

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**  
**MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE**  
**UNIVERSITE MOHAMED BOUDIAF - M'SILA**

**FACULTE DE TECHNOLOGIE**  
**DEPARTEMENT D'ELECTRONIQUE**  
N° :



**DOMAINE : SCIENCES ET TECHNOLOGIE**  
**FILIERE : ELECTRONIQUE**  
**OPTION : ELECTRONIQUE DES SYSTEMES**  
**EMARQUEES**

**Mémoire présenté pour l'obtention**  
**Du diplôme de Master Académique**  
**Intitulé**

**DEEP LEARNING POUR LA SEGMENTATION**  
**D'IMAGES**

**Présenté Par :**

Mr. BOUBAYA Semail.  
Mr. BERBIT Djamel.

**Soutenu devant le jury composé de :**

Mr.	Université de M'SILA	Président
Mr. LAHOUAOUI Lalaoui	Université de M'SILA	Encadreur
Mr.	Université de M'SILA	Examinatrice

**Promotion : Juin 2021**

## **Remerciements**

*Je remercie en premier lieu, Dieu tout puissant qui m'a donné la force, la volonté et surtout la patience pour accomplir ce travail*

*Je tiens essentiellement et particulièrement à exprimer ma sincère et profonde gratitude à l'égard de Monsieur **Dr. LAHOVAOVI Lalaoui**, d'avoir eu l'aménité d'accepter de diriger ce travail, qui s'est montré d'une générosité sans bornes sur tous les plans, scientifique et humain ; voué d'une indulgence, et d'une patience inouïes ; je lui serais toujours reconnaissant pour le constant support intellectuel et ses directives, ce fut à la fois une chance et un grand honneur d'être votre étudiant, veuillez trouver dans ce travail mon respect et ma reconnaissance intarissable.*

*Je remercie également Monsieur **Dr. BRJK Youcef** pour son soutien durant l'élaboration de mon mémoire ainsi que le président et les membres du jury qui nous font l'honneur d'accepter de juger mon travail. Sans oublier bien entendu tous les enseignants qui ont contribué par leur savoir et leurs encouragements le long de nos parcours.*

*Je tiens à remercier également les responsables et tout le personnel du département d'électronique de la faculté de technologie à l'université Mohamed Boudiaf –M'sila*

*Mes remerciements à tous mes amis ainsi qu'à toutes les personnes que j'ai connues, qui m'ont aidé, soutenu et encouragé.*

*Un grand merci à ma famille, qui m'aurez apporté la stabilité et les encouragements pour persévérer ; c'est grâce à votre présence que je sois toujours à la hauteur.*

**BERBIT Djamel**

### *Dédicace*

*Je dédie en premier lieu ce modeste travail*

*A mes chers ma mère qui est toujours présents et continuent de l'être pour faire mon bonheur.*

*A ma femme pour son soutien et encouragement*

*A mes frères, mes sœurs, mes neveux et mes nièces*

*A toute la famille BERBIT et BENSEFFA*

*A tous mes collègues de travail et tous les étudiants de la spécialité électronique des systèmes embarqués promotion (2021)*

*A toute personne ayant contribué à ce travail de près ou de loin.*

*BERBIT Djamel*

### *Dédicace*

*Je dédie en premier lieu ce modeste travail*

*A mes chers ma mère qui est toujours présents et continuent de l'être pour faire mon bonheur.*

*A ma femme pour son soutien et encouragement*

*A mes frères, mes sœurs, mes neveux et mes nièces*

*A toute la famille BOUBAYA*

*A tous mes collègues de travail et tous les étudiants de la spécialité électronique des systèmes embarqués promotion (2021)*

*A toute personne ayant contribué à ce travail de près ou de loin.*

**BOUBAYA Semail**

# Table des matières

Table des matières .....	IV
Liste des Figures.....	VII
Liste de Tableaux.....	IX
Abréviation.....	X
Introduction générale.....	1

## Chapitre I : Segmentation d'images

1.1. Introduction.....	3
1.2. Définition de la segmentation d'images.....	3
1.3. Différentes approches de segmentation.....	4
1.3.1. Approches contour .....	4
1.3.2. Approche région.....	7
1.3.2.1. Croissance de région (région GROWING) .....	7
1.3.2.2. Segmentation par fusion de régions (MERGE) .....	8
1.3.2.3. Segmentation par division de régions (SPLIT) .....	8
1.3.2.4. Segmentation par division-fusion (Split and Merge) .....	9
1.3.3. La segmentation basée sur la classification des pixels .....	9
1.3.3.1. Classification de pixels non supervisée .....	9
1.3.3.2. Classification de pixels supervisée .....	10
1.4. Quelques algorithmes de segmentation d'images .....	10
1.4.1. Algorithmes de classification de pixels non-supervisée .....	10
1.4.1.1. Algorithme des k-moyennes .....	10
1.4.1.2. Algorithme des C-moyennes floues (FCM) .....	11
1.4.1.3. Algorithme de Fisher .....	11
1.4.2. Algorithmes de classification de pixels supervisée .....	12
1.4.2.1. Algorithme des k-plus proches voisins .....	12
1.4.2.2. Algorithme de Bayes .....	13
1.4.2.3. Algorithme des Réseaux de Neurones Multi Couches .....	14
1.5. Segmentation sémantique d'images .....	14
1.6. Conclusion .....	15

## Chapitre II : DEEP LEARNING

1.	Introduction.....	16
2.	Définition de l'apprentissage profond (deep learning) .....	16
3.	Fonctionnement du deep learning.....	17
4.	Pour quoi le deep learning ? .....	17
4.1.	Domaines d'application de l'apprentissage profonde .....	18
4.2.	Architectures de réseaux de neurones profonds .....	19
4.2.1.	Les réseaux de neurones convolutifs .....	19
4.2.2.	L'entraînement d'un nouveau CNN est difficile .....	20
4.2.2.1.	La couche de convolution (CONV) .....	20
4.2.2.2.	Couche de pooling (POOL) .....	21
4.2.2.3.	Couche de correction (ReLU) .....	21
4.2.2.4.	Couche entièrement connectée(FC) .....	22
4.2.2.5.	Couche de perte (LOSS) .....	22
4.2.2.6.	Perceptron.....	22
4.2.2.7.	Choix des hyperparamètres.....	23
4.2.2.7.1.	Nombre de filtres .....	23
4.2.2.7.2.	Forme de filtres.....	23
4.2.2.7.3.	Forme de Max pooling.....	24
4.2.3.	Réseau de neurones récurrents .....	24
4.2.3.1.	C'est quoi un RNN? .....	24
4.3.	Avantages de CNNs.....	25
5.	Conclusion : .....	26

## Chapitre III : Description de l'imagerie médicale

1.	Introduction .....	27
2.	Quelques techniques d'imagerie médicale.....	27
2.1.	Les Images par Résonance Magnétique.....	27
2.2.	Les Images Ultrasonores : L'Echographie.....	28
2.3.	Image rayons X.....	29
2.4.	La radiographie.....	30
2.5.	L'imagerie nucléaire.....	31
2.6.	Tomodensitométrie (scanner) .....	32

2.7. L'imagerie thermique.....	32
2.8. L'imagerie satellitale.....	33
2.9. L'imagerie hyper spectrale.....	33
3. Conclusion.....	34

**Chapitre VI :**

1. Introduction : .....	35
2. Base de données .....	35
3. U-net .....	35
4. Logigramme de segmentation d'image IRM .....	37
5. Test .....	38
5.1. Affichage de l'image d'IRM .....	38
5.2. Configuration U-Net 3D .....	39
5.3. Mesures d'évaluation de la segmentation.....	43
5.3.1. Coefficient de similarité Dice .....	43
5.3.2. Résultat de calcul Coefficient de similarité (Dice) .....	44
6. Conclusion.....	45

<b>Conclusion Générale .....</b>	<b>46</b>
----------------------------------	-----------

**LES REFERENCES BIBLIOGRAPHIQUES**

**Résumé**

# Liste des figures

Figure 1.1: Segmentation d'une image couleur.....	3
Figure 1.2: Quelques modèles de contours.....	4
Figure 1.3: contour et ses dérivées. ....	5
Figure 1.4: Détection de contour par les différents filtres. ....	5
Figure 1.5: image originale (à gauche), contour détecté par LAPLACIEN.....	6
Figure1.6: Exemple d'application du filtre de CANNY.....	6
Figure 1.7: Exemple d'un SNAKE qui épouse le corps de l'araignée.....	7
Figure 1.8: Croissance progressive des régions. ....	7
Figure 1.9: Décompositions successives des blocs. ....	8
Figure 1.10: Agrégation itérative des blocs similaires au bloc 1. ....	9
Figure 1.11: (a) l'image originale, (b) l'image classée par K-means avec $K=3$ , (c) l'image classée par k-means avec $K=5$ .....	11
Figure 1.12: Représentation d'une image dans l'espace RGB.....	11
Figure 1.13: Classification pixellaire par l'algorithme des k - P P V.....	12
Figure 1.14: Classification pixellaire par l'algorithme de Bayes.....	13
Figure 1.15: Classification pixellaire par l'algorithme des SVM.....	13
Figure 1.16: Classification pixellaire par l'algorithme des RNMC.....	14
Figure 1.17: Image et étiquette des pixels.....	15
Figure 2.1: La relation entre l'intelligence artificielle, le ML et le deep learning.....	16
Figure 2.218: Le procédé du ML classique comparé à celui du Deep Learning.....	18
Figure 2.3: Les réseaux de neurones convolutifs. ....	20
Figure 2.4: Exemple d'une convolution 2D.....	20
Figure 2. 5: (à gauche) Average pooling même couleur, (à droite) Max pooling : chaque case correspond.....	21
Figure 2.6: Un réseau de neurones convolutif qui reçoit une image 2D comme entrée et qui est composé d'une couche convolutive, une fonction d'activation non linéaire, une couche MAX pooling et enfin un perceptron multi Couche.....	23
Figure2.7: Exemple de max pooling (2x2). ....	24
Figure 2.8: (à gauche) Un RNN. (à droite) Sa version déroulé Source : Nature.....	25
Figure 3.1: Image médicale IRM.....	27
Figure 3.2: Image d'échographie de grossesse.....	29

Figure 3.3: Image médicale rayons x.....	30
Figure 3.4: Image médicale radiographie.....	31
Figure 4. 1: L'architecture de la solution U-net.....	36
Figure 4. 2: Logigramme de segmentation d'image IRM.....	37
Figure 4. 3: a- image originale IRM.....	38
Figure 4. 4: b- image original 3D.....	38
Figure 4. 5: Volume d'image étiqueté.....	39
Figure 4. 6: a-U net 3D architecture.....	40
Figure 4. 7: b-U net 3D architecture.....	41
Figure 4. 8: Comparaison entre l'image de référence et l'image test après segmentation.....	42
Figure 4. 9: Volume étiqueté de test.....	42
Figure 4. 10: Volume étiqueté de référence.....	43
Figure 4. 11: Statistiques scores coefficient de similarité (Dice).....	45

## Liste de Tableaux

Le Tableau 4.1 : les valeurs de Dice des cinq volumes. ....	44
---	----

## **Abréviations**

**k-P P V** : k-plus proches voisins.

**AI** : Artificielle Intelligence

**ML** : Machine learning

**NTIC** : nouvelles technologies de l'information et de la communication

**CNN** : Convolutional Neural Network

**ReLU** : abréviation de Unités Rectifié linéaires

**MLP**: Multilayer perceptron

**RNN**: Recurrent Neural Network

**TDNN** : réseau neuronal à retard temporel

**GPU**: Graphics processing unit

**IRM**: Imagerie par résonance magnétique

**TAO**: Tomographe axial assisté par ordinateur

**BraTS**: BrainTumour Segmentation

**U-Net**: réseau de neurones à convolution développé pour la segmentation d'images biomédicales.

**3D** : trois dimension.

**Dice** : Coefficient de similarité

# Introduction générale

## **Introduction générale :**

Au cours des dernières décennies, l'imagerie médicale est devenue indispensable dans le diagnostic et la thérapie des maladies.

Avec l'enrichissement des bases d'images médicales, de nouvelles méthodes sont requises pour mieux traiter cet énorme volume de données.

Cependant, en raison des grandes variations et de la complexité des données d'imagerie médicale, il est généralement difficile de déduire des solutions analytiques ou des méthodes simples pour décrire et représenter des objets tels que des lésions et des anatomies dans des données.

Par conséquent, les tâches en imagerie médicale nécessitent un apprentissage à partir d'exemples, et cela constitue un des intérêts clés du domaine de l'apprentissage machine.

L'apprentissage machine (automatique) est devenu l'un des principaux outils pour l'analyse d'images médicales. Les techniques d'apprentissage automatique constituent des solutions pour développer des outils permettant d'aider les médecins à diagnostiquer, à prédire le risque d'atteinte de maladies et à les prévenir avant qu'il ne devienne trop tard. L'apprentissage profond (En anglais, Deep Learning) est un nouveau volet du domaine de l'apprentissage automatique qui englobe un large éventail d'architectures de réseaux élaborés pour exécuter plusieurs tâches.

La première utilisation de réseaux de neurones pour l'analyse d'images médicales remonte à plus de vingt ans, leur utilisation a augmenté de plusieurs ordres de grandeur au cours des ces dernières années. Des revues récentes ont mis en évidence l'application de l'apprentissage profond à un large éventail de tâches d'analyse d'images médicales (segmentation, classification, détection, enregistrement, reconstruction d'image, amélioration, etc.).

Le cerveau est organisé en plusieurs zones, qui gèrent chacune des activités spécifiques : le langage, l'équilibre du corps, les battements du cœur, la circulation du sang, la mémoire, etc. Une tumeur peut entraîner des troubles très différents selon la zone dans laquelle elle se développe.

On appelle tumeur du cerveau toutes les tumeurs qui se développent à l'intérieur du crâne. Elles peuvent se développer dans n'importe quelle zone du cerveau : les hémisphères, le cervelet, le tronc cérébral, l'hypophyse, etc.

Il existe des dizaines de tumeurs cérébrales différentes, que l'on distingue en fonction de trois caractéristiques :

- La localisation de la tumeur
- Le type de tumeur
- Le degré d'agressivité de la tumeur

La localisation de la tumeur est également un élément essentiel pour le choix des traitements. Une tumeur située à la surface du cerveau par exemple, est généralement plus facile à extraire qu'une tumeur située au centre du cerveau.

Selon leur type (les médecins parlent de type histologique), les tumeurs ne se comportent pas de la même manière. Certaines se développent plus vite que d'autres ou ont un risque accru de récurrence par exemple.

L'IRM est le premier examen à réaliser pour détecter une tumeur cérébrale. Pour des raisons pratiques, elle est souvent précédée ou complétée par un scanner (pour lequel on obtient souvent un rendez-vous plus rapidement).

Dans notre travail nous voulons appliquer une méthode de segmentation des types de tumeur de cerveau par l'apprentissage profond (Deep Learning) en utilisant un réseau de neurone conventionnel CNN (U net)

Le travail décrit dans ce mémoire est organisé en quatre chapitres suivis d'une conclusion générale, chacun représente une partie du travail de ce projet qu'elle soit théorique ou pratique (simulation sous MATLAB) :

Dans un premier chapitre, les notions théoriques nécessaires pour la segmentation des images et la compréhension du projet sont présentées. D'abord, plusieurs modèles de Deep learning sont expliqués dans le deuxième chapitre. Certains sont naïfs, d'autres raffinés, certains sont théoriques et d'autres empiriques. Puis dans le troisième chapitre nous avons présenté une description des images médicales. Enfin nous avons fini par exposer les résultats et le test de différentes images.

# Chapitre I

## 1.1. Introduction

Le traitement d'images est une discipline de l'informatique et des mathématiques appliquées qui étudie les images numériques et leurs transformations, dans le but d'améliorer leur qualité ou d'en extraire de l'information.

La segmentation d'images est une étape importante et primordiale dans le processus de traitement et d'analyse d'images qui a pour but de partitionner une image en régions homogènes et regrouper des pixels ayant des propriétés communes suivant des critères prédéfinis. La figure suivante représente un exemple de segmentation d'images.

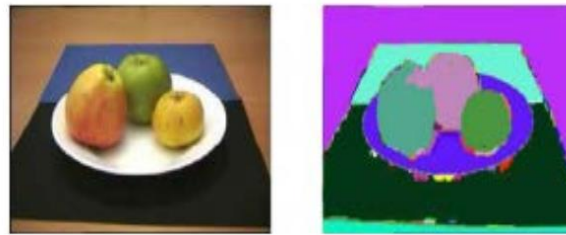


Figure 1.1: Segmentation d'une image couleur

Sur cette figure, on voit bien que chaque objet de l'image se voit attribuer une couleur, d'où la séparation en régions dites homogènes. Cependant on voit qu'il peut y avoir des défauts de reconnaissance et donc y avoir des confusions entre les régions comme c'est le cas ici entre la pomme du milieu et celle se situant à sa gauche car elles ont une zone en violet commune.

## 1.2. Définition de la segmentation d'images

La segmentation est une des étapes critiques de l'analyse d'images qui conditionne la qualité des mesures effectuées ultérieurement. C'est généralement une première étape d'un traitement plus complexe comme la reconnaissance de formes. Elle permet de cerner les formes des objets sur lesquels doit porter l'analyse. de délimiter des régions (l'intérêt et de les extraire du fond). Une bonne méthode de segmentation sera celle qui permettra d'arriver à une bonne interprétation.

Elle devra donc avoir simplifié l'image sans pour autant en avoir trop réduit le contenu.

La segmentation est une décomposition de l'image  $I$  en  $n$  régions  $R_i$  tel que :

1.  $\bigcup_i R_i = I$ .
2.  $R_i \cap R_j = \emptyset$ .
3.  $P(R_i) = \text{vrai}$ .
4.  $P(R_i \cap R_j) = \text{faux}$ .

En conclusion La segmentation consiste à :

- Regrouper les pixels de l'image qui partagent une même propriété pour former des régions homogènes.
- Répartir l'ensemble de pixels de l'image en différents groupes.
- Découper l'image en région. Une région est caractérisée par contours et par homogénéité (par exemple, même couleur).
- Partitionner une image en un ensemble de régions connexes et disjointes.
- La recherche de zones de l'image possédant des attributs communs, comme la luminosité, la couleur ou plus rarement la texture.

### 1.3. Différentes approches de segmentation :

Généralement, les méthodes de segmentation sont regroupées en trois approches chacune ayant des avantages et ses domaines d'application et elles sont parfois complémentaires, ces approches sont :

1. Segmentation basée sur les contours (en anglais : edge-based segmentation).
2. Segmentation basée sur les régions (en anglais : régions-based segmentation).
3. Segmentation en utilisant la classification.

#### 1.3.1. Approches contour :

L'approche contour consiste à identifier les changements entre les régions. En général, un élément de contours est un point de l'image appartenant à la frontière de deux ou plusieurs objets ayant des niveaux de gris différents.

La figure suivante montre quelques modèles de contours.

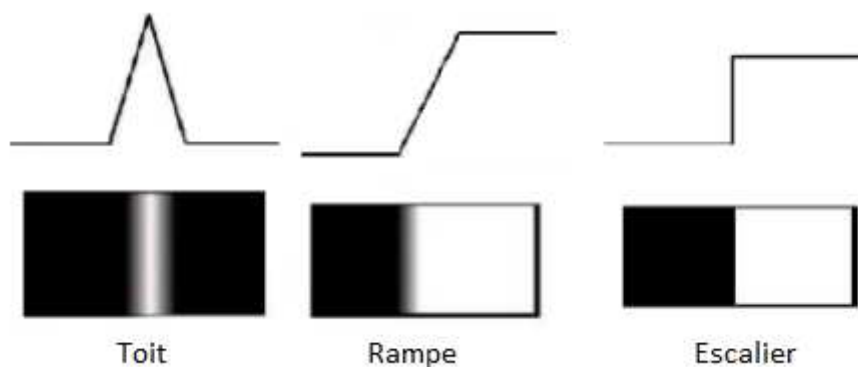


Figure 1.2 : Quelques modèles de contours

- Marche d'escalier : le contour est net (contour idéal).
- Rampe : le contour est plus flou.
- Toit : il s'agit d'une ligne sur un fond uniforme.

Plusieurs méthodes ont été adaptées pour la détection des contours, on distingue principalement les méthodes dérivatives, les méthodes analytiques et les contours déformable.

Les méthodes dérivatives sont les plus utilisées pour détecter des transitions d'intensité par différenciation numérique première ou deuxième dérivé (voir figure suivante).

A chaque position, un opérateur est appliqué afin de détecter les transitions significatives au niveau de l'attribut de discontinuité choisi. Le résultat est une image binaire constituée de points de contours et de points non-contours.

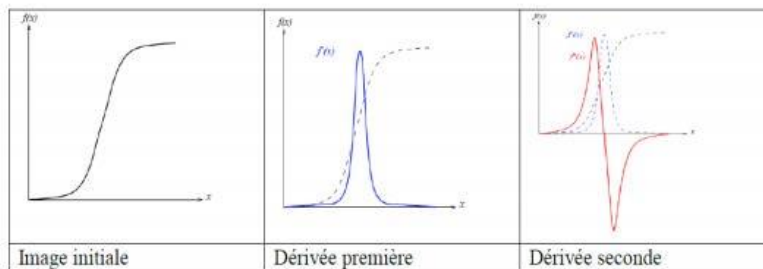


Figure 1.3 : contour et ses dérivées.

Il existe plusieurs opérateurs de gradient (première dérivée), parmi eux il ya les masques de ROBERT, PREWITT et SOBEL [1].

La figure ci-dessous fournie une image à laquelle sont appliqués ces operateurs.

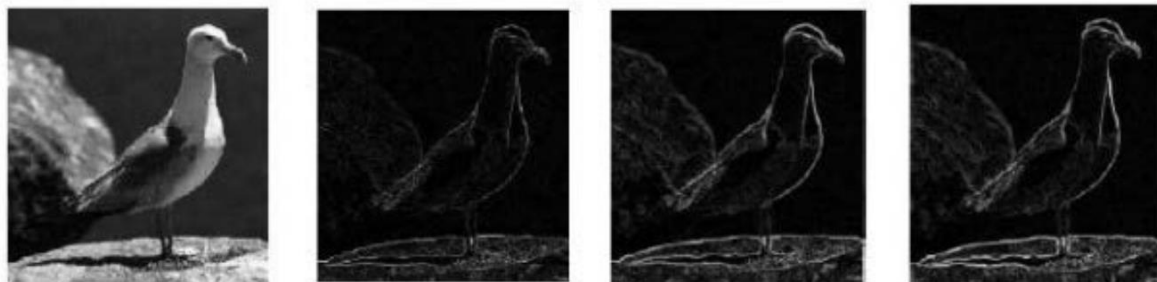


Figure 1.4 : Détection de contour par les différents filtres.

La figure suivante représente la détection de contour en utilisant l'opérateur LAPLACIEN qui est basé sur la seconde dérivée.



Figure 1.5 : image originale (à gauche), contour détecté par LAPLACIEN

Les Méthodes analytiques consistent à trouver un filtre optimal satisfaisant les trois contraintes suivantes :

- **Une bonne détection** : faible probabilité d'oublier un vrai point de contour et une faible probabilité de marquer un point image comme contour alors qu'il ne l'est pas.
- **Une bonne localisation** : les points contours doivent être le plus près possible de leur position réelle dans l'image.
- **Une réponse unique** : à un contour unique : un point de contour ne doit être détecté qu'une seule fois par le filtre mis en œuvre.

Le détecteur de contour de CANNY est le plus utilisé [2].

Au filtre de CANNY, DERICHE a proposé un autre filtre (condition initiale différente) qui permet une simplification de son implémentation.

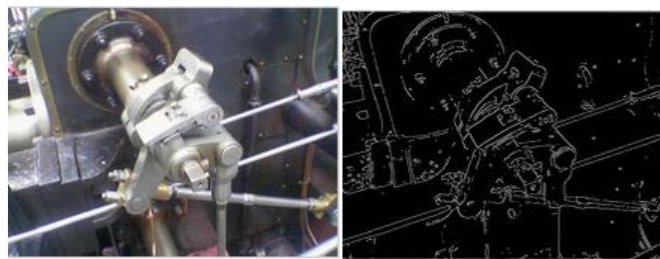


Figure1.6 : Exemple d'application du filtre de CANNY

Les modèles déformables, introduits par KASSONT sont aussi connus sous les noms de « SNAKES » ou « contours actifs ».

L'intérêt principal des contours actifs est de détecter des objets dans une image en utilisant les techniques d'évolution de courbes. L'idée est de partir d'une courbe initiale, généralement un carré ou un cercle, et de la déformer jusqu'à obtenir le contour de l'objet.

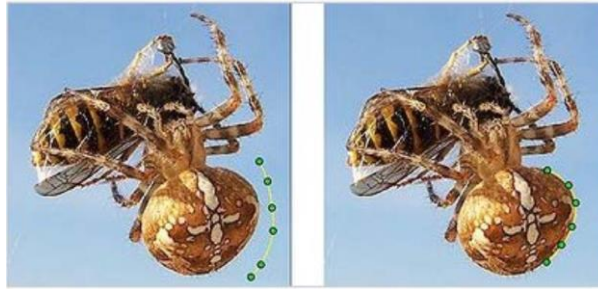


Figure 1.7 : Exemple d'un SNAKE qui épouse le corps de l'araignée

### 1.3.2. Approche région :

La segmentation d'image par l'approche région consiste à découper l'image en régions. Les pixels adjacents sont regroupés en régions distinctes selon un critère d'homogénéité ou de similarité donnée. Ce critère peut être, par exemple, le niveau de gris, couleur, texture...etc.

Un processus de groupement est répété jusqu'à ce que tous les pixels dans l'image soient inclus dans des régions. Cette approche vise, donc, à segmenter l'image en se basant sur des propriétés intrinsèques des régions.

Il existe plusieurs méthodes telles que la segmentation par croissance de région, par division de région, et par fusion de région que nous présentons ci-dessous.

#### 1.3.2.1. Croissance de région (région GROWING) :

Cette technique consiste à faire progressivement accroître les régions autour de leur point de départ.

Le principe de l'agrégation de pixel est le suivant : on choisit 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 itère le processus.

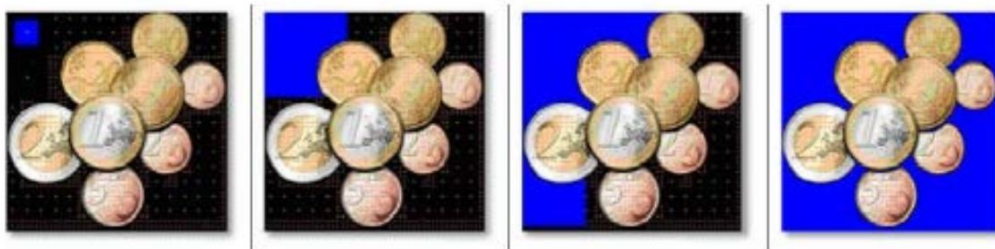


Figure 1.8 : Croissance progressive des régions.

Parmi les avantages de cette technique, nous pouvons citer :

- La simplicité et la rapidité de la méthode.
- La segmentation d'objet à topologie complexe.
- La préservation de la forme de chaque région de l'image.

Cependant, il existe plusieurs inconvénients comme :

- Une mauvaise sélection des germes ou un choix du critère de similarité mal adapté peuvent entraîner des phénomènes de sous-segmentation ou de sur-segmentation.
- Il peut y avoir des pixels qui ne peuvent pas être classés.

### 1.3.2.2. Segmentation par fusion de régions (MERGE) :

Les techniques de réunion (région merging) sont des méthodes ascendantes où tous les pixels sont visités. Pour chaque voisinage de pixel, un prédicat P est testé. S'il est vérifié les pixels

correspondants sont regroupés dans une région.

Les inconvénients de cette méthode se situent à deux niveaux :

- Cette méthode dépend du critère de fusion qui peut influencer sur le résultat final de la segmentation.
- Elle peut introduire l'effet de sous-segmentation.

### 1.3.2.3. Segmentation par division de régions (SPLIT) :

La division consiste à partitionner l'image en régions homogènes selon un critère donné. Le principe de cette technique est de considérer l'image elle-même comme région initiale, qui par la suite est divisée en régions. Le processus de division est réitéré sur chaque nouvelle région (issue de la division) jusqu'à l'obtention de classes homogènes

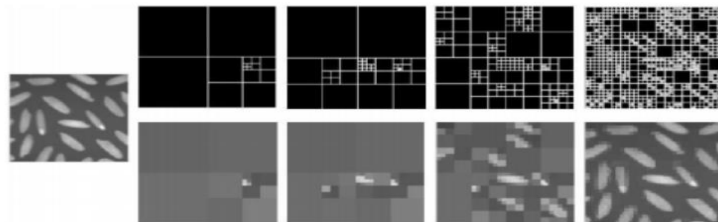


Figure 1.9 : Décompositions successives des blocs.

Cette méthode présente un inconvénient majeur qui est la sur-segmentation. Toutefois, ce problème peut être résolu en utilisant la méthode de division-fusion que nous présentons dans ce qui suit.

#### 1.3.2.4. Segmentation par division-fusion (Split and Merge) :

Ces méthodes combinent les deux méthodes décrites précédemment, 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.

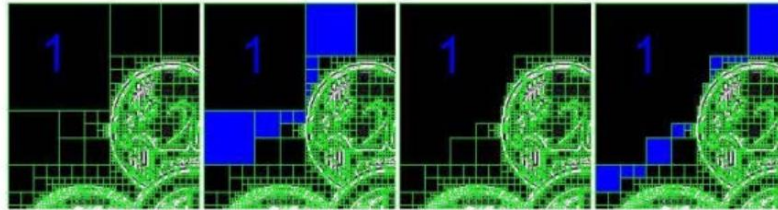


Figure 1.10 : Agrégation itérative des blocs similaires au bloc 1.

Les inconvénients de cette méthode se situent à trois niveaux :

- Les limites des régions obtenues sont habituellement imprécises et ne coïncident pas exactement aux limites des objets de l'image.
- La difficulté d'identifier les critères pour agréger les pixels ou pour fusionner et diviser les régions.

#### 1.3.3. La segmentation basée sur la classification des pixels :

De nombreux travaux sur la classification des pixels qui consiste à affecter à chaque pixel de l'image une classe qui définit les régions à extraire de l'image.

Un classifieur désigne tout outil de reconnaissance qui pour un vecteur reçu en entrée, donne des informations sur sa classe d'appartenance. Cet outil peut s'écrire sous la forme d'une fonction, qui à l'aide de descripteur d'un vecteur  $x$  à reconnaître, qui à l'aide de descripteur d'un vecteur  $x$  à reconnaître, attribue à  $x$  la classe  $C_i$  parmi  $k$  classe possible ( $i \in [1, k]$ ) Nous pouvons alors définir un classifieur par la relation suivante où l'ensemble  $K = (C_1, \dots, C_k)$

$$e : x \in \mathbb{R}^n \rightarrow K \quad (1.1)$$

La classification des pixels peut être supervisées ou non supervisée.

##### 1.3.3.1. Classification de pixels non supervisée :

La classification de pixels non supervisée appelée aussi classification de pixels sans apprentissage consiste à découper l'espace de représentation en zones homogènes selon un

---

critère de vraisemblance entre les individus. Cette approche est utilisée pour effectuer une classification de pixels en aveugle c'est-à-dire sans connaissance a priori sur l'image et ne nécessite donc pas de phase d'apprentissage.

### **1.3.3.2. Classification de pixels supervisée :**

La classification de pixels supervisée appelée aussi classification de pixels avec apprentissage consiste à définir une fonction de discrimination effectuant un découpage de l'espace de représentation à partir d'une connaissance a priori de l'image. Ce type de classification nécessite la création d'une base d'apprentissage faisant intervenir une segmentation de référence.

## **1.4. Quelques algorithmes de segmentation d'images :**

### **1.4.1. Algorithmes de classification de pixels non-supervisée :**

Nous allons à présent présenter trois des algorithmes de classification de pixels non-supervisée

à savoir :

1. L'algorithme des k-moyennes.
2. L'algorithme des C-moyennes floues.
3. L'algorithme de Fisher.

#### **1.4.1.1. Algorithme des k-moyennes :**

L'un des algorithmes les plus connus, pour la classification est l'algorithme K-means largement adopté en traitement d'images vu sa simplicité de mise en œuvre et sa capacité à fournir une bonne approximation de la segmentation recherchée. C'est un algorithme itératif qui minimise la somme des distances entre chaque pixel et le centre des classes. Ces centroïdes sont initialement placés le plus loin possible les uns des autres afin d'optimiser la qualité des résultats obtenus. Le principe de cet algorithme consiste à échanger des pixels entre deux classes jusqu'à ce que la somme des distances intra classes ne puisse plus diminuer. Le résultat idéal serait un ensemble de classes compacts et clairement séparés. Néanmoins cette méthode nécessite comme unique paramètre un nombre de classes K prédéfini a priori par l'utilisateur [3]. La figure 9 présente le résultat obtenu de l'application de l'algorithme K-means sur une image test où le nombre de classe K a été choisi arbitrairement. Sur cette image, nous remarquons qu'un mauvais choix de la valeur de K conduira à un résultat qui n'a pas de rapport avec l'image originale.



Figure 1.11 : (a) l'image originale, (b) l'image classée par K-means avec  $K=3$ , (c) l'image classée par k-means avec  $K=5$

Les résultats de classification établis par l'algorithme des k-moyennes peuvent fluctuer selon les paramètres d'entrés qui sont [4]

#### 1.4.1.2. Algorithme des C-moyennes floues (FCM) :

L'algorithme des C-moyennes floues (FUZZY C-MEANS) diffère peu de l'algorithme des k-moyennes si ce n'est qu'il est basé sur une classification floue. Il n'associe pas directement une classe  $C_i$  à un pixel  $x$  de l'image mais plutôt un degré d'appartenance à une classe (compris entre 0 et 1).

Il est basé sur l'optimisation d'un critère quadratique où la somme des écarts quadratiques intra-classes doit être minimale [4].

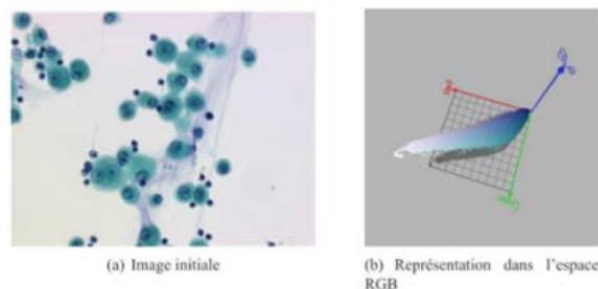


Figure 1.12 : Représentation d'une image dans l'espace RGB

Les deux paramètres d'entrées de cet algorithme devant être définis au préalable sont respectivement le paramètre  $m$  et le nombre de classes  $k$ .

#### 1.4.1.3. Algorithme de Fisher :

L'algorithme de Fisher consiste à effectuer une classification de pixels de l'image en  $k$  classes, en utilisant le partitionnement d'un histogramme de niveaux de gris en  $k$  classes disjointes tel que la somme des variances des classes soit minimale. Dans le cadre de la couleur, il convient d'appliquer cet algorithme séparément sur les trois composantes de l'espace de représentation couleur.

### 1.4.2. Algorithmes de classification de pixels supervisée :

Nous allons à présent présenter quatre algorithmes de classification de pixels supervisée à savoir :

1. Algorithme des k-plus proches voisins.
2. Algorithme de Bayes.
3. Algorithme des Machines à support de vecteurs.
4. Algorithme des Réseaux de Neurones Multi Couches.

#### 1.4.2.1. Algorithme des k-plus proches voisins :

La méthode des k-plus proches voisins (k-P P V) est une méthode d'estimation non paramétrique de densité [5,6].

Elle consiste à rechercher à partir d'une base d'apprentissage et d'une distance définie sur l'espace des données les K plus proches voisins d'un élément. Ce qui revient à calculer l'estimation de densité  $r_{ii}$  de la classe  $C_i$  au point  $x$  défini par la relation suivante :

$$r_i = \frac{k_i(x)}{n_i \times V(x)} \quad (1.2)$$

Où  $k_i(x)$  est le nombre de points de  $C_i$  appartenant aux k - P P V de  $x$ ,  $n_i$  est le cardinal de la classe  $C_i$  et  $V(x)$  le volume de la plus petite boule contenant le k - P P V de  $x$ .

La probabilité d'appartenance à une classe de cet élément  $x$  est alors proportionnelle au nombre d'exemples de cette classe parmi ses k - P P V.

La classe de l'élément  $x$  est alors celle ayant la plus grande probabilité. Les résultats dépendant fortement de la valeur de  $k$ , il convient de choisir une valeur minimisant le risque d'erreur déterminé à partir d'une base de tests

La figure suivante illustre une classification pixellaire effectuée par l'algorithme des k - P P V sur deux images d'une base de tests.

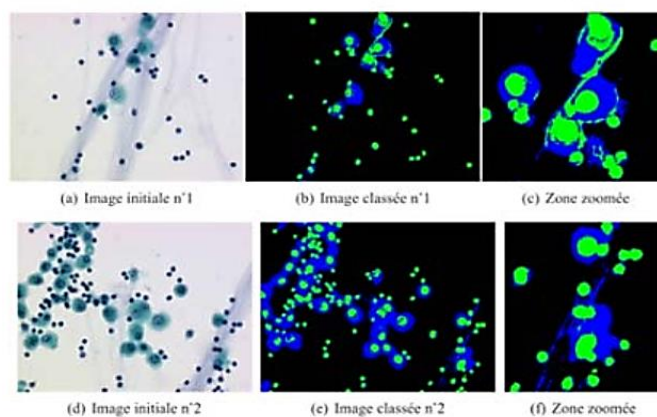


Figure 1.13 : Classification pixellaire par l'algorithme des k - P P V

### 1.4.2.2. Algorithme de Bayes :

Cet algorithme de classification est basé sur la théorie de décision Bayésienne [7]. La méthode utilisée est une approche statistique supposant que le problème de classification peut être exprimé en termes probabilistes.

L'hypothèse de base est généralement que la fonction de densité de probabilité d'un pixel  $x$  d'appartenir à une classe  $C_i$  est de forme gaussienne.

L'algorithme de Bayes cherche à déterminer pour chaque élément  $x$  la classe  $C_i$  qui maximise la probabilité de contenir cet élément.

$$f(x, C_i) = -\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) - \frac{1}{2} \log |\Sigma_i| - \log p_i + \frac{k}{2} \log 2\pi \quad (1.3)$$

Où  $k$  est le nombre de classes,  $\mu_i$  la moyenne des éléments de la classe  $C_i$ ,  $\Sigma_i$  la matrice de variance-covariance et  $p_i$  la probabilité a priori de la classe  $C_i$ .

La figure suivante illustre une classification pixellaire effectuée par l'algorithme de Bayes sur deux images d'une base de tests.

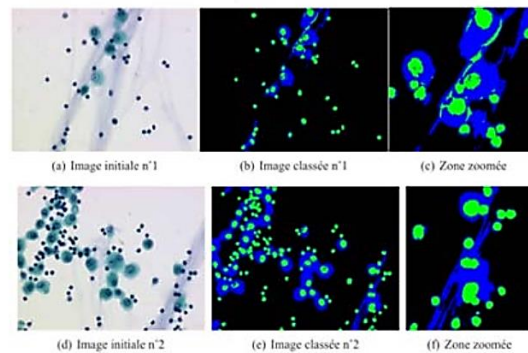


Figure 1.14 : Classification pixellaire par l'algorithme de Bayes

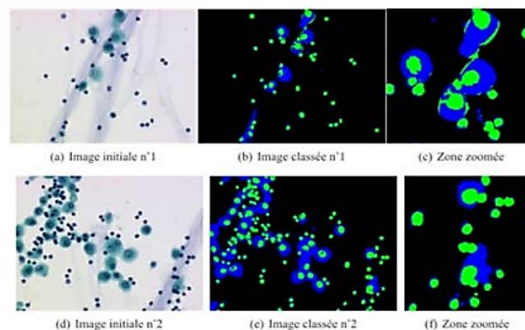


Figure 1.15 : Classification pixellaire par l'algorithme des SVM

### 1.4.2.3. Algorithme des Réseaux de Neurons Multi Couches :

Les réseaux de neurones multicouches (MLP : Multi Layer Perceptron) sont utilisés depuis de nombreuses années dans le domaine de la classification étant donné leurs bons résultats.

L'idée principale des MLP est de grouper des neurones par couche et de connecter complètement les neurones des couches adjacentes.

Typiquement, les couches sont organisées de la façon suivante :

- Une couche d'entrée (paramètres caractérisant un objet), une ou plusieurs couches cachées (augmentant les possibilités d'apprentissage), et une couche de sortie (fournissant la classe trouvée pour un objet)
- La phase d'apprentissage consiste à modifier les poids reliant les neurones de façon à ce que la classe en sortie corresponde à celle de l'objet présenté en entrée. Cette modification est effectuée par un algorithme de rétropropagation [8].

Afin d'obtenir une bonne généralisation, il reste deux paramètres à régler : la durée de l'apprentissage et le nombre de neurones cachés.

Ils sont choisis de façon à minimiser le risque d'erreur déterminé à partir d'une base de tests.

La figure suivante illustre une classification pixellaire effectuée par l'algorithme des Réseaux de Neurons Multi Couches sur deux images d'une base de tests.

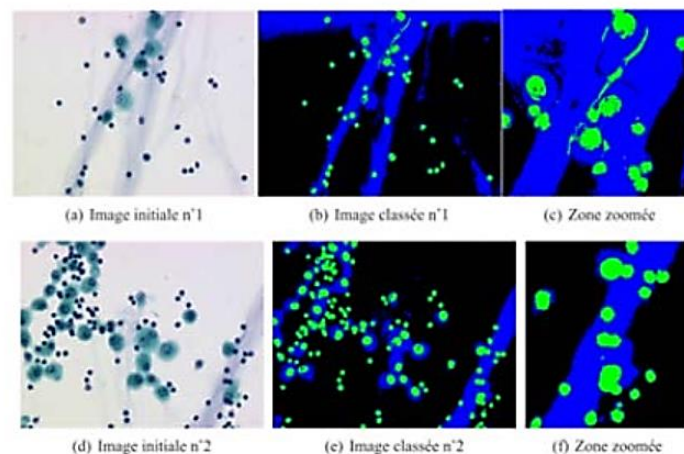


Figure 1.16 : Classification pixellaire par l'algorithme des RNMC

### 1.5. Segmentation sémantique d'images :

La segmentation sémantique associe une étiquette ou une catégorie à chaque pixel d'une image. Elle permet de reconnaître un ensemble 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 suivante, 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



Figure 1.17 Image et étiquette des pixels

La segmentation sémantique étiquette les pixels d'une image c'est ce qui la rend utile dans des applications de divers domaines :

**Conduite autonome** : pour identifier un parcours conduisez pour les véhicules en distinguant la route des obstacles tels que les piétons, trottoirs, poteaux et autres véhicules. **Contrôles industriels** : pour détecter les défauts dans des matériaux, comme le contrôle des composants électroniques.

**Imagerie satellite** : Pour identifier les montagnes, les rivières, les déserts et autres terrains.

**Imagerie médicale** : Pour analyser et détecter les anomalies cancéreuses dans les cellules.

**Vision robotique** : Pour identifier les objets et le terrain et s'y déplacer.

## 1.6. Conclusion

Dans ce chapitre, nous avons définis la segmentation d'images en étudiants développant les différentes approches. Dans le chapitre suivant nous allons détailler la technique de l'apprentissage profond (DEEP LEARNING) qui sera utilisée pour la segmentation d'images sémantique.

# Chapitre II

## 1. Introduction

L'intelligence artificielle est une discipline scientifique recherchant des méthodes de solution de problèmes à forte complexité logique ou algorithmique.

Le deep learning ou apprentissage profond est un type d'intelligence artificielle dérivé du machine learning (apprentissage automatique) où la machine est capable d'apprendre par elle-même, contrairement à la programmation où elle se contente d'exécuter à la lettre des règles prédéterminées. L'apprentissage automatique champ d'étude de l'intelligence artificielle. Par conséquent, L'apprentissage profond (en anglais deep learning, deep structured learning, hierarchical learning) est un ensemble de méthodes d'apprentissage automatique tentant de modéliser avec un haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non linéaires.

Dans ce chapitre nous allons présenter tout d'abord les notions en relation avec l'apprentissage profond.

Le Deep Learning est un nouveau domaine de recherche du ML, qui a été introduit dans le but de rapprocher le ML de son objectif principal : l'intelligence artificielle.

Il concerne les algorithmes inspirés par la structure et le fonctionnement du cerveau. Ils peuvent apprendre plusieurs niveaux de représentation dans le but de modéliser des relations complexes entre les données.

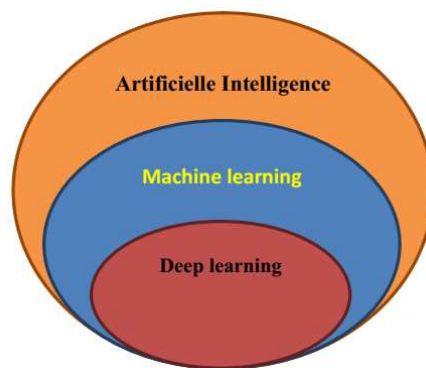


Figure 2.1 : La relation entre l'intelligence artificielle, le ML et le deep learning

## 2. Définition de l'apprentissage profond (deep learning)

L'apprentissage profond (deep learning) est un ensemble de techniques d'apprentissage automatique qui a permis des avancées importantes en intelligence artificielle dans les dernières années.

Dans l'apprentissage automatique, un programme analyse un ensemble de données afin de tirer des règles qui permettront de tirer des conclusions sur de nouvelles données. L'apprentissage

profond est basé sur ce qui a été appelé, par analogie, des « réseaux de neurones artificiels », composés de milliers d'unités (les « neurones ») qui effectuent chacune de petites opérations simples. Les résultats d'une première couche de « neurones » servent d'entrée aux calculs d'une deuxième couche et ainsi de suite

Par exemple, pour la reconnaissance visuelle, des premières couches d'unités identifient des lignes, des courbes, des angles... des couches supérieures identifient des formes, des combinaisons de formes, des objets, des contextes...

Les progrès de l'apprentissage profond ont été possibles notamment grâce à l'augmentation de la puissance des ordinateurs et au développement de grandes bases de données (« big data »).[9]

### 3. Fonctionnement du Deep Learning ?

Comme à l'intérieur du cerveau humain, les signaux voyagent entre les neurones du cerveau artificiel. Le secret de cette prouesse repose en grande partie sur les algorithmes. Dans le cas de la reconnaissance visuelle, pour être performant, l'algorithme du deep learning doit être capable d'identifier toutes les formes existantes et dans tous les angles.

Ainsi, il sera capable de détecter une voiture sur la route au milieu du paysage. Ceci n'est possible que si la machine a suivi un entraînement poussé. Et ceci passe par la visualisation de milliers de photographies sur lesquelles apparaissent une voiture, de toutes les formes et dans tous les angles possibles.

Lorsque l'image nouvelle apparaît, elle est envoyée au réseau de neurones qui se charge de les analyser et de déterminer si l'objet au milieu du cliché est bel et bien une voiture. La machine a-t-elle gagné son pari ? Elle garde sa bonne réponse au chaud, car elle l'aidera à résoudre d'autres situations similaires le jour où elle devra reconnaître une autre voiture.

### 4. Pourquoi le deep learning ?

Les algorithmes de ML décrits dans la première partie fonctionnent bien pour une grande variété de problèmes.

Cependant ils ont échoués à résoudre quelques problèmes majeurs de l'IA telle que la reconnaissance vocale et la reconnaissance d'objets.

Le développement du deep learning fut motivé en partie par l'échec des algorithmes traditionnels dans de telle tâche de l'IA.

Mais ce n'est qu'après que de plus grandes quantités de données ne soit disponibles grâce notamment au Big Data et aux objets connectés et que les machines de calcul soient devenues plus puissantes qu'on a pu comprendre le potentiel réel du Deep Learning.

Une des grandes différences entre le Deep Learning et les algorithmes de ML traditionnelles c'est qu'il s'adapte bien, plus la quantité de données fournies est grande plus les performances d'un algorithme de Deep Learning sont meilleures.

Contrairement à plusieurs algorithmes de ML classiques qui possèdent une borne supérieure à la quantité de données qu'ils peuvent recevoir des fois appelée "plateau de performance", les modèles de Deep Learning n'ont pas de telles limitations (théoriquement) et ils sont même allés jusqu'à dépasser la performance humaine dans des domaines comme l'image processing.

Autre différence entre les algorithmes de ML traditionnelles et les algorithmes de Deep Learning c'est l'étape de l'extraction de caractéristiques.

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

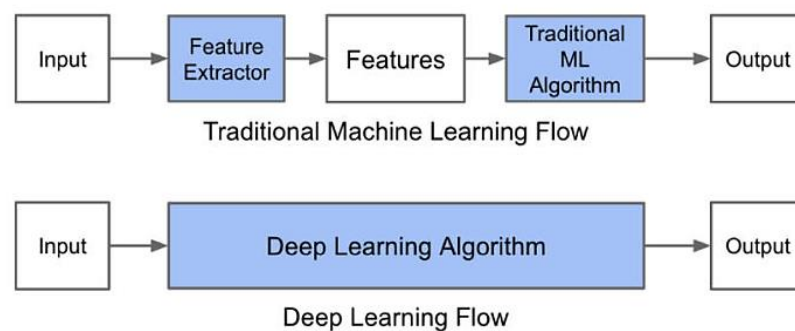


Figure 2.2 : Le procédé du ML classique comparé à celui du Deep Learning

#### 4.1. Domaines d'application de l'apprentissage profonde

Ces techniques se développent dans le domaine de l'informatique appliquée aux NTIC (reconnaissance visuelle par exemple d'un panneau de signalisation par un robot ou une voiture autonome et vocale notamment) à la robotique, à la bio-informatique la reconnaissance ou comparaison de formes, la sécurité, la santé, et c... , la pédagogie assistée par l'informatique, et plus généralement à l'intelligence artificielle.

L'apprentissage profond peut par exemple permettre à un ordinateur de mieux reconnaître des objets hautement déformables et/ou analyser par exemple les émotions révélées par un visage photographié ou filmé, ou analyser les mouvements et position des doigts d'une main, ce qui peut être utile pour traduire le langage des signes, améliorer le positionnement automatique d'une caméra, etc... Elles sont utilisées pour certaines formes d'aide au diagnostic médical (ex. : reconnaissance automatique d'un cancer en imagerie

médicale), ou de prospective ou de prédiction (ex. : prédiction des propriétés d'un sol filmé par un robot)

## 4.2. Architectures de réseaux de neurones profonds

Il existe un grand nombre de variantes d'architectures profondes. La plupart d'entre elles sont dérivées de certaines architectures. Il n'est pas toujours possible de comparer les performances de toutes les architectures, car elles ne sont pas toutes évaluées sur les mêmes ensembles de données.

Le Deep Learning est un domaine à croissance rapide, et de nouvelles architectures, variantes ou algorithmes apparaissent toutes les semaines [10].

### 4.2.1. Les réseaux de neurones convolutifs

Les réseaux de neurones convolutifs sont à ce jour les modèles les plus performants pour classer des images. Désignés par l'acronyme CNN, de l'anglais Convolutional Neural Network, ils comportent deux parties bien distinctes. En entrée, une image est fournie sous la forme d'une matrice de pixels. Elle a deux dimensions pour une image aux niveaux de gris.

La couleur est représentée par une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales [Rouge, Vert, Bleu].

La première partie d'un CNN est la partie convolutive à proprement parler. Elle fonctionne comme un extracteur de caractéristiques des images. Une image est passée à travers d'une succession de filtres, ou noyaux de convolution, créant de nouvelles images appelées cartes de convolutions.

Certains filtres intermédiaires réduisent la résolution de l'image par une opération de maximum local. En fin, les cartes de convolutions sont mises à plat et concaténées en un vecteur de caractéristiques, appelé code CNN.

Ce code CNN en sortie de la partie convolutive est ensuite branché en entrée d'une deuxième partie, constituée de couches entièrement connectées. Le rôle de cette partie est de combiner les caractéristiques du code CNN pour classer l'image.

La sortie est une dernière couche comportant un neurone par catégorie. Les valeurs numériques obtenues sont généralement normalisées entre 0 et 1, de somme 1, pour produire une distribution de probabilité sur les catégories.



Figure 2.3 : Les réseaux de neurones convolutifs.

#### 4.2.2. L'entraînement d'un nouveau CNN est difficile

Une architecture CNN est formée par un empilement de couches de traitement indépendant :

##### 4.2.2.1. La couche de convolution (CONV)

Trois hyper paramètres permettent de dimensionner le volume de la couche de convolution (aussi appelé volume de sortie) : la 'profondeur', le 'pas' et la 'marge'.

1. 'Profondeur' de la couche : nombre de noyaux de convolution (ou nombre de neurones associés à un même champ récepteur).
2. 'Le pas' contrôle le chevauchement des champs récepteurs. Plus le pas est petit, plus les champs récepteurs se chevauchent et plus le volume de sortie sera grand.
3. 'La marge (à 0)' ou 'zéro padding' : parfois, il est commode de mettre des zéros à la frontière du volume d'entrée. La taille de ce 'zero -padding' est le troisième hyperparamètre. Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée.

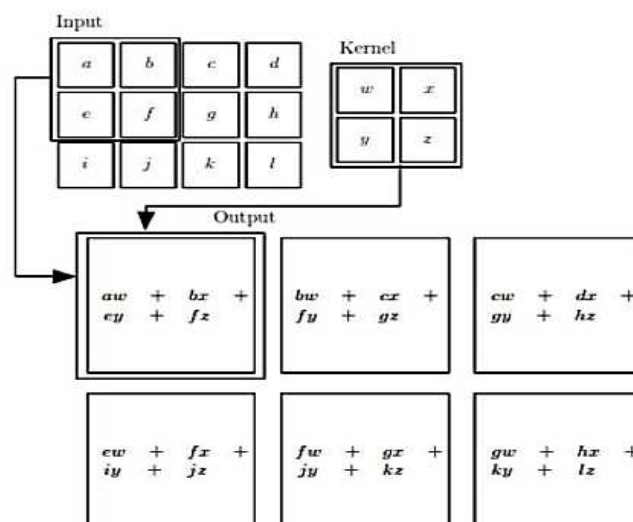


Figure 2.4 : Exemple d'une convolution 2D

### 4.2.2.2. Couche de pooling (POOL)

Une architecture atypique d'un réseau convolutif se compose de trois types de couches différentes. D'abord une couche convolutive pour générer un ensemble d'activations linéaires ensuite, on les fait passer à travers une couche d'activation non linéaire telle que Rectified Linear Unit (ReLU), enfin on utilise la fonction pooling.

Il permet de réduire progressivement la taille des représentations afin de réduire la quantité de paramètres et de calcul dans le réseau et, par conséquent de contrôler également le surapprentissage.[10]

Il permet l'invariance aux petites translations

Utile lorsque on préfère savoir si une caractéristique est présente plutôt que la région de sa présence.

Plusieurs type de pooling différent (MAX pooling (très populaire), AVG pooling, ...)

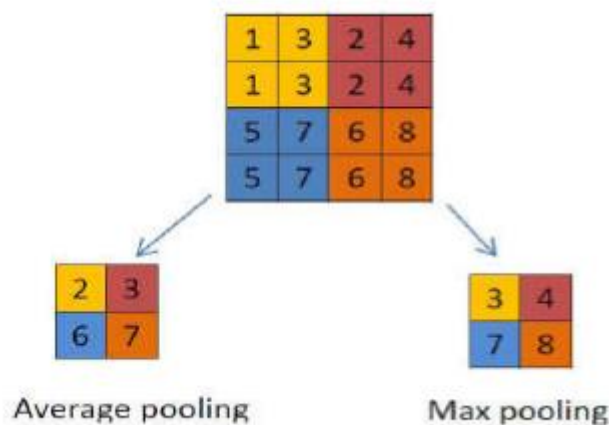


Figure 2. 5 : (à gauche) Average pooling même couleur, (à droite) Max pooling : chaque case correspond

### 4.2.2.3. Couche de correction (ReLU)

Souvent, 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. On a notamment :

- La correction ReLU (abréviation de Unités Rectifié linéaires) :  $f(x) = \max(0, x)$ . Cette fonction, appelée aussi « fonction d'activation non saturante », augmente les propriétés non linéaires de la fonction de décision et de l'ensemble du réseau sans affecter les champs récepteurs de la couche de convolution.
- La correction par tangente hyperbolique  $f(x) = \tanh(x)$ .

- La correction par la tangente hyperbolique saturante :  $f(x) = |\tanh(x)|$ .
- La correction par la fonction sigmoïde.

Souvent, la correction Relu est préférable, car il en résulte la formation de réseau neuronal plusieurs fois plus rapide, sans faire une différence significative à la généralisation de précision.

#### 4.2.2.4. Couche entièrement connectée (FC)

Après plusieurs couches de convolution et de max-pooling, le raisonnement de haut niveau dans le réseau neuronal se fait via des couches entièrement connectées. Les neurones dans une couche entièrement connectée ont des connexions vers toutes les sorties de la couche précédente (comme on le voit régulièrement dans les réseaux réguliers de neurones). Leurs fonctions d'activations peuvent donc être calculées avec une multiplication matricielle suivie d'un décalage de polarisation.

#### 4.2.2.5. Couche de perte (LOSS)

La couche de perte spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et réel. Elle est normalement la dernière couche dans le réseau. Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées. La perte « Soft max » est utilisée pour prédire une seule classe parmi K classes mutuellement exclusives. La perte par entropie croisée sigmoïde est utilisée pour prédire K valeurs de probabilité indépendante dans [0,1].

La perte euclidienne est utilisée pour régresser vers des valeurs réelles.[10]

#### 4.2.2.6. Perceptron

Après avoir extrait les caractéristiques des entrées, on attache à la fin du réseau un perceptron ou bien un MLP. Le perceptron prend comme entrée les caractéristiques extraites et produit un vecteur de N dimensions ou N est le nombre de classe ou chaque élément est la probabilité d'appartenance à une classe.

Chaque probabilité est calculée à l'aide de la fonction softmax dans le cas où les classes sont exclusivement mutuelles.

$$\text{Softmax}(z)_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (2.1)$$

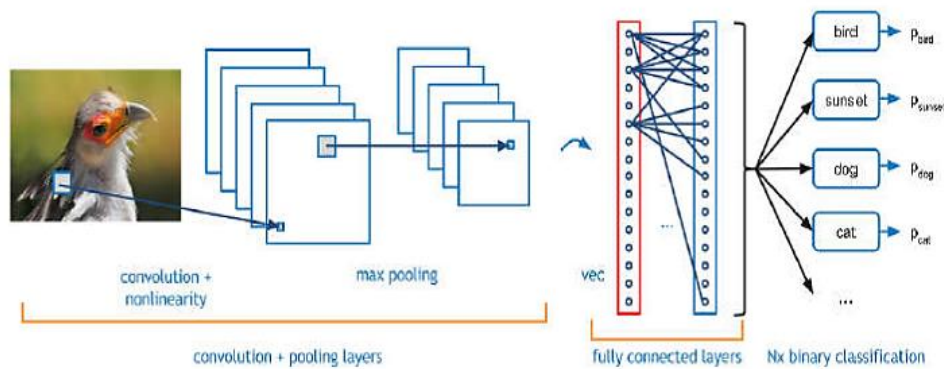


Figure 2.6 : Un réseau de neurones convolutif qui reçoit une image 2D comme entrée et qui est composé d'une couche convolutive, une fonction d'activation non linéaire, une couche MAX pooling et enfin un perceptron multi Couche

#### 4.2.2.7. Choix des hyperparamètres

Les CNNs utilisent plus d'hyperparamètres qu'un MLP standard. Même si les règles habituelles pour les taux d'apprentissage et des constantes de régularisation s'appliquent toujours, il faut prendre en considération les notions de nombre de filtres, leur forme et la forme du max pooling.

##### 4.2.2.7.1. Nombre de filtres

Comme la taille des images intermédiaires diminue avec la profondeur du traitement, les couches proches de l'entrée ont tendance à avoir moins de filtres tandis que les couches plus proches de la sortie peuvent en avoir davantage. Pour égaliser le calcul à chaque couche, le produit du nombre de caractéristiques et le nombre de pixels traités est généralement choisi pour être à peu près constant à travers les couches. Pour préserver l'information en entrée, il faudrait maintenir le nombre de sorties intermédiaires (nombre d'images intermédiaire multiplié par le nombre de positions de pixel) pour être croissante (au sens large) d'une couche à l'autre. Le nombre d'images intermédiaires contrôle directement la puissance du système, dépend du nombre d'exemples disponibles et la complexité du traitement.

##### 4.2.2.7.2. Forme de filtres

Les formes de filtre varient grandement dans la littérature. Ils sont généralement choisis en fonction de l'ensemble de données. Les meilleurs résultats sur les images de MNIST (28x28) sont habituellement dans la gamme de 5x5 sur la première couche, tandis que les ensembles de données d'images naturelles (souvent avec des centaines de pixels dans chaque dimension) ont

tendance à utiliser de plus grands filtres de première couche de 12x12, voire 15x15.

Le défi est donc de trouver le bon niveau de granularité de manière à créer des abstractions à l'échelle appropriée et adaptée à chaque cas.

#### 4.2.2.7.3. Forme de Max pooling

Les valeurs typiques sont 2x2. De très grands volumes d'entrée peuvent justifier un pooling 4x4 dans les premières couches. Cependant, le choix de formes plus grandes va considérablement réduire la dimension du signal, et peut entraîner la perte de trop d'information.

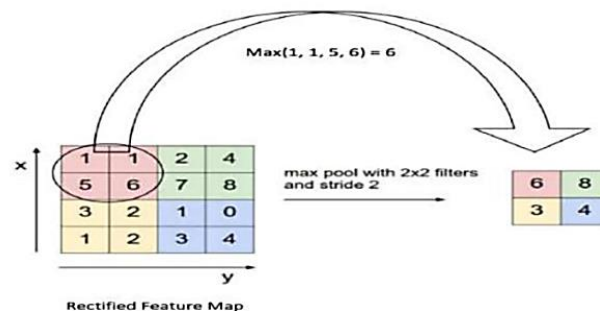


Figure 2.7 : Exemple de max pooling (2x2).

### 4.2.3. Réseau de neurones récurrents

Les humains ne commencent pas leurs pensées à zéro à chaque seconde. Lorsqu'on lit un livre, on comprend chaque mot en fonction de la compréhension des mots précédents. On n'oublie pas tout et recommençons à réfléchir à nouveau. Nos pensées ont une persistance. RNN (les réseaux de neurones récurrents) traitent ce problème. Ce sont des réseaux avec des boucles, permettant aux informations de persister.

#### 4.2.3.1. C'est quoi un RNN?

L'idée derrière les RNN (Récurrent Neural Network) est d'utiliser des informations séquentielles. Dans un réseau neuronal traditionnel, nous supposons que toutes les entrées (et les sorties) sont indépendantes les unes des autres. Mais pour de nombreuses tâches, c'est une très mauvaise idée. Si on veut prédire le prochain mot dans une phrase, il faut connaître les mots qui sont venus avant. Les RNN sont appelés récurrents car ils exécutent la même tâche pour chaque élément d'une séquence, la sortie étant dépendante des calculs précédents. Une autre façon de penser les RNN est qu'ils ont une « mémoire » qui capture l'information sur ce qui a été calculé jusqu'ici.

En théorie, les RNN peuvent utiliser des informations dans des séquences arbitrairement

longues, mais dans la pratique, on les limite à regarder seulement quelques étapes en arrière. Voici à quoi ressemble un RNN typique :

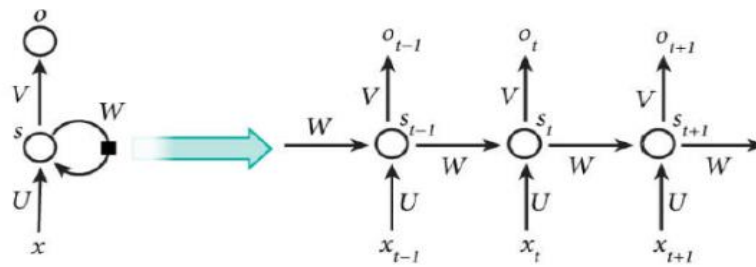


Figure 2.8 : (à gauche) Un RNN. (à droite) Sa version déroulée Source : Nature

Le schéma ci-dessus montre un RNN déroulé. En déroulant, nous signifions simplement qu'on montre le réseau pour la séquence complète. Par exemple, si la séquence qui nous intéresse est une phrase de 5 mots, le réseau serait déroulé en un réseau de neurones de 5 couches, une couche pour chaque mot. Les formules qui régissent les calculs dans un RNN sont les suivantes :

- $X_t$  est l'entrée au moment  $t$ .
- $U$  ;  $V$  ;  $W$  sont les paramètres que le réseau va apprendre des données de l'apprentissage.
- $S_t$  est l'état caché au moment  $t$ . C'est la « mémoire » du réseau.  $s_t$  est calculé en fonction de l'état

Caché précédent et de l'entrée à l'étape actuelle :

- $$S_t = f(Ux_t + Ws_{t-1}) \quad (2.2)$$

Où  $f$  est une fonction non linéaire telle que : ReLu ou Hyperbolic tangent (tanh).

- $O_t$  est la sortie au moment  $t$ . Par exemple, si on veut prédire le prochain mot dans une phrase, ce serait un vecteur de probabilités dans un vocabulaire.
- $$O_t = \text{softmax}(Vs_t) \quad (2.3)$$

### 4.3. Avantages de CNNs

Un avantage majeur des réseaux convolutifs est l'utilisation d'un poids unique associé aux signaux entrant dans tous les neurones d'un même noyau de convolution. Cette méthode réduit l'empreinte mémoire, améliore les performances et permet une invariance du traitement par translation. C'est le principal avantage du CNN par rapport au MLP, qui lui considère chaque neurone indépendant et donc affecte un poids différent à chaque signal entrant. Lorsque le volume d'entrée varie dans le temps (vidéo ou son), il devient intéressant de rajouter un paramètre de temporisation (delay) dans le paramétrage des neurones. On parlera dans ce cas

de réseau neuronal à retard temporel (TDNN).

Comparés à d'autres algorithmes de classification de l'image, les réseaux de neurones convolutifs utilisent relativement peu de pré-traitement. Cela signifie que le réseau est responsable de faire évoluer tout seul ses propres filtres (apprentissage sans supervision), ce qui n'est pas le cas d'autres algorithmes plus traditionnels. L'absence de paramétrage initial et d'intervention humaine est un atout majeur des CNN.

## **5. Conclusion :**

Dans ce chapitre on a présenté les notions importantes qui sont en relation avec l'apprentissage profond (définition, Architectures....etc).

Aussi qu'une vision générale sur l'apprentissage profond, tout en donnant en détail la méthode choisie dans notre travail de recherche qui est le CNNs.

Le prochain chapitre, traite les détails de la conception, ainsi que la méthode et les outils utilisés pour la réalisation de notre application.

# Chapitre III

## 1. Introduction

L'imagerie médicale est un ensemble de techniques consistant à mettre en image différentes régions ou différents organes de l'organisme. Il existe plusieurs types d'imageries médicales qui sont plus ou moins adaptées en fonction des zones à étudier.

On distingue notamment la radiologie, qui utilise les rayons X pour explorer les os, les poumons, l'abdomen ou le système digestif, l'échographie, qui se sert des ultrasons pour explorer les organes pleins et qui permet d'observer le fœtus au cours de la grossesse, l'imagerie par résonance magnétique (IRM) qui permet d'obtenir des images en coupes du corps, le scanner qui donne également des images en coupe et qui permet de réaliser la représentation d'organes en 3D.

## 2. QUELQUE TECHNIQUES D'IMAGERIE MEDICALE

### 2.1. L'imagerie par résonance magnétique

L'imagerie par résonance magnétique (IRM) est issue d'une série de découvertes scientifiques faites tout au long du 20<sup>ème</sup> siècle, qui exploite les propriétés magnétiques des tissus, basée sur les propriétés magnétiques des noyaux atomiques, généralement des protons, qui sont abondants dans les tissus biologiques riches en eau et en graisses. Sur les tissus mous, elle offre de très bons contrastes qui peuvent être modulés grâce à l'utilisation de séquences spécifiques.

L'imagerie par résonance magnétique offre une gamme variée d'examen permettant de caractériser la fibrose et les tumeurs hépatiques par différentes modalités. Un même examen permet alors de quantifier différents paramètres physiologiques. Les temps d'acquisition importants sont une des principales limitations de l'IRM, restreignant les possibilités de réaliser l'ensemble des examens IRM potentiellement utiles, bien que les évolutions méthodologiques et technologiques laissent espérer une marge de progression conséquente.



Figure 3. 1: Image médicale IRM [11]

L'imagerie par résonance magnétique (IRM) est une modalité d'imagerie qui s'est imposée aujourd'hui comme une technique de choix à la fois dans le cadre clinique et dans le cadre recherche. Les informations extraites des images IRM peuvent être anatomiques ou fonctionnelles comme en imagerie de diffusion ou de perfusion.

L'IRM est une imagerie sensible à des artéfacts dont les sources sont multiples. Parmi ces artéfacts, l'artéfact de susceptibilité magnétique est l'un de ceux qui pose le plus de problèmes.

L'origine de cet artéfact est la différence de susceptibilité magnétique entre deux milieux. L'artéfact de susceptibilité magnétique est présent dans les principales familles de séquences écho. [11]

## 2.2. Les Images Ultrasonores : L'Echographie

L'Echographie est une technique médicale consistant à visualiser certains organes à l'aide de sons à haute fréquence (les Ultrasons). Les sons réfléchis par les organes sont analysés par ordinateur de façon à produire une image sur un écran ou une photographie. Les sons sont émis par un cristal à oscillation rapide dont la fréquence se situe entre 18 et 20 KHz.

Ces vibrations du sonde, en contact étroit avec la peau, pour émettre les sons et recevoir les échos. La peau est également enduite de gel pour améliorer l'acoustique.

L'air, les os et les tissus calcifiés absorbent la quasi-totalité des faisceaux d'ultrasons, Cet examen n'est donc pas utile pour diagnostiquer les atteintes osseuses ou pulmonaires.

En revanche, les fluides sont de bons conducteurs d'ultrasons, si bien que cette technique est utilisée pour examiner les kystes et des organes tels la vessie, le système biliaire ou visualiser le fœtus dans le sac amniotique.

L'Echographie peut également servir aux examens du réseau artériel, du cœur du pancréas, de la cavité péritonéale, de l'appareil urinaire, des ovaires, du système veineux, du cerveau et de la moelle épinière.

L'examen du fœtus pendant la grossesse est pourtant son utilisation la plus courante.

L'échocardiographie est l'application de la technique des ultrasons à l'examen du cœur. Elle est utilisée pour étudier les maladies cardiaques congénitales, les affections coronariennes, les tumeurs cardiaques et d'autres troubles du cœur.

L'échographie est également utilisée pour guider les interventions chirurgicales comme l'amniocentèse ou au cours des biopsies délicates.

A la différence des rayons X, l'échographie n'est pas contre-indiquée pendant la grossesse puisqu'elle ne présente aucun risque, ne pour la mère, ni pour l'enfant.

Elle permet de suivre la croissance et le développement du fœtus, de s'assurer de sa bonne santé et de préciser le terme du bébé car il est possible de déterminer l'âge exact du fœtus en mesurant son tour de tête. [12]



Figure 3. 2 : Image d'échographie de grossesse

### 2.3. Image rayons X

Les rayons X sont une forme de rayonnement électromagnétique, au même titre que la lumière visible, l'ultra-violet, l'infra-rouge, les micro-ondes, les ondes radio ou les rayons gamma. Les premières versions de tubes à rayons X voient le jour au début du XX<sup>ème</sup> siècle et, depuis maintenant un siècle, le principe physique régissant leur fonctionnement reste le même. Les matériaux et la technologie ont, par contre, largement évolué et permettent des régimes de fonctionnement plus soutenus, ainsi que de meilleurs rendements de conversions énergétiques.

Peuvent être produits de deux manières différentes. La première manière est celle liée aux changements d'orbite des électrons.

Des rayons X sont produits par des transitions électroniques qui font intervenir les couches internes proches du noyau.

Ces transitions peuvent être données lors d'un rayonnement incident de rayons X vers l'atome ou bien par un bombardement d'électrons, fait qui provoquera l'excitation de l'atome (éjection des électrons).

Si l'électron éjecté est proche du cœur, un électron de la périphérie va occuper sa position (processus de désexcitation) en émettant un photon qui va appartenir au domaine X. La deuxième façon de produire des rayons X est basée sous le principe d'accélération (freinage et changement de trajectoire) d'électrons.

Un tube à rayons X sera alors nécessaire. Le principe est le suivant : Lorsqu'on applique une haute tension électrique (de l'ordre de 20 à 400 kV) entre deux électrodes, un courant d'électrons, de la cathode vers l'anode (appelée aussi anticathode ou cible), est produit. Lorsque le faisceau d'électrons avance vers la cible, ils sont freinés par les atomes de celle-ci, en

provoquant un rayonnement continu de freinage ou de Bremsstrahlung (phénomène décrit plus bas), dont une partie du spectre est dans le domaine des rayons X.

En même temps, les rayons résultant provoquent l'excitation des atomes de la cible qui, telle que décrit au début de cette section, réémettent un rayonnement X grâce au phénomène de la fluorescence X.

Le tube donnera un spectre résultant de la superposition du rayonnement de freinage et de la fluorescence X de la cible [11]



Figure 3. 3 : Image médicale rayons X

#### 2.4. La radiographie

La radiographie est la formation, sur un film photographique, d'une image d'un objet qui a été interposé entre le film et une source de Rayon X. Le radiogramme est l'image obtenue de cette manière. Le terme radiologie (terme issu du latin radius : Rayon et du grec logos : discours) est interne plus global qui désigne l'utilisation des Rayons X dans un but diagnostic radiographique ou thérapeutique (cancer).

Le terme de radiographie est maintenant utilisé dans un sens plus large d'imagerie médicale (remplaçant de ce fait la radiologie diagnostique).

Elle est notamment utilisée en orthopédie, en rhumatologie et en orthodontie ou elle permet d'étudier les traumatismes osseux (fractures, etc.), les déformations du squelette ou les implantations dentaires.

La pneumologie y a aussi recours (radio de poumons).

Chez la femme, la radiographie du sein (mammographie) est devenue un examen systématique de prévention du cancer du sien.

Il est aussi possible de visualiser certains organes ou parties creuses, habituellement invisibles aux rayons X, en les « remplissant » d'un produit de contraste, opaque aux rayons X : c'est la radiographie de contraste. [13]



Figure 3. 4 : Image médicale radiographie

### 2.5. L'imagerie nucléaire

L'invention de la médecine nucléaire s'est reposée sur la découverte des isotopes radioactifs par Frédérique et Irène Joliot-Curie en 1934 et sur le développement de la première gamma-caméra en 1957 par Hal Anger[4]. La médecine nucléaire regroupe plusieurs techniques comme la scintigraphie : on injecte au patient un produit radioactif, porté par une molécule ayant des propriétés d'association aux tissus biologiques. En effet, ces isotopes s'accumulent dans certains organes précis que le médecin cherche à étudier. Ils émettent alors des rayonnements, gamma le plus souvent (mais d'autres techniques peuvent être utilisées), détectés par une gamma-caméra à scintillation.

Les radio-isotopes sont des éléments non-stables subissant des désintégrations ou des désexcitations pour arriver à un état stable. Ils sont donc des isotopes radioactifs. Le plus utilisé en médecine est le technétium de faible demi-vie.

Il provient de la désintégration du molybdène, un métal de transition. Une fois injecté au patient, le technétium se désexcite en technétium stable et émet des rayons gammas d'énergie 140 Kev -idéale puisque assez énergétique pour traverser les tissus vivants, et donc pouvant être captés par la gamma-caméra-. Tout d'abord, les photons passent par un collimateur à trous parallèles qui permet leur sélection. En effet, seuls les photons frappant le collimateur de façon perpendiculaire à sa surface sont utiles pour la formation de l'image. Sans collimateur, il est impossible de déterminer l'origine des photons et ainsi de reconstruire une image. Ensuite, le cristal scintillant convertit les photons gammas ou béta en lumière visible. Enfin des photomultiplicateurs augmentent le nombre de photons et convertissent le signal optique en signal numérique.

L'électronique et l'informatique calculent la position et l'énergie des photons gammas et reconstruisent numériquement l'image. On peut améliorer la résolution de l'image par une

collimation énergétique : seuls les photons ayant l'énergie caractéristique du radio-traceur sont retenus pour la formation de l'image. [14]

## **2.6. Tomodensitométrie (scanner)**

Le scanner appelé tomographe axial assisté par ordinateur (TAO), est un dispositif de radiographie associant rayon X, Est une image des plans de coupe d'un objet, en particulier du corps humain. Inventé par le Britannique Godfrey Hounsfield dans les années 1970.

Le scanner est une technique permettant d'établir une succession d'images «de coupe » de la zone scannée.

Ces images sont le résultat d'un traitement et d'une reconstruction informatique de l'information reçue par les capteurs de l'appareil.

Le scanner permet de visualiser des modifications de volume ainsi que des anomalies de structure au niveau des tissus ou des organes. Cette technique est utilisée dans divers domaines d'application :

- La cancérologie pour la détection de nodules, le contrôle de la réponse à la chimiothérapie et le bilan d'extension des cancers (pulmonaires, bronchiques, pelviens, ORL),
- La chirurgie pour localiser précisément la zone où l'on envisage l'intervention,
- La traumatologie pour l'étude de traumatismes et de fractures.

Le scanner aujourd'hui continue à évoluer pour tenter d'aller plus loin dans le diagnostic médical. Afin d'avancer vers cet objectif, cette modalité a du s à ouvrir à la présentation en trois démontions.

Cette nouvelle visualisation apporte beaucoup de bénéfices en faveur du diagnostic médical. Mais cette nouvelle technologie n'est possible qu'avec l'utilisation de consoles de post-traitement. Ces consoles sont des ordinateurs utilisant des algorithmes spécifiques afin de pouvoir reconstruire les images scannées en trois dimensions. Elles ont aujourd'hui, un rôle de plus en plus important et sont à la base de l'exploitation des données acquises par le scanner.[13]

## **2.7. L'imagerie thermique**

L'imagerie thermique est largement utilisée dans le domaine militaire pour ses capacités de vision diurne et nocturne et sa longue portée d'observation. Cette technologie est basée sur la détection passive dans l'infrarouge. En conditions météorologiques dégradées ou quand la cible est partiellement dissimulée par du feuillage ou des filets de camouflages militaires, elle devrait être à court terme de plus en plus complémentée par un système d'imagerie active. Cette technologie est essentielle pour l'imagerie à longue portée.

La technique d'imagerie dite flash 2D est basée sur une source laser impulsionnel qui illumine la scène et sur une caméra rapide synchronisée qui constitue le système d'imagerie.

Ces deux technologies sont bien éprouvées en présence de conditions météorologiques claires. Les modèles TRM4 (imagerie thermique) et PERFIMA (imagerie active) sont capables de prédire correctement les performances de tels systèmes par beau temps. En revanche, en conditions dégradées telle que la pluie, le brouillard ou la neige, ces modèles deviennent non pertinents.[15].

## **2.8. L'imagerie satellitaire**

L'imagerie satellitaire est incontestablement la technique de télédétection qui représente le plus d'avenir eu égard aux progrès constants, tant en matière de lanceurs qu'en matière de capteurs. En cartographie topographique, les images provenant des satellites d'observation de la terre viennent heureusement compléter dans ce domaine les photographies aériennes, mais pour le moment leur pouvoir de résolution limité les empêche de les remplacer totalement. Il est vrai que dans un avenir assez proche les progrès de la technique combleront cette différence et risquent de reléguer quelque peu la photo aérienne au second plan. En tout état de cause, les images satellitaires deviennent de plus en plus incontournables en cartographie topographique, en particulier dans les régions difficilement accessibles, même pour des survols en avions, mais également dans les autres régions, où elle fournissent des données complémentaires par rapport aux photos aériennes.

L'utilisation de l'imagerie satellitaire rend le processus d'élaboration de la carte rapide, précis et économique Rapide, parce que un satellite de télédétection est capable d'établir la carte d'un vaste massif ou d'une ville entière dans le temps nécessaire à une équipe de techniciens pour décharger ses équipements sur le site, ou à un pilote pour effectuer sa visite pré-vol. Précis : la camera et le détecteur ne peuvent mentir.

Dans la mesure où aucune intervention humaine n'est requise pour créer une image satellitaire brute, les informations contenues constituent une représentation précise, objective et impartiale des objets et caractéristiques de la surface de la terre. [16]

## **2.9. L'imagerie hyper spectrale**

L'imagerie hyper spectrale est une technique combinant l'imagerie et la spectroscopie où chaque image est prise pour une bande étroite du spectre électromagnétique. Par exemple, l'œil humain voit la lumière en trois bandes (rouge, vert et bleu) tandis que l'imagerie hyper-spectrale permet de 'voir' en un très grand nombre de bandes allant typiquement du visible au proche infrarouge.

Le terme hyper-spectral fait référence au nombre de bandes qu'il est possible de prendre à travers le spectre électromagnétique. Par comparaison, l'imagerie multi spectrale fait l'acquisition de bandes discrètes du spectre tandis que l'imagerie hyper-spectrale fait l'acquisition de bandes continues de moins de 10 nm.

Grâce à notre technologie unique, nous sommes en mesure d'obtenir des largeurs de bandes aussi faibles que 2-4nm et même 0,3nm. Après que les mesures aient été traitées, il est possible d'avoir l'information spatiale et spectrale de chaque pixel de l'image et ainsi identifier une molécule, un matériau ou un tissu grâce à sa signature spectrale unique.

Les images hyper spectrales sont présentées sous forme d'un cube hyper-spectral contenant toute l'information spatiale et spectrale d'un échantillon.

Les dimensions dans ces cubes sont celle de l'espace (axes x et y) et la troisième est la dimension spectrale (z), correspondant à l'intensité lumineuse en fonction de la longueur d'onde.[17]

### **3. Conclusion**

L'imagerie médicale est certainement l'un des domaines de la médecine qui a le plus progressé ces vingt dernières années. Il regroupe l'ensemble des techniques permettant de visualiser une partie des corps humains.

Dans ce chapitre nous avons présenté une étude sur l'imagerie médicale de façon générale, ses caractéristiques, les différents types existants (image résonance magnétique, rayon x, imagerie nucléaire,...etc.) De ce fait, dans le prochain chapitre, une description générale et détaillée sur les différentes méthodes de segmentation d'images.

# Chapitre VI

## 1. Introduction :

L'imagerie médicale permet d'analyser les tissus par des médias extrêmement divers, leur exploitation et leur interprétation sert à établir plus finement un diagnostic médical. Actuellement, la segmentation constitue une étape très importante dans le traitement et l'interprétation des images médicales, de ce fait une multitude de logiciels assurent cette quête de manière plus performante grâce à l'avancement de la science. Dans ce chapitre nous allons présenter les logiciels les plus connus et les plus utilisés de nos jours.

## 2. Base de données :

Dans notre conception on utilise l'ensemble de données **BraTS**. L'ensemble de données **BraTS** contient des IRM des tumeurs cérébrales. La taille du fichier de données est d'environ 7 Go, Pour télécharger les données **BraTS**, L'ensemble de données peut être téléchargé de site web **Medical Segmentation Decathlon "Task01\_BrainTumour.tar"**, L'ensemble de données contient 750 corps en quatre dimensions, dont chacun représente un ensemble d'images en trois dimensions. La taille de chaque volume en quatre dimensions est de  $240*240*155*4$ , et les trois premières dimensions correspondent à la hauteur, la largeur et la profondeur de l'image de volume en trois dimensions. La quatrième dimension correspond aux différentes formes de numérisation. L'ensemble de données est divisé en 484 volumes d'images d'apprentissage avec des étiquettes de voxels et 266 volumes d'images de test.

Le volume d'image de test n'a pas d'étiquettes, donc cet exemple n'utilise pas ces données de test. Au lieu de cela, l'exemple divise 484 volumes d'images d'apprentissage en trois ensembles de données indépendants pour l'apprentissage, la vérification et les tests.

## 3. U-net :

U-Net est un réseau de neurones à convolution développé pour la segmentation d'images biomédicales au département d'informatique de l'université de Fribourg en Allemagne. Son architecture est basée sur le réseau entièrement convolutionnel, elle a été modifiée et étendue pour fonctionner avec moins d'images d'entraînement et pour permettre une segmentation plus précise. La segmentation d'une image  $512 \times 512$  prend moins d'une seconde sur un GPU récent.

Le réseau se compose d'une partie contractante et une voie expansive, ce qui lui confère une architecture en forme de « U ». La partie contractante est un réseau de convolution typique qui consiste en une application répétée de convolutions, chacune suivie d'une unité linéaire

rectifiée (ReLU) et d'une opération de pooling maximum. Pendant la contraction, les informations spatiales sont réduites tandis que les informations sur les caractéristiques sont augmentées. La voie expansive combine les informations de caractéristiques géographiques et spatiales à travers une séquence de convolutions et concaténations ascendantes avec des fonctionnalités de haute résolution issues de la voie contractante, donc l'architecture U-Net est séparée en trois parties :

1. La voie de contraction / sous-échantillonnage
2. goulot d'étranglement
3. La voie d'extension / suréchantillonnage.

## Network Architecture

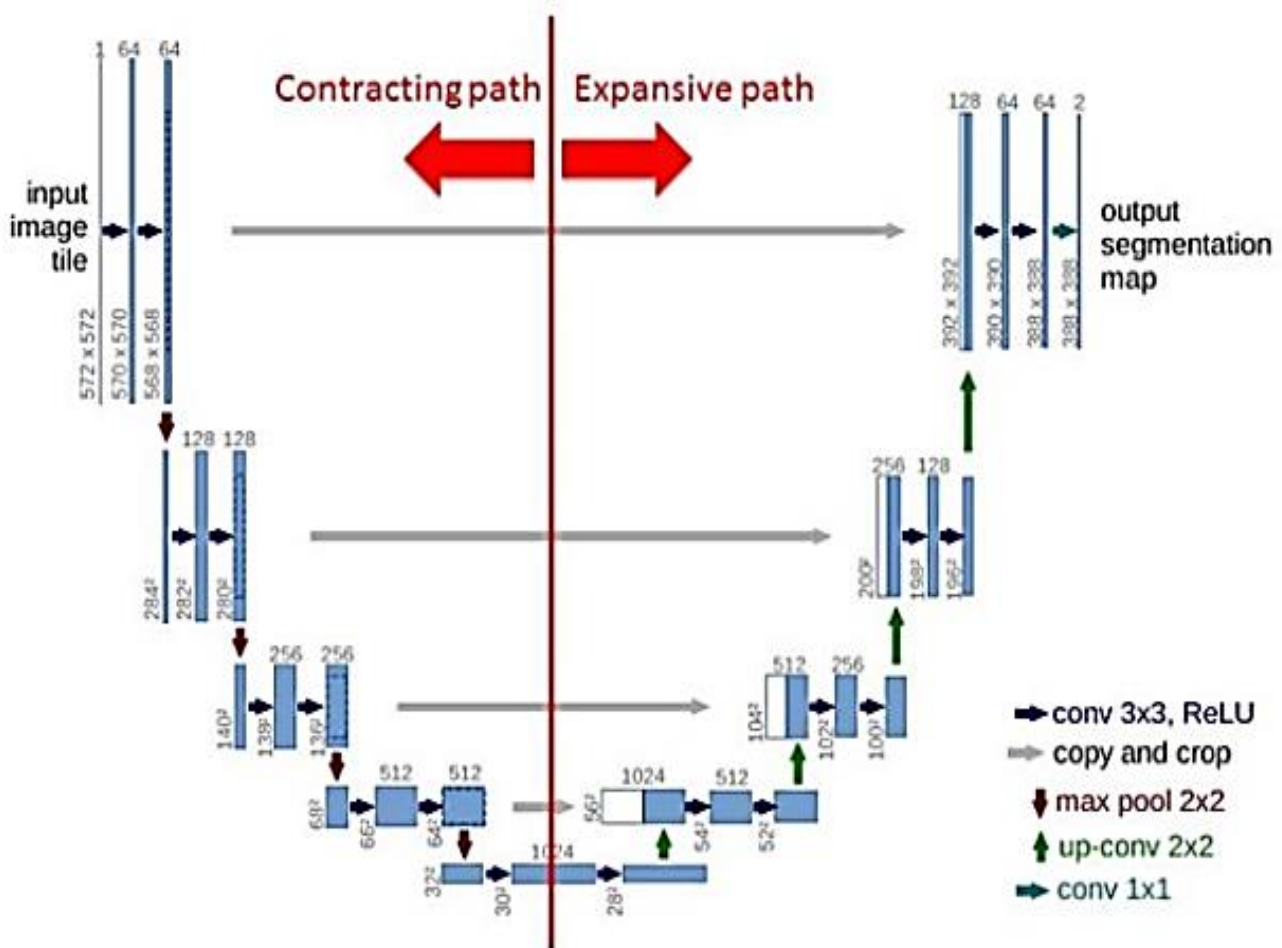


Figure 4. 1: L'architecture de la solution U-net

#### 4. Logigramme de segmentation d'image IRM :

Nous avons présenté les différentes phases de système implémenté comme suite :

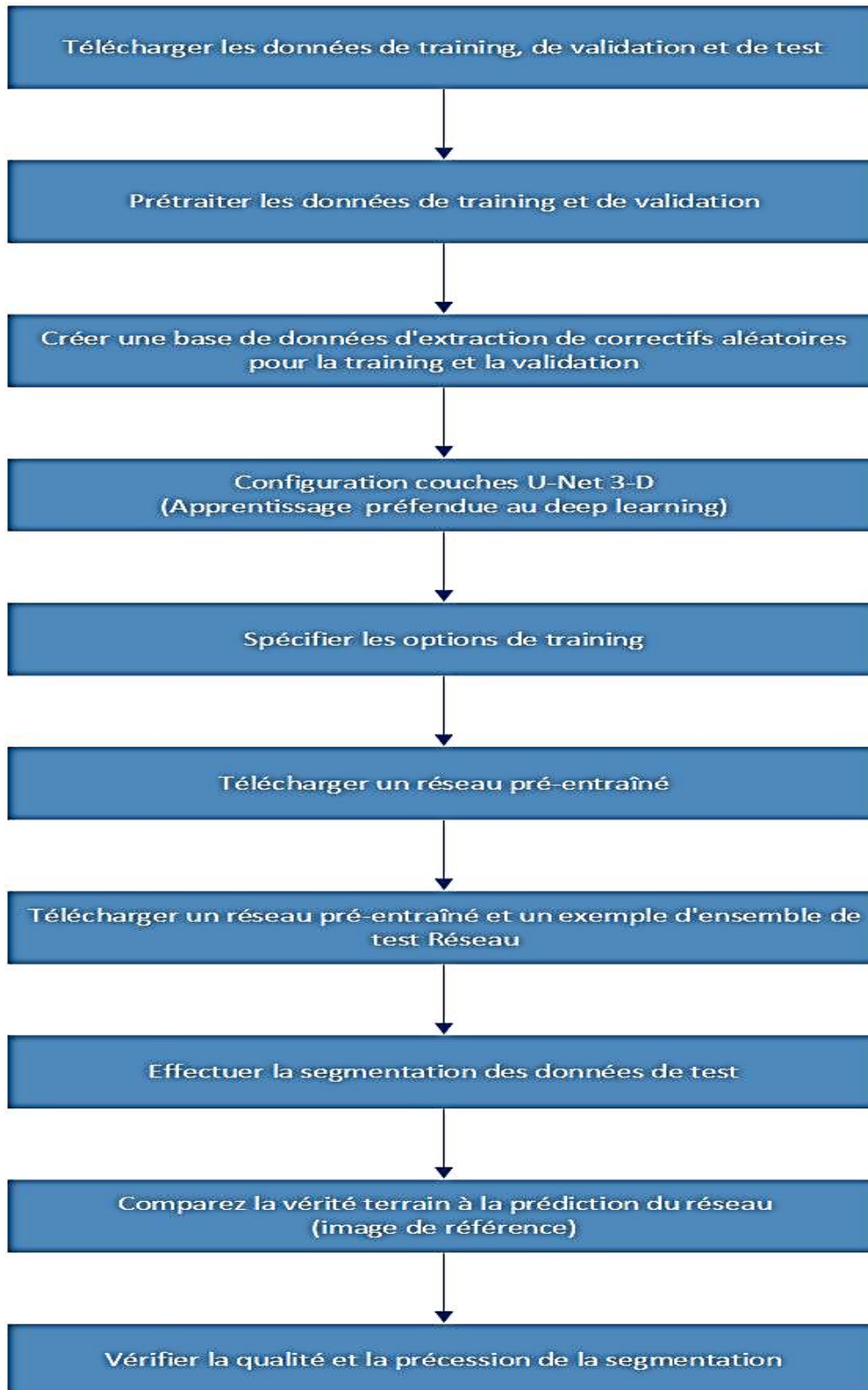


Figure 4. 2: Logigramme de segmentation d'image IRM

## 5. Test

Pour le test, nous avons choisi cinq volumes d'image qui font parties de la base de données .

### 5.1. Affichage de l'image d'IRM

Dans la figure ci-dessous montre l'image originale qui on va traiter avec la première image (a) en vue le cerveau avec la tumeur dans trois coté (frontal, horizontal et vertical) alors dans la deuxième image (b) en voir l'image en 3D.



Figure 4. 3: a- image originale IRM

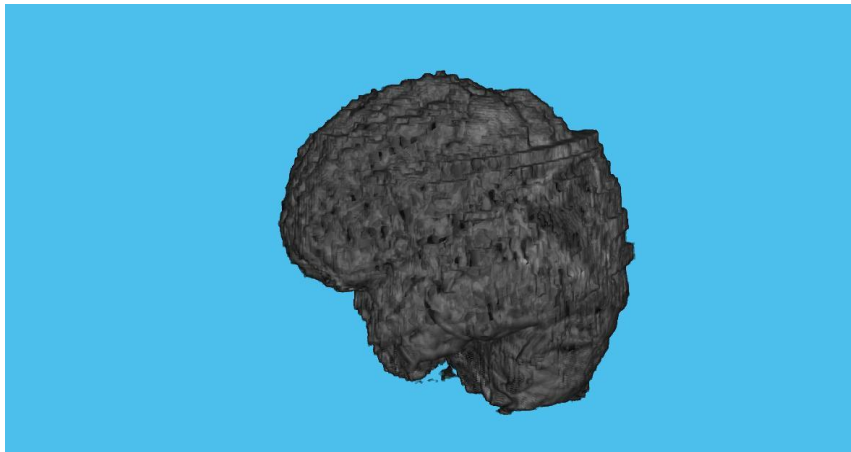


Figure 4. 4: b- image original 3D

Pour visualiser un volume d'image étiqueté. On met l'arrière-plan complètement transparent en définissant la visibilité de l'étiquette d'arrière-plan.

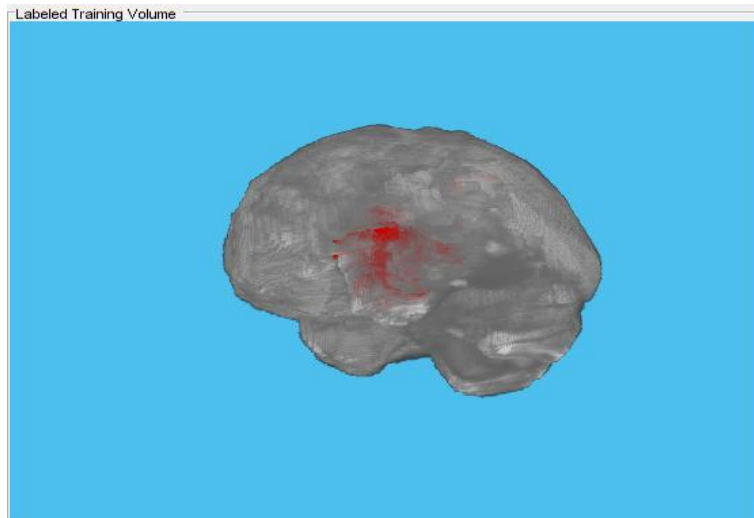


Figure 4. 5: Volume d'image étiqueté

## 5.2. Configuration U-Net 3D

Dans ce programme on utilise un le réseau 3-D U-Net. Dans le réseau U-Net, la série initiale de couches convolutionnelles est entrecoupée de couches de pooling max, diminuant successivement la résolution de l'image d'entrée.

Ces couches sont suivies d'une série de couches convolutionnelles entrecoupées d'opérateurs de suréchantillonnage, augmentant successivement la résolution de l'image d'entrée.

Une couche de normalisation par lots est introduite avant chaque couche ReLU. Le nom U-Net vient du fait que le réseau peut être dessiné avec une forme symétrique comme la lettre U.

Téléchargez une version pré-entraînée de 3-D U-Net et cinq échantillons de volumes de test et leurs étiquettes correspondantes à partir de l'ensemble de données BraTS. Le modèle pré-entraîné et les échantillons de données vous permettent d'effectuer une segmentation sur des données de test sans avoir à télécharger l'ensemble des données ou à attendre que le réseau s'entraîne.

Pour effectuer la segmentation des données de test un GPU est fortement recommandé pour effectuer la segmentation sémantique des volumes d'images

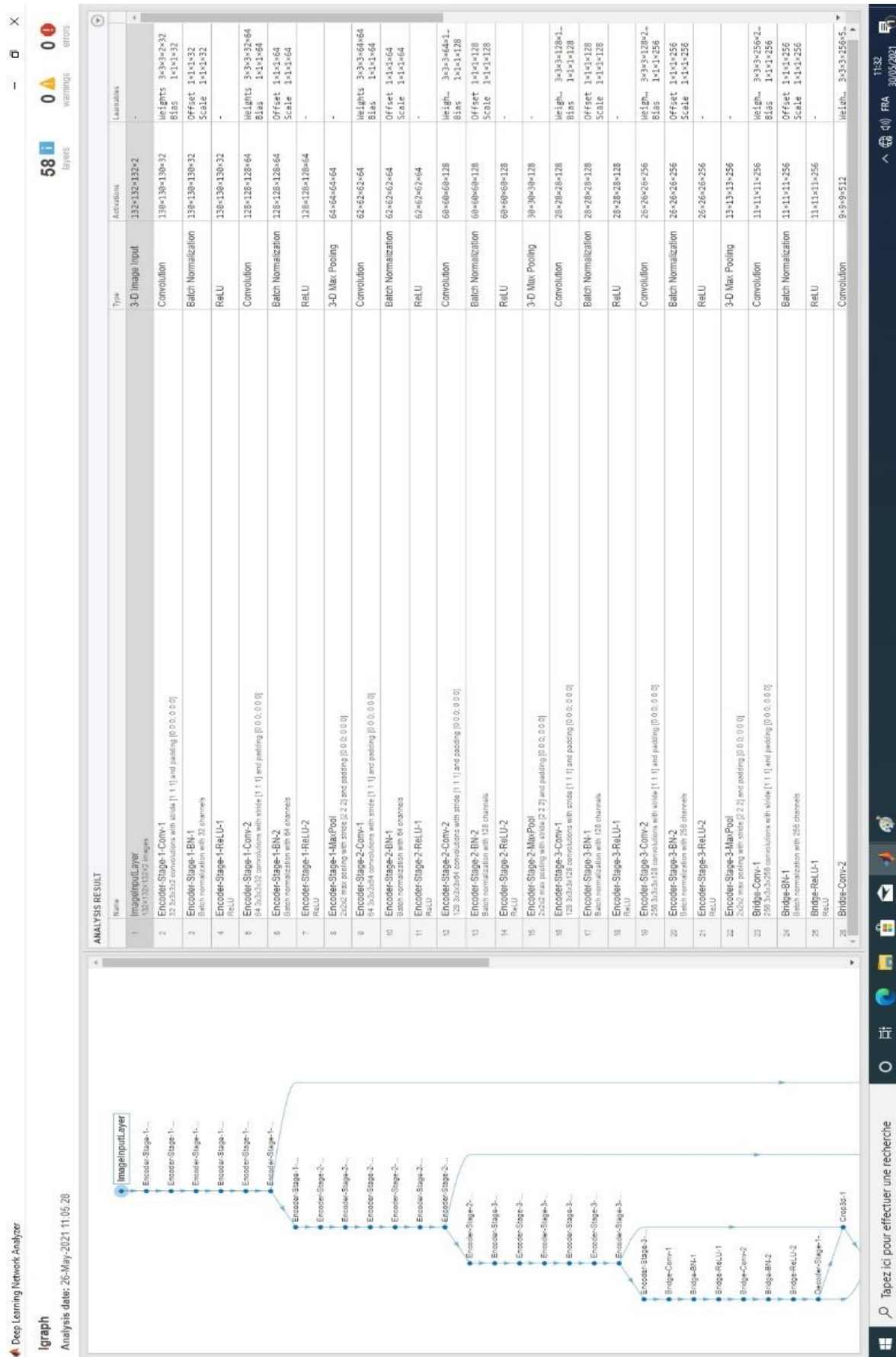


Figure 4. 6: a-U net 3D architecture

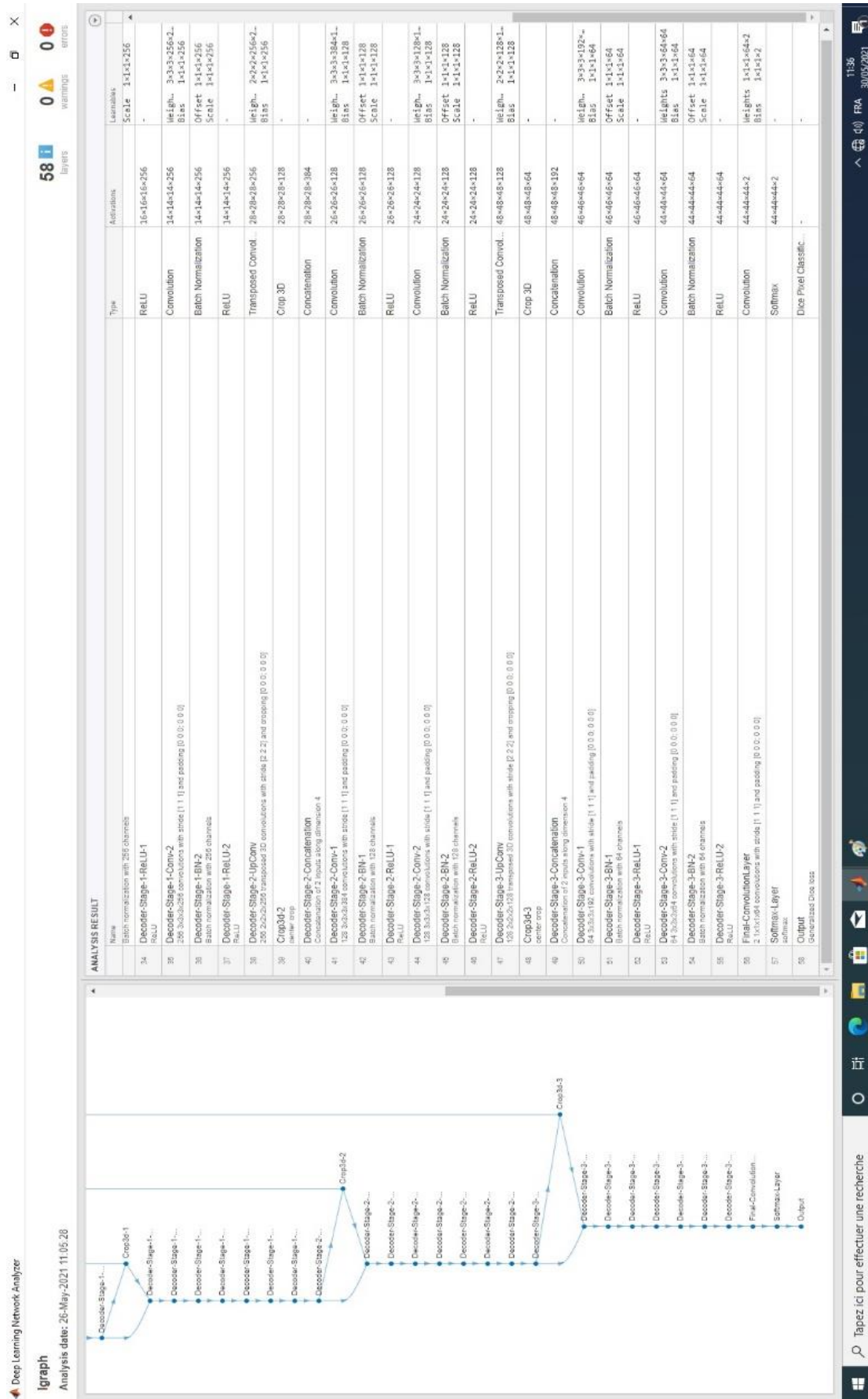


Figure 4. 7: b-U net 3D architecture

Comparaison entre l'image de référence du programme Matlab et résultat du l'image test après segmentation

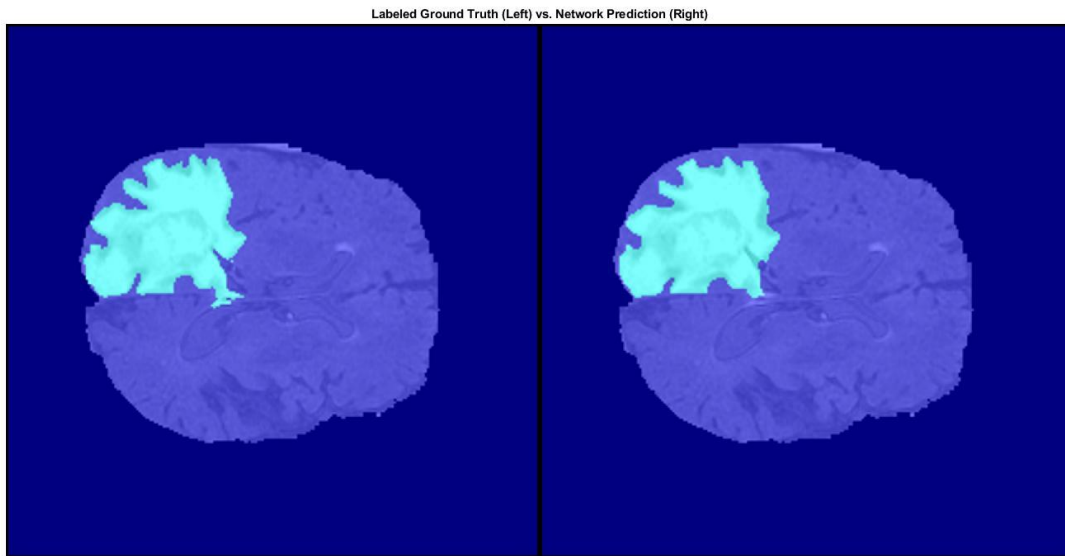


Figure 4. 8: Comparaison entre l'image de référence et l'image test après segmentation

Pour l'affichage du volume étiqueté de référence, Rendez l'arrière-plan totalement transparent en réglant la visibilité de l'étiquette d'arrière-plan. Comme la tumeur se trouve à l'intérieur du tissu cérébral, rendez certains voxels du cerveau transparents, de sorte que la tumeur soit visible. Pour rendre certains voxels du cerveau transparents, spécifiez le seuil de volume comme un nombre dans la plage [0, 1]. Toutes les intensités de volume normalisées inférieures à cette valeur seuil sont totalement transparentes. Dans ce programme, le seuil de volume est inférieur à 1 afin que certains pixels du cerveau restent visibles, pour donner un contexte à l'emplacement spatial de la tumeur dans le cerveau.

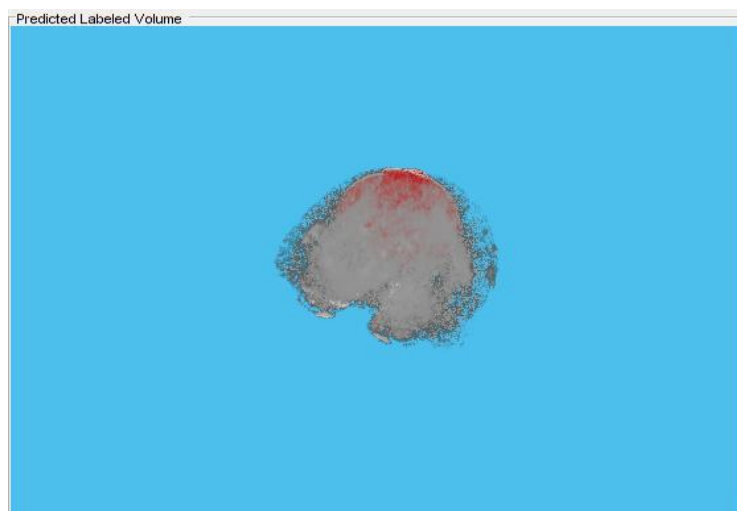


Figure 4. 9: Volume étiqueté de test

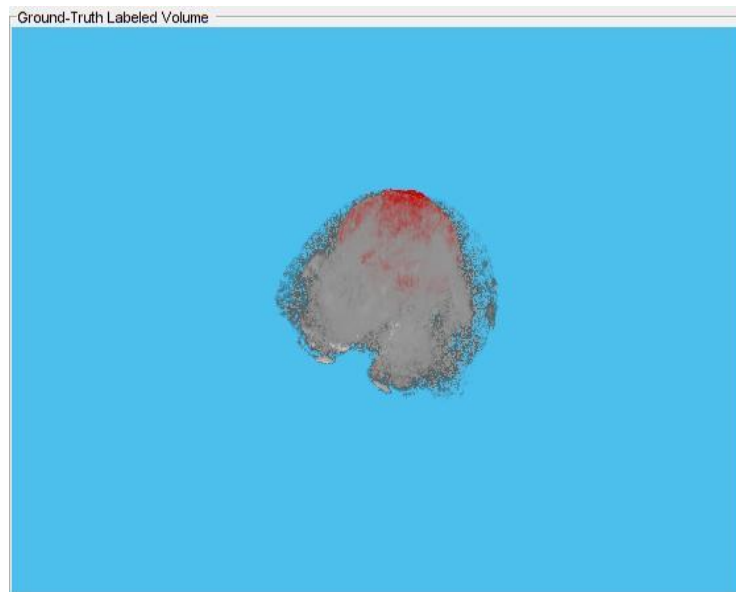


Figure 4. 10: Volume étiqueté de référence

### 5.3. Mesures d'évaluation de la segmentation

Pour évaluer la performance d'un algorithme de segmentation, on peut se référer à diverses mesures de similitude (ou dissimilitude).

Dans la littérature, nous pouvons trouver des mesures supervisées et non supervisées.

Toutefois, nous nous consacrons seulement aux mesures de similitude de la première catégorie.

Les mesures supervisées évaluent la performance de l'algorithme en le comparant à une image de référence de logiciel Matlab 2019b (vérité terrain).

Dans notre travail nous avons évalué notre approche en nous basant sur le coefficient de **Dice**

#### 5.3.1. Coefficient de similarité Dice

Le coefficient de similarité Dice de deux ensembles A et B est exprimé par :

$$\text{dice}(A, B) = 2 * |\text{intersection}(A, B)| / (|A| + |B|) \quad (4.1)$$

où  $|A|$  représente le cardinal de l'ensemble A .

L'indice Dice peut également être exprimé en termes de

- vrais positifs ( TP )
- faux positifs (FP)
- faux négatifs ( FN )

comme suit :

$$\text{dice}(A, B) = 2 * TP / ( 2 * TP + FP + FN ) \quad (4.2)$$

L'indice Dice est lié à l'indice Jaccard selon :

$$\text{dice}(A, B) = 2 * \text{jaccard}(A, B) / (1 + \text{jaccard}(A, B)) \quad (4.3)$$

### 5.3.2. résultat de calcul Coefficient de similarité (Dice)

Le Tableau 4.1 présente les valeurs des critères d'évaluation de résultat obtenue par le calcul de Coefficient de similarité (Dice) avec cinq volume d'images (arrière-plan et tumeur dans le cerveau)

volId	Résultat Coefficient de similarité (Dice)	
	Dice d'arrière-plan	Dice de la tumeur dans le cerveau
01	0.9992	0.9603
02	0.9994	0.9555
03	0.9991	0.9537
04	0.9993	0.9649
05	0.9995	0.9581

Le Tableau 4.1 : les valeurs de Dice des cinq volumes.

On trouve que la valeur de ce critère Coefficient de similarité (Dice) est proche de la valeur Un « 1 » qui indique la grande similarité entre l'image segmentée et l'image vérité terrain c-à-d l'image segmentation par le réseaux CNN Unet et presque similaire avec l'image de référence (vérité terrain).

On a aussi :

- La valeur moyenne des dés de fond (d'arrière-plan) sur 5 volumes Id de test égale à 0,9993.
- La valeur moyenne du Coefficient de similarité (Dice) de la tumeur sur 5 volumes Id de test égale à 0,9585.

D'après ces valeurs moyennes on constate que l'erreur est minimale dans l'opération de segmentation.

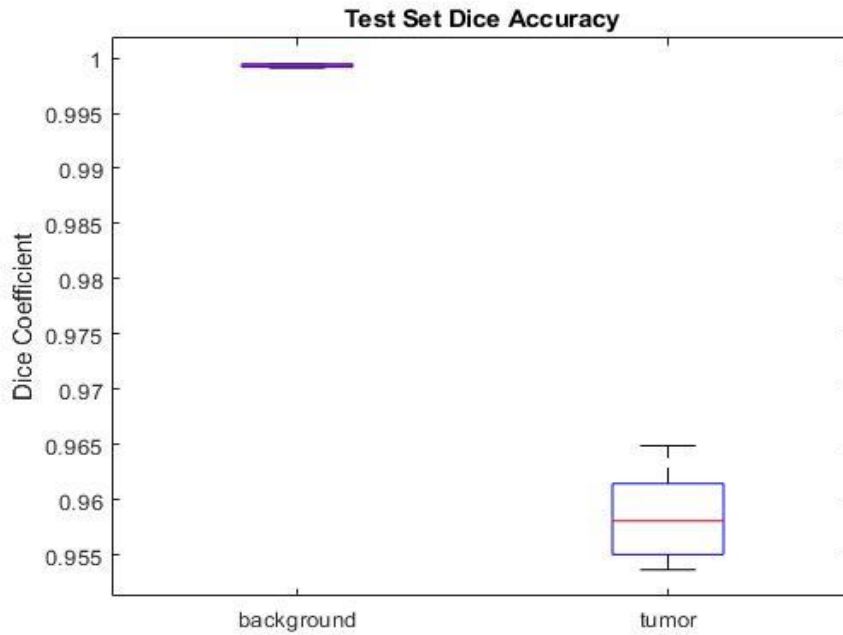


Figure 4. 11: Statistiques scores coefficient de similarité (Dice )

La figure 4.10 montre un boxplot (Statistics and Machine Learning Toolbox) qui visualise les statistiques sur les scores coefficient de similarité (Dice) pour l'ensemble des cinq volumes de test. Les lignes rouges du graphique indiquent la valeur médiane de coefficient de similarité (Dice) pour les classes. Les limites supérieure et inférieure de la boîte bleue indiquent les 25ème et 75ème percentiles, respectivement.

Les moustaches noires s'étendent jusqu'aux points de données les plus extrêmes qui ne sont pas considérés comme des valeurs aberrantes.

## 6. Conclusion

Sachant que le temps de la formation du réseau et de l'étiquetage de sa propre base de données est des inconvénients majeurs pour réaliser la segmentation sémantique nous avons préféré exploiter.

Dans ce chapitre nous avons détaillé les différentes étapes pour réaliser la segmentation sémantique avec le Deep Learning en utilisant Unet.

Nous avons introduit les réseaux de neurones convolutionnels utilisés en présentant les différents types de couches utilisées pour réaliser la segmentation.

# Conclusion générale

## **Conclusion générale**

Dans ce travail nous avons discuté la segmentation d'image et du Deep Learning médical (IRM) basé sur les réseaux de neurones convolutionnels en particulier Unet.

Nous avons introduit ces réseaux de neurones convolutionnels en présentant les différents types de couches utilisées dans la classification : la couche convolutionnelle, la couche de rectification, la couche de pooling et la couche fully connected. Nous avons exploité et fait maîtriser la segmentation sémantique.

Dans la phase d'implémentation, l'utilisation d'un CPU fait que le temps d'exécution soit trop. Afin de régler ce problème il est nécessaire d'utiliser des réseaux de neurones convolutionnels profonds déployés sur un GPU au lieu d'un CPU.

Comme perspectives, nous suggérons aux futures promotions de faire des études et des travaux de simulation sur d'autres model d'imagerie médicale en utilisant une autre technique de classification du CNN, et des autres algorithmes, en collaboration avec des centre d'imagerie médicale et des programme plus performante.

Pour réaliser sa propre segmentation sémantique, il est nécessaire d'utiliser sa propre base étiquetée, Ceci est un travail fastidieux mais qui est nécessaire dans le cas d'un domaine d'utilisation bien défini au préalable.

# Références

## LES REFERENCES BIBLIOGRAPHIQUES

- [1] C.Houassine, segmentation d'images par une approche biomimétique hybride. Université universite m'hamed bougara- boumerdes. 2012.
- [2] J. Canny, A, computational approach to edge detection, IEEE trans. on pattern analysis and machine intelligence, vol. 8, n°6, pp. 679-698, novembre 1986.
- [3] Sarah GHANDOUR. Segmentation d'images couleurs par morphologie mathématique : application aux images microscopiques. PhD thesis, Université de Toulouse III – Paul Sabatier, 2010.
- [4] Mr Cyril Meurie. Segmentation d'images couleur par classification pixellaire ethiérarchie de partitions. PhD thesis, Université de CAEN/BASSENORMANDIE, 2005
- [5] D. Michie, D. Spiegelhalter, and C. Taylor. Machine learning, neural and statistical classication, 1994. vol. 6, pages 84106.
- [6] R. Duda, P. Hart, and D. Stork. Pattern classication, 2001. Xiley Interscience 2e édition.
- [7] J. Cocquerez and S. Philipp. Analyse d'images : ltrage et segmentation, 1995. Paris Masson
- [8] J. Héroult and C. Jutten. Réseaux neuronaux et traitement du signal, 1994.
- [09 ] <http://www.psychomedia.qc.ca/lexique/definition/apprentissage-profond>
- [10] Mr. Mokri Mohammed Zakaria., » Classification des images avec les réseaux de neurones convolutionnels ": mémoire de projet de fin d'étude. Université Abou Bakr Belkaid Tlemcen.
- [11] R. BENTATA, « segmentation d'images tomographiques par émission de positions ». Mémoire de magister, département d'informatique, 2011.
- [12] Jean-Pergrale-Echographie Médicale : principes et application–Philips Medical Systems Research Paris-2005.
- [13] S. OUDDANE, « Compression des images médicales 3D ». Mémoire de magister département d'informatique, USTO-MB, 2009.
- [14] Site <https://dematerialisation.au.service.du.vivant.wordpress.com/les-differentetechniques-dimagerie-medicale/>
- [15] Erwan BERNARD «Comparaison théorique et expérimentale de l'imagerie active et de l'IR2 après traitement en conditions dégradées».23 novembre 2015
- [16] Imagerie Satellitale: Documentions produite par Business Image Group et Spot Image CNES1999.
- [17] <http://www.photonetc.com/fr/imagerie-hyperspectrale>

## Résumé :

La segmentation d'image est un sujet clé dans le traitement d'image et la vision par ordinateur avec des applications telles que la compréhension de scènes, l'analyse d'images médicales, la perception robotique, la vidéosurveillance, la réalité augmentée et la compression d'images, et autres. Divers algorithmes de segmentation d'images ont été développés dans la littérature. Dans ce travail nous avons proposé les modèles d'apprentissage en profondeur (deep learning models) pour segmenter les images.

**Mot clés :** Segmentation, IRM, deep learning , images médicales, CNN, Réseau de neurones, Unet

## Abstract

Image segmentation is a key topic in image processing and computer vision with applications such as scene understanding, medical image analysis, robotic perception, video surveillance, augmented reality and video surveillance. compression of images, and others. Various image segmentation algorithms have been developed in the literature. In this work we have proposed deep learning models to segment images.

**Keywords :** Segmentation, MRI, deep learning, medical images, CNN, Neural network, Unet

## ملخص

تجزئة الصور هو موضوع رئيسي وذو أهمية قصوى في معالجة الصور بالحاسوب في عدة مجالات وتطبيقات لفهم وتحليل المشاهد ونذكر على سبيل المثال تحليل الصور الطبية والإدراك الآلي والمراقبة بالفيديو والواقع المعزز والمراقبة بالفيديو وضغط الصور وغيرها من الخوارزميات لتجزئة الصورة تم تطويرها في وقتنا الحالي. ونقدم في هذه الدراسة نماذج التعلم العميق لتجزئة الصور .

**المفاتيح :** التجزئة ، التصوير بالرنين المغناطيسي ، التعلم العميق ، الصور الطبية ، الشبكة العصبية .