

République Algérienne Démocratique et Populaire

الجمهورية الجزائرية الديمقراطية الشعبية

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

وزارة التعليم العالي و البحث العلمي



المدرسة العليا الآلية للإعلام سيدي بلعباس

École nationale Supérieure d'Informatique

Mémoire de fin d'études

Pour l'obtention du diplôme master 2 en Informatique

Option : Systèmes Informatiques et Web

Comparaison entre les approches d'apprentissage profond pour la classification de cancer de la peau

Réalisé par :

Mme. Guitt Bahia

Mme. Yagoub Fatima Zohra

Encadré par :

M. Chaib Soulaymane (ESI)

M. Benammar Abdessalem(CRTI)

Promotion : 2020/2021

Remerciements

Tout d'abord, je remercie Allah le tout puissant de m'avoir donné le courage et la patience nécessaires à mener ce travail à son terme.

Je tiens à remercier tout particulièrement mon encadreur **M. CHAIB Sulaimen**, pour l'aide compétente qu'il m'a apporté, pour sa patience et son encouragement. Son œil critique m'a été très précieux pour structurer le travail et pour améliorer la qualité des différentes sections.

Je tiens à remercier également mon promoteur **M. BENAMMAR Abdessalam** pour son aide immense, la qualité de son suivi ainsi que pour tous les conseils et les informations qu'il m'a prodigués avec un degré de patience et de professionnalisme sans égal.

Un très grand remerciement et une très grande reconnaissance sont destinés à **Mme. Ziouech** de m'avoir aidé à obtenir ce stage de fin d'études chez CRTI.

Je désire remercier également **M. BOUZAR Abderrahmane** pour les renseignements précieux qu'il m'a fournis ainsi que pour son encouragement.

Que les membres de jury trouvent, ici, l'expression de mes sincères remerciements pour l'honneur qu'ils me font en prenant le temps de lire et d'évaluer ce travail.

Je souhaite aussi remercier l'équipe pédagogique et administrative de l'ESI SBA pour leurs efforts dans le but de nous offrir une excellente formation.

Je tiens à remercier **M. BELFEDHAL paddinour seaAla** disponibilité et ses orientations.

Pour finir, je souhaite remercier toute personne ayant contribué de près ou de loin à la réalisation de ce travail.

Dédicace

“

Je dédie ce modeste travail À mes chers parents, pour tous leurs sacrifices, leur amour, leur encouragement et leurs prières tout au long de mes études,

*À ma chère sœur **Asma** source de joie et de bonheur,*

*À mes chers frères, **Badreddine** et **Alaaddine**,*

*À ma deuxième famille, mes chères copines **Ikhlassé**,
Ikram, **Houda**, **Ilhem**, **Amina** et **Rachida**,*

*À tous mes amis qui m'ont toujours encouragé, tout particulièrement **Oussama** pour sa contribution de près ou de loin tout au long de ce parcours,*

*Son oublier mon binôme **Bahou** pour sa collaboration, sa patience et sa compréhension dans la réalisation de ce projet,*

Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infaillible, Merci d'être toujours là pour moi.

”

Dédicace

“

Avec l'expression de ma reconnaissance, je dédie ce modeste travail à ceux qui, quels que soient les termes embrassés, je n'arriverais jamais à leur exprimer mon amour sincère.

*À l'homme, mon précieux offre du dieu, qui doit ma vie, ma réussite et tout mon respect :mon père **Ahmed**.*

*À la femme qui a souffert sans me laisser souffrir, qui n'a jamais dit non à mes exigences et qui n'a épargné aucun effort pour me rendre heureuse : mon adorable mère **Malika**.*

*À mes chères soeurs **Hilda,Zahida,Imane,Yousra** et mon frère **Salah Eddin** qui n'ont pas cessée de me conseiller,encourager et soutenir tout au long de mes études, que Dieu les protège et leurs offre la chance et le bonheur.*

*À mes chères amis **Ikram,Houda,Ikhlas,Ilham,Amina,Islam** avec qui j'ai partagé que les moments inoubliables.*

*À mon cher binome **Fatima Zohra** pour son soutien moral,sa patience et sa compréhension tout au long de ce sujet*

”

Résumé

L'analyse automatique d'images médicales s'impose aujourd'hui comme un outil incontournable pour le diagnostic voire même d'aide à la décision dans le domaine médical.

L'objectif de ce sujet de Master est d'effectuer une synthèse des travaux de l'état de l'art portant sur les approches d'apprentissage profond ayant été proposées pour l'analyse des images médicales du cancer de la peau. De plus, Nous avons proposé un modèle pour améliorer les performances de classification des lésions cutanées en utilisant les réseaux de neurones convolutifs.

Mots clés : Cancer de la peau , la classification du cancer de la peau ,L'apprentissage profond ,Les images médicales,Cnn .

ملخص

يعد التحليل التلقائي للصور الطبية في الوقت الحاضر أداة أساسية للتشخيص وحتى دعم القرار في المجال الطبي .

الهدف من موضوع هذا البحث هو تجميع أحدث الأعمال على مناهج التعلم العميق التي تم اقتراها لتحليل الصور الطبية لسرطان الجلد. بالإضافة إلى ذلك ، اقترحنا نموذجاً لتحسين أداء تصنيف الآفات الجلدية باستخدام الشبكات العصبية التلافيية .

كلمات مفتاحية : سرطان الجلد ، تصنیف سرطان الجلد ، التعلم العميق ، مناهج التعلم العميق.

Abstract

The automatic analysis of medical images is nowadays an essential tool for diagnosis and even decision support in the medical field.

The objective of this Master thesis is to carry out a synthesis of the state of the art work on deep learning approaches that have been proposed for the analysis of medical images of skin cancer. Moreover, we proposed a model to improve the classification performance of skin lesions using convolutional neural networks.

Keywords : Skin cancer ,Skin cancer classification, Deep learning,Deep learning approaches.

Table des matières

Remerciements	I
Dédicace	II
Dédicace	III
Résumé	IV
V	ملخص
Abstract	VI
Introduction générale	1
1 Analyse et reconnaissance des images médicales	3
1.1 Introduction	4
1.2 Contexte médical	4
1.2.1 L'imagerie médicale	4
1.2.2 Les types d'imagerie médicale	4
1.2.3 Les différentes technologies d'imagerie médicale	5
1.2.4 Les Caractéristiques d'une image médicale	7
1.2.5 Anatomie de la peau	10
1.3 Cancer de la peau	11
1.3.1 Les carcinomes	12
1.3.2 Les melanomes	13
1.4 Classification des images médicales	14
1.4.1 L'apprentissage profond	14
1.4.2 Les réseaux de neurones convolutifs	16
1.4.3 Les techniques d'apprentissage profond	19
1.5 Conclusion	21
2 État de l'art	22
2.1 Introduction	23
2.2 Réseaux de neurones entraînés de bout en bout	24
2.2.1 Réseaux de neurones entraînés en partant de zéro	24
2.2.2 Réseaux de neurones avec apprentissage par transfert	25
2.3 Extraction des caractéristiques par CNN	28
2.4 Conclusion	34

Table des matières

3 Approche adaptée pour la classification d'images médicales	35
3.1 Introduction	36
3.2 Description de l'architecture adaptée	36
3.3 Description détaillée de l'approche proposée	38
3.4 Évaluations expérimentales	38
3.4.1 Métriques de performance	38
3.4.2 Ensemble des données ISIC	39
3.4.3 Résultats et discussion	42
3.4.4 Synthèse	44
3.5 Conclusion	45
Conclusion et perspectives	47

Table des figures

1.1	Illustration sur l'imagerie morphologique [10].	5
1.2	Illustration sur l'imagerie fonctionnelle [10].	5
1.3	Illustration sur la radiographie [24].	6
1.4	Illustration sur la scannographie [43].	6
1.5	Illustration sur l'échographie [11].	7
1.6	Illustration sur l'IRM [22].	7
1.7	Représentation matricielle d'une image numérique [7].	8
1.8	Image numérique avec différentes résolutions [19]	8
1.9	Problème d'élimination de bruit : à gauche image bruitée , à droite image restaurée [9].	9
1.10	Décomposition d'une image couleur en ses trois composantes rouge, verte et bleue [51].	9
1.11	Une image en niveaux de gris, agrandissement d'une zone de l'image, et affichage des valeurs constituant la matrice image. [52].	10
1.12	Anatomie de la peau [47].	10
1.13	Les différents types du cancer de la peau [53].	11
1.14	Illustration du cancer basocellulaire. [54].	12
1.15	Des photos réels du cancer carcinome basocellulaire [55].	13
1.16	Des photos réels du Carcinome épidermoïde [56].	13
1.17	Des photos réels du cancer mélanome [57].	14
1.18	Structure d'un neurone artificiel ou neurone formel [38]	15
1.19	Architecture de Perceptron multicouche [49]	15
1.20	Architecture d'un réseau de neurone convolutif [35]	17
1.21	Schéma d'une opération de pooling avec un noyau MaxPool de taille 2*2 et d'un pas de 2 [13].	18
1.22	Taxonomie proposée pour la classification du cancer de la peau.	19
1.23	représentation graphique de la méthode d'extraction des caractéristiques .	21
2.1	Taxonomie repsrésentant les travaux cités dans notre état de l'art.	24
2.2	Architecture proposée par Nasr et al [31]	25
2.3	Architecture VGG16 [60].	26
2.4	Architecture Inception [27].	27
2.5	Architecture MobileNetV2 [42].	27
2.6	Architecture proposée par Rehan et al [5].	28
2.7	Architecture proposée par Pomponiu et al [36].	29
2.8	Architecture proposée par Pham et al [34].	30
2.9	Organigramme représente la méthodologie proposée par Maia et al [29].	31
2.10	Architecture proposée par Rezaoana et al [39].	32

Table des figures

3.1	Représentation de l'architecture MobileNet [37].	36
3.2	Illustration représente à gauche une convolution séparable en profondeur , et à droite une convolution standard. [8].	37
3.3	Illustration représente les couches d'un bloc de convolution séparable en profondeure. [37].	37
3.4	Illustration des difficultés des images d'ensemble de données ISIC	40

Liste des tableaux

1.1	Les fonctions d'activation	16
2.1	Tableau représentant les résultats obtenus par la combinaison des différentes architectures avec les différents classificateurs sur le dataset PH2 [29]	31
2.2	Tableau des travaux concernant la classification du cancer de la peau . . .	33
3.1	Tableau représente les ensembles de données d'images dermatoscopiques accessibles au public [50].	41
3.2	Résultats obtenus par la première approche.	42
3.3	La matrice de confusion de la première approche	42
3.4	Résultats obtenus par la deuxième approche.	43
3.5	La matrice de confusion de la deuxième approche	43
3.6	Résultats obtenus par la troisième approche.	44
3.7	Tableau représente les performances obtenues pour chaque base de données	44
3.8	Tableau représente les performances obtenues pour chaque architecture . . .	44

Liste des sigles et acronymes

AC	<i>Accuracy</i>
ANN	<i>Artifiel Nneural Nnetwork</i>
CNN	<i>Convolution Neural Nnetwork</i>
DL	<i>Deep Learning</i>
GAP	<i>Global Average Pooling</i>
IRM	<i>Imagerie par Résonance Magnétique</i>
ISIC	<i>International Skin Image Classification</i>
KNN	<i>K-Nearst Neibor</i>
ReLU	<i>Rectified Linear Unit</i>
RGB	<i>Red Green Blue</i>
RE	<i>Recall</i>
RF	<i>Random Forest</i>
RL	<i>Regression Logistic</i>
RSB	<i>Rapport Signal Bruit</i>
SE	<i>SEnsitivity</i>
SVM	<i>Support Vector Machine</i>
SCC	<i>Squamous Cell Carcinoma</i>
UV	<i>Ultrat Violet</i>

Introduction générale

Le domaine médical est un domaine très sensible qu'il nécessite une grande attention. Son efficacité est sans appel : elle n'est plus à démontrer. Les résultats d'un bon diagnostic, afin d'éviter des catastrophes qui, probablement, mènent jusqu'à la mort, est capital. Il est d'un grand intérêt. Pour ce faire, il y a lieu de faire des recherches profondes pour résoudre ce problème. Le traitement automatique des images médicales est donc arrivé à point nommé pour répondre à cette question cruciale.

Le cancer de la peau est l'une des maladies qui fait appel à des compétences très élevées pour la détection de la maladie. Les dermatologues ne sont pas en mesure de diagnostiquer le cancer de la peau rapidement, en raison du faible contraste et de l'obscurcissement entre les zones touchées sur la peau et les régions cutanées normales. Plusieurs artefacts, y compris les poils, les veines, les marques de règle et la couleur l'étalonnage, peuvent brouiller et occire les lésions cancéreuses de la peau. Ces raisons pourraient réduire les performances de la justesse du diagnostic du cancer de la peau.

A l'instar du développement auquel nous assistons, notamment dans le domaine de l'intelligence artificielle et précisément dans le domaine de l'apprentissage profond et sans impact dans tous les domaines. L'apprentissage profond a prouvé son existence sur le terrain de la médecine et surtout dans l'imagerie médicale à cause des raisons suivantes : premièrement, il offre une précision de reconnaissance du cancer de la peau, très élevée par rapport à l'œil nu, de l'être humain avec une grande économie d'argent et de temps. En second lieu, le « deep learning » est capable d'analyser des données stockées sous un format non structuré comme les photos et les vidéos. En troisième phase, Il est capable d'extraire autant de caractéristiques que possible de la cible, y compris des caractéristiques abstraites difficiles ou impossibles à décrire.

Dans le cadre de ce Master, nous sommes fixées pour objectif principal l'exploration de la problématique de classification d'images médicales, en particulier celles du cancer de la peau, à base d'apprentissage profond.

La structure de notre thèse est formée comme suit :

Dans le premier chapitre “**L'apprentissage profond dans le domaine médicale**”. Nous définissons ce qu'est une image médicale et nous expliquons l'anatomie de la peau. Nous détaillons ensuite les types de cancer de la peau en précisant leurs classes et mettons l'accent sur le cancer le plus mortel, à savoir le mélanome. De plus, nous donnons des statistiques sur le taux d'incidence de cette maladie.

Dans le deuxième chapitre “**État de l'art**”, nous parcourons les principaux travaux

Introduction générale

de la littérature qui traitent de la classification d'images médicales et nous en présentons une synthèse. En particulier, nous proposons une taxonomie des travaux d'apprentissage profond qui traitent de la classification du cancer de la peau et nous effectuons une discussion autour de ces travaux.

Dans le troisième chapitre “ **Approche proposée pour la classification d'images médicales**”, nous décrivons de manière détaillée notre proposition relative à la classification du cancer de la peau. Nous commençons par présenter le principe global de cette proposition. Nous détaillons l'architecture de classification de l'état de l'art retenue pour nos travaux . Après, nous présentons nos approches suivies pour l'entraînement de nos modèles. Puis nous présentons les résultats de nos approches. Pour cela, nous donnons d'abord les métriques d'évaluation utilisées pour la classification des images. Nous décrivons ensuite l'ensemble de données choisi pour notre évaluation en mettant en avant ses caractéristiques et son avantage dans notre étude par rapport à d'autres ensembles des données. Nous présentons ensuite nos résultats pour chaque approche tout en proposant une discussion de ces résultats. Enfin, nous terminons par une synthèse qui résume les enseignements retenus de notre travail.

Chapitre 1

Analyse et reconnaissance des images médicales

1.1 Introduction

Dans ce chapitre, nous introduisons quelques concepts pour l'imagerie médicale en général et pour le cancer de la peau en particulier. Nous commençons par définir l'imagerie médicale en citant le rôle et les types de ces images. Ensuite, nous mettons l'accent sur l'anatomie de la peau et expliquons les différentes parties de cet organe. Après, nous expliquons le terme de tumeur en donnant des informations détaillées sur ce dernier. Cette explication est suivie par des définitions des cancers de la peau les plus connus, à savoir le carcinome basocellulaire, le Carcinome épidermoïde et le mélanome. A la fin, nous montrons l'utilité de l'apprentissage profond dans la classification des images médicales.

1.2 Contexte médical

Dans cette section, nous présentons quelques généralités sur le contexte médical. Nous commençons par citer les différents types des images médicales ainsi leurs caractéristiques. Ensuite, nous décrivant les couches essentielles de la peau.

1.2.1 L'imagerie médicale

L'imagerie médicale est un moyen utilisé comme un outil pour affiner un diagnostic en vue pratique pour traiter les patients avec plus d'efficacité ou pour suivre l'évolution du traitement à travers la création d'une représentation visuelle lisible d'une information à caractère médical.

L'imagerie médical a connu une évolution spectaculaire, elle permet actuellement d'examiner non seulement les os mais aussi les articulations, les viscères, le cœur en mouvement ou encore l'activité cérébrale.

1.2.2 Les types d'imagerie médicale

En imagerie médicale, on distingue deux types d'imagerie :

- Imagerie morphologique : qui fournit des informations relatives à l'anatomie des organes(voir figure 1.1)

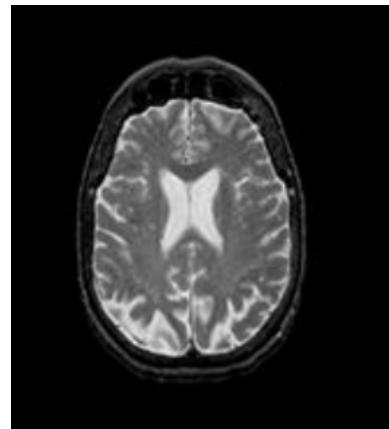


Figure 1.1 : Illustration sur l'imagerie morphologique [10].

- L'imagerie fonctionnelle : qui permet d'étudier des processus métaboliques, biochimiques et physiologiques (voir figure 1.2).

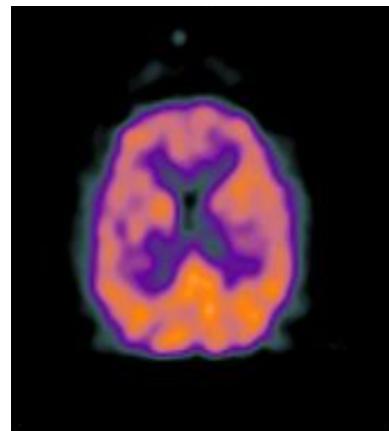


Figure 1.2 : Illustration sur l'imagerie fonctionnelle [10].

1.2.3 Les différentes technologies d'imagerie médicale

- **La radiographie** : Repose sur l'utilisation des rayons X qui ont la propriété de traverser les tissus de manière plus ou moins importante selon leur densité (voir figure 1.3).



Figure 1.3 : Illustration sur la radiographie [24].

- **La scannographie :** Permet d'obtenir des images 3D grâce à une rotation simultanée de la source émettrice de rayons X et du détecteur autour du corps (voir figure 1.4).

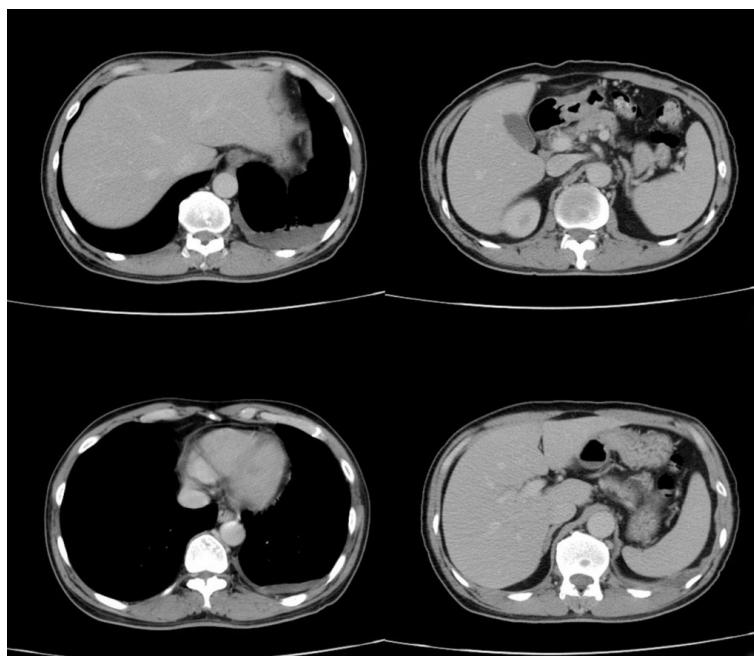


Figure 1.4 : Illustration sur la scannographie [43].

- L'échographie : Est une technique d'exploration de l'intérieur du corps, basée sur les ultra-sons. L'image obtenue est une coupe de l'organe à étudiée(voir figure 1.5).



Figure 1.5 : Illustration sur l'échographie [11].

- L'imagerie par résonance magnétique (IRM) : Permet de visualiser des détails invisibles sur les radiographies standards, l'échographie ou le scanner. Schématiquement, cette technique utilise un gros aimant et analyse la réaction des différents tissus du corps à ces champs magnétiques (voir figure 1.6).[21]

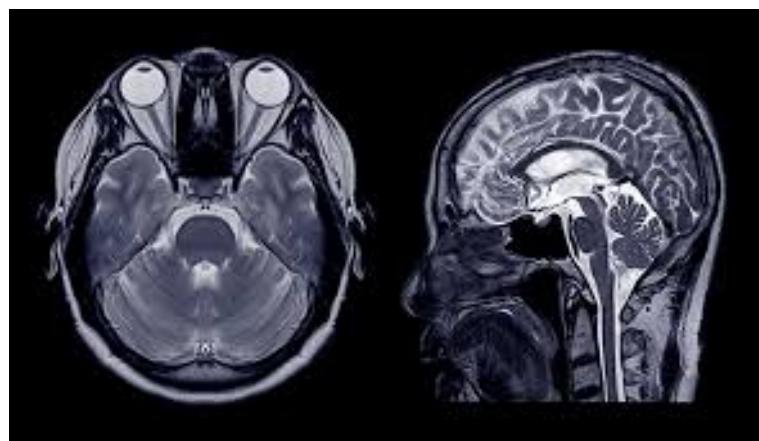


Figure 1.6 : Illustration sur l'IRM [22].

- L'endoscopie : Permet de visualiser l'intérieur des organes, un conduit naturel ou une cavité. En faisant pénétrer un tube ayant à sa tête une caméra pour que le spécialiste visionne.

1.2.4 Les Caractéristiques d'une image médicale

L'image est un ensemble structuré d'informations caractérisées par les paramètres suivants :

A) Dimension :

C'est la taille de l'image. Cette dernière se présente sous forme de matrice dont

chaque case représente un pixel qui contient une valeur numérique bien précise (voir figure 1.7). Le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre total de pixels dans une image.

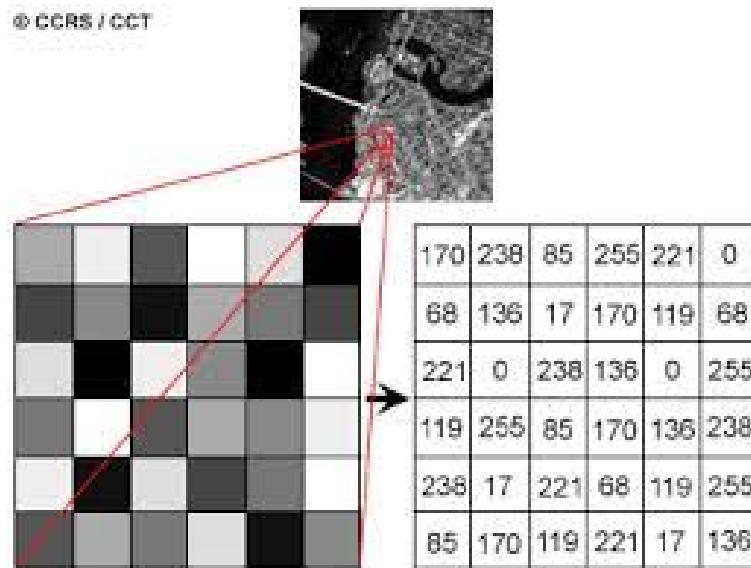


Figure 1.7 : Représentation matricielle d'une image numérique [7].

B) Résolution :

Définit la netteté et la qualité d'une image. Cette dernière se présente par le nombre de pixels contenus dans l'image par unité de longueur. Plus la résolution est grande, plus l'image est précise dans les détails (voir figure 1.8).



Figure 1.8 : Image numérique avec différentes résolutions [19]

C) Bruit :

Pour mesurer le niveau du bruit, on calcule le rapport signal-sur-bruit (RSB) :

$$RSB = \frac{S}{\sigma} \quad (1.1)$$

où S est l'intensité du signal et σ est le bruit.

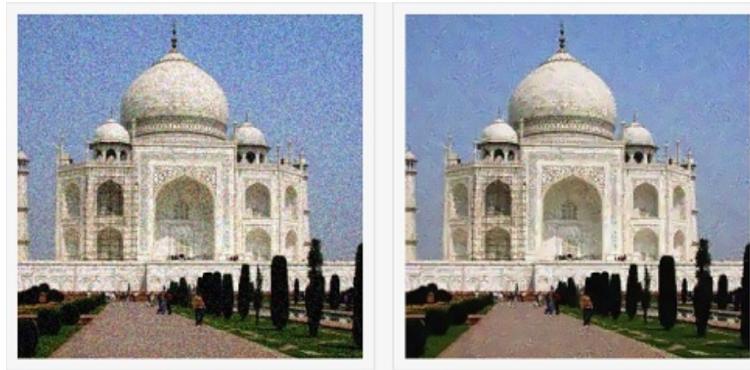


Figure 1.9 : Problème d'élimination de bruit : à gauche image bruitée , à droite image restaurée [9].

D) Luminance

:

Moyenne des différents canaux d'une image, qui indique la puissance lumineuse du pixel.

E) Contraste

Quantité informelle qui indique la différence entre les zones claires et les zones sombres d'une image. Le contraste est défini par l'équation suivante :

$$C = \frac{S_1 - S_2}{S_1 + S_2} \quad (1.2)$$

où C est le contraste. S_1 est le signal de la lésion. S_2 est le signal de fond.

F) Image en couleurs :

Une image couleur est en réalité composée de trois images, afin de représenter le rouge, le vert, et le bleu (voir la figure 1.10). Chacune de ces trois images s'appelle un canal. Cette représentation en rouge, vert et bleu mime le fonctionnement du système visuel humain.

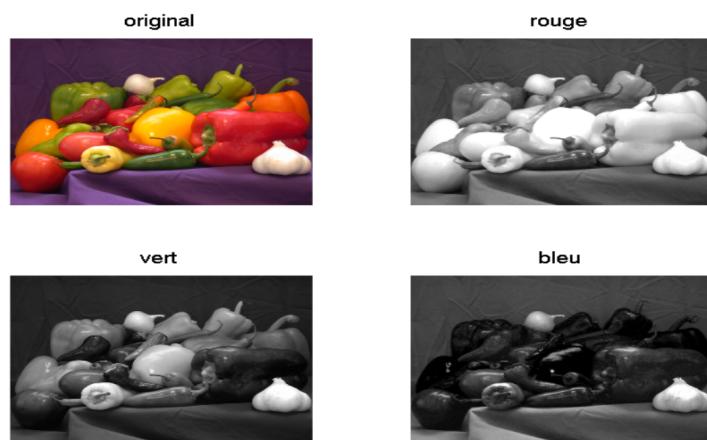


Figure 1.10 : Décomposition d'une image couleur en ses trois composantes rouge, verte et bleue [51].

G) Image à niveaux gris

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires (voir la figure 1.11).

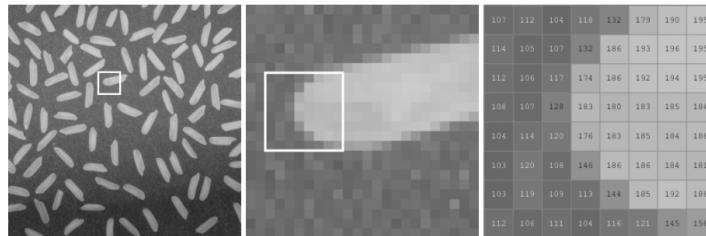


Figure 1.11 : Une image en niveaux de gris, agrandissement d'une zone de l'image, et affichage des valeurs constituant la matrice image. [52].

1.2.5 Anatomie de la peau

La peau est l'organe le plus étendu du corps humain. Elle mesure près deux mètres carrés et épaisseur de 2 mm en moyenne. Elle représente 16% du poids total de l'organisme. C'est notre allié contre les agressions du monde extérieur. Une barrière contre la pollution, les chocs, les microbes, les rayons UV du soleil mais aussi contre les maladies [47].

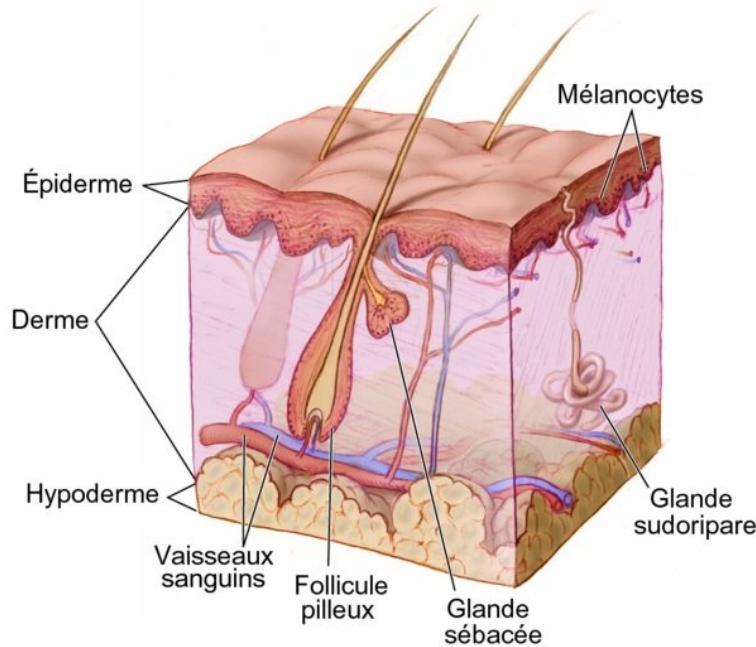


Figure 1.12 : Anatomie de la peau [47].

Selon la figure 1.12, la peau est constituée de trois tissus superposés : le tissu le plus externe est l'épiderme, le tissu intermédiaire est le derme et le tissu le plus profond est l'hypoderme.

A) L'épiderme

Tissu mince formé de couches de cellules juxtaposées qui recouvre le corps. Sa fonction principale est la protection de l'organisme contre les agressions extérieures. Les cellules qui composent l'épiderme sont :

- **Les cellules kératinocytes** : représentent 90 à 95% des cellules épidermiques. Elles produisent de la kératine.
- **Les cellules mélanocytes** : produisent de la mélanine dans la couche basale.
- **Les cellules de Langerhans** : représentent 2 à 4 % des cellules épidermiques. [23]

B) Le derme

Tissu épais qui se situe entre l'épiderme et l'hypoderme. Sa fonction principale est la protection et la réparation des tissus endommagés. Les principales cellules du derme sont fibroblastes, qui regroupent deux types de fibres protéiques : le collagène et l'élastine [25]

C) L'hypoderme

Tissu le plus épais qui se situe sous le derme, composé des cellules graisseuses et adipocytes. Son rôle principal est de diminuer les pressions auxquelles la peau est soumise. Il protège également l'organisme des variations de température.[26]

1.3 Cancer de la peau

Le cancer de la peau est une croissance anormale des cellules cutanées causée par des dommages non réparés de l'ADN qui déclenchent des mutations. C'est un type de cancer difficile à dépister et pouvant être grave. Il apparaît à n'importe quel âge. Il touche principalement les zones de peau qui sont les plus exposées aux ultra-violets (visage, bras...) [3].

Il existe deux grands types de cancers de la peau : les **carcinomes** (bas cellulaires ou épidermoïdes) qui représentent 90% des cancers cutanés et les **mélanomes** qui sont plus rares, mais aussi plus graves. [2]

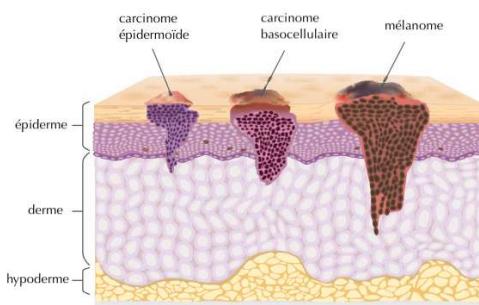


Figure 1.13 : Les différents types du cancer de la peau [53].

1.3.1 Les carcinomes

Les cancers cutanés non mégalonocytaires, qui regroupent les carcinomes basocellulaires (basaliomes) et les carcinomes épidermoïdes (spinocellulaires), constituent les néoplasies les plus fréquentes dans les populations à peau claire.[15] Les carcinomes cutanés se développent la plupart du temps au niveau de la région cervico-faciale [14]. Il y aurait plus de dix millions de carcinomes basocellulaires et près de 2,9 millions de carcinomes épidermoïdes diagnostiqués chaque année dans le monde.[28] Ils se divisent en 2 catégories :

A) Le carcinome basocellulaire

Le carcinome basocellulaire représente environ 70 des cas de cancer de la peau non mélanome. Néanmoins, il s'agit de cancers de faible gravité en raison de leur développement lent (au fil des mois ou des années) et localisé (Il se propage rarement à d'autres parties du corps). Comme son nom l'indique, ce type de cancer de la peau se développe à partir des cellules basales de l'épiderme (couche extérieure dans l'anatomie de la peau). [41]

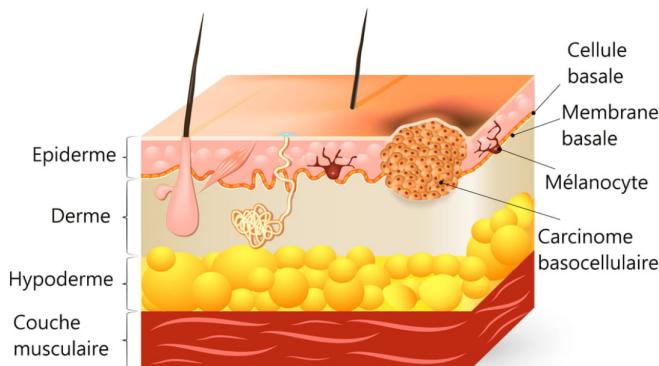


Figure 1.14 : Illustration du cancer basocellulaire. [54].

Ce carcinome peut prendre trois formes cliniques différentes :

- le carcinome basocellulaire nodulaire
- le carcinome sclérodermiforme
- le carcinome superficiel



Figure 1.15 : Des photos réels du cancer carcinome basocellulaire [55].

B) Les carcinomes épidermoïdes

Les carcinomes épidermoïdes sont plus rares que les carcinomes basocellulaires avec un pourcentage de 20 des cas de cancer cutané. Ils sont aussi plus agressifs mais généralement facile à traiter. Ils se développent rapidement sur plusieurs semaines ou mois. Ils peuvent apparaître n'importe où mais on les trouve surtout sur les zones photo exposées car les UV sont un facteur de risque important de ces carcinomes. S'il n'est pas traité, le SCC peut se propager à d'autres parties du corps.[33]



Figure 1.16 : Des photos réels du Carcinome épidermoïde [56].

1.3.2 Les melanomes

Le mélanome est un cancer grave et potentiellement mortel. Il peut apparaître dans n'importe quelle partie du corps contenant des mélanocytes.[44] Le mélanome ne représente qu'un faible pourcentage des cancers de la peau (4). Néanmoins, il est responsable de 75 de cas des décès par ce dernier [18]. Si le mélanome est détecté et traité à un stade précoce, les chances de guérison sont élevées, mais si le diagnostic est tardif, le mélanome peut se développer plus profondément dans la peau et se propager à d'autres parties du corps et il devient alors difficile à traiter.[44]. On distingue quatre formes de mélanome :

- Le mélanome superficiel extensif
- Le mélanome de dubreuil ou mélanome solaire

- Le mélanome acrolentigineux
- Le mélanome nodulaire



Figure 1.17 : Des photos réels du cancer mélanome [57].

1.4 Classification des images médicales

Depuis des décennies, le problème de la classification des lésions cutanées a été introduit chez la communauté d'apprentissage automatique. La classification automatisée des lésions peut à la fois aider les médecins à leur routine clinique quotidienne et permettre un accès rapide et bon marché à des diagnostics vitaux, même en dehors de l'hôpital, par installation d'applications sur des appareils mobiles. Ces raisons justifient l'intérêt de développer des approches puissantes de classification.

1.4.1 L'apprentissage profond

L'apprentissage profond ou le DEEP LEARNING est une branche des architectures d'apprentissage automatique (Machine Learning) qui tentent de modéliser des abstractions de haut niveau dans les données à l'aide de plusieurs couches de traitement.

A) Les neurones

Un neurone artificiel est un ensemble d'opérations mathématiques. Tout d'abord un poids et un biais sont appliqués de manière affine à une valeur d'entrée : en analyse d'images celle-ci est la valeur d'un pixel. Puis, une fonction d'activation est appliquée au résultat intermédiaire pour représenter les données dans l'espace des données de cette fonction. Souvent, cette fonction d'activation est non-linéaire, car elle permet de représenter des données complexes où la combinaison linéaire ne fonctionne pas.[12] les neurones sont Les briques élémentaires de l'apprentissage profond sont les réseaux de neurones, qui sont combinés pour former les réseaux de neurones profonds.

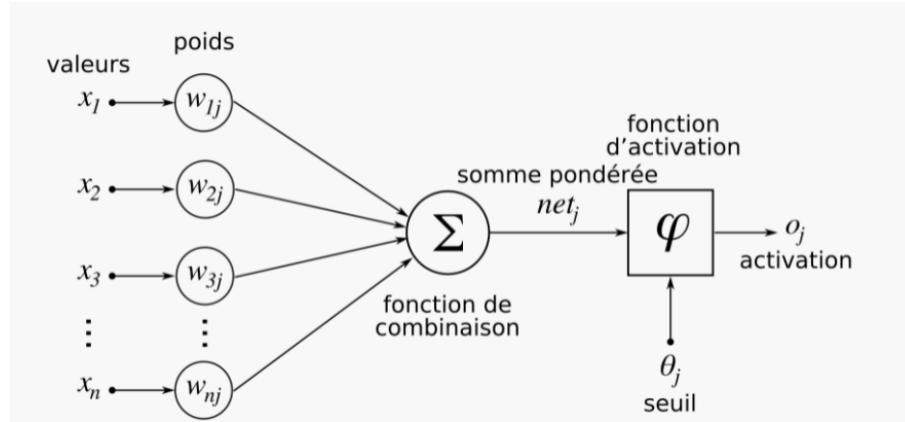


Figure 1.18 : Structure d'un neurone artificiel ou neurone formel [38]

B) Perceptron multicouche

Perceptron multicouche est l'un des modèles du réseau neurones le plus largement utilisé. Il est un réseau neuronal à propagation directe (feed forward). Il utilise la technique de rétro propagation pour l'apprentissage. Dans la figure suivante, la couche qui se trouve à gauche représente la couche d'entrée de neurones. Elle n'effectue aucun calcul contrairement aux éléments des autres couches.

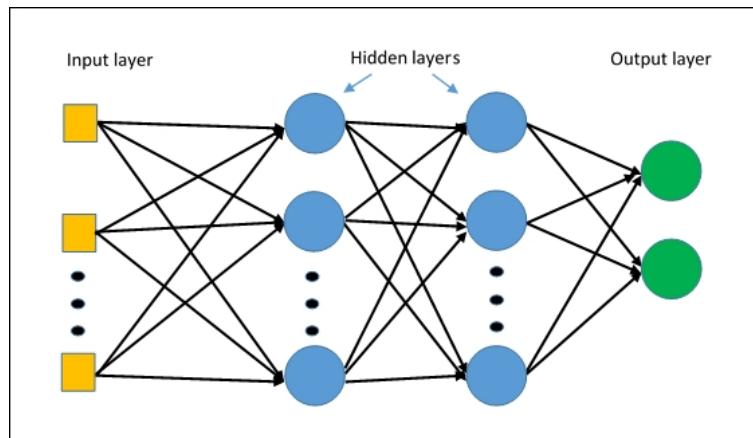


Figure 1.19 : Architecture de Perceptron multicouche [49]

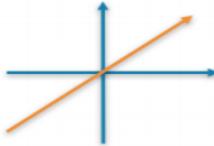
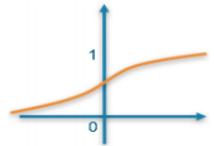
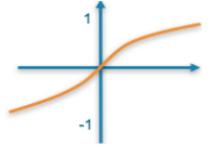
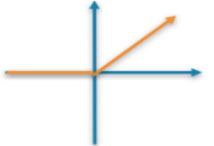
La couche centrale contient une ou plusieurs couches cachées de neurones qui calculent les données et subissent des répétitions. La couche de droite représente la couche de sortie qui produit les résultats. Chaque couche contient plusieurs neurones qui ne sont pas interconnectés avec des neurones de la même couche, mais tous les neurones d'une couche donnée sont entièrement connectés aux neurones dans les couches adjacentes. Les interconnexions sont associées à des valeurs numériques qui sont les poids synaptiques. Ces derniers sont multipliés par les neurones de la couche précédente, puis additionnés aux neurones de la couche actuelle.

C) Fonction d'activation

La fonction d'activation est une fonction mathématique utilisée par toutes les architectures de l'apprentissage profond. Elle permet de calculer le résultat de sortie du

neurone. On distingue plusieurs fonctions d'activation. Chaque fonction est détaillée dans le tableau ci-dessus

TAB. 1.1 : Les fonctions d'activation

Nom de la fonction	La formule	Graphe
Linéaire	$F(x) = ax$	
Sigmoïde	$F(x) = \frac{1}{1 + e^{-x}}$	
Tangent hyperbolique	$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	
Relu	$F(x) = \max 1, 0$	

1.4.2 Les réseaux de neurones convolutifs

Les CNN sont des réseaux de neurones multicouches avec une architecture spécifique qui s'est avérée très puissante dans des domaines tels que le traitement d'image médicale. Ils ont été démontrés pour mieux identifier les visages, les objets et les panneaux de signalisation mieux que les êtres humains.

Il existe différents types d'architectures convolutives tels que les réseaux de neurones convolutifs (CNNs), Les réseaux de neurones convolutifs profonds (DCNNs), Les réseaux de neurones convolutifs très profonds (Very DCNNs).

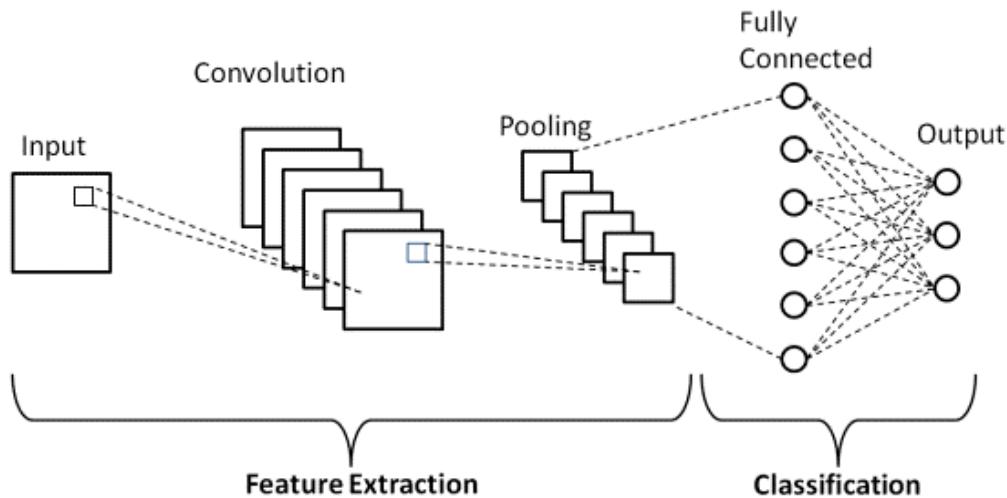


Figure 1.20 : Architecture d'un réseau de neurone convolutif [35]

Une architecture typique d'un réseau de neurones convolutif comporte quatre types de couches :

- **Couche de convolution**

Permet de passer l'image à travers un ensemble de filtres balayant quelques pixels à la fois et créant des cartes de caractéristiques (features map en anglais). Chaque filtre est multiplié par une portion de l'image. Les valeurs obtenues sont ensuite additionnées pour donner une seule valeur. La valeur résultante est placée dans la matrice de caractéristiques. C'est une carte indiquant où la caractéristique a été trouvée. Les valeurs des filtres ne sont pas connues au préalable, elles sont apprises dans la phase de l'apprentissage, c'est là où le réseau convolutif apprend à extraire des caractéristiques significatives. Définir une couche de convolution revient à déterminer quatre hyper paramètres :

- **Nombre de filtres** : correspond à la profondeur de la carte de caractéristiques. Plus ce nombre augmente, plus le nombre de caractéristiques que le réseau peut apprendre, augmente. Seulement, dans ce cas, l'apprentissage risque d'être très lent et on risque de tomber sur le problème de sur-apprentissage.
- **Taille de filtres** : définit la taille des cartes de caractéristiques en sortie.
- **Stride** : ce paramètre défini de combien de pixels le filtre de convolution se déplace à chaque étape.
- **Padding** : consiste à ajouter P zéros tout autour de l'image.

- **Couche de RELU**

Après chaque couche de convolution, le réseau applique une fonction mathématique sur les cartes de caractéristiques mises en évidence. Elle consiste à remplacer les valeurs négatives par des zéros.

- **Couche de pooling**

L'opération de pooling permet de remplacer un carré de pixels par une seule valeur en choisissant la valeur maximale de la grille (Max pooling) ou bien en calculant la moyenne de la grille (Average pooling) . Afin de réduire la taille des images et éviter le sur-apprentissage tout en préservant leurs caractéristiques importantes.

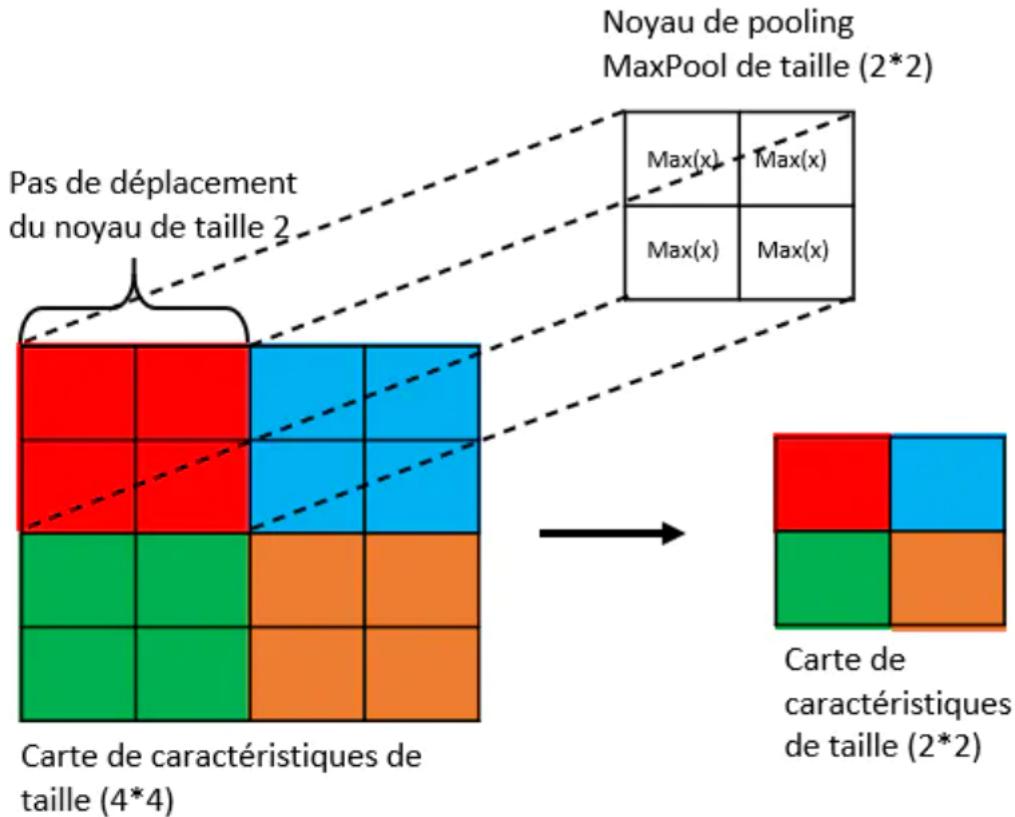


Figure 1.21 : Schéma d'une opération de pooling avec un noyau MaxPool de taille 2×2 et d'un pas de 2 [13].

- **Couche entièrement connectée**

Permet de classifier l'image en entrée du réseau : elle renvoie un vecteur de taille N, où N est le nombre de classes dans notre problème de classification d'images. Chaque élément du vecteur indique la probabilité pour l'image en entrée d'appartenir à une classe.

L'idée des réseaux neuronaux convolutifs profonds a été publiée à l'origine par Hinton, Krizhevsky et Sutskever et a été utilisée pour atteindre plus de performance dans le défi de classification ImageNet.[1] Le fonctionnement de CNN est similaire au fonctionnement de DCNN. Ils ne diffèrent que dans le nombre de couches. Puisque le nombre de couches de l'architecture CNN est entre 5 et 10 couches tandis que celles de l'architecture de DCNN est entre 50- 100 couches.

Un réseau de neurone convolutif très profond est une architecture similaire à celle du CNN et Deep CNN, ils ne diffèrent que dans :

- **Le nombre de couche** Le nombre de couche de very deep cnn est très grand par rapport aux architectures CNN et Deep CNN.
- **La combinaison de la couche** l'architecture Very Deep CNN contient une combinaison de ConvNet + Max pooling + Batch Normalization au lieu d'une combinaison de ConvNet et Max pooling seulement.
 - **Batch Normalization** Est une technique qui permet de faire facilement l'apprentissage des réseaux neurones très profonds en normalisant les entrées d'une couche pour chaque mini-batch.
La normalisation des entrées permet de stabiliser le processus d'apprentissage et, par conséquent, réduire considérablement le nombre d'époques d'apprentissage nécessaires pour former des réseaux profonds.[1]
 - **Dropout** Est une technique de régularisation qui aide à réduire le sur-apprentissage.

1.4.3 Les techniques d'apprentissage profond

Pour la classification des cancers de la peau, les réseaux de convolution peuvent être utilisés de deux manières différentes représentées dans la figure 1.22. La première consiste à utiliser les réseaux CNN comme **un extracteur des caractéristiques**. La deuxième manière consiste à entraîner les réseaux CNN de bout en bout. Dans ce cas, le modèle CNN est entraîné pour apprendre la relation entre les données en entrée et ses classes grâce à un apprentissage de bout en bout. Ici, l'extraction des caractéristiques et la classification forment un tout et ne sont pas une étape de prétraitement sur les données brutes. Cette deuxième méthode peut être opérée de deux façons, soit d'apprendre de modèle **à partir de zéro** ou avec **apprentissage par transfère**.

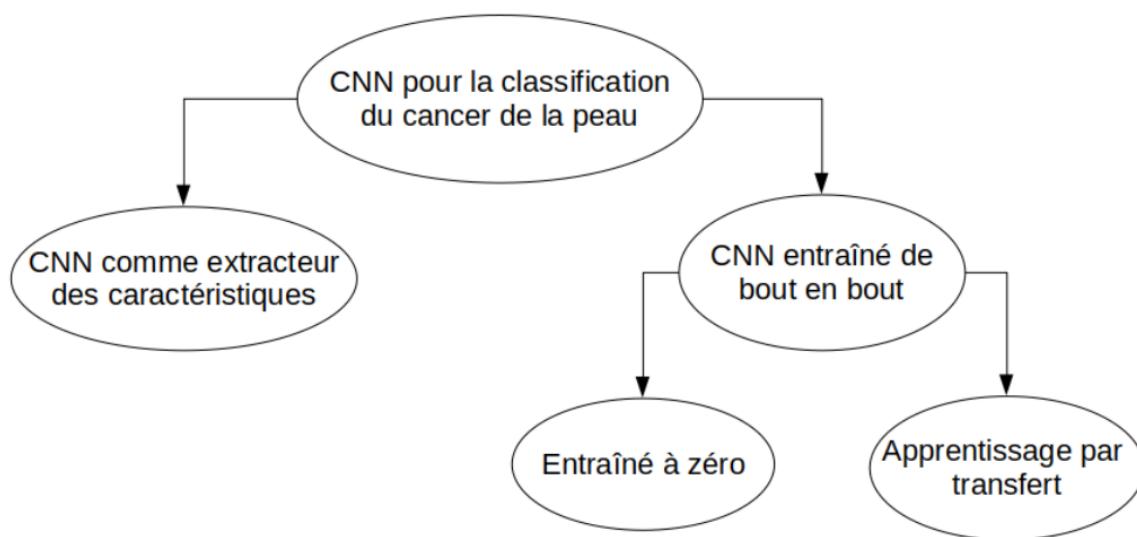


Figure 1.22 : Taxonomie proposée pour la classification du cancer de la peau.

A) l'apprentissage à partir de zéro

C'est une méthode d'entraînement des modèles dans l'apprentissage profond dans laquel, les poids et les paramètres sont initialisés de manière incertaine. Pour que le modèle puisse atteindre des bons résultats, on aura besoin de :

- Stabiliser le gradient avec la méthode Batch Normalization.
- un grand temps pour que le modèle converge.
- Des données suffisantes pour que le modèle s'entraîne bien. [62]

B) l'apprentissage par transfert

Les DCNNs ont généralement besoin de grands ensembles de données d'images pour atteindre une précision prédictive élevée. Mais, dans de nombreux domaines, l'acquisition de telles données est difficile .C'est pour cela que l'utilisation de caractéristiques de DCNNs déjà établis (tels que VGG-16, AlexNet et GoogLeNet) pré-entraînés sur des ensembles de données d'images à grande échelle (tels que ImageNet), s'est avérée très utile pour résoudre les problèmes de classification d'images grâce au concept apprentissage par transfert. [45] L'idée derrière l'apprentissage par transfert est qu'il est moins coûteux et plus efficace d'utiliser des modèles d'apprentissage profond entraînés sur des ensembles de données d'images " big data " (comme ImageNet) et de " transférer " leur capacité d'apprentissage à un nouveau modèle de classification plutôt que d'entraîner un classificateur DCNN à partir de zéro [6]. En général, il est préférable de faire l'apprentissage avec les poids obtenus lors d'un apprentissage précédent plutôt que de le commencer avec des poids initialisés de façon incertaine [32] "As the difference between the pre-training dataset and the Target task data increases, the transferability will decrease" Yosinski et al.[61]

Deux scénarios différents peuvent être étudiés pour réussir à entraîner un modèle pré-entraîné sur les images des lésions cutanées.Ces scénarios dépendent de la taille des données et de la similarité entre eux. Le premier scénario est celui où nous disposons d'un petit ensemble de données et d'une faible similarité des données. Dans ce cas, il est recommandé de fixer les couches initiales du réseau, disons k couches, et de recycler les les n-k couches restantes. Le second scénario est celui où nous avons un petit ensemble de données et une grande similarité des données. Dans ce cas, nous supprimons la dernière couche du réseau, appelée couche entièrement connectée ou couche dense, et fixons l'ensemble du réseau. Nous modifions ensuite la dernière couche pour classifier soit bénin ou malin. Les couches fixées sont utilisées comme extracteur de caractéristiques, et la dernière couche comme classificateur de caractéristiques.

C) Extraction des caractéristiques

Dans ce cas, le réseau CNN est préentraîné sur un autre grand dataset et la classification est effectuée par un autre classifieur tel que les k plus proches voisins(KNN), les machines à vecteurs de support (SVM) ou les réseaux de neurones artificiels (ANN).

Cette méthode initialise le modèle avec les poids d'un autre modèle entraîné déjà sur un dataset plus grand (tel que ImageNet), le modèle pré-entraîné appris des caractéristiques qui pourraient être utiles pour la tâche de classification .Cela correspond, en pratique, à l'utilisation d'une ou de plusieurs couches sélectionnées du

CNN pré-entraîné comme un extracteur de caractéristiques fixes, ce qui peut être réalisé en gardant tous les blocs convolutionnels et en entraînant uniquement les couches entièrement connectées sur le nouveau dataset(voir figure 1.23).

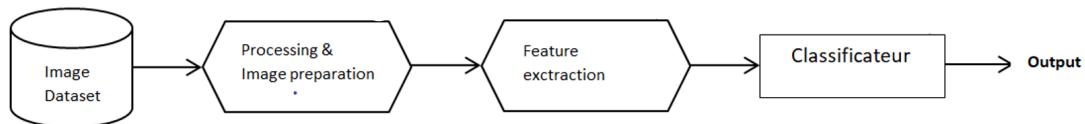


Figure 1.23 : représentation graphique de la méthode d'extraction des caractéristiques

1.5 Conclusion

Dans ce chapitre, nous avons donné un aperçu du contexte médical concernant le mélanome. Ce dernier se développe à partir des cellules de cancer de type mélanocytes. Nous avons également vu que le diagnostic visuel de mélanome est une tâche difficile en raison de la similitude entre les cellules bénignes et malignes. Par ailleurs, nous avons constaté que l'automatisation des systèmes de diagnostic a amélioré la précision de détection de mélanome notamment ceux qui ont utilisé des approches d'apprentissage profond pour la tâche de classification.

Chapitre 2

État de l'art

2.1 Introduction

Il existe aujourd’hui plusieurs recherches autour de la problématique de classification des images médicales. En effet, de par son intérêt grandissant pour des fins de traitement clinique et d’enseignement, plusieurs approches de classification d’images ont été proposées dans les différents domaines médicaux ayant besoin d’une automatisation de la tâche de catégorisation des images et cela comme un outil d’aide au diagnostic.

Récemment, plusieurs approches traitant de la problématique de classification d’images médicales ont été conçues sur la base des réseaux neuronaux profonds. Ces derniers s’imposent aujourd’hui comme une méthode d’apprentissage au potentiel prometteur pour différentes tâches de classification. En particulier, et lorsqu’on traite des images, les réseaux de neurones de convolution sont souvent cités comme donnant les meilleurs résultats.

Dans ce chapitre, nous allons discuter les travaux réalisés à base d’apprentissage profond pour la classification du cancer de la peau. En effet, après avoir donné des éléments du contexte sur l’application de l’apprentissage profond dans l’imagerie médicale, nous nous focalisons ici sur les travaux directement liés à notre thématique à savoir la classification des images de la peau dans une perspective de dépistage les types du cancer de peau le plus répandu. Pour ce faire, nous allons d’abord citer quelques statistiques et prédictions concernant le cancer de la peau. Ensuite, nous allons proposer une taxonomie pour décrire les travaux réalisés dans ce domaine à base des réseaux de convolution. À la fin, nous allons discuter les techniques et les algorithmes utilisés.

Au cours des dernières années, le nombre de cas de personnes atteintes du cancer de la peau a augmenté de 53%. Ceci est en partie dû à une exposition accrue aux rayons Ultra-Violet. Bien que le mélanome soit l’un des types de cancer de la peau les plus meurtriers, un diagnostic précoce peut permettre de facilement le guérir.

En outre, les statistiques de l’année 2019 aux états unis montrent que plus de 5.4 millions de personnes ont un cancer de la peau. De plus, des prédictions pour l’année 2020 ont estimé que le nombre de nouveaux cas de cancer de la peau diagnostiqués augmentera de près de 2% et le nombre de décès à cause de ce cancer augmentera de 3 à 5 %.

La première étape du diagnostic d’une lésion maligne par un dermatologue est un examen visuel de la zone cutanée suspecte. Un diagnostic précis est important en raison des fortes similitudes entre certains types de lésions. De plus, la précision du diagnostic est corrélée fortement avec l’expérience professionnelle du médecin. Sans support technique supplémentaire, les dermatologues ont un taux de précision de 65% à 80% dans le diagnostique du cancer de la peau . Dans des cas suspects, l’inspection visuelle est complétée par des images dermatoscopiques prises avec une caméra spéciale de haute résolution. Avec cette technique de soutien, la précision du diagnostique des lésions cutanées peut être augmentée de 49%. La combinaison de l’inspection visuelle et des images dermatoscopiques se traduit finalement par une absolue précision de détection du cancer de la peau de 75% à 84% par les dermatologues.

Depuis quelque temps, le problème de la classification des lésions cutanées a été introduit chez la communauté d’apprentissage automatique. La classification automatisée des lésions peut à la fois aider les médecins à leur routine clinique quotidienne et permettre un accès rapide et bon marché à des diagnostics vitaux, même en dehors de l’hôpital,

par installation d'applications sur des appareils mobiles. Ces raisons justifient l'intérêt de développer des approches puissantes de classification.

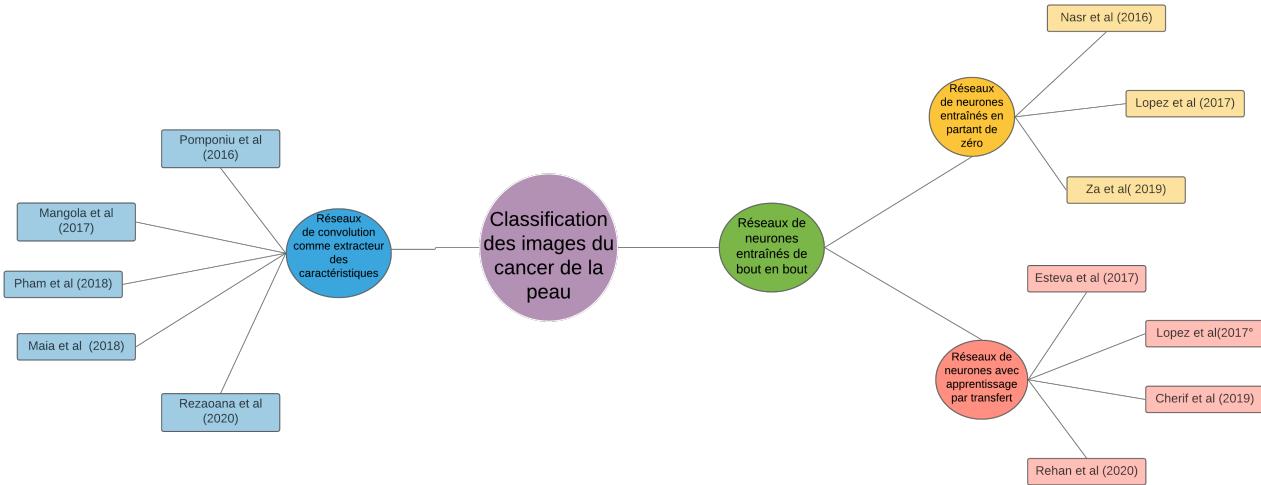


Figure 2.1 : Taxonomie représentant les travaux cités dans notre état de l'art.

2.2 Réseaux de neurones entraînés de bout en bout

Dans cette partie, nous allons discuter des travaux qui ont utilisé différentes architectures CNN entraînées de bout en bout pour la classification du cancer de la peau. Ces architectures suivent deux stratégies. La première consiste à entraîner le CNN en partant de zéro (from scratch). Cela veut dire que les paramètres de l'architecture sont initialisés aléatoirement au début. La deuxième stratégie consiste à entraîner le CNN avec les techniques d'apprentissage par transfert. Dans ce deuxième cas, les paramètres de l'architecture sont initialisés à des valeurs obtenues suite à un apprentissage préalable sur d'autres datasets comportant un nombre important d'images. Les valeurs de ces paramètres seront peaufinées sur le dataset du cancer de la peau.

2.2.1 Réseaux de neurones entraînés en partant de zéro

Concernant cette méthode, Nasr et al [31] ont utilisé une architecture CNN entraînée de bout en bout pour la classification du mélanome en utilisant des images .

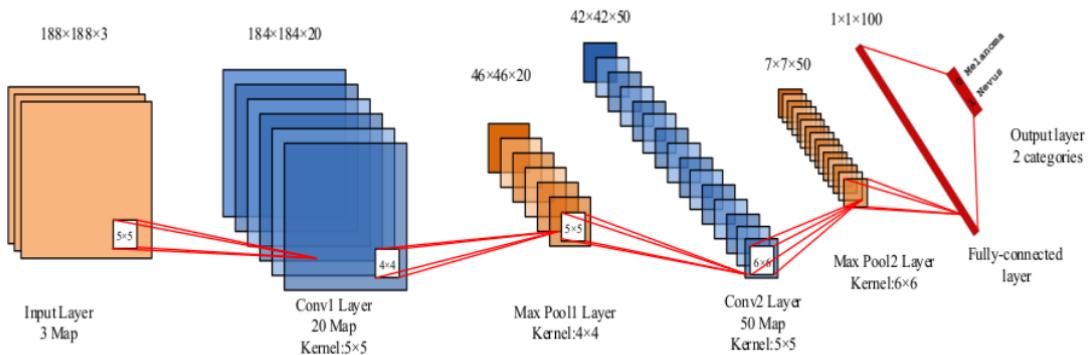


Figure 2.2 : Architecture proposée par Nasr et al [31]

Comme illustrée dans la Figure 2.2, l’architecture proposée par [31] est composée de 2 couches de convolution suivie d’une couche de max pooling pour l’extraction des caractéristiques. Une couche complètement connectée est rajoutée à la fin du réseau pour déterminer la classe de l’image en entrée. De plus, ils ont utilisé d’autres techniques de traitement d’images avant entraînement pour aider à extraire les caractéristiques. Parmi ces techniques on cite la correction d’éclairage, la génération de masque ainsi que l’application d’un filtre Gaussian. Pour tester les performances de cette architecture, ils ont utilisé un ensemble de données de petit taille. Cet ensemble de données est divisé en 136 images pour l’entraînement et 34 pour le test. La précision réalisée est de l’ordre de 81%.

En outre, Romero Lopez et al [40] ont choisi l’architecture VGG16 (voir la figure 2.3) entraînée de bout en bout (from scratch) sur le dataset fourni par la compétition ISIC-2016 pour la classification binaire du cancer de la peau (mélanome et Non mélanome), De plus, ils ont utilisé d’autres techniques de traitement d’images avant l’entraînement pour aider à extraire les caractéristiques. Parmi ces techniques on cite le basculement horizontal, la rotation, le cisaillement et le zoom. Cet ensemble de données est composée de 900 images d’entraînement et 379 images de test. Les auteurs ont achevé une précision équivalente à 66%.

Par contre, Md Zahangir Alom et al [4] ont appliqué l’architecture IRRCNN pour la classification du cancer de la peau. Cette architecture est composée de trois unités IRR suivies d’une couche de pooling. A la fin de ce modèle, une couche de GAP est ajoutée et suivie par une couche softmax. Pour l’entraînement de cette architecture, les auteurs ont utilisé l’ensemble de données ISIC 2018 qui contient 10015 images dermatoscopiques de 7 classes. Après le redimensionnement de ces images de 650×450 à 192×192 et la normalisation, Il y a eu 7010 images pour l’entraînement et 3005 pour le test. Les auteurs ont comparé les performances de cette architecture avec et sans augmentation. L’architecture avec augmentation a obtenu une meilleure précision équivalente à 87,09%.

2.2.2 Réseaux de neurones avec apprentissage par transfert

À cause du manque important des données et la difficulté de trouver un ensemble de données suffisant pour la problématique considérée, plusieurs travaux utilisent la tech-

nique d'apprentissage par transfert. Cette technique a prouvé une efficacité et donne les meilleures performances obtenues jusqu'à là pour le problème de classification.

Citons d'abord le travail de ROMERO-LOPEZ et al. [27]. Les auteurs ont comparé entre trois méthodes pour la classification binaire du cancer de la peau (mélanome contre non-mélanome). La première méthode consiste à entraîner l'architecture avec des paramètres initialisés aléatoirement. La deuxième méthode consiste à appliquer l'apprentissage par transfert sur la partie extraction des caractéristiques seulement et entraîner la partie classifieur sur le dataset choisi. La dernière méthode consiste à appliquer l'apprentissage par transfert et optimiser tous les paramètres de l'architecture sur le nouveau dataset. Pour ce faire, ils ont choisi l'architecture VGG16 (Fig 2.3) [46] initialiser aléatoirement pour la première méthode, préentraînée sur le dataset ImageNet [58] pour la deuxième et la troisième méthodes . Ils ont travaillé sur le dataset fourni par la compétition ISIC-2016 [59] qui est composée de 346 images d'entraînement et 150 images de test. Pour l'entraînement ils ont procédé comme suit. Ils ont d'abord effectué une normalisation des images et ont rendu la valeur de chaque pixel dans $[0; 1]$. Ils ont ensuite effectué un recadrage des images suivi d'un redimensionnement des images à la taille 224x224. Pour le test, la dernière méthode a donné les meilleures performances avec une précision de 81.33% et a démontré le rôle d'apprentissage par transfert avec optimisation de tous les paramètres de l'architecture. Par contre, la deuxième méthode a donné les plus faibles résultats malgré l'application d'apprentissage par transfert. Cela peut être expliqué par le type de données inadaptées utilisées pour le transfert.

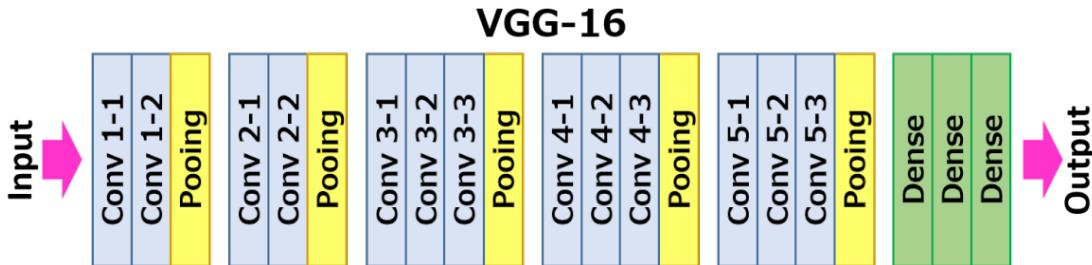


Figure 2.3 : Architecture VGG16 [60].

De plus, Esteva et al.[17] sont les premiers à avoir entraîné un modèle sur un grand dataset composé de 129450 images cliniques dont 3374 sont des images obtenues des appareils dermatoscopiques. Ce dataset a été divisé sur 2032 lésions de la peau. Esteva et al.[17] ont travaillé sur deux problèmes de classifications binaires. Le premier consiste en la classification entre carcinomes kératinocytaires et kératose séborrhéique bénigne qui permettent d'identifier le cancer le plus commun. Le deuxième est la classification entre mélanomes malins contre naevus bénins qui permettent d'identifier le cancer le plus dangereux. Les auteurs ont travaillé avec l'architecture Inception v3 [48] préentraîné avec une approche différente (voir Fig 2.4).

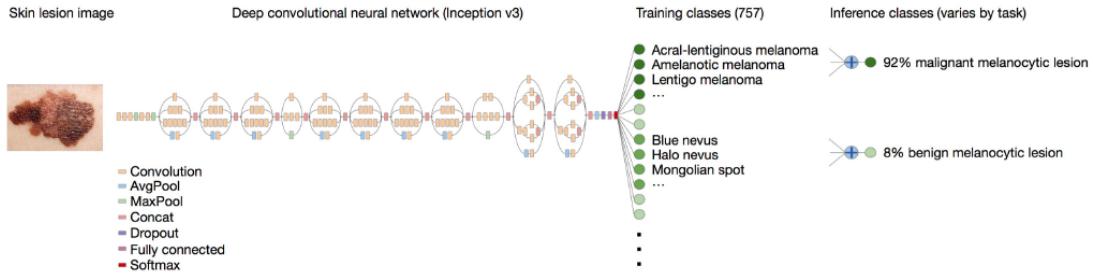


Figure 2.4 : Architecture Inception [27].

Pour valider les performances des leurs CNN, les auteurs ont comparé leurs résultats avec 21 dermatologues certifiés sur l'épiderme et la mélanocytose. Pour la première tâche de classification, le réseau de convolution a eu une précision de 72.1% par contre les dermatologues ont une moyenne de 66%. Concernant la deuxième tâche, CNN a réussi à avoir une précision 55.5% en comparaison aux médecins qui ont eu 54.3 %.

Un autre travail à inscrire dans cette catégorie est celui proposé par ECH-CHERIF et al.[16] qui ont développé une architecture de convolution pour la classification.

Les auteurs ont entraîné manuellement un modèle CNN appelé MobileNetV2 (Fig 2.5) en utilisant l'apprentissage par transfert pour la classification binaire des lésions cutanées en classes bénignes ou malignes. Pour tester les performances de cette architecture, les auteurs ont utilisé un dataset de 48 373 images dermoscopiques collectées à partir de trois archives différentes, étiquetées et validées par des dermatologues experts. Le modèle entraîné a donné une précision globale de 91,33 %.

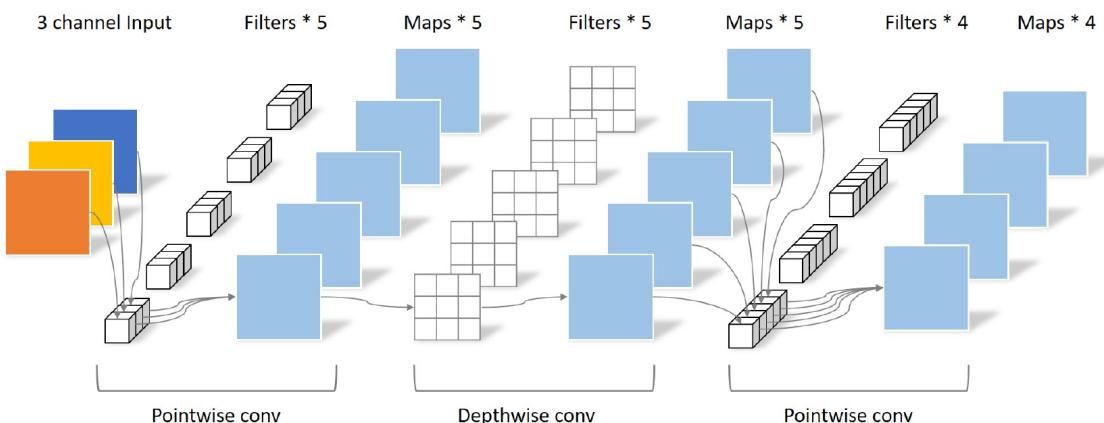


Figure 2.5 : Architecture MobileNetV2 [42].

Rehan et al .[5] ont contribué dans la classification du cancer de la peau en collectant des datasets des images dermoscopiques et cliniques fournis par les deux datasets Derm-Quest et DermIS. Les images sont divisées en deux parties, à savoir le « training » (77%) et le « test » (23%). et les partitions des deux parties sont garanties d'avoir à la fois des cas de mélanome et de non mélanome. Ils ont entraîné le réseaux de neurones AlexNet en utilisant l'apprentissage par transfert pour la classification binaire. Dans le modèle AlexNet, il y a un total de huit couches dont 5 couches convolutionnelles et trois couches

entièrement connectées. Il y a une couche de max pooling entre les deux premières couches convolutives. Une couche de pooling est ajoutée après les couches convolutives pour réduire la taille de la carte de caractéristiques. Pendant la procédure d'apprentissage, un petit taux d'apprentissage est utilisé parce qu'il donne des changements plus petits pour mettre à jour le poids. Donc ,il y a cinq couches transférées et trois nouvelles couches pour l'adaptation. Seule la nouvelle couche est entraînée pour une classification précise du cancer. Les auteurs ont entraîné ce modèle sur deux datasets différents avec et sans augmentation. Les meilleures performances ont été obtenues par le modèle entraîné sur les images ROI du dataset DermIS avec augmentation dont la précision atteint 97.9%.

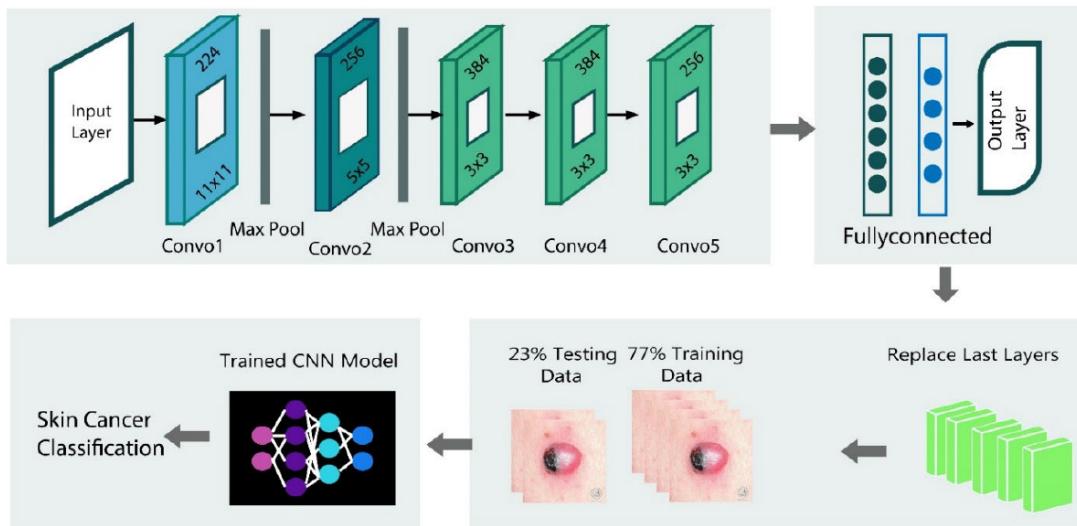


Figure 2.6 : Architecture proposée par Rehan et al [5].

2.3 Extraction des caractéristiques par CNN

Dans cette partie, nous allons détailler les travaux se basant sur des architectures CNN comme extracteur de caractéristiques pour la classification d'images du cancer de la peau. Nous allons citer chronologiquement les travaux qui ont gardé seulement la partie d'extraction des réseaux existants et ont uniquement remplacé la dernière couche, c'est-à-dire la couche complètement connectée, par un autre type de classificateur.

L'architecture proposée par Pomponiu et al.[36] est un autre exemple d'architecture s'inscrivant dans cette catégorie. Les auteurs ont proposé l'architecture présentée dans la Fig 2.7 Cette architecture est composée d'un réseau de convolution AlexNet également préentraîné sur ImageNet et un classifieur de type K-plus proche voisin (Knn). Les auteurs ont placé trois classificateurs dans des positions différentes. Ils ont choisi la similarité cosinus comme fonction de distance car elle est robuste et la plus utilisée dans travaux précédents. Pour l'évaluation, ils ont utilisé seulement 399 images capturées par une caméra digitale avec des différentes résolutions. Ces données sont issues du datasets DermQuest et DermIS qui comporte 217 images bénignes et 182 malignes. En vue du faible nombre de données, les auteurs ont eu besoin de recourir à des techniques d'augmentation de données. Le

nouvel ensemble d'apprentissage comportait 10000 images avec lequel une précision de 93.64% a été atteinte.

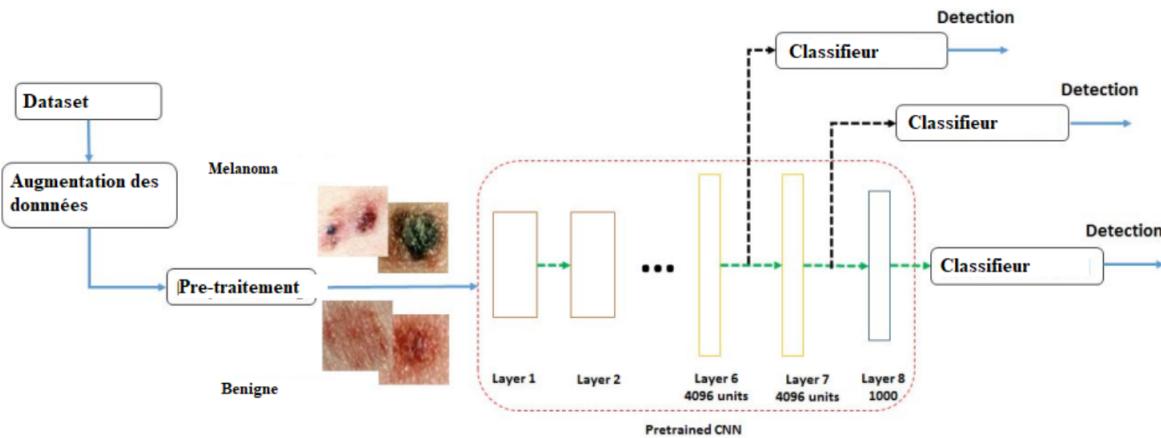


Figure 2.7 : Architecture proposée par Pomponiu et al [36].

Dans un autre registre, Menegola et al.[30] ont proposé de montrer le rôle d'apprentissage par transfert sur la classification du cancer de la peau. Les auteurs ont proposé d'utiliser deux datasets relatifs à deux domaines différents. Pour cela, ils ont utilisé, en plus du dataset ImageNet, le dataset Retinopathy. Il s'agit d'un dataset dans le domaine médical qui comporte plus de 35000 images pour la rétinopathie diabétique. Ils ont également utilisé deux datasets pour le cancer de la peau, l'un fourni par la compétition ISIC-2016 et l'autre dénommé le dataset Atlas. Pour l'extraction des caractéristiques, les auteurs ont comparé deux modèles VGG16 et VGG-M pour démontrer le rôle de la profondeur du réseau de neurones. Ils ont déployé un classifieur SVM pour trois types de classification à savoir mélanome et bénigne, maligne, bénigne et Mélanome et enfin bénigne et carcinome. Pour tester le rôle de l'apprentissage par transfert, ils ont essayé plusieurs scénarios dont un modèle initialisé aléatoirement, un modèle préentraîné sur le dataset Retinopathy , un modèle préentraîné sur le dataset ImageNet . Ils ont obtenu la meilleure précision, à savoir 83.8% , avec le modèle VGG-16 préentraîné sur le dataset ImageNet. Ils ont comparé les performances obtenues et ils ont abouti à deux principaux résultats. Premièrement, l'application d'apprentissage par transfert donne de meilleurs résultats en utilisant, pour le transfert, un dataset appartenant à un domaine autre que celui du mélanome et cela par rapport à l'utilisation d'un dataset dans le même domaine. Deuxièmement, la profondeur du réseau utilisé influe sur les performances du modèle de sorte que plus le modèle est profond plus il donnera de meilleurs résultats.

Pham et al.[34] a démontré l'application des techniques d'augmentation des données sur la classification du cancer de la peau pour augmenter les performances des modèles profonds de convolution et explorer l'influence de ces techniques sur trois classificateurs différents. Pour cela, ils ont travaillé sur un dataset qui contient 6762 images dermatoscopiques de deux classes mélanome et bénigne. Il y a eu 6162 images pour l'entraînement et la validation et 600 images pour le test. Ces données sont fournies par la compétition ISIC 2017.

Pour résoudre le problème du déséquilibre entre le nombre de données des classes, ils

ont utilisé différentes méthodes pour l'augmentation des données en gardant la sémantique des images. Parmi ces méthodes, on cite les transformations géométriques, comme la rotation aléatoire, la translation, miroir, changement d'échelle, cisaillement et recadrage, et l'augmentation de couleur en utilisant la normalisation des données.

La Figure 2.8 représente la stratégie utilisée par Pham et al.[34] pour la classification. Pour l'extraction des caractéristiques, les auteurs ont d'abord choisi le réseau de neurones de convolution InceptionV4 pré entraîné sur le dataset ImageNet sans utiliser le classificateur. Ensuite pour la classification, ils ont choisi trois (03) différents classificateurs pour démontrer l'effet des techniques d'augmentation des données. Il s'agit des machines à vecteurs de supports (SVM), des forêts des arbres décisionnels (Random forest) et les réseaux de neurones (NN). Le modèle avec un réseau de neurones comme classificateur a obtenu les meilleurs résultats avec une précision de 89%. Les performances des classificateurs SVM et RF ont eu un effet néfaste lorsque nous avons modifié l'augmentation des données de DAUG 50 à DAUG 100.

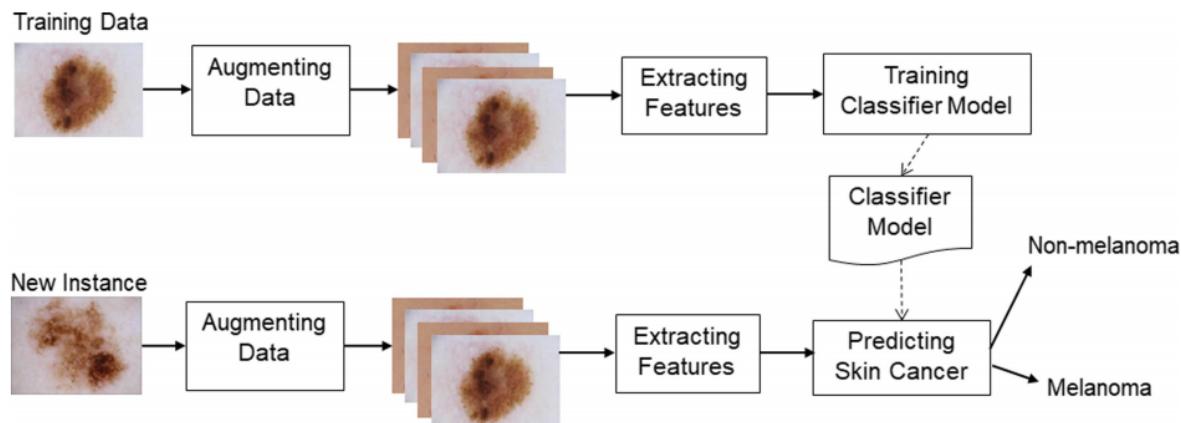


Figure 2.8 : Architecture proposée par Pham et al [34].

Un autre travail à inscrire dans cette catégorie est celui proposé par Maia et al [29], Les acteurs on comparés entre les principaux modèles de CNN en ce qui concerne la tâche de classification du cancer de la peau. Ses modèles sont pré entraînés sur le dataset ImageNet pour l'extraction des caractéristiques .Quatre ensembles de vecteurs caractéristiques ont été générés et testés séparément avec six types de classificateurs pour la tâche de diagnostic de mélanome : régression Logistique, machines à vecteurs de supports (SVM) avec des noyaux linéaires et radiaux, Naive Bayes, adaptive boosting (AdaBoost) et les forêts des arbres décisionnels.

Les auteurs [29] ont utilisés le dataset de petite taille PH2 pour l'évaluation. Ce dataset comporte 200 images dermatoscopiques dont 40 images de type mélanome, 160 images non mélanome.

D'après le tableau 2.1 ci-dessous,Nous remarquons que l'architecture VGG19 avec une régression logistique comme classificateur a obtenu les meilleurs résultats avec une précision de 92,5%.

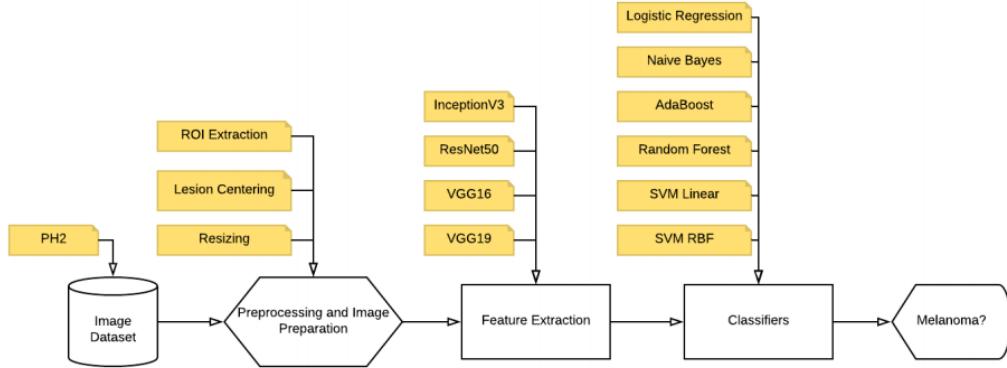


Figure 2.9 : Organigramme représentant la méthodologie proposée par Maia et al [29].

TAB. 2.1 : Tableau représentant les résultats obtenus par la combinaison des différentes architectures avec les différents classificateurs sur le dataset PH2 [29]

Architecture	Classificateur	Precision
InceptionV3	RL	0.9250
	SVM Linear	0.9200
	Naive Bayes	0.7800
	SVM RBF	0.8700
	AdaBoost	0.7800
	Random Forest	0.8250
ResNet50	SVM Linear	0.9100
	RL	0.9000
	Random Forest	0.8850
	Naive Bayes	0.8100
	SVM RBF	0.8850
	AdaBoost	0.8300
VGG16	RL	0.9150
	SVM Linear	0.9100
	SVM RBF	0.8950
	Random Forest	0.8950
	Naive Bayes	0.8400
	AdaBoost	0.8700
VGG19	RL	0.9250
	SVM Linear	0.9050
	Random Forest	0.9100
	SVM RBF	0.8950
	Naive Bayes	0.8150
	AdaBoost	0.8450

Rezaoana et al [39] ont développé une nouvelle architecture de convolution présentée dans la figure 2 pour la classification du cancer de la peau. Cette architecture est com-

posée de deux sections. la première section est pour l'extraction des caractéristiques, elle contient deux branches parallèles concaténées à la fin. Chaque branche contient cinq (5) unités convolutives dilatées suivie d'une couche de max-pooling . La deuxième section est pour la classification et la détection. De plus, ils ont utilisé d'autres techniques de traitement d'images avant l'entraînement pour aider à extraire les caractéristiques. Parmi ces techniques on cite le basculement horizontal, la rotation, le cisaillement et le zoom.

Pour tester les performances de cette architecture, les auteurs ont utilisé un dataset contient 25 780 images bénignes et malignes classés selon neuf (9) types de cancer. Ce dataset est divisé en 19537 images pour l'entraînement, 2167 pour la validation et 3799 pour le test.

Ils ont comparé leurs résultats avec les architectures VGG16 et VGG19. Leur architecture a obtenu une meilleure précision égale à 79.45%.

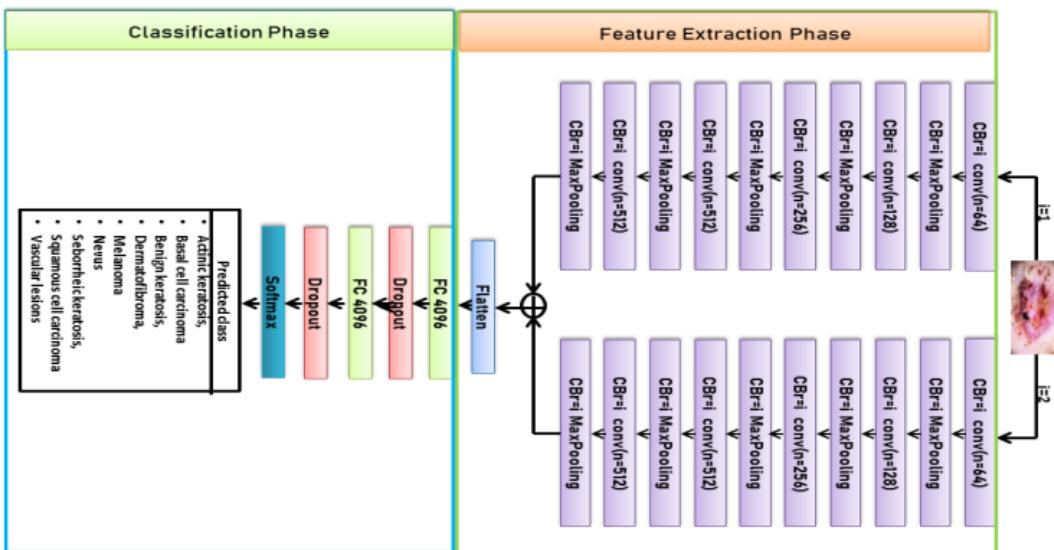


Figure 2.10 : Architecture proposée par Rezaoana et al [39].

À la fin de cette section, nous résumons les travaux existants dans la Table 2.2 ainsi que les techniques et les datasets sur lesquels ils étaient évalué.

TAB. 2.2 : Tableau des travaux concernant la classification du cancer de la peau

Type	Titre	Dataset	Architecture	précision
CNN Comme extracteur des caractéristiques	Pomponiu et al(2016) [36]	DermQuest+ DermIS	AlexNet+KNN	93,64%
	Menegola et al(2017) [30]	ISIC 2016 + Atlas	VGG16 et VGGM + SVM	83,8%
	Pham et al(2018)[34]	ISIC 2017	InceptionV4 + (SVM,RF,NN)	89%
	Maia et al(2018) [29]	PH2	VGG16,VGG19, ResNet50 et Inception v3 + (SVM+NN+ LRN+RF+AD)	92.5%
CNN à partir de zéro	Rezaoana et al(2020) [39]	25 780 images	Proposée	79 ,45%
	Nasr et al (2016) [31]		CNN	81%
	Lopez et al (2017) [40]	ISIC-2016	VGG16	66%
	ZA et al (2019) [4]	ISIC 2018	IRRCNN	87.09%
CNN avec apprentissage par transfert	Lopez et al (2017) [27]	ISIC-2016	VGG16	81.33%
	Esteva et al(2017) [17]	ISIC archive+Dermofit	Inception V3	72.1%
	Cherif et al(2019) [16]	DermNet+ISIC archive+Dermofit	MobileNetV2	91.33%
	Rehan et al (2020) [5]	DermIS+ DermQuest	AlexNet	97.9%

Au terme de notre présentation des travaux traitants de l'analyse des images médicales, en particulier images cutanées, nous avons réalisé que l'apprentissage profond a permis d'avoir des résultats prometteurs pour ce domaine. De plus, les auteurs ont utilisé

différentes techniques et architectures d'apprentissage profond mais il s'est avéré que les réseaux de convolution sont les plus adaptés pour la classification des images soit comme extracteur des caractéristiques ou bien entraîné de bout en bout .

D'après les résultats affichés dans le tableau comparatif, On constate que les auteurs qui ont utilisés la méthode d'apprentissage à partir de zéro, on obtenu les mauvaises performances par rapport aux autre travaux.

Par contre, l'apprentissage par transfert a permis d'avoir de meilleurs résultats et devenu plus utilisé dans ce domaine pour répondre au problème de manque de données qui est un facteur important pour valoriser les travaux. Par exemple [[5],[29]] ont des modèles avec de très bonnes performances mais sur un ensemble de données de taille réduite.

En outre, nous constatons aussi que le choix d'architecture a un impact important sur les résultats obtenus, par exemple [16], [17] ont utilisés deux différentes architectures sur le même dataset, mais la précision de [16] était plus élevée que celle de [17].

L'approche de « bout en bout » extraire que les caractéristiques standards comme la couleur, la texture et la forme. tandis que, « features extraction » permet d'extraire les caractéristiques selon l'image traitée pour une meilleure exploitation de cette dernière. Cela a mené d'atteindre des résultats plus performants. Par exemple, on utilisons le même modèle d'architecture sur le même dataset,[30] a obtenu des résultats mieux que celle [40].

2.4 Conclusion

Dans ce chapitre, Nous avons commencé par présenté des statistiques et des prédictions concernant la propagation de cette maladie et les résultats des diagnostics avec les méthodes classiques. Ensuite, Nous avons cité quelques travaux récents réalisés dans domaine médical à l'aide des techniques d'apprentissage profond pour la détection du cancer de la peau.

A la fin, Nous avons établi un tableau comparatif et rédigé une discussion critique sur les travaux réalisés dans ce domaine. D'après les résultats cités, les techniques d'apprentissage profond ont une grande importance dans la classification du cancer de la peau. De plus, les réseaux de convolution sont les plus exploités pour cette maladie avec différentes techniques. L'apprentissage par transfert a donné des résultats remarquables et il permet d'améliorer les performances des modèles en exploitant d'autres ensembles de données plus large.

Chapitre 3

Approche adaptée pour la classification d'images médicales

3.1 Introduction

Sur la base de notre étude des travaux de l’état de l’art et des discussions qui se sont dégagées, nous proposons une nouvelle architecture à base des techniques d’apprentissage profond pour répondre au problème de la classification du cancer de la peau. Notre contribution tire profit de différents concepts retrouvés en apprentissage profond. En effet, Nous proposons d’appliquer les différentes techniques d’apprentissage pour cette problématique particulière et les évaluer pour mesurer le degré de validité de nos hypothèses.

Pour ce faire, Nous décrivons l’architecture d’apprentissage profond utilisée comme composante dans notre proposition dans la 2ème section. Ensuite, nous décrivons en détails notre proposition et ses variétés dans la 3ème section. A la fin, Nous évaluons nos approches proposés dans la 4ème section .Tout d’abord, nous citons les différentes métriques utilisées pour mesurer la performance. Ensuite, nous décrivons l’ensemble des données utilisé, en mettant en relief les difficultés de cet ensemble. Puis, Nous allons justifier le choix d’opérer avec ce dataset par des comparaisons avec d’autres datasets. Nous présentons les résultats de nos approches en illustrant avec des graphes d’apprentissage et les performances obtenues sur l’ensemble de test. Par ailleurs, nous discutons ces résultats en justifiant les choix de nos approches afin de valider la meilleure approche. A la fin, Nous présentons une synthèse de notre travail dans laquelle nous faisons sortir tout l’intérêt de nos propositions sur la base des résultats expérimentaux ainsi obtenus.

3.2 Description de l’architecture adaptée

MobileNet est un réseau de neuronal artificiel réalisé par Google, spécialement conçue pour fonctionner sur les appareils mobiles et les systèmes embarqués avec des contraintes de ressource mémoire, calcul et énergétique. Elle basé sur une architecture simplifiée qui utilise des convolutions séparables en profondeur (depthwise separable convolutions) au lieu des convolutions standards pour construire des réseaux neuronaux profonds légers.

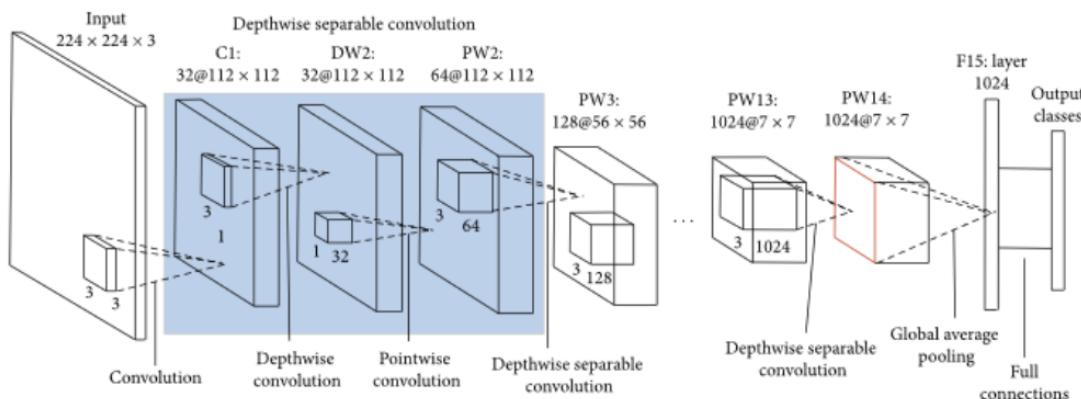


Figure 3.1 : Représentation de l’architecture MobileNet [37].

Les convolutions séparables en profondeur sont une forme de convolutions factorisées, où une convolution standard est factorisée en une convolution en profondeur et une convolution ponctuelle. Cette factorisation réduit drastiquement le temps de calcul et la taille du modèle.[20]

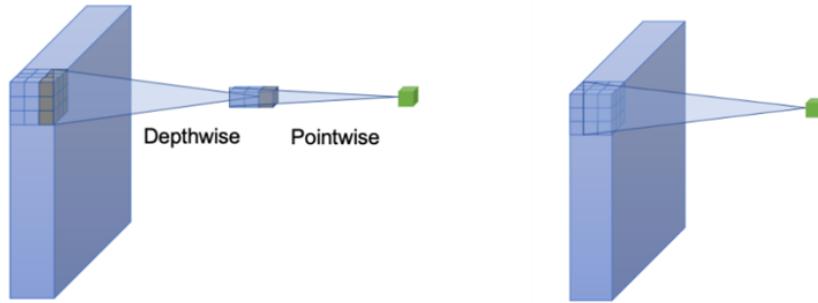


Figure 3.2 : Illustration représente à gauche une convolution séparable en profondeur , et à droite une convolution standard. [8].

- **Convolution en profondeur** : Consiste à appliquer un filtre sur chaque canal, contrairement à la convolution classique qui applique un filtre sur l’ensemble des canaux.
- **Convolution ponctuelle** : Consiste à combiner les sorties de la Depthwise Convolution, elle est aussi appelée convolution 1×1 .

D’après la figure 3.1, L’architecture MobileNet est composée de 13 couches de convolution séparable en profondeur sauf la première couche qu’est une convolution simple. D’après la figure 3.3 une couche de convolution séparable en profondeur contient une sorte de “moyenne pondérée” sur chaque canal, puis une normalisation de batch, ensuite le Relu est appliqué et finalement une convolution de 1×1 est utilisée pour obtenir une map sur laquelle on applique encore une normalisation de batch et un Relu. À la fin, elle se termine par une couche des neurones qui représente chaque classe de prédiction.

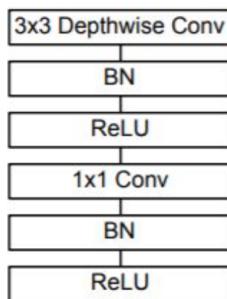


Figure 3.3 : Illustration représente les couches d’un bloc de convolution séparable en profondeur. [37].

3.3 Description détaillée de l’approche proposée

Nous avons proposé trois (3) variantes pour répondre au problème de classification du cancer de la peau. La première consiste à entraîner l’architecture MobileNet à partir de zéro.

La deuxième consiste à appliquer l’apprentissage par transfert sur l’architecture MobileNet. Nous allons utiliser l’architecture MobileNet pré entraînée sur l’ensemble de données ImageNet qui est composée de plus d’un million images naturelles réparties sur 1000 classes utilisées pour la détection des objets dans la nature. D’abord, nous avons remplacés les cinq (5) dernières couches du model par un couche Dense qui contient sept classes pour correspondre à notre problématique. Ensuite, nous avons gardé les poids de toutes les couches du model pré entraîné sauf les 23 dernières couches qui seront entraînés sur notre dataset (Fine tuning). Car les premières couches contiennent des variables très génériques (e.g. détection des angles ou des couleurs) et donc pertinentes pour de nombreuses tâches différentes.

Nous avons ensuite affiné l’apprentissage de cette architecture de bout en bout sur notre ensemble de données. Autrement dit, nous allons optimiser les valeurs des paramètres du modèle récupérés d’apprentissage sur l’ensemble des données ImageNet avec les nouvelles valeurs calculées d’après l’ensemble de données ISIC. Alors, les images sont redimensionnées vers 224×224 pour correspondre à l’entrée de l’architecture. L’image parcourt ensuite les différentes couches de MobileNet de sorte à permettre au modèle de calculer l’erreur des paramètres. À la fin, la rétropropagation va permettre de corriger tous les paramètres du l’architecture (partie extracteur des caractéristiques et la partie de classifieur) pour arriver à converger vers les meilleures performances.

Et la dernière, consiste à appliquer l’architecture MobileNet comme un extracteur de caractéristiques pré entraînée sur l’ensemble de données ImageNet. Ensuite pour la classification, nous avons choisi deux différents classificateurs : des forêts des arbres décisionnels (Random forest) et K plus proches voisins (KNN) pour démontrer l’effet des classificateurs sur les résultats obtenus.

3.4 Évaluations expérimentales

Dans cette section, nous proposons d’évaluer nos approches proposées pour la classification du cancer de la peau. Cette évaluation va nous permettre de mesurer le degré de validité de nos hypothèses de départ et donc de répondre aux questions posées dans la problématique de ce travail.

3.4.1 Métriques de performance

Nous définissons les critères de performance utilisés pour l’évaluation des approches de classification des images qui sont très simples à lire et à comprendre. Elles permettent de visualiser très rapidement les données et les statistiques afin d’analyser les performances

d'un modèle et d'identifier les tendances qui peuvent aider à modifier les paramètres. Cependant, nous citons quelques définitions utiles avant de définir ces métriques :

- **Vrai positive (VP)** : les cas où la prédiction est positive, et la valeur réelle est effectivement positive.
- **Vrai négative (VN)** : les cas où la prédiction est négative, et la valeur réelle est effectivement négative.
- **Faux positive (FP)** : les cas où la prédiction est positive, mais la valeur réelle est négative.
- **Faux négative (FN)** : les cas où la prédiction est négative, mais la valeur réelle est positive.
- **Accuracy (AC)** : Le nombre des prédictions correctes par rapport à toutes les prédictions, elle représente le taux de succès de modèle c'est à dire le nombre des images bien classées sur la totalité des images.

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN} \quad (3.1)$$

- **Précision (sensibilité)** : : Le nombre des cas positifs que le classificateur a correctement prédit, sur tous les cas positifs dans les données.

$$Precision = \frac{VP}{VP + FP} \quad (3.2)$$

- **Rappel** : Le nombre des prédictions positives correctes, elle représente le taux de succès de détecter que l'image appartient à la classe correcte.

$$Rappel = \frac{VP}{VP + FN} \quad (3.3)$$

3.4.2 Ensemble des données ISIC

Donnons un aperçu sur l'ensemble des données utilisé pour entraîner nos modèles. Il s'agit du dataset ISIC qui représente l'ensemble des données des images dermatoscopiques du cancer de la peau.

L'organisme ISIC a mis au point puis a étendu une archive open source accessible au public d'images de peau pour tester et valider les approches proposées. Cette archive sert de ressource publique d'images pour l'entraînement, la validation et le test de modèles du diagnostic automatisé comme celui que nous proposons dans ce projet.

Au cours des dernières années, les grands défis de l'ISIC ont été organisés pour des fins d'analyse des lésions cutanées jusqu'à la détection du cancer de la peau, présentant des problèmes de segmentation des lésions, de détection des schémas de diagnostics cliniques et de classification des lésions.

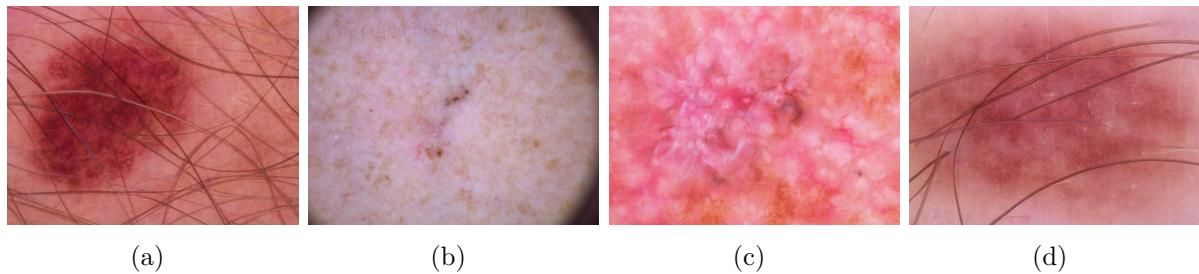


Figure 3.4 : Illustration des difficultés des images d'ensemble de données ISIC

Ces défis ont donné lieu à des nouvelles conclusions des nombreuses publications. Il ont été acceptés comme norme de référence de facto par d'autres groupes de recherche. Il y a eu depuis plusieurs améliorations de cette archive, dont ISIC Challenge2016, ISIC Challenge 2017, ISIC Challenge 2018 et ISIC Challenge 2019. Pour chaque compétition, ISIC augmente la taille de l'ensemble de données pour donner plus d'informations à utiliser pour l'apprentissage. Cela permet d'une part d'augmenter le nombre de données d'apprentissage et d'autre part, de rajouter de nouveaux cas complexes à l'ensemble de test.

Malgré la richesse de cet ensemble de données, il contient des problèmes majeurs pour être exploités dans l'apprentissage profond. D'abord, les images n'ont pas les mêmes dimensions, alors que l'entrée de l'architecture MobileNet est 224x224. Donc, nous sommes obligés de retransformer ces images pour être compatibles avec cette architecture. Cette transformation va créer la première limite qui se résume dans la déformation de tumeur qui est un critère très important pour la détermination du type de maladie.

La deuxième limite est la difficulté de détecter la classe d'une image donnée même pour les techniques standards ou bien pour un médecin expert. Cette difficulté est due à la disparition de la maladie sur toute l'image, à l'apparition des cheveux de certaines images ou bien à la similarité des tumeurs pour des classes différentes, où deux images de classes différentes ont la même forme de tumeur ou bien la même couleur (voir Fig 3). Tous ces critères vont être détectés comme bruit pour le modèle d'apprentissage profond et ils représenteront des informations qui vont baisser les paramètres de modèle.

Une dernière limite consiste en un déséquilibre de la quantité des données entre les classes. Par exemple, le nombre des images de la classe df représente que 2% de la totalité des images. Cette limite va provoquer une convergence artificielle du modèle vers cette classe. C'est-à-dire pour le modèle, cette classe représente la minorité et elle n'influencera pas dans la phase d'apprentissage (deviens presque négligeable) et alors il sera difficile de la détecter.

Nous avons trouvé d'autres ensembles des données qui traitent la classification de cancer de la peau. Nous citons notamment Ph2 dataset [4], ensemble de 200 images dermatoscopiques fournis par le service de l'hôpital Pedro Hispano et Dermofit dataset [05, 06], un ensemble de 1300 images divisé en dix(10) types du cancer de la peau, Dermquest dataset [05] et Dermis dataset [06]...

Le problème est la taille de ces ensembles des données qui ne donne pas des résultats

Chapitre 3. Approche adaptée pour la classification d'images médicales

robustes dans l'apprentissage profond. En outre, le dataset ISIC rassemble la majorité de ces ensembles pour fournir un large ensemble de données.

Pour notre travail, nous avons utilisés le jeu de données Ham10000 comme un ensemble d'entraînement. Ce dernier a été fourni aux participants du challenge ISIC 2018 pour résoudre le problème de la classification du cancer de la peau.

Description de l'ensemble de données

Les images dermatoscopiques du HAM10000 ont été collectées sur une période de 20 ans sur deux sites différents, le département de dermatologie de l'université médicale de Vienne, en Autriche, et le cabinet de Cliff Rosendahl, spécialisé dans le cancer de la peau, à Queensland, en Australie. Le site australien a stocké les images et les méta-données dans des fichiers PowerPoint et des bases de données Excel. Le site autrichien a commencé à collecter des images avant et a stocké les images et les métadonnées dans différents formats au cours de différentes périodes. [50] Ham10000 est composé de 10015 images de 7 classes différentes (Actinic keratoses and intraepithelial carcinoma (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv) and vascular lesions (vasc)). Nous avons choisi cet ensemble des données pour différentes raisons. Il fournit un nombre d'exemples raisonnable (par rapport aux travaux cités dans l'état de l'art) à exploiter pour l'apprentissage profond, De plus le HAM10000 englobe la majorité des types de cancer qui existent non seulement si le cancer est bénigne ou malin cela peut aider les dermatologues à donner le traitement exacte pour chaque lésion.

Le tableau 3.1 représente une comparaison entre quelques ensembles de données selon des critères différents :

TAB. 3.1 : Tableau représente les ensembles de données d'images dermatoscopiques accessibles au public [50].

Dataset	License	Total images	Pathologic verification (%)	akiec	bcc	bkl	df	mel	nv	vasc
PH2	Research&Education ^a	200	20.5%	-	-	-	-	40	160	-
Atlas	No license	1024	unknown	5	42	70	20	275	582	30
ISIC 2017 ^b	CC-0	13786	26.3%	2	33	575	7	1019	11861	15
Rosendahl	CC BY-NC 4.0	2259	100%	295	296	490	30	342	803	3
ViDIR Legacy	CC BY-NC 4.0	439	100%	0	5	10	4	67	350	3
ViDIR Current	CC BY-NC 4.0	3363	77.1%	32	211	475	51	680	1832	82
ViDIR MoleMax	CC BY-NC 4.0	3954	1.2%	0	2	124	30	24	3720	54
HAM10000	CC BY-NC 4.0	10015	53.3%	327	514	1099	115	1113	6705	142

3.4.3 Résultats et discussion

Dans cette section, nous allons présenter les résultats de nos approches proposées pour la classification de cancer de la peau. Nous allons ensuite discuter les résultats obtenus. A la fin, nous donnerons des arguments en discutant les performances de chaque approche et nous allons mettre l’accent sur la meilleure approche. Pour information, nous avons affecté aux hyper-paramètres d’apprentissage les valeurs suivantes : 30 epochs, 10 images par batch et taux d’apprentissage égal à 0.01. Nous avons lancé le processus d’apprentissage sur une machine fournis par Google Colab pro avec un GPU Tesla P100-PCIE-16GB de 25 Go de RAM.

Pour la première approche qui consiste à entraîner l’architecture MobileNet de bout en bout pour la classification de cancer de la peau. Nous avons obtenu les résultats présentés dans la Table 3.2

TAB. 3.2 : Résultats obtenus par la première approche.

	AC	SE	RE	akiec	SE	RE
MobileNet	87.95	86	87	bcc	0.68	0.77
				blk	0.64	0.40
				df	0.25	0.17
				mel	0.40	0.56
				nv	0.93	0.96
				vasc	0.92	1.00
				avg	0.86	0.87

TAB. 3.3 : La matrice de confusion de la première approche

	akiec	bcc	blk	df	mel	nv	vasc
akiec	6	5	7	1	5	2	0
bcc	3	23	1	0	2	1	0
blk	0	0	30	0	13	32	0
df	0	0	0	1	0	5	0
mel	0	1	3	0	22	13	0
nv	3	5	6	2	13	721	1
vasc	0	0	0	0	0	0	11

Le premier constat est dû la difficulté de produire des résultats mieux que ceux de ZA et al [4]. qui ont entrainé une architecture de bout en bout sur le même ensemble de données que nous. Nous estimons que cela peut être dû à l’initialisation aléatoire des paramètres du modèle et aux différentes conditions d’apprentissage. Un autre constat consiste à la performance du modèle pour la bonne classification de chaque image. Nous remarquons que notre modèle a une précision moyenne de 86%, mais ce n’est pas vraiment le cas pour chaque classe (NV : 93%, DF : 25%) Ce résultat est plutôt faible pour une tâche dans un domaine aussi critique. Nous pensons que cette faiblesse est due au nombre de

données d’apprentissage pour chaque classe, nous avons déjà mentionné que l’ensemble de données est fortement déséquilibré, avec 6705 pour la classe NV et seulement 115 images pour la classe DF.

Pour améliorer les résultats précédents, nous avons proposé d’utiliser l’apprentissage par transfert sur l’ensemble des données ImageNet (détailé dans la partie section précédente). Nous avons obtenu les résultats présentés dans la Table 3.4.

TAB. 3.4 : Résultats obtenus par la deuxième approche.

	AC	SE	RE
MobileNet	91	89	90

	SE	RE
akiec	0.62	0.62
bcc	0.77	0.77
bkl	0.69	0.65
df	0.75	0.50
mel	0.55	0.41
nv	0.95	0.97
vasc	0.91	0.97
avg	0.89	0.90

TAB. 3.5 : La matrice de confusion de la deuxième approche

	akiec	bcc	bkl	df	mel	nv	vasc
akiec	16	2	5	0	2	1	0
bcc	1	23	3	0	1	2	0
bkl	3	1	49	0	5	17	0
df	0	0	0	3	0	3	0
mel	2	1	4	0	16	16	0
nv	4	3	10	1	5	727	1
vasc	0	0	0	0	0	1	10

Cette approche a permis d’avoir de meilleures performances avec une précision moyenne de 89%. Aussi nous avons remarqué une augmentation dans la précision de chaque classe, par exemple la précision de la classe DF a augmenté de 25% à 75%. Selon les résultats représentés dans la table 3.4, nous avons remarqué une amélioration dans la détection des classes. Cette amélioration est principalement due à l’initialisation des paramètres transférés du modèle pré-entraîné sur un grand ensemble des données (ImageNet). Cela permet de rendre notre modèle plus robuste et il offre de meilleurs résultats.

Pour la troisième approche ”Feature extraction”. Nous avons atteint Les performances présentés dans la Table 3.6

TAB. 3.6 : Résultats obtenus par la troisième approche.

	AC	SE	RE
MobileNet	84	79	83
akiec	0.64	0.63	
bcc	0.75	0.75	
blk	0.61	0.63	
df	0.68	0.55	
mel	0.53	0.49	
nv	0.91	0.91	
vasc	0.88	0.91	
avg	0.79	0.83	

D’après les résultats atteint de chaque approche, nous constatons que l’approche d’apprentissage par transfert a eu les meilleures performances par rapport aux autres, et pour savoir si notre modèle est robuste, on l’a testé sur différentes ensembles de données, et on a obtenu de bons résultats représentés dans le tableau suivant :

TAB. 3.7 : Tableau représentant les performances obtenues pour chaque base de données

	Accuracy
PH2	79%
DB1000	87%
HAM10000	91%

Durant la réalisation de notre master, Nous avons testé d’autres architectures. Les résultats obtenus sont présentés dans le tableau suivant :

TAB. 3.8 : Tableau représentant les performances obtenues pour chaque architecture

	Accuracy
VGG16	82%
ResNet50	96%
MobileNet	91%

3.4.4 Synthèse

Les algorithmes d’apprentissage profond peuvent être utilisés dans les différents domaines de l’imagerie médicale en particulier la classification des images médicales. Ces algorithmes ont été exploités pour aider le diagnostic des maladies. Nous citons en particulier la tâche de classification des images potentielles de cancer de la peau, qui représente la problématique que nous nous sommes proposés d’étudier. De plus, les techniques d’apprentissage profond ont obtenu de bonnes performances dans ce domaine. Pour certaines problématiques, les modèles produits dépassent les capacités des experts. Ces performances vont permettre d’accélérer le diagnostic des maladies, si elles existent, et donc ouvre la porte pour plusieurs opportunités anticipant l’aggravation de certaines maladies.

Plusieurs propositions qui traitent de la thématique de classification dans le domaine médical ont été conçues à base des réseaux de neurones de convolutions (CNN). Les réseaux de neurones de convolution ont été exploités dans la majorité des travaux qui traitent de la classification des cancers de la peau. Ces architectures ont montré des performances prometteuses pour la classification des images et ont été largement introduites dans le domaine médical.

Les performances de ces réseaux de neurones dépendent de la quantité d’information utilisée en entraînement. Ce critère représente un problème majeur dans le domaine médical où les images médicales ont besoin de temps pour les collecter et d’expérience pour les annoter afin d’utiliser pour la classification. Par contre, plusieurs sociétés ont commencé à fournir des ensembles des données d’imagerie médicale et organiser des compétitions afin de réaliser des tâches précises dans ce domaine. Nous avons cité la compétition ISIC qui traite la classification de cancer de la peau où les organisateurs ont permis de collecter une archive de plus de 10 000 images dermatoscopiques réparties sur sept types de cancer de la peau.

Plusieurs techniques ont été adoptées pour répondre au problème de manque de données en apprentissage profond, en particulier l’apprentissage par transfert. Cette technique consiste à transférer l’apprentissage d’un modèle entraîné sur un large ensemble de données pour l’utiliser sur un autre ensemble restreint sans autant être du même domaine. Le transfert peut être affiné sur tous les paramètres du modèle ou pour un nombre précis des couches et faire entraîner le reste, ou bien transférer l’apprentissage pour la partie d’extraction des caractéristiques et entraîner le classifieur seulement.

Cette méthode était utilisée pour plusieurs travaux pour la détection de cancer de la peau, car elle permet d’améliorer les performances des modèles et accélérer la phase d’entraînement. Cette amélioration est démontrée par notre deuxième approche où nous avons constaté que l’application de cette technique a surpassé les performances de modèle de la première approche entraînée de bout en bout.

3.5 Conclusion

Dans ce chapitre, nous avons expliqué notre contribution proposée à base des techniques d’apprentissage profond au premier lieu. Notre contribution répond à la problématique de classification du cancer de la peau. Nous avons adopté l’architecture MobileNet pré-entraînée sur l’ensemble de données ImageNet puis l’affiner sur l’ensemble des images dermatoscopiques de la peau ISIC.

Au deuxième lieu, Nous avons évalué notre contribution en commençant par définir les métriques utilisées afin d’évaluer les performances des modèles pour la classification des images. Ensuite, nous avons donné un aperçu sur l’ensemble des données utilisé proposé par la compétition ISIC 2018. Nous avons signalé les difficultés de cet ensemble et nous avons cité d’autres ensembles des données. Après nous avons présenté nos résultats obtenus par chaque approche, nous avons aussi discuté ces résultats en donnant des explications.

Notre meilleure approche consiste appliquer la technique de transfert learning sur

Chapitre 3. Approche adaptée pour la classification d'images médicales

l'architecture MobilNet. Cette approche a réussi d'améliorer les performances de notre modèle et d'avoir une bonne précision de 91 %.

Conclusion et perspectives

Conclusion générale

Le travail réalisé dans le cadre de ce Master s'inscrit dans le domaine d'application des approches d'apprentissage profond pour la classification dans le domaine médical. Notre travail consistait principalement à appliquer des approches d'apprentissage profond pour la classification des images de cancer de la peau. Nous avons effectué, en premier lieu, une synthèse des travaux qui traitent de la classification à base d'apprentissage profond dans le domaine médical. Nous avons ensuite proposé une taxonomie selon l'architecture de l'état de l'art des travaux sur la classification du cancer de la peau. Après nous avons cité également des travaux réalisés concernant les approches définies. Nous avons mis l'accent sur le rôle d'apprentissage par transfert dans le domaine médical et en particulier pour la classification du cancer de la peau. À la suite de cela, nous avons présenté nos contributions sous forme d'approche de classification. Nous avons d'abord présenté et expliqué les architectures que nous avons utilisées dans notre contribution en l'occurrence MobileNet avec ses différentes variantes. À la suite de cela nous avons détaillé nos approches proposées avec des illustrations des architectures.

Ces propositions ont fait l'objet d'évaluations expérimentales. Nous avons d'abord défini les métriques d'évaluation utilisées pour la classification des images. Nous avons également décrit l'ensemble des données utilisé. Après, nous avons ensuite présenté les résultats de nos propositions et les avons discutés suivant différents aspects. À la fin, nous avons clôturé par une synthèse ce qu'il faut retenir de nos travaux.

Notre travail peut être amélioré davantage, et cela de plusieurs façons. Nous donnons ci-après quelques pistes :

- Introduire les mécanismes d'attention sur notre architecture pour se focaliser sur la zone la plus importante de l'image et avoir plus de précision.
- Entrainer notre model sur un grand ensemble de données .

Bibliographie

- [2] Philippe ABIMELEC. “Cancer de la peau (Carcinomes)”. In : ().
- [3] Marwan Ali ALBAHAR. “Skin lesion classification using convolutional neural network with novel regularizer”. In : *IEEE Access* 7 (2019), p. 38306-38313.
- [4] Md Zahangir ALOM et al. “Skin cancer segmentation and classification with NABLA-N and inception recurrent residual convolutional networks”. In : *arXiv preprint arXiv :1904.11126* (2019).
- [5] Rehan ASHRAF et al. “Region-of-Interest based transfer learning assisted framework for skin cancer detection”. In : *IEEE Access* 8 (2020), p. 147858-147871.
- [6] Yaniv BAR et al. “Chest pathology detection using deep learning with non-medical training”. In : *2015 IEEE 12th international symposium on biomedical imaging (ISBI)*. IEEE. 2015, p. 294-297.
- [8] Francois CHOLLET. “Xception : Deep Learning With Depthwise Separable Convolutions”. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Juil. 2017.
- [14] R DERAEMAECKER. “Les carcinomes cutanés”. In : *Thérapeutique du cancer*. Springer, 2011, p. 685-708.
- [15] Thomas L DIEPGEN et V MAHLER. “The epidemiology of skin cancer”. In : *British Journal of Dermatology* 146 (2002), p. 1-6.
- [16] Ahmed ECH-CHERIF, Mohammed MISBHAUDDIN et Mohammed ECH-CHERIF. “Deep Neural Network based mobile dermoscopy application for triaging skin cancer detection”. In : *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*. IEEE. 2019, p. 1-6.
- [17] Andre ESTEVA et al. “Dermatologist-level classification of skin cancer with deep neural networks”. In : *nature* 542.7639 (2017), p. 115-118.
- [18] Fredrik GEORGSSON. “Tor-BjörnHolmström”. Thèse de doct. Master Thesis in Computing Science “A Survey et Evaluation of Features for ...”, 2005.
- [20] Andrew G HOWARD et al. “Mobilennets : Efficient convolutional neural networks for mobile vision applications”. In : *arXiv preprint arXiv :1704.04861* (2017).
- [27] Adria Romero LOPEZ et al. “Skin lesion classification from dermoscopic images using deep learning techniques”. In : *2017 13th IASTED international conference on biomedical engineering (BioMed)*. IEEE. 2017, p. 49-54.
- [28] Robyn LUCAS et al. *Solar ultraviolet radiation : global burden of disease from solar ultraviolet radiation*. World Health Organization, 2006.

Bibliographie

- [29] Lucas Bezerra MAIA et al. "Evaluation of melanoma diagnosis using deep features". In : *2018 25th international conference on systems, signals and image processing (IWSSIP)*. IEEE. 2018, p. 1-4.
- [30] Afonso MENEGOLA et al. "Knowledge transfer for melanoma screening with deep learning". In : *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE. 2017, p. 297-300.
- [31] Ebrahim NASR-ESFAHANI et al. "Melanoma detection by analysis of clinical images using convolutional neural network". In : *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2016, p. 1373-1376.
- [32] Maxime OQUAB et al. "Learning and transferring mid-level image representations using convolutional neural networks". In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, p. 1717-1724.
- [33] Adeline PERRINAUD. "Carcinomes épidermiques (ou spinocellulaires)". In : *La Presse Médicale* 37.10 (2008), p. 1485-1489.
- [34] Tri-Cong PHAM et al. "Deep CNN and data augmentation for skin lesion classification". In : *Asian Conference on Intelligent Information and Database Systems*. Springer. 2018, p. 573-582.
- [35] PHUNG et RHEE. "A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets". In : *Applied Sciences* 9 (oct. 2019), p. 4500.
- [36] Victor POMPONIU, Hossein NEJATI et N-M CHEUNG. "Deepmole : Deep neural networks for skin mole lesion classification". In : *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2016, p. 2623-2627.
- [39] Noortaz REZAOANA, Mohammad Shahadat HOSSAIN et Karl ANDERSSON. "Detection and Classification of Skin Cancer by Using a Parallel CNN Model". In : *2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*. IEEE. 2020, p. 380-386.
- [40] Adria ROMERO LOPEZ et al. "Skin lesion classification from dermoscopic images using deep learning techniques". In : *2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*. 2017, p. 49-54.
- [41] Venura SAMARASINGHE et Vishal MADAN. "Nonmelanoma skin cancer". In : *Journal of cutaneous and aesthetic surgery* 5.1 (2012), p. 3.
- [42] Mark SANDLER et al. "Mobilenetv2: Inverted residuals and linear bottlenecks". In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, p. 4510-4520.
- [44] Mariam A SHEHA, Mai S MABROUK et Amr SHARAWY. "Automatic detection of melanoma skin cancer using texture analysis". In : *International Journal of Computer Applications* 42.20 (2012), p. 22-26.
- [45] Hoo-Chang SHIN et al. "Deep convolutional neural networks for computer-aided detection : CNN architectures, dataset characteristics and transfer learning". In : *IEEE transactions on medical imaging* 35.5 (2016), p. 1285-1298.

Bibliographie

- [46] Karen SIMONYAN et Andrew ZISSERMAN. “Very deep convolutional networks for large-scale image recognition”. In : *arXiv preprint arXiv :1409.1556* (2014).
- [48] Christian SZEGEDY et al. “Inception-v4, inception-resnet and the impact of residual connections on learning”. In : *Thirty-first AAAI conference on artificial intelligence*. 2017.
- [50] Philipp TSCHANDL, Cliff ROSENDAHL et Harald KITTLER. “The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions”. In : *Scientific data* 5.1 (2018), p. 1-9.
- [61] Jason YOSINSKI et al. “How transferable are features in deep neural networks ?” In : *arXiv preprint arXiv :1411.1792* (2014).
- [62] Kai ZHAO, Yan ZHOU et Xin CHEN. “Object Detection : Training From Scratch”. In : *IEEE Access* 8 (2020), p. 157520-157529.

Webographie

- [1] *A Layman's Guide to Deep Convolutional Neural Networks / by Jojo John Moolayil / Towards Data Science.* fr. URL : <https://towardsdatascience.com/a-laymans-guide-to-deep-convolutional-neural-networks-7e937628605f/>.
- [7] *Caractéristiques des images.* fr. URL : <https://www.rncan.gc.ca/cartes-outils-publications/imagerie-satellite-photos-aer/tutoriels-sur-la-teledetection/introduction/caracteristiques-des-images/14642/>.
- [9] *CIRCUIT. Using Route Optimization To Cut Delivery Costs.* en. URL : <https://www.photowizi.com/enlever-bruit-photo/> (visité le 29/05/2013).
- [10] Philippe CIUCIU. *Introduction à L'imagerie Biomédicale.* fr. URL : <https://slideplayer.fr/slide/1645867/>.
- [11] *Comment se passe une échographie pelvienne ? : tout savoir - Top Santé.* fr. URL : <https://www.topsante.com/medecine/examens-medicaux/comment-se-passe-une-echographie-pelvienne/>.
- [12] *Deep learning : les réseaux de neurones convolutifs pour la classification d'images médicales.* fr. URL : <https://www.imaios.com/fr/Societe/blog/Classification-des-images-medicales-comprendre-le-reseau-de-neurones-convolutifs-CNN/>.
- [13] *Deep learning : les réseaux de neurones convolutifs pour la classification d'images médicales.* fr. URL : <https://www.imaios.com/fr/Societe/blog/Classification-des-images-medicales-comprendre-le-reseau-de-neurones-convolutifs-CNN/>.
- [19] HERVÉ. *Le mythe de la résolution de 72 dpi pour les images web.* fr. URL : <https://www.luzphotos.com/logiciels/resolution-image-web/>.
- [21] *Imagerie médicale.* fr. URL : [https://www.doctissimo.fr/html/sante/imagerie/imagerie_sommaire.htm/](https://www.doctissimo.fr/html/sante/imagerie/imagerie_sommaire.htm).
- [22] *IRM / Centre de radiologie Sud Rhône Imagerie.* fr. URL : <https://www.sud-rhone-imagerie.fr/examens/irm/>.
- [23] *La cellule de Langerhans.* fr. URL : <https://biologiedelapeau.fr/spip.php?article11/>.
- [24] *La radiographie.* fr. URL : https://fr.wikipedia.org/wiki/R%C3%A9seau_de_neurones_artificiels.
- [25] *La structure de la peau - Noviderm Belgique.* fr. URL : <https://www.noviderm.com/fr-fr/information/13563/la-structure-de-la-peau/>.

- [26] *Le rôle et la structure de la peau.* fr. URL : https://www.doctissimo.fr/html/dossiers/peau_boutons/sa_3884_peau.htm.
- [37] Abhijeet PUJARA. *Image Classification With MobileNet.* en. URL : <https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470/>.
- [38] *Réseau de neurones artificiels.* fr. URL : <http://rayons-x-imagerie-medicale.emonsite.com/pages/cat-2/la-radiographie.html>.
- [43] *Scanner médical au cabinet de radiologie d'Obernai en Alsace.* fr. URL : <https://www.radiologie-piemont.com/scanner/>.
- [47] *Structures et rôles de la peau - De quoi la peau est-elle composée ?* fr. URL : <https://sante.lefigaro.fr/mieux-etre/beaute/structures-roles-peau/quoi-peau-est-elle-composee>.
- [49] *TensorFlow - Apprentissage de Perceptron multicouche.* fr. URL : <https://www.hebergementwebs.com/tutoriel-tensorflow/tensorflow-apprentissage-de-perceptron-multicouche/>.
- [51] *Using Route Optimization To Cut Delivery Costs.* en. URL : <https://www.pfl-cepia.inra.fr/index.php?page=tutoImg-images-couleurs/>.
- [52] *Using Route Optimization To Cut Delivery Costs.* en. URL : <https://www.pfl-cepia.inra.fr/index.php?page=tutoImg-niveaux-de-gris/>.
- [53] *Using Route Optimization To Cut Delivery Costs.* URL : <https://www.futura-sciences.com/sante/questions-reponses/cancer-peau-difference-carcinome-melanome-14892/>.
- [54] *Using Route Optimization To Cut Delivery Costs.* URL : <https://sante.journaldesfemmes.fr/fiches-maladies/2614543-carcinome-definition-symptomes-causes-traitements/>.
- [55] *Using Route Optimization To Cut Delivery Costs.* URL : <https://www.msdmanuals.com/fr/professional/troubles-dermatologiques/cancers-cutan%C3%A9es/carcinome-basocellulaire>.
- [56] *Using Route Optimization To Cut Delivery Costs.* URL : <https://nuage1962.wordpress.com/tag/carcinome-epidermoide/>.
- [57] *Using Route Optimization To Cut Delivery Costs.* URL : <https://fr.wikipedia.org/wiki/M%C3%A9lanome>.
- [58] *Using Route Optimization To Cut Delivery Costs.* en. URL : <https://www.image-net.org/>.
- [59] *Using Route Optimization To Cut Delivery Costs.* en. URL : <https://isicarchive.com/>.
- [60] *Using Route Optimization To Cut Delivery Costs.* en. URL : <https://neurohive.io/en/popular-networks/vgg16/> (visité le 20/11/2018).