

N° d'ordre :
N° de série :



RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



UNIVERSITE ECHAHID HAMMA LAKHDAR - EL OUED
FACULTÉ DES SCIENCES EXACTES
Département D'Informatique

Mémoire de Fin D'étude
Présenté pour l'obtention du Diplôme de

MASTER ACADEMIQUE

Domaine : **Mathématique et Informatique**
Filière : **Informatique**
Spécialité : **Systemes Distribués et Intelligence Artificielle**

Présenté par :

- **Mouchira Bouziane**
- **Sabrine Berouba**

Thème

**Hybridation de plusieurs caractéristiques de
Deep Learning pour l'amélioration de la
classification du cancer de la peau.**

Soutenu le 20-06- 2021 Devant le jury :

M.	Ismail kertiou	MCA	Président
M.	Mohammad Anouar naoui	MAA	Rapporteur
Dr.	Abdel Kamel Ben Ali	MCA	Encadreur

Année Universitaire : 2020/2021

Remerciements

Nous tenons à exprimer nos sincères remerciements à Dieu Tout-Puissant pour ses bonnes actions et sa gratitude pour nous avoir donné le courage et la patience pour accomplir cet humble travail.

Nous tenons à exprimer nos sincères remerciements et notre gratitude à l'estimé superviseur **Dr. "Abdellamel Ben Ali"** pour son aide, ses précieux conseils, sa gentillesse, ses encouragements, sa présence avec nous et sa confiance qui nous ont permis de continuer à poursuivre ce travail.

Naturellement, nous remercions nos familles pour leur aide, leur générosité et leur soutien moral qui a été pour nous une source de courage et de confiance.

Nous remercions également chaleureusement les membres estimés du jury qui ont accepté l'évaluation.

Ce travail, merci à tous les professeurs et amis

Pour leur soutien et leur patience enfin, un grand merci à tous ceux qui, en un mot, m'ont donné

La force de continuer à travailler vers mes objectifs.

Dédicaces

Nous dédions notre travail à : Les deux êtres les plus chères au monde, les bougies qui m'ont toujours guidé vers le bon chemin : Ma mère, la lumière de notre vie, qui a tout fait pour notre réussite et notre bonheur. Mon père, à lui, nous devons offrir tout le respect et l'amour pour son soutien et sa tendresse. A nos sœurs pour leur aide et courage
A nos frères, sœurs À mes amies,
A tous les membres de la famille et à toutes les amies.

Mouchira Bouziane

Dédicaces

Nous dédions notre travail à : Les deux êtres les plus chères au monde, les bougies qui m'ont toujours guidé vers le bon chemin : Ma mère, la lumière de notre vie, qui a tout fait pour notre réussite et notre bonheur. Mon père, à lui, nous devons offrir tout le respect et l'amour pour son soutien et sa tendresse. A nos sœurs pour leur aide et courage

A nos frères, sœurs À mes amies,

A tous les membres de la famille et à toutes les amies.

Sabrine Berouba

Résumé

Le cancer de la peau est une forme courante de cancer, et une détection précoce augmente le taux de survie.

Dans cette note, nous avons créé un système qui nous permet de classer les images médicales du cancer de la peau (bénigne ou maligne).

L'objectif principal de ce sujet est de démontrer la capacité de la machine et du deep learning à classer les images médicales en déterminant les pourcentages de précision.

Cette idée était d'extraire les caractéristiques des images et de les alimenter dans plusieurs modèles CNN pré-entraînés à l'aide d'un apprentissage en profondeur, puis de les combiner sous la forme d'un rayon final et de les transformer pour l'apprentissage à l'aide de l'algorithme d'apprentissage automatique SVM qui aide à prédire la peau cancer avec une précision améliorée.

Et nous avons, à notre tour, incarné cette idée dans le domaine médical sur le cancer de la peau en particulier, car nous nous sommes appuyés sur les modèles CNN suivants (VGG 16 ; VGG 19 ; ResNet ; Inception V3) et nous avons obtenu des taux de précision proches et en fusionnant les quatre modèles ensemble, nous avons obtenu le pourcentage (91 %) comme taux de précision le plus élevé, ce qui est considéré comme un bon pourcentage. Et nous avons remarqué qu'il y a une amélioration des taux de précision lors de la fusion des modèles entre eux.

Mots clés : Apprentissage automatique, Apprentissage profond (ou *Deep learning*), Réseaux de neurones convolutifs (ou CNN), Caractéristiques hybrides, Classification, Cancer de la peau

Abstract

Skin cancer is a common form of cancer, and early detection increases the survival rate.

In this memo, we have created a system that allows us to classify medical images of skin cancer (benign or malignant).

The main objective of this subject is to demonstrate the capacity of the machine and of deep learning to classify medical images by determining the percentages of precision.

This idea was to extract the features of the images and feed them into several pre-trained CNN models using deep learning, then combine them as a final layer and transform them for learning using SVM machine learning algorithm which helps predict skin cancer with improved accuracy.

And we, in turn, embodied this idea in the medical field on skin cancer in particular, because we relied on the following CNN models (VGG 16; VGG 19; ResNet; Inception V3) and we got close accuracy rates and by merging the four models together we got the percentage (91%) as the highest accuracy rate, which is considered a good percentage. And we noticed that there is an improvement in accuracy rates when merging models together.

Keywords: Machine learning, Deep learning, Convolutional neural networks (or CNN), Hybrid characteristics, Classification, Skin cancer

الملخص

سرطان الجلد هو شكل شائع من السرطانات ، والاكتشاف المبكر عنه يزيد من معدل البقاء على قيد الحياة.

في هذه المذكرة قمنا بإنشاء نظام يمكننا من تصنيف الصور الطبية لسرطان الجلد (حميد او خبيث) .حيث يتمثل الهدف الاساسي من هذا الموضوع في ابراز قدرة التعلم الالي و العميق في تصنيف الصور الطبية عن طريق تحديد نسب الدقة.

و تمثلت هذه الفكرة في استخراج الميزات من الصور و ادخالها الى عدة نماذج الشبكات العصبية التلافيفية CNN مدربة مسبقا باستخدام التعلم العميق ثم دمجها على شكل شعاع نهائي و تحويلها للتعلم باستخدام خوارزمية التعلم الالي SVM التي تساعد في التنبؤ بسرطان الجلد بدقة محسنة.

و نحن بدورنا جسدنا هذه الفكرة في المجال الطبي على سرطان الجلد بشكل خاص, حيث اعتمدنا على نماذج ال CNN التالية (VGG 16 ;VGG 19 ; RasNet ; Inception V3) و حصلنا على نسب دقة متقاربة و عند دمج النماذج الاربعة معا حصلنا على النسبة (91%) كأعلى نسبة دقة وهي تعتبر نسبة جيدة. و لاحظنا ان هناك تحسن في نسب الدقة عند دمج النماذج مع بعضها البعض.

الكلمات المفتاحية: التعلم الآلي، التعلم العميق، الشبكات العصبية التلافيفية (أو CNN)، الخصائص الهجينة، التصنيف، سرطان الجلد.

Table des matières

Remerciements	i
Dédicaces	ii
Dédicaces	iii
Résumé	iv
Abstract	v
المخلص	vi
Table des matières	vii
Table des figures	x
Liste des tableaux	xi
Liste des acronymes	xii
Introduction générale	13
Contexte de recherche	1
Objectifs du travail mémoire.....	2
Organisation du mémoire	2
Chapitre 1	4
Extraction de caractéristiques et classification d'images	4
1.1. Introduction	4
1.2. Caractéristique de bas niveaux.....	4
1.2.1. Approches globales	4
1.2.2. Approches locales.....	8
1.3. Caractéristique de plus hauts niveaux	10
1.4. Types d'apprentissage automatique	10
1.4.1. Apprentissage supervisé	10
1.4.2. Apprentissage non supervisé	11
1.4.3. Classification	12
1.5. Méthodes de classification	13

1.5.1. K plus proches voisins	13
1.5.2. Naïf de Bayes	13
1.5.3. Arbre de décision.....	14
1.5.4. Foret aléatoire	15
1.5.5. SVM	16
1.6. Conclusion.....	18
Chapitre 2.....	19
Apprentissage en profondeur "Deep Learning"	19
2.1. Introduction	18
2.2. Apprentissage automatique (Machine Learning)	18
2.3. Apprentissage en profondeur (<i>Deep Learning</i>).....	18
2.4. Relation entre ML et DL	19
2.5. Domaines d’application de DL	19
2.6. Les différents types de modèles	20
2.6.1. Machines Boltzmann restreintes.....	20
2.6.2. Réseau de neurones "Neural Networks"	20
2.6.3. Réseaux de neurones convolutifs (CNN)	21
2.6.4. Paramètres CNN	23
2.7. Projet Image Net.....	25
2.8. Architecture de CNN.....	25
2.8.1. Quelques réseaux convolutifs célèbre	26
2.8.2. Scénarios de classification basée sur CNN.....	30
2.9. Conclusion.....	31
Chapitre 3.....	32
Classification d’images médicales par ML et DL	32
3.1 Introduction	32
3.2. Imagerie médicale	32
3.3. Apprentissage profond pour l’analyse d’images médicales	33
3.4. Apprentissage automatique pour aider au diagnostic médical du cancer de la peau .	34
3.4.1. Cancer de la peau.....	34
3.4.2. Photographie pour vérifier et détecter le cancer de peau.....	35
3.4.3. Détection et classification assistées par ordinateur (CADe/CADx) des photographique cancer de peau	36

3.4.4. Détection de tumeur assistée par ordinateur (CADe).....	37
3.4.5. Diagnostic des tumeurs assisté par ordinateur (CADx).....	37
3.5. Quelques travaux récents utilisant ML et/ou DL	38
3.6. Conclusion.....	40
Chapitre 4.....	41
Hybridation de caractéristiques de DL pour améliorer la classification du cancer de peau	41
4.1. Introduction	41
4.2. Problématique étudié.....	41
4.3. Data Set	42
4.4. Méthode proposée	43
4.4.1. Classification de maladies grâce à des méthodes d'apprentissage automatique	44
4.4.2. Représentations hybrides de caractéristiques de DL pour améliorer la classification de maladies	45
4.5. Présentation des outils.....	48
4.5.1. Logiciel.....	48
4.5.2. Matériel.....	49
4.5.3. Langages et outils de développement.....	50
4.6. Explication du code.....	50
4.7. Résultats et discussion.....	52
4.7.1. Analyse des résultats	52
4.7.2. Discussion des résultats	53
4.8. Conclusion.....	53
Conclusion générale	54
Références	56

Table des figures

Figure 01 : Schéma apprentissage supervise [17].	11
Figure 02 : Schéma apprentissage non supervisé [17].	12
Figure 03 : Exemple d'arbre décision [26].	15
Figure 04 : Exemple de forêts aléatoires.	16
Figure 05 : Le procédé du ML classique comparé à celui du DL [28].	19
Figure 06 : Schéma d'un neurone à gauche et représentation d'un neurone formel à droit [32].	21
Figure 07 : Exemple de couches de réseau neuronal convolutif expliquées.	22
Figure 08 : L'architecture du modèle VGG-16 [41].	27
Figure 09 : Architecture <i>Google-Net</i> [41].	27
Figure 10 : Exemples d'images dataset d'apprentissage du mélanome.	34
Figure 11 : Exemples d'images dataset d'apprentissage du CBC.	35
Figure 12 : Exemples d'images dataset d'apprentissage du CSC.	35
Figure 13 : Exemple de Data Set.	43
Figure 14 : Schéma général de notre méthode de classification du cancer de la peau.	44
Figure 15 : Vecteur de caractéristiques extrait de la 1ère et de la 2ième couches FC.	46
Figure 16 : Extraction de caractéristiques.	51
Figure 17 : Le travail de la SVM.	51

Liste des tableaux

Tableau 01 : Model de CNN [43].....	29
Tableau 02 : Résultats de certaines applications de Deep Learning dans l'analyse d'images médicales [46].....	33
Tableau 03 : Tailles des vecteurs de caractéristiques DL combinées.	47
Tableau 04 : Rapports de performances TS (%) pour les modèles CNNs.	53

Liste des acronymes

LBP: Local Binary Pattern.

GLCM: Gray-Level Co-occurrence Matrix

SVM: Support Vector Machine.

HOG: histogram of oriented gradients

SIFT: Scale Invariant Feature Transform

DSIFT: Dense Scale Invariant Feature Transform.

ReLU: Rectified Linear Unit

FV: Fisher Vector

VLAD: Vector of locally aggregated descriptors.

CAD: Computer Aided Design.

CADe: Computer Aided Detection.

CADx: Computer Aided Diagnosis.

SK learn : Scikit-learn.

Introduction générale

Les travaux de ce mémoire de Master s'intéressent à la classification et à la prédiction des types de cancer de la peau en utilisant des stratégies basées sur des réseaux de neurones à convolution.

Contexte de recherche

Ces dernières années, l'**imagerie médicale** est devenue très importante dans le diagnostic et le traitement des maladies. Avec l'enrichissement des bases de données d'images médicales, de nouvelles méthodes sont nécessaires pour mieux gérer cet énorme volume de données. En raison de la grande différence entre les données d'imagerie médicale et leur complexité, les tâches de l'imagerie médicale nécessitent un apprentissage, des exemples (formation) et c'est l'une des préoccupations de l'apprentissage automatique.

L'**apprentissage automatique** (ou ML pour l'anglais *Machine Learning*) est devenu l'un des principaux outils d'analyse d'images médicales. C'est un domaine très actif de l'intelligence artificielle. L'objectif principal est de trouver des solutions pour développer des outils pour aider les cliniciens à diagnostiquer, prédire le risque de maladie et à le prévenir avant qu'il ne soit trop tard. C'est grâce à la capacité stéréoscopique d'apprendre, d'interpréter et de prendre des décisions avec référence à une classe d'algorithmes informatiques qui construisent des modèles pour la classification. ML était à la base des systèmes de détection ou de diagnostic assistés par ordinateur (CAO) qui ont été développés pour aider les radiologues à améliorer la précision du diagnostic.

L'**apprentissage en profondeur** (ou DL pour l'anglais *Deep Learning*) est également une nouvelle partie de ML car il englobe un large éventail d'architectures réseaux conçues pour effectuer plusieurs tâches. Les réseaux de neurones sont utilisés pour analyser les images médicales depuis 20 ans. Ces dernières années, il s'est concentré sur l'application de DL à un large éventail de tâches d'analyse d'images médicales (segmentation, classification, détection, reconstruction d'image, optimisation, etc.) [1] [2]. Récemment, les **réseaux de neurones à convolution** ou **CNN** (pour l'anglais *the convolutional neural network*) a démontré qu'il était capable de dériver automatiquement des représentations de caractéristiques puissantes à partir des images pour la classification optimale d'images dans plusieurs domaines, contrairement aux représentations globales conventionnelles.

Objectifs du travail mémoire

Le **cancer de la peau** est une maladie cutanée causée par les rayons ultraviolets résultant d'une exposition excessive au soleil ou à une moindre quantité de mélanine (mélanine, le pigment protecteur présent dans la couche supérieure de la peau (l'épiderme)). Chaque année, plus de 5,4 millions de nouveaux cas de cancer de la peau sont diagnostiqués chez plus de 3,3 millions de personnes aux États-Unis. Un diagnostic précis et rapide est la meilleure solution pour lutter contre cette maladie. En effet, les méthodes normales mettent plus de temps à donner un résultat et à donner des résultats inexacts. Les chercheurs se sont intéressés à ce type de cancer en testant des algorithmes qui impliquent un apprentissage en profondeur pour diagnostiquer la présence ou l'absence de cancer de la peau.

Récemment, Settou, Kholadi et Ben Ali, des chercheurs de département d'informatique à l'université d'El Oued, ont proposé et évalué une nouvelle **approche hybride d'extraction de caractéristiques** pour la classification des dommages sismiques [33]. Cette approche est basée sur des représentations des caractéristiques profondes (ou de DL) extraites par plusieurs modèles CNN pré-entraînés et à la fois de deux premières couches entièrement connectées (FC1 et FC2) de chaque modèle CNN utilisé, en vue d'améliorer la performance de classification. L'objectif de notre travail de mémoire est de réappliquer et tester cette approche dans le contexte d'imagerie médicale et plus précisément à la classification du cancer de la peau.

Organisation du mémoire

Ce mémoire est organisé de la manière suivante :

- Le premier chapitre présente d'abord les concepts et les notions de base en analyse d'images puis les méthodes d'extraction des caractéristiques.
- Le deuxième chapitre est consacré au DL. Nous expliquerons la relation entre ML et DL, puis nous présenterons les applications et les différentes approches de DL. En plus, nous exposerons l'architecture de réseau CNN et plusieurs différents modèles.
- Le troisième chapitre est dédié à la présentation d'un état de l'art sur la classification d'images médicales par des méthodes de ML et DL.

- Le quatrième chapitre est consacré à la présentation de l’approche générale d’hybridation de plusieurs vecteurs de caractéristiques DL et à la validation expérimentale de cette approche pour la classification des types du cancer de la peau.

Enfin, nous finissons ce mémoire par une conclusion générale donnant des améliorations envisageables ainsi que des perspectives d’avenir.

Chapitre 1

Extraction de caractéristiques et classification d'images

1.1. Introduction

L'analyse est l'étape la plus importante de la photographie, car la principale information pour interpréter un message visuel par un observateur humain est le contour ou la texture. En outre, l'extraction se fait à partir des domaines informatiques de la navigation, de la recherche et de la récupération d'images à partir de bases d'images. En ce qui concerne la classification, nous exploitons l'information visuel pour regrouper les images en plusieurs catégories.

L'analyse d'image est l'extraction d'un certain nombre de caractéristiques et leur expression sous une forme limite, où l'étape d'extraction précède l'étape de prise de décision pour répondre aux questions connexes. Comme on le trouve dans notre cas d'étude, cette texture biologique est-elle saine ou malade ?

La texture est une caractéristique visuelle fondamentale de la vision par ordinateur. Le principal problème réside dans le fait qu'il est facile pour le système visuel humain de faire la distinction entre les structures compositionnelles.

Dans ce chapitre, nous présenterons des concepts et notions liés à l'extraction caractéristique de bas niveaux, de niveau intermédiaire et de hauts niveaux, à l'apprentissage automatique, et aux méthodes de classification.

1.2. Caractéristique de bas niveaux

1.2.1. Approches globales

L'objectif de cette section est de lister tous les descripteurs utilisables dans un cadre de reconnaissance d'images, et de présenter leur spécificité. Il présente comment intégrer des informations dans un système de reconnaissance d'objets basé sur des informations de couleur, de texture ou de forme, afin de comparer des images.

1.2.1.1. *Descripteurs de couleurs*

Cohérence spatiale : dans certains systèmes, tels que les vagues d'eau ou l'optique, les états ondulatoires peuvent s'étendre sur une ou deux dimensions. La cohérence spatiale décrit la corrélation (ou la relation prévisible) entre les ondes à différents points de l'espace,

soit latéralement, soit longitudinalement. La cohérence temporelle décrit la corrélation entre les ondes observées à différents moments dans le temps. La plage de séparation entre les deux points sur lesquels il y a une interférence significative définit le diamètre de la zone de cohérence. Il est également utilisé dans les systèmes d'imagerie optique et en particulier dans divers types de télescopes d'astronomie. Parfois, les gens utilisent également la « cohérence spatiale » pour désigner la visibilité lorsqu'un état ondulatoire est combiné à une copie spatialement décalée de lui-même [01].

Corrélogrammes : est un graphique de statistiques de corrélation et il est une représentation graphique mettant en évidence une ou plusieurs corrélations entre des séries de données. Par exemple, dans l'analyse des séries chronologiques, un graphique des autocorrélations de l'échantillon par rapport (les décalages temporels) est une auto corrélogramme. Si une corrélation croisée est tracée, le résultat est appelé un corrélogramme croisé. Le corrélogramme est un outil couramment utilisé pour vérifier le caractère aléatoire d'un ensemble de données. Si elles sont aléatoires, les autocorrélations doivent être proches de zéro pour toutes les séparations à décalage temporel. S'il n'est pas aléatoire, alors une ou plusieurs des autocorrélations seront significativement différentes de zéro. De plus, des corrélogrammes sont utilisés dans l'étape d'identification du modèle pour les modèles de séries chronologiques à moyenne mobile autorégressive de Box-Jenkins. Les autocorrélations doivent être proches de zéro pour le caractère aléatoire ; si l'analyste ne vérifie pas le caractère aléatoire, alors la validité de bon nombre des conclusions statistiques devient suspecte. Le corrélogramme est un excellent moyen de vérifier ce caractère aléatoire. Parfois, les corrélogramme matrices de couleurs des forces de corrélation en analyse multivariée [02].

1.2.1.2. Descripteurs des textures

Il n'y a pas de définition cohérente et précise du concept de texture. En effet, les textures naturelles ne sont pas uniformes et n'ont pas une définition précise, malgré leur présence complète sur les photographies (images médicales et aériennes). Il existe quelques définitions de la texture, notamment :

La définition donnée par le dictionnaire, affirme qu'une texture est la reproduction spatiale d'un motif de base dans plusieurs directions. La texture est une propriété de la surface ou de la structure de l'objet [03].

Les caractéristiques de texture extraites sont bénéfiques pour reconnaître et distinguer diverses structures. Une texture est une répétition d'éléments avec une certaine fréquence et caractérisée par différentes statistiques (moyenne, variance, histogramme) [04].

Dans la littérature, deux grandes familles de texture s'imposent, les textures déterministes et les textures stochastiques.

Une texture est dite *déterministe* si elle est formée par un agencement régulier d'un motif géométrique. Ce type de texture est entièrement caractérisé par un motif et sa règle déterministe d'agencement.

Les textures dites *stochastiques* se caractérisent par un contenu irrégulier dont il est parfois difficile d'en extraire une primitive de base (motif non localisable). Elles sont décrites par des lois ou des modèles statistiques et sont considérées comme une réalisation d'un champ aléatoire bidimensionnel homogène [05].

Le but de la classification est d'assigner chaque région uniformément texturée à la classe de texture à laquelle elle appartient suivant un certain critère de similitude.

Le but de la segmentation de texture de découper une image en régions homogènes uniformément texturées ou possédant des propriétés proches ou identiques en faisant ressortir les frontières entre les classes et les descripteurs de texture. Il existe plusieurs méthodes, notamment LBP, GLCM, Filtre de Gabor, transformée en ondelettes et HOG. Et chacune d'elles en dépend en raison de ses avantages, et nous évitons également certaines d'entre elles pour leurs défauts.

Méthode LBP : est basée sur le signe de la différence entre le pixel central et ses voisins. Une fois que le symbole LBP est déterminé pour chaque pixel de l'image, cette méthode est facile à mettre en œuvre, stabilité aux changements d'éclairage monotones, stabilité de rotation et faible complexité de calcul. La LBP est également l'une des méthodes les plus utilisées pour l'analyse des tissus.[06][07]

Méthode GLCM : représente un histogramme de valeurs de niveaux de gris simultanées à un décalage horizontal spécifié sur l'image [08].

Méthode du filtre de Gabor : Filtre de Gabor Filtre linéaire dont la réponse est une onde impulsionnelle modulée par une fonction gaussienne ; Les filtres de Gabor sont largement utilisés dans l'indexation, pour décrire la texture.

L'utilisation des filtres de Gabor consiste à analyser indépendamment différentes parties de l'espace de Fourier, à l'aide de plusieurs filtres. Chacun est défini par des valeurs plus élevées de coefficients à une fréquence et une direction particulière et est très sensible aux fréquences situées au centre de la région. Tous les filtres sont conçus selon la même loi, qui est gaussienne, avec des paramètres différents. Les filtres de Gabor sont aujourd'hui une référence pour l'indexation d'images et la recherche par texture [09].

Transformée en ondelettes : Méthode de transformation d'onde : La transformation d'onde est un ensemble infini de transformations différentes, avec de nombreuses utilisations dans différents domaines, similaire à la transformée de Fourier avec diverses fonctions de maturation.

La différence entre les transformées en quatre et en ondelettes est que la transformée de Fourier décompose le signe en sinus et cosinus, c'est-à-dire pour localiser les fonctions dans l'espace de Fourier ; Comme pour la transformée en ondelettes, elle utilise des fonctions qui existent à la fois dans l'espace de Fourier et dans l'espace cognitif [10].

Pour calculer la transformation en ondelettes, on n'a besoin que de deux filtres donc on utilise une paire de filtres pour chaque niveau de résolution (filtre passe-haut et filtre passe-bas) : au lieu de calculer le produit scalaire de l'ondelette par le signal, on perçoit la sortie asymétrique du signal à l'aide de ces filtres.

Méthode HOG : est une caractéristique utilisée en vision par ordinateur. La technique calcule des histogrammes locaux de l'orientation du gradient sur une grille dense, c'est-à-dire sur des zones régulièrement réparties sur l'image. Parce qu'est que l'apparence et la forme locale d'un objet dans une image peuvent être décrites par la distribution de l'intensité du gradient ou la direction des contours. Ceci peut être fait en divisant l'image en régions adjacentes de petite taille, appelées cellules, et en calculant pour chaque cellule l'histogramme des directions du gradient ou des orientations des contours pour les pixels à l'intérieur de cette cellule.[11]

1.2.1.3. Descripteurs de forme

La forme est un descripteur très important dans l'indexation des images. La forme désigne l'aspect général d'un objet, son contour. Il existe plusieurs de méthodes utilisées permettant de reconnaître une forme donnée dans une image comme :

Les descripteurs géométriques de région de forme permettent de distinguer les différents types de forme que peuvent prendre les objets d'une scène. Ils nécessitent une segmentation en région préalable de l'image calculés sur les différentes régions de l'image, surface relative est le nombre de pixels. Ces descripteurs très simples permettent d'obtenir des informations sur la géométrie des régions de l'image.

Les moments géométriques permettent de décrire une forme à l'aide de propriétés statistiques. Itérée pour toutes les formes d'une image, cette méthode permet finalement de relever toutes les formes communes à deux images [12].

1.2.2. Approches locales

1.2.2.1. Détecteur de coins de Harris

Les pièces Harris sont découvertes par Harris en raison des limitations du détecteur *Maravoc*. Parmi les limitations prises en compte par Harris, citons :

1. La réponse du détecteur est anisotrope en raison de la nature discrète des directions de changement qui peuvent être effectuées et pour l'optimisation, nous nous appuyons sur l'élargissement de Taylor de la fonction de densité.
2. La réponse du détecteur de *Moravec* est bruitée du fait du voisinage considéré. Pour améliorer cela, Harris et Stephen ont suggéré d'utiliser un filtre gaussien.
3. Le détecteur de *Moravec* est très réactif aux courbes de niveau, même si seul un minimum de E doit être pris en compte à partir de chaque pixel afin de respecter le comportement [13].

1.2.2.2. SIFT

David Lowe a développé l'algorithme SIFT en 1999. Cet algorithme a été breveté aux États-Unis et est le propriétaire de l'Université de la Colombie-Britannique.

Méthode SIFT (Fixed Feature Conversion of Scale = Transformation de caractéristique visuelles invariante à l'échelle) : Cette méthode est représentée en convertissant l'image en un ensemble de caractéristiques vectorielles qui sont fixées par les transformations géométriques habituelles (proportionnelle et rotation). Le point fort de cette méthode est la possibilité de faire correspondre des points forts distants avec de grandes formes de caméras.

Applications SIFT : Représenté dans la recherche de techniques d'extraction d'informations de l'image à travers des algorithmes nécessitant une vision artificielle, parmi les techniques qui consistent à trouver une transformation géométrique qui permet le passage d'une image (la source) à une autre (la cible) image. Il existe plusieurs applications comme : Assistance des conducteurs de voitures, la photographie digitale, le suivi des visages, les films et les vidéos ... etc.

Algorithmes SIFT : Cet algorithme se concentre sur deux étapes de base, à savoir :

La première étape : convertir l'image en propriétés vectorielles. Cette transformation comporte des étapes à suivre, à savoir : calcule des points d'intérêts et des descripteurs, mise en correspondance.

La deuxième étape : comparer les descripteurs des deux images, afin de trouver la transformation qui a subi l'image, les étapes suivantes doivent être suivies ; construction de l'espace des échelles, localisation des extrema dans les DoG, localisation précises des points clés/Élimination des points à faible contraste et des points sur les arêtes, affectation d'orientation aux points clés et calcule des descripteurs.

1.2.2.3. SURF

Il s'agit d'un algorithme de détection de caractéristiques et d'un descripteur ou d'un algorithme de caractéristiques accélérées robuste qui a été introduit par les chercheurs de l'ETH Zurich et de la Katholieke Universiteit Leuven en 2006 pour la première fois et en 2008, il a été révisé. Ses utilisations se situent dans le domaine de la vision par ordinateur, pour découvrir des objets ou des tâches de reconstruction 3D.

SURF est une inspiration partielle du SIFT, qui, selon ses auteurs, est plus rapide et, selon ses auteurs, plus puissant pour différentes transformations d'images. SURF s'est basé sur des agrégats bidimensionnels de réponse en ondelettes de *Haar* et a également utilisé efficacement des images intégrées (comme caractéristique principale).

L'approximation de l'onde de *Haar* est utilisée pour un détecteur de point dépendant d'un déterminant [14] et le hessien est également utilisé parce que les descripteurs doivent fonctionner dans un modèle spatial non linéaire, "les réponses dérivées sont calculées et résumées dans un vecteur de descripteur de point clé", puis "le vecteur est centré au point de base " et dans le dernier " le descripteur est normalisé au vecteur unitaire" [15].

1.3. Caractéristique de plus hauts niveaux

Le traitement de haut niveau consiste à reconnaître l'objet afin de le classer en objet en fonction de la connaissance relative des éléments de l'image. Les méthodes d'intelligence artificielle sont appliquées dans le domaine de l'analyse et de l'interprétation d'images, car cette analyse de haut niveau dépend et se concentre sur la simulation de la perception humaine.

1.4. Types d'apprentissage automatique

1.4.1. Apprentissage supervisé

Dans l'apprentissage supervisé, l'algorithme a accès à ce qu'il essaie de prédire, lorsqu'il existe un ensemble de données étiquetées ou des exemples qui ont été liés à la classe par un enseignant ou un expert. Cet ensemble d'exemples constitue la base de l'apprentissage, cela peut être, par exemple, la présence ou l'absence d'une maladie ou la gravité des symptômes. Ensuite, les méthodes d'apprentissage supervisé ont pour objectif global de construire à partir de la base d'apprentissage ou des fonctions de classification. Cette fonction permet, à partir de la description de l'objet, d'identifier un attribut spécifique, la classe. L'utilisation d'un algorithme est axée sur la recherche de la fonction optimale qui capture la relation entre l'entrée et la variable cible. La raison pour laquelle ce type d'apprentissage est appelé « supervisé » est que l'algorithme en a une connaissance préalable. Les valeurs de sortie sont (par exemple : tumeur maligne vs tumeur bénigne). L'algorithme est entraîné à l'aide de plusieurs exemples. Dans ce sens, l'apprentissage est un processus itératif de prédictions et d'ajustements ultérieurs, de sorte que la différence entre la sortie et les attentes cibles soit réduite autant que possible. Du côté des performances, il s'agit d'un calcul qui est mesuré en comparant les prédictions de l'algorithme aux valeurs cibles réelles dans les nouvelles données [16].

Les méthodes d'apprentissage supervisé ont pour objectif global de construire à partir de la base d'apprentissage ou des fonctions de classification. Nous exposons sur la figure suivante une schématisation résumant l'apprentissage supervisé.

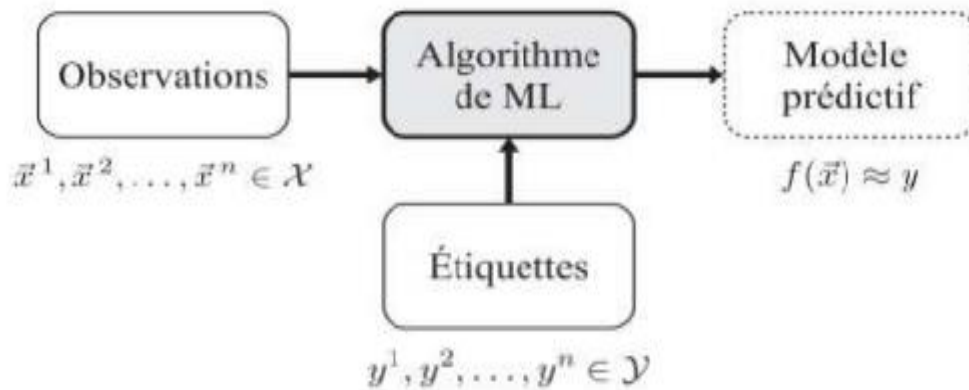


Figure 01 : Schéma apprentissage supervisé [17].

1.4.2. Apprentissage non supervisé

L'apprentissage non supervisé, apprentissage par observation ou découverte, consiste à déterminer une classification « raisonnable » à partir d'un ensemble d'objets ou de situations particulières (exemples non nommés). Il n'y a pas de valeur cible claire dans la tâche d'apprentissage non supervisée. Le but de ce type est de découvrir des structures de base dans les données. La tâche de l'apprentissage non supervisé est de trouver des solutions aux problèmes d'assemblage.

Elle concerne notamment les sciences naturelles, les classifications de documents et de livres, mais aussi la classification des sciences élaborée au cours des siècles par les philosophes [16].

L'automatisation des bâtiments de classification est aujourd'hui un véritable domaine de recherche. L'idée principale utilisée dans la création de classes d'objets est une mesure de similarité entre les objets. Les catégories ou les concepts sont construits de manière à augmenter la similitude au sein de la classe et à réduire la similitude entre les classes. L'apprentissage non supervisé correspond également à une catégorisation conceptuelle.

Nous exposons sur la figure suivante une schématisation résumant l'apprentissage non supervisé.

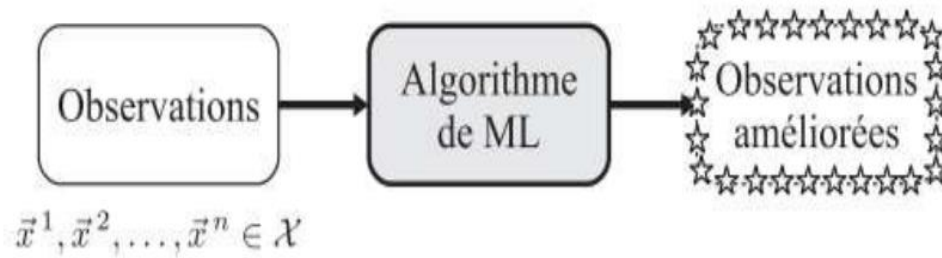


Figure 02 : Schéma apprentissage non supervisé [17].

1.4.3. Classification

La classification d'images est une série de méthodes unies par la théorie pour pouvoir utiliser des images pour des analyses complémentaires ou pour la cartographie. L'objectif de la classification est de développer un système capable d'attribuer une classification automatique des images, la fatigue et le temps requis pour une grande quantité de données d'image.

Le modèle de classification représente un modèle ML dont la sortie appartient à un ensemble limité de valeurs (par exemple : bon, moyen, mauvais) et il appartient également à l'un des types d'apprentissage supervisé ou non supervisé [18].

Il existe deux types de classification :

Classification hiérarchique : Il s'agit d'une méthode de classification itérative dont le principe est simple, car elle consiste à regarder la hiérarchie d'un niveau inférieur à un niveau supérieur et cela peut être dans le sens inverse. Où le niveau suivant est le niveau qui est atteint par en passant par le chapitre attendu dans le niveau actuel.

L'ensemble des classifications en cascade réalisées produit un arbre où chaque nœud d'un niveau correspond à la classe attendue à ce niveau.

Classification multi-classes : Une classification multi-classes définit une tâche de classification avec plus de deux classes (k classe ou $k > 2$), le but est de définir des instances de l'une des k classes potentielles. La classification hiérarchique résout le problème de la classification multi-classes en divisant l'espace de sortie, par exemple dans un arbre.

Chaque nœud maître est divisé en plusieurs nœuds enfants et le processus se poursuit jusqu'à ce que chaque nœud enfant ne représente qu'une seule classe [19].

1.5. Méthodes de classification

Il existe plusieurs méthodes de classification dont les suivantes :

1.5.1. K plus proches voisins

L'algorithme de k plus proches voisins (noté parfois KNN pour l'anglais *K-Nearest Neighbors*) fait partie des méthodes de classification supervisée les plus couramment utilisées, en particulier en analyse d'images. Il a l'avantage d'être à la fois non-paramétrique, simple et efficace.

L'algorithme KNN permet de traiter des nuages de points non linéairement séparables [20]. Le principe de l'algorithme KNN est de définir une classification des nouvelles données x c'est-à-dire la classification de la liste k plus proche des voisins. Quant à la détermination de l'appariement ou de la similitude, on s'appuie sur des mesures de distances (la distance euclidienne est la plus utilisée. C'est des premières formes d'utilisation des données étiquetées en imagerie médicale était l'algorithme KNN [21].

L'avantage majeur de KNN est d'être simple à concevoir. Cependant, son inconvénient majeur est qu'il est sensible aux bruits. Un autre point négatif de KNN réside dans le fait que pour un nombre de variable prédictives très grands, le calcul de la distance devient très coûteux.

1.5.2. Naïf de Bayes

Naive Bayes (NB) est une collection d'algorithmes de classification basés sur le théorème de Bayes. Il est une technique de classification supervisée. Couramment utilisé dans l'apprentissage, Il s'agit d'une technique de classification basée sur le théorème de Bayes avec une hypothèse d'indépendance parmi les prédicteurs. En termes simples, un classificateur NB suppose que la présence d'une caractéristique particulière dans une classe n'est pas liée à la présence d'une autre caractéristique.

Le classifieur NB est un algorithme supervisé probabiliste qui suppose que l'existence d'une caractéristique pour une classe, est indépendante de l'existence d'autres caractéristiques, raison pour laquelle on utilise l'adjectif « naïf ». Une personne peut être considérée comme un homme s'il pèse un certain poids et mesure une certaine taille. Même si

ces caractéristiques sont liées dans la réalité, un classifieur BN détermine que la personne est un homme en considérant indépendamment ces caractéristiques de taille et de poids.

Malgré des hypothèses de base extrêmement simplistes, ce classifieur conduit à de très bons résultats dans beaucoup de situations réelles complexes. En 2004, un article a montré qu'il existe des raisons théoriques derrière cette efficacité inattendue [22]. Toutefois, une autre étude de 2006 montre que des approches plus récentes (arbres renforcés, forêts aléatoires) permettent d'obtenir de meilleurs résultats [23].

1.5.3. Arbre de décision

Les arbres de décision (ou DT pour l'anglais *Decision Tree*) sont un outil très populaire de classification. Ils appartiennent aux méthodes d'apprentissage supervisés non paramétriques, qui sont largement utilisées en classification et en régression [24].

Les DT se fondent sur une approche algorithmique et se visualisent sous forme d'un arbre avec des règles identifiant les techniques de fractionnement d'un ensemble de données. L'ultime enjeu est la création d'un modèle ayant comme objectif la prédiction de la valeur d'une variable cible en se basant sur les règles de décision [25].

Le principe de DT repose sur la construction d'un arbre de taille limitée. La racine constitue le point de départ de l'arbre et représente l'ensemble des données d'apprentissage. Puis ces données sont segmentées en plusieurs sous-groupes, en fonction d'une variable discriminante (un des attributs). Une fois l'arbre construit à partir des données d'apprentissage, on peut prédire un nouveau cas en le faisant descendre le long de l'arbre, jusqu'à une feuille. Comme la feuille correspond à une classe, l'exemple sera prédit comme faisant partie de cette classe.

Exemple, sur la Figure 03, la première variable discriminante est l'*outlook*. Elle divise la population en trois sous-groupes : *Sunny*, *Overcast* et *Rain*. Le processus est ensuite réitéré au deuxième niveau de l'arbre, où les sous-populations sont segmentées à leur tour en fonction d'une autre valeur discriminante.

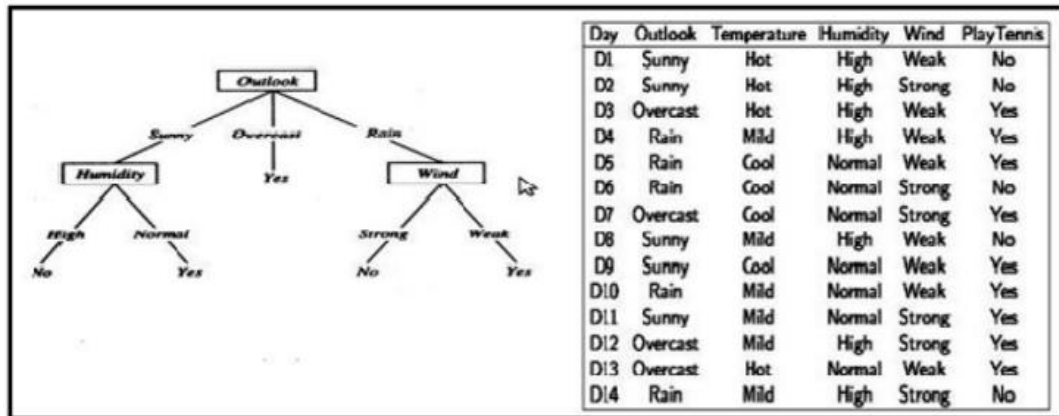


Figure 03 : Exemple d'arbre décision [26].

Les DT a les avantages suivants : simple à comprendre et à interpréter, peu de préparation des données, les variables prédictives en entrée peuvent être aussi bien qualitatives que quantitatives, et performant sur de grands jeux de données.

1.5.4. Forêt aléatoire

L'algorithme des « forêts aléatoires » (ou *Random Forest*, *RF*, parfois aussi traduit par forêt d'arbres décisionnels) est un algorithme de classification qui réduit la variance des prévisions d'un arbre de décision seul, améliorant ainsi leurs performances. Pour cela, il combine de nombreux arbres de décisions dans une approche de type *bagging*. L'idée générale derrière la méthode est la suivante : au lieu d'essayer d'obtenir une méthode optimisée en une fois, on génère plusieurs prédicteurs avant de mettre en commun leurs différentes prédictions.

L'algorithme de RF peut se diviser en deux étapes. Premièrement, dans lequel n arbres aléatoires sont créés, cela forme la forêt aléatoire. Dans la deuxième étape, le résultat pour la même caractéristique de test de tous les arbres de décision est combiné. Ensuite, la prédiction finale est dérivée en évaluant les résultats de chaque arbre de décision ou simplement en faisant une prédiction qui apparaît le plus souvent dans les arbres de décision (RF est illustré sur Figure 04) [28].

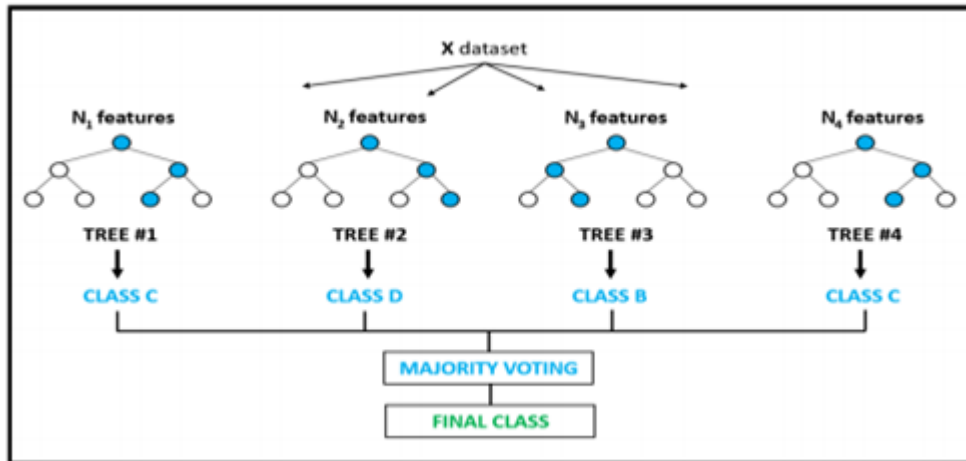


Figure 04 : Exemple de forêts aléatoires.

RF a les avantages suivants :

- C'est un des meilleurs algorithmes pour ce qui est de la précision.
- Reconnaissance très rapide ;
- Multi-classes par nature ;
- Efficace sur inputs de grande dimension ;

Ses inconvénients sont :

- Apprentissage souvent long ;
- Une implémentation difficile ;
- Valeurs extrêmes souvent mal estimées dans cas de régression.

1.5.5. SVM

Une machine à vecteurs de support (SVM) est un classificateur discriminant formellement défini par un hyperplan séparateur. SVM est un ensemble de méthodes d'apprentissage supervisé utilisées pour la classification, la régression et la détection des valeurs aberrantes. L'algorithme produit un hyperplan optimal qui catégorise les nouveaux exemples. Dans un espace à deux dimensions, cet hyperplan est une ligne divisant un plan en deux parties où dans chaque classe se trouvait de chaque côté. Il s'agit d'un ensemble d'équations mathématiques réglées pour vous donner la réponse la plus précise possible le plus rapidement possible. Un SVM peut être utilisé pour détecter les cellules cancéreuses sur la base de millions d'images ou pour prédire les futurs itinéraires de conduite avec un modèle de régression bien ajusté.

Un SVM est capable de séparer des frontières très complexes où les données ne sont pas linéairement séparables, ceci en utilisant une fonction de transformation d'espace (fonction noyau). En effet, les SVMs utilisent divers types de noyaux pour transformer les données non linéairement séparables en données linéairement séparables au biais du changement de l'espace de représentation des données d'entrées en un espace de plus grande dimension, dans lequel il est probable d'exister une séparation linéaire.

Il existe deux types différents de SVM ; chacun est utilisé pour des choses différentes :

- ✓ SVM simple : est généralement utilisé pour les problèmes de régression linéaire et de classification.
- ✓ Kernel SVM : offre plus de flexibilité pour les données non linéaires, car on peut ajouter plus de fonctionnalités pour s'adapter à un hyperplan au lieu d'un espace bidimensionnel.

Le principe de SVM repose sur deux notions clés : la notion de marge maximale et la notion de noyau. La marge est la distance séparant la frontière de séparation (l'hyperplan) et les prototypes les plus proches de cette frontière. Ces prototypes sont appelés vecteurs supports. Dans les SVM, la frontière de séparation est choisie comme étant celle qui maximise la marge. Parmi les modèles des SVM, il y a le cas linéairement séparable et les cas non linéairement séparable. Les premiers sont les plus simples car ils permettent de trouver facilement le classifieur linéaire.

Les SVMs permettent de traiter des problèmes de classification non linéaire complexe. Un autre avantage réside dans le fait que les SVM constituent une alternative aux réseaux de neurones car plus faciles à entraîner. L'inconvénient commun de SVMs est qu'elles sont souvent moins performantes que les forêts aléatoires.

1.6. Conclusion

Dans ce chapitre, nous expliquons toutes les caractéristiques d'une image. L'analyse est une étape très importante en imagerie. En effet, les principales informations pour l'interprétation d'un message visuel pour un observateur humain demeurent les contours ou bien les textures.

Dans l'analyse d'image, la texture présente un descripteur très important pour la description du contenu de l'image. Dans ce sens, de nombreuses approches d'analyse de la texture ont été développées et validé à travers les travaux de la littérature.

Durant un processus de classification, une bonne précision obtenue est directement liée aux caractéristiques des descripteurs et à leur capacité à discriminer les images dans les différentes classes.

A cet effet, les chercheurs sont toujours à la recherche de descripteurs faciles à calculer tout en ayant le pouvoir de discriminer les classes, pour contribuer considérablement à la performance de classification. Nous avons discuté des types d'apprentissage supervisé et non supervisé. Nous avons traité de l'illustration d'algorithmes de classification

Dans le deuxième chapitre, nous éclairerons l'apprentissage profond et ses architectures.

Chapitre 2

Apprentissage en profondeur

"Deep Learning"

2.1. Introduction

Récemment, l'utilisation de l'intelligence artificielle (IA) est devenue de plus en plus courante dans tous les domaines de la vie pratique, en raison de sa quête continue de moyens et de solutions à des problèmes d'une grande complexité logique ou algorithmique, en utilisant l'apprentissage en profondeur et l'apprentissage automatique, en particulier dans l'avenir des soins de santé. L'apprentissage profond est un groupe de méthodes d'apprentissage automatique qui tentent de modéliser à un niveau élevé d'abstraction de données via l'architecture CNN.

Dans ce chapitre, nous mettrons en évidence l'apprentissage en profondeur, l'apprentissage automatique et toutes les architectures CNN.

2.2. Apprentissage automatique (Machine Learning)

ML est l'une des applications de l'IA, qui donne aux systèmes la capacité d'apprendre et de s'améliorer automatiquement à partir de l'expérience afin de devenir plus efficaces dans la prédiction des résultats sans être explicitement programmés pour le faire.

Le processus d'apprentissage commence par observations ou des données ou une expérience directe ou des instructions afin de prendre de meilleures décisions à l'avenir. L'objectif principal est de permettre aux ordinateurs d'apprendre automatiquement sans intervention humaine ou d'assistance et d'ajuster les actions en conséquence [26].

2.3. Apprentissage en profondeur (*Deep Learning*)

DL est une classe d'algorithmes de ML qui utilisent plusieurs couches pour extraire des fonctions d'un niveau progressivement plus élevé à partir d'une entrée élémentaire. Par exemple, dans le traitement d'image, les couches inférieures peuvent définir des bords, tandis que les couches supérieures peuvent définir des concepts humains tels que des nombres, des lettres ou des visages.

DL repose sur ce que l'on appelle, par analogie, des "réseaux de neurones artificiels", constitués de milliers d'unités "neurones" qui effectuent chacune de petites et simples opérations [27].

2.4. Relation entre ML et DL

DL est une forme spécialisée d'apprentissage automatique. Un flux de travail d'apprentissage automatique commence par l'extraction manuelle des fonctionnalités pertinentes des images. Les fonctionnalités sont ensuite utilisées pour créer un modèle qui catégorise les objets de l'image. Avec un flux de travail d'apprentissage en profondeur, les fonctionnalités pertinentes sont automatiquement extraites des images. En outre, l'apprentissage en profondeur effectue un « apprentissage de bout en bout » - où un réseau reçoit des données brutes et une tâche à effectuer, telle que la classification, et il apprend à le faire automatiquement.

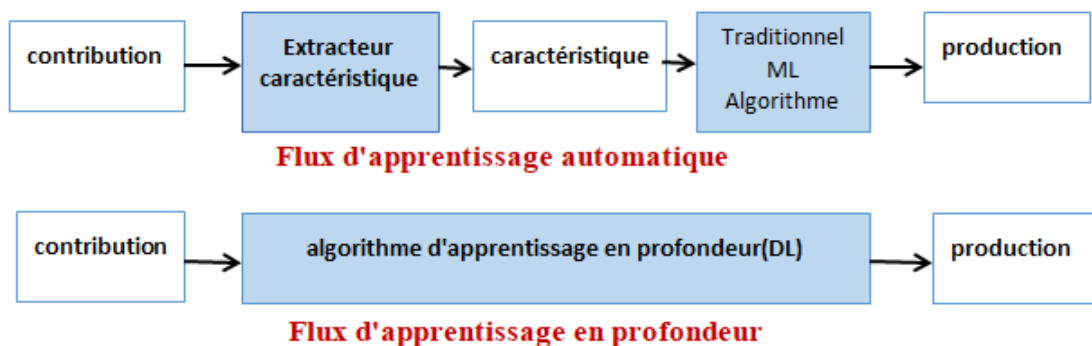


Figure 05 : Le procédé du ML classique comparé à celui du DL [27].

Différence clé est que les algorithmes de DL s'adaptent aux données, tandis que l'apprentissage superficiel converge. L'apprentissage superficiel fait référence aux méthodes d'apprentissage automatique qui plafonnent à un certain niveau de performance lorsque vous ajoutez plus d'exemples et de données d'entraînement au réseau.

Le principal avantage des réseaux de DL est qu'ils continuent souvent de s'améliorer à mesure que le volume de vos données augmente.

2.5. Domaines d'application de DL

Ces techniques se développent dans le domaine de l'informatique qu'appliquée aux NTIC à la robotique, à la bio-informatique, etc.

L'apprentissage profond est utilisé dans de nombreux domaines l'intelligence artificielle comme sont :

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

2.6. Les différents types de modèles

2.6.1. Machines Boltzmann restreintes

Une machine de Boltzmann contrainte a été définie comme un type de réseau neuronal artificiel pour un apprentissage non supervisé basé sur l'apprentissage automatique. Il a été inventé à l'origine comme Harmonium en 1986 par Paul Smolensk.

Cet outil est utilisé pour obtenir une estimation de la distribution de probabilité d'un ensemble de données.

La machine de Boltzmann se compose d'une couche de neurones d'entrée avec une couche cachée de neurones. Nous appelons la machine de Boltzmann une machine de restriction, et c'est dans le cas de tous les neurones d'une même couche indépendants les uns des autres. C'est un cas particulier de la machine Boltzmann [29][30].

2.6.2. Réseau de neurones "Neural Networks"

Réseaux de neurones Les méthodes d'apprentissage profond utilisent des architectures de réseaux de neurones, ce qui explique pourquoi les modèles d'apprentissage profond sont souvent appelés réseaux de neurones profonds (*Deep Neural Networks* – DNN). Les premières formulations du réseau de neurones remontent à 1943. L'idée est de reproduire le

fonctionnement d'un neurone humain, comme représenté sur la figure 06. Le fonctionnement d'un neurone formel est simple : c'est une somme pondérée d'entrées à laquelle on applique une fonction d'activation. Les coefficients de pondération sont appelés poids synaptiques et la fonction d'activation utilise un seuil : la sortie valant 1 si la somme pondérée dépasse le seuil et 0 sinon. On note b la valeur du seuil [28] :

$$Y = \begin{cases} 1 & \text{si } \sum_k W_k i_k > b \\ 0 & \text{sinon} \end{cases}$$

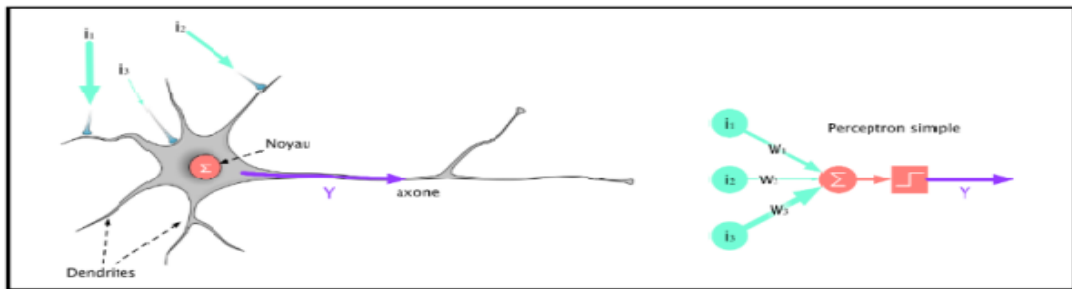


Figure 06 : Schéma d'un neurone à gauche et représentation d'un neurone formel à droite [31].

La fonction d'entrée-sortie peut varier. Donc un neurone peut être :

- **Une unité linéaire :** la sortie est proportionnelle à l'entrée pondérée totale ;
- **Une unité de seuil :** la sortie est fixée à l'un des deux niveaux, dépendamment de la puissance totale, si elle est supérieure à une valeur déterminée ou non ;
- **Une unité sigmoïde :** la sortie varie de façon continue, mais pas de façon linéaire.

2.6.3. Réseaux de neurones convolutifs (CNN)

2.6.3.1. Définition de CNN

Un réseau de neurones convolutifs (CNN ou ConvNet) est une classe de réseaux de neurones profonds, le plus couramment utilisé dans l'analyse d'images visuelles. Un CNN contient une ou plusieurs couches convolutives, souvent avec une couche de sous-échantillonnage chacune capable de reconnaître des formes plus complexes, ce qui est suivi d'une ou plusieurs couches connectées comme dans le réseau neuronal standard, le cortex, dans le cerveau. La couche de convolution CNN fait la même chose que les cellules humaines, chaque entité de la couche collecte les données d'un ensemble d'entités placées dans une petite classe de la couche précédente appelée champ de réception local.

Les réseaux de neurones convolutifs ont la capacité de détecter des motifs dans une image d'entrée, tels que des lignes, des dégradés, des cercles ou même des yeux et des visages. L'utilisation de CNN est devenue l'une des dernières technologies dans les applications visuelles telles que la classification d'images et le traitement du langage naturel [32].

2.6.3.2. Conception de réseaux neuronaux convolutifs

L'architecture d'un réseau de neurones convolutifs est un réseau de neurones à rétroaction multicouche, réalisé en empilant de nombreuses couches cachées les unes sur les autres en séquence. C'est cette conception séquentielle qui permet aux réseaux de neurones convolutifs d'apprendre des caractéristiques hiérarchiques.

Les couches cachées sont généralement des couches convolutives suivies de couches d'activation, certaines d'entre elles étant suivies de couches de regroupement.

Un simple réseau de neurones convolutifs qui facilite la compréhension des principes de conception de base est le premier réseau de neurones convolutifs LeNet-5, publié par Yann LeCun en 1998. LeNet est capable de reconnaître les caractères manuscrits [33].

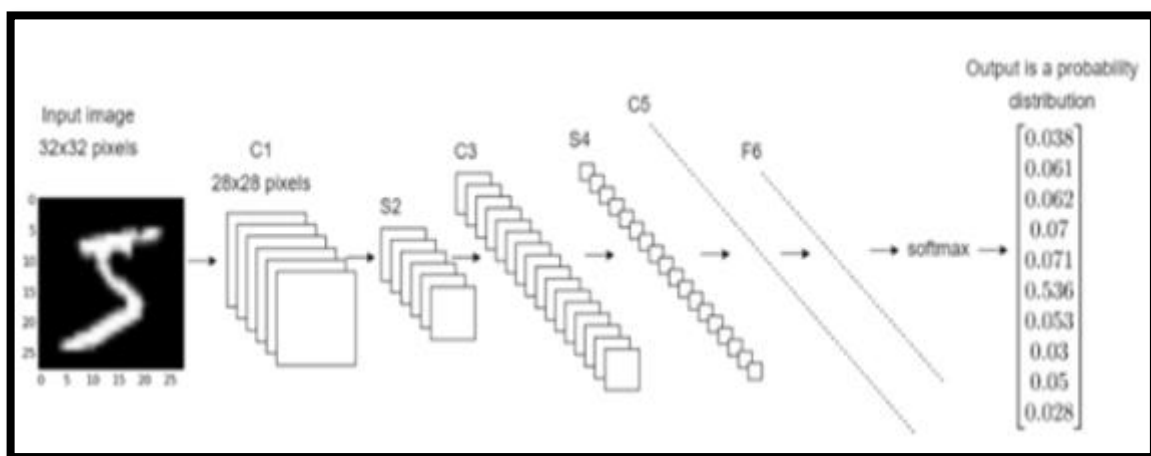


Figure 07 : Exemple de couches de réseau neuronal convolutif expliquées.

2.6.3.3. Couche convolutif

Elles sont considérées comme les plus importantes pour l'extraction de caractéristiques, chaque couche convolutive possède généralement un certain nombre de filtres (noyaux ou *Weights*) qui génère le même nombre de cartes d'entités en faisant glisser les filtres à travers les cartes d'entités de la couche précédente. L'objectif d'une couche convolutif est d'extraire les caractéristiques du volume d'entrée [32] [33].

2.6.3.4. Couches Pooling

Ces couches sont généralement placées entre deux couches convolution elles successives. Ils effectuent une opération de sous-échantillonnage pour réduire la taille des cartes d'entités et le nombre de paramètres réseau via la mise en commun maximale ou moyenne. Dans cette couche un sous-échantillonnage spatial est appliqué [34].

2.6.3.5. Couches entièrement connectées

Les couches entièrement connectées connectent chaque neurone d'une couche à chaque neurone d'une autre couche. Ces couches suivent généralement la convolution et l'interrogation [35].

2.6.3.6. Couche de correction (ReLU)

L'efficacité du traitement peut être améliorée en superposant des couches de traitement qui exécutent une fonction d'activation (fonction mathématique) via les signaux de sortie. Elle peut être considérée comme la fonction d'activation non saturée où la correction ReLU est représentée par l'équation mathématique $f(x) = \max(0, x)$. La couche de correction ReLU remplace toutes les valeurs négatives reçues en entrée par des zéros. Il joue le rôle d'une fonction d'activation. Il fait également référence à la vraie fonction non linéaire.

Elle peut être corrigée par la tangente hyperbolique $f(x) = \tanh(x)$ ou corrigée par le losange hyperbolique saturé : $f(x) = |\tanh(x)|$ ou correction par la fonction x [34][35].

2.6.4. Paramètres CNN

Il y a quatre paramètres, qui sont les suivants :

2.6.4.1. Fonction d'activation

Les fonctions d'activation sont des fonctions mathématiques et non linéaires qui sont appliquées à un signal à la sortie d'un neurone artificiel dans les réseaux de neurones pour calculer la somme pondérée des entrées et des biais, et sont utilisées pour déterminer si un neurone peut ou non se déclencher. Il traite les données présentées par un traitement hiérarchique qui est généralement une descente hiérarchique, puis produit une sortie pour le réseau de neurones, qui contient les paramètres des données [36].

Softmax : est une fonction mathématique qui convertit un vecteur numérique en un vecteur de probabilité, où les probabilités de chaque valeur sont proportionnelles à l'échelle relative de chaque valeur dans le vecteur. ; La fonction de cette couche est de transformer un vecteur de valeurs K réelles en un vecteur de valeurs K réelles dont la somme est 1. Les valeurs d'entrée peuvent être positives, négatives, nulles ou supérieures à un, mais softmax convertit à des valeurs comprises entre 0 et 1, afin qu'elles puissent être interprétées comme des possibilités. Nous utilisons une couche softmax pour convertir les scores en une distribution de probabilité normale [36].

Sigmoïde : Parfois appelée fonction logistique. Les résultats de la recherche de la fonction sigmoïde ont produit trois types de focus sigmoïde, qui sont utilisés dans les applications d'apprentissage en profondeur. Il est non linéaire et est couramment utilisé dans les réseaux de neurones à action directe. Elle est représentée par une fonction réelle dérivable, définissant les valeurs d'entrée réelles, avec un certain degré de lissage et des dérivées positives partout [36].

2.6.4.2. Époque (Epoch)

En chronologie, une ère est un moment dans le temps choisi comme origine d'une ère calendaire particulière. L'"époque" agit comme un point de référence par rapport auquel le temps est mesuré. L'époque est déterminée par correspondance. C'est un ensemble de cycles utilisés pour améliorer les résultats [37].

2.6.4.3. Taille du lot (Batch size)

La taille du lot est un hyper paramètre qui spécifie le nombre d'échantillons à traiter avant la mise à jour des paramètres du modèle interne. Considérez un lot comme une boucle pour itérer sur un ou plusieurs échantillons et faire des prédictions. À la fin du lot, les prédictions sont comparées aux variables de sortie attendues et l'erreur est calculée. A partir de cette erreur, un algorithme de mise à jour est utilisé pour améliorer le modèle, par exemple pour descendre le gradient d'erreur.

L'ensemble de données d'apprentissage peut être divisé en un ou plusieurs lots.

Avantages de l'utilisation de la taille du lot < nombre de tous les échantillons :

Il nécessite moins de mémoire. Étant donné que vous entraînez le réseau avec moins d'échantillons, la procédure d'entraînement exhaustive nécessite moins de mémoire.

Inconvénients de l'utilisation de la taille du lot <nombre de tous les échantillons :

Plus le lot est petit, moins l'estimation du gradient est précise [37].

2.6.4.4. Taux d'apprentissage (Learning rate)

Le taux d'apprentissage est un paramètre de réglage dans un algorithme d'optimisation qui détermine la taille du pas de chaque itération au fur et à mesure qu'il se rapproche de la fonction de perte minimale. C'est un paramètre super-configurable utilisé dans la formation des réseaux de neurones qui a une petite valeur positive, souvent dans le compris entre 0,0 et 1,0. Le taux d'apprentissage contrôle la rapidité avec laquelle le modèle s'adapte au problème. C'est aussi la vitesse à laquelle le modèle d'apprentissage automatique « apprend ».

2.7. Projet Image Net

Image Net est un grand projet de base de données visuelle utilisé pour rechercher un logiciel de reconnaissance visuelle d'objets. Au niveau de ce projet, plus de 14 millions d'images sont expliquées manuellement afin de faire référence aux objets.

Image Net contient plus de 20 000 catégories, avec un modèle inclus.

Le widget de données contient les légendes d'URL pour les images tierces gratuites directement à partir d'image Net.

En 2010, un concours de logiciels annuel a été organisé sous le nom de "Le défi de reconnaissance visuelle à grande échelle sur l'image Net (ILSVRC)". Ces programmes se disputent la classification et la découverte correctes des objets et des scènes. Le défi utilise une liste de "segments" "sur un millier de catégories qui ne se chevauchent pas [38].

Le but de ce concours est d'obtenir les meilleurs résultats basés sur la reconnaissance d'images image Net.

2.8. Architecture de CNN

Les CNN sont des architectures d'apprentissage en profondeur largement utilisées, ce qui a conduit à une révolution technique vision par ordinateur. Un modèle CNN typique est

construit en empilant plusieurs couches, y compris des couches convolutionnelles, des couches de regroupement et des couches entièrement connectées [32].

Au fil des ans, des variantes d'architectures CNN ont été développées, conduisant à des avancées étonnantes dans le domaine de l'apprentissage profond. Les taux d'erreur dans les compétitions telles que le défi ILSVRC Image Net constituent une bonne mesure de ces progrès. Dans ce concours, le taux d'erreur des 5 premiers pour la classification des images est passé de plus de 26% à moins de 3%. Dans cet article, nous examinerons certaines des architectures CNN populaires qui se sont démarquées dans leur approche et ont considérablement amélioré les taux d'erreur par rapport à leurs prédécesseurs.

Ces architectures servent de lignes directrices de conception générales que les praticiens de l'apprentissage automatique adapteront ensuite pour résoudre diverses tâches de vision par ordinateur.

Ces architectures servent d'extracteurs de fonctionnalités riches qui peuvent être utilisés pour la classification d'images, la détection d'objets, la segmentation d'images et de nombreuses autres tâches plus avancées.

2.8.1. Quelques réseaux convolutifs célèbres

2.8.1.1. VGG-Net

Le réseau VGG a été présenté par les chercheurs du Visual Graphics Group à Oxford. Ce réseau est pyramidal, où les couches inférieures les plus proches de l'image sont larges, tandis que les couches supérieures sont profondes. VGG contient des couches convolutives ultérieures suivies par des couches de regroupement. Les couches de regroupement sont responsables de rendre les couches plus étroites. Dans l'architecture VGG, tous les filtres ont une taille de 3×3 et des dimensions de 224×224 . Comme le montre la figure 08 suivante.

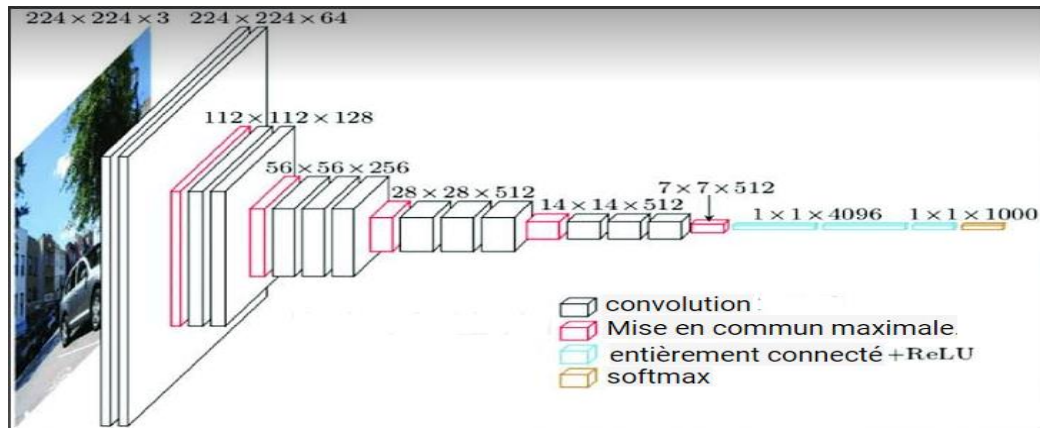


Figure 08 : L'architecture du modèle VGG-16 [39].

2.8.1.2. Google-Net

GoogleNet (ou Foundation Network) est une architecture conçue par les chercheurs de Google. Il contient 22 couches par rapport à VGG qui en contient 19, et ses trains sont plus rapides que VGG, et quant à sa taille précédemment apprise, il est relativement plus petit que la taille de VGG.

La structure finale contient de nombreuses unités constitutives, comme le montre la figure 09.

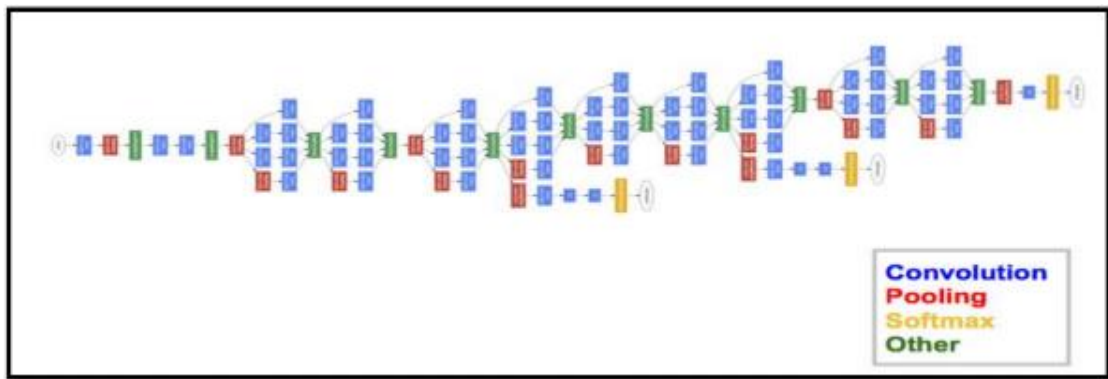


Figure 09 : Architecture Google-Net [39].

2.8.1.3. Res-Net

ResNet est meilleur qu'Alex-Net en termes de performances. ResNet est très populaire en raison de sa grande communication et de la facilité avec laquelle il peut former ce type de réseau, En 2015, cette architecture a remporté le concours Image Net. Parmi les caractéristiques de cette architecture Res Net, on trouve :

Écart de performance (c'est-à-dire réduire de moitié le taux d'erreur le plus élevé 5)

Cette architecture démontre une percée architecturale par rapport aux autres réseaux.

L'utilisation de la méthode de communication ResNet par saut, où cette idée a été inventée par l'auteur de ResNet, en raison du problème AlexNet, qui était proportionnel à la performance, c'est-à-dire que plus le réseau est profond, meilleures sont les performances et plus le réseau, les performances se sont détériorées en général, la difficulté de la formation.

La fonction de saut signifie que le réseau est lié les uns aux autres par des connexions d'identité. Où cette connexion représente une version simple des valeurs d'entrée ajoutées à la sortie de cette couche elle-même.

2.8.1.4. Alex-Net

AlexNet est la première architecture profonde ; il a été développé par Alex Krizhevsky et al. En 2012 pour participer au concours Image Net, ImageNet 2012) a convaincu une grande partie de la communauté de la vision par ordinateur de se pencher sérieusement sur l'apprentissage en profondeur pour les tâches de vision par ordinateur. Il a atteint un taux d'erreur de 17% dans le top 5 tandis que le deuxième meilleur n'atteignait que 26% !

ALEX NET s'appelle le modèle d'apprentissage par transfert et se compose de couches qui combinent quelques couches convolutives avec des couches entièrement connectées [32].

Ces couches sont divisées comme suit : couche entièrement connectée, couche softmax et couche de sortie de classification ; le transfert d'apprentissage préserve les valeurs des paramètres des couches précédentes du réseau précédemment formé. Le taux d'apprentissage initial est réglé sur une petite valeur pour ralentir la couche de transport.

L'une des différences les plus significatives entre Alex Net et les autres algorithmes de détection d'objets est l'utilisation de ReLU pour la partie non linéaire au lieu de la fonction Sigmond ou de Tanh comme les réseaux de neurones traditionnels. Alex Net tire parti de la formation plus rapide de ReLU pour rendre son algorithme plus rapide.

2.8.1.5. Inception-V3

Ce type d'architecture représente une architecture de réseau de neurones convolutifs de la famille Inception. Il s'intéresse aux améliorations car il utilise le lissage d'étiquette et la convolution (7 * 7) et utilise également un classificateur supplémentaire pour diffuser les informations.

Le tableau 01 résume plusieurs model CNN pré-entraîne. Nous choisissons les quatre model (VGG16, VGG19, ResNet 50, Inception) dans l'étude nous.

Tableau 01 : Model de CNN [40].

Model	Size	Top-1 Accuracy	Top-5 Accuracy	Paramètres	Depth
<u>Xception</u>	88 MB	0.790	0.945	22,910,480	126
<u>VGG16</u>	528 MB	0.713	0.901	138,357,544	23
<u>VGG19</u>	549 MB	0.713	0.900	143,667,240	26
<u>ResNet50</u>	98 MB	0.749	0.921	25,636,712	-
<u>ResNet101</u>	171 MB	0.764	0.928	44,707,176	-
<u>ResNet152</u>	232 MB	0.766	0.931	60,419,944	-
<u>ResNet50V2</u>	98 MB	0.760	0.930	25,613,800	-
<u>ResNet101V2</u>	171 MB	0.772	0.938	44,675,560	-
<u>ResNet152V2</u>	232 MB	0.780	0.942	60,380,648	-
<u>InceptionV3</u>	92 MB	0.779	0.937	23,851,784	159
<u>InceptionResNetV2</u>	215 MB	0.803	0.953	55,873,736	572
<u>MobileNet</u>	16 MB	0.704	0.895	4,253,864	88
<u>MobileNetV2</u>	14 MB	0.713	0.901	3,538,984	88
<u>DenseNet121</u>	33 MB	0.750	0.923	8,062,504	121
<u>DenseNet169</u>	57 MB	0.762	0.932	14,307,880	169
<u>DenseNet201</u>	80 MB	0.773	0.936	20,242,984	201
<u>NASNetMobile</u>	23 MB	0.744	0.919	5,326,716	-
<u>NASNetLarge</u>	343 MB	0.825	0.960	88,949,818	-
<u>EfficientNetB0</u>	29 MB	-	-	5,330,571	-
<u>EfficientNetB1</u>	31 MB	-	-	7,856,239	-
<u>EfficientNetB2</u>	36 MB	-	-	9,177,569	-
<u>EfficientNetB3</u>	48 MB	-	-	12,320,535	-
<u>EfficientNetB4</u>	75 MB	-	-	19,466,823	-
<u>EfficientNetB5</u>	118 MB	-	-	30,562,527	-
<u>EfficientNetB6</u>	166 MB	-	-	43,265,143	-
<u>EfficientNetB7</u>	256 MB	-	-	66,658,687	-

2.8.2. Scénarios de classification basée sur CNN

Il existe trois scénarios spéciaux de classification, qui sont [41] :

2.8.2.1. Apprentissage de CNN à partir de zéro

L'architecture CNN est construite à partir de zéro en utilisant des couches convolutif ; couches Pooling ; couches entièrement connectées ; couche de correction (ReLU). Cependant, la conception du réseau, c'est-à-dire le choix du nombre de couches et du nombre de paramètres dans chaque couche est une tâche difficile. C'est parce qu'il n'y a pas de définition standard et dépend complètement de nombreux des facteurs, tels que les caractéristiques de l'image d'entrée (par exemple, la taille, la complexité, etc.), la complexité de l'application, le nombre de classes et le nombre d'échantillons d'apprentissage disponibles. En général, les ensembles de couches et leurs hyper-paramètres conviennent à des applications définies empiriquement sur la base de l'ensemble de données. Cette approche nécessite une quantité de données d'apprentissage pour éviter le sur-apprentissage, et ne peut donc être adoptée que lorsqu'une grande quantité d'un échantillon d'apprentissage est disponible.

2.8.2.2. Réglage d'un modèle CNN pré-entraîné

Obtenu en apprenant des fonctionnalités spécifiques à l'application en adoptant un modèle pré-entraîné qui a déjà été entraîné à l'aide d'un grand nombre d'échantillons (généralement des millions) de domaines connexes, et en ajustant leurs poids à l'aide d'échantillons d'entraînement de l'application étudiée celles-ci.

Les types de modèles CNN pré-entraînés sont de plus en plus disponibles (MatConvNet, 2016). De plus, l'adaptabilité de ceux-ci sont des modèles pré-entraînés pour effectuer un transfert d'apprentissage à distance.

En adoptant le modèle pré-entraîné de CNN pour l'apprentissage par transfert, il est nécessaire de modifier la couche finale entièrement connectée, dont la dimension doit être égale au nombre de classes spécifié pour la classification. La structure du réseau peut également être modifiée en ajoutant ou en supprimant des couches, si nécessaire. Cependant, dans la plupart des cas, la structure du modèle pré-entraîné est largement préservée car les caractéristiques qu'ils ont apprises sont souvent plus générales et contiennent des informations importantes pour toute application de classification d'images, y compris la télédétection. À cette fin, le modèle est modifié avant l'apprentissage de plusieurs manières et

évalué simultanément (c'est-à-dire en supprimant ou en ajoutant des couches et une évaluation simultanée à l'aide de nos ensembles de données). Le modèle prévoyait que des résultats supérieurs sont pris en charge ici.

2.8.2.3. *Modèles pré-entraînés sans réglage*

Dans cette approche, un modèle CNN pré-entraîné sert d'outil d'extraction de caractéristiques, dans lequel la sortie de l'une des couches entièrement connectées dans une architecture CNN est considérée comme la représentation des caractéristiques d'un patch d'image donné. Ceci est basé sur la supposition sur la généralisation des caractéristiques CNN apprises par un modèle CNN pré-entraîné utilisant un grand nombre d'images génériques avec des propriétés très variables. Les caractéristiques extraites de la couche entièrement connectée peuvent être utilisées dans le processus de classification en la combinant avec n'importe quel classificateur comme SVM.

2.9. Conclusion

Dans ce chapitre, nous avons expliqué les concepts les plus importants liés au *deep learning* (tels que sa définition, ses structures ... etc.) et nous avons également expliqué les réseaux CNN et certains de leurs modèles, que nous avons adoptés dans cette recherche. Dans le prochain chapitre, nous traiterons des détails les plus importants de la conception, ainsi que des méthodes que nous adopterons et des outils utilisés pour réaliser leur application.

Chapitre 3

Classification d'images médicales par ML et DL

3.1. Introduction

Ces dernières années, avec le développement technologique dans le domaine de l'intelligence artificielle. Des méthodologies d'apprentissage automatique (ML), d'apprentissage en profondeur (DL) et de traitement d'image ont été utilisées pour améliorer la précision de détection et donner de meilleurs résultats. Elles sont également adoptées pour classer les images médicales (qu'il s'agisse d'une personne malade ou en bonne santé). En conséquence, l'ordinateur CAD (pour l'anglais *Computer-Aided Design*) est devenu le composant principal du diagnostic médical, car il a formé une deuxième opinion pour aider le médecin à prendre sa décision de classer (radio d'imagerie médicale).

L'objectif principal de ML et DL est le besoin d'intelligence et d'apprentissage supplémentaire pour traiter les opinions non confirmées afin que les résultats de l'éternuement soient plus précis et cohérents. Nous utilisons ce domaine d'apprentissage pour des raisons impératives, qui sont d'aider à la prise de décision, fournir de la précision et une réponse immédiate dans les situations d'urgence, et réduire les efforts des médecins, car la performance des systèmes médicaux est liée à la CAD par ordinateur pour sa force à distinguer et précision dans le processus de classification des images concernées pour l'individu.

Dans ce chapitre, nous expliquerons les éléments les plus importants grâce auxquels nous appliquerons ML et DL pour aider à créer un système intelligent qui permet aux médecins de donner des résultats précis.

3.2. Imagerie médicale

Les images médicales sont devenues une pierre clé dans le diagnostic et le suivi de presque toutes les maladies. Ces images offrent aux médecins la possibilité de participer à l'examen et au traitement des patients.

L'imagerie médicale est certainement l'un des domaines de la médecine qui a le plus progressé ces vingt dernières années. Ces récentes découvertes permettent non seulement un meilleur diagnostic mais offrent aussi de nouveaux espoirs de traitement pour de nombreuses maladies comme le cancer, épilepsie, tuberculose, etc. L'identification précise de la lésion

facilite déjà le recours à la chirurgie. De telles techniques permettent également de mieux comprendre le fonctionnement de certains organes encore mystérieux, comme le cerveau.

L'image médicale est une représentation discrète d'une partie du corps qui est décrite par une matrice à 3 dimensions [42][43].

3.3. Apprentissage profond pour l'analyse d'images médicales

Les chercheurs estiment que les images médicales représentent actuellement au moins 90% de toutes les données médicales, ce qui en fait la plus importante source de données dans l'industrie des soins de santé. Cela devient une quantité écrasante à l'échelle humaine, quand on considère que les radiologues dans les salles d'urgence de certains hôpitaux reçoivent quotidiennement des milliers d'images. De nouvelles méthodes sont donc nécessaires pour extraire et représenter les données de ces images plus efficacement.

Bien que l'une des applications les plus courantes d'apprentissage automatique dans le domaine médical ait été en réalité l'imagerie médicale, ce n'est que récemment que des algorithmes d'apprentissage profond ont été introduits.

Le tableau 01 ci-dessous présente des résultats de certaines applications de DL dans l'analyse d'images médicales. La première colonne contient les tâches traitées, la seconde contient les noms des auteurs ayant réalisé le travail en question. La troisième colonne présente la performance atteinte par les différents travaux en termes de précision.

Tableau 02 : Résultats de certaines applications de DL dans l'analyse d'images médicales [44].

	Auteurs de travaux	Précision (%)
Rétinopathie diabétique	Kathirvel	96
	Pratt et al.	75
	Haloï et al.	98
Détection d'éléments histologiques et microscopiques	Quinn et al.	100 pour le paludisme 99 pour la tuberculose
	Dong et al.	98.66
Alzheimer et Parkinson	Suk	95.35
	Sarraf et al.	98.84
	Liu et al.	87.76

3.4. Apprentissage automatique pour aider au diagnostic médical du cancer de la peau

3.4.1. Cancer de la peau

Le cancer de la peau est un type de tumeur maligne qui attaque les cellules de la peau. Certains types de cancer peuvent être classés comme bénins ou superficiels, mais ils doivent être surveillés en permanence. Le cancer de la peau peut être causé par une exposition excessive aux rayons ultraviolets, qu'il s'agisse de rayons naturels (le soleil) ou artificiels (lits de bronzage).

Il existe trois types de cancer :

Le *mélanome* (également appelé mélanome malin) : ce type commence dans les mélanocytes et se propage rapidement à certaines parties du corps s'il n'est pas diagnostiqué et traité tôt. Le cancer de la peau peut commencer par une taupe [45].

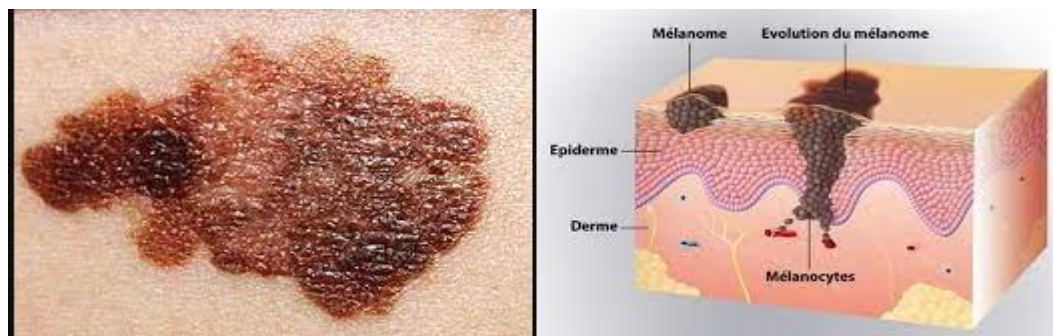


Figure 10 : Exemples d'images dataset d'apprentissage du mélanome.

Le *carcinome bas cellulaire* (CBC) : Il s'agit du type de cancer de la peau le plus courant, autre que le cancer de la peau. Cela commence par la couche basale, qui est la couche la plus profonde de l'épiderme. Il se développe généralement dans les zones exposées au soleil telles que la tête et le cou. Les cellules sanguines entières se développent lentement et se propagent à peine aux autres parties du corps.



Figure 11 : Exemples d'images dataset d'apprentissage du CBC.

Le *carcinome spin cellulaire* (CSC) ou carcinome squameux : ce type commence dans les cellules squameuses - des cellules minces et plates qui ressemblent aux écailles de poisson et se trouvent dans la peau et les tissus tapissant les parties du corps. Ces cancers ne se propagent parfois qu'à d'autres tissus. Le CCN est la deuxième forme la plus courante de cancer de la peau. [45][46]

Les méthodes de détection du cancer de la peau sont la photographie, la dermatoscopie et l'échographie.

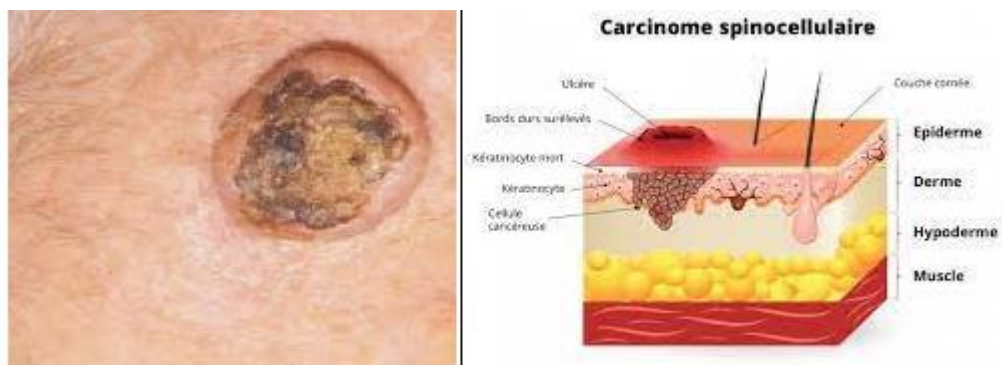


Figure 12 : Exemples d'images dataset d'apprentissage du CSC.

3.4.2. Photographie pour vérifier et détecter le cancer de peau

Photographie du cancer de la peau est ensemble des techniques d'enregistrement de rayonnements électromagnétiques par des procédés photochimiques. Il est conçu pour détecter les anomalies précocement avant qu'elles ne provoquent des symptômes cliniques. La dermatographie n'est pas seulement pratiquée dans les campagnes de dépistage du cancer de la peau, mais elle est également pratiquée pour le diagnostic ainsi que la localisation lors d'interventions chirurgicales telles que la transplantation cutanée.

L'avantage de cet examen est qu'il permet d'examiner toutes les cellules de la peau, le radiologue examine les images cutanées pour rechercher les signes d'une tumeur maligne ou bénigne en observant, comme s'il y avait l'un des objectifs suivants

Une grande tache brune avec une tache plus foncée, une taupe qui change de couleur, de taille ou de sensation, ou qui saigne petite lésion avec des bords irréguliers et des parties qui apparaissent rouges, des lésions sombres sur la paume des mains, la plante des pieds, le bout des doigts ou des orteils, ou sur les muqueuses qui tapissent la bouche ou le nez, et d'autres symptômes.

L'interprétation et l'analyse d'images médicales est l'un des processus les plus difficiles dans les applications avancées de reconnaissance de formes et de vision par ordinateur. En photographie des cancers cutanés, la principale difficulté est de faire la distinction entre les tumeurs bénignes et malignes.

Dans le cas de la photographie des cellules de la peau, le patient sera par conséquent exposé à plusieurs tests coûteux et incontournables.

En ce sens, les chercheurs ont exploité les techniques de ML pour tirer parti des limites de ces technologies traditionnelles grâce à l'assistance des médecins avec un deuxième avis, réduisant ainsi les erreurs humaines potentielles.

3.4.3. Détection et classification assistées par ordinateur (CADe/CADx) des photographiques cancer de peau

Les techniques d'apprentissage automatique et les applications de traitement d'image ont contribué à la tâche de diagnostic médical et d'identification des maladies de la peau, car elles ont abouti à des diagnostics plus efficaces et plus précis en termes de résultats. Par conséquent, dans le passé, il y avait un grand intérêt pour le développement d'un système de détection et de diagnostic assisté par ordinateur (CADe pour la détection et CADx pour le diagnostic).

C'est la deuxième opinion que les photologies ont lors de l'interprétation des images photographiques des cellules de la peau.

Les systèmes d'assistance au diagnostic médical (CAD) ont été développés à l'aide de méthodes de ML et DL.

Nous nous appuyons également sur l'analyse des images pour identifier la maladie à travers la structure de l'image afin de déterminer si cette peau est normale ou anormale. Nous déterminons également si la peau est maligne ou bénigne.

La CAD est conçue par des développeurs à des fins de traitement d'image, et elle s'appuie sur les caractéristiques des techniques d'extraction pour distinguer différentes structures en les traduisant en valeurs numériques [47].

3.4.4. Détection de tumeur assistée par ordinateur (CADe)

La détection des tumeurs assistée par ordinateur (CADe) est la première étape essentielle du diagnostic des radiographies cutanées ou cutanées.

Les algorithmes couramment appliqués pour l'extraction de la zone d'intérêt (ROI) sont basés sur l'analyse des pixels. Pour que les pixels qui composent une texture satisfaisante aient des descripteurs différents pour les autres pixels qui composent la texture sans aucune blessure. Les descripteurs utilisés pour détecter une tumeur peuvent être constitués de valeurs de gris ou de mesures calculées par tissu.

Dans le cas où cette peau est détectée comme contenant une blessure, elle est entrée dans l'étape de classification basée sur l'apprentissage automatique du système d'assistance au diagnostic médical. Cancer de la peau (CAD).

3.4.5. Diagnostic des tumeurs assisté par ordinateur (CADx)

Le diagnostic tumoral assisté par ordinateur (CADx) caractérise les lésions détectées par un spécialiste en imagerie ou par CADe. La sortie de ces systèmes est Classer le type de lésion ou de lésion selon qu'elle est maligne ou bénigne.

Les systèmes de diagnostic assisté par ordinateur (CADx) effectuent une extraction de descripteur distincte dans des régions suspectes segmentées. Ensuite, les descripteurs extraits sont utilisés comme indicateurs d'entrée pour un classificateur, et le modèle prédictif est formé en ajustant les poids des différents descripteurs en fonction des propriétés statistiques du classificateur. Ensemble d'apprentissage pour estimer la probabilité qu'une zone d'intérêt appartienne à une classe.

3.5. Quelques travaux récents utilisant ML et/ou DL

Dans cette section, nous décrivons quelques travaux antérieurs visant à la classification du cancer de la peau.

A) Etude dans [45] :

Technique d'apprentissage : Apprentissage profond avec CNN

DataSet : Images de peau (5846 images).

Méthode et résultats : Dans cet article, l'auteur a mené une étude afin de classer les types de cancer de la peau en termes d'exactitude, de sensibilité et de spécificité, où la classification est représentée en six classe à travers lesquelles nous définissons ce cancer de la peau à quelle classe il appartient (MM, BCC, SK, Nevus, H/H, SL) Ce diagnostic a été posé par région. Concernant la formation et le test : Six classe ont été suivis pour chacun des FRCNN et des Dermatologues (BCD et TRN), et les résultats ont été représentés comme suit (200 questions * 10 tests), les taux de précision pour classer les six classes FRCNN étaient de 86,2% et pour la classification des BCD, il était de 79,5%. Pour l'industrie TRN, le taux d'exactitude était de 75,1%.

Grâce à cette comparaison, il a été constaté que FRCNN avait le taux de précision le plus élevé des dermatologues, car l'écart type des catégories FRCNN était de 2,80 %, tandis que pour les dermatologues, l'écart type était de 4,41 %. La précision de la classification FRCNN était statistiquement significativement plus élevée que celle des BCD.

Performance : $ACC(FRCN) = 91.5\%$.

B) Etude dans [46] :

Technique d'apprentissage : Apprentissage profond avec Algorithme de gradient aléatoire (SGD).

Data Set : Dermatoscopique (10015 images) ; Mélanome (Mel, 1113 échantillons), naevus mélanocytaires (NV, 6705 échantillons), carcinome basocellulaire (BCC, 514 échantillons), kératose actinique et carcinome intraépithélial (AKIEC, 327 échantillons), kératose bénigne (BKL, 1099 échantillons), dermato-fibrome (DF, 115 prélèvements), lésions vasculaires (VASC, 142 prélèvements).

Méthode et résultats : Dans cet article, un modèle robuste est proposé pour améliorer le diagnostic et l'identification du mélanome à l'aide d'une caméra standard. L'expérience a été

menée entre un ensemble d'images formées à l'aide de modèles CNN et une par des dermatologues. L'architecture de ces modèles a été recyclée sur son jeu de données, optimisant toutes les couches et remplaçant les couches supérieures par un milieu de refroidissement, deux entièrement connectées et enfin une couche *softmax* qui permet 8 classes de diagnostic. Le taux d'apprentissage a été fixé à 0,0007 pour le modèle Inception v3. Dans le cas des autres modèles groupés, le taux d'apprentissage était de 0,0006 ; en comparant les résultats, dans le cas de la validation du mélanome, un taux de précision de 94,4% a été atteint avec le modèle ResNet 152, contre 82,26% pour le dermatologue.

Dans le cas du carcinome bas cellulaire, il a atteint le niveau le plus élevé dans le modèle DenseNet 201 avec 99,3 % par rapport aux dermatologues qui avaient une précision de 88,82 %. Dans ce dernier, il s'est avéré qu'il y avait une différence significative entre les résultats adoptés par les modèles CNN qui avaient de meilleurs résultats que ceux obtenus par les dermatologues.

Performance : Norme ROC AUC = 99.3%.

C) Etude dans [47] :

Technique d'apprentissage : Apprentissage profond Studio DLS (Deep learning studio) avec Algorithme de ECOE SVM

Data Set : Cancer de la peau (129450 images cliniques) dont (3374 images) dermatoscopiques

Méthode et résultats : dans cette article, Deux expériences ont été menées pour calculer le taux de précision de la classification du cancer de la peau, car le MODÈLE 1 a été conçu à l'aide de réseaux convolutifs bidimensionnels et ce modèle a été développé en utilisant ECOE SVM avec AlexNet L'apprentissage en profondeur, tandis que le reste des modèles utilisait pré-réseaux CNN formés Une norme générale représente une base de données pour l'apprentissage automatique, En ce qui concerne le MODÈLE 1, un tableau de bord a été utilisé qui montre les progrès, les changements de précision et les temps d'achèvement pendant l'entraînement, car il a la capacité de s'arrêter s'il y a une erreur dans la course.

En comparant les résultats des deux expériences qui ont été établies à l'aide de DLS avec de MODÈLE 1 et HAM10000 pour le reste des modèles, il a été constaté que la valeur ROC plus élevée indique que le modèle de classificateur diagnostique rarement le cancer à tort (cancer / non-cancer). Le DLS créé a une résolution de 96%, ce qui est supérieur aux autres résultats,

qui ont une résolution de 95,1%. Après cette étude, il est devenu clair que les modèles DLS ont atteint une AUC de 99,77% dans la détection des cellules cancéreuses à partir d'images de cellules cancéreuses.

Performance : AUC = 99.77%.

3.6. Conclusion

L'apprentissage automatique est un domaine de l'intelligence artificielle qui gagne en importance dans la recherche sur les tumeurs cutanées. En effet, ce domaine nécessite un nombre considérable d'observations et de connaissances et ce sujet est désormais abordé par un nombre croissant d'initiatives de protection et de stockage des données médicales.

Le défi principal est d'identifier et de créer des modèles à partir de ces données et d'utiliser ces modèles pour faire des prédictions sur des données nouvelles et non découvertes.

L'apprentissage automatique est également un domaine en pleine croissance avec une pléthore d'approches parmi lesquelles choisir. Dans ce chapitre, nous expliquons le rôle de l'apprentissage profond et automatique dans le traitement du cancer de la peau. Dans le chapitre suivant, nous clarifierons la problématique dessinée dans ce domaine en plus du cadre méthodologique de l'apprentissage automatique utilisé pour réaliser l'approche proposée.

Chapitre 4

Hybridation de caractéristiques de DL pour améliorer la classification du cancer de peau

4.1. Introduction

Ce dernier chapitre est consacré à la présentation de notre proposition pour améliorer la classification des images médicales, ainsi que les résultats de validation que nous obtiendrons. Dans un premier temps, nous présenterons le problème abordé et le jeu de données du cancer de la peau. Par la suite, nous exposerons la structure générale du système proposé, puis nous détaillerons chacune de ses deux phases principales : phase d'extraction des caractéristiques et la phase de classification. Ensuite, nous décrirons l'environnement de travail et les différents outils utilisés. Enfin, nous présentons notre étude expérimentale, puis analysons et interprétons les résultats numériques obtenus.

4.2. Problématique étudié

L'apprentissage automatique (ML) et l'apprentissage profond (DL) sont deux branches de l'intelligence artificielle (IA). L'apprentissage automatique offre une myriade d'opportunités pour améliorer et souligner l'importance du diagnostic médical intelligent dans le domaine de la médecine en général. Ceci, grâce à une analyse objective et approfondie des données médicales, afin d'identifier de nouveaux schémas et facteurs de risque à long terme pour aider au diagnostic. En plus, L'IA, utilisant des systèmes de santé commercialisables, peut fournir une solution pour les traitements appropriés pour chaque patient.

Pour l'analyse d'images de peau, nous adoptons l'extraction de descripteurs comme première étape, étant donné l'importance de cette étape lors de l'application de techniques d'apprentissage automatique.

Les formes malignes et saines de la tumeur sont importantes, selon le radiologue, car elles reposent sur la règle selon laquelle plus le bord irrégulier de la tumeur est grand, plus elle est observée en rapport avec une tumeur maligne. De plus, ces *cellules malignes* sont appelées cellules cancéreuses, et en cas d'irritation cutanée, elles sont transférées vers d'autres organes du corps, ce qui augmente le risque d'infection. C'est pour les cellules malignes. Quant aux *cellules bénignes*, elles sont de forme bien définie et grandes par rapport à la taille des cellules malignes, et elles sont très petites.

Les symptômes du cancer de la peau varient selon le type de tumeur et le degré de sa progression. Pour le mélanome : une lésion pigmentée apparaît sur la peau qui est irrégulière

en bords et en surface. De petites bosses ou des ulcères peuvent apparaître. Lésions cutanées pigmentées bénignes de couleur uniforme et de couleur principalement claire. Le diamètre d'une lésion pigmentaire maligne est supérieur à 6 mm car les lésions plus petites sont généralement bénignes et le carcinome épidermoïde de la peau est l'apparition d'une masse de couleur claire sur la peau. Des cloques peuvent apparaître avec des saignements dans les zones de la peau les plus exposées au soleil, et ce type de cancer est moins susceptible de se propager et varie de 0,3% à 5%. Symptômes du mélanome basal : plaies pigmentées ou décolorées qui provoquent généralement des démangeaisons.

Différents descripteurs de forme et de texture ont été proposés pour classer les bosses détectées sur les images cutanées comme malignes ou bénignes. Une bonne classification est directement liée aux caractéristiques des descripteurs et à la capacité de distinguer les images dans ces deux catégories. Par conséquent, les chercheurs sont toujours à la recherche de descripteurs simples qui a la capacité de distinguer les items à calculer, ce qui contribue de manière significative à la classification des performances.

Compte tenu de ce que nous avons rencontré dans ce mémoire, nous avons étoffé notre problème de recherche dans le but ultime de contribuer à une technique puissante qui nous permet d'extraire des caractéristiques et de combiner des modèles entre eux pour parvenir à une classification du type de cancer (bénin/malin) avec précision.

4.3. Data Set

Dans les milieux cliniques, un nombre limité de grains de beauté sont généralement obtenus à l'aide de la photographie et sont disponibles sous forme de diapositives 2D. L'ensemble de données d'imagerie a été obtenu par le développeur Claudio Fanconi, candidat à la maîtrise en technologies de l'information et génie électrique (EET) de l'ETH Zürich et formé en tant que chercheur chez IBM [48].

Nous avons collecté 2637 diapositives auprès d'un certain nombre de patients, contenant un ensemble d'images didactiques réparties comme suit : 1197 malins et 1440 bénins. Un autre ensemble de test est réparti comme suit : 300 malins et 360 bénins. Les images sont aux dimensions de 224 * 224. La taille est de 163 Mo.

La figure 13 contient un ensemble de données équilibré d'images de grains de beauté bénins et malins.

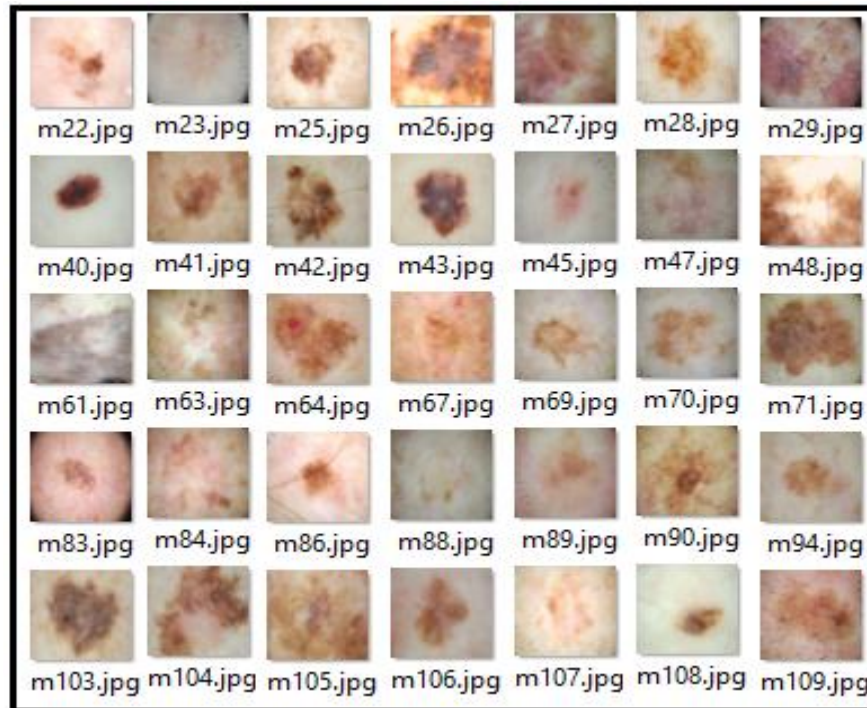


Figure 13 : Exemple de Data Set.

4.4. Méthode proposée

Le schéma de la figure 14 présente une conception générale de notre méthode proposée, qui consiste en plusieurs étapes par lesquelles un résultat est atteint, qui est la classification du cancer de la peau, en extrayant des vecteurs de caractéristiques de plusieurs architectures CNNs et en intégrant toutes les caractéristiques CNNs pour produire une représentation complète de l'image d'entrée. Dans l'étape finale, le vecteur de caractéristiques combiné (ou *géant*) est transmis à un classificateur supervisé de type SVM pour obtenir le résultat de la classification.

Dans notre travail, nous avons utilisé l'algorithme **SVM** car c'est l'algorithme de classification le plus puissant et le plus connu dans le domaine de l'apprentissage automatique. Nous avons également utilisé les modèles de CNN suivant : **VGG16**, **VGG19**, **Inception V3** et **RasNet**, car ils avaient la plus grande précision dans **ImageNet**.

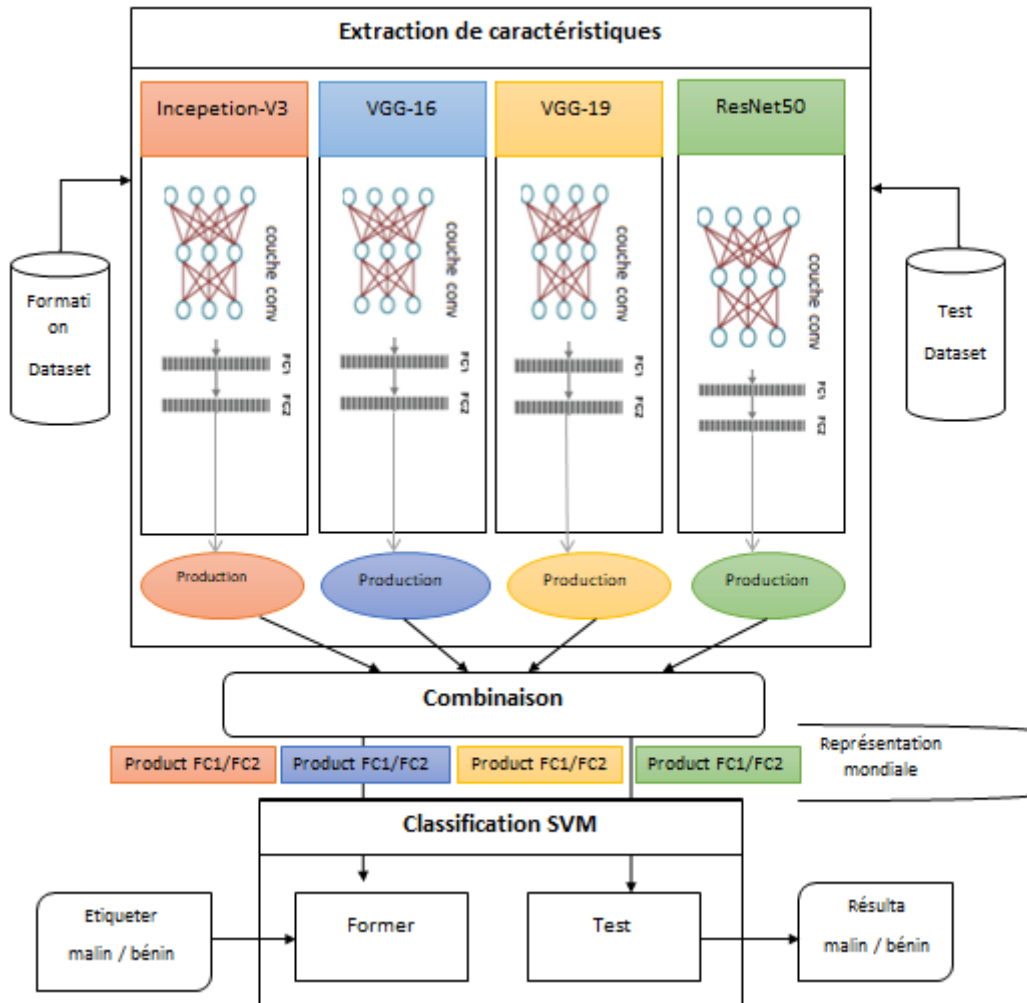


Figure 14 : Schéma général de notre méthode de classification du cancer de la peau.

4.4.1. Classification de maladies grâce à des méthodes d'apprentissage automatique

Les méthodes utilisant des techniques d'apprentissage supervisé sont avérées utiles pour détecter la maladie. Ces méthodes se composent de deux unités principales, à savoir le module d'extraction de caractéristiques et le module de classification. Sur la base des fonctionnalités extraites, le module de classification est conçu pour utiliser un classificateur existant (par exemple un SVM) pour séparer la classe en un patient ou un non-patient

A montré que les CNNs ont la capacité d'extraire des caractéristiques profondes (ou on dit de DL) par rapport au diagnostic d'un clinicien de cancer de la peau pour faire la distinction entre malin et bénin. Les activations intermédiaires apprises à l'aide de réseaux profonds CNN préalablement formés sur de grands ensembles de données : tâches de

classification avec des données d'entraînement limitées, y compris la classification de la maladie.

4.4.2. Représentations hybrides de caractéristiques de DL pour améliorer la classification de maladies

Dans cette étude, nous développons un système pour la construction d'un modèle éducatif supervisé pour la classification des cancers à l'aide de caractéristiques de DL hybrides. L'objectif principale est de tester la possibilité de combiner plusieurs vecteurs de caractéristiques extraits par différents modèles CNN afin d'améliorer les performances. Cette idée a été originalement proposé et validée par Settou et *al.* [32], des auteurs de département d'informatique à l'université de El Oued, dans pour l'amélioration de la classification de dommages. Notre objectif dans ce travail est de réappliquer cette idée dans le contexte de classification d'images médicales. Le système proposé se compose de deux phases principales :

- 1) La **phase d'extraction des caractéristiques** qui est cruciale dans le procédé de notre architecture et s'exécute basé sur une collection de modèles CNN ;
- 2) La **phase de classification des maladies** qui est implémenté par un classificateur SVM afin de séparer les images d'entrée en deux classes différents malin et bénin.

La figure 01 montre un schéma général de notre méthode proposée. Les vecteurs de caractéristiques sont d'abord extraits de plusieurs architectures CNN. Ensuite, toutes les caractéristiques CNN sont combinées pour produire une représentation globale de l'image d'entrée. En dernière étape, le vecteur de caractéristiques combiné (ou *géant*) est passé au classificateur supervisé SVM pour obtenir le résultat de classification.

4.4.2.1. Phase d'extraction de caractéristiques

L'extraction de caractéristiques est le processus de mappage des pixels de l'image dans une entité adaptée à l'espace. Le choix d'une méthode d'extraction de caractéristiques appropriée affecte fortement les performances du classificateur. Etant donné que les images médicales sont très complexes et très diverses dans leurs propriétés, la méthode d'extraction des caractéristiques joue un rôle très important dans l'amélioration des performances de classification. Plus précisément, les performances de classification des maladies dépendent principalement de la puissance de la méthode d'extraction de caractéristiques.

Notre objectif est d'examiner des approches d'extraction de caractéristiques hybrides pour la classification des maladies. Nos approches hybrides s'appuient sur des représentations des caractéristiques de plusieurs modèles CNN pré-entraînés. Dans ce paragraphe, nous expliquons d'abord comment nous avons extrait le vecteur de caractéristiques d'une image d'entrée de chaque modèle CNN utilisé. Ensuite, nous décrivons notre approche de construction de caractéristiques hybrides.

A- Caractéristiques extraites d'un modèle CNN particulier

Pour l'extraction des caractéristiques des images d'entrée, nous avons évalué quatre modèles CNN, qui sont : Resnet, Inception V3, VGG-16, VGG-19. Ces modèles ont été utilisés avec succès dans de nombreuses applications. Nous avons utilisé ces modèles CNNs comme des outils pour extraire des caractéristiques en exploitons la dernière couche entièrement connectée (ou FC pour l'anglais *full connected*) qui est ensuite insérée dans la fonction d'activation pour la classification (généralement un vecteur d'activation de l'une des deux premières couches complètement connectées (FC1 ou FC2) est utilisé directement comme une représentation finale des caractéristiques de l'image).

Ensuite, nous avons appliqué l'activation FC1 et FC2 pour un modèle particulier, comme illustré dans la figure 03. Après cela, nous combinons ces deux vecteurs de caractéristiques pour obtenir un vecteur de caractéristiques de chaque modèle CNN donné.

Pour chaque image d'entrée, nous obtenons quatre vecteurs de caractéristiques de tailles suivantes : 3048, 5096, 5096 et 3048 pour Inception, VGG-16, VGG-19 et ResNet, respectivement.

Enfin, chaque image d'entrée est envoyée en un vecteur de caractéristiques de spécifique taille.

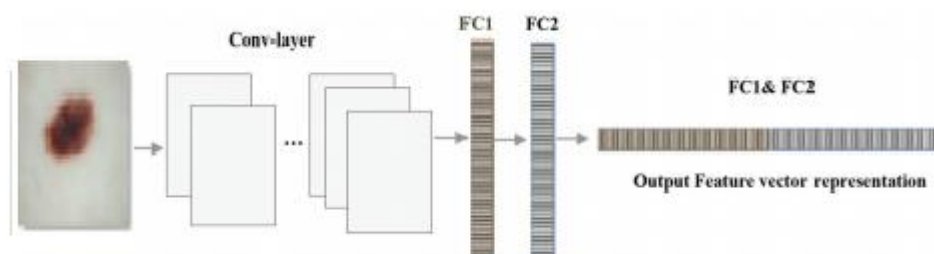


Figure 15 : Vecteur de caractéristiques extrait de la 1ère et de la 2ième couches FC.

B- Représentation hybride proposée

Le vecteur de caractéristiques final est un regroupement de caractéristiques de DL extraites pour chaque image d'entrée par les modèles suivants : Inception, VGG-16, VGG-19 et ResNet. Nous obtenons des vecteurs combinés de différentes tailles en fonction du nombre de vecteurs initiaux regroupés comme indiqué dans le tableau 03 ci-dessous.

Après cette étape, chaque vecteur hybride est passé à un SVM pour la classification.

Tableau 03 : Tailles des vecteurs de caractéristiques DL combinées.

Variante	Modèles CNN combinés	Taille du vecteur de caractéristiques DL de modèles combinés		
		Somme de tailles de vecteurs FC1	Somme de tailles de vecteurs FC2	Somme de tailles de vecteurs FC1&FC2
V1-1	VGG16	1000	4096	5096
V1-2	VGG19	1000	4096	5096
V1-3	Inception V3	1000	2048	3048
V1-4	RasNet	1000	2048	3048
V2-1	VGG16 + VGG19	2000	8192	10192
V2-2	VGG16 + Inception V3	2000	6144	8144
V2-3	VGG16 + RasNet	2000	6144	8144
V2-4	VGG19 + Inception V3	2000	6144	8144
V2-5	VGG19 + RasNet	2000	6144	8144
V2-6	Inception V3 + RasNet	2000	4096	6096
V3-1	VGG16 + VGG19 + Inception V3	3000	10240	13240
V3-2	VGG16 + VGG19 + RasNet	3000	10240	13240
V3-3	VGG16 + Inception V3 + RasNet	3000	8192	11192
V3-4	VGG19 + Inception V3 + RasNet	3000	8192	11192
V4	VGG16 + VGG19 + Inception V3 + RasNet	4000	12288	16288

4.4.2.2. Phase de classification des maladies

Après l'extraction des caractéristiques et la construction des vecteurs hybrides, l'étape suivante est la classification. Dans cette l'étape, le vecteur de caractéristiques hybride est

passé à un classificateur supervisé. Étant donné que ce vecteur de caractéristiques hybride à des dimensions élevées, le classificateur utilisé doit être capable d'adapter aux données d'entrée hyper-dimensionnelles. Dans cette étude, nous avons appliqué un classificateur SVM pour sa capacité à manipuler des espaces de caractéristiques de grandes tailles.

Un SVM linéaire sont formé à l'aide d'images infectées et non infectées. En fait, la classification des types de cancer par SVM se compose de deux phases : apprentissage et test. Dans la phase d'apprentissage, le SVM prend les vecteurs hybrides de caractéristiques CNNs avec les étiquettes d'image d'entrée correspondantes (infectées et non infectées). Dans la phase de test, le SVM ne prend que le vecteur hybride des caractéristiques sans aucune étiquette et le compare avec les modèles existants dans la base des modèles obtenue à partir de la phase d'apprentissage pour faire la classification.

4.5. Présentation des outils

Dans cette section, nous exposons les différents outils logiciels qui nous avons utilisées pour réaliser nos modèles et expériences.

4.5.1. Logiciel

4.5.1.1. *TensorFlow*

Il s'agit d'une bibliothèque open source compatible Python pour les comptes numériques à grande échelle, créée par l'équipe Google Brain qui rend DL et ML plus rapides et plus faciles. TensorFlow peut entraîner et exploiter des réseaux neuronaux profonds pour la reconnaissance d'images, des réseaux neuronaux répétitifs, la classification des nombres et des modèles séquence à séquence pour la traduction automatique [49].

4.5.1.2. *Keras*

Keras est une bibliothèque de logiciels open source qui fournit une interface Python pour les réseaux de neurones artificiels. Il agit comme une interface pour la bibliothèque TensorFlow. Nous incluons également notre modèle CNN (Inception V3 ; ResNet ; VGG16 ; VGG19).

4.5.1.3. *Framework*

Il est écrit et entretenu par Francis Chollet, un autre membre de l'équipe Google Brain. Il permet aux utilisateurs de choisir si les modèles qu'ils construisent sont exécutés sur Theano ou TensorFlow. Conçu pour permettre une expérimentation rapide avec des réseaux de neurones profonds, il se concentre sur le fait d'être convivial, modulaire et extensible [49].

4.5.1.4. *SK learn*

Scikit-learn est une bibliothèque d'apprentissage automatique de logiciels gratuits pour le langage de programmation Python [3]. Il comporte divers algorithmes de classification, de régression et de clustering, y compris SVM, et est conçu pour interagir avec les bibliothèques numériques et scientifiques Python NumPy et SciPy. Dans cet algorithme, nous appliquons plusieurs algorithmes chacun avec une fonction spécifique [49].

4.5.1.5. *Open CV*

Open CV a été conçu pour fournir une infrastructure commune pour les applications de vision par ordinateur et pour accélérer l'utilisation de la perception de la machine dans les produits commerciaux. En tant que produit sous licence BSD, Il permet aux entreprises d'utiliser et de modifier le code [50]. Nous utilisons cet algorithme pour lire et traiter l'image.

4.5.1.6. *Jupyter notebook*

Il s'agit d'une application Web open source qui permet de créer et de partager des documents contenant du code en direct, des équations et des visualisations avec Python. Ses utilisations sont de nettoyer les données et de les convertir sous forme de lots pour exécution. Nous avons choisi *Jupyter* au lieu de *Pycharm*, ce qui fait que l'avantage qu'il contient, qui est l'exécution par lots, limite le temps de correction de l'erreur qui peut survenir lors de l'exécution.

4.5.2. **Matériel**

Nous avons implémenté notre approche et réalisé nos expériences sur un ordinateur DELL, sous le système d'exploitation : Windows 10 Pro 64 bits, Processeur : Intel(R) Core(TM) i5-3337U CPU @ 1.80 GHz et mémoire installée (RAM): 4096 MB.

4.5.3. Langages et outils de développement

4.5.3.1. Langage Python

Python est un langage de programmation interprété de haut niveau et un programme orienté objet développé par Python Software Foundation ; Sa mise en œuvre a commencé en décembre 1989 à être créé par Guido van Rossum.

Python a plusieurs versions de Python 2.x ou Python 3.x. Nous avons utilisé la dernière version de Python 3.6.7 [51].

4.5.3.2. Pourquoi python

Il est facile de travailler avec et d'accéder à des bibliothèques de haute qualité. C'est le plus utilisé dans le domaine de l'intelligence artificielle et de l'apprentissage automatique. Pandas, image-Net, pascal et mlxtend tout supporte Python. Ses interprètes sont disponibles pour de nombreux systèmes d'exploitation. C'est-à-dire le plus des bibliothèques célèbres dans le domaine de l'intelligence artificielle sont fabriquées par Python comme cv2 et Scikit-learn. Guides Scikit-learn pour utiliser le langage Python [51].

4.6. Explication du code

Tout d'abord, nous avons commencé à coder pour extraire les propriétés de l'image.

- Nous prenons une image puis la convertissons en une liste unidimensionnelle.
- Nous extrayons les caractéristiques de chaque modèle.
- Nous le regroupons dans une liste finale avec son étiquette.

```
for img in tqdm(os.listdir(TRAIN_DIR)):
    if ('b' in img):
        label_ = 0 #benign
    else:
        label_ = 1 #malignant
    l=0
    path = os.path.join(TRAIN_DIR+'\\', img)
    M = []
    for i in models: # Models [InceptionV3, ResNet50, VGG16, VGG19]
        if (l==0):
            image = load_img(path, target_size=(299, 299)) # InceptionV3
        else:
            image = load_img(path, target_size=(224, 224)) # ResNet50 , VGG16 , VGG19
        l=l+1
        image = img_to_array(image)
        image = image.reshape((1, image.shape[0], image.shape[1], image.shape[2]))
        image = preprocess_input(image)
        model_extractfeatures_1 = Model(i.input, i.layers[-1].output)
        model_extractfeatures_2 = Model(i.input, i.layers[-2].output)
        fc1_features = model_extractfeatures_1.predict(image)
        fc2_features = model_extractfeatures_2.predict(image)
        M = np.append(M ,fc2_features[0])
        M = np.append(M ,fc2_features[0])
    M = np.append(M,label_)
```

Figure 16 : Extraction de caractéristiques.

Deuxièmement, que se passe-t-il au niveau SVM ?

- Les caractéristiques extraites du code précédent sont divisées en deux parties, pour les deux phases formation et test ;
- Le modèle est entraîné à l'aide de SVM.
- Obtention d'un pourcentage (%) dans l'entraînement et le test (train et test).
- Enfin, le model est enregistré.

```
X_train, X_validation, Y_train, Y_validation = model_selection.train_test_split(X, Y, test_size=0.25) #split
Model_SVM = SVC(C=80)
Model_SVM.fit(X_train, Y_train) # Training
y_pred = Model_SVM.predict(X_train)
train_accuracy = accuracy_score(Y_train, y_pred)
y_pred = Model_SVM.predict(X_validation)
test_accuracy = accuracy_score(Y_validation, y_pred)
filename = 'E://Model.sav'
pickle.dump(clf_bag, open(filename, 'wb')) #save model
```

Figure 17 : Le travail de la SVM.

4.7. Résultats et discussion

Dans cette section, nous évaluons la performance de notre méthode proposée. La performance de classification de chaque essai est évaluée par le taux de succès (%) qui est calculé par la division de nombre d'images correctement classés sur le nombre total d'images de test.

4.7.1. Analyse des résultats

Sur la base des résultats que nous avons obtenus, il est difficile de comparer les résultats entre eux à la fois. Nous avons donc divisé la comparaison entre les modèles selon le nombre de modèles CNNs combinés, comme indiqué dans le tableau 02.

A) Un modèle CNN unique

Nous notons que le taux de précision pour VGG19 est supérieur aux autres pourcentages et atteint 0,85%, et nous notons également que Inception V3 a donné le taux de précision le plus bas et atteint 0,77%.

B) Deux modèles CNNs combinés"

On note que le taux de précision varie de 0,84% à 0,89%, ce qui améliore les performances du cas d'un modèle unique. On remarque également que la combinaison de VGG16 et ResNet a donné le taux de précision le plus élevé de 0,89%, et que le cas de VGG19 combiné avec Inception V3 a donné taux de précision le plus faible de 0,84%.

C) Trois modèles CNNs combinés

On note que le taux de précision varie entre 0,86% à 0,83%, où le taux de précision atteint 0,86%, qui est le pourcentage le plus élevé en regroupant les modèles suivants VGG 16 et ResNet et VGG19, et à 0,83% qui est le plus faible dans le cas de VGG 16, ResNet et Inception V3.

D) Quatre modèles CNNs combinés

Il représentait le taux de précision le plus élevé, et ce pourcentage atteignait 0,91 %.

En comparant la moyenne des pourcentages de précision pour toutes les sections, nous notons que la section qui contient quatre modèles CNN donne le pourcentage le plus élevé et atteint 0,91%, suivie par le cas de deux modèles ensemble et son pourcentage atteint 0,87%

puis le cas de trois modèles et le taux de précision atteint 0,86%, et la plus faible précision dans le cas d'un modèle unique, avec un taux de 0,83%.

4.7.2. Discussion des résultats

Nous concluons par la convergence entre les rapports de précision que SVM a atteint sa limite d'optimisation (c'est-à-dire qu'il a atteint le stade de saturation), et nous avons également constaté que le programme a atteint le stade de saturation lorsque deux modèles sont combinés. Ce programme ne donne pas de différence de résultats s'il y a fusion de plus de deux modèles.

Tableau 04 : Rapports de performances TS (%) pour les modèles CNNs.

Combinaison	Variante	Modèles CNN combinés	TS (%)
Modèle CNN unique	V1-1	VGG16	0.8304
	V1-2	VGG19	0.8508
	V1-3	Inception V3	0.7739
	V1-4	RasNet	0.8383
Combinaison de deux CNNs	V2-1	VGG16 + VGG19	0.8681
	V2-2	VGG16 + Inception V3	0.8540
	V2-3	VGG16 + RasNet	0.8995
	V2-4	VGG19 + Inception V3	0.8477
	V2-5	VGG19 + RasNet	0.8901
	V2-6	Inception V3 + RasNet	0.8477
Combinaison de trois CNNs	V3-1	VGG16 + VGG19 + Inception V3	0.8556
	V3-2	VGG16 + VGG19 + RasNet	0.8697
	V3-3	VGG16 + Inception V3 + RasNet	0.8367
	V3-4	VGG19 + Inception V3 + RasNet	0.8634
Combinaison de quatre CNNs	V4	VGG16 + VGG19 + Inception V3 + RasNet	0.9112

4.8. Conclusion

Tout au long de ce chapitre, nous avons présenté l'approche proposée et validé les résultats illustrés par les taux de précision spécifiques à la classification de types de tumeur. En effet, les résultats de ce travail ont été comparés entre eux par sections. Cependant, nous avons pu obtenir de bons résultats en termes de précision de classification. Nous nous sommes appuyés sur des modèles combinés pour CNN et SVM pour la classification. Ceci dans le but de réaliser une classification photographique des cancers de la peau.

Conclusion générale

L'apprentissage en profondeur est actuellement utilisé pour l'analyse des images médicales du cancer de la peau et l'extraction de caractéristiques, la détection des nodules et la classification du type de tumeur (maligne ou bénigne). La plupart des algorithmes d'apprentissage automatique doivent comprendre les données pour extraire les caractéristiques pertinentes pour la classification. L'architecture d'apprentissage en profondeur a été utilisée pour classer différentes parties de la peau. Il s'avère que la sélection de fonctionnalités d'algorithmes d'apprentissage en profondeur tels que CNN affecte considérablement les performances du système. C'est pourquoi CNN réussit si bien dans la plupart des applications de reconnaissance d'images. CNN est principalement utilisé pour le traitement d'images. Les avantages de l'utilisation de l'architecture CNN sont de classer les types de tumeurs cutanées. CNN est l'approche la plus largement utilisée dans le dépistage du mélanome et surpasse également les modèles conventionnels.

- L'extraction et la classification des caractéristiques sont combinées par torsion, regroupement et connectivité complète.
- Il a été démontré que les modèles CNN fournissent de meilleurs résultats de classement pour le mélanome.
- La sélection d'entités n'est pas non plus nécessaire dans CNN car les entités sont définies par couche de déformation.

Le diagnostic médical assisté par ordinateur (CAD) est d'une grande utilité à l'appui des examens médicaux préventifs en photogravure des cellules de la peau ainsi que dans d'autres domaines tels que la neurologie ou les sciences cardiovasculaires.

Enfin, les modèles CNN (VGG16, VGG19, ResNet, Inception V3) combinés ont atteint des taux de précision proches, et grâce à eux, nous sommes arrivés à l'observation que lors de la fusion des modèles, nous pouvons obtenir un bon taux de précision. Par exemple, lors de la fusion des quatre modèles ensemble, nous avons obtenu un taux de précision de 0,91 pour détecter s'il s'agit d'un cancer de la peau bénin ou malin à partir du total des images prises à partir de Data Set.

À travers notre étude, nous suggérons quelques façons de travailler avec eux à l'avenir, telles que :

- Exploiter la dernière couche de modèles CNN, où nous la mettons dans SVM ;
- Appliquer d'autres méthodes d'apprentissage automatique à la place de SVM ;

- Adopter d'autres modèles CNN pour extraire les caractéristiques telles que ZIPF ou GABOR et les insérer dans le SVM ;
- Nous suggérons également d'utiliser des GPU pour réduire et accélérer le temps de calcul.

Références

- [01] [https://en.wikipedia.org/wiki/Coherence_\(physics\)?fbclid=IwAR00I900kRmzEmPxmVYEJ7LJ89zqFkLxBVu5q3yuJ1b6YohEI-TVJ_j2kFw#Spatial_coherence](https://en.wikipedia.org/wiki/Coherence_(physics)?fbclid=IwAR00I900kRmzEmPxmVYEJ7LJ89zqFkLxBVu5q3yuJ1b6YohEI-TVJ_j2kFw#Spatial_coherence); 11-06-2021
- [02] Kamel Houari," RECHERCHE D'IMAGES PAR LE CONTENU" Thèse De Doctorat, Université Mentouri Constantine, Juin 2010.
- [03] Hai, J.; Tan, H.; Chen, J.; Wu, M.; Qian, K.; Xu, J., et al.; Multi-level features combined end-to-end learning for automated pathological grading of breast cancer on digital mammograms. Computer zed Médical Imaging and Graphics (2019)
- [04] Chahnez Hadj S, Khadidja B: Filtre de Gabor. Université Abou Bekr Belkaid – Tlemcen. 2014/2015
- [05] Mme OUSLIMANI ép. OUTOUDERT Farida," Etude comparative des techniques de codage d'images en vue d'une segmentation ", Thèse De Doctorat, UNIVERSITÉ MOULOUD MAMMERY, TIZI-OUZOU, 21/10/ 2018.
- [06] Öztürka Ş. , Akdemir B. : " Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and GLCM", 2018 , Available online at www.sciencedirect.com .
- [07] A Brief View on Medical Diagnosis Applications with Deep Learning. U. Kose et al. 2021,
- [08] Site : https://support.echoview.com/WebHelp/Windows_and_Dialog_Boxes/Dialog_Boxes/Variable_properties_dialog_box/Operator_pages/GLCM_Texture_Features.htm , consulté le 29-04-2021.
- [09] Site: https://fr.wikipedia.org/wiki/Filtre_de_Gabor, consulté le 29-04-2021.
- [10] Site: <http://gwyddion.net/documentation/user-guide-fr/wavelet-transform.html> ,consulté le 29-04-2021.
- [11] Mention Électronique Spécialité Electronique des systèmes embarqués ESE -Par Hadj Attou Abdelkader et BOUIZZOUL ABD ELDJALAL
- [12] Kamel Houari," RECHERCHE D'IMAGES PAR LE CONTENU" Thèse De Doctorat, Université Mentouri Constantine, Juin 2010.
- [13] Deep Learning pour la classification des images –Par Moualek Djaloul Youcef.
- [14] Herbert Bay, Andreas Ess, Tinne Tuytelaars et Luc Van Gool, « SURF: Speeded Up Robust Features », Computer Vision and Image Understanding, vol. 110, no 3, p. 346 359, 2008.
- [15] Contribution à la détection faciale à l'aide du descripteur AKAZE- Par: MAHI Kheireddine et TRARI MEDJAOUI Nabil
- [16] Analyse de la texture des images mammaires par une fusion des lois de Zipf et des filtres Gabor dans un processus de classification des tumeurs mammaires –Par: HADFI Chiraz et TOUANSA Insaf.
- [17] Chloé-Agathe, A. Introduction au Machine Learning (Dunod, 2018).
- [18] L'apprentissage profond (Deep Learning) pour la classification et la recherche d'images par le contenu –Par: Boughaba Mohammed et Boukhris Brahim.
- [19] Analyse d'images médicales, application à la catégorisation des types de tuberculose pulmonaire –Par: YAGOUB Djamel Eddine.
- [20] Mme OUSLIMANI ép. OUTOUDERT Farida," Etude comparative des techniques de codage d'images en vue d'une segmentation ", Thèse De Doctorat, UNIVERSITÉ MOULOUD MAMMERY, TIZI-OUZOU, 21/10/ 2018.
- [21] Ashnil, K. Lei, B. Jinman, K. David Dagan, F. Machine learning in medical imaging. Biomédical Information Technologie 2020, pp 167-196

- [22] V. N. Vapnik and S. Kotz, Estimation of dependences based on empirical data, vol. 40. Springer Verlag New York, 1982.
- [23] V. N. Vapnik and S. Kotz, Estimation of dependences based on empirical data, vol. 40. Springer Verlag New York, 1982.
- [24] Analyse de la texture des images mammaires par une fusion des lois de Zipf et des filtres Gabor dans un processus de classification des tumeurs mammaires –Par: HADFI Chiraz et TOUANSA Insaf.
- [25] Rachid, M. Application des techniques d'apprentissage automatique pour la prédiction
- [26] https://blogs.msdn.microsoft.com/big_data_france/2014/06/05/lapprentissage-automatiquemachine-learning-comment-a-marche - Consulté le 12/06/2021
- [27] MACHINE LEARNING, An Algorithmic Perspective, Ralf Herbrich, 2015, International Standard Book Number-13: 978-1-4665-8333-7.
- [28] Site : https://fr.wikipedia.org/wiki/Machine_de_Boltzmann_restreinte - Consulté le 28-04-2021
- [29] (en) Paul Smolensky, David E. Rumelhart (dir.) et James L. McClelland (dir.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations, MIT Press, 1986, 194–281 p. (ISBN 0-262-68053-X, [lire en ligne \[archive\]](#)), « Chapter 6: Information Processing in Dynamical Systems: Foundations of Harmony Theory »
- [31] LAWRENCE, Steve, GILES, C. Lee, TSOI, Ah Chung, et al. Face recognition: A convolutional neural-network approach. IEEE transactions on neural networks, 1997, vol. 8, no 1, p. 98-113.
- [32] Settou T., Kholadi M.K. et Ben Ali A.: "Improving damage classification via hybrid deep learning feature representations derived from post-earthquake aerial images", International Journal of Image et Data Fusion, DOI:10.1080/19479832.2020.1864787,2020.
- [33] A. Krizhevsky, I. Sutskever et G. E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks », Advances in neural Processing Systems de traitement. 2012.
- [34] Graham, Benjamin. Fractional Max-Pooling. 18 décembre 2014.
- [35] L'apprentissage profond (Deep Learning) pour la classification et la recherche d'images par le contenu –Par : Boughaba Mohammed et Boukhris Brahim
- [36] Site: <https://deeplylearning.fr/cours-theoriques-deep-learning/fonction-dactivation/> consulté le 10-06-2021.
- [37] Site : <https://en.wikipedia.org/wiki/Epoch> consulté le 10-06-2021.
- [38] Site : <https://en.wikipedia.org/wiki/ImageNet> consulté le 12-06-2021.
- [39] Site : https://medium.com/@siddharthdas_32104/cnns-architectures-lenet-alexnet-vgggooglenet-resnet-and-more-666091488df5. - Consulté le 14/05/2021
- [40] Site. <https://keras.io/api/applications/#usage-examples-for-image-classification-models> - consulté le 17-04-2021.
- [41] Vetrivel A., Gerke M. , Kerle N., Nex F. ,Vosselman G. : " Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images , and multiple –kernel -learning",2017,journal homepage : www.elsevier.com/locate/isprsjprs ,isprsjprs.2017.03.001.
- [42] Site: <https://en.wikipedia.org/wiki/ImageNet> - consulté le 19-04-2021.
- [43] Brunelle F., Brunelle P.:" Intelligence artificielle et imagerie médicale : définition, état des lieux et perspectives Artificial intelligence and medical imaging : definition, state of the art and perspectives ",2019. This manuscript version is made available under the Elsevier user license <https://www.elsevier.com/open-access/userlicense/1.0/,S0001407919350745>.
- [44] Analyse d'images médicales, application à la catégorisation des types de tuberculose pulmonaire –Par: YAGOUB Djamel Eddine.
- [45] Jinnai S.h., Yamazaki N., Hirano H., Sugawara Y., Ohe Y. et Hamamoto R.:" The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning ", Biomolecules 2020, 10, 1123; doi:10.3390/biom10081123.

- [46] Rezvantab A., Safigholi H. et Karimijeshni S.: "Dermatologist Level Dermoscopy Skin Cancer Classification Using Different Deep Learning Convolutional Neural Networks Algorithms ",
- [47] Kadampur M.A.et Al Riyaae S.:" Skin cancer detection: Applying a deep learning-based model driven architecture in the cloud for classifying dermal cell images", journal homepage: <http://www.elsevier.com/locate/imu> , j.imu.2019.100282.
- [48] Site: <https://dsc.community.dev/u/myw58k/> - consulté le 31-05-2021
- [49] C. Arial et al. – “Query-focused abstractive summarization using neural networks”, These, Lethbridge, Alta.: University of Lethbridge, Dept. of Mathematics and . . ., 2019.
- [50] OpenCV, retrieved from site: <https://opencv.org/about/>;10-06-2021
- [51] Site : <https://www.lung.org/lung-health-diseases/lung-disease-lookup/pneumonia/symptoms-and-diagnosis> –consulté le 03-05-2021.