



INSTITUT FRANCOPHONE INTERNATIONAL

## TRAVAIL PERSONNEL ENCADRE

TPE : DEEP LEARNING POUR LA CLASSIFICATION DES IMAGES  
MEDICALES

---

# Rapport TPE

---

*Auteur: Soumana  
H.Abdourahmane*

*Encadrant: Dr.Ho T.Vinh*

January 18, 2020

## Contents

<b>1</b>	<b>INTRODUCTION</b>	<b>2</b>
1.1	ANALYSE DU SUJET . . . . .	3
1.2	CONTEXTE . . . . .	3
1.3	PROBLÉMATIQUE ET OBJECTIF . . . . .	4
1.4	Notion sur la pneumonie . . . . .	4
<b>2</b>	<b>Notion d'une image</b>	<b>6</b>
2.1	Définition . . . . .	6
2.2	Caractéristiques de l'image: . . . . .	6
<b>3</b>	<b>Les approches de classification d'images</b>	<b>9</b>
3.1	l'approche supervisée . . . . .	9
3.2	L'approche non supervisée . . . . .	9
<b>4</b>	<b>DEEP LEARNING</b>	<b>9</b>
4.1	POURQUOI DEEP LEARNING? . . . . .	10
4.2	Les différents types de modèles . . . . .	11
4.2.1	les réseaux de neurones convolutifs . . . . .	12
4.2.2	Les Couches de Réseaux de Neurone Convolutifs . . . . .	13
<b>5</b>	<b>Solution proposée et Implémentation</b>	<b>14</b>
5.1	Architecture générale de notre modèle . . . . .	15
5.2	Logiciels et bibliothèques Utilisés dans l'implémentation . . . . .	15
5.3	Configuration Utilisé dans l'implémentation . . . . .	15
5.4	Architecture de notre réseau . . . . .	16
5.4.1	Architecture du modèle 01 . . . . .	16
5.4.2	Architecture du modèle 02 . . . . .	17
<b>6</b>	<b>Expérimentation et résultats obtenus</b>	<b>19</b>
6.1	Résultats obtenus pour le modèle 01 : . . . . .	19
6.2	Résultats obtenus pour le modèle 02 . . . . .	20
6.2.1	Tableau de comparaison des résultats . . . . .	22
6.3	La planification de tâches . . . . .	23
6.4	Quelques travaux . . . . .	23
6.5	Conclusion . . . . .	25
6.6	Conclusion et perspective . . . . .	25
<b>7</b>	<b>REFERENCES</b>	<b>25</b>

# 1 INTRODUCTION

Dans le monde actuel,nous observons des milliers d'images qu'on croit identique, alors qu'ils existent des différences qui nous échappent.Pour pallier à ce problème étendu,Nous avons souhaiter utiliser la classification des images dans le domaine médical ,qui consistera à identifier les images à partir de leurs caractéristiques. (couleur, forme,contour,etc...) a fin de catégoriser ces dernières. Le Deep Learning est un nouveau domaine de recherche du ML, qui a été introduit dans le but de rapprocher le Machine Learning de son objectif principal : l'intelligence artificielle. Il concerne les algorithmes inspirés par la structure et le fonctionnement du cerveau. L'émergence récente de l'apprentissage automatique et des méthodes d'apprentissage en profondeur de l'image médicale, l'analyse a permis le développement des systèmes de diagnostic intelligents basés sur l'imagerie médicale pouvant aider les médecins à prendre de meilleures décisions concernant la santé d'un patient,notre travail consiste à classer des images du thorax(Pneumonie Normale).[2].[13]

## CHAPITRE 1

### 1.1 ANALYSE DU SUJET

Au cœur des évolutions technologiques, l'intelligence Artificielle est devenue en quelques années un des enjeux scientifiques majeurs du secteur de la santé, la perspective des nouvelles applications dans le domaine de l'imagerie médicale. L'intelligence artificielle (IA) soutient l'expansion de la médecine de précision, la transformation des parcours de soins et l'amélioration de l'expérience patient. Depuis plus de vingt ans, Siemens Healthineers place son expertise au service du développement de solutions enrichies par l'intelligence artificielle. Et aujourd'hui, les nouveaux algorithmes de deep learning constituent une aide significative à la classification des images médicales. En effet, l'objectif de la classification d'images est d'élaborer un système capable d'affecter une classe automatiquement à une image. Ainsi, ce système permet d'effectuer une tâche d'expertise qui peut s'avérer coûteuse à acquérir pour un être humain en raison notamment de contraintes physiques comme la concentration, la fatigue ou le temps nécessité par un volume important de données d'images.

Notre sujet de Travail Personnel Encadré (TPE), porte sur le problème de la détection automatique des pathologies, en particulier de la pneumonie, en appliquant l'analyse et la classification à partir d'image DME à l'aide d'une approche basée sur l'apprentissage en profondeur (Deep Learning en anglais).

### 1.2 CONTEXTE

Ce travail personnel encadré porte sur la classification d'images médicales (Une étude pathologique sur la pneumonie) en utilisant l'apprentissage en profondeur. Il existe un besoin de concevoir de nouvelles applications d'analyse et de classification de la pathologie pneumonie de manière spécifiques. La pneumonie est une infection des poumons causée soit par une bactérie, soit par un virus. L'infection bactérienne la plus souvent identifiée, et qui représente 2/3 des pneumopathies bactériennes est provoquée par un pneumocoque. Chaque année, un enfant sur six de moins de cinq ans décède de la pneumonie, soit plus que le VIH/Sida, le paludisme et la méningite réunis. La quasi-totalité de ces décès se produit dans des pays à faible et moyen revenu. Pour bien visualiser ce qu'est cette maladie, utilisons par exemple une métaphore. Les poumons sont comme une grappe de raisins. Les tiges étant les bronches et les grains, les alvéoles. Les bronchites se sont les tiges qui sont touchées et la pneumonie se sont les alvéoles qui sont touchés et cela peut donc être beaucoup plus grave. L'état du patient peut nécessiter la prise d'antibiotiques pour éviter

une surinfection ou pour prévenir la bactérie du pneumocoque qui peut entraîner le décès si ce n'est pas bien soigné. Les gestes de prévention sont simples. Il faut faire attention à bien se laver les mains et à avoir une hygiène irréprochable. Pour cela, notre défi à relever est de créer un outil permettant d'analyser et de classer automatiquement la pneumonie à partir d'images de DME (Document Médical Électronique). C'est dans cette optique que nous nous proposons d'apporter une solution partielle dans le cadre notre TPE (Travail Personnel Encadré), en nous appuyant sur des données existantes afin de créer un outil d'analyse et de classification des images médicales.

### 1.3 PROBLÉMATIQUE ET OBJECTIF

Le but de la recherche en Intelligence Artificielle est de créer une technologie qui permette aux ordinateurs et aux machines de fonctionner d'une manière intelligente. Le problème général de la création d'une intelligence a été divisé en plusieurs sous problèmes. Celles-ci consistent en des capacités que les chercheurs espèrent qu'un système intelligent pourra exécuter. Dans ce travail, nous nous fixons pour objectif de développer une application basée sur l'apprentissage en profondeur (le Deep Learning en anglais) pour résoudre le problème de la classification d'une image pneumonie ou Normal.

### 1.4 Notion sur la pneumonie

La pneumonie est une infection aiguë d'un ou des deux poumons, due à une bactérie, un virus ou un champignon. Chez l'adulte, la pneumonie est le plus souvent due à une bactérie. Et la bactérie le plus fréquemment en cause est le Pneumocoque (*Streptococcus pneumoniae*, pneumonies bactériennes). Quelque fois aussi, la pneumonie est consécutive à une fausse route ; du liquide gastrique ou des aliments passent dans les bronches, atteignent les poumons et provoquent une infection. On parle de pneumonie d'inhalation. Les bactéries présentes dans le nez et la gorge sont l'une des causes les plus fréquentes de pneumonie communautaire ; le Pneumocoque étant responsable de la plupart des cas. Certaines pneumonies se transmettent fréquemment d'une personne à l'autre. Dans le cas du Pneumocoque, la transmission se fait souvent d'enfant à adulte ; l'enfant porteur sain (qui abrite la bactérie dans le nez et la gorge sans symptômes) transmet la bactérie à une personne plus fragile (par exemple un individu de plus de 65 ans) qui développera la maladie. Notre étude se basera principalement sur deux classes : Pneumonie BACTÉRIE ou Virus et NORMAL. **Pneumonie BACTÉRIE** : Si le système immunitaire est affaibli par un virus présent dans le pharynx, ces bactéries peuvent envahir les voies respiratoires.

La pneumonie bactérienne est le plus souvent causée par la bactérie *Streptococcus pneumoniae* (le pneumocoque)

**Normal:** L'état sans pneumonie.[13]

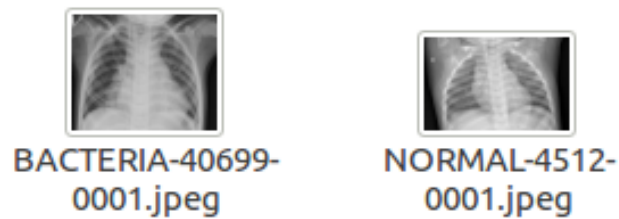


Figure 1: image jeu de donnée

## CHAPITRE 2

# 2 Notion d'une image

## 2.1 Définition

Une image médicale peut être définie comme la représentation des variations spatiales et temporelles d'une ou de plusieurs propriétés physiques d'un sujet sous investigation. On parle d'images statiques lorsque la variation temporelle de la propriété n'est pas prise en considération, autrement on parle d'images dynamiques. Les images sont soit des projections planaires du sujet soit des coupes tomographiques (du grec "tranche") à travers lui. La coupe peut être transversale, sagittale ou coronale selon son orientation dans le repère. L'image est considérée aussi comme un ensemble de points auquel est affectée une grandeur physique (luminance, couleur). Ces grandeurs peuvent être continues (image analogique) ou bien discrètes (images digitales). Mathématiquement, l'image représente une fonction continue  $IF$ , appelée fonction image, de deux variables spatiales représentée par  $IF(x, y)$  mesurant la nuance du niveau de gris de l'image aux coordonnées  $(x, y)$ . La fonction Image peut se représenter sous la forme suivante :  $IF : \mathbb{R}^2 \rightarrow \mathbb{R}$  Avec  $\mathbb{R}$ : ensemble des réelles.  $(x, y) \rightarrow IF(x, y)$   $x$  et  $y$  : Deux variables réelles[1] Image couleur RVB

L'œil humain analyse la couleur à l'aide de trois types de cellules photo 'les cônes'. Ces cellules sont sensibles aux basses, moyennes, ou hautes fréquences (rouge, vert, bleu). Pour représenter la couleur d'un pixel, il faut donc donner trois nombres, qui correspondent au dosage de trois couleurs de base : Rouge, Vert, Bleu. On peut ainsi représenter une image couleur par trois matrices chacune correspondant à une couleur de base. Image d'intensités C'est une matrice dans laquelle chaque élément est un réel compris entre 0 (noir) et 1 (blanc). On parle aussi d'image en niveaux de gris, car les valeurs comprises entre 0 et 1 représentent les différents niveaux de gris. Image binaire Une image binaire est une matrice rectangulaire dans l'élément valent 0 ou 1. Lorsque l'on visualise une telle image, les 0 sont représentés par du noir et les 1 par du blanc.

## 2.2 Caractéristiques de l'image:

L'image est un ensemble structuré d'information caractérisé par les paramètres suivants :[1]

**Pixel** Le pixel est l'abréviation du mot « Picture élément », est une unité de surface permettant de définir la base d'une image numérique. Il matérialise un point donné  $(x,y)$  du plan de l'image. L'information présentée par le pixel est le niveau de gris (ou la couleur) prélevée à l'emplacement correspondant dans l'image réelle. La différence entre image monochrome et image couleur réside dans la quantité d'informations contenue dans chaque pixel, par exemple dans une image couleur (RVB : Rouge, Vert, Bleu) la valeur d'un pixel est représentée sur trois octets pour chaque couleur.[1]

**Dimension-Résolution** La dimension est la taille de l'image. Elle se présente sous forme d'une matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels). Le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre total de pixels dans une image. Par contre, la résolution est la clarté ou la finesse de détails atteinte par un moniteur ou une imprimante dans la production d'images. Sur les moniteurs d'ordinateur, la résolution est exprimée en nombre de pixels par unité de mesure (pouce ou centimètre). On utilise aussi le mot résolution pour désigner le nombre total de pixels horizontaux et verticaux sur un moniteur. Plus ce nombre est grand, plus la résolution est meilleure.

**Voisinage** Le plan de l'image est divisé en termes de formes rectangulaires ou hexagonales permettant ainsi l'exploitation de la notion de voisinage (voir figure 1). Le voisinage d'un pixel est formé par l'ensemble des pixels qui se situent autour de ce même pixel. On définit aussi l'assiette comme étant l'ensemble de pixels définissant le voisinage pris en compte autour d'un pixel. On distingue deux types de voisinage : Voisinage à 4: On ne prend en considération que les pixels qui ont un côté commun avec le pixel considéré. Voisinage à 8: On prend en compte tous les pixels qui ont au moins un point en liaison avec le pixel considéré.

$p(i-1,j)$	$p(i,j-1)$	$p(i+1,j)$		$p(i-1,j-1)$	$p(i,j-1)$	$p(i+1,j-1)$
	$p(i,j)$			$p(i-1,j)$	$p(i,j)$	$p(i+1,j)$
	$p(i,j+1)$			$p(i-1,j+1)$	$p(i,j+1)$	$p(i+1,j+1)$

**Voisinage a 4****voisinage a 8**

Figure 2: Voisinages

**Niveau de gris** C'est la valeur d'intensité lumineuse d'un pixel. Cette valeur peut aller du noir (0) jusqu'au blanc (255) en passant par les nuances qui sont contenues dans l'intervalle  $[0, 255]$ . Elle correspond en fait à la quantité de la lumière réfléchie. Pour 8 bits, on dispose de 256 niveaux de gris dont 40 sont reconnus à l'œil nue. Plus le nombre de bit est grand plus les niveaux sont nombreux et plus la représentation est fidèle.



Contraste C'est l'opposition marquée entre deux régions d'une image. Une image contrastée présente une bonne dynamique de la distribution des valeurs de gris sur tout l'intervalle des valeurs possibles, avec des blancs bien clairs et des noirs profonds. Au contraire une image peu contrastée a une faible dynamique, la plupart des pixels ayant des valeurs de gris très proches. Si  $L_1$  et  $L_2$  sont les degrés de luminosité respectivement de deux zones voisines  $A_1$  et  $A_2$  d'une image, le contraste est défini par le rapport :  $C=L_1L_2/L_1+L_2$ . Luminance C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface, pour un observateur lointain, le mot luminance est substitué au mot brillance, qui correspond à l'éclat d'un objet. Une bonne luminance se caractérise par :

- . Des images lumineuses (brillantes);
  - . Un bon contraste : il faut éviter les images où la gamme de contraste tend vers le blanc ou le noir; ces images entraînent des pertes de détails dans les zones sombres ou lumineuses.
  - . L'absence de parasites. Bruit Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur. C'est un parasite qui représente certains défauts (poussière, petits nuages, baisse momentanée de l'intensité électrique sur les capteurs, ...etc.). Il se traduit par des taches de faible dimension et dont la distribution sur l'image est aléatoire.
- Contour Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentant une différence significative. Dans une image numérique, les contours se situent entre les pixels appartenant à des régions ayant des intensités moyennes différentes; il s'agit de contours de type « saut d'amplitude ». Un contour peut également correspondre à une variation locale d'intensité présentant un maximum ou un minimum; il s'agit alors de contour « en toit ».
- Présentation et orientation des images Par convention, les images sont présentées de la même façon au sein d'une modalité, qu'elle soit en projection ou en coupes: en projection de face ou en coupe coronale (ou frontale), la droite du patient est présentée à gauche de l'écran : comme si on regardait simplement le patient de face; sur une vue axiale (ou transverse) la droite du patient est toujours située à gauche de l'écran: comme si on regardait le patient – allongé sur le dos – depuis ses pieds. Sur une vue sagittale ou de profil, l'image est présentée avec la partie antérieure du patient à gauche de l'écran, comme si on regardait le patient de profil gauche (à l'exception des images d'échographie sur lesquelles le pôle crânial du patient est présenté à gauche de l'écran). Les images de scintigraphie planaire (projections) sont toujours représentées comme si l'observateur avait le point de vue de la caméra (cf. « Médecine nucléaire »).[9]

### 3 Les approches de classification d'images

La classification des images consiste à répartir systématiquement des images selon des classes établies au préalable, classer une image lui fait correspondre une classe, marquant ainsi sa parenté avec d'autres images c'est à dire regrouper deux ou plusieurs images en fonction de leurs catégories à l'aide d'un système informatique. Ainsi de nombreuses approches ont été consacrées à la classification, elles peuvent être séparées en deux grandes catégories :

#### 3.1 l'approche supervisée

[5] En classification supervisée, vous ne laissez pas l'ordinateur créer les classes, vous les créez vous-mêmes et laissez à l'ordinateur l'étape suivante, c'est-à-dire l'assignation des pixels aux classes. Cela signifie que vous devez d'abord déterminer le nombre et la nature des classes que vous voulez utiliser. Suivez les étapes suivantes pour une classification supervisée :

#### 3.2 L'approche non supervisée

La classification non-supervisée consiste à laisser l'ordinateur calculer automatiquement les classes sur la base de plusieurs (en tout cas plus d'une) bandes de fréquences de votre image. Cela vous laisse la tâche d'identifier le bon nombre et la nature réelle des classes obtenues. Le but de ces différentes méthodes de classification est d'identifier les classes auxquelles appartiennent des objets à partir de certains traits descriptifs. Elles s'appliquent à un grand nombre d'activités humaines et conviennent en particulier au problème de la prise de décision automatisée. La procédure de classification sera extraite automatiquement à partir d'un ensemble d'exemples.[5]

## 4 DEEP LEARNING

Le Deep Learning est un nouveau domaine de recherche du ML, qui a été introduit dans le but de rapprocher le ML de son objectif principal : l'intelligence artificielle. Il concerne les algorithmes inspirés par la structure et le fonctionnement du cerveau. Ils peuvent apprendre plusieurs niveaux de représentation dans le but de modéliser des relations complexes entre les données.

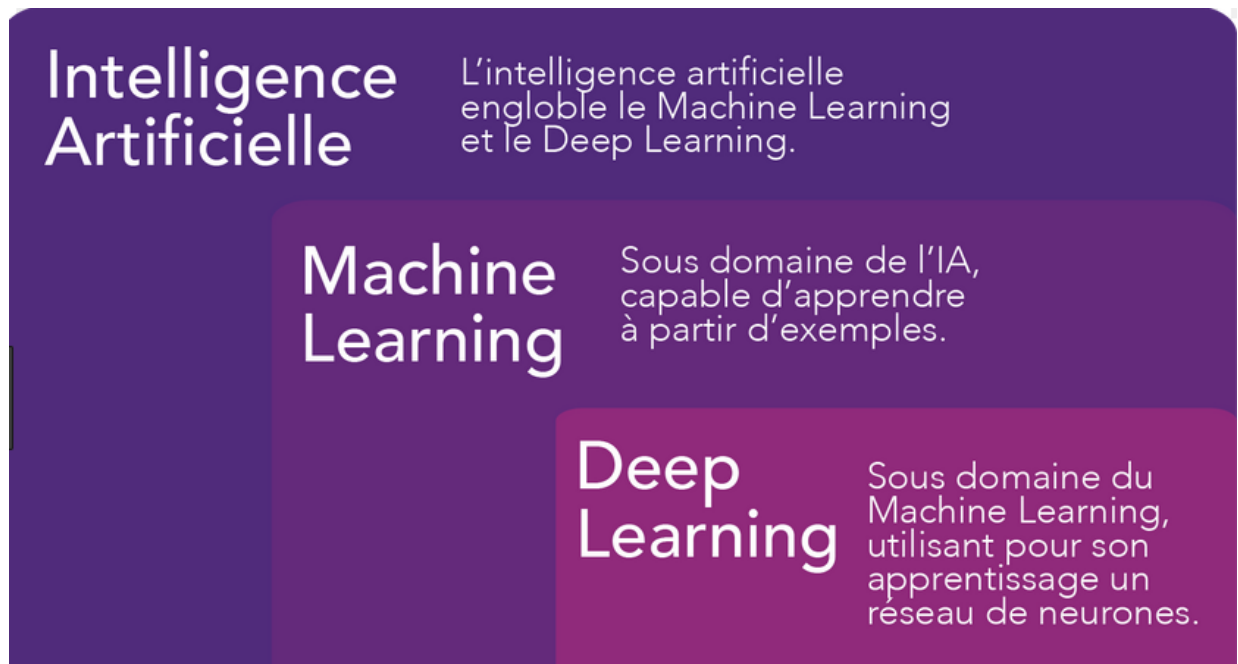


Figure 3:  
Relation entre Intelligence Artificielle, Machine Learning, Deep Learning

Le Deep Learning est basé sur l'idée des réseaux de neurones artificielles et il est taillé pour gérer de larges quantités de données en ajoutant des couches au réseau. Un modèle de deep learning a la capacité d'extraire des caractéristiques à partir des données brutes grâce aux multiples couches de traitement composé de multiples transformations linéaires et non linéaires et apprendre sur ces caractéristiques petit à petit à travers chaque couche avec une intervention humaine minime .

## 4.1 POURQUOI DEEP LEARNING?

Les algorithmes de ML décrits dans la première partie fonctionnent bien pour une grande variété de problèmes. Cependant ils ont échoués à résoudre quelques problèmes majeurs de l'IA telle que la reconnaissance vocale et la reconnaissance d'objets. Le développement du deep learning fut motivé en partie par l'échec des algorithmes traditionnels dans de telle tâche de l'IA. Mais ce n'est qu'après que de plus grandes quantités de données ne soit disponibles grâce notamment au Big Data et aux objets connectés et que les machines de calcul soient devenues plus puissantes qu'on a pu comprendre le potentiel réel du Deep Learning.

Une des grandes différences entre le Deep Learning et les algorithmes de ML tra-

ditionnelles c'est qu'ils'adapte bien, plus la quantité de données fournie est grande plus les performances d'un algorithme de Deep Learning sont meilleurs. Contrairement à plusieurs algorithmes de ML classiques qui possèdent une borne supérieure a la quantité de données qu' ils peuvent recevoir des fois appelée "plateau de performance", les modèles de Deep Learning n'ont pas de telles limitations (théoriquement) et ils sont même allés jusqu'à dépasser la performance humaine dans des domaines comme l'image processing.

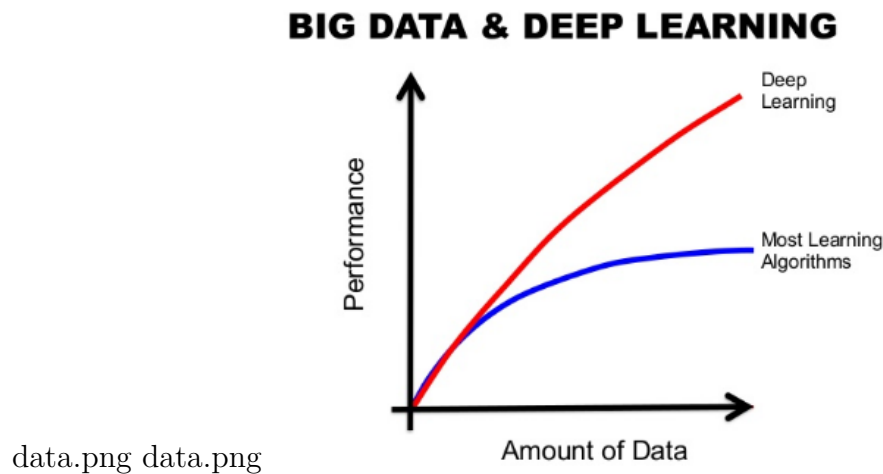


Figure 4:

La différence de performance entre le Deep Learning et la plupart des algorithmes de ML en fonction de la quantité de données.

Autre différence entre les algorithmes de ML traditionnelles et les algorithmes de Deep Learning c'est l'étape de l'extraction de caractéristiques. Dans les algorithmes de ML traditionnelles l'extraction de caractéristiques est faite manuellement, c'est une étape difficile et coûteuse en temps et requiert un spécialiste en la matière alors qu'en Deep Learning cette étape est exécutée automatiquement par l'algorithme.

## 4.2 Les différents types de modèles

Il existe un grand nombre de variantes d'architectures profondes. La plupart d'entre elles sont dérivées de certaines architectures parentales originales. Il n'est pas toujours possible de comparer les performances de toutes les architectures, car elles ne sont pas toutes évaluées sur les mêmes ensembles de données. Le Deep Learning est un domaine à croissance rapide, et de nouvelles architectures, variantes ou algorithmes apparaissent toutes les semaines.

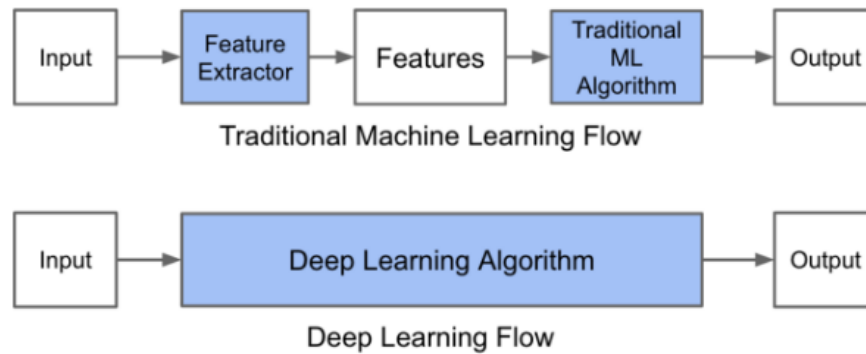


Figure 5:

Le procédé du ML classique comparé à celui du Deep Learning.

#### 4.2.1 les réseaux de neurones convolutifs

(réseaux de neurones convolutifs) sont un type de réseau de neurones spécialisés pour le traitement de données ayant une topologie semblable à une grille. Les exemples comprennent des données de type série temporelle, qui peuvent être considérées comme une grille 1D en prenant des échantillons à des intervalles de temps réguliers et des données de type image, qui peuvent être considérées comme une grille 2D de pixels. Les réseaux convolutifs ont connu un succès considérable dans les applications pratiques. Le nom « réseau de neurones convolutif » indique que le réseau emploie une opération mathématique appelée convolution. La convolution est une opération linéaire spéciale. Les réseaux convolutifs sont simplement des réseaux de neurones qui utilisent la convolution à la place de la multiplication matricielle dans au moins une de leurs couches.

#### 4.2.2 Les Couches de Réseaux de Neurone Convolutifs

Ci dessous quelques couches de réseaux de RNC.

Les couches de RNC	Définitions	Position dans le modèle
Couche convolution	Elle est la composante clé de RNC, but est de repérer la présence d'un ensemble des features des images reçues en entrée.	Première couche dans le modèle
Couche pooling	Elle reçoit en entrée plusieurs features maps et applique à chacune d'entre elle le pooling. Pooling: consiste à réduire la taille des images, tout en préservant leurs caractéristiques importantes.	Elle est souvent placée entre deux couches de convolution
Couche de correction ReLu (Rectified linear units)	Elle remplace toutes les valeurs négatives reçues en entrée par zéro. elle joue le rôle de fonction d'activation.	elle est appliquée très souvent en sortie d'une couche de convolutions pour diverses raisons.
Couche Full y-Connected	Elle permet de classifier l'image en entrée du réseau, elle renvoie un vecteur de taille N. ou N est le nombre de classe dans notre problème.	Cette couche constitue toujours la dernière couche d'un Réseau de Neurone convolutif ou non, donc elle n'est pas caractéristique pour le CNN.

Figure 6:  
Les différentes couches de notre réseau.

## CHAPITRE 3

# 5 Solution proposée et Implémentation

Dans le cadre de la mise en place d'un outil d'analyse et de classification d'images de la pneumonie et tenant Compte de l'étude des méthodes de l'apprentissage en profondeur (Deep learning en anglais), nous avons tiré des conclusions pour lesquelles nous essayerons d'utiliser le réseau de neurones convolutionnels (CNN) en tant que modèle de classification. La solution que nous proposons consiste :

- A créer notre propre réseau de neurones convolutionnel (CNN) ;
- Procéder à son entraînement en utilisant les images de notre jeu de données ;
- Augmenter les données d'apprentissage pour avoir les meilleurs performances possibles du modèle de classification ;

— Nous allons travailler avec les bibliothèques Tensorflow et Keras pour la classification et pour améliorer les performances du modèles on va utiliser une technique simple et efficace comme l'augmentation des données.

**Le Jeu de donnée:** Nous utilisons les données de DME (Document Médical Électronique). C'est un effort international visant à améliorer le diagnostic de la pneumonie, parrainé par University of California San Diego et la Cell Press qui est une maison d'édition basée au états-unis et Filiale du groupe éditorial.elle est spécialisée dans la publication de revue biomédicales. Daniel S. Kermany, Michael Goldbaum et ses collègues ont travaillé sur l'article et ont obtenus des bons résultats. <https://doi.org/10.1016/j.cell.2018.02.010> et le lien de jeu de donnée: <https://data.mendeley.com/datasets/rscbjbr9sj/3>.

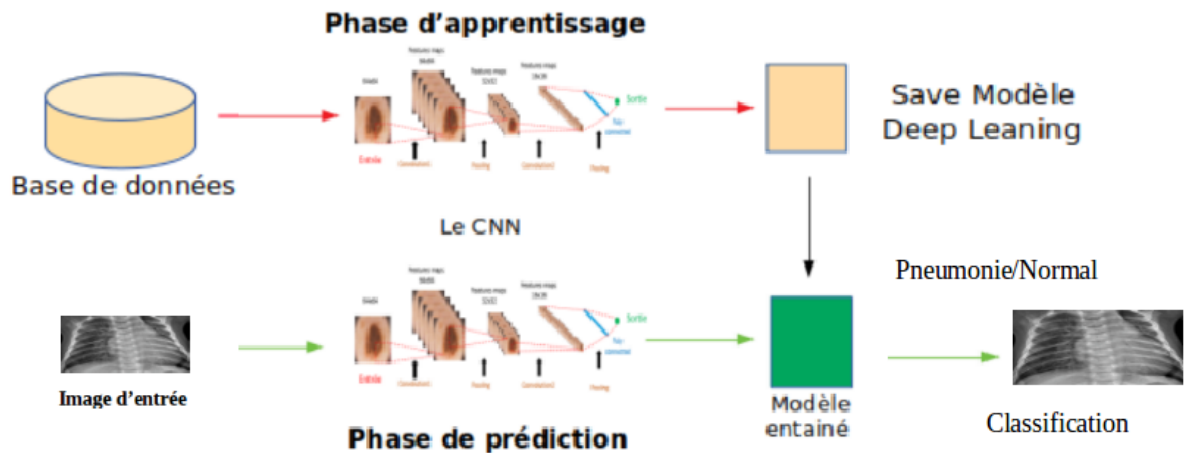
Notre jeu de données est composé de **5858 images** de taille différente dont **1584 Normal et 4274 Pneumonial (ayant des bactéries)** que nous allons repartir en deux (2) échantillons à savoir :

- **5233 images d'apprentissages ou Training** dont **1349 Normal et 3884 pneumonie** ;
- **625 images de test/validation** dont **235 Normal et 390 pneumonie**.

Toutes les images sont nommées en utilisant le schéma

«NORMAL-id-0001.jpeg »,«BACTERIA-id-0001.jpeg» .

## 5.1 Architecture générale de notre modèle



## 5.2 Logiciels et bibliothèques Utilisés dans l'implémentation



### 5.3 Configuration Utilisé dans l'implémentation

La configuration du matériel utilisé dans notre implémentation est :

- Un ordinateur portable Dell i5 CPU 2.40 GHZ
- Carte graphique Intel
- RAM de taille 4 GO
- Disque dur de taille 500 GO
- Système d'exploitation Linux Ubuntu version 16.04



## 5.4 Architecture de notre réseau

Au cours de nos expérimentations, nous avons créé deux modèles (modèle 1, modèle 2) avec différentes architectures, où on a appliqué les deux modèles sur notre base d'image. Dans ce qui suit on présente l'architecture des deux modèles :

### 5.4.1 Architecture du modèle 01

Le premier modèle CNN que nous présentons dans la Figure 7 est composé de quatre couches de convolution et quatre couches de maxpooling et de deux couches de fully connected. L'image en entrée est de taille  $299 \times 299$ , l'image passe d'abord à la première couche de convolution, cette couche repérera la présence d'un ensemble des features des images reçues en entrée, elle est composée de 32 filtres de taille  $3 \times 3$ , la fonction d'activation ReLU est utilisée, cette fonction d'activation force les neurones à retourner des valeurs positives, après cette convolution 32 feature maps seront créés de taille  $32 \times 32$ . Ensuite, les 32 feature maps qui sont obtenus ils sont données en entrée de la deuxième couche de convolution qui est composée aussi de 32 filtres. La fonction d'activation ReLU est appliquée sur cette couche (pour remplacer les valeurs négatives reçues en entrée par zéro). Le Maxpooling est appliqué après pour réduire la taille de l'image en conservant les caractéristiques importantes de l'image. Nous répétons la même chose avec les couches de convolutions restantes (deux, trois, quatre). Après ces quatre couches de convolution, nous utilisons un réseau de neurones composé de deux couches fully connected. La première couche est composée de 256 neurones où la fonction d'activation utilisée est le ReLU, la deuxième couche utilise la fonction softmax qui permet de calculer la distribution de probabilité des 2 classes (nombre de classe dans notre base d'images).

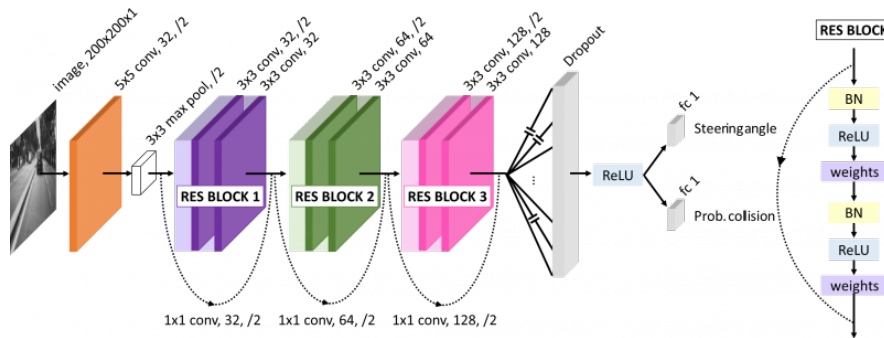


Figure 7:  
Architecture du modèle 01

Layer (type)	Output Shape	Param #
conv2d_119 (Conv2D)	(None, 299, 299, 4)	52
max_pooling2d_29 (MaxPooling)	(None, 149, 149, 4)	0
conv2d_120 (Conv2D)	(None, 149, 149, 8)	136
max_pooling2d_30 (MaxPooling)	(None, 74, 74, 8)	0
dropout_25 (Dropout)	(None, 74, 74, 8)	0
conv2d_121 (Conv2D)	(None, 74, 74, 12)	396
max_pooling2d_31 (MaxPooling)	(None, 37, 37, 12)	0
dropout_26 (Dropout)	(None, 37, 37, 12)	0
conv2d_122 (Conv2D)	(None, 37, 37, 16)	784
max_pooling2d_32 (MaxPooling)	(None, 18, 18, 16)	0
dropout_27 (Dropout)	(None, 18, 18, 16)	0
flatten_7 (Flatten)	(None, 5184)	0
dense_13 (Dense)	(None, 256)	1327360
dropout_28 (Dropout)	(None, 256)	0
dense_14 (Dense)	(None, 2)	514
Total params: 1,329,242		
Trainable params: 1,329,242		
Non-trainable params: 0		

Figure 8:  
Configuration du modèle 01

#### 5.4.2 Architecture du modèle 02

Le premier modèle CNN que nous présentons dans la Figure 9 est composé de quatre couches de convolution et quatre couches de maxpooling et de deux couches de fully connected. L'image en entrée est de taille 299\*299, l'image passe d'abord à la première couche de convolution, cette couche repérera la présence d'un ensemble des features des images reçues en entrée, elle est composée de 32 filtres de taille 3\*3, la fonction d'activation ReLU est utilisée, cette fonction d'activation force les neurones à retourner des valeurs positives, après cette convolution 32 features maps seront créés de taille 32\*32. Ensuite, les 32 feature maps qui sont obtenus ils sont données en entrée de la deuxième couche de convolution qui est composée aussi de 32 filtres. La fonction d'activation ReLU est appliquée sur cette couche (pour remplacer les valeurs négatives reçues en entrée par zéro). Le Maxpooling est appliqué après pour réduire la taille de l'image en conservant les caractéristiques importantes de l'image. Nous répétons la même chose avec les couches de convolutions restantes (deux, trois, quatre). Après ces quatre couches de convolution, nous utilisons un réseau de neurones composé de deux couches fully connected. La première

couche est composée de 256 neurones où la fonction d'activation utilisée est le ReLU, la deuxième couche utilise la fonction softmax qui permet de calculer la distribution de probabilité des 2 classes (nombre de classe dans notre base d'images).

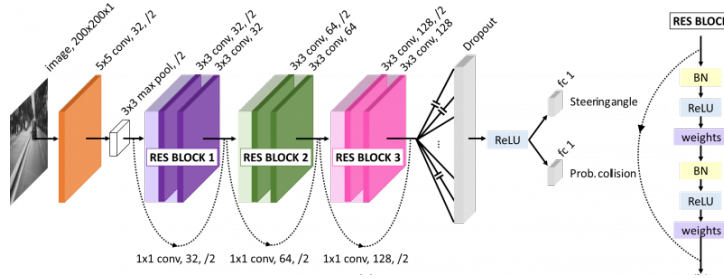


Figure 9:  
Architecture du modèle 02

Layer (type)	Output Shape	Param #
conv2d_135 (Conv2D)	(None, 299, 299, 4)	52
max_pooling2d_45 (MaxPooling)	(None, 149, 149, 4)	0
conv2d_136 (Conv2D)	(None, 149, 149, 16)	272
max_pooling2d_46 (MaxPooling)	(None, 74, 74, 16)	0
dropout_41 (Dropout)	(None, 74, 74, 16)	0
flatten_10 (Flatten)	(None, 87616)	0
dense_19 (Dense)	(None, 256)	22429952
dropout_42 (Dropout)	(None, 256)	0
dense_20 (Dense)	(None, 2)	514
Total params: 22,430,790		
Trainable params: 22,430,790		
Non-trainable params: 0		

Figure 10:  
Configuration du modèle 02

## CHAPITRE 4

# 6 Expérimentation et résultats obtenus

Afin de montrer les résultats obtenus pour les deux modèles, on illustre dans ce qui suit les résultats en termes de précision et d'erreur ainsi qu'une matrice de confusion pour chacun des deux modèles. Notons que les deux modèles sont entraînés sur collab c'est-à-dire dans le cloud GPU.

### 6.1 Résultats obtenus pour le modèle 01 :

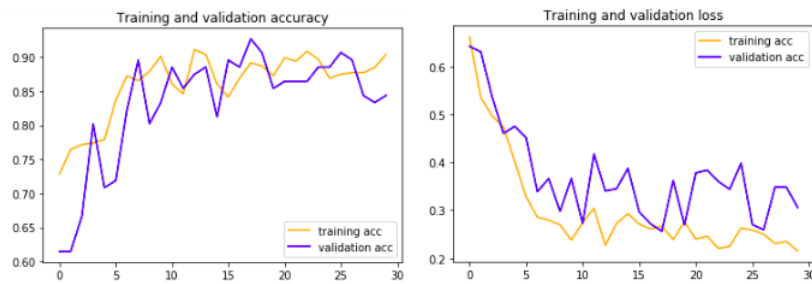


Figure 11:  
Précision et Erreur pour le modèle 01

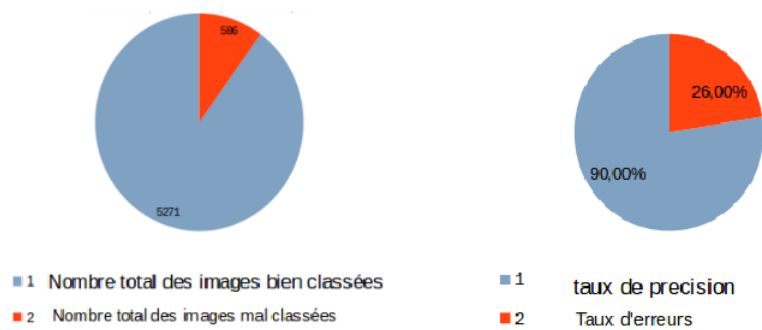


Figure 12:  
Nombre d'images mal et bien classées Taux d'erreur et de précision

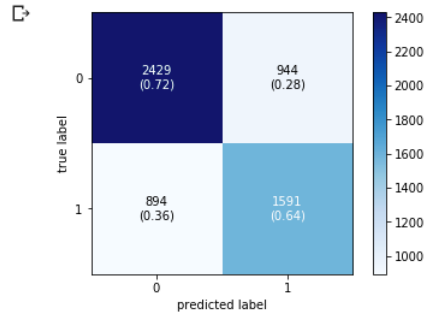


Figure 13:  
la matrix de confusion du modèle 01

D'après cette Figure, nous avons 5271 images bien classées et 586 images mal classées sur l'ensemble de données, et nous avons enregistré une précision de 90 pourcent et une perte de 26 pourcent, et également la matrice de confusion nous donne une visibilité sur les images bien et mal classées sur les données de training et test que nous pouvons voir sur la figure. De plus, nous avons observé sur le modèle 01, la précision de l'apprentissage et de la validation augmente avec le nombre d'époque, ceci reflète qu'à chaque époque le modèle apprend plus d'informations. De même, l'erreur d'apprentissage et de la validation diminue avec le nombre d'époque.

## 6.2 Résultats obtenus pour le modèle 02

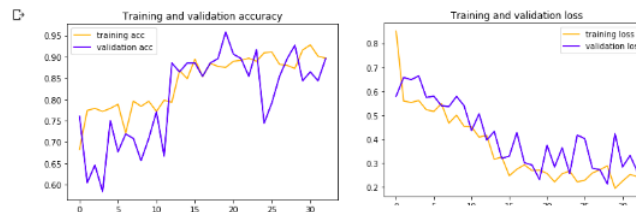


Figure 14:  
Précision et Erreur pour le modèle 02

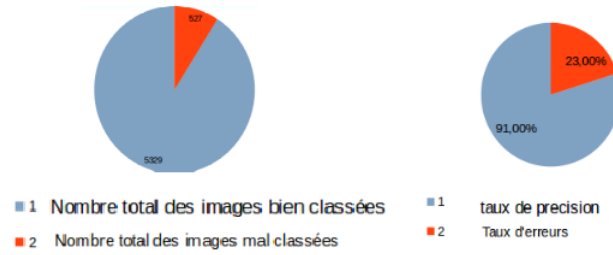


Figure 15:  
Nombre d'images mal et bien classé Taux d'erreur et de précision

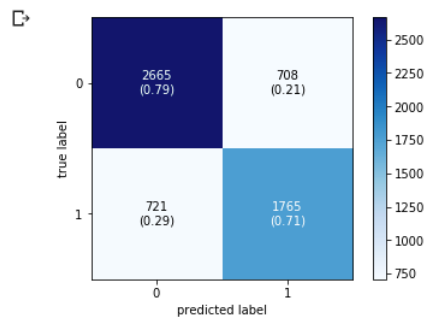


Figure 16:  
la matrix de confusion du modèle 02

D'après cette Figure, nous avons 5329 images bien classées et 527 images mal classées sur l'ensemble de données, et nous avons enregistré une précision de 91% et une perte de 23 %, et également la matrice de confusion nous donne une visibilité sur les images bien et mal classées sur les données de training et test que nous pouvons voir sur la figure. De plus, nous avons observé sur le modèle 02, la précision de l'apprentissage et de la validation augmente avec le nombre d'époque, ceci reflète qu'à chaque époque le modèle apprend plus d'informations. De même, l'erreur d'apprentissage et de la validation diminue avec le nombre d'époque.

### 6.2.1 Tableau de comparaison des résultats

Modèles	Architecture utilisée			Nombre époque	Précision obtenus sur les données d'apprentissage	Précision obtenus sur les données de validation	Erreur
	convolution	Max poling	Fully connected				
Modèle 01	04	04	02	30	72%	64%	30%
Modèle 02	06	06	02	40	79%	71%	20%

Figure 17:  
tableau de comparaison

Le tableau ci dessus nous montre l'architecture utilisée dans chaque modèle ainsi que leurs nombre d'époques. Les résultats obtenus sont exprimés en termes de précision d'apprentissage, de validation et d'erreur. Le temps d'exécution est trop coûteux, c'est pourquoi on a jugé utile d'utiliser le google colab(Cloud). D'après les résultats que nous avons enregistré sur notre tableau de comparaison le modèle 2 a présenté les meilleurs résultats par rapport au premier modèle, en effet nous avons augmenté le nombre de couches et également le nombre d'époques que vous pouvez voir dans le tableau ci dessus et nous avons enregistré les résultats de modèle 2. Les résultats de nos deux modèles nous montre une variation au niveau de l'apprentissage de (72 à 79% ) et pour la précision de validation les pourcentages sont de (64 à 71 %) est également les erreurs ont variées (30 à 20 %). D'une manière générale, on peut dire que Plus on augmente les couches, on varie les nombres des époques, on joue sur les valeurs de batch, plus on enregistre un bon résultat.

### 6.3 La planification de taches

Travaux	Descriptions	Durée estimée
Familiarisation aux nouvelles librairies et langages	Familiarisations avec les librairies et les langages	1 mois 1 semaines
Préparation de l'entraînement	Installation de librairies	1 semaine
Implémentation et apprentissage des modèles	-Organisation du jeu de donnée' -codage des modèles; -application du modèle au jeu de donnée; -application du modèle en cloud; -Présentation a l'encadreur et correction des suggestions.	2 mois 1 semaine
Test et évaluations des résultats	-Test de résultats obtenus; -Présentation des résultats obtenus a l'encadrant; -Correction et amélioration si possible;	10 jours
Rédaction du rapport final	-Assemblages de différents sous rapports; -soumission du rapport et correction par l'encadreur; -soumission du rapport final;	15 jours

Figure 18:  
la planification des taches

### 6.4 Quelques travaux

L'imagerie médicale a fait d'immenses progrès ces dernières années. C'est principalement dû à la numérisations de l'imagerie diagnostique, qui améliore la qualité des données et des images, ainsi qu'à la rapidité de mise à disposition des images. Ensemble de données sur des maladies pathologiques sur le thorax . En outre, dans la comparaison des radiographies thoraciques présentant une pneumonie par rapport à la normale, elles ont atteint une précision de 92,8%, avec une sensibilité de 93,2% et une spécificité de 90,1% . La surface sous la courbe ROC pour la détection de pneumonie était de 96,8% .De même, les groupes de Google et Andrew Y. Ng ont également obtenu d'excellents résultats en utilisant des systèmes **d'IA** pour le diagnostic des maladies oculaires et de la pneumonie. Les résultats de ces études démontrent la récente. L'apprentissage en profondeur est efficace pour classer les images OCT de la dégénérescence maculaire liée à l'âge normal Objectif: L'avènement des dossiers médicaux électroniques (DME) avec de grandes bases de données d'imagerie électronique avec les avancées dans les réseaux de neurones profonds et l'apprentissage automatique a fourni une occasion unique d'atteindre des jalons dans l'analyse automatisée des images. La tomographie par cohérence optique est la modalité d'imagerie la plus courante en ophtalmologie et représente un ensemble de données dense et riche lorsqu'elle est combinée à des étiquettes dérivées du



REM. Nous avons cherché à déterminer si l'apprentissage en profondeur pouvait être utilisé pour distinguer les images OCT normales des images provenant de patients atteints de dégénérescence maculaire liée à l'âge (DMLA).Après l'apprentissage,ils ont obtenus des résultats excellents.

## 6.5 Conclusion

Dans cette partie, nous avons décrit la démarche suivie afin d'obtenir le classifieur le plus précis possible. Nous avons commencé par un réseau de neurones convolutif classique et analysé ses défauts en observant son comportement grâce aux graphes et nous avons fait changer les paramètres un à un afin de comprendre l'impacte de chacun sur la performance finale. Après avoir trouvé l'architecture qui nous convient le mieux, nous avons décidé de tirer le maximum du réseau en laissant l'apprentissage durer longtemps jusqu'à obtenir la performance maximale que peut atteindre notre modèle.

## 6.6 Conclusion et perspective

Grâce aux Deep Learning, l'avenir de l'intelligence artificielle est prometteur. Dans ce travail, nous avons exploré le domaine de la classification d'image qui comme tous les autres domaines de l'intelligence artificielle ont connue une évolution majeure depuis l'apparition du Deep Learning. Afin d'aboutir à ces résultats, nous avons passé beaucoup de temps à lire et à étudier les publications et les articles pour voir ce qui se fait de mieux en matière de classification et pour pouvoir concevoir notre propre modèle. Pour finir, avant de passer aux perspectives, ce travail nous a permis de mettre en pratique nos connaissances sur les réseaux de neurones et d'en acquérir d'autres et le temps passé à lire des articles nous a servi d'une bonne initiation à la recherche. Comme perspectives nous pouvons citer :

- Implémentation des techniques qui font le succès des modèles qui participent au challenge Imagenet (ResNet ...).
- Tester sur de nouvelles bases.
- Segmentation et génération des images.

## 7 REFERENCES

1. <http://dspace.univ-tlemcen.dz/bitstream/112/12583/1/Deep-Learning-pour-la-classification..pdf> accessed 29 Novembre 2018]"
2. <https://www.zdnet.fr/actualites/imagerie-medicale-le-deep-learning-peut-il-faire-mieux-qu-un-medecin-39858256.htm> accessed 26 Novembre 2018]"
3. <https://www.healthcare.siemens.fr/news/intelligence-artificielle-imagerie-medicale.html> accessed 23 Novembre 2018]"

4. <https://ch.mathworks.com/fr/discovery/deep-learning.html> accessed 30 Novembre 2018].
5. <https://grass.osgeo.org/gdp/grass5tutor/HTMLfr/c1625.html>;
6. <https://www.gralon.net/articles/materiel-et-consommables/materiel-informatique-et-consommable-informatique/article-comprendre-le-deep-learning-en-6-questions-10956.htm>;
7. <http://dspace.univ-tlemcen.dz/bitstream/112/322/11/ChapitreII.pdf>;
8. <https://tel.archives-ouvertes.fr/tel-01704480/>;
9. <https://medium.com/@CharlesCrouspeyre/comment-les-r>;
10. <https://eu-ireland-custom-media-prod.s3-eu-west-1.amazonaws.com/France/Download/CERF475>
11. <http://mag.monchval.com/les-differentes-techniques-dimagerie-medicale/>.
12. <http://dspace.univ-tlemcen.dz/bitstream/112/12583/1/Deep-Learning-pour-la-classification..pdf>
13. <https://www.sciencedirect.com/science/article/pii/S0092867418301545figs6>.