable.
- **Gestion des variations d'échelle** Dans certaines tâches de classification, il peut être important de détecter des motifs ou des objets à différentes échelles. En redimensionnant les images de manière cohérente, nous facilitons la détection de ces motifs à différentes échelles, ce qui peut améliorer les performances du modèle.

4.6.2 Mis les images en niveau de gris

Les images en niveaux de gris sont souvent utilisées dans les modèles de classification pour plusieurs raisons

- **Réduction de la complexité** Les images en niveaux de gris contiennent une seule composante de couleur (l'intensité lumineuse), contrairement aux images en couleur qui ont trois composantes (rouge, vert, bleu). En utilisant des images en niveaux de gris, on réduit la complexité du modèle en ne considérant qu'une seule dimension de couleur, ce qui peut faciliter l'entraînement et réduire les exigences de puissance de calcul.
- **Élimination des variations de couleur** Les modèles de classification peuvent se concentrer davantage sur les motifs et les caractéristiques structurelles des images lorsqu'ils sont en niveaux de gris. Les variations de couleur peuvent introduire des informations supplémentaires mais aussi du bruit, ce qui peut rendre plus difficile la détection des motifs pertinents.
- **Réduction de la dimensionnalité** En convertissant les images en couleur en niveaux de gris, on réduit la dimensionnalité des données. Cela peut être bénéfique lorsque les ressources de calcul sont limitées ou lorsque les données contiennent des informations redondantes dans les composantes de couleur.

4.6.3 Encodage des données

Afin de pouvoir effectuer des calculs arithmétiques, les données utilisées pour l'apprentissage doivent être converties en format numérique. Étant donné que notre jeu de données contient des colonnes textuelles, nous devons encoder les données. Plusieurs techniques sont disponibles, notamment le Label Encoder, le One Hot Encoder, et bien d'autres. Dans notre projet, nous avons employé ces techniques pour résoudre ce problème.

One Hot Encoder

Le One Hot Encoder est une méthode qui permet de convertir chaque valeur catégorique discrète, telle que des classes d'objets ou des catégories

de texte, en un tableau. La longueur du tableau correspond au nombre total de valeurs possibles dans la colonne. Chaque valeur est associée à un index spécifique, qui est fixé à 1, tandis que toutes les autres valeurs sont fixées à 0. Cette technique est utilisée pour encoder les étiquettes dans un modèle multiclasse, afin de prendre en compte un ensemble de possibilités. La figure 3.5 illustre le fonctionnement du One Hot Encoder.

4.6.4 Normalisation des données

La normalisation des données est le processus qui consiste à transformer les valeurs d'un ensemble de données afin qu'elles se situent dans une plage commune, tout en préservant les relations complexes (patterns) entre les données. C'est une pratique courante de prétraitement des données qui est essentielle pour assurer le bon fonctionnement des algorithmes d'apprentissage automatique. Il existe plusieurs méthodes de normalisation des données, telles que le "Min-Max Scaling" et la "Z-Score Standardization". Dans notre cas, nous avons choisi d'utiliser la méthode de la Z-Score Standardization, car elle gère efficacement les valeurs aberrantes

Z-Score Standardization

Le Z-Score Standardization est une méthode de normalisation des données qui se révèle efficace pour résoudre le problème des valeurs aberrantes. La formule suivante démontre la mise en uvre du processus de normalisation avec le Z-Score Standardization.

4.6.5 Régularisation des données

La régularisation est une méthode utilisée pour améliorer les modèles d'apprentissage automatique en réduisant la fonction de perte ajustée et en évitant les problèmes de surajustement et de sous-ajustement. Parmi les techniques courantes de régularisation, on retrouve la régression Ridge, la régres-

sion Lasso et la régularisation par élagage.

4.6.6 Subdiviser les données pour l'entraînement et le test

Il est nécessaire de diviser l'ensemble de données en deux parties distinctes. Nous utilisons généralement 80 de l'ensemble pour entraîner le modèle, tandis que les 20 restants sont utilisés pour tester le modèle après l'entraînement. Cette division est illustrée dans la figure 4.3.

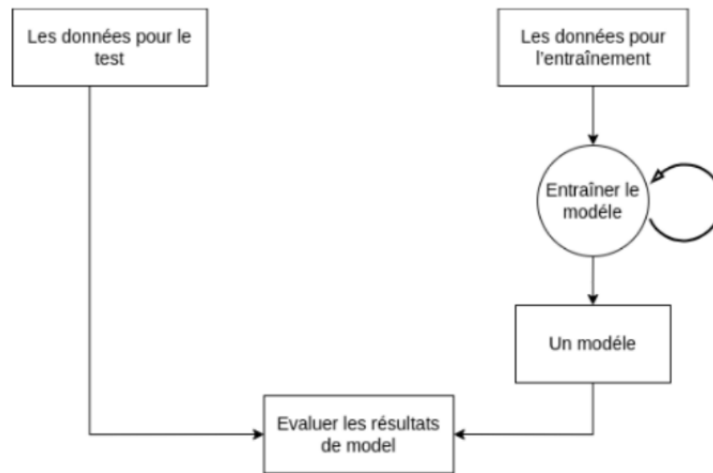


FIGURE 4.3 – Subdiviser les données

4.7 L'entraînement de modèle

La détermination du nombre optimal de couches et de paramètres d'un modèle est un processus itératif qui dépend des données spécifiques, de la complexité du problème et des ressources disponibles. Il n'existe pas de méthode universelle pour déterminer précisément ces valeurs. Une approche courante consiste à entraîner et évaluer plusieurs modèles avec des architectures différentes afin de trouver la configuration optimale. Nous commençons

généralement par des architectures simples comportant un nombre limité de couches et de paramètres, puis nous augmentons progressivement la complexité du modèle pour évaluer son impact sur les performances, Dans notre cas le tableau 4.1 contient les couches et le nombre des parametres

La couche	Le type de couche	le nombre du paramètres
1	Convolutional 2D.	256
2	Max Pooling 2D.	0
3	Flatten.	0
4	Dense.	77888
5	Droupout.	0
6	Dense.	903
Totale		78,209

TABLE 4.1 – l’architecture du modèle.

4.7.1 Le fonctionnement

Chaque nud est lié à une fonction d’activation arithmétique, et son état actif ou non-actif lors de l’exécution dépend des résultats de ses données d’entrée et de la fonction d’activation correspondante. Le schéma ci-dessous illustre les étapes d’entraînement d’un modèle au niveau du nud :

4.7.2 La fonction d’activation

La fonction d’activation est utilisée pour introduire de la non-linéarité dans le modèle et modéliser les relations complexes entre les variables d’entrée et de sortie. Elle est caractérisée par des paramètres initialement assignés à des valeurs aléatoires, appelés poids. Lors de l’entraînement du réseau de neurones, ces poids associés à chaque nud sont ajustés de manière itérative pour minimiser une fonction de perte qui mesure la différence entre les prédictions du modèle et les valeurs cibles. Les algorithmes d’optimisation sont utilisés pour effectuer ces ajustements de poids.

Il existe plusieurs fonctions d'activation couramment utilisées, telles que la fonction sigmoïde, la fonction tangente hyperbolique (\tanh), la fonction Rectified Linear Unit (ReLU) et ses variantes telles que Leaky ReLU et Parametric ReLU (PReLU), ainsi que la fonction softmax utilisée pour la classification multiclasse. Chaque fonction d'activation possède ses propres propriétés et avantages, et le choix de la fonction dépend du problème spécifique et des caractéristiques des données.

Dans notre projet, nous utilisons les fonctions d'activation ReLU et softmax. La fonction ReLU est utilisée dans toutes les couches, à l'exception des couches de sortie. La fonction softmax est utilisée dans la couche de sortie du modèle de classification.

La fonction d'activation ReLU

La fonction d'activation ReLU (Rectified Linear Unit) est l'une des fonctions les plus populaires utilisées en Deep Learning. Elle est définie comme suit [27] :

$$f(x) = \max(0, x)$$

Dans cette équation

- "x" représente la variable d'entrée
- "f(x)" correspond au résultat de la fonction.
- "max" indique que la fonction retourne la valeur maximale entre 0 et "x".

En d'autres termes, lorsque "x" est positif ou nul, la fonction "f(x)" renvoie la valeur de "x". Cependant, si "x" est négatif, la fonction "f(x)" renvoie 0, car le maximum entre 0 et "x" serait 0. Cette fonction est extrêmement efficace pour introduire de la non-linéarité dans les modèles de réseaux de neurones, car elle permet de modéliser des relations complexes entre les variables d'entrée et de sortie [27].

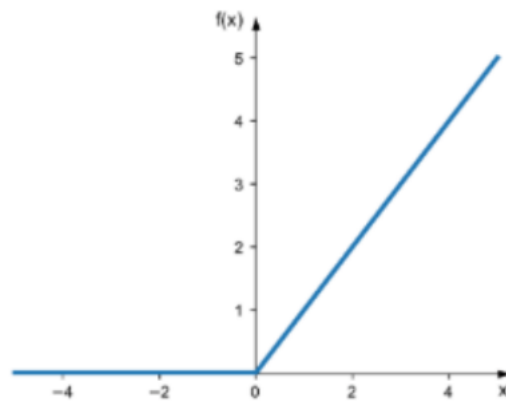


FIGURE 4.4 – La fonction RELU

La fonction d'activation softmax

La fonction d'activation softmax est couramment utilisée dans les réseaux de neurones pour la classification multiclasse. Elle est principalement employée en tant que dernière couche d'un réseau de neurones afin de générer une distribution de probabilités sur les différentes classes de sortie [28].

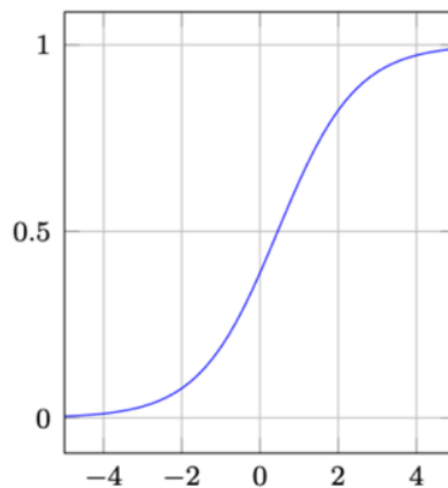


FIGURE 4.5 – La fonction Softmax

4.7.3 La fonction de perte

Dans notre cas, nous utilisons les fonctions de perte "categorical cross entropy". Ces fonctions sont des outils mathématiques qui permettent de mesurer la capacité d'un modèle d'apprentissage automatique à prédire le résultat attendu pour une entrée spécifique.

categorical cross entropy

Il est fréquent d'utiliser la fonction de perte "categorical cross entropy" (ou entropie croisée catégorique) lors de la classification multi-classe. Cette fonction de perte quantifie la disparité entre la distribution de probabilité prédite par le modèle et la distribution de probabilité réelle des classes. Plus précisément, elle calcule l'entropie croisée entre ces deux distributions de probabilité.

4.7.4 L'algorithme d'optimisation

Pour notre projet, nous utilisons l'algorithme Adam pour trouver l'ensemble des paramètres qui minimisent la fonction de perte d'un modèle donné. L'algorithme d'optimisation met en général à jour de manière itérative les paramètres du modèle jusqu'à ce qu'un minimum de la fonction de perte soit atteint.

Optimiseur Adam

L'optimiseur Adam (Adaptive Moment Estimation) est un algorithme d'optimisation alternatif qui fournit des poids de réseau neuronal plus efficaces en effectuant des cycles répétés d'estimation adaptative du moment. Adam étend la descente de gradient stochastique pour résoudre plus rapidement les problèmes non convexes tout en utilisant moins de ressources que de nombreux autres algorithmes d'optimisation. Il est particulièrement efficace pour les ensembles de données extrêmement volumineux en maintenant des

gradients plus étroits sur de nombreuses itérations d'apprentissage. Adam combine les avantages de deux autres techniques de gradient stochastique, les gradients adaptatifs et la propagation quadratique moyenne, pour créer une nouvelle approche d'apprentissage permettant d'optimiser divers réseaux neuronaux. [53] Voici quelques avantages de l'algorithme Adam [54] :

- Facile à mettre en œuvre
- Calcul efficace de l'élan (momentum)
- Faibles exigences en mémoire
- Bien adapté aux problèmes à grande échelle en termes de données et/ou de paramètres
- Convient aux objectifs non stationnaires
- Les hyperparamètres ont des interprétations intuitives et nécessitent généralement un réglage minimal

4.8 Plateforme matérielle et logicielle utilisée

Afin de construire un modèle d'apprentissage profond, il est essentiel d'avoir des ressources appropriées pour soutenir le processus. Ces ressources peuvent être classées en deux catégories : le matériel et le logiciel, qui sont décrits en détail ci-dessous :

4.8.1 Plateforme matérielle

Afin de répondre à ces exigences matérielles, nous avons utilisé un ordinateur doté de performances adaptées à l'apprentissage profond. Cependant, pour bénéficier des ressources nécessaires, nous avons opté pour l'utilisation de Google Colab, qui propose les fonctionnalités mentionnées dans le Tableau 4.2.

Le tableau 4.3 présente les caractéristiques de chaque machines utilisée lors de la réalisation de ce projet.

Le matériel	Les caractéristiques
CPU	Intel Xeon CPU 2.30 GHz
GPU	NVIDIA Tesla T4
RAM	13 GB
Stockage	100 GB

TABLE 4.2 – La Configuration du Colab.

les caractéristiques du machine	le PC01	le PC02
nom du PC	DELL LATITUDE E5450 .	Lenovo G50
RAM	8GB.	4GB.
GPU	Intel(R) Core(TM) i7-5600U 2.60 GHZ.	Intel(R) Core(TM) i3-4700U 2.20 GHZ.
le système d'exploitation	Windows 10 64bits.	Windows 10 64bits.

TABLE 4.3 – Les caractéristiques des machines utilisées

4.8.2 Plateforme logicielle

Dans cette section, nous présentons la plateforme que nous avons utilisée pour effectuer l'implémentation, ainsi que toutes les bibliothèques et applications nécessaires qui ont contribué à la réalisation de ce travail.

Python

Python est le langage de programmation open source le plus largement utilisé par une vaste communauté d'informaticiens et de développeurs. En tant que langage de programmation open source très répandu, Python permet aux développeurs de se concentrer sur leur tâche plutôt que sur les détails techniques. Il offre une liberté aux développeurs en éliminant les contraintes formelles présentes dans les langages plus anciens. Par conséquent, le développement de code en Python est plus rapide et plus facile par rapport à d'autres langages [29].



FIGURE 4.6 – Python logo

Keras

Keras, écrit en langage Python, est une API de réseau neuronal. Il s'agit d'une bibliothèque open source conçue pour être modulaire, rapide et conviviale. Keras est largement utilisé comme API de réseau neuronal pour le développement et la validation de modèles neuronaux. Il permet de créer facilement des couches et de mettre en place des architectures complexes pour les réseaux neuronaux. Il propose plusieurs modules entièrement configurables qui peuvent être combinés pour créer de nouveaux modèles. La modularité de Keras permet d'ajouter facilement de nouvelles fonctionnalités sous forme de modules distincts. Ainsi, Keras offre une grande flexibilité et convient parfaitement à la recherche et à l'innovation [30].



FIGURE 4.7 – Keras logo

Numpy

NumPy est une bibliothèque de calcul numérique en Python qui permet de manipuler efficacement des tableaux multidimensionnels. Elle offre des fonctionnalités avancées pour effectuer des opérations mathématiques et statistiques sur de grandes quantités de données. NumPy est largement utilisé dans des domaines scientifiques tels que l'apprentissage automatique, l'apprentissage profond, l'analyse de données et la simulation numérique. Cette

bibliothèque constitue une base fondamentale pour de nombreuses autres bibliothèques Python axées sur le calcul scientifique [31].



FIGURE 4.8 – Numpy logo

Pandas

Pandas est une bibliothèque open-source spécialement conçue pour la manipulation efficace des données relationnelles ou étiquetées. Elle propose des structures de données flexibles, telles que les DataFrames, qui sont adaptées à l'organisation et à l'analyse des données. Pandas offre des fonctionnalités avancées pour le nettoyage, la transformation et l'exploration de données numériques et de séries temporelles. En exploitant les fonctionnalités de NumPy, Pandas offre des performances élevées tout en offrant une syntaxe intuitive, ce qui en fait un outil essentiel pour l'analyse de données en Python [32].



FIGURE 4.9 – Pandas logo

Google collaboratory

Google Collaboratory, également connu sous le nom de Colab, est un service cloud développé par Google Research. Il s'agit d'un environnement de développement intégré (IDE) qui permet à tous les utilisateurs d'écrire du code source dans leur éditeur et de l'exécuter directement depuis leur navigateur. Colab offre un environnement de calcul gratuit avec un accès

facile à des ressources de traitement telles que les unités de traitement graphique (GPU) et les unités de traitement tensoriel (TPU), ce qui permet une exécution rapide de calculs intensifs. Il offre également des fonctionnalités intégrées pour la visualisation de données, l'utilisation de bibliothèques populaires comme TensorFlow et PyTorch, ainsi que la collaboration en temps réel avec d'autres utilisateurs. Google Colab est largement utilisé dans les domaines de l'apprentissage automatique, de l'analyse de données et de la recherche scientifique [33].



FIGURE 4.10 – Google colab logo

Tensor flow

TensorFlow est un framework open source conçu par des chercheurs de Google. Son objectif principal est de prendre en charge les tâches d'apprentissage automatique, d'apprentissage profond, ainsi que d'autres tâches liées à l'analyse statistique et prédictive. En offrant des fonctionnalités avancées, TensorFlow vise à simplifier le développement et la gestion d'applications analytiques avancées pour différents utilisateurs tels que les phytotechniciens, les statisticiens et les modélisateurs prédictifs. Son architecture flexible et ses outils puissants en font un choix populaire pour les projets de machine learning et d'analyse de données [34].



FIGURE 4.11 – Tensorflow logo

Scikit-Learn

Scikit-learn, également appelé sklearn, est une bibliothèque puissante pour l'apprentissage automatique en Python. Elle propose une gamme d'outils efficaces pour la modélisation statistique en apprentissage automatique, tels que la classification, la régression et le regroupement, le tout à travers une interface Python cohérente. Cette bibliothèque est principalement construite sur NumPy, SciPy et Matplotlib, et elle est largement adoptée par la communauté de l'apprentissage automatique. Scikit-learn se distingue par ses fonctionnalités avancées, sa documentation complète et sa réputation en termes de robustesse et de flexibilité [35].



FIGURE 4.12 – Scikit logo

4.9 L'analyse du résultat

Après avoir terminé l'entraînement du modèle, il devient impératif d'effectuer des tests de qualité sur les résultats obtenus afin de s'assurer que le modèle fonctionne correctement et produit des résultats fiables. Ces tests sont essentiels pour vérifier l'efficacité et la précision du modèle avant de l'utiliser dans des applications concrètes.

4.9.1 Résultat de l'entraînement

Un modèle de classification basé sur le Deep Learning a été mis en place. Ce modèle a été entraîné à plusieurs reprises afin d'atteindre les meilleures performances possibles. Enfin, il est nécessaire d'analyser les résultats de l'entraînement pour garantir la fiabilité et l'efficacité du système.

d'après l'entraînement nous avons obtenue les résultats illustrés dans la figure 4.1.

```
46/46 [=====] - 0s 6ms/step - loss: 0.9520 - accuracy: 0.8793
```

FIGURE 4.13 – La résultat d'évaluation du modèle

4.9.2 Évaluation des résultats de la perte (Loss)

Les résultats présentés dans la figure illustrent l'évolution de la perte (Loss) en fonction du nombre d'itérations lors des phases de validation et d'entraînement.

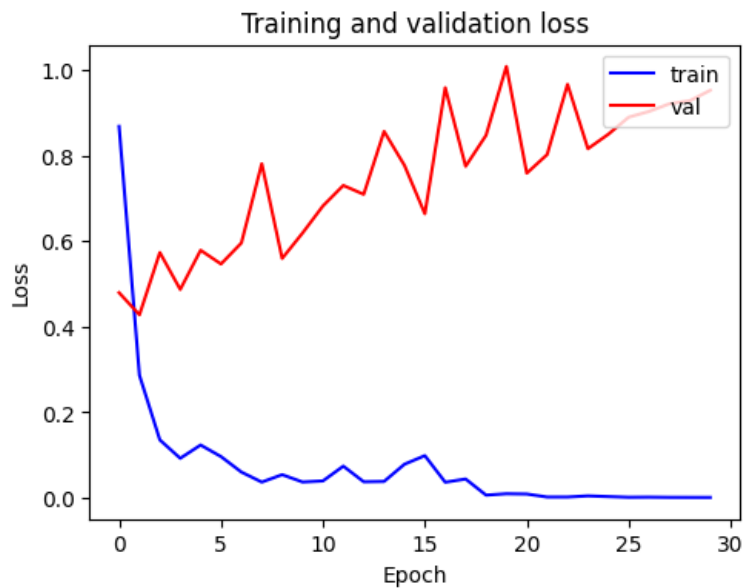


FIGURE 4.14 – La résultat d'évaluation du modèle

D'après les illustrations de la figure 4.14, il est évident que les pertes d'entraînement et de validation maintiennent des valeurs optimales, ce qui indique que le modèle apprend correctement. Les courbes présentent cependant une certaine irrégularité, ce phénomène étant attribué à l'utilisation du

mécanisme de Drop-Out pour prévenir le surapprentissage. On notons que la valeur calculée de la perte est **0.9520**.

4.9.3 Évaluation des résultats de la précision (Accuracy)

Le résultat de précision de modèle, testé sur une partie de 20% des ensembles de données réservée à cet effet, est : **87.93**

Le résultat de précision (**Accuracy**) en fonction du nombre d'itérations lors de la validation et de l'entraînement est illustré dans la figure 4.15.

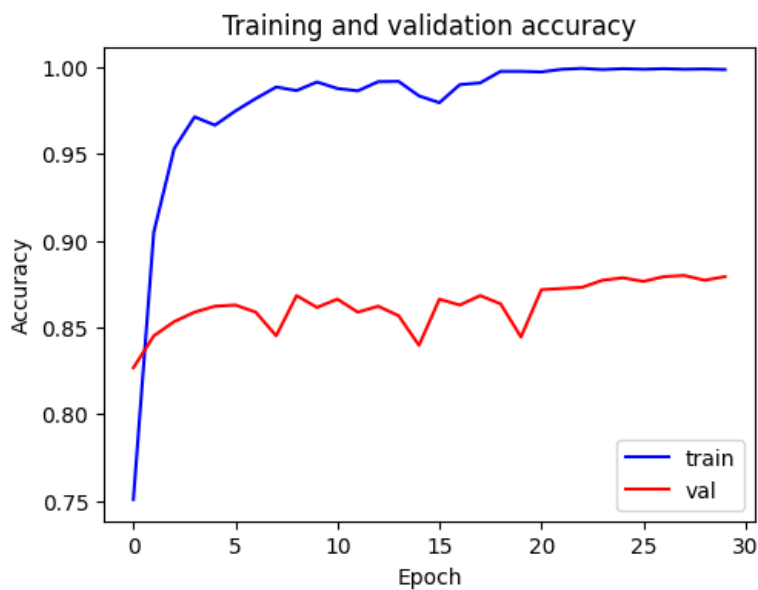


FIGURE 4.15 – La résultat d'évaluation du modèle

4.9.4 Evaluation du résultat de la matrices de confusion

La matrice de confusion illustrée dans la figure 4.16, dans chaque cellule représente une combinaison entre une prédiction de classe et une vraie classe. Les valeurs affichées dans les cellules indiquent que le nombre d'échantillons appartenant à cette combinaison.

Les cellules situées le long de la diagonale principale (qui commence en haut à gauche) représentent les combinaisons où la prédiction et la classe réelle porte les mêmes valeurs. Par conséquent, les cellules de la diagonale principale contiennent le nombre de prédictions correctes.

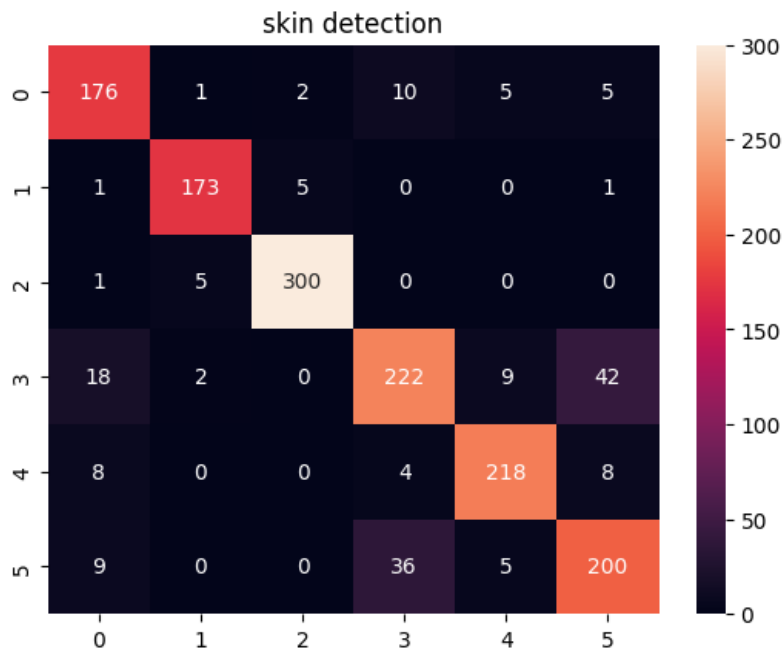


FIGURE 4.16 – La résultat d'évaluation du modèle

4.10 Conclusion

En conclusion, nous avons discuté de la sélection des données, du modèle de classification utilisé et du prétraitement des données, notamment le redimensionnement des images, la conversion en niveaux de gris, l'encodage, la normalisation et la régularisation. Nous avons également abordé la subdivision des données pour l'entraînement et les tests, l'entraînement du modèle, son fonctionnement, ainsi que les fonctions d'activation, de perte et l'algorithme d'optimisation. De plus, nous avons examiné la plateforme matérielle et logicielle utilisée pour le développement du système. Enfin, nous avons analysé les résultats obtenus, notamment les résultats d'entraînement, l'évaluation de la perte, la précision et la matrice de confusion

Conclusion générale et perspectives

En conclusion, ce mémoire a exploré le domaine des systèmes de prédiction des maladies de la peau en utilisant les réseaux de neurones convolutifs (CNN). Notre objectif principal était de développer un système précis, fiable et efficace pour aider à la prédiction des affections cutanées.

Tout au long de cette recherche, nous avons examiné l'état de l'art des méthodes de prédiction des maladies de la peau, en mettant l'accent sur l'utilisation des réseaux de neurones convolutionnels. Nous avons réalisé une analyse approfondie des différentes étapes du processus, de la sélection et du prétraitement des données à l'entraînement et à l'évaluation des modèles.

Les résultats obtenus ont démontré la valeur de notre approche. Le système développé a montré des performances prometteuses dans la prédiction des maladies de la peau, avec une précision élevée et une capacité à reconnaître un large éventail de conditions dermatologiques. Ces résultats offrent des perspectives encourageantes pour l'amélioration des soins de santé dans le domaine dermatologique.

Ce mémoire a également souligné l'importance de la disponibilité de ressources matérielles et logicielles appropriées pour soutenir les calculs intensifs nécessaires à l'entraînement et à l'inférence des modèles. Une attention particulière a été accordée à l'optimisation des performances du système en termes de temps d'exécution et de capacité de traitement des données.

Il convient de noter que ce système de prédiction des maladies de la

peau est un outil complémentaire à l'expertise médicale. Il ne remplace pas le diagnostic clinique effectué par des professionnels de la santé qualifiés, mais peut servir de guide pour soutenir leur prise de décision et améliorer l'efficacité des soins.

La présente étude a ouvert de nouvelles perspectives passionnantes pour la recherche et le développement dans le domaine de la prédiction des maladies de la peau. Cependant, il reste encore de nombreux aspects à explorer et des opportunités d'amélioration à considérer. Dans cette section, nous discuterons des perspectives futures de ce domaine.

1. **Intégration de données multimodales** : L'intégration de sources de données multimodales, telles que l'historique du patient, les notes cliniques ou les informations génétiques, avec les images permettrait une compréhension plus complète des maladies de la peau. Les recherches devraient se concentrer sur des techniques efficaces de fusion et d'apprentissage de représentation pour ces sources de données hétérogènes.
2. **Déploiement et validation dans le monde réel** : Il est essentiel de valider les performances du système dans des environnements cliniques réels. Les recherches futures devraient mener des études de validation approfondies impliquant des dermatologues et des professionnels de santé pour évaluer l'efficacité, la fiabilité et la scalabilité du système. La collaboration avec des institutions de santé et son intégration dans l'infrastructure existante faciliteront son déploiement pratique.

En conclusion, ce mémoire a contribué à l'avancement des systèmes de prédiction des maladies de la peau en fournissant une méthodologie solide, des résultats probants et une base solide pour de futures recherches et développements. L'application pratique de ces avancées technologiques pourrait avoir un impact significatif sur le domaine de la dermatologie, en permettant une détection précoce, des diagnostics précis et une amélioration globale des soins cutanés.

Bibliographie

- [1] *What is an image ?* <https://www.techtarget.com/whatis/definition/image>.
- [2] *Neuron-SEM-2.png (800E600)*, <https://upload.wikimedia.org/wikipedia/commons/c/c2/Neuron-SEM-2.png>, (Accessed on 06/27/2023).
- [3] *Images*, <https://www.f-legrand.fr/scidoc/docmml/image/niveaux/images/images.html>, (Accessed on 06/27/2023).
- [4] *Image numérique - Représentation des couleurs*, <https://www.techno-science.net/glossaire-definition/Image-numerique-page-2.html>.
- [5] *Photo - caractéristique des images*, <http://hazmat.free.fr/7-photo/photo-images.html>.
- [6] *sous-pixel Wiktionnaire*, <https://fr.wiktionary.org/wiki/sous-pixel>, (Accessed on 06/27/2023).
- [7] *Quest-ce que la "résolution" d'une image ? | Canson Infinity*, <https://www.canson-infinity.com/fr/faq/qu-est-ce-que-la-resolution-d-une-image#:~:text=La%20r%C3%A9solution%20d%27une%20image%20est%20le%20nombre%20de,il%20y%20aura%20d%27information%20dans%20l%27image%20%28plus%20pr%C3%A9cise%29.>
- [8] *Profondeur d'une image*, <http://pixelavenue.com/blog/91-lexique-profondeur-image>.
- [9] *Definition, Picture size, DPI and PPI | FMEdda |*, https://fmedda.com/en/article/dpi_ppi.
- [10] *Defining Image Quality Thomas Fitzgerald Photography*, <https://blog.thomasfitzgeraldphotography.com/blog/2019/1/defining-image-quality>.
- [11] V. SHANKAR, V. KUMAR, U. DEVAGADE, V. KARANTH et K. ROHITAKSHA, "Heart disease prediction using CNN algorithm," *SN Computer Science*, t. 1, n° 3, p. 170, 2020.

- [12] F. RUSTAM, A. ISHAQ, K. MUNIR, M. ALMUTAIRI, N. ASLAM et I. ASHRAF, "Incorporating CNN Features for Optimizing Performance of Ensemble Classifier for Cardiovascular Disease Prediction," *Diagnostics*, t. 12, n° 6, p. 1474, 2022.
- [13] A. S. GOWDA, "A Voting Classifier based Hybrid Model for Heart Disease Prediction," *Journal of Engineering Sciences*, t. 14, n° 04, 2023.
- [14] A. SHUKLA, R. TIWARI et S. TIWARI, "Review on alzheimer disease detection methods : Automatic pipelines and machine learning techniques," *Sci*, t. 5, n° 1, p. 13, 2023.
- [15] P. F. O. RAYER, *Traité théorique et pratique des maladies de la peau : avec un atlas...* Chez JB Baillière, 1835, t. 3.
- [16] *Dermatologie : tout savoir - Santé sur le Net*, <https://www.sante-sur-le-net.com/maladies/dermatologie/>, (Accessed on 06/27/2023).
- [17] *Prediction of skin disease using a new cytological taxonomy based on cytology and pathology with deep residual learning method / Scientific Reports*, <https://www.nature.com/articles/s41598-021-92848-y>.
- [18] M. CREPY, "Eczéma professionnel, actualités 2012," *Revue Française d'Allergologie*, t. 52, n° 3, p. 212-217, 2012.
- [19] *Le diagnostic des infections fongiques invasives*, <https://www.revmed.ch/revue-medicale-suisse/2005/revue-medicale-suisse-13/le-diagnostic-des-infections-fongiques-invasives>.
- [20] Y. WANG, L. ZHANG, M. YANG et al., "Development of a predictive model for screening patients with psoriasis at increased risk of psoriatic arthritis," *Dermatology and Therapy*, p. 1-15, 2022.
- [21] *Dermatologist-level classification of skin cancer with deep neural networks / Nature*, <https://www.nature.com/articles/nature21056>.
- [22] *Automated Dermatological Diagnosis : Hype or Reality ? - PubMed*, <https://pubmed.ncbi.nlm.nih.gov/29864435/>.
- [23] *Automatic eczema classification in clinical images based on hybrid deep neural network - PubMed*, <https://pubmed.ncbi.nlm.nih.gov/35809409/>.
- [24] *Découvrez les différentes couches d'un CNN - Classez et segmentez des données visuelles - OpenClassrooms*, <https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5083336-decouvrez-les-differentes-couches-dun-cnn>.

- [25] *Deep learning : les réseaux de neurones convolutifs pour la classification d'images médicales*, <https://www.imaios.com/fr/ressources/blog/classification-des-images-medicales-comprendre-le-reseau-de-neurones-convolutifs-cnn>.
- [26] *DataScienceToday - Les réseaux de neurones convolutifs*, <https://www.datasciencetoday.net/index.php/fr/deep-learning/173-les-reseaux-de-neurones-convolutifs>.
- [27] *Que signifie ReLU en matière d'intelligence artificielle ? - ActuIA*, <https://www.actuia.com/faq/que-signifie-relu-en-matiere-dintelligence-artificielle/#:~:text=ReLU%20est%20une%20fonction%20d,les%20r%C3%A9sultats%20n%C3%A9gatifs%20par%20z%C3%A9ro>.
- [28] *A quoi sert la fonction Softmax ? - ActuIA*, <https://www.actuia.com/faq/a-quoi-sert-softmax/#:~:text=La%20fonction%20d'activation%20softmax,couche%20dans%20un%20intervalle%20donn%C3%A9>.
- [29] *Introduction to Python*, https://www.w3schools.com/python/python_intro.asp.
- [30] *Keras : tout savoir sur l'API de Deep Learning*, <https://datascientest.com/keras>.
- [31] *What is NumPy ? NumPy v1.25 Manual*, <https://numpy.org/doc/stable/user/whatisnumpy.html>.
- [32] *Introduction to Pandas in Python - GeeksforGeeks*, <https://www.geeksforgeeks.org/introduction-to-pandas-in-python/>.
- [33] *Google Colab ou Google Colaboratory : qu'est-ce que c'est / Matériel gratuit*, <https://www.hwlibre.com/fr/colaboratoire-google/>.
- [34] *What is TensorFlow ? | Definition from TechTarget*, <https://www.techtarget.com/searchdatamanagement/definition/TensorFlow>.
- [35] *Scikit-Learn : guide de démarrage rapide en Machine Learning avec Python*, <https://www.data-transitionnumerique.com/scikit-learn-python/>.