

الجمهورية الجزائرية الديمقراطية الشعبية  
République algérienne démocratique et populaire  
وزارة التعليم العالي والبحث العلمي  
Ministère de l'enseignement supérieur et de la recherche scientifique  
جامعة عين تموشنت بلحاج بوشعيب  
Université –Ain Temouchent- Belhadj Bouchaib  
Faculté des Sciences et de Technologie  
Département Mathématiques et informatique



Projet de Fin d'Etudes  
Pour l'obtention du diplôme de Master en Informatique  
Domaine : Mathématiques et Informatique  
Filière : Informatique  
Spécialité : Réseaux et Ingénierie des Données (RID)  
Thème

**SEGMENTATION ET INTERPRÉTATION D'IMAGES MÉDICALES  
PAR APPRENTISSAGE AUTOMATIQUE ET SYSTÈME MULTI-AGENTS**

**Présenté Par :**

- 1) Mr. Benameur Habib
- 2) Melle. Dahmane Hayat

**Devant le jury composé de :**

Mr. Messaoudi Mohamed Amine	MAA	UAT.B.B (Ain Temouchent)	Président
Mr. Saidi Mohamed Réda	MCB	UAT.B.B (Ain Temouchent )	Examineur
Mme. Belgrana Fatima Zohra	MCA	UAT.B.B (Ain Temouchent )	Encadrante

*Année Universitaire 2021/2022*

# Table des matières

Table des figures	iv
Liste des tableaux	vi
Introduction générale	1
<b>1 Traitement d'images et imagerie médicale</b>	<b>3</b>
1.1 Introduction	4
1.2 Partie 01 : Traitement d'images	4
1.2.1 Image numérique	4
1.2.2 Caractéristiques de l'image	4
1.2.3 Système de traitement d'images	7
1.2.4 Acquisition et numérisation	8
1.2.5 Le système de visualisation	8
1.2.6 Les techniques de traitement d'images numériques	8
1.3 Partie 02 : L'imagerie médicale	9
1.3.1 Quelques éléments d'anatomie cérébrale	10
1.3.2 Quelques éléments d'anatomie du thorax	12
1.4 Conclusion	12
<b>2 La segmentation des images médicales et l'apprentissage automatique : état de l'art</b>	<b>14</b>
2.1 Introduction :	15
2.2 Partie 01 : La segmentation des images	15
2.2.1 Définition	15
2.2.2 Type de segmentation	16
2.2.3 Techniques de segmentation	17
2.3 Partie 02 : la segmentation et l'apprentissage automatique	19
2.3.1 Réseaux de neurones artificiels	19
2.3.2 Apprentissage Profond : AP (Deep Learning :DL)	23
2.3.3 Réseaux de Neurones Convolutifs (RNC)	24
2.4 D'autres travaux de segmentation des images médicales via l'apprentissage profond	33
2.5 Conclusion	34

<b>3</b>	<b>Les Systèmes Multi-Agents et la segmentation des images médicales</b>	<b>35</b>
3.1	Introduction . . . . .	36
3.2	Notions de base . . . . .	36
3.2.1	L'agent . . . . .	36
3.2.2	Types d'agents . . . . .	37
3.2.3	L'environnement . . . . .	37
3.2.4	Interaction entre les agents . . . . .	38
3.3	Les systèmes Multi-Agents . . . . .	38
3.3.1	Caractéristiques des systèmes Multi-Agents . . . . .	38
3.3.2	Les approches SMAs destinées à la segmentation . . . . .	39
3.4	Conclusion . . . . .	42
<b>4</b>	<b>Approche proposée, résultats et discussions</b>	<b>43</b>
4.1	Introduction . . . . .	44
4.2	Nos approches proposées . . . . .	44
4.3	Présentation de nos approches proposées . . . . .	46
4.3.1	La première contribution : U-Net 1 . . . . .	46
4.3.2	La deuxième contribution : ResNet50_U-Net . . . . .	47
4.3.3	La troisième contribution : combinaison de six modèles profonds . . . . .	48
4.3.4	La quatrième contribution : double U-Net_SMA . . . . .	49
4.4	Ajustement des paramètres . . . . .	51
4.5	Résultats obtenus, discussions et comparaisons . . . . .	52
4.5.1	Les bases d'images utilisées . . . . .	52
4.5.2	Estimation de performances . . . . .	53
4.5.3	Résultats et discussions : . . . . .	53
4.6	Implémentation . . . . .	58
4.6.1	Environnement de travail . . . . .	58
4.6.2	Présentation de notre application . . . . .	60
4.7	Conclusion . . . . .	62
	<b>Conclusion générale</b>	<b>62</b>
	<b>Appendices</b>	<b>64</b>
.1	Python . . . . .	65
	<b>Bibliographie</b>	<b>66</b>

# Table des figures

1.1	Groupe de pixel formant la lettre A. . . . .	4
1.2	La résolution. . . . .	6
1.3	Histogramme et palette associés à une image [Kaidi, 2017]. . . . .	7
1.4	Schéma d'un système de traitement numérique. . . . .	8
1.5	Image du thorax issue de la radiographie . . . . .	9
1.6	Image du cerveau issue d'une IRM . . . . .	10
1.7	Les composants de cerveau. (a) L'encéphale (b) Une coupe de cerveau [Azouza, 2018].	11
1.8	Système respiratoire . . . . .	13
2.1	Segmentation sémantique .[Kifadji, 2020] . . . . .	16
2.2	Segmentation par instance.[ ichi.pro, 2022] . . . . .	17
2.3	Segmentation panoramique [Ichi.pro., 2022] . . . . .	17
2.4	Croissance progressive des régions. . . . .	18
2.5	Agrégation itérative des blocs similaires au bloc 1. . . . .	19
2.6	Modélisation d'un neurone artificiel. . . . .	20
2.7	Graphe de fonction Sigmoidale . . . . .	20
2.8	Graphe de fonction tangente hyperbolique. . . . .	21
2.9	Exemple d'un réseau de neurones artificiel à propagation avant . . . . .	22
2.10	Perceptron monocouche [Dahmane, 2020]. . . . .	22
2.11	Les Réseaux de Neurones à fonction de Base Radiale RN-FBR . . . . .	23
2.12	Comparaison de méthodes de classification de véhicules de Machine Learning (gauche) et d'apprentissage profond(droite). . . . .	24
2.13	Opération d'une convolution . . . . .	25
2.14	Exemple de fonctionnement de Max pooling et Average pooling . [Asif et Amjad, 2020] . . . . .	26
2.15	Graphe de fonction ReLu. . . . .	26
2.16	opération d'aplatissement. . . . .	27
2.17	Architecture de VGG-16 [Theckedath et Sedamkar, 2020] . . . . .	29
2.18	Bloc convolutif résiduel [Audebert, 2018] . . . . .	30
2.19	Architecture du réseau U-Net, de type encodeur-décodeur. . . . .	31
2.20	Mise à l'échelle composée [Tan & Le, 2019] . (a) . exemple de réseau de base.(b) - Largeur (w) ,(c) Profondeur(d), (d)Résolution (r), (e) Mise à l'échelle composée . . . . .	32
2.21	L'architecture proposée dans [Dolz et al., 2020]en utilisant une fusion précoce. . . . .	33

2.22	La structure de la segmentation proposée dans [Salama & Aly, 2021] . . . . .	34
3.1	Architecture générale d'un agent . . . . .	36
3.2	Système Multi-Agents vu selon différents niveaux de détail . . . . .	39
3.3	Boucle de perception-action des agents avec l'environnement et interactions entre des agents régions et contours qui coopèrent. [Porquet et al., 2003] . . . . .	40
3.4	Le système multi-agents proposé [Semchedine, 2007] . . . . .	41
3.5	l'interaction entre agents dans l'approche développée. [Boulkhiout, 2010] . . . . .	42
4.1	Organigramme des approches proposées . . . . .	45
4.2	L'architecture de notre modèle proposé U-Net1. . . . .	46
4.3	l'architecture de modèle ResNet50_U-Net. . . . .	48
4.4	Architecture de notre troisième contribution "combinaison de résultats issus de six modèles profonds". . . . .	49
4.5	Architecture de notre deuxième modèle hybride proposé Double U-Net_Sma. . . . .	49
4.6	L'architecture de notre modèle proposée U-Net 2. . . . .	50
4.7	Modélisation de l'interaction entre agent 2 et agent 3 dans notre approche proposée double U-Net_SMA avec le réseau de pétri. . . . .	51
4.8	Exemple d'augmentation de donnée sur une image de la base Br35H. . . . .	52
4.9	Comparaison de résultats de segmentation d'une image de la base Br35 [kaggle1, 2022]. via les modèles proposées U-Net 1, ResNet50_U-Net. (A) : image originale ,(B) : U-Net 1,(C) : ResNet50_U-Net , (D) :Masque(vérité terrain). . . . .	54
4.10	Résultats de segmentation d'une image de la base Br35 [kaggle1, 2022]. via le modèle proposée combinaison de six modèles. (A) : image original ,(B) : Masque(vérité terrain) ,(C) : notre combinaison proposée . . . . .	55
4.12	Résultats de segmentation par le modèle proposé Resnet50_U-Net sur des images 3D de la base CT Chest [kaggle2., 2022].(A) : verité terrain, (B) segmentation via approche hybride proposée ResNet50_U-Nnet . . . . .	56
4.13	Résultat de segmentation obtenu via notre approche hybride proposée Double U-Net_SMA sur une image 2D de la base Brats 2020, (A) : Resnet50, (B) : segmentation via le model double U-Net_SMA, (C) : vérité terrain . . . . .	56
4.14	Résultat de segmentation via notre approche hybride proposée Double U-Net_SMA sur une image 3D de la base Brats 2020. (A) : Vérité terrain, (B) Model double U-Net_SMA. . . . .	57
4.15	Interface d'accueil de notre application. . . . .	60
4.16	Chargement d'une image de la base CT Chest segmentation. . . . .	60
4.17	Affichage des résultats de segmentation obtenus via notre modèle ResNet50 et le modèle EfficientNet sur des images de la base CT Chest segmentation. . . . .	61
4.18	Affichage des résultats obtenus de notre modèles ResNet50 et le modèle Double U-Net sur des images de la base BRATS2020. . . . .	61

# Liste des tableaux

4.1	Comparaison des résultats obtenus sur les deux modèles proposés : U-Net 1 et ResNet50_U-Net avec différentes époques. . . . .	54
4.2	Résultats de segmentation par le modèle ResNet50_U-Net sur des images 2D de la base CT Chest [kaggle2., 2022]. (A) : vérité terrain, (B) : segmentation via approche hybride proposée ResNet50_U-Net . . . . .	55
4.3	Comparaison des résultats obtenus sur les deux modèles de base (Efficient net et Resnet50)	56
4.4	Comparaison des résultats de segmentation obtenus via nos deux approches hybrides proposées Double U-Net_SMA et ResNet50_U-Net. . . . .	57
4.5	Comparaison en matière de temps d'exécution du modèle Double U-Net avec et sans SMA . . . . .	57

# Remerciements

Louanges tout d'abord à ALLAH qui est l'origine de toute réussite dans notre vie.

Nous souhaitons adresser nos sincères remerciements à notre encadrante, madame Fatima Zohra Belgrana pour ses conseils et son encadrement. Après à tous ceux en qui par leur présence, leur soutien, leur disponibilité et leurs conseils nous avons trouvé courage afin d'accomplir ce projet.

Nous voulons aussi remercier les membres de jury d'avoir accepté d'examiner et d'évaluer notre travail. Enfin, nos remerciements s'adressent aussi à tous ceux qui ont participé, de près ou de loin, à l'élaboration de ce projet de fin d'études et en particulier à nos familles et nos amis.

# LISTE DES ABREVIATION

- AA** Apprentissage Automatique
- AP** Apprentissage profond
- Adam** Adaptive Moment Estimation
- CONV** Couche de convolution
- DL** Deep learning
- IA** Intelligence artificielle
- IAD** Intelligence artificielle distribuée
- IRM** L'imagerie par résonance magnétique
- IoU** L'intersection sur l'union
- MVS** Machines à Vecteurs de Support
- PR** Precision
- RNA** Réseaux de neurones artificiels
- RN- FBR** Réseaux Neurone à Fonction de Base Radiales
- RNC** Réseaux de Neurones Convolutifs
- RNN** Réseaux de neurones récurrents
- ReLU** Rectification Linéaire Unit
- SMA** Les systèmes multi agents
- TFP** Taux de Faux Positifs

# Introduction générale

Le traitement d'images est une étape importante dans la visualisation, l'analyse, l'interprétation, le stockage et la transmission des images. Elle intervient dans divers domaines tels que de la sécurité (le tatouage numérique, la biométrie, . . .), le domaine militaire (télé-détection, la simulation, . . .) ou encore le domaine médical (détection de tumeurs, suivis de maladies, . . .), ce dernier représente notre domaine d'application.

Avec le développement technologique des systèmes d'acquisition d'images, notamment dans le domaine médical, l'imagerie médicale est en constante évolution ces dernières années, fournissant un nombre croissant de données. Il est question alors, d'assister l'expert dans l'analyse qualitative et quantitative des images afin de lui faciliter leur interprétation. Pour d'aboutir à cela plusieurs méthodes et approches issues du traitement et d'analyse d'images existent, parmi, on retrouve, la segmentation, une étape de haut niveau dans ce processus de traitement des images.

Nous nous sommes situés tout au long de notre étude, dans une approche de segmentation sémantique, qui consiste à attribuer chaque pixel à une classe d'objets (étiqueter le pixel). Cette technique est devenue un élément essentiel de l'analyse d'image de nos jours car elle facilite la détection, la description et la visualisation des zones d'intérêt dans l'image. Le domaine médical est très sensible, il est question par conséquent de maximiser les performances de la segmentation.

Ces dernières années avec la révolution de l'apprentissage profond, de nombreux problèmes de classification et de segmentation sémantique ont été résolus à l'aide ces d'architectures dites profondes, en particulier via les réseaux de neurones convolutifs, qui ont surpassé de nombreuses autres méthodes en matière de précision et d'efficacité, raison pour laquelle nous avons opté dans notre travail pour cette approche

La notion de temps est aussi importante surtout dans le cas des tranches d'images 3D, pour pallier à cela nous avons fait recours aux Systèmes Multi-Agents (SMA), il s'agit d'un paradigme comportant plusieurs entités qui ont un comportement mutuel comme la coopération, la coordination et la négociation. Il offre une architecture de contrôle, une distribution d'expertise et un parallélisme. Ils sont largement utilisés dans divers domaines tels que la robotique et la vision par ordinateur.

Dans ce mémoire nous proposons quatre contributions de segmentation sémantique des images médicales basée sur un apprentissage profond, plus particulièrement en utilisant une architecture d'un réseau de neurones convolutif nommée U-Net ainsi que le modèle de base ResNet. Nos approches proposées sont entraînées et testées avec trois bases données contenant des images du thorax et du cerveau, en 2D et 3D.

Notre mémoire est structuré en quatre chapitres, nous commençant tout d'abord par une introduction générale afin d'introduire notre problématique et nos objectifs.

Dans le premier chapitre, nous avons introduit quelques notions de base sur le traitement d'images, ensuite nous présentons les anatomies des organes dont nous avons effectué une segmentation à savoir, le cerveau et le thorax.

Le deuxième chapitre est consacré à la description des approches de segmentation d'images ainsi qu'aux techniques de l'apprentissage automatique et d'apprentissage profond, où nous avons présenté un état de l'art dans ce contexte. Dans le troisième chapitre, nous avons donné un aperçu sur les systèmes Multi-Agents.

Dans le quatrième et dernier chapitre, nous avons illustré nos approches hybrides et modèles proposées, puis nous discuter nos différents résultats obtenus. Nous clôturons ce mémoire par une conclusion générale.

# Chapitre 1

## Traitement d'images et imagerie médicale

## 1.1 Introduction

La vision par ordinateur est une branche de l'informatique et plus précisément de l'intelligence artificielle, qui est consacrée au développement d'algorithmes pour traiter des informations visuelles (images) qui peuvent être obtenues en utilisant diverses techniques (cameras, microscopes, scanner ...). Or, le traitement des images médicales permet d'apporter une aide précieuse quant au dépistage de maladies. La segmentation est une étape importante dans ce processus.

Dans ce premier chapitre, nous abordons les notions de base nécessaires à la compréhension des techniques du traitement d'images. Ensuite, nous allons donner un aperçu sur les différents éléments d'anatomie cérébrale et du thorax, domaine d'application que nous avons opté pour.

## 1.2 Partie 01 : Traitement d'images

### 1.2.1 Image numérique

L'image numérique est l'image dont la surface est divisée en éléments de tailles fixes appelés cellules ou pixels, ayant chacun comme caractéristique un niveau de gris ou une couleur prélevée à son emplacement correspondant. La numérisation d'une image est la conversion de celle-ci de son état analogique en un état numérique représenté par une matrice bidimensionnelle de valeur numérique  $f(x,y)$  [Adoum & Daoussa, 2014]

### 1.2.2 Caractéristiques de l'image

Comme nous l'avons vu, l'image est un ensemble structuré d'informations, parmi ses caractéristiques nous pouvons citer les paramètres suivants :

#### 1.2.2.1 Le pixel

Un pixel est une valeur numérique qui représente des intensités lumineuses, c'est le plus petit élément que le matériel et le logiciel peuvent manipuler sur une image. Exemple, La lettre A est affichée comme un groupe de pixels dans la figure ci-dessous.

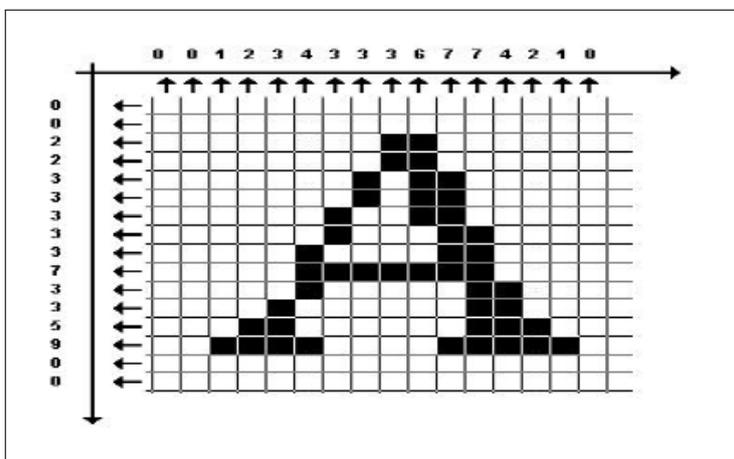


FIGURE 1.1 – Groupe de pixel formant la lettre A.



### 1.2.2.2 Résolution et définition

Les termes résolution et définition sont liés à la notion de pixel, ils définissent la netteté et la qualité d'une image [Guesmia, 2019].

- La définition est le nombre de pixels que comporte une image numérique en largeur et en hauteur. Par exemple une image dont la définition est  $1500 \times 1200$  correspond à une image de 1500 pixels en largeur et 1200 pixels en hauteur.
- La résolution d'une image est le nombre de pixels contenu dans l'image par unité de longueur, elle s'exprime le plus souvent en PPP (Point Par Pouce) ou dpi (dots per inch), et parfois en points par cm. Sachant qu'un pouce est égal à 2,54 cm .

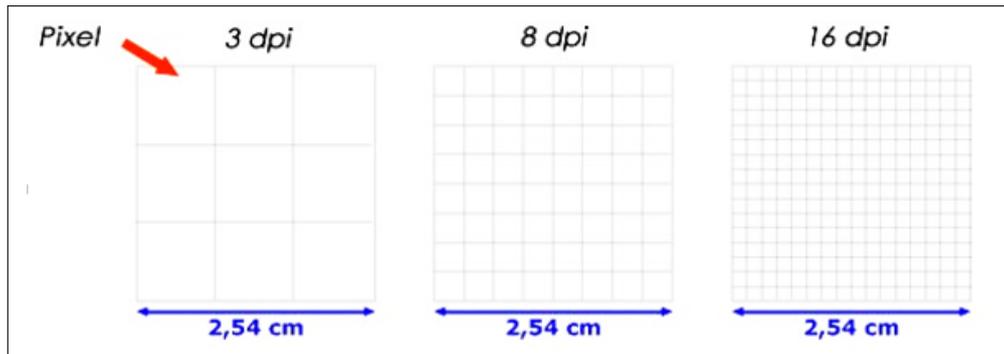


FIGURE 1.2 – La résolution.

### 1.2.2.3 La texture

La texture décrit les propriétés d'une région dans une image numérique, qui a des caractéristiques homogènes, ces caractéristiques sont par exemple un motif basique qui se répète [Koudri & Mohammedi, 2021]. Ils existent trois principales familles de texture à savoir, la texture structurale, la texture aléatoire et la textures directionnelles [Majdoulayane, 2009].

### 1.2.2.4 Le 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 [Koudri & Mohammedi, 2021].

### 1.2.2.5 La luminance

La luminance (brillance) est définie comme la moyenne de l'intensité de tous les pixels de l'image, elle représente le quotient de l'intensité lumineuse d'une surface par l'aire appartenant de cette surface [Azouza, 2018].

### 1.2.2.6 Histogramme

L'histogramme des niveaux de gris ou des couleurs d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (ou couleur) dans l'image. Il permet de donner un grand nombre

d'informations sur la distribution des niveaux de gris (ou couleur) et de voir entre quelles bornes est répartie la majorité des niveaux de gris (ou couleur) dans le cas d'une image trop claire ou d'une image trop foncée [Kaidi, 2017].

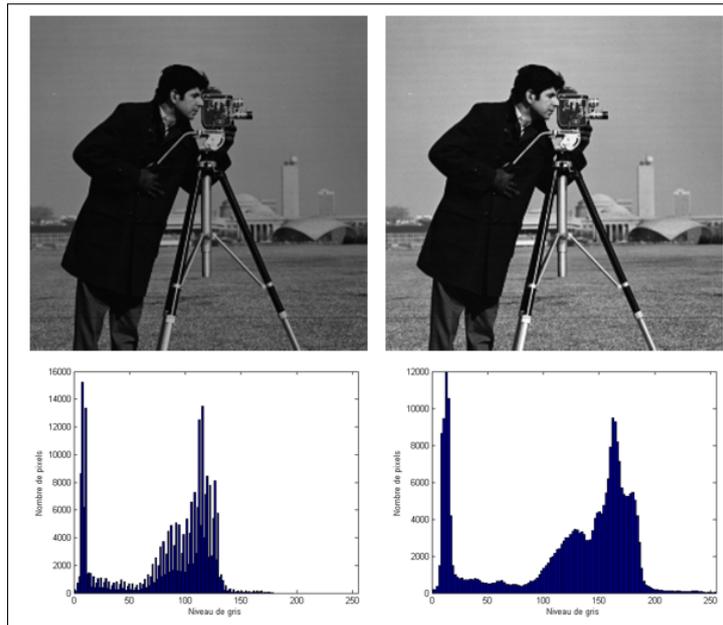


FIGURE 1.3 – Histogramme et palette associés à une image [Kaidi, 2017].

### 1.2.2.7 Le contraste

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images. Si  $L1$  et  $L2$  sont les degrés de luminosité respectivement de deux zones voisines  $A1$  et  $A2$  d'une image, le contraste  $C$  est défini par le rapport exprimé comme suit [Mokri, 2017] :

$$C = \frac{L1 - L2}{L1 + L2} \quad (1.1)$$

### 1.2.3 Système de traitement d'images

Un système de traitement d'images est généralement composé des unités suivantes [Koudri & Mohammedi, 2021] :

- Un système d'acquisition et de numérisation qui permet d'effectuer l'échantillonnage et la quantification d'une image ;
- Une mémoire de masse pour stocker les images numérisées ;
- Un système de visualisation ;
- Une unité centrale permettant d'effectuer les différentes opérations de traitement d'images.

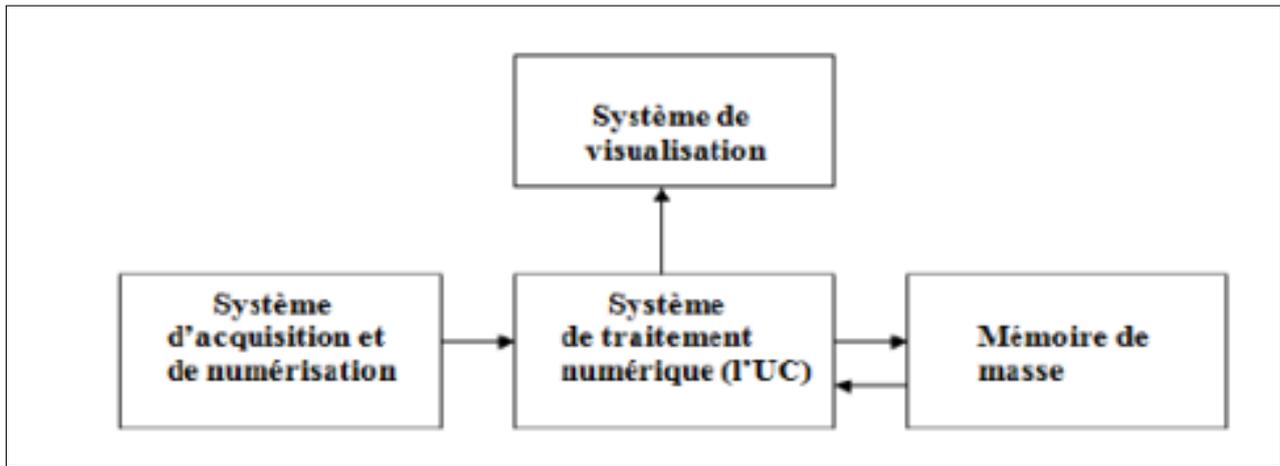


FIGURE 1.4 – Schéma d'un système de traitement numérique.

### 1.2.4 Acquisition et numérisation

L'acquisition des images constitue un des maillons essentiels de toute chaîne de conception et de production d'images. Pour pouvoir manipuler une image sur un système informatique, il est avant tout nécessaire de lui faire subir une transformation qui la rendra lisible et manipulable par ce système. Le passage de cet objet externe (l'image d'origine) à sa représentation interne (dans l'unité de traitement) se fait grâce à une procédure de numérisation [Benfriha & Hamel, 2016].

### 1.2.5 Le système de visualisation

Tout système de traitement d'image est doté d'un dispositif de visualisation qui permet l'affichage des images. L'utilisation de différents types de reconstituteurs permet de transformer le signal numérique qu'est la matrice image en un signal analogique visible par l'œil de l'observateur. Pour cela, différents types de supports peuvent être employés : moniteur vidéo, clichés photographiques, impression sur papier. Dans tous les cas et pour chaque échantillon de l'image numérique, on recrée un nouvel élément d'image ou un nouveau pixel dont on choisit la forme de façon à reconstituer une image analogique qui soit la plus proche possible de l'image avant numérisation compte tenu des erreurs introduites lors de l'acquisition, de la numérisation et de la transmission [Benfriha & Hamel, 2016].

### 1.2.6 Les techniques de traitement d'images numériques

Les techniques de traitement sont destinées à l'exploitation des informations contenues dans les images, ceci dans le but d'améliorer la qualité des images et de les rendre plus facilement interprétables, en d'autres termes elles permettent d'augmenter la qualité visuelle de l'image. Parmi les techniques de traitement d'images [Benierbah & Khamadja, 2017] nous avons :

- 1) **Le prétraitement** : les opérations de prétraitement sont toutes les techniques visant à améliorer la qualité d'une image ;
- 2) **Le restauration** : amélioration de l'image ayant subi des dégradations (une caméra qui a bougé) ;
- 3) **La segmentation** : procédure de partitionnement de l'image en régions homogènes au sens d'un certain critère, représentant différents objets d'intérêt.

## 1.3 Partie 02 : L'imagerie médicale

L'imagerie médicale regroupe différentes techniques permettant d'observer l'intérieur du corps humain. Il existe de multiples méthodes d'acquisition et de restitution d'images médicales, adaptées à l'observation de diverses maladies et permettant de réaliser différents objectifs. Nous allons voir dans ce qui suit quelques-unes.

### a) La radiographie :

La radiographie permet de visualiser un segment du corps en deux dimensions. Son principal inconvénient est l'irradiation du patient, bien que les techniques actuelles diminuent considérablement la quantité d'irradiation délivrée [Mabrouk, 2013].

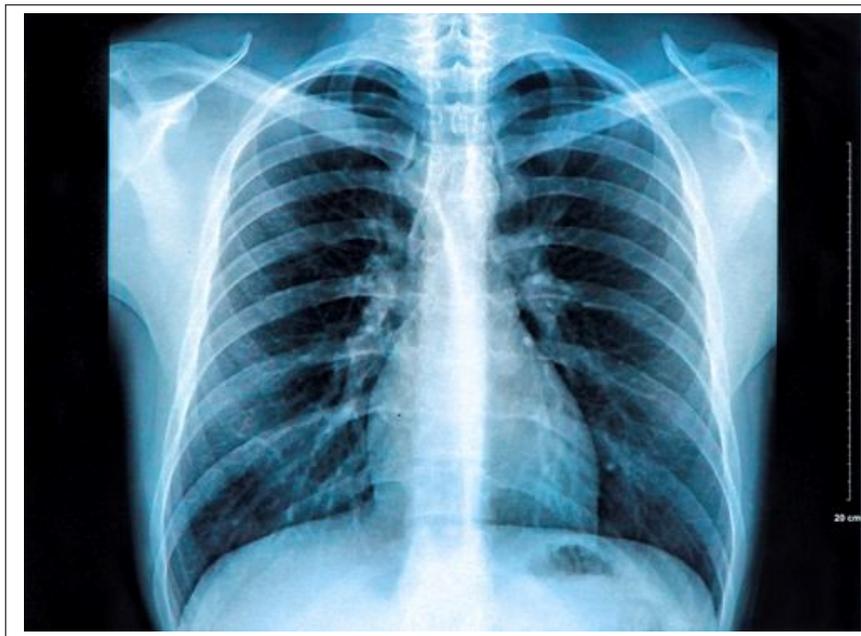


FIGURE 1.5 – Image du thorax issue de la radiographie .

### b) Le scanner :

Le scanner permet de visualiser une partie du corps en faisant défiler les coupes, il est utilisé principalement pour étudier le cerveau, le thorax l'abdomen et le pelvis. Il nécessite parfois l'injection d'un produit de contraste à base d'iode afin de permettre une meilleure étude des structures observées [Mabrouk, 2013].

c) **L'Imagerie par Résonance Magnétique (IRM) :** L'imagerie par résonance magnétique est une technique récente qui permet de produire des images anatomiques dans des plans multiples et qui peut fournir des informations sur la composition des tissus. La première acquisition d'IRM a été produite en 1973 par Lauterbur et Damadian [Seynaeve & Broos, 1995]. Cette technique a permis d'arriver à des avancées technologiques et scientifiques majeures notamment pour le dépistage des pathologies [Laguel, 2010].

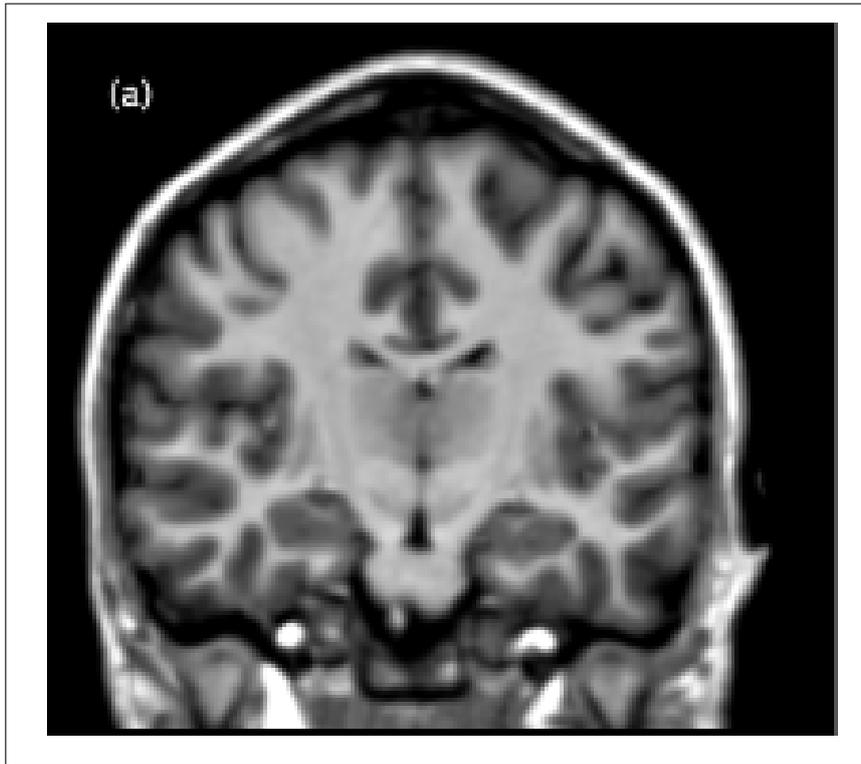


FIGURE 1.6 – Image du cerveau issue d'une IRM .

### 1.3.1 Quelques éléments d'anatomie cérébrale

L'encéphale est la partie du système nerveux qui loge dans la boîte crânienne et repose à la base du crâne. Son poids moyen est de 1400 à 1800 grammes. Il est enveloppé et protégé par des membranes appelées méninges. Il comprend le cerveau, le cervelet et le tronc cérébral (voir la figure (1.7 (a))).

#### a) Le cerveau

Le cerveau est la partie la plus volumineuse du système nerveux central. Il est constitué essentiellement de deux hémisphères, séparés par la scissure interhémisphérique, et reliés par différentes structures telles que le corps calleux, le thalamus et l'hypothalamus [Laguel, 2010].

#### b) Le cervelet

Le cervelet situé en parallèle sur deux grandes voies nerveuses (celles qui amènent les messages sensoriels vers les zones du cortex qui en font l'analyse et celles qui partent du cortex et descendent vers les muscles pour les faire contracter). Le cervelet est en partie impliqué dans la coordination des mouvements du corps [Azouza, 2018].

#### c) Le tronc cérébral

Représente un carrefour majeur du système nerveux central entre le cerveau et le cervelet. Il est responsable en particulier de la régulation de la respiration et du rythme cardiaque. C'est également un centre de passage des voies motrices et sensibles, ainsi qu'un centre de contrôle de la douleur [Azouza, 2018].

### 1.3.1.1 Les composants du cerveau

Si on regarde une coupe du cerveau (voir la figure (1.7 (b))), on remarque un ruban gris qui borde la surface, c'est la substance grise ou cortex. Les neurones envoient leurs prolongements, ou axones, vers l'intérieur du cerveau. Comme ces prolongements apparaissent plus clairs sur la coupe, ils constituent ce qu'on appelle la substance blanche. Ces axones relient une région cérébrale à une autre, se dirigent vers l'extérieur pour acheminer les commandes motrices ou amènent au cerveau les informations sensorielles (vision, audition,...). Les régions motrices, qui commandent nos mouvements, et les régions tactiles, qui ressentent le contact sur la peau, se trouvent au centre de la tête de chaque côté du grand sillon qui sépare les régions frontales à l'avant et les régions pariétales à l'arrière (le sillon central). Ces grands faisceaux sensoriels et moteurs sont croisés, c'est-à-dire que l'hémisphère gauche commande le côté droit du corps, et reçoit les informations tactiles du côté droit [Azouza, 2018]. Le cerveau est alors composé de trois principales substances à savoir la matière blanche (axones), la matière grise (cortex) et le liquide céphalo-rachidien.

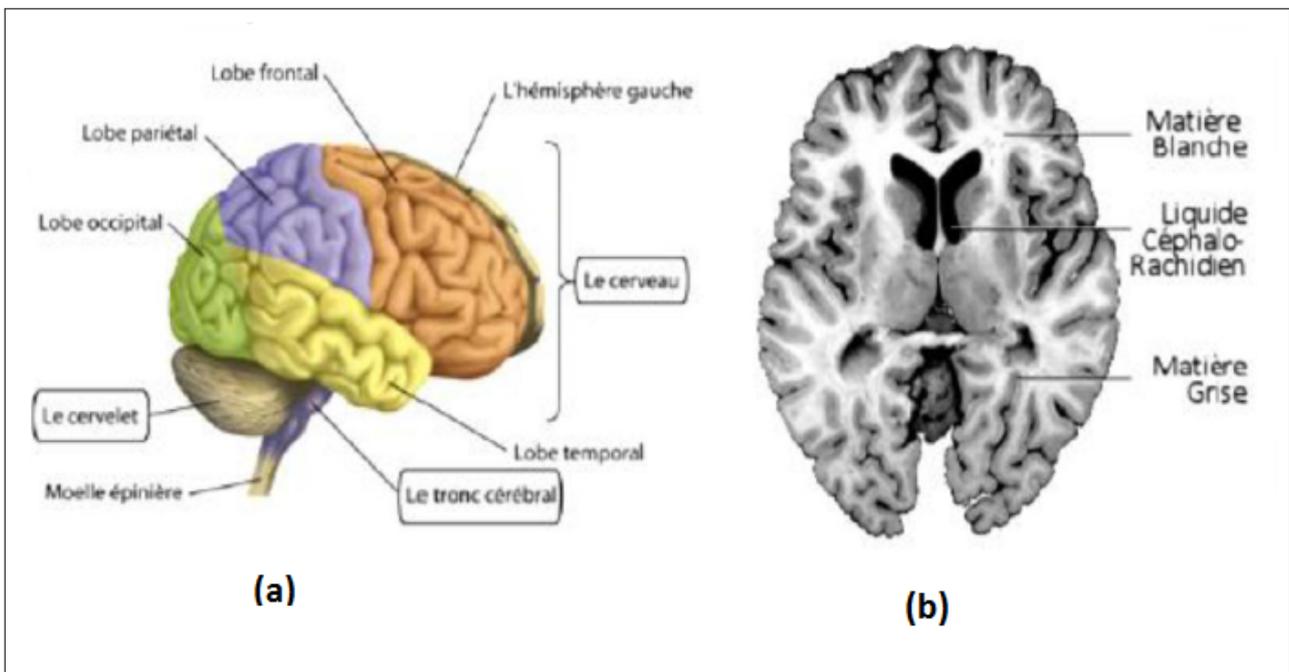


FIGURE 1.7 – Les composants de cerveau. (a) L'encéphale (b) Une coupe de cerveau [Azouza, 2018].

### 1.3.1.2 Les tissus cérébraux

a) **Le Liquide Céphalo-Rachidien (LCR)** : Le LCR est le liquide dans lequel baignent le cerveau et le cervelet, d'un volume moyen de 150 ml, il est absorbé par le système veineux cérébral et continuellement renouvelé. Ses principaux rôles sont [Scherrer, 2008] :

- la protection mécanique du système nerveux central contre les chocs par amortissement des mouvements ;
- la protection contre les infections, car il contient les médiateurs de l'immunité humorale et cellulaire
- le transport des hormones entre les différentes régions éloignées du cerveau.

b) **La matière grise (MG) :**

On peut distinguer plusieurs structures composées de la MG, d'une part, on trouve le cortex, situé à la périphérie du cerveau et d'autre part les noyaux gris centraux. La matière grise est essentiellement composée de neurones et constitue donc le siège de l'activité cérébrale [Laguel, 2010].

c) **La matière blanche (MB) :**

La MB (ou substance blanche) se situe dans la partie interne du cerveau. Elle correspond à la gaine de myéline qui recouvre les axones des neurones pour en accélérer la conduction. Ces axones myélinisés s'assemblent en faisceaux (l'équivalent des nerfs) pour aller établir des connexions avec d'autres groupes de neurones [Laguel, 2010].

### 1.3.2 Quelques éléments d'anatomie du thorax

Le thorax est la région du corps située entre le cou et le diaphragme. Sa composante squelettique est appelée la cage thoracique. La cage thoracique est une paroi osseuse expansive pour permettre la respiration. Elle est composée de cartilage et d'os, avec une structure en forme de cage qui lui permet d'enfermer et de protéger les principaux organes de circulation sanguine et de respiration, le cœur et les poumons [Moreau, 2018].

#### 1.3.2.1 Poumons, Cœur et Trachée

- a) **Les poumons :** Les poumons, entourés chacun par une membrane appelée la plèvre, sont les principaux organes de la respiration permettant l'échange des gaz vitaux en fournissant l'oxygène à tout le corps et en éliminant le dioxyde de carbone [Bodo, 2017].
- b) **La trachée :** C'est un conduit fibro-cartilagineux aérifère, impair et médian. Elle fait suite au larynx et se termine dans le thorax en se bifurquant en 02 bronches principales (bronches souches). Elle a pour rôle de véhiculer l'air vers les poumons [Bensalem, 2021].
- c) **Le cœur :** Le cœur est un organe vital creux et musculaire, qui assure la circulation sanguine dans le corps grâce à des contractions régulières. Il est composé de quatre cavités : deux oreillettes droite et gauche et deux ventricules droit et gauche, entourés dans une enveloppe, le péricarde. Les oreillettes fonctionnent comme des réservoirs pour le sang, tandis que les ventricules pompent le sang à travers le système vasculaire en se contractant et en se relaxant d'une manière rythmée [Bodo, 2017]. Le cœur fonctionne comme une pompe qui, grâce à ses contractions régulières, propulse le sang dans tout l'organisme et assure ainsi l'alimentation en oxygène du corps entier.

## 1.4 Conclusion

Nous avons présenté dans ce chapitre quelques concepts liés au domaine du traitement d'images, où nous avons défini des notions élémentaires utilisées dans cette discipline. Plusieurs applications sont issues de cette dernière, telles que le prétraitement, la restauration, l'interprétation et la segmentation.

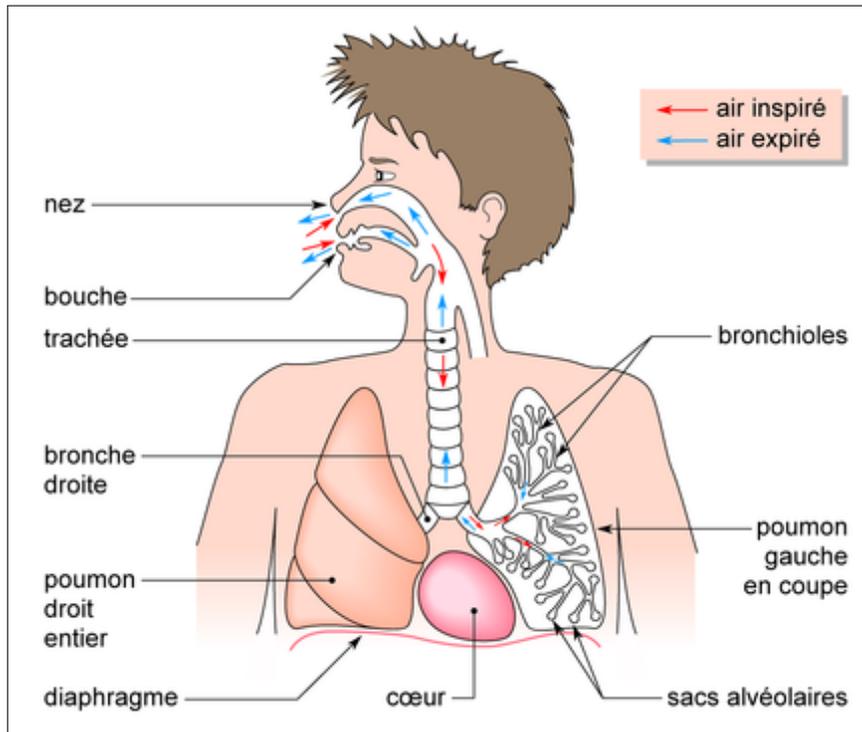


FIGURE 1.8 – Système respiratoire .

Ces applications interviennent dans plusieurs domaines, notamment celui de la médecine, or, l'imagerie médicale joue un rôle important lors de l'établissement du diagnostic, nous avons par conséquent présenté quelques techniques utilisées dans ce domaine à savoir l'IRM, le scanner et la radiographie.

Dans le chapitre suivant, nous allons présenter un état de l'art sur la segmentation d'images, une technique issue du traitement d'images et qui représente le sujet de notre étude. Dans ce prochain chapitre, nous allons focaliser sur la segmentation des images médicales via l'apprentissage automatique et l'apprentissage profond plus particulièrement.

## Chapitre 2

La segmentation des images médicales et  
l'apprentissage automatique : état de l'art

## 2.1 Introduction :

La segmentation des images médicales joue un rôle primordial dans l'élaboration d'un système d'aide au diagnostic afin de fournir au médecin des informations de haut niveau sur le contenu de l'image.

Un des objectifs principaux est alors d'aboutir à des résultats aussi performants que possible, tout en précisant les différents volumes anatomiques qu'ils soient normaux ou pathologiques.

La segmentation permet alors d'apporter une aide pour la détection des affections tumorales telles que le cancer ou au diagnostic des maladies telles que l'Alzheimer ou la sclérose en plaques.

L'apprentissage profond peut être considéré comme un sous-ensemble de l'apprentissage automatique. C'est un domaine basé sur l'apprentissage et l'amélioration par lui-même en examinant les algorithmes informatiques. Il a aidé à la classification des images, à la traduction de la langue et à la reconnaissance vocale. Il peut être utilisé pour résoudre tout problème de reconnaissance ; de formes et sans intervention humaine.

Ce chapitre est consacré à un état de l'art sur la segmentation des images médicales via l'apprentissage profond. Il est divisé en deux parties, dans une première partie, nous aborderons les différents types et techniques de segmentation d'images. Dans une deuxième partie, nous verrons l'apprentissage automatique où nous présenterons quelques notions de base concernant les réseaux de neurones artificiels, puis nous exposerons quelques concepts théoriques liés l'apprentissage en profondeur, plus précisément les réseaux de neurones convolutifs.

## 2.2 Partie 01 : La segmentation des images

### 2.2.1 Définition

La segmentation d'image est une phase d'importance capitale dans tout processus de vision par ordinateur. Le processus de segmentation consiste à découper l'image en région dans un but de simplifier sa représentation, pour la rendre plus facile à analyser et à interpréter. La segmentation d'images est typiquement utilisée pour localiser des objets ou reconnaître des contours d'objets. Techniquement, c'est un processus d'étiquetage des pixels couvrant la totalité de l'image, où un ensemble de contours délimitant des régions représentant une certaine homogénéité de couleur, d'intensité et de texture [Bendaoud, 2017].

Il existe plusieurs domaines d'application du traitement d'images, tels que l'imagerie satellitaire [Neupane et al., 2021], l'identification et vérification de personnes [Mekruksavanich & Jitpattanakul, 2021], les systèmes de contrôle de la circulation [Seo et al., 2022] ainsi l'imagerie médicale [Mirsky et al., 2019].

## 2.2.2 Type de segmentation

Il existe trois principaux types de segmentation d'images :

- a) **Segmentation sémantique** La segmentation sémantique est une tâche particulièrement importante en vision par ordinateur. Plusieurs jeux de données ont été introduits pour comparer différentes méthodes de segmentation sur des images multimédias, de conduite autonome, de télédétection et médicales [Audebert et al., 2018]. La segmentation sémantique associe une étiquette ou une catégorie à chaque pixel d'une image. Elle permet de reconnaître des ensembles de pixels qui forment des catégories distinctes.

La séparation d'images en deux classes est un exemple simple de segmentation sémantique. Par exemple, à la figure (2.1), une image présentant une personne à la plage est associée à une version montrant les pixels de l'image segmentés en deux classes distinctes : la personne et l'arrière-plan [Kifadji, 2020].

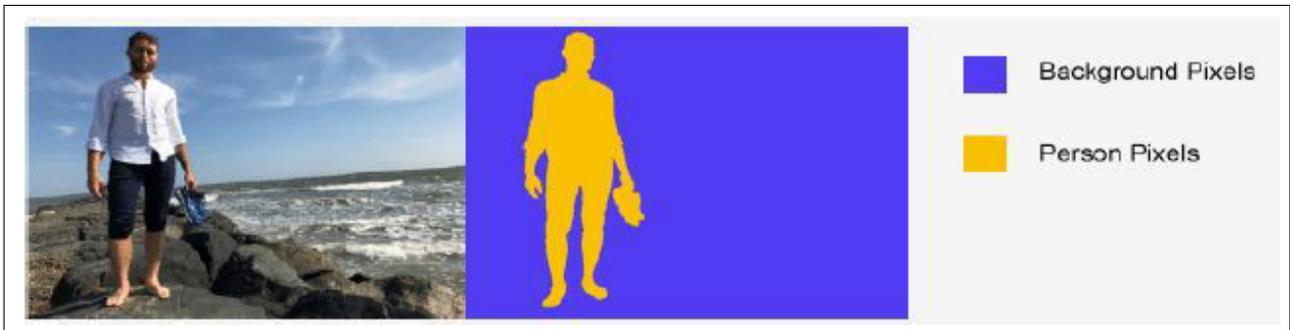


FIGURE 2.1 – Segmentation sémantique .[Kifadji, 2020]

- b) **Segmentation par instance** Il s'agit d'une annotation combinée de la détection de cible et de la segmentation sémantique. La détection de la cible vient en premier, puis chaque pixel est étiqueté (segmentation sémantique). Une segmentation par instance est illustrée par la figure 2.2, où la personne est prise comme objet cible. Par contre dans une segmentation sémantique, il n'y a pas une distinction entre les différentes instances dans la même catégorie. [Ichi.pro., 2022].

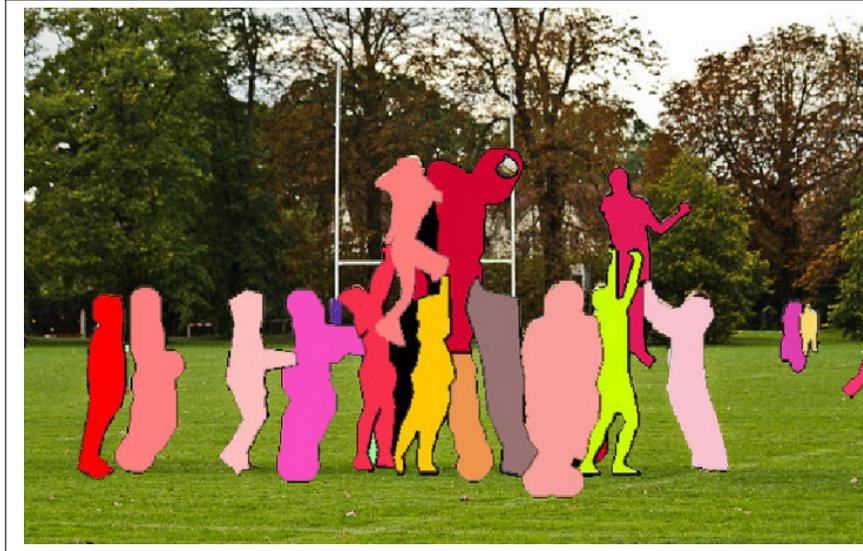


FIGURE 2.2 – Segmentation par instance.[ ichi.pro, 2022]

- c) **Segmentation panoramique** Une autre combinaison d’annotation de segmentation sémantique et de segmentation d’instance, mais plus complexe. La segmentation panoramique signifie que toutes les cibles doivent être détectées et que différentes instances de la même catégorie doivent être distinguées (voir figure 2.3).

La segmentation panoramique consiste à détecter et à segmenter tous les objets de l’image, y compris l’arrière-plan, et à distinguer les différentes instances (différentes couleurs sont utilisées) [Ichi.pro., 2022].

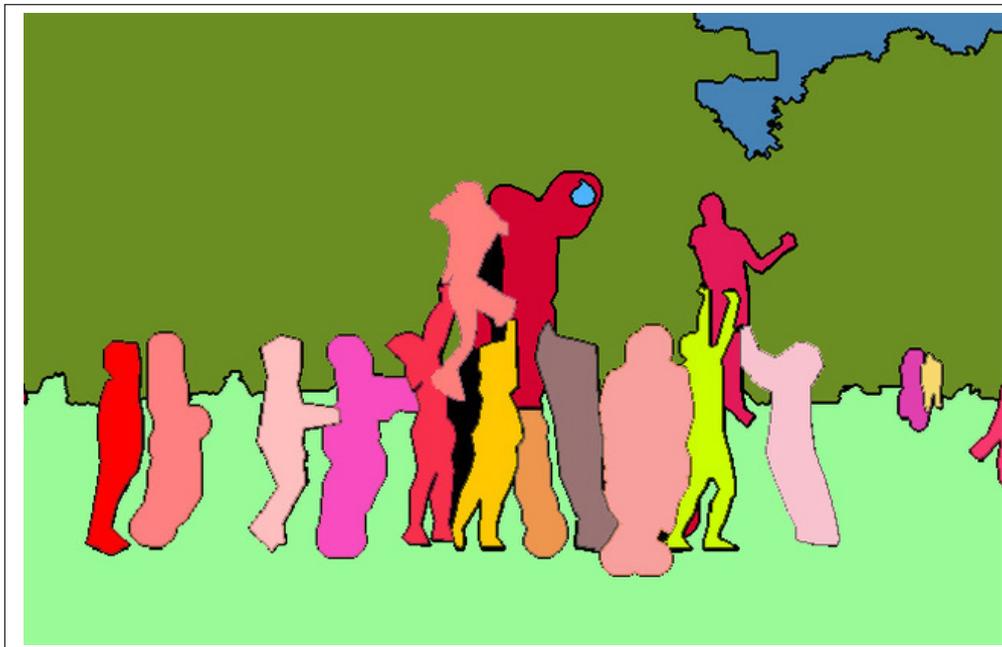


FIGURE 2.3 – Segmentation panoramique [Ichi.pro., 2022]

### 2.2.3 Techniques de segmentation

La segmentation est une étape primordiale en traitement d’images, à ce jour, il existe de nombreuses méthodes de segmentation, que l’on peut regrouper en trois principales classes :

- Segmentation par région :

Les approches de segmentation régions visent à créer une partition de l'image en un ensemble de régions homogènes au sens d'un ou plusieurs critères pour avoir des ensembles de pixels partageant des propriétés communes. Les régions sont différenciées entre elles par des propriétés élémentaires fondées sur des critères locaux tels que le niveau de gris de chaque pixel, ou bien sur un attribut estimé dans le voisinage du pixel tel que la valeur moyenne du niveau de gris, la variance, ou des paramètres de texture [Kifadji, 2020].

- Seuillage :

Le seuillage d'une image, est la méthode la plus simple de segmentation de cette image, le seuillage est une technique qui repose sur une mesure quantitative d'une grandeur. Il permet de classer les pixels selon un seuil fixé, si un pixel a une valeur supérieure à ce seuil, il prendra la valeur 1 (blanc) sinon il prendra la valeur zéro (noir) [Bendaoud, 2017].

- Croissance de régions :

Cette technique consiste à faire progressivement accroître les régions autour de leur point de départ. Le principe de l'agrégation de pixel consiste à choisir un germe (le point de départ est le choix d'un ensemble de pixels appelés "germes" et on fait croître ce germe tant que des pixels de son voisinage vérifient le test d'homogénéité. Lorsqu'il n'y a plus de pixels candidats dans le voisinage, on choisit un nouveau germe et on répète le processus [Larbi, 2018].



FIGURE 2.4 – Croissance progressive des régions.

- Division/ fusion :

ces méthodes combinent deux méthodes , la division de l'image en de petites régions homogènes, puis la fusion des régions connexes et similaires au sens d'un prédicat de regroupement. Deux régions seront fusionnées si elles répondent aux critères de similarité des niveaux de gris et d'adjacence de régions. On s'arrête quand le critère de fusion n'est plus vérifié [Benfriha & Hamel, 2016].

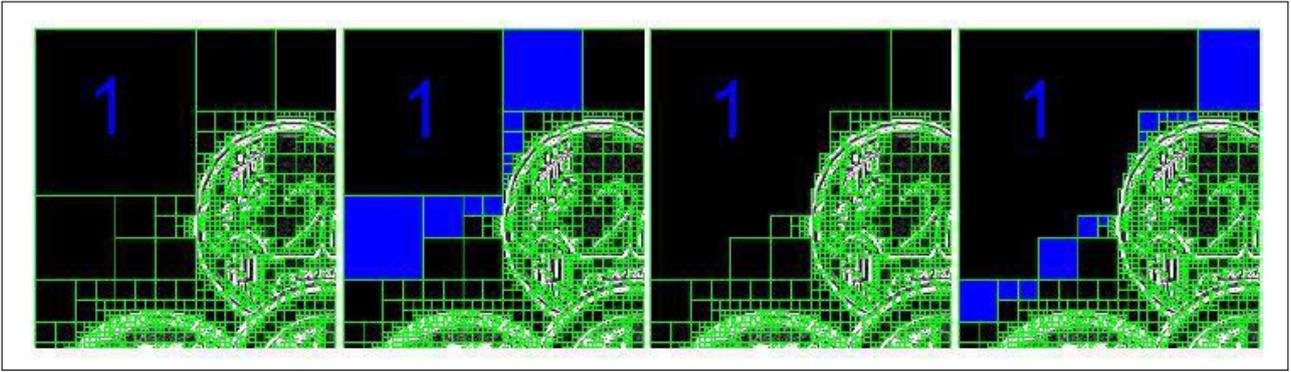


FIGURE 2.5 – Agrégation itérative des blocs similaires au bloc 1.

- Morphologie mathématique :

La morphologie mathématique est une théorie essentiellement non linéaire, particulièrement utile pour filtrer, segmenter et quantifier des images. Initialement destinée au traitement des images binaires. La morphologie mathématique offre un grand nombre d'outils très puissants de traitement et d'analyse d'images que l'on retrouve sous différentes rubriques dans certains logiciels d'analyse d'images et même de retouches d'images, dont le but est l'étude des objets en fonction de leurs formes, de leurs tailles, des relations avec leurs voisinages (en particulier topologiques), de leurs textures et de leurs niveaux de gris ou de leurs couleurs. Par les transformations qu'elle propose, elle se situe à différents niveaux du traitement d'images (filtrage, segmentation, mesure, analyse de texture) et définit ainsi des outils pour la reconnaissance des formes [Adoum & Daoussa, 2014].

- Segmentation par classification :

La segmentation par classification consiste à partitionner une image en un ensemble de classes disjointes, pour chaque pixel de l'image on attribue une étiquette parmi l'ensemble des étiquettes qui correspondent chacune à une classe. On dit qu'une classification est supervisée si le nombre de classes est connu a priori sinon il s'agit d'une classification non supervisée [Guesmia, 2019].

## 2.3 Partie 02 : la segmentation et l'apprentissage automatique

### 2.3.1 Réseaux de neurones artificiels

Les réseaux de neurones artificiels (RNA) sont une classe d'algorithmes effectuant un apprentissage automatique. Ils sont inspirés par le système nerveux biologique.

#### 2.3.1.1 Principe de base

L'unité de base du calcul dans un réseau de neurones artificiel est le neurone qui reçoit des entrées de certains autres neurones ou d'une source externe ayant des valeurs numériques  $x_1, x_2, \dots, x_n$  auxquels il est connecté par des synapses et calcule une sortie  $y$ . Chaque entrée  $x_i$  a un poids associé  $w_i$ , qui est attribué en fonction de son importance relative par rapport aux autres entrées. La valeur d'entrée  $x$  du neurone correspond à la somme pondérée de ses entrées en ajoutant une autre entrée ayant un poids

b appelé biais. Ensuite, le neurone applique une fonction f sur cette somme, comme illustré à la figure 2.6 [Dahmane, 2020]. La sortie du neurone y, appelée activation de sortie, est calculée selon la formule

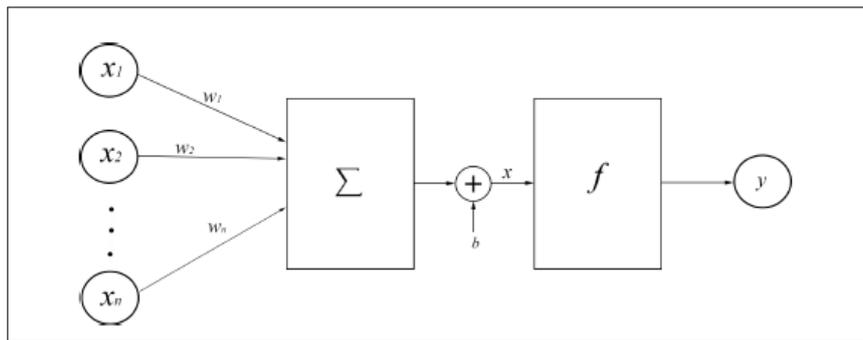


FIGURE 2.6 – Modélisation d’un neurone artificiel.

suivante :

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right) \tag{2.1}$$

Où f est la fonction d’activation non linéaire, elle a pour but d’introduire la non-linéarité dans la sortie d’un neurone. Ceci est important, car la plupart des données du monde réel sont non linéaires et donc les neurones apprendront ces représentations non-linéaires. Chaque fonction d’activation prend un nombre unique et effectue une certaine opération mathématique sur lui Il existe plusieurs fonctions d’activation utilisées dans la pratique à savoir :

- a) **La fonction Sigmoidé** : elle prend une entrée réelle et la réduit entre 0 et 1 via la formule suivante :

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2.2}$$

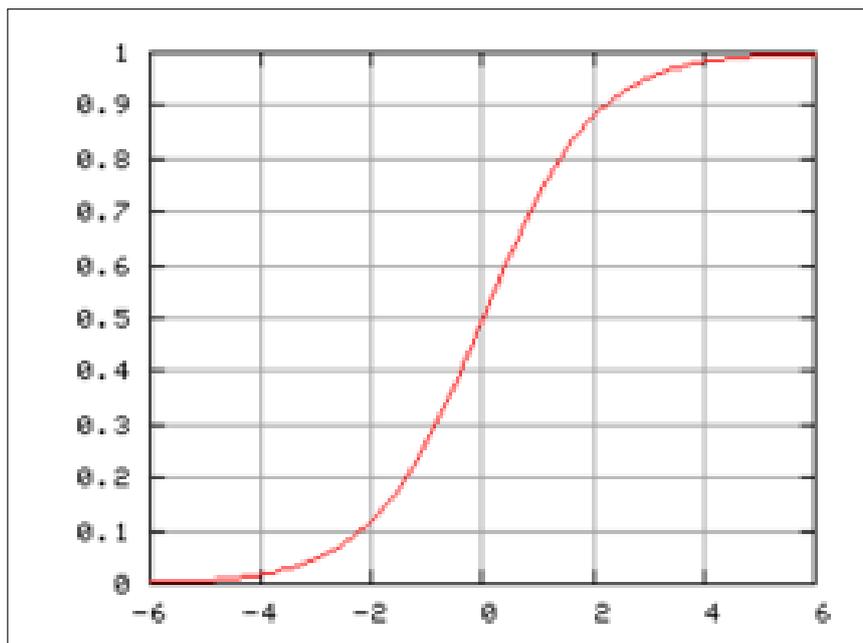


FIGURE 2.7 – Graphe de fonction Sigmoidé

- b) La fonction tangente hyperbolique : cette fonction est similaire à la fonction sigmoïde avec une différence de sortie, variante dans l'intervalle  $[1, 1]$  au lieu de  $[0, 1]$ . Elle est définie comme suit :

$$\tanh(x) = \frac{2}{1 + e^{1-2x}} \quad (2.3)$$

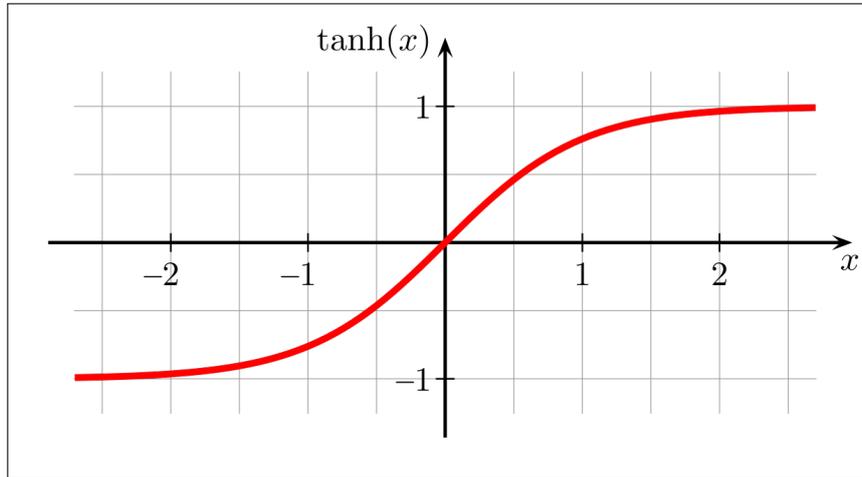


FIGURE 2.8 – Graphe de fonction tangente hyperbolique.

- c) La fonction Softmax : le prend un vecteur N-dimensionnel de valeurs réelles et le transforme en vecteur de valeurs réelles dans un intervalle  $[0,1]$  de telle sorte que la somme soit égale à 1. Étant donné une matrice d'entrée  $X_n$  la fonction est définie comme suit :

$$\sigma(X_n, k) = \frac{e^{X_n, k}}{\sum_{j=1}^k e^{X_n, j}} \quad (2.4)$$

### 2.3.1.2 Quelques types de réseaux de neurones

Il existe plusieurs catégories de réseau de neurones considérant différents critères à savoir la topologie des connexions entre les neurones, la fonction d'agrégation, la fonction d'activation ou encore l'algorithme d'apprentissage adoptés. Nous allons présenter dans ce qui suit quelques types de RNA.

#### a) Le réseau de neurones à propagation avant

Le réseau de neurones à propagation avant (feedforward neural network) était le premier et le plus simple réseau de neurones artificiel mis au point. Il contient plusieurs neurones disposés en couches, les neurones des couches adjacentes ont des connexions entre eux et toutes ces connexions ont des poids qui leur sont associés Un réseau de neurones à propagation avant peut-être constitué de trois types de neurones (voir Figure 2.9) [Dahmane, 2020] :

**neurones d'entrée** : les neurones d'entrée fournissent des informations du monde extérieur au réseau et construisent ensemble une « couche d'entrée ». Aucun calcul n'est effectué dans les nœuds d'entrée, ils transmettent simplement les informations aux nœuds cachés.

**neurones cachés** : les neurones cachés n'ont pas de connexion directe avec le monde extérieur (d'où le nom "caché"), ils effectuent des calculs et transfèrent les informations des neurones

d'entrée aux neurones de sortie, ensemble ils forment une "couche cachée".

**neurones de sortie** : les neurones de sortie construisent ensemble une "couche de sortie". Ils sont responsables des calculs et du transfert des informations du réseau vers le monde extérieur.

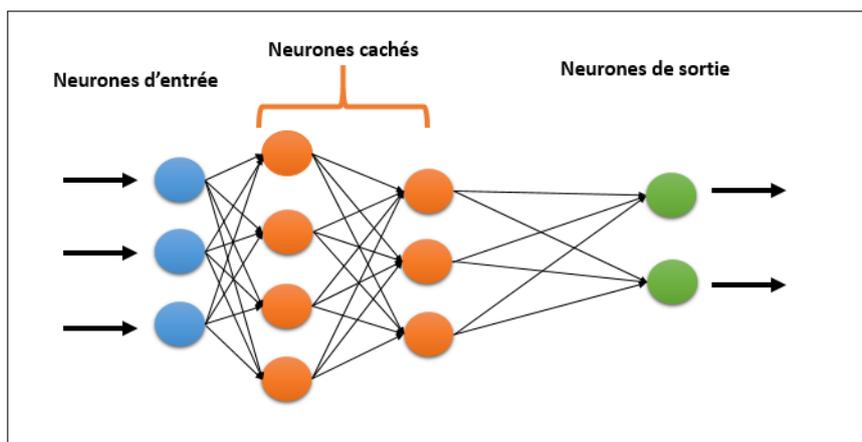


FIGURE 2.9 – Exemple d'un réseau de neurones artificiel à propagation avant .

- **Perceptron monocouche** : il s'agit d'un réseau de neurones à propagation avant le plus simple, il contient une seule couche cachée. Son algorithme repose sur le calcul du gradient de sortie puis sur la rétro propagation de celui-ci à travers la fonction de seuil puis des poids. C'est pourquoi la fonction de seuil doit être dérivable. La fonction d'activation de Heaviside est donc remplacée par des fonctions d'activation lui ressemblant et qui sont dérivables. Les principales fonctions d'activation sont : la fonction linéaire, la fonction sigmoïde et la fonction tangente hyperbolique [Mokri, 2017], son architecture générale est représentée dans la figure 2.10.

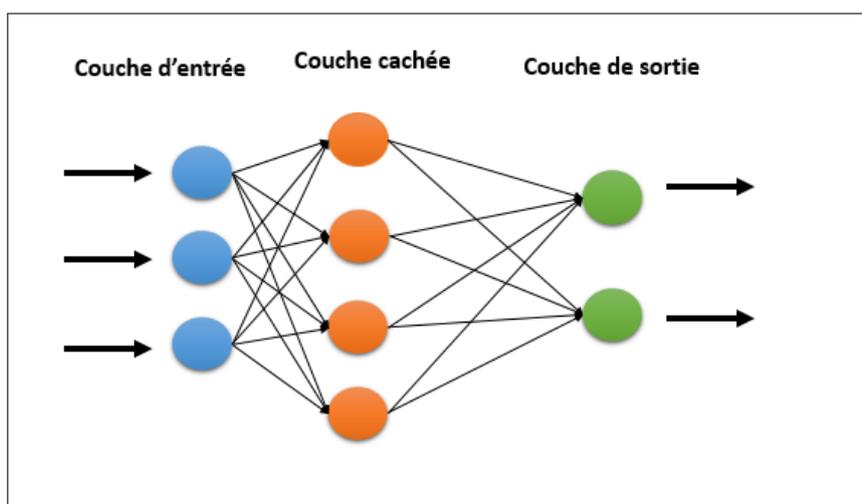


FIGURE 2.10 – Perceptron monocouche [Dahmane, 2020].

- **Perceptron multicouches** Ce sont les réseaux de neurones les plus connus, ils sont aussi du type propagation en avant, ses fonctions d'activation sont du type sigmoïde et son apprentissage est de type supervisé (par correction des erreurs). Le perceptron multicouche est un des réseaux de neurones les plus utilisés pour des problèmes d'approximation, de classification et de prédiction. Il est habituellement constitué de deux ou trois couches de neurones totalement connectés (voir figure 2.9) [Khodja, 2011].

- b) **Les Réseaux Neurone à Fonction de Base Radiales (RN- FBR)** Ce sont les réseaux de neurones les plus connus, ils sont aussi du type propagation en avant, ses fonctions d'activation sont du type sigmoïde et son apprentissage est de type supervisé (par correction des erreurs). Le perceptron multicouche est un des réseaux de neurones les plus utilisés pour des problèmes d'approximation, de classification et de prédiction. Il est habituellement constitué de deux ou trois couches de neurones totalement connectés (voir figure 2.9) [Khodja, 2011].

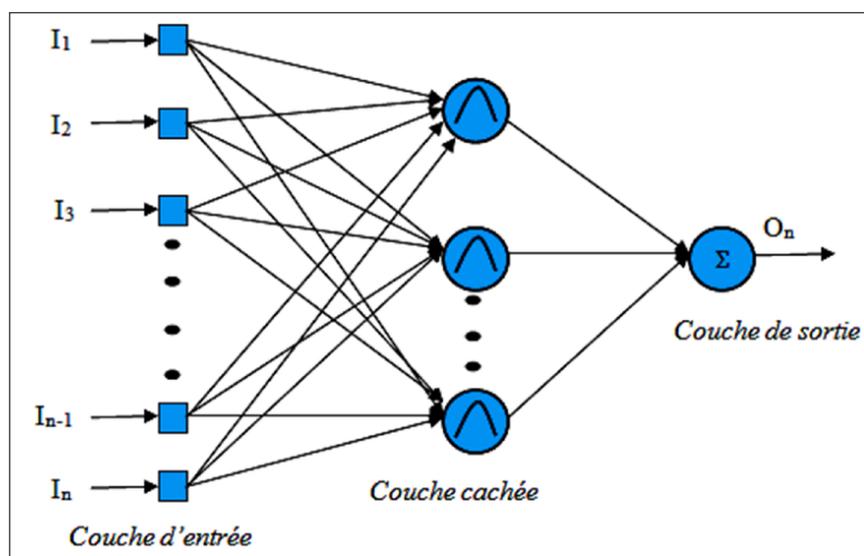


FIGURE 2.11 – Les Réseaux de Neurones à fonction de Base Radiale RN-FBR .

- c) **Les réseaux de neurones récurrents** Les RNRs (ou RNN pour Recurrent Neural network), ces réseaux sont dit récurrents car il comporte des cycles dans leurs graphes de connectivités, ce qui change la dynamique du réseau l'amenant à l'auto-entretenir [Senani, 2019]. De nombreux types de RNN ont été développés au cours des 30 dernières années tels que les réseaux d'Elman, les réseaux de Jordan et les Echo StateNetworks [Gregory,2017].

## 2.3.2 Apprentissage Profond : AP (Deep Learning :DL)

### 2.3.2.1 Historique

Depuis 2012, les algorithmes à base d'apprentissage profond semblent prêts à résoudre bien des problèmes : reconnaître des visages comme le propose DeepFace, synthétiser des textes, bientôt permettre la conduite de voitures autonomes ou encore la recherche de cellules cancéreuses.

Les fondements de ces méthodes ne sont pas si récents, l'apprentissage profond a été formalisé en 2007 à partir des nouvelles architectures de réseaux de neurones dont les précurseurs sont McCulloch et Pitts

[McCulloch & Pitts, 1943] qui ont été suivi par de nombreux développements comme les Réseaux de Neurones Convolutifs RNC ( CNN pour Convolutional neural network) de Yann Le Cun et Yoshua Bengio [LeCun et al., 1995], et les réseaux de neurones profonds qui en découlent en 2012 et ouvrent la voie à de nombreux champs d'applications comme la vision, le traitement du langage ou la reconnaissance de la parole [Larbi, 2018].

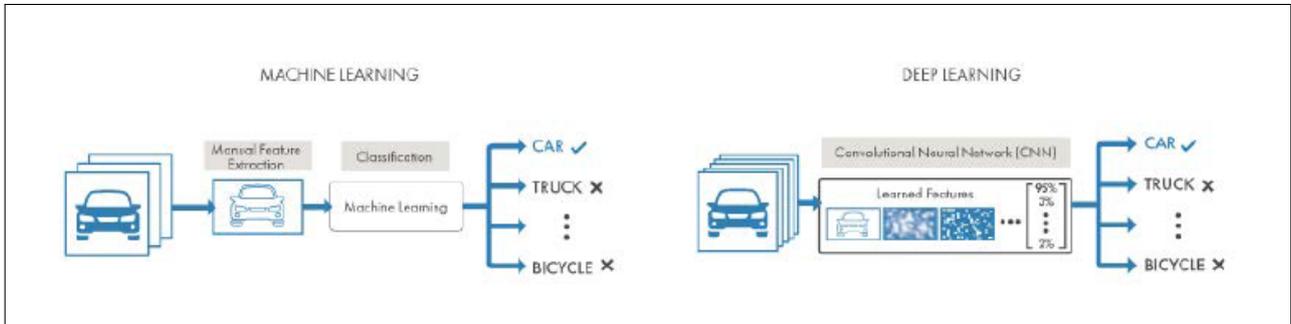


FIGURE 2.12 – Comparaison de méthodes de classification de véhicules de Machine Learning (gauche) et d'apprentissage profond(droite).

### 2.3.2.2 Définition

L'apprentissage profond est un ensemble de méthodes d'apprentissage automatique (Machine Learning) tentant de modéliser avec un haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non-linéaires. Dans un processus d'apprentissage profond, l'extraction de caractéristiques pertinentes de l'image est automatique contrairement à l'apprentissage automatique où cette extraction est manuelle. En outre, l'apprentissage profond effectue un apprentissage "de bout en bout" : à partir de données brutes, un réseau se voit assigner des tâches à accomplir (une classification, par exemple) et apprend comment les automatiser.

Un des avantages majeurs des réseaux d'apprentissage profond réside dans leur capacité à continuer à s'améliorer en même temps que le volume de vos données augmente [Larbi, 2018].

## 2.3.3 Réseaux de Neurones Convolutifs (RNC)

Les réseaux de neurones Convolutifs ou (CNN pour Convolutional Neural Network en anglais), sont des réseaux de neurones multicouches spécialisés dans la classification des images et la reconnaissance de formes. Ils comportent deux parties distinctes, première partie c'est la partie convolutionnelle, son architecture repose sur des couches de convolution alternant avec des couches d'agrégation (pooling). L'image en entrée est fournie sous forme de matrice de pixels, elle passe par une succession de filtres au niveau de chaque couche de convolution, créant de nouvelles images appelées cartes de convolutions. La couche d'agrégation (pooling) réduit la taille de l'image par une fonction de max pooling [Belhadjer & Sarouer, 2018].

### 2.3.3.1 Architecture générale des RNCs

Une architecture RNC est formée par un empilement de couches effectuant des traitements indépendants [Mokri, 2017] :

- La couche de convolution (CONV) qui traite les données d'un champ récepteur.
- La couche de pooling (POOL), qui permet de compresser l'information en réduisant la taille de l'image intermédiaire (souvent par sous-échantillonnage).
- La couche de correction (ReLU), souvent appelée par abus 'ReLU' en référence à la fonction d'activation (Unité de rectification linéaire, ou Rectification Linéaire Unit en anglais).
- La couche entièrement connectée (FC), qui est une couche de type perceptron.
- La couche de perte (LOSS).

### 1) Les couches convolutionnelles

Reconnaître des objets dans des images revient à comparer l'image pixel par pixel, pour trouver des caractéristiques. Bien qu'efficaces pour le traitement d'images, les réseaux de neurones multicouches, ont beaucoup de mal à gérer des images de grande taille, ce qui est dû à la croissance exponentielle du nombre de connexions avec la taille de l'image. Par exemple une image de taille 200 x 200 x 3 (200 de large, 200 de haut, 3 canaux de couleur), conduirait ainsi à traiter 120 000 entrées par neurone ce qui, multiplié par le nombre de neurones, devient énorme [Belhadjer & Sarouer, 2018].

Une couche de convolution permet de reconnaître des formes sur une image, généralement avec des filtres de taille 2 x 2 ou 3 x 3 pixels, ce que limite le nombre d'entrées. Pour chaque filtre, et pour tout carré de la taille du filtre sur la matrice d'entrée, les cases correspondantes de la matrice de sorties sont égales à la somme de toutes les valeurs du filtre, multipliées respectivement par ceux des carrés d'entrée. Ainsi, la taille de la matrice de sortie est égale à celle de l'entrée moins deux lignes et deux colonnes (voir figure 2.13).

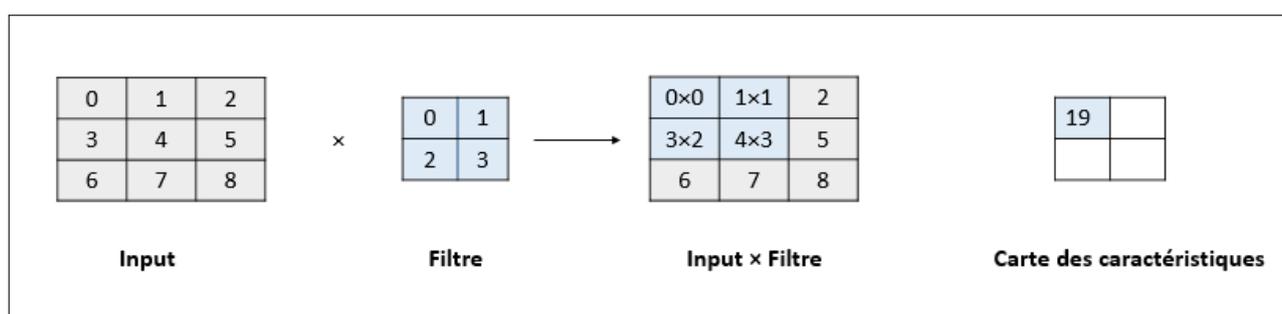


FIGURE 2.13 – Opération d'une convolution .

### 2) Les couches de pooling

Placées entre deux couches de convolution, ces couches abaissent les données de l'image extraites par les couches convolutionnelles pour réduire la dimension de la carte des caractéristiques afin de diminuer le temps de traitement. Une couche de pooling permet de réduire la taille de la matrice d'entrée, d'augmenter la vitesse tout en gardant les informations les plus importantes [Belhadjer & Sarouer, 2018]. Il existe deux types de pooling (voir figure 2.14).

- **Le Maximum Pooling** : il est question ici dans ce type de pooling de diviser la taille de l'entrée par 2 en prenant le maximum des carrés de taille 2 × 2 pixels.

- **L’Average Pooling** : consiste à diviser la taille de l’entrée par 2 en calculant la moyenne des carrés de taille 2 x 2 pixels. Le maximum pooling est le plus utilisé surtout pour la reconnaissance des objets de grandes tailles et complexes [Mokri, 2017].

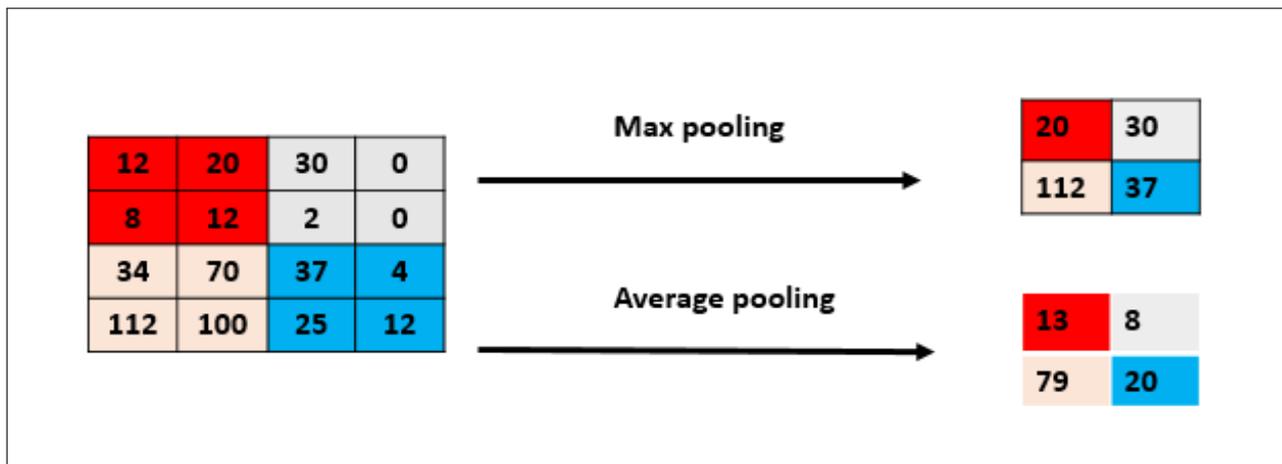


FIGURE 2.14 – Exemple de fonctionnement de Max pooling et Average pooling .  
[Asif et Amjad, 2020]

- 3) **Unités linéaires rectifiées** Il est possible d’améliorer l’efficacité du traitement en intercalant entre les couches de traitement une couche qui va opérer une fonction mathématique (fonction d’activation) sur les signaux de sortie. La fonction ReLU (pour Rectified Linear Unit) force les neurones à retourner des valeurs positives (voir figure 2.15) [Belhadjer & Sarouer, 2018]. Elle est définie comme suit :

$$Y(x) = \max(0, x) \quad (2.5)$$

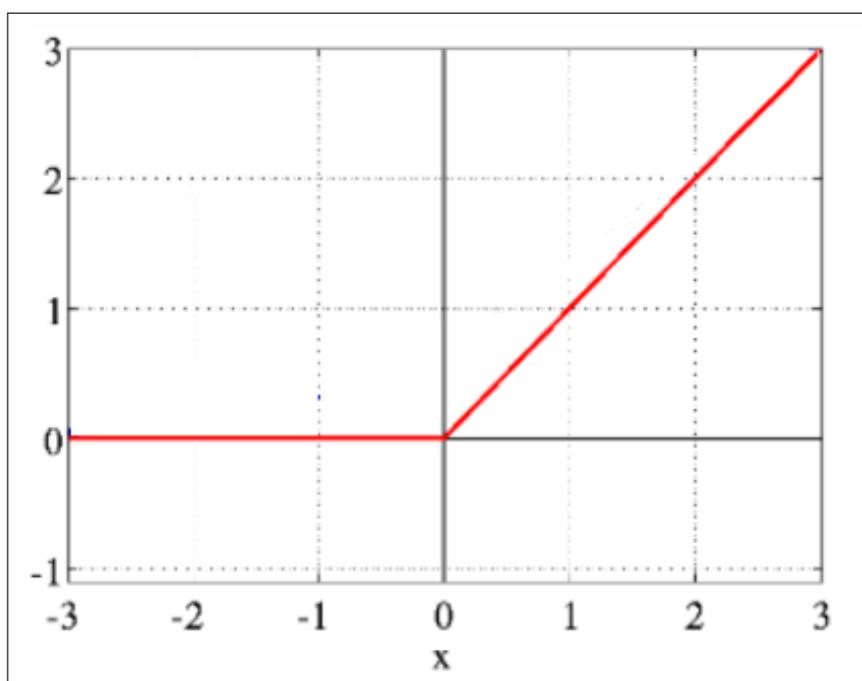


FIGURE 2.15 – Graphe de fonction ReLu.

- 4) **Les couches entièrement connectées**

La couche entièrement connectée est un perceptron multicouche qui utilise une fonction d'activation Softmax dans la couche de sortie. Le terme «entièrement connecté» implique que chaque neurone de la couche précédente est connecté à chaque neurone de la couche suivante.

Dans le fonctionnement entièrement connecté d'un réseau de neurones, la représentation d'entrée est aplatie dans un vecteur de caractéristiques et transmise à travers un réseau de neurones afin de prédire les probabilités de sortie. La figure 2.16 présentée ci-dessous décrit l'opération d'aplatissement [Senani, 2019].

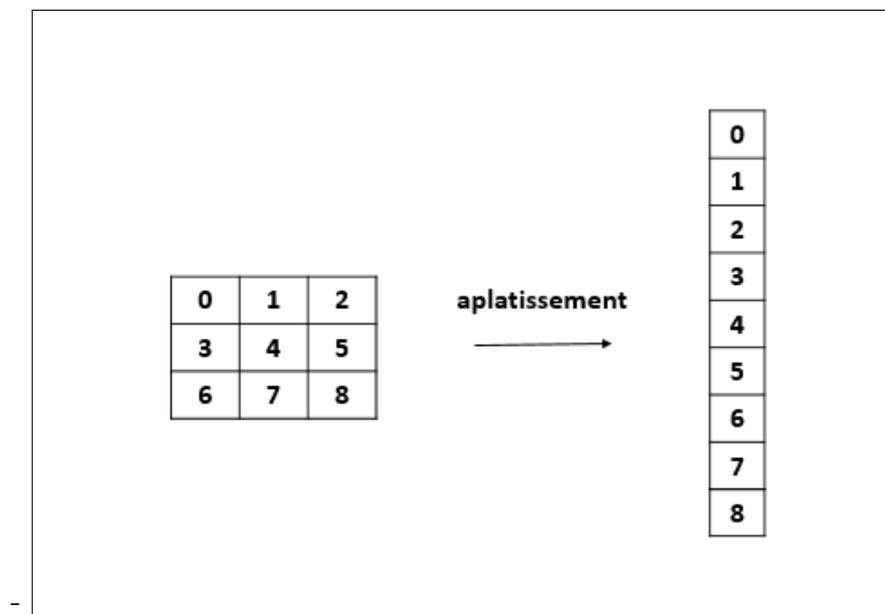


FIGURE 2.16 – opération d'aplatissement.

Les lignes sont concaténées pour former un vecteur long. Si plusieurs couches d'entrée sont présentes, ses lignes sont également concaténées pour former un vecteur de caractéristiques encore plus long. Le vecteur de caractéristique est ensuite passé à travers plusieurs couches denses (entièrement connectées). À chaque couche dense, le vecteur de caractéristiques subit les mêmes opérations que dans un réseau de neurones classique. La couche entièrement connectée aide à mapper la représentation entre l'entrée et la sortie [Senani, 2019].

- 5) **Couche de perte (LOSS)** La couche de perte spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et le signal réel. Elle est normalement la dernière couche dans le réseau [Mokri, 2017].

Diverses fonctions de perte adaptées à différentes tâches peuvent être utilisées parmi eux nous citons :

- a) **L'entropie croisée catégorique (categorical cross entropy)** : définit la perte lorsque les résultats de la catégorie sont non-binaires, c'est-à-dire deux résultats possibles : Oui/Non/Peut-être) ou (Type 1/ Type 2/... Type n) [Bellahmer, 2020].
- b) **La fonction Dixeloss** : elle est utilisée comme métrique de performance en segmentation, comme fonction de coût à optimiser. La normalisation permet d'obtenir de meilleures performances en cas de fort déséquilibre entre les classes. En segmentation, cela correspond aux cas où les objets à segmenter sont petits par rapport à l'image [Couteaux, 2021].

- c) **La fonction Focal loss** : cette fonction de coût a pour but, par rapport à une entropie croisée binaire classiquement utilisée pour les tâches de classification, de défavoriser les ancres faciles, c'est-à-dire les ancres que le réseau classe correctement avec une forte confiance.

L'hypothèse est que le réseau arrivera facilement à classer correctement les exemples négatifs avec une forte confiance, et qu'une simple entropie croisée ne pénaliserait pas assez les faux négatifs par rapport aux vrais négatifs [Couteaux, 2021].

- **Optimiseurs**

La fonction d'optimisation est un algorithme mathématique qui utilise les dérivées, les dérivées partielles et la règle de la chaîne dans le calcul pour comprendre l'ampleur des changements que le réseau observera dans la fonction de perte en apportant une petite modification au poids des neurones. La modification de la fonction de perte, qui serait une augmentation ou une diminution, aide à déterminer la direction du changement nécessaire dans le poids de la connexion.[Bellahmer, 2020]..

- **\* Estimation adaptative des moments** : ou Adam pour "Adaptive Moment Estimation" est une technique d'optimisation qui calcule un taux d'apprentissage adaptatif pour chaque paramètre.

Elle définit la dynamique et la variance du gradient de la perte et exploite un effet combiné pour mettre à jour les paramètres de poids [Bellahmer, 2020].

### 2.3.3.2 Quelques modèles de réseaux de neurones convolutifs

Les performances des réseaux de neurones profonds dans la détection d'objets reposent sur une grande quantité de données d'apprentissage ainsi que sur la puissance de calcul. La collecte et l'annotation d'un ensemble de données (Dataset) de taille suffisante sont un processus mouvementé et qui prend du temps. Par conséquent, la plupart des recherches sur la détection d'objets sont effectuées sur des ensembles de données accessibles au public. Il existe plusieurs ensembles de données de ce type. Nous présenterons dans ce qui suit quelques architectures RNC les plus courantes et récentes.

#### 1) Le modèle VGG16

VGG16 est un réseau à 16 couches utilisé par le groupe de géométrie visuelle (VGG) de l'Université d'Oxford dans le cadre du concours ILSVRC (ImageNet) en 2014. Le modèle atteint un taux de succès de 92,5% dans le top 5 sur l'ensemble de validation. Il prend en entrée une image en couleurs de taille  $224 \times 224$  px et la classe dans une des 1000 classes. Il renvoie donc un vecteur de taille 1000, qui contient les probabilités d'appartenance à chacune des classes. L'architecture de VGG-16 est illustrée par la figure 2.17. Le VGG16 est utilisable dans le domaine de l'imagerie médicale. [Bouabdallaoui, 2019]

#### 2) Le réseau de neurones résiduel (ResNet)

Le réseau de neurones résiduel connu par le nom de ResNet (pour Residual Network- plus) apparaît suite au problème de rétropropagation du gradient et l'augmentation de l'erreur d'apprentissage. En effet, lorsque le réseau de neurones est trop profond, le gradient se réduit à zéro

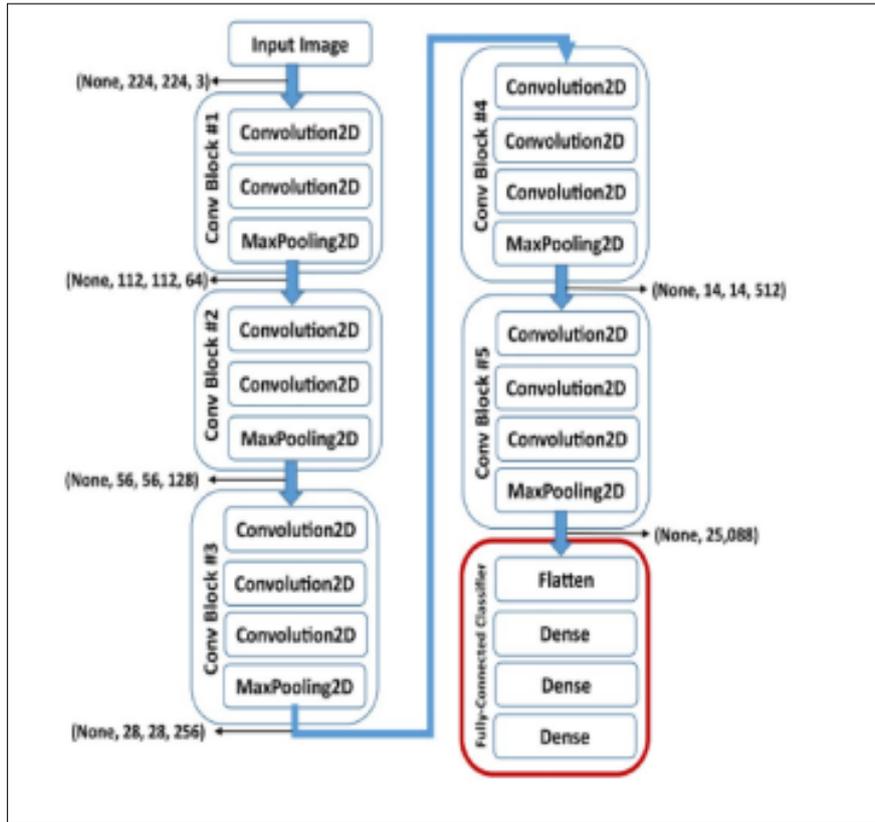


FIGURE 2.17 – Architecture de VGG-16 [Theckedath et Sedamkar, 2020]

et devient très faible pour mettre à jour les poids des différentes couches d'où la difficulté de l'apprentissage. Avec ResNet, l'optimisation des réseaux profonds est assurée en utilisant des connexions résiduelles. Cela permet aux gradients de passer dans deux couches de convolution, mais aussi de passer directement à travers un saut aux couches suivantes en reliant l'entrée de la  $n$ ème couche à la sortie de la  $(n + a)$ ème couche [Dahmane, 2020].

Al-haija et Adebajo [Al-haija et Adebajo, 2020] proposent un modèle qui utilise des techniques d'apprentissage par transfert à partir d'un modèle ResNet-50 RNC pré-entraînés avec la base ImageNet afin de traiter et classer l'ensemble de données BreakHis. Les résultats de simulation montrent que leur modèle proposé a atteint une excellente précision de classification de 99 %, surpassant les autres modèles de comparaison entraînés sur le même ensemble de données.

Il existe plusieurs versions de ResNet, telles que ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet152. La différence entre ces différentes architectures se situe au niveau du nombre des couches (18 couches, 34 couches, etc)[Dahmane, 2020].

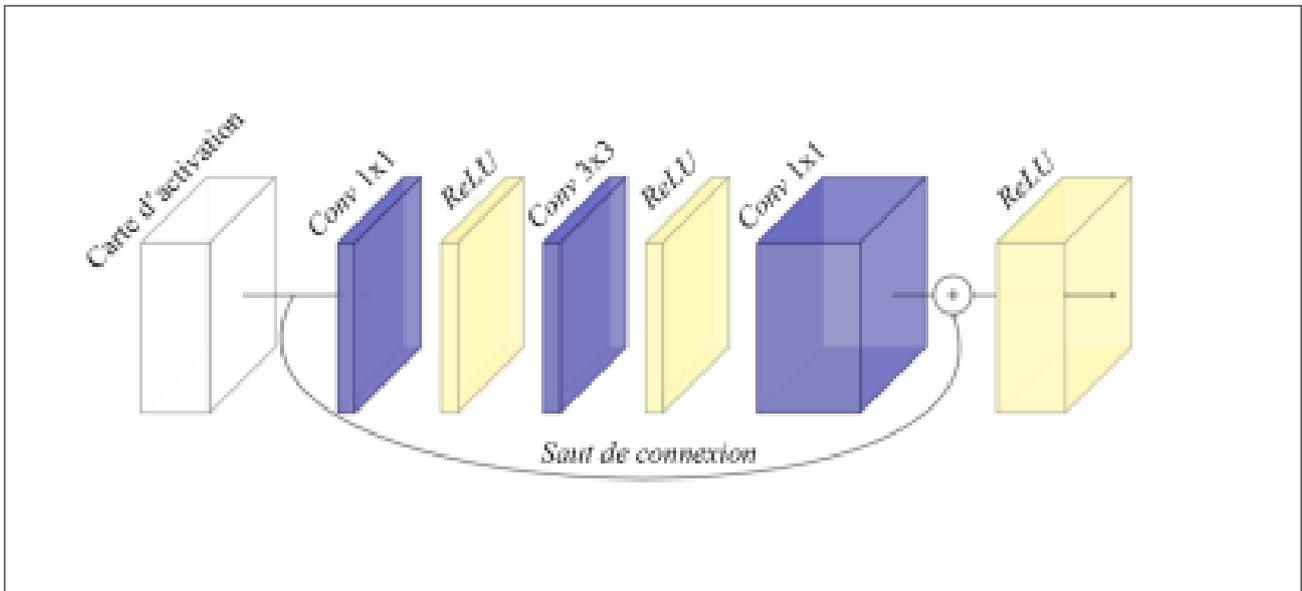


FIGURE 2.18 – Bloc convolutif résiduel [Audebert, 2018]

### 3) Le modèle Unet (Architecture encodeur- décodeur)

L'architecture encodeur-décodeur est constituée de deux parties. L'encodeur est une représentation pyramidale classique semblable à celle des RNCs. La différence réside dans la deuxième partie où à défaut d'avoir une seule opération de sur-échantillonnage pour retrouver seulement la dimension spatiale, le décodeur est symétrique à l'encodeur mais remplace le sous-échantillonnage par le sur-échantillonnage.

Dans cette partie, le but avant d'arriver à retrouver la dimension spatiale est de récupérer les détails de l'image pour atteindre une résolution complète en associant les caractéristiques globales et locales via des sauts de connexions longues en provenance de l'encodeur. Comme suite logique au succès que l'Unet a suscité, diverses approches de segmentation d'images médicales y ont eu recours et certaines ont apporté des modifications à son architecture pour contourner ses limites [Guerroumi, 2019].

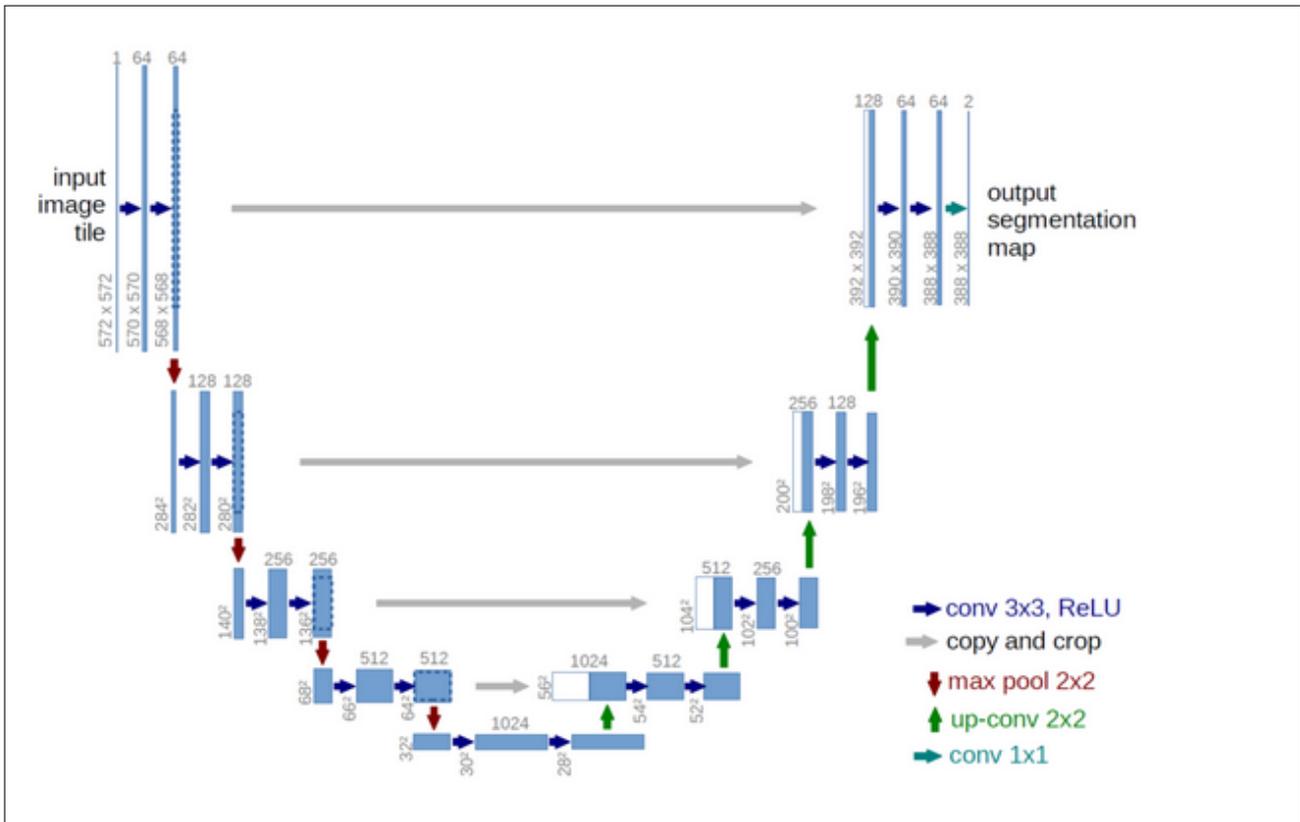


FIGURE 2.19 – Architecture du réseau U-Net, de type encodeur-décodeur.

Le réseau U-Net (figure 2.19) a été développé pour résoudre des problématiques de segmentation biomédicale et connaît aujourd’hui un succès fort avec des applications dans de nombreux domaines en biologie, santé et dans l’industrie. Il se distingue par l’utilisation de connexions résiduelles à chaque résolution, qui retransmettent les cartes de descripteurs du bloc encodeur vers le bloc décodeur correspondant, pour améliorer la localisation de descripteurs de haut niveau [Ganaye, 2019].

#### 4) Le modèle EfficientNet

EfficientNet [Tan & Le, 2019] est une architecture et une méthode de mise à l’échelle qui met à l’échelle uniformément toutes les dimensions de profondeur/ largeur/résolution à l’aide d’un coefficient composé. Contrairement à la pratique conventionnelle qui met ces facteurs à l’échelle de manière arbitraire, la méthode de mise à l’échelle EfficientNet met à l’échelle uniformément la largeur, la profondeur et la résolution du réseau avec un ensemble de coefficients de mise à l’échelle fixes.

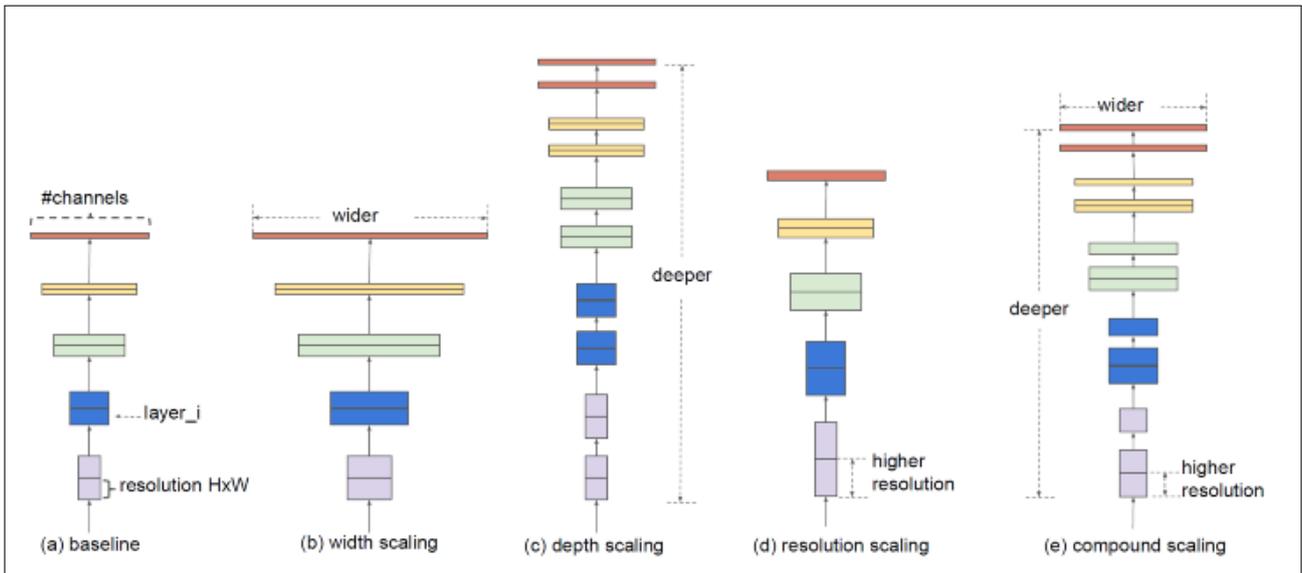


FIGURE 2.20 – Mise à l'échelle composée [Tan & Le, 2019] . (a) . exemple de réseau de base.(b) - Largeur ( $w$  ),(c) Profondeur(d), (d)Résolution ( $r$ ), (e) Mise à l'échelle composée

Notant toutes fois que la :

- Largeur ( $w$ ) : la mise à l'échelle de la largeur du réseau est couramment utilisée pour les modèles de petite taille. Les réseaux plus larges ont tendance à être capables de capturer des fonctionnalités plus fines et sont plus faciles à former. Cependant, les réseaux extrêmement larges mais peu profonds ont tendance à avoir des difficultés à capturer des caractéristiques de niveau supérieur.
- Profondeur ( $d$ ) : la mise à l'échelle de la profondeur du réseau est le moyen le plus couramment utilisé par de nombreux ConvNets. Cependant, mettre à l'échelle un modèle de base avec un coefficient de profondeur  $d$  différent, ce qui suggère en outre le retour de précision décroissant pour les ConvNets très profonds.
- Résolution( $r$ ) : avec des images d'entrée de résolution plus élevée, les ConvNets peuvent potentiellement capturer des motifs plus fins .Des résolutions plus élevées améliorent la précision, mais le gain de précision diminue pour les résolutions très élevées.

Intuitivement, pour des images de résolution plus élevée, l'augmentation de la profondeur du réseau permet d'obtenir des champs de réception plus grands qui peuvent aider à capturer des fonctionnalités similaires qui incluent plus de pixels dans des images plus grandes. En conséquence, il faudrait également augmenter la largeur du réseau lorsque la résolution est plus élevée afin de capturer des motifs plus fins.

Il est alors judicieux de penser à une mise à l'échelle composée, où il sera question de coordonner et équilibrer différentes dimensions de mise à l'échelle plutôt que la mise à l'échelle conventionnelle à une dimension.

## 2.4 D'autres travaux de segmentation des images médicales via l'apprentissage profond

L'apprentissage profond a permis des progrès importants et rapides dans divers domaines à savoir l'analyse du signal sonore ou visuel, la reconnaissance vocale, le traitement automatisé du langage, ainsi que la vision par ordinateur et plus particulièrement la segmentation des images médicales.

Khairandish et son équipe [Khairandish et al., 2021] proposent un nouveau système hybride (RNC-MVS) basé sur les Réseaux de Neurons Convolutifs et les machines à vecteurs de support dédié à la détection et la classification des images en IRM. Les RNCs ont été adoptés par les auteurs en raison de leur puissante capacité d'extraction de caractéristiques et de leur propriété de classification robuste.

Dolz et al. [Dolz et al., 2020] proposent une nouvelle méthode de segmentation du cerveau de nourrissons iso intenses dans des images IRM multimodales basée de sur un ensemble de RNC (deep CNN). Les auteurs dans ce réseau entièrement convolutif prennent en compte le contexte spatial 3D des données volumétriques et modélise les informations locales et globales dans la segmentation. Dans leur modèle, les caractéristiques de chaque couche convolutive sont agrégées en entrée de la première couche entièrement connectée. De plus, deux stratégies différentes ont été étudiées pour combiner des modalités d'images à entrées multiples, en utilisant une fusion précoce ou tardive de ces modalités. La performance de la méthode proposée a été évaluée dans le Grand challenge MICCAI iSEG-2017 sur la segmentation IRM cérébrale du nourrisson de 6 mois. Leurs résultats obtenus démontrent également l'avantage de combiner la prédiction de plusieurs RNC, avec des améliorations de performances par rapport à l'utilisation d'un seul RNC.

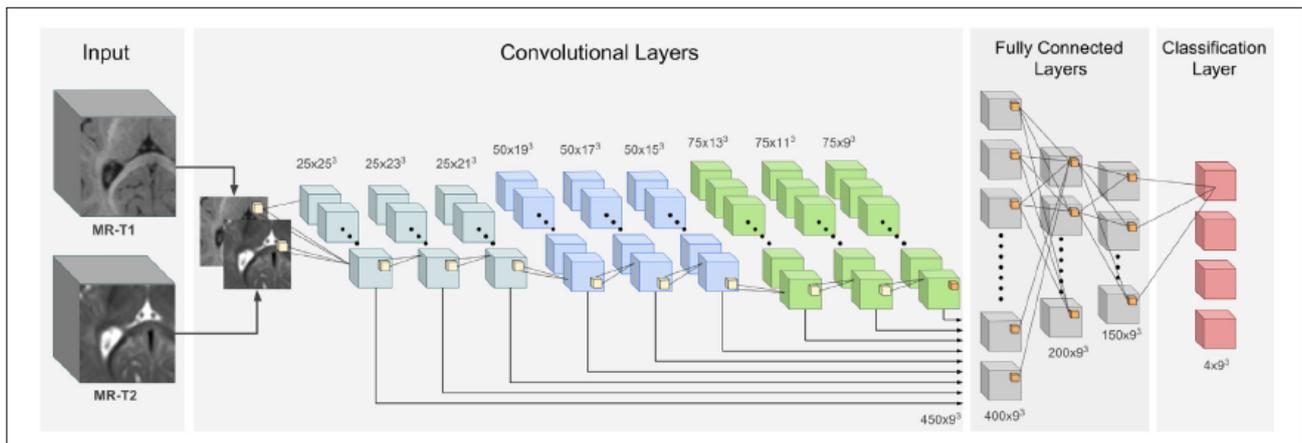


FIGURE 2.21 – L'architecture proposée dans [Dolz et al., 2020] en utilisant une fusion précoce.

Salama et Aly [Salama & Aly, 2021] ils proposent une nouvelle méthode de segmentation sur la base de différents modèles d'apprentissage en profondeur, y compris InceptionV3, DenseNet121, ResNet50, VGG16 et MobileNetV2 en vue d'un diagnostic du cancer du sein, avec plus de précision et un moindre temps sur des images de mammographies numérisées. Les auteurs utilisent une version modifiée du modèle U-Net pour le processus de segmentation. La performance du diagnostic est évaluée en termes d'intersection sur l'union (Intersection-Over-Union : IoU), exactitude, sensibilité, précision, aire sous la courbe (Area Under the Curve : AUC), score F1, ainsi que le temps d'exécution. Salama et Aly adoptent dans leur travail transfert d'apprentissage, ils utilisent notamment une augmentation des données (pour plus de variation), ce qui améliore la généralisation et atténue le sur-apprentissage.

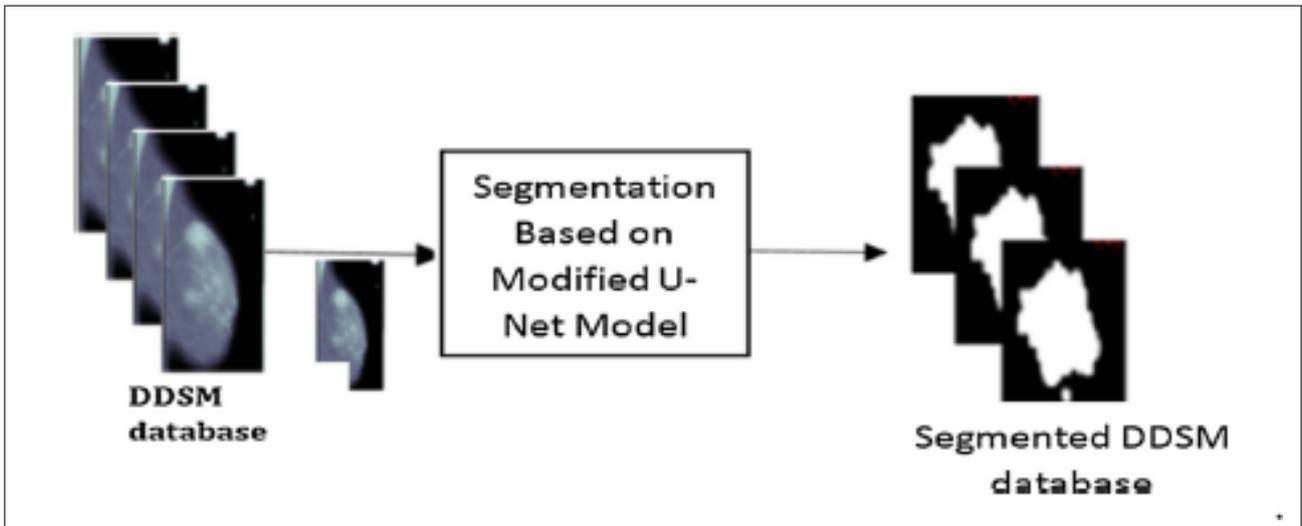


FIGURE 2.22 – La structure de la segmentation proposée dans [Salama & Aly, 2021]

## 2.5 Conclusion

Dans ce chapitre, nous avons présenté les différentes approches et techniques de segmentation d'images qui représente notre domaine d'application. Nous avons également donné quelques concepts de base quant à l'apprentissage automatique et notamment l'apprentissage profond qui ont révolutionné ces dernières années le domaine de la reconnaissance des formes en général, et celui de la segmentation des images médicales en particulier, où nous avons illustré ceci en citant quelques travaux de la littérature.

Le chapitre suivant sera consacré au Système Multi-Agents (SMA), un paradigme issu de l'Intelligence Artificielle Distribuée (IAD), qui était à son tour utilisé dans le domaine de segmentation autant qu'architecture de contrôle.

## Chapitre 3

# Les Systèmes Multi-Agents et la segmentation des images médicales

### 3.1 Introduction

Les systèmes Multi-Agents (SMA) sont actuellement un champ de recherche très actif, avec une base d'applications en constante expansion. Cette discipline relie plusieurs domaines, notamment l'intelligence artificielle et la distribution. C'est une discipline qui s'intéresse aux comportements collectifs résultant de l'interaction de plusieurs entités autonomes et flexibles appelées agents.

La segmentation à son tour a fait appel aux SMAs qui permettent une certaine distribution d'expertise offrant une architecture de contrôle. Dans ce chapitre, nous présenterons les concepts de base utilisés dans le domaine des SMAs, et nous verrons également quelques travaux sur la segmentation d'images via ce paradigme.

### 3.2 Notions de base

#### 3.2.1 L'agent

La description d'un système Multi-Agents passe nécessairement par celle d'un agent. De manière générale, un agent représente "une entité qui agit". Agir signifie être capable d'entrer en action et de produire un effet. Autrement dit c'est une entité, complexe ou très simple, capable de modifier son environnement [Soltani, 2007]. Un agent est une entité physique ou virtuelle autonome capable d'interagir avec son environnement dont il possède une représentation partielle. Il peut également interagir avec les autres agents partageant cet environnement. Un agent possède des compétences et connaissances propres. Il peut les combiner pour offrir des services [Inguere, 2018].

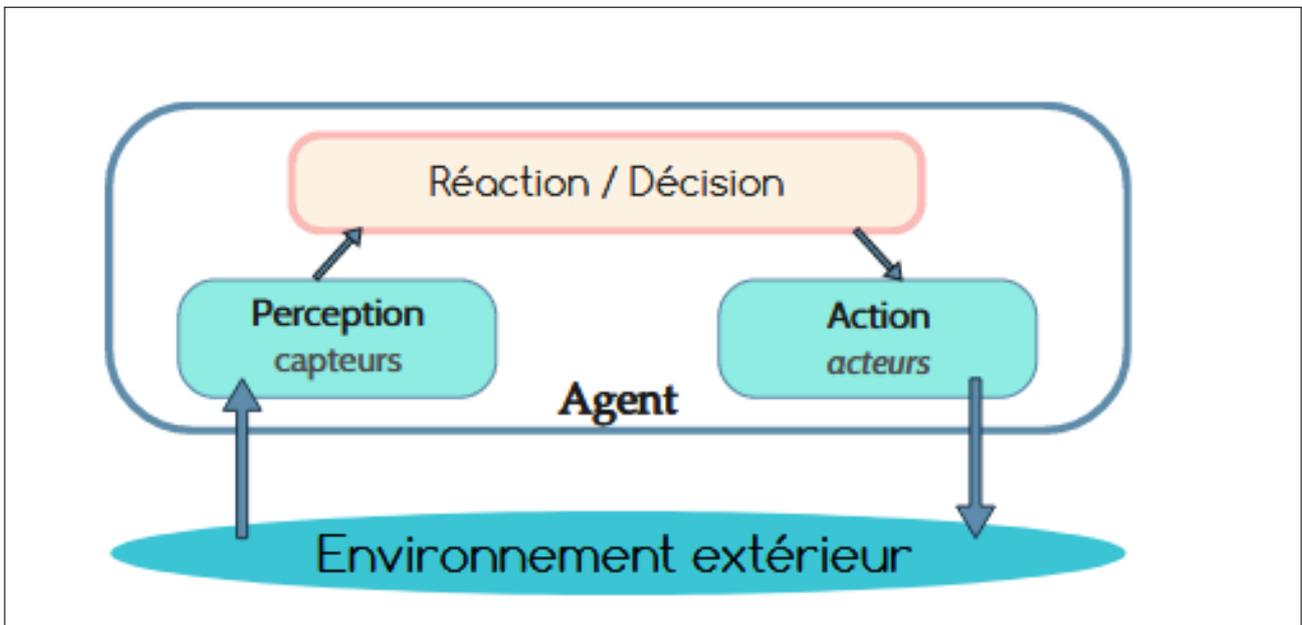


FIGURE 3.1 – Architecture générale d'un agent

### 3.2.2 Types d'agents

Il existe 3 types d'agents dans les SMAs [Chebout, 2014] :

- a) **Les agents réactifs** : sont des agents ayant la capacité de répondre uniquement à la loi stimulus/réponse. Ils sont extrêmement simples et disposent d'un processus de raisonnement procédural sans planification, d'un protocole de communication indirecte.
- b) **Les agents cognitifs** : quant à eux, sont des agents à base de connaissance se caractérisant par leur capacité de raisonnement et par la complexité de leur structure. Chaque agent dispose d'une base de connaissances liées à son domaine d'expertise et à son environnement. Ils sont généralement intentionnels et ils sont parfois amenés à négocier pour résoudre leurs conflits.
- c) **Les agents hybrides** : les architectures précédentes présentent certaines faiblesses. En effet, les architectures purement réactives ont un comportement assez simpliste alors que les architectures délibératives utilisent des mécanismes de raisonnement qui ne sont pas faciles à manipuler et qui ne sont pas suffisamment réactifs. Afin d'apporter une réponse à ces imperfections, des architectures hybrides en couches ont été proposées. L'idée principale est de structurer les fonctionnalités d'un agent en deux ou plusieurs couches hiérarchiques qui interagissent entre elles afin d'atteindre un état cohérent de l'agent.

### 3.2.3 L'environnement

L'environnement est l'endroit d'immersion dans lequel évoluent les agents, c'est une première classe d'abstraction fournissant les conditions environnementales aux agents pour exister et sert d'intermédiaire pour les interactions entre les agents ainsi que pour l'accès aux ressources [Moussa, 2011]. Un environnement multi-agents est caractérisé par plusieurs attributs comme [Marir & Silem, 2017] :

- a) **L'accessibilité** : dans un environnement accessible les agents peuvent accéder à son état intégral. Par contre, la portée des actions de l'agent et de sa perception est locale dans un environnement inaccessible.
- b) **Le déterminisme** : l'état d'un environnement déterministe est lié seulement à son état précédent et à l'action. Dans un environnement indéterministe, les résultats de la même action et dans le même contexte peuvent être différents.
- c) **Le dynamisme** : un changement de l'état dans un environnement statique se produit exclusivement par les actions des agents. Cependant, l'état d'un environnement dynamique peut être modifié sans l'intervention des agents.
- d) **La continuité** : si le nombre de perceptions et d'actions possible de l'agent dans un environnement est illimité, on dit que l'environnement est continu. Sinon, on le considère discret.

### 3.2.4 Interaction entre les agents

L'interaction entre agents dans un SMA peut être définie comme la mise en relation dynamique de deux ou plusieurs agents par le biais d'un ensemble d'actions. Ces interactions prennent deux formes : directes et indirectes [Moussa, 2011].

- **Les interactions directes** : sont en général des caractéristiques des agents cognitifs qui, travaillant pour un but précis, sont capables de communiquer intentionnellement. Ce type d'interaction se rapproche d'un acte de langage.
- **Les interactions indirectes** : sont les caractéristiques des agents réactifs qui répondent à des stimuli de l'environnement par le dépôt d'informations relatives à des critères environnementaux. Dans ce cas-là, les interactions entre les agents sont modélisées comme suit :
  - coopération : travailler ensemble dans un but commun, ceci est fait au travers de l'environnement sans aucune communication directe entre les agents ;
  - coordination : organiser la solution d'un problème de telle sorte que les interactions nuisibles soient évitées ou que les interactions bénéfiques soient exploitées.

## 3.3 Les systèmes Multi-Agents

Un SMA et un ensemble d'agents évoluant dans un environnement commun qui interagissent entre eux selon une certaine organisation. C'est un ensemble d'entités qui coordonnent leurs connaissances, buts, expériences et plans pour agir ou résoudre des problèmes, incluant le problème de la coordination inter-agent lui-même. Les agents composent un système, car ils constituent un ensemble cohérent autour d'un objet commun. L'objet qui les relie peut prendre diverses formes, il peut être un objectif commun (par exemple, résoudre un problème), un langage commun, ou tout simplement un environnement commun. L'ensemble des agents forme un groupe selon un réseau d'acointances leur permettant ainsi d'avoir des interactions. Selon les objectifs qu'ils poursuivent et leurs capacités à les atteindre, l'interaction entre les membres du groupe peut être coopérative, conflictuelle, une coordination ou une simple communication directe qu'elle soit ou indirecte dépendant à son tour aux types d'agents utilisés [Soltani, 2007].

### 3.3.1 Caractéristiques des systèmes Multi-Agents

Les SMAs possèdent plusieurs caractéristiques, à savoir [Semchedine, 2007] :

- a) Chaque agent a des informations ou des capacités de résolution de problèmes limités, ainsi chaque agent a un point de vue partiel, offrant ainsi une distribution d'expertise.
- b) Le contrôle dans ces systèmes repose sur des protocoles de communication et sur les capacités de raisonnement de chaque agent.
- c) La résolution d'un problème complexe est guidée par la recherche d'un compromis entre plusieurs intelligences ayant chacune des points de vue différents.

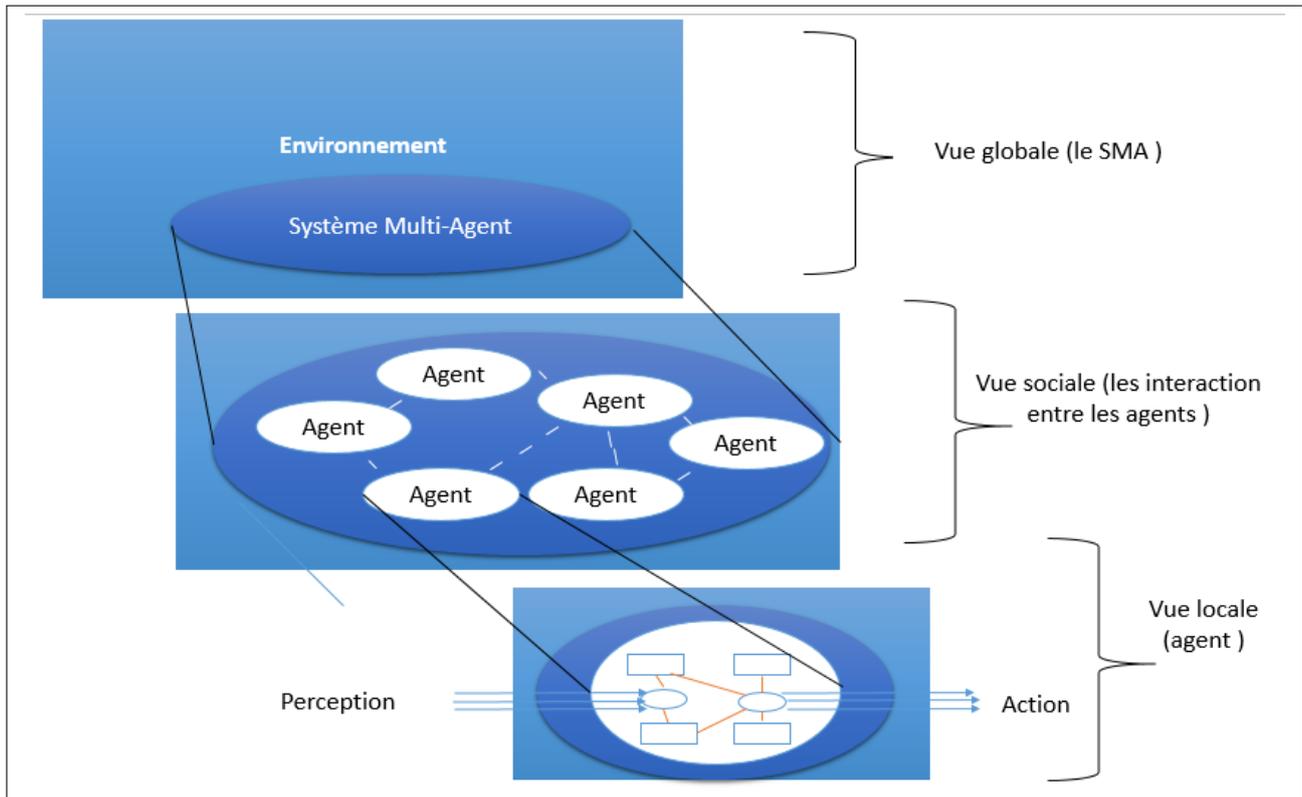


FIGURE 3.2 – Système Multi-Agents vu selon différents niveaux de détail

- d) La stratégie de résolution des problèmes se construit de façon dynamique, de la coopération des agents et de la résolution des conflits qui peuvent se produire.

### 3.3.2 Les approches SMAs destinées à la segmentation

Le principal objectif de l'utilisation des SMAs dans la segmentation et de profiter pleinement des caractéristiques de ces derniers, qui permettent une distribution d'expertises, un parallélisme qui contribue à son tour à la réduction du temps d'exécution, mais aussi une architecture de contrôle. Les agents peuvent toutefois être situés et déployés sur l'image.

Plusieurs travaux ont été proposés en vue d'une segmentation d'images; Porquet et al [Porquet et al., 2003] proposent une plate-forme permettant d'étudier les différentes manières de mettre en œuvre une coopération contour-région. Les auteurs ont choisi une approche multi-agent où les régions et les contours sont des agents situés qui coopèrent et négocient pour optimiser la segmentation de l'image. Les stratégies de coopération / négociation sont implémentées sous forme d'automates appliquant divers critères de segmentation.

Quant à Semchedine [Semchedine, 2007], il propose une approche qui consiste à faire la coopération entre l'algorithme des C-moyens flous (FCM pour Fuzzy C-Means) dont la contrainte d'appartenance d'un individu à une classe est gérée d'une manière relative et l'algorithme C-moyennes possibilistes (PCM Possibilistic C-Means) qui relâche cette contrainte reflétant ainsi de manière exacte la réalité de la distribution des individus. Afin de diminuer l'influence des données aberrantes et ambiguës, il a introduit deux concepts de rejet : le rejet d'ambiguïté qui concerne les individus situés entre les frontières des différentes classes et le rejet d'aberrance concernant les individus qui sont très éloignés par rapport à toutes les classes. Pour contrôler et/ou coordonner l'exécution des différentes phases

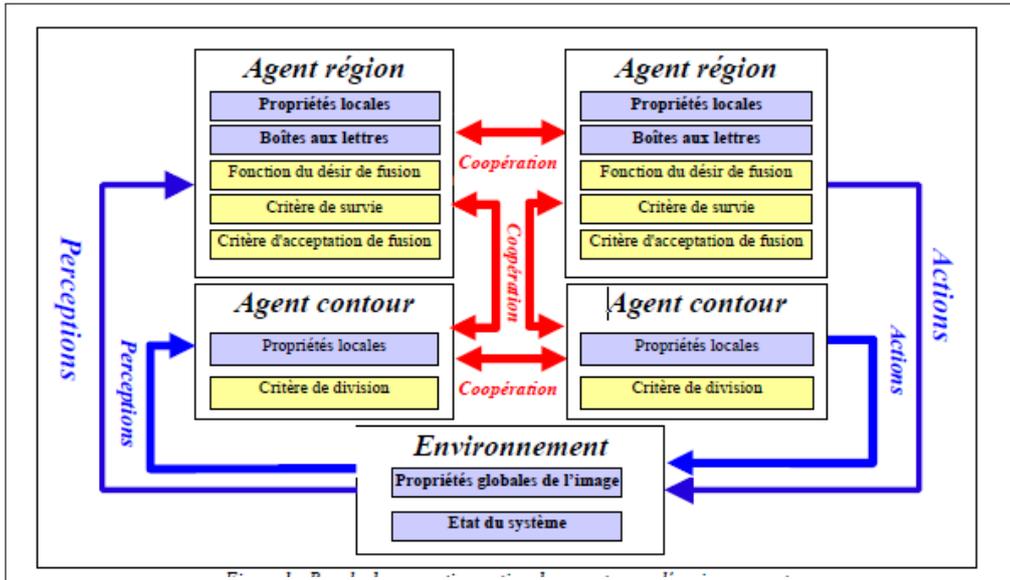


FIGURE 3.3 – Boucle de perception-action des agents avec l’environnement et interactions entre des agents régions et contours qui coopèrent. [Porquet et al., 2003]

du processus de segmentation, les auteurs ont fait recours aux SMAs qui possèdent la faculté d’apprentissage de nouvelles compétences pour gérer l’incertitude liée aux informations manipulées (images IRM). L’architecture du SMA proposée est présentée dans la figure 3.4.

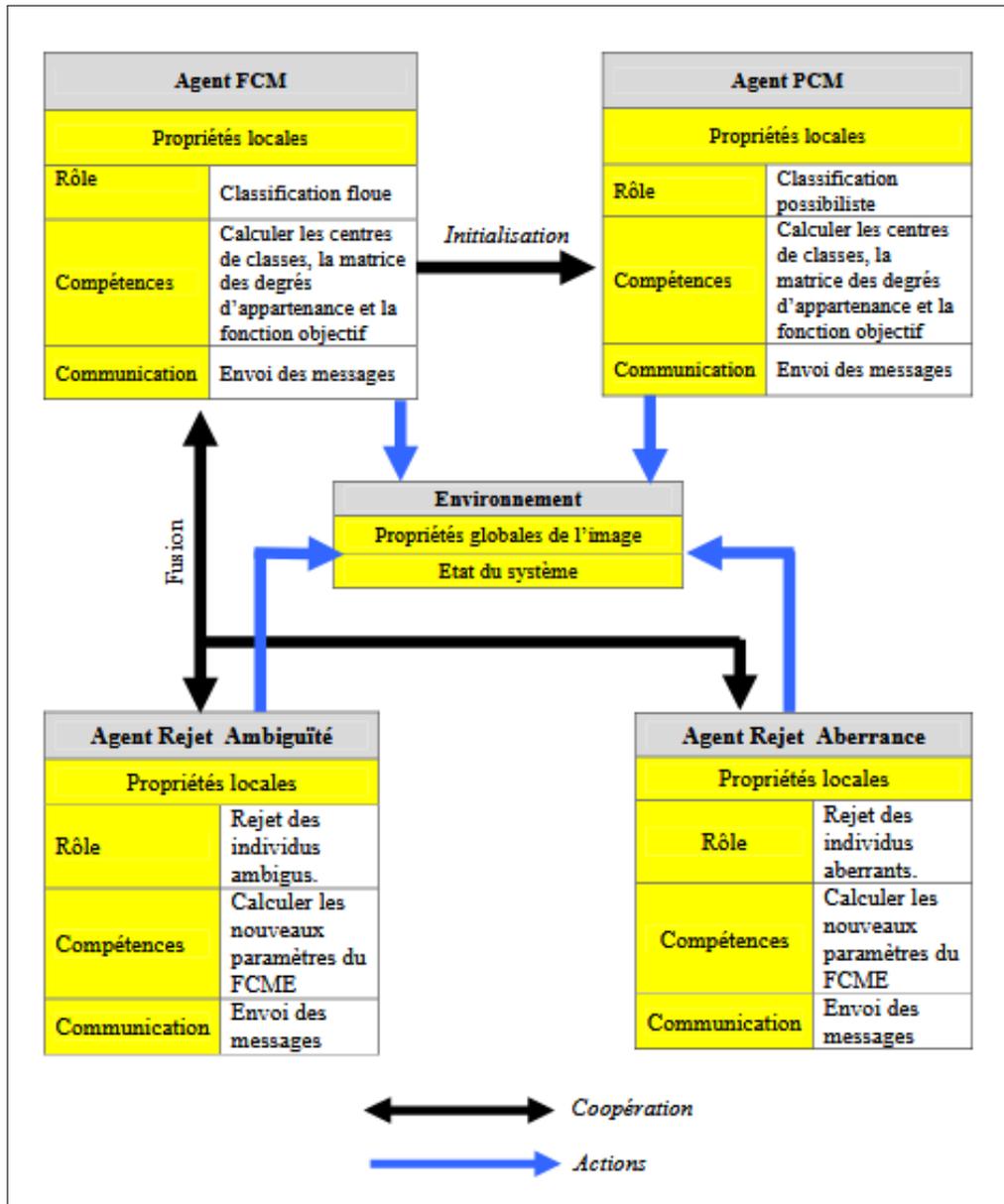


FIGURE 3.4 – Le système multi-agents proposé [Semchedine, 2007]

Boulkhiout [Boulkhiout, 2010] a développé une architecture multi-agents adaptative qui combine entre un algorithme supervisé qui est la Machine à Vecteur de Support (MVS) pour la détection de contour et un algorithme non supervisé basé sur les fourmis artificielles pour la détection et la localisation des régions homogènes (AntClustering). Leur objectif était de profiter des avantages de chacun, afin que la segmentation caractérise mieux les différentes structures cérébrales (MG, MB, LCR), l'auteur réalise aussi une comparaison entre ces deux types de méthodes.

Dans son approche, les agents sont les entités de base qui permettent de localiser des structures dans les images médicales. Dans chaque algorithme de l'approche, une population d'agents est lancée, où chaque agent segmente une partie de l'image (c'est-à-dire classe un ou plusieurs pixels). L'image segmentée résulte de la fusion des résultats des agents. Dans cette approche les agents permettent d'effectuer la classification des principaux tissus cérébraux, son architecture est présentée dans la figure 3.5.

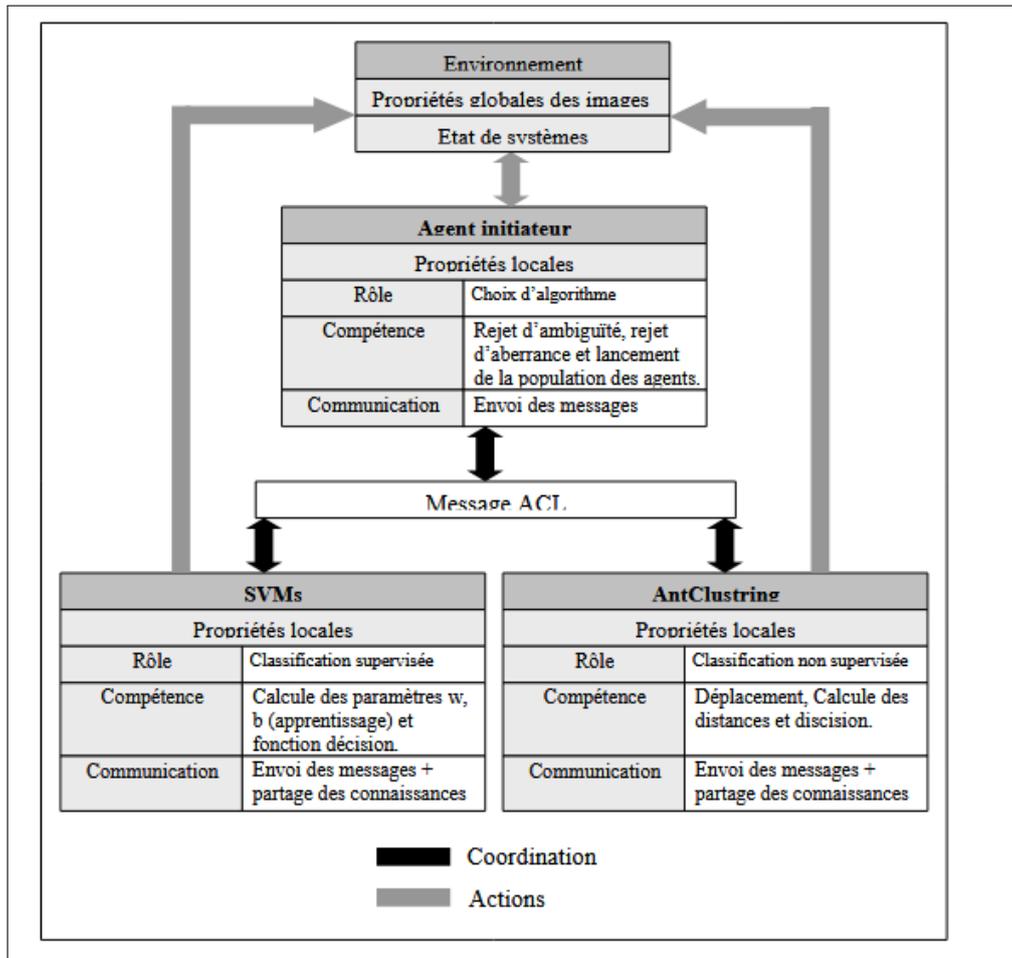


FIGURE 3.5 – l'interaction entre agents dans l'approche développée. [Boulkhiout, 2010]

### 3.4 Conclusion

Dans ce chapitre nous avons introduit certains des éléments de base concernant les SMAs. Nous avons également montré quelques exemples d'applications de ce paradigme dans le domaine de la segmentation d'images. Nous avons pu constater que les SMAs permettent plusieurs avantages, notamment une meilleure organisation, une diversité d'expertise, une architecture de contrôle puis aussi du parallélisme qui signifie une réduction en temps d'exécution. Dans le chapitre suivant nous allons exposer notre approche proposée, les approches que nous avons opté pour à savoir l'apprentissage profond et les SMAs, nous présenterons également nos résultats obtenus ainsi que notre application.

## Chapitre 4

### Approche proposée, résultats et discussions

## 4.1 Introduction

La première étape dans le processus de compréhension d’une image est le plus souvent l’extraction des objets qui la composent. Or, la segmentation est primordiale en analyse d’images, elle permet d’apporter une aide au diagnostic dans le domaine médical.

Le but de notre travail est alors de proposer une approche de segmentation sémantique des images médicales, tout en améliorant les performances de cette dernière et en réduisant le temps d’exécution. Pour cela, dans un premier nous avons opté pour des méthodes issues de l’apprentissage profond qui ont déjà montré leur efficacité en reconnaissance des formes et en détection.

Qu’il s’agisse de diagnostic, de planification, ou d’ordonnancement, les problèmes réels sont souvent abordés par un groupe d’entités. Sur ce, et dans un deuxième temps nous avons fait recours aux SMAs, qui offrent une certaine distribution physique et d’expertise, permettant ainsi d’aboutir à une meilleure résolution du problème. Pour valider notre travail nous avons utilisé plusieurs bases d’images à savoir Br35H [, 2022], CT Chest[kaggle2., 2022] et BRATS2020 [Med, 2022].

Dans ce chapitre, nous avons illustré nos quatre approches proposées, où nous avons présenté les méthodes et modèles que nous avons adoptés ainsi que les architectures des modèles de l’AP que nous avons introduit, pour par la suite exposer les résultats obtenus et quelques discussions. Nous clôturons le chapitre par une partie implémentation, où nous présentons quelques détails quant à cette dernière à savoir les environnements de travail, le langage et les bibliothèques, l’interface de notre application, puis finalement une conclusion.

## 4.2 Nos approches proposées

Afin d’aboutir à notre objectif principal, qui est la segmentation sémantique des images médicales, nous avons réalisé quatre contributions, la première en utilisant l’architecture U-Net, la deuxième est une hybridation entre un modèle résiduel de base qui est ResNet 50 et U-Net, la troisième une combinaison de résultats issus de six modèles profonds, est la quatrième et la dernière approche est constituée d’un double U-Net et un SMA .

Durant notre travail, nous nous sommes hélas confrontés à un problème, où il s’agit de trouver une base de données qui soit étiquetée, vu que nous utilisons des méthodes issues de l’AP appliquant un apprentissage supervisé. Malheureusement plusieurs de ces bases, soit elles sont d’un accès restreint, non étiquetées ou encore comportant quelques défauts.

Il faudrait toutefois noter que le choix du modèle, des paramètres à ajuster ou en encore le type d’hybridation dépend fortement de la base utilisée et par conséquent de l’objectif tracé. Car s’il s’agit d’une segmentation binaire ou multi classes, des images en IRM ou autres, les résultats ne peuvent être les mêmes, raison pour la quelle nous avons introduit ces quatre contributions.

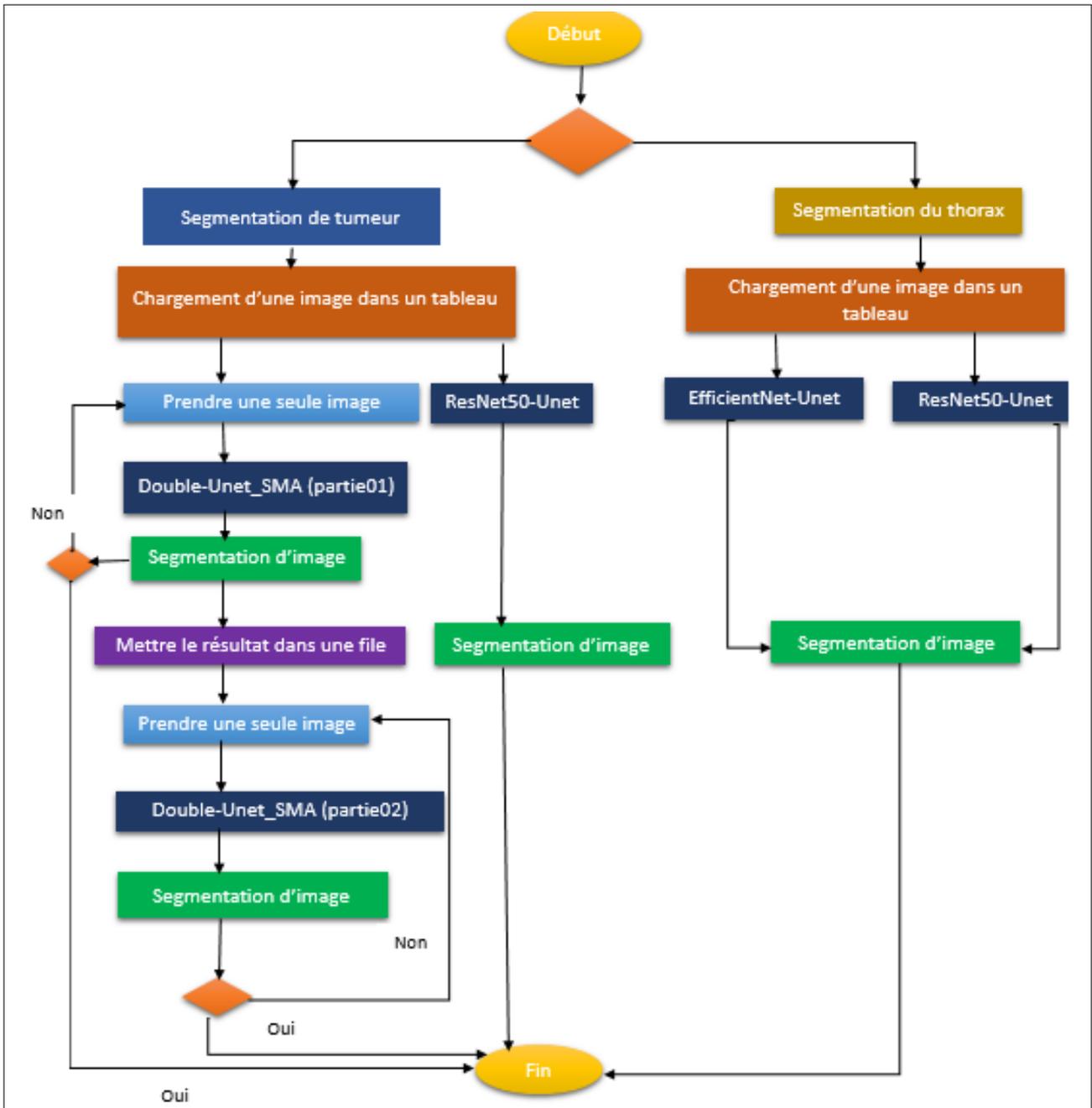


FIGURE 4.1 – Organigramme des approches proposées

## 4.3 Présentation de nos approches proposées

### 4.3.1 La première contribution : U-Net 1

L'architecture d'encodeur-décodeur est l'une des architectures qui sont largement utilisées dans la segmentation d'images médicales. L'encodeur prend l'entrée et la réduit en calculant la représentation des caractéristiques, et produit des cartes de caractéristiques (feature map) qui contiennent des informations codées de l'image d'entrée. Dans la partie décodeur, ces cartes sont échantillonnées, nous utilisons une architecture U-Net, incluons une méthode de connexion de type saute-couche (skip connection) pour partager les informations du codeur aux couches du décodeur [Ganaye, 2022].

Dans notre architecture, l'encodeur suit l'architecture typique d'un réseau convolutionnel. Il consiste à l'application répétée de deux convolutions  $3 \times 3$ , chacune suivie d'une unité linéaire rectifiée (ReLU) et d'une opération de max pooling de  $2 \times 2$  avec stride 2 pour le sous-échantillonnage. À chaque étape de sous-échantillonnage, nous doublons le nombre de canaux de fonction. Chaque étape de décodeur consiste en un sur-échantillonnage de la carte des caractéristiques, suivi d'une convolution  $2 \times 2$  qui réduit de moitié le nombre de canaux caractéristiques, et deux convolutions de  $3 \times 3$ , chacune suivit d'un ReLU. Le recadrage est nécessaire en raison de la perte de pixels de bordure dans chaque convolution. Au niveau final, une convolution  $1 \times 1$  est réalisée avec n filtre où chaque filtre représente une classe.

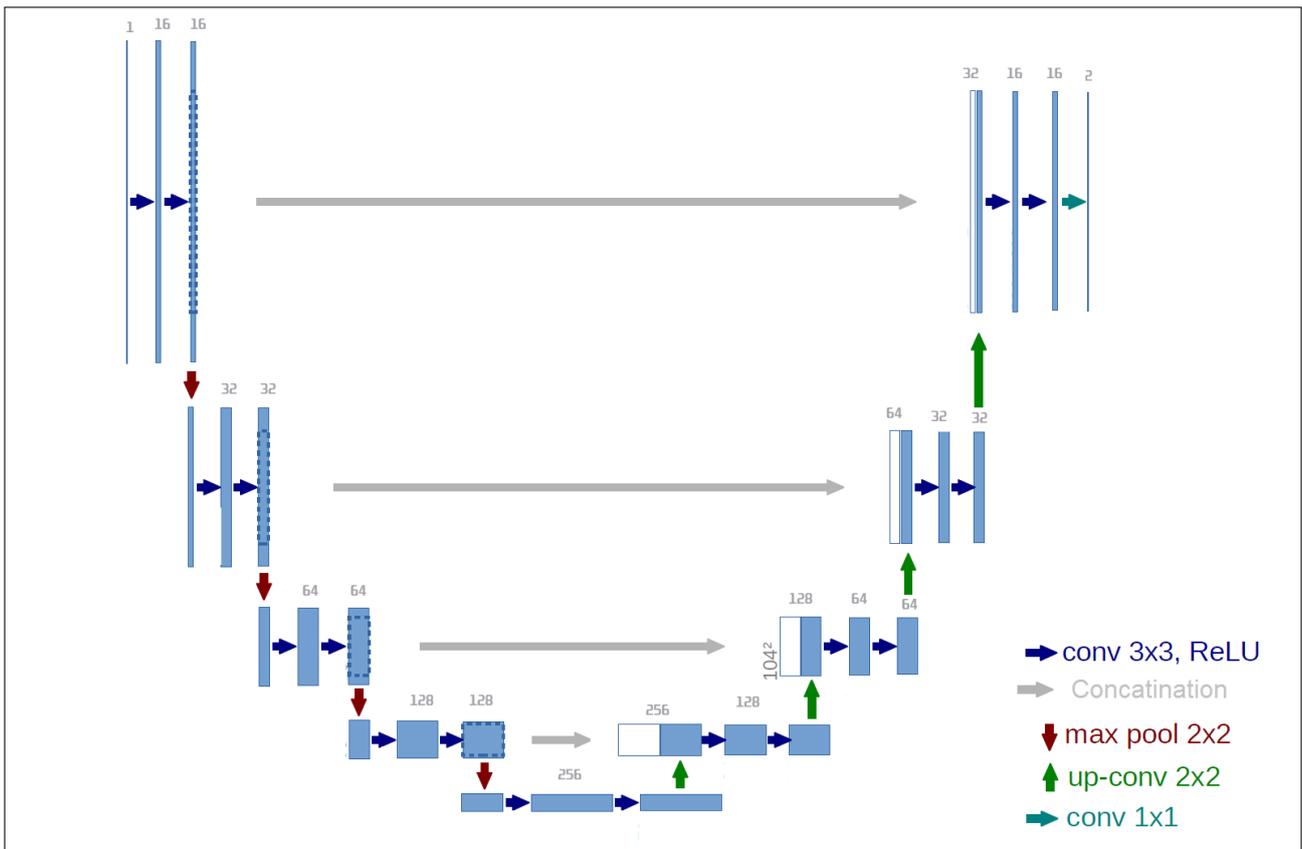


FIGURE 4.2 – L'architecture de notre modèle proposé U-Net1.

### 4.3.2 La deuxième contribution : ResNet50\_U-Net

Les modèles issus de l'apprentissage supervisé sont nombreux, nous avons adopté dans cette deuxième approche une hybridation entre U-Net et ResNet50. Il s'agit dans cette deuxième approche, de réaliser une segmentation multi-classes sur des images du thorax, où il est question de segmenter sémantiquement l'image en trois classes le cœur, la trachée et les poumons. Le ResNet permet de la création de réseaux de neurones très profonds, à la précision meilleure que ceux ayant des architectures linéaires car il a la capacité d'extraire davantage d'informations et d'avoir ainsi une analyse plus avancée des images [makina-corpus, 2022]. Nous utilisons ResNet-50 comme réseau fédérateur et fusionnons progressivement les fonctionnalités à différentes échelles.

De plus, la convolution dilatée est intégrée dans les cartes de caractéristiques multi-échelles pour extraire des caractéristiques de contexte riches.

#### \* Architecture de modèle proposé : ResNet50\_U-Net

Nous avons utilisé ResNet50 [Al-haijaet, 2020] comme backbone, la partie qui transforme une image en cartes de caractéristiques d'U-Net pour notre modèle.

L'utilisation d'un encodeur déjà entraîné aide le modèle à converger facilement. L'image d'entrée est introduite dans l'encodeur Res-Net50, composé d'une série de blocs résiduels comme composant principal. Ces blocs résiduels aident l'encodeur extrait les caractéristiques importantes de l'image d'entrée, qui sont ensuite transmises au décodeur.

Le décodeur démarre une convolution de transposition qui met à l'échelle les cartes de caractéristiques entrantes dans la forme souhaitée. Ensuite, ces cartes de caractéristiques mises à l'échelle sont concaténées avec les cartes de caractéristiques de forme spécifiques de l'encodeur entraîné via des connexions de type saute-couche. Ces connexions aident le modèle à obtenir toutes les informations sémantiques de bas niveau de l'encodeur, ce qui permet au décodeur de générer les cartes d'entités souhaitées. Après cela, il est suivi par les deux couches de convolution  $3 \times 3$ , où chaque couche est suivie d'une couche de normalisation par lots (batch normalization layer) et d'une couche de correction ReLU (RectifiedLinearUnits). La sortie du dernier bloc du décodeur est transmise à une couche de convolution  $1 \times 1$ , puis ensuite transmise à une fonction d'activation sigmoïde, générant finalement le masque binaire souhaité.

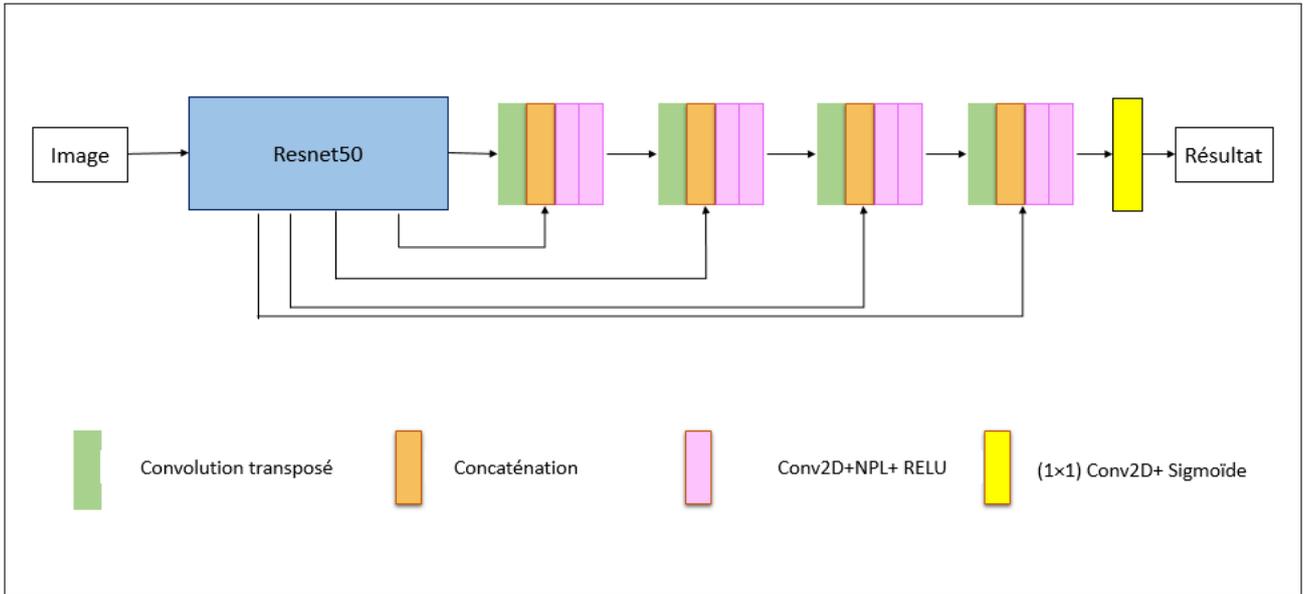


FIGURE 4.3 – l'architecture de modèle ResNet50\_U-Net.

### 4.3.3 La troisième contribution : combinaison de six modèles profonds

Dans cette contribution nous avons combiné les six modèles ci-dessous, dans le but d'améliorer encore plus les performances de segmentation :

- ResNet\_U-Net avec 50 époques ;
- ResNet\_U-Net avec meilleur point dans 120 époques ;
- Resnet \_U-Net avec meilleur point dans 120 époques avec augmentation de données ;
- U-Net avec 50 époques avec augmentation de données ;
- U-Net : meilleur point. dans 120 époques ;
- U-Net meilleur point dans 120 époques avec augmentation de données.

Nous avons pris le résultat des 06 modèles et calculé leur moyenne :

$$Image[i][j] = \frac{\sum_{i=1}^6 Imagen[i][j]}{6} \quad (4.1)$$

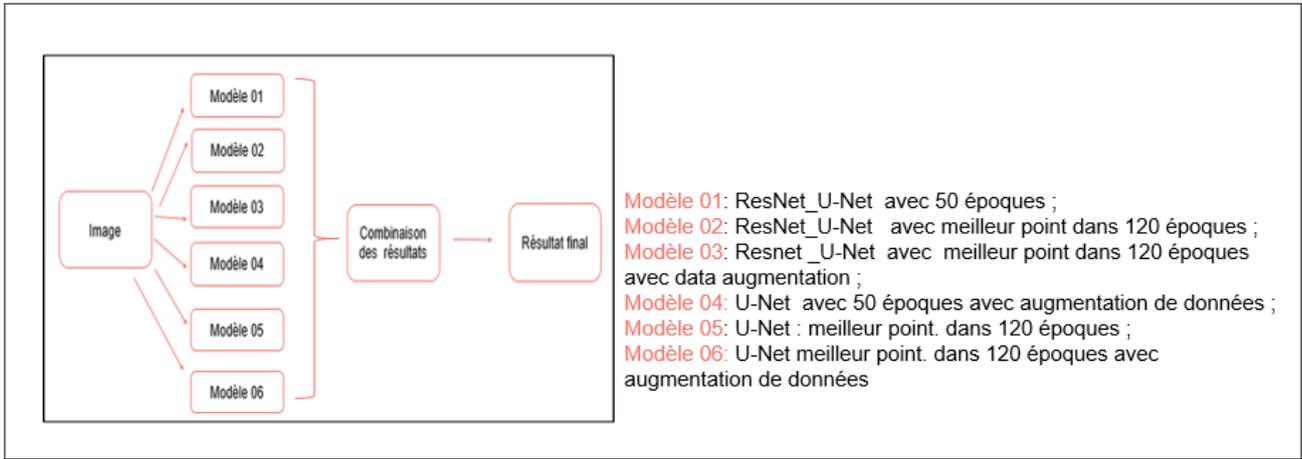


FIGURE 4.4 – Architecture de notre troisième contribution ”combinaison de résultats issus de six modèles profonds”.

### 4.3.4 La quatrième contribution : double U-Net\_SMA

Notre dernière contribution est caractérisée par l’introduction des SMAs et cela à deux niveaux le premier niveau dans l’architecture U-Net elle-même, le deuxième au niveau de la série d’images du même tranche.

Nos choix sont justifiés par les caractéristiques de la base de données utilisée cette fois-ci, où il s’agit de la base BRATS [Med, 2022], un challenge pour les programmeurs dans le monde de l’apprentissage profond. Cette base contient une série d’images pour chaque tranche, une fois que la segmentation est terminée, une reconstruction 3D est réalisée.

#### \* Architecture du modèle proposé double U-Net

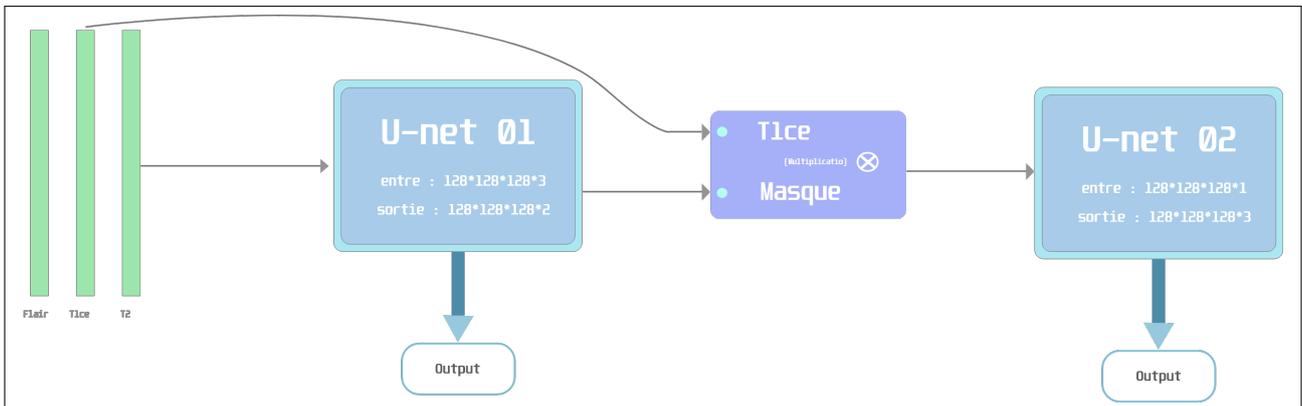


FIGURE 4.5 – Architecture de notre deuxième modèle hybride proposé Double U-Net\_Sma.

Sachant que le premier réseau U-Net 1 est déjà illustré dans la figure 4.2, le deuxième réseau U-Net 2 est présenté dans la figure 4.6.

Dans cette architecture, dans la partie d’encodeur, il consiste à l’application répétée de deux convolutions  $3 \times 3 \times 3$ , chacune suivie d’une unité linéaire rectifiée (ReLU) et d’une opération de max pooling de  $2 \times 2$ .

À chaque étape de sous-échantillonnage, nous doublons le nombre de canaux de fonction. Chaque étape

de décodeur consiste en un sur-échantillonnage de la carte des caractéristiques, suivi de deux convolutions de 3x3x3, chacune suivit d'un ReLU. Le recadrage est nécessaire en raison de la perte de pixels de bordure dans chaque convolution. Au niveau final, une convolution 1x1x1 et une fonction d'activation softmax.

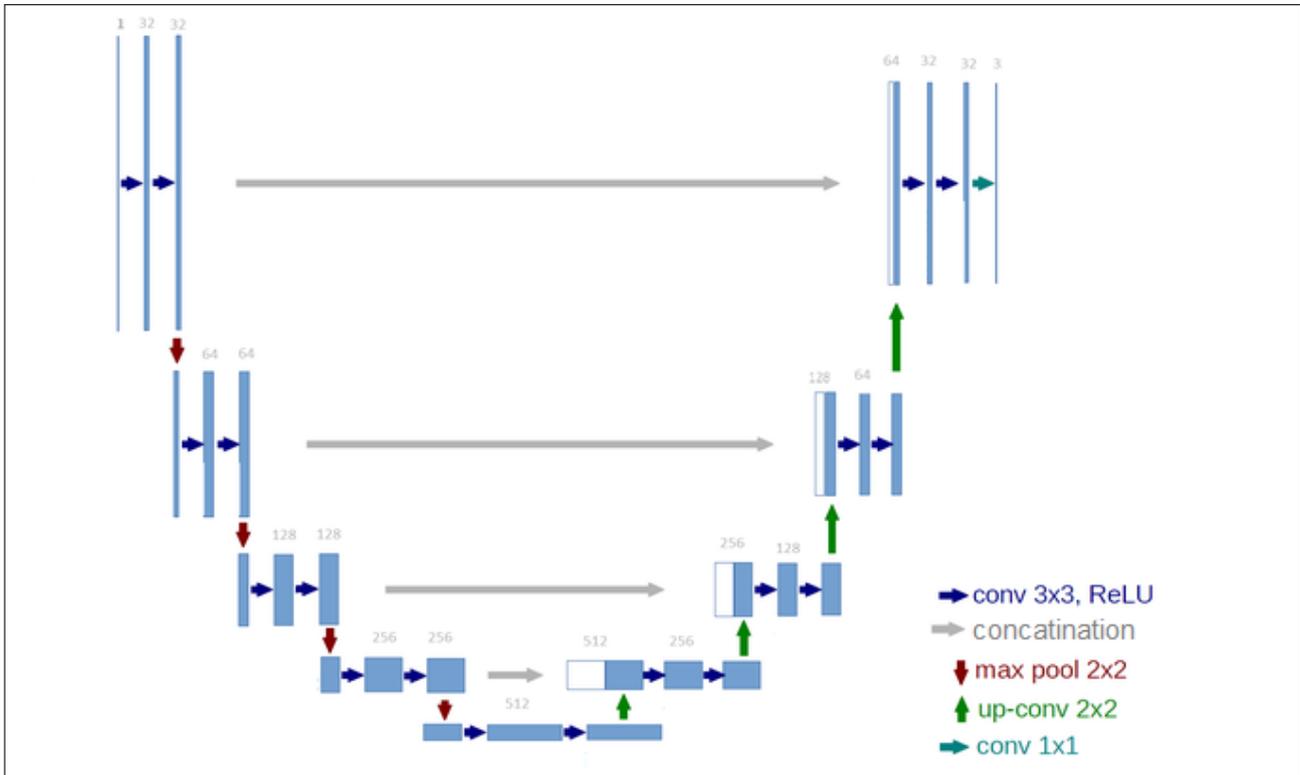


FIGURE 4.6 – L'architecture de notre modèle proposée U-Net 2.

\* **Architecture du SMA proposée** Nous avons proposé une architecture multi-agents dans le but de réaliser la tâche de segmentation des images en IRM de la base d'images (Brats 2020). Ce paradigme permet d'apporter une distribution d'expertises et de réduire notamment le temps d'exécution.

Nous avons utilisé plusieurs agents caractérisés par une coopération de tâche et de résultats. La communication entre les agents est indirecte par partage de ressources, ils communiquent à travers une zone de données commune. La figure 4.7 illustre une modélisation de l'interaction entre les agents dans l'approche développée sur une série d'images d'une tranche3D (slice).

Dans notre approche Double U-Net\_SMA nous avons utilisé 3 types d'agents :

- **Agent interface** : c'est l'agent qui s'exécute en premier lors du lancement de l'application, parmi ses tâches principales : le chargement de l'image, lancement des autres agents et l'affichage des résultats
- **Agent U-Net 1** : son rôle est de segmenter l'image initiale en détectant la tumeur entière.
- **Agent U-Net2** : il reçoit le résultat de l'agent U-Net 1 et réalise une deuxième segmentation, en déterminant la tumeur centrale (core tumor) et la tumeur rehaussé (enhancing tumor).

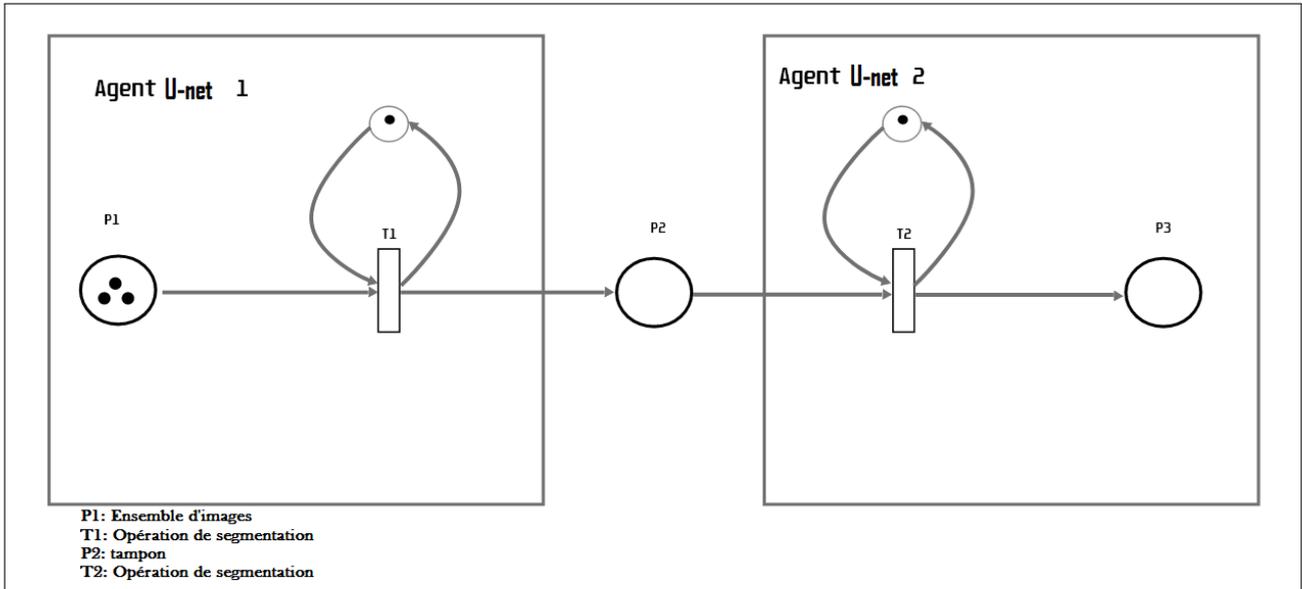


FIGURE 4.7 – Modélisation de l’interaction entre agent 2 et agent 3 dans notre approche proposée double U-Net\_SMA avec le réseau de pétri.

## 4.4 Ajustement des paramètres

Nous avons été confrontés à un problème d’ajustement de paramètres lors de la réalisation nos modèles de l’AP, parmi ces paramètres nous citons :

### \* Nombre d’époque (epoch)

Quand les données sont trop volumineuses pour être traitées en même temps, on divise la population d’apprentissage en lots (batch). Seules les données du lot sont présentées au réseau de neurones et les poids sont adaptés en fonction de ces derniers. On parle d’une époque (epoch) lorsque toutes les données du lot ont été présentées [Durand, 2019].

Il n’existe pas de méthode rigoureuse pour déterminer le nombre d’époque, par conséquent nous avons essayé plusieurs valeurs et calculer par conséquent les performances (voir table 4.1).

### \* Point de contrôle(Checkpointing )

Les points de contrôlesont configurés pour enregistrer les poids du réseau uniquement lorsqu’il y a une amélioration de la précision de la classification sur l’ensemble de données de validation.

\* **L’augmentation de données** C’est un ensemble de techniques permettant d’augmenter artificiellement la quantité de données en générant de nouvelles de données à partir de données existantes. Nous avons adopté le zoom.

### \* L’apprentissage par transfert

Consiste à prendre des modèles pré-entraînés sur une base de données de référence standard de vision par ordinateur (tel que : ImageNet, COCO, FMNIST, etc.), il est question alors de congeler les couches inférieures pour ne pas être entraînés à nouveau et à réutiliser ses poids en plus d’ajouter quelques couches en haut [Mebrouki, 2021].Les modèles réseduels tels que ResNet-50 que nous avons adopté permettent un apprentissage par transfert.

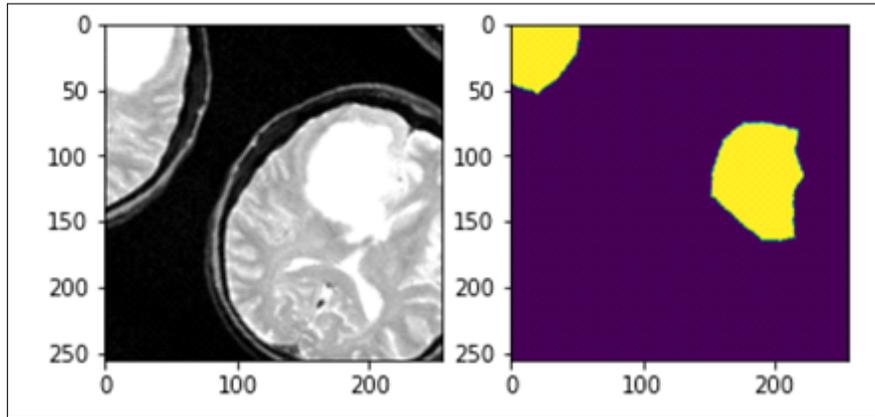


FIGURE 4.8 – Exemple d’augmentation de donnée sur une image de la base Br35H.

## 4.5 Résultats obtenus, discussions et comparaisons

Afin d’étudier la performance de nos approches proposées, nous avons entraîné nos modèles de l’AP sur des images issues de trois bases à savoir Br35H [kaggle1, 2022], CT Chest [kaggle2., 2022] et BRATS2020 [Med, 2022].

Et pour montrer l’efficacité de nos approches de segmentations proposées, nous avons réalisé des comparaisons en utilisant plusieurs métriques.

### 4.5.1 Les bases d’images utilisées

- **Br35H Brain TumorDetection 2020**[kaggle1, 2022] : est un ensemble d’images de 801 en format JPEG, ils sont étiquetés dans un fichier JSON qui contient des points représentant le contour de la tumeur. Cette base est partitionnée en 500 images pour l’entraînement, 201 images pour la validation et 100 images pour le test.
- **Chest CT Segmentation** : est un ensemble d’images étiquetés de 16708 images, ces images sont réparties en 3 classes : poumon, cœur , trachée. Cet ensemble de données est public [kaggle2., 2022]
- **Brain Tumor Segmentation (BRATS 2020)** : BRATS est un ensemble d’images étiquetées, il existe plusieurs versions de la base à savoir BRATS 2018 [kaggle3.com, 2022] et BRATS2020 [Med, 2022]. La base utilise des IRMs préopératoires multi-institutionnelles et se concentre principalement sur la segmentation des tumeurs cérébrales intrinsèquement hétérogènes (en apparence, forme et histologie), à savoir les gliomes. Tous les scans multimodaux BraTS sont disponibles sous forme de fichiers NIFTI (.nii.gz) et décrivent : a) natif (T1) et b) post-contraste pondéré en T1 (T1Gd), c) pondéré en T2 (T2) et d) T2 fluide volumes de récupération d’inversion atténuée (T2-FLAIR), et ont été acquis avec différents protocoles cliniques et divers scanners de plusieurs institutions (n = 19), mentionnées ici comme contributeurs de données.

Tous les jeux de données d’imagerie ont été segmentés manuellement, par un à quatre évaluateurs, suivant le même protocole d’annotation, et leurs annotations ont été approuvées par des neuro-radiologues expérimentés [google, 2022].

## 4.5.2 Estimation de performances

Le processus d'estimation des performances est indispensable pour illustrer le degré d'efficacité des méthodes et approches choisis. Nous avons tout au long de notre étude, opté pour quatre estimateurs qui à leur tour utilisent les métriques de la matrice de confusion à savoir [Dahmane, 2020] :

- **VN (TN)** : nombre des pixels correctement rejetés ;
- **VP (TP)** : nombre des pixels correctement identifiés ;
- **FP (FP)** : nombre des pixels incorrectement identifiés ;
- **FN (FN)** : nombre des pixels incorrectement rejetés.

Les estimateurs de performance utilisés sont les suivants :

- **Précision PR (Accuracy)** : Elle est calculée comme suit via la formule suivante :

$$précision = PR = \frac{VP + VN}{VP + VN + FP + FN} \quad (4.2)$$

- **La Sensibilité ou Taux de Vrai Positif TVP (Positive PredictiveRate)** :

Elle est calculée comme suit :

$$Sensibilité = TVP = \frac{VP}{VP + FP} \quad (4.3)$$

- **Le Taux de Faux Positifs (TFP) (False Predictive Rate)** :

Ile est calculée comme suit :

$$= \frac{FP}{VP + FP} \quad (4.4)$$

- **L'intersection sur l'union (IoU : Means-Intersection On Union)** Également connue sous le nom de coefficient de similarité deJaccard(Jaccard Similarity Coefficient), est la métrique la plus utilisée. Pour chaque classe, l'IoU est le rapport entre les pixels correctement classés et le nombre total de pixels réels et prévus dans cette classe.Il est calculé comme suit :

$$IoU = \frac{VP}{VP + FP + FN} \quad (4.5)$$

## 4.5.3 Résultats et discussions :

Afin d'étudier les performances de nos approches proposées, et de démontrer l'efficacité des hybridations réalisées, nous avons appliqué nos algorithmes développés sur trois bases. Le choix du modèle ainsi que de l'hybridation proposée dépend principalement des objectifs tracés et des caractéristiques de la base.

Où dans un premier temps, il était question d'une segmentation binaire, nous avons alors utilisé la base Br35 [kaggle1, 2022]. Puis par la suite, nous avons effectué une segmentation multi- classes, alors nous avons utilisé les bases CT Chest [kaggle2., 2022] et BRATS2020[Med, 2022] qui représentent des séries d'images qui composent des tranches d'images 3D .

a) **Présentation et comparaison de résultats de segmentation entre modèle U-Net1, ResNet50\_U-Net, et combinaison de 6 modèles**

- **Application sur la base Br35 [kaggle1, 2022]** : Nous avons appliqué nos trois premières contributions dans un premier temps sur la base Br35 [kaggle1, 2022](voir section 4.5.1), contenant des images de cerveau avec tumeur ainsi que leur vérité terrain .Il s’agit ici d’une segmentation sémantique binaire en deux classes, la tumeur et le fond.

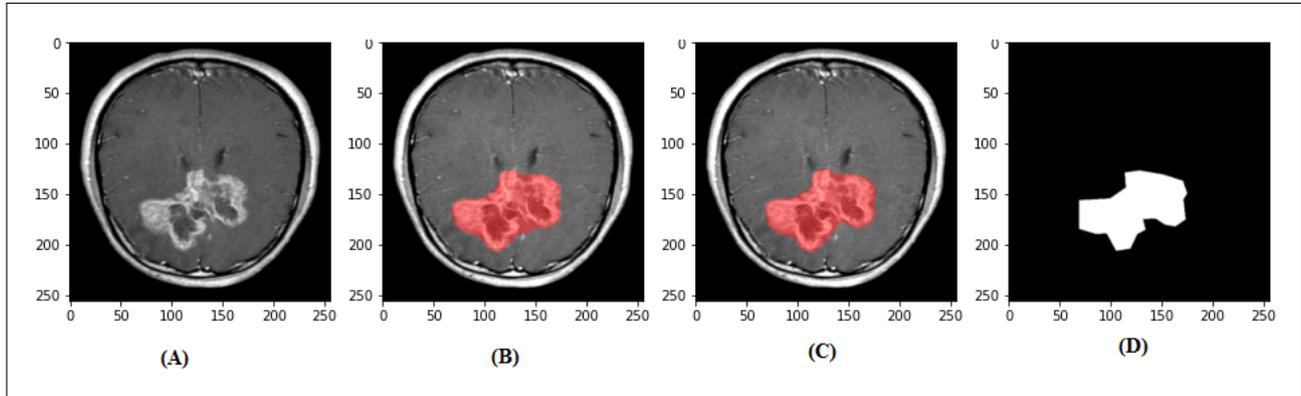


FIGURE 4.9 – Comparaison de résultats de segmentation d’une image de la base Br35 [kaggle1, 2022]. via les modèles proposées U-Net 1, ResNet50\_U-Net. (A) : image originale ,(B) : U-Net 1,(C ) : ResNet50\_U-Net , (D ) :Masque(vérité terrain).

Une comparaison quantitative via les estimateurs de performances de segmentation à savoir la PR, IoU, TVP et TFP sont présentés dans la table 4.1

TABLE 4.1 – Comparaison des résultats obtenus sur les deux modèles proposés : U-Net 1 et ResNet50\_U-Net avec différentes époques.

	U-Net 1 (Contribution1)				Resnet 50-UNet (Contribution2)			
	PR(%)	IoU(%)	TVP(%)	TFP(%)	PR(%)	IoU(%)	TVP(%)	TFP(%)
50 époques	97.58	65.20	83.39	16.60	98.43	76.44	93.58	06.41
120 époques	98.28	75.44	87.45	12.54	98.61	78.94	92.24	07.75
120 époque + AD	98.64	79.74	90.70	09.29	98.83	82.48	91.49	08.50

La table 4.1 illustre une première comparaison entre les résultats de nos deux approches U-Net 1 et ResNet50\_U-Net .Les résultats obtenus sont exprimés en termes de précision, de TVP, de TFP, et de IoU. Nous remarquons que le modèle **ResNet50-UNet** a de meilleurs résultats que le modèle **U-Net** avec 98.83% de **précision** , 82.48% de **IOU** , 91.49% et 08.50% de **TFP** .

Afin d’améliorer encore plus le résultat nous avons proposé notre troisième contribution ”Combinaison de six modèles”.La figure 4.10 illustre les résultats de segmentation obtenus.

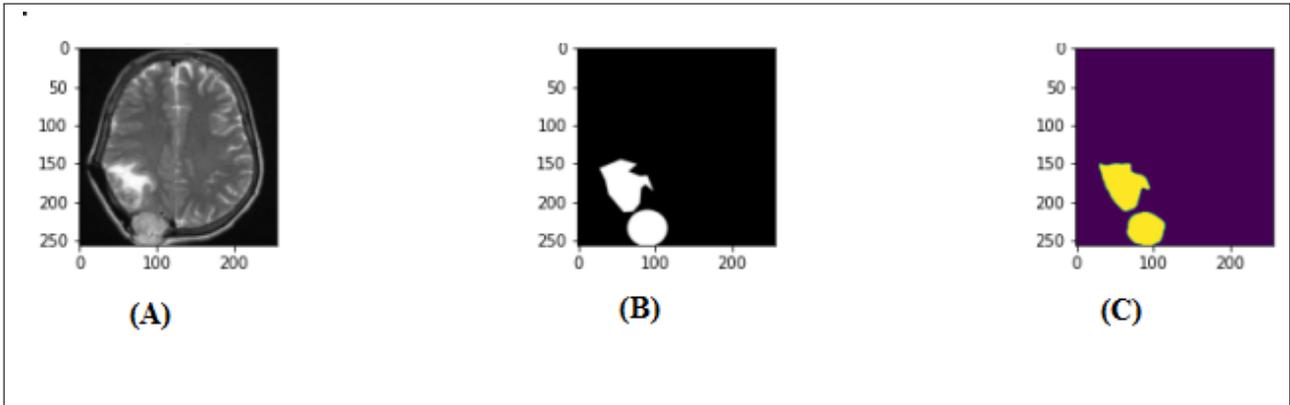


FIGURE 4.10 – Résultats de segmentation d’une image de la base Br35 [kaggle1, 2022]. via le modèle proposée combinaison de six modèles. (A) : image original ,(B) : Masque(verité terrain) ,(C ) : notre combinaison proposée

TABLE 4.2 – Résultats de segmentation par le modèle ResNet50\_U-Net sur des images 2D de la base CT Chest [kaggle2., 2022]. (A) : verité terrain, (B) : segmentation via approche hybride proposée ResNet50\_U-Net

Approche	PR(%)	IoU(%)	TVP(%)	TFP(%)
<b>Contribution3</b>	98,89%	83,24%	92,62%	07,37%

Les résultats de segmentation obtenus et illustrés dans la table 4.2 montrent l’efficacité de cette deuxième approche hybride proposée qui consiste à une combinaison de résultat entre les six modèles de ResNet50\_U-Net et U-Net 1. Nous remarquons une amélioration par rapport aux résultats précédents présentés dans la table 4.1. Nous avons obtenu une augmentation sur le taux de Précision de 98,89%, l’IOU de 83,24%, le TVP de 92,62%, ainsi qu’une réduction du TFP avec un taux de 07,37%.

- **Application sur la base Chest CT segmentation**

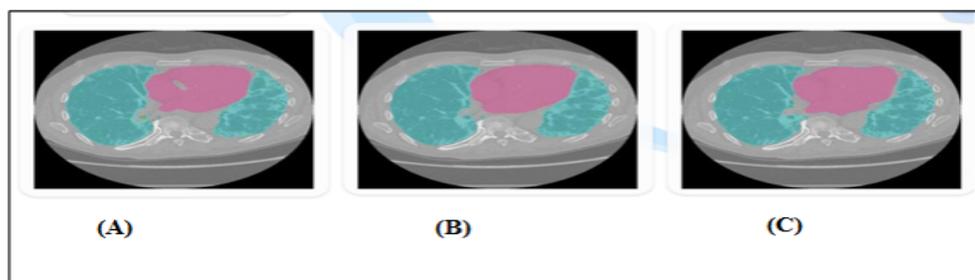


FIGURE 4.11 – Résultats de segmentation par le modèle ResNet50\_U-Net et EfficientNet-U-net sur des images 2D de la base CT Chest[kaggle2., 2022]. (A) : EfficientNet\_U-Net, (B) ResNet50\_U-Net, (C) : verité terrain

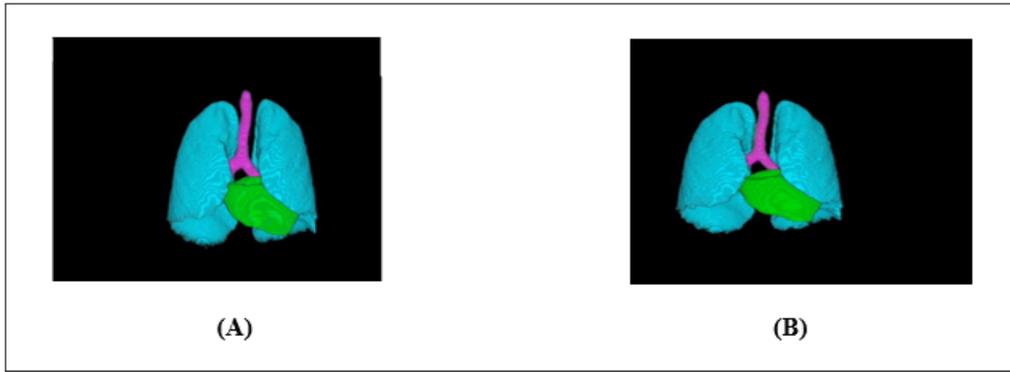


FIGURE 4.12 – Résultats de segmentation par le modèle proposé Resnet50\_U-Net sur des images 3D de la base CT Chest [kaggle2., 2022].(A) : vérité terrain, (B) segmentation via approche hybride proposée ResNet50\_U-Nnet

TABLE 4.3 – Comparaison des résultats obtenus sur les deux modèles de base (Efficient net et Resnet50)

	PR(%)	IoU(%)	TVP(%)	TFP(%)
<b>ResNet50_U-Net</b>	99.42%	89.22%	95.14%	04.85%
<b>EfficientNet_U-Net</b>	99.40%	89.22%	94.05%	05.94%

D’après la table 4.3, on observe que le modèle ResNet50 a de meilleurs résultats par rapport au modèle efficiente Net en matière de précision, TVP et TFP.

b) **Présentation et comparaison de résultats de segmentation entre modèle Double U-Net\_SMA et ResNet50\_U-Net**

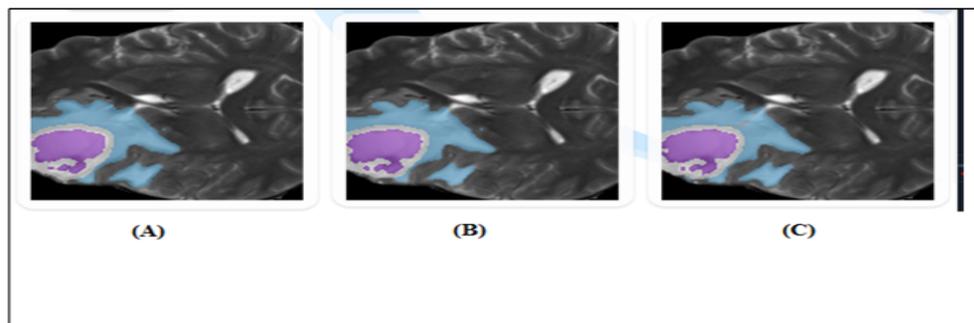


FIGURE 4.13 – Résultat de segmentation obtenu via notre approche hybride proposée Double U-Net\_SMA sur une image 2D de la base Brats 2020, (A) : Resnet50, (B) : segmentation via le model double U-Net\_SMA, (C) : vérité terrain

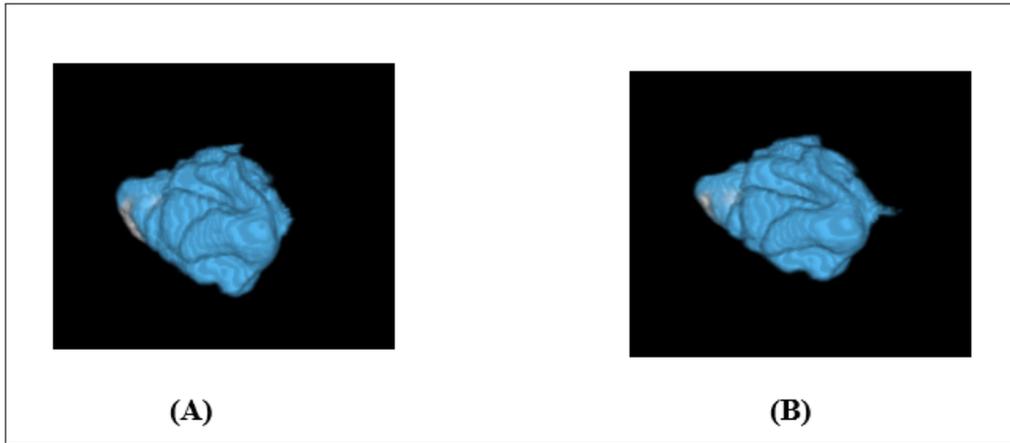


FIGURE 4.14 – Résultat de segmentation via notre approche hybride proposée Double U-Net\_SMA sur une image 3D de la base Brats 2020. (A) : Vérité terrain, (B) Model double U-Net\_SMA.

Les résultats illustrés dans les figures 4.13 et 4.14 montrent une similarité entre les images segmentées via notre approche hybride Double U-Net\_SMA et la vérité terrain. Pour plus de précision en matière de résultats, nous calculons les performances de segmentation via les estimateurs PR, IoU, TVP et TFP (voir table 4.4).

TABLE 4.4 – Comparaison des résultats de segmentation obtenus via nos deux approches hybrides proposées Double U-Net\_SMA et ResNet50\_U-Net.

Approche/performance	PR(%)	IoU(%)	TVP(%)	TFP(%)
<b>Double U-Net_SMA Contribution 4</b>	98.87%	72.52%	87.74%	12.25%
<b>Resnet50_U-Net Contribution 2</b>	98.80%	70.16%	83.96%	16.03%

D’après la table 4.4, nous remarquons que le modèle Double U-Net\_SMA permet d’aboutir à de meilleurs résultats que ceux du modèle U-Net. Avec 98.87% de précision ,72.52% de IOU , 87.74% TVP et 12.25%de TFP.

Afin de prouver l’efficacité de l’utilisation des SMAs, nous calculons le temps d’exécution avec et sans usage de ce paradigme. Les résultats sont illustrés dans la table 4.5 présentée ci-dessous.

TABLE 4.5 – Comparaison en matière de temps d’exécution du modèle Double U-Net avec et sans SMA .

Approche	Temps d’exécution (s) (CPU-PC)	Temps d’exécution (s) (Colab+ GPU )
<b>Double U-Net_SMA</b>	192.6	11.01
<b>Double U-Net</b>	201	16.49

La table 4.5 illustre la différence entre les deux modèles en matière de temps d’exécution, après l’exécution de 10 images d’une tranche 3D , nous remarquons que double Unet-SMA est plus rapide.

## 4.6 Implémentation

### 4.6.1 Environnement de travail

\* **Environnement matériel** La configuration du matériel utilisé dans notre implémentation est :

- Un pc portable
- Carte graphique GTX 960 M
- CPU I7
- RAM de taille 8 Go
- Disque dur de taille 128 Go ssd
- System d'exploitation Windows 10 64 bits

\* **Environnement logiciel**

• **Google Colaboratory** : Colaboratory ou 'Colab'. Permet d'écrire et d'exécuter le code Python de son choix par le biais du navigateur. Offert par Google (gratuit), basé sur Jupyter Notebook et destiné à la formation et à la recherche dans l'apprentissage automatique. Cette plateforme permet d'entraîner des modèles de l'apprentissage directement dans le Cloud. Colab permet [Djaballah, 2021] :

- D'améliorer les compétences de codage en langage de programmation Python.
- De développer des applications en Deep Learning en utilisant des bibliothèques Python populaires telles que Keras, TensorFlow, PyTorch et OpenCV.
- D'utiliser un environnement de développement (Jupyter Notebook) qui ne nécessite aucune configuration, mais la fonctionnalité qui distingue Colab des autres services est l'accès à un processeur graphique GPU, totalement gratuit.

• **Keras** :

Keras est une API de réseaux de neurones de haut niveau, écrite en Python et capable de fonctionner sur TensorFlow ou Theano. Il a été développé en mettant l'accent sur l'expérimentation rapide, être capable d'aller de l'idée à un résultat avec le moins de délai possible est la clé pour faire de bonnes recherches. Il a été développé dans le cadre de l'effort de recherche du projet ONEIROS (Open-ended Neuro-Electronic Intelligent Robot Operating System) et son principal auteur et mainteneur est François Chollet, un ingénieur Google[Mokri, 2017].

• **Python** :

Python est un langage de programmation de haut niveau interprété (il n'y a pas d'étape de compilation) et orienté objet avec une sémantique dynamique. Il est très sollicité par une large communauté de développeurs et de programmeurs. Python est un langage simple, facile à apprendre et permet une bonne réduction du coût de la maintenance des codes. Les bibliothèques (packages) python encouragent la modularité et la réutilisabilité des codes. Python et ses bibliothèques sont disponibles (en source ou en binaires) sans charges pour la majorité des plateformes et peuvent être redistribués gratuitement.[Mokri, 2017]

• **Pycharm** :

Pycharm est un environnement de développement intégré utilisé pour programmer en python.il

permet l'analyse de code et contient un débogueur graphique. Il permet également la gestion des tests unitaires, l'intégration de logiciel de gestion de versions, et supporte le développement web avec Django, développé par l'entreprise tchèque JetBrains, c'est un logiciel multi-plate-forme qui fonctionne sous Windows, Mac OS X et GNU/Linux. Il est décliné en édition professionnelle, diffusé sous licence propriétaire et en édition communautaire diffusé sous licence Apache.[wikipedia, 2022]

- **Tkinter :**

est une bibliothèque d'interface utilisateur graphique portable (GUI) open source conçue pour être utilisée dans les scripts Python. Tkinter repose sur la bibliothèque Tk, la bibliothèque GUI utilisée par Tcl/Tk et Perl, qui est à son tour implémentée en C. Par conséquent, on peut dire que Tkinter est implémenté en utilisant plusieurs couches[Djaballah, 2021].

## 4.6.2 Présentation de notre application

- **Interface d'accueil** L'interface d'accueil de notre application permet d'accéder à toutes ses principales fonctionnalités. Elle contient les boutons :
  - la bouton **poumon segmentation** : c'est pour l'importation de l'image originale de la base CT ChestSementation.
  - la bouton **Tumeur segmentation** : pour l'importation de l'image originale de la base BRATS2020

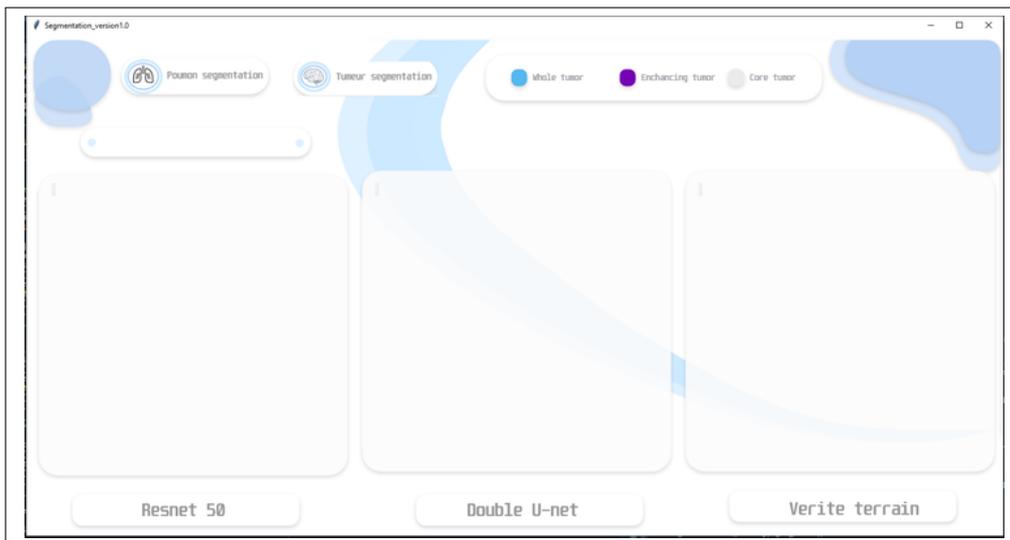


FIGURE 4.15 – Interface d'accueil de notre application.

- **Chargement d'une image**

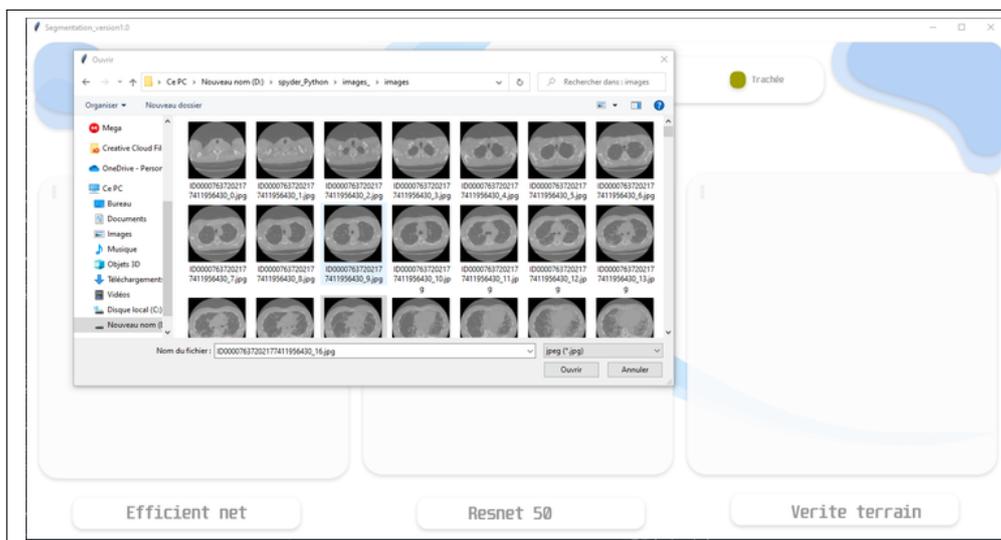


FIGURE 4.16 – Chargement d'une image de la base CT Chest segmentation.

- Affichage de résultat

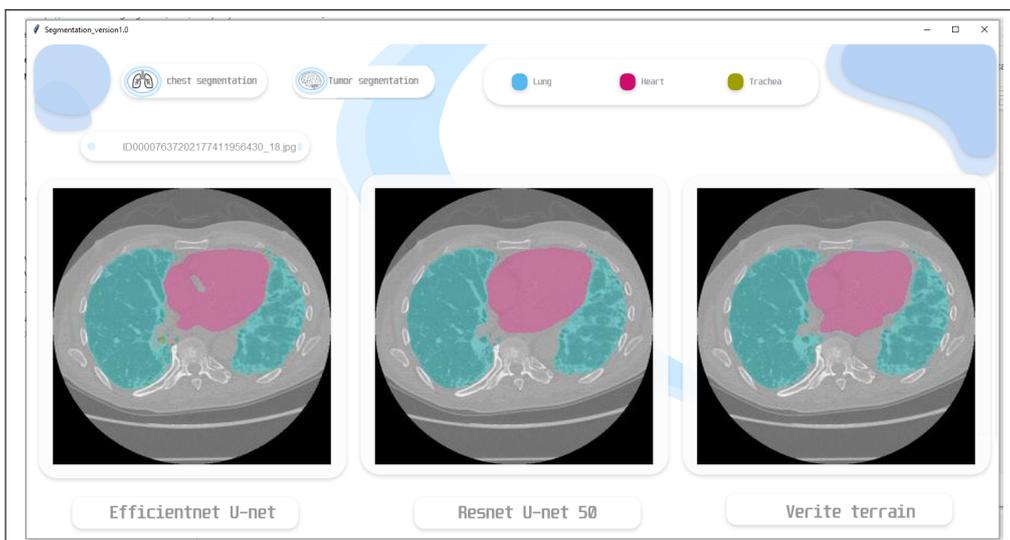


FIGURE 4.17 – Affichage des résultats de segmentation obtenus via notre modèle ResNet50 et le modèle EfficientNet sur des images de la base CT Chest segmentation.

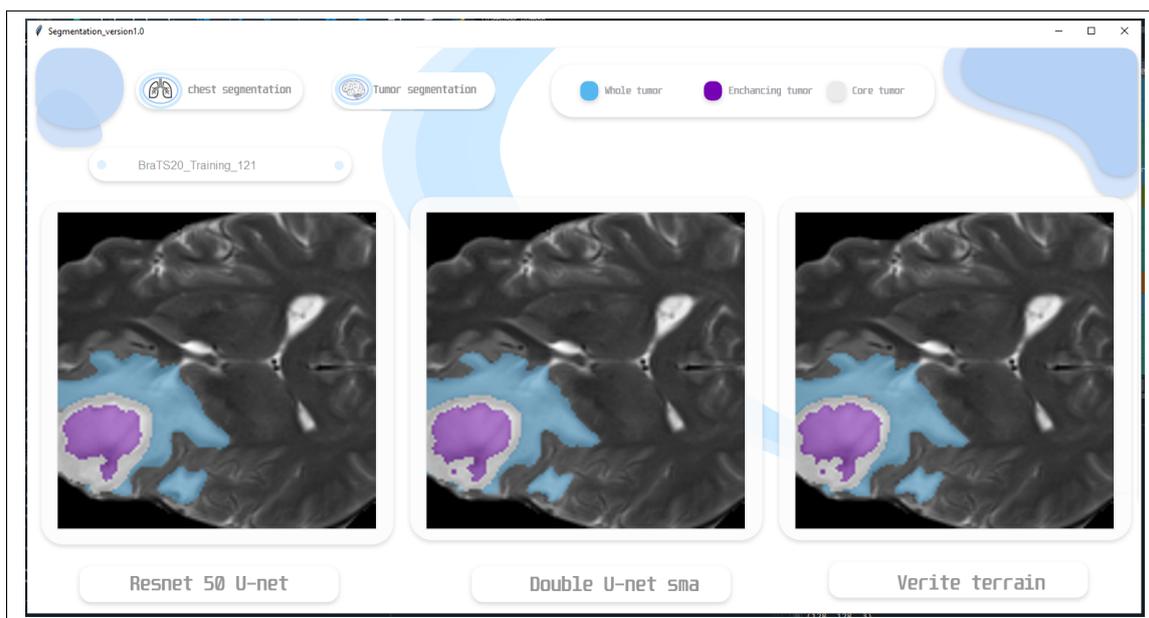


FIGURE 4.18 – Affichage des résultats obtenus de notre modèles ResNet50 et le modèle Double U-Net sur des images de la base BRATS2020.

## 4.7 Conclusion

La segmentation d'images médicales joue un rôle primordial dans les systèmes de diagnostic, il est question alors d'essayer d'aboutir à de meilleurs résultats possibles en matière de performance.

Tout au long de notre étude, nous nous sommes situés dans une approche d'apprentissage profond, qui a montré son efficacité dans ce domaine. Nous avons proposé quatre contributions, à savoir U-Net1, approche hybride ResNet50-U-Net, une approche hybride via une combinaison de résultat de 6 modèles et finalement une approche hybride double U-Net\_SMA. Chaque approche est dotée de sa propre architecture et ses propres paramètres, permettent d'obtenir des résultats différents. Pour entraîner, valider nos approches, nous avons utilisé trois bases d'images à savoir Br35H (Brain tumor 2020) [kaggle1, 2022], CT Chest [kaggle2., 2022] et BRATS2020 [Med, 2022].

L'utilisation des SMA a permis de réduire le temps d'exécution. Les résultats obtenus sont très encourageants, nous avons pu améliorer à chaque fois les performances de segmentation. Les résultats obtenus via nos quatre contributions ont montré que l'approche double U-Net\_SMA est le plus robuste et efficace en termes de taux de précision, d'IoU, TVP et TFP.

# Conclusion générale

L'analyse d'images est un domaine très vaste qui touche plusieurs disciplines notamment la médecine, il peut fournir une aide précieuse dans la détection et le suivi des maladies.

La segmentation sémantique est l'une des opérations les plus importantes dans un système de traitement d'images. Elle aide à fournir au médecin des informations de haut niveau sur le contenu de l'image.

Dans ce mémoire, nous nous sommes intéressés à la segmentation sémantique des images médicales en faisant recours à l'apprentissage profond. Nous avons alors présenté dans un premier chapitre, quelques notions de base sur le traitement d'images et l'imagerie médicale. Dans le deuxième chapitre nous avons introduit les différents types et techniques de la segmentation, nous avons par la suite abordé l'apprentissage profond (Deep Learning ) en détaillant le principe des réseaux de neurones artificiels et les réseaux de neurones convolutifs nous avons également présenté un état de l'art dans ce contexte.

Dans le troisième chapitre, nous avons présenté quelques concepts généraux sur les systèmes Multi-agents ainsi que quelque approche utilisant les SMAs en vue d'une segmentation.

Le quatrième chapitre est consacré aux approches proposées, à l'implémentation et aux résultats obtenus. Nous avons proposé quatre approches à savoir U-Net, ResNet50\_U-Net, une combinaison de six modèles convolutionnels, et finalement une approche Double U-Net\_SMA. Nos approches sont appliquées sur trois différentes bases de données, les résultats obtenus sont satisfaisants. Le choix du modèle dépend principalement du type d'image 2D ou 3D, ainsi que de l'objectif tracé à savoir une segmentation binaire ou multi-classes.

Afin d'améliorer les résultats, nous avons appliqué une augmentation de données, et effectuer plusieurs tests pour choisir le bon nombre d'époques.

Une comparaison entre ces différentes contributions amène à dire que le modèle Double U-Net\_SMA donne de meilleures performances en matière de précision, d'indice de Jaccard, taux des vrais et des faux positifs.

# Annexes

## .1 Python

Nous avons utilisé :

**Matplotlib** : c'est une bibliothèque utilisée pour afficher et lire et écrire des images en format png.

**OpenCv** : pour les traitements des images tel que resize création des vidéos avec un ensemble d'images.

**Nibabel et SimpleITK** : Ce package utilisé pour lire et écrire les images en format .nii .nifty fichier pour la base d'images Brats2020.

**NumPy** : On trouve dans le module NumPy :

- les outils de manipulation des tableaux pour le calcul numérique.
- Nombreuses fonctions de manipulation de tableaux.

Les tableaux NumPy sont plus rapides que les listes Python pour les raisons suivantes

- Un tableau est une collection de types de données homogènes qui sont stockés dans des emplacements de mémoire contigus.
- D'autre part, une liste en Python est une collection de types de données hétérogènes stockés dans des emplacements de mémoire non contigus.

**to\_categorical** : c'est une méthode utilisée pour transformer les labels en format one hot and coding, elle sert à transformer les données d'entraînement avant de les transmettre à notre modèle. Si les données d'entraînement utilisent des classes comme nombres, `to_categorical` transformera ces nombres en vecteurs appropriés à utiliser avec des modèles.

**sklearn** : C'est une bibliothèque elle comporte divers algorithmes tels que la machine à vecteurs de support, les forêts aléatoires, aussi permet de tracer la matrice de confusion.

**MinMaxScaler** : se trouve dans la bibliothèque sklearn preprocessing pour la normalisation des images rendre les valeurs des images entre 0 et 1.

**keras.metrics** : c'est une fonction utilisée pour évaluer les performances de notre modèle.

# Bibliographie

- Adoum, A. & Daoussa, A. (2014). Segmentation des images médicales par les méthodes de détection de contour. Master's thesis, Centre Universitaire d'Ain Témouchent.
- Audebert, N. (2018). *Classification de données massives de télédétection*. PhD thesis, Université Bretagne Sud.
- Audebert, N., Boulch, A., Le Saux, B., & Lefèvre, S. (2018). Segmentation sémantique profonde par régression sur cartes de distances signées. In *Reconnaissance des Formes, Image, Apprentissage et Perception (RFIAP)*.
- Azouza, N. (2018). Segmentation par classification floue : application à l'imagerie par résonance magnétique. Master's thesis, Université Akli Moand Oulhadje, Bouira.
- Belhadjer, H. & Sarouer, B. (2018). *Classification des images avec les réseaux de neurones convolutionnels*. PhD thesis, Université Mouloud Mammeri.
- Bellahmer, H. (2020). Implémentation et évaluation d'un modèle d'apprentissage automatique pour l'estimation de la valeur marchande de propriétés immobilières proposé. Master's thesis, Université Mouloud Mammeri De Tizi-Ouzou.
- Bendaoud, M. (2017). *Développement de méthodes d'extraction de contours sur des images à niveaux de gris*. PhD thesis, Université Mohamed Boudiaf des Sciences et de la Technologie-Mohamed Boudiaf . . . .
- Benfriha, S. & Hamel, H. (2016). Segmentation d'image par coopération région-contours », mémoire master professionnel. Master's thesis, Université Kasdi Merbah, Ouargla.
- Benierbah, H. & Khamadja, M. (2017). Traitement d'images, master : Systèmes des télécommunications. Master's thesis, Université des frères Mentouri (Constantine 1).
- Bensalem, H. (2021). Anatomie de la trachée. Université Ferhat Abbas, Sétif.
- Bodo, M. (2017). *Étude de la réponse biomécanique du thorax soumis à des sollicitations dynamiques dans un contexte civil et militaire par la méthode des éléments finis*. PhD thesis, Université Bourgogne Franche-Comté.
- Bouabdallaoui, K. (2019). Détecter, identifier et localiser des oiseaux ravageurs des cultures. Master's thesis, Université de Franche-Comté.

- Boulkhiout, Y. (2010). *Architecture Multi-agents adaptative pour La détection et la localisation de structures : application à la segmentation d'images*. PhD thesis.
- Chebout, M. S. (2014). Evaluation de l'impact du refactoring aspect dans un contexte multi-agents.
- Couteaux, V. (2021). *Apprentissage profond pour la segmentation et la détection automatique en imagerie multi-modale : application à l'oncologie hépatique*. PhD thesis, Institut Polytechnique de Paris.
- Dahmane, K. (2020). *Analyse d'images par méthode de Deep Learning appliquée au contexte routier en conditions météorologiques dégradées*. PhD thesis, Université Clermont Auvergne.
- Djaballah, M. (2021). Système de prédiction de la consommation d'énergie basé deep learning. Master's thesis, Université de 8Mai 1945 Guelma.
- Dolz, J., Desrosiers, C., Wang, L., Yuan, J., Shen, D., & Ayed, I. B. (2020). Deep cnn ensembles and suggestive annotations for infant brain mri segmentation. *Computerized Medical Imaging and Graphics*, 79, 101660.
- Durand, P. (2019). *Réseaux de neurones convolutifs en médecine nucléaire : applications à la segmentation automatique des tumeurs gliales et à la correction d'atténuation en TEP/IRM*. PhD thesis, Université Paris Descartes.
- Ganaye, P. (2019). *A priori et apprentissage profond pour la segmentation en imagerie cérébrale*. PhD thesis, Université de Lyon.
- Guerroumi, N. (2019). *Segmentation automatique par apprentissage profond de la colonne vertébrale scoliotique à partir d'images de résonance magnétique*. PhD thesis, École de technologie supérieure.
- Guesmia, H. (2019). algorithme de segmentation par chan-vese pour les différentes modalités des images. Master's thesis, Université de M'sila Mohamed-Boudiaf, Msila.
- Ichi.pro. (2022). <https://ichi.pro>. Date de consultation : 29 /05/2022.
- Inguere, T. (2018). *Intégration des systèmes multi-agents aux systèmes embarqués pour la délégation de tâches*. PhD thesis, Université du Maine.
- kaggle1 (2022). <https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection>. Date de consultation : 29 /05/2022,.
- kaggle2. (2022). <https://www.kaggle.com/code/polomarco/chest-ct-segmentation-lung-heart-trachea-data>. Date de consultation : 29 /05/2022,.
- Kaidi, D. (2017). Classification non supervisée de pixels d'images couleur par analyse d'histogrammes tridimensionnels. Master's thesis, Université Mouloud Mammeri ,Tizi-Ouzou.
- Khairandish, M., Sharma, M., Jain, V., Chatterjee, J., & Jhanjhi, N. (2021). A hybrid cnn-svm threshold segmentation approach for tumor detection and classification of mri brain images. *IRBM*.

- Kifadji, A. (2020). Segmentation des images de drones par apprentissage profond ou deep learning. Master's thesis, Université Badji Mokhtar . Annaba.
- Koudri, F. & Mohammedi, M. (2021). Segmentation d'image irm cérébrale en utilisant l'intensité et l'information contextuelle par le champ aléatoire de markov. Master's thesis, Université de M'sila Mohamed-Boudiaf, Msila.
- Laguel, H. (2010). Déploiement sur une plateforme de visualisation d'un algorithme coopératif pour la segmentation d'images irm basé sur les systèmes multi-agents. *Projet de Fin d'Étude pour l'obtention du diplôme d'ingénieur d'état de l'université des Sciences et de la Technologie Houari Boumediene*.
- Larbi, N. (2018). Segmentation d'images avec le deep learning. Master's thesis, Université Mouloud Mammeri, Tizi-Ouzou.
- LeCun, Y., Bengio, Y., et al. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10), 1995.
- Mabrouk, S. (2013). Segmentation des images médical par contour actif. Master's thesis, Université des Sciences et de la Technologie Oran Mohamed-Boudiaf ,Oran.
- Majdoulayane, H. (2009). *Extraction de caractéristiques de texture pour la classification d'images satellites*. PhD thesis.
- Marir, T. & Silem, A. (2017). La programmation des systèmes multi-agents normatifs—une approche basée sur la plateforme.
- McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115–133.
- Mebrouki, M et Sekkil, H. (2021). Etude comparative entre les différentes architectures des réseaux de neurones convolutifs (cnns) pour la détection de la fatigue du conducteur. Master's thesis, Ecole Supérieure En Sciences Appliquées, Tlemcen.
- Med (2022). <https://www.med.upenn.edu/cbica/brats2020/data.html>. Date de consultation : 29 /05/2022,.
- Mekruksavanich, S. & Jitpattanakul, A. (2021). Biometric user identification based on human activity recognition using wearable sensors : An experiment using deep learning models. *Electronics*, 10(3), 308.
- Mirsky, Y., Mahler, T., Shelef, I., & Elovici, Y. (2019). {CT-GAN} : Malicious tampering of 3d medical imagery using deep learning. In *28th USENIX Security Symposium (USENIX Security 19)* (pp. 461–478).
- Mokri, M. (2017). Classification des images avec les réseaux de neurones convolutionnels. Master's thesis, Université Abou Bakr Belkaid ,Tlemcen.

- Moreau, B. (2018). *Modélisation statistique de la géométrie 3D de la cage thoracique à partir d'images médicales en vue de personnaliser un modèle numérique de corps humain pour la biomécanique du choc automobile : Statistical modeling of the 3D geometry of the rib cage from medical images to personalize a numerical human body model for the biomechanics of car crash*. PhD thesis, Montpellier.
- Moussa, R. (2011). *Ségmentation Multi-agents en Imagerie Biologique Et Médicale : Application Aux IRM 3D*. PhD thesis, Université Bordeaux I; Université Sciences et Technologies-Bordeaux I.
- Neupane, B., Horanont, T., & Aryal, J. (2021). Deep learning-based semantic segmentation of urban features in satellite images : A review and meta-analysis. *Remote Sensing*, 13(4), 808.
- Porquet, C., Settache, H., Ruan, S., & Revenu, M. (2003). Une plate-forme multi-agent pour la segmentation d'images. étude des stratégies de coopération contour-région. In *ORASIS'2003* (pp. 413–422).
- Salama, W. M. & Aly, M. H. (2021). Deep learning in mammography images segmentation and classification : Automated cnn approach. *Alexandria Engineering Journal*, 60(5), 4701–4709.
- Scherrer, B. (2008). *Segmentation des tissus et structures sur les IRM cérébrales : agents markoviens locaux coopératifs et formulation bayésienne*. PhD thesis, Institut National Polytechnique de Grenoble-INPG.
- Semchedine, M. (2007). *Système coopératif hybride de classification dans un SMA : application à la segmentation d'images IRM*. PhD thesis.
- Senani, S. (2019). *Réseaux de neurones convolutionnels pour la détection précoce de la rétinopathie diabétique*. PhD thesis, Université Mouloud Mammeri.
- Seo, S., Chen, D., Kim, K., Kang, K., Doo, D., Chae, M., & Park, H. (2022). Temporary traffic control device detection for road construction projects using deep learning application. In *Construction Research Congress (CRC)*, Arlington VA.
- Seynaeve, P. & Broos, J. (1995). The history of tomography. *Journal belge de radiologie*, 78(5), 284–288.
- Soltani, L. (2007). *Les Systèmes Multi-Agent pour le Contrôle de Production*. PhD thesis, Université de Batna 2.
- Tan, M. & Le, Q. (2019). Efficientnet : Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105–6114). : PMLR.
- wikipedia (2022). <https://fr.wikipedia.org/wiki/PyCharm>. Date de consultation : 29 /05/2022,.

# Résumé

Dans un système de traitement d'images, la segmentation est primordiale. A ce jour, il n'existe pas de méthode universelle de segmentation d'image. Toute technique n'est efficace que pour un type d'image donné, pour un type d'application donné, et dans un contexte informatique donné. Le but de ce travail est de réaliser la segmentation sémantique des images médicales à l'aide de l'apprentissage automatique issu de l'Intelligence Artificielle (IA). Il s'agit précisément d'implémenter plusieurs approches en tirant parti des complémentarités entre différents modèles de l'Apprentissage Profond (AP) et augmenter ainsi la robustesse du processus de segmentation. Tout l'enjeu de cette coopération se base sur la segmentation par classification de pixels afin d'extraire une information utile et pertinente. Nous avons proposé quatre approches de segmentation sémantique en utilisant différents modèles des architectures Unet et Resnet50. Une implémentation est alors mise en œuvre autour d'un système multi-agents (SMA), apportant une distribution d'expertise et une réduction en temps d'exécution sur les séquences d'images 3D.

**Mots clés :** Imagerie Médicale, segmentation, Apprentissage Automatique(AA), Apprentissage Profond (AP), Intelligence Artificiel (IA), Systèmes Multi-Agents(SMA).

## Abstract

In an image processing system, segmentation is primordial. Until now, there is no universal image segmentation method. Any technique is effective only for a given type of image, for a given type of application, and in a given context. The main objective of this work is to achieve the semantic segmentation of medical images using machine learning, a category Artificial Intelligence (AI) algorithm.

It is precisely about implementing several approaches by taking advantage of the complementarities between different models of Deep Learning (DL). This allows to increase the robustness of the segmentation process. The challenge of this cooperation is based on segmentation by pixel classification in order to extract useful and relevant information. We have proposed four semantic segmentation approaches using different models of the Unet and Resnet50 architectures.

An implementation is then done around a multi-agent system (MAS), providing a distribution of expertise and a reduction in execution time on 3D image sequences (slice). **Keywords :** Medical Imaging, segmentation, Machine Learning (ML), Artificial Intelligence (AI), Deep Learning (DL), Multi-Agent Systems (MAS).