

République Algérienne Démocratique Et Populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université Larbi Tebessi
Faculté des sciences exactes et sciences de la nature et de la vie
Département de Mathématiques et Informatique



*Mémoire de fin d'étude en vue d'obtention
d'un diplôme de Master en Informatique*

Domaine : Informatique
Spécialité : Systèmes et Multimédias

Thème :

**Segmentation des images médicales par
apprentissage profond**

Présenté Par :

Sahli Aoulia

Devant le jury :

<i>Dr.</i>	<i>Zeggari ahmed</i>	<i>MCB</i>	<i>Université Larbi Tébessi</i>	<i>Président</i>
<i>Dr.</i>	<i>Gharbi Abdelhakim</i>	<i>MAA</i>	<i>Université Larbi Tébessi</i>	<i>Examineur</i>
<i>Dr..</i>	<i>akram bennour</i>	<i>MCA</i>	<i>Université Larbi Tébessi</i>	<i>Encadreur</i>
<i>Dr..</i>	<i>Rachid mahmoudi</i>	<i>MAA</i>	<i>Université Larbi Tébessi</i>	<i>Co-Encadreur</i>

Année Universitaire 2020/2021

Dédicaces

A mes parents, ` a qui je dois trop. `
sahli aoulia.

Remerciements

Nous tenons à remercier en premier lieu notre grand dieu qui nous a donné la force et le courage pour continuer et nous avoir aidées et éclairer le chemin pour réaliser de ce modeste travail

Je voudrais adresser toute ma gratitude aux directeurs de ce mémoire, dr.Akram Bennour et Dr. Rachid Mahmoudi enseignants au département de mathématiques et d'informatique, faculté des sciences exactes et sciences de la nature et de la vie de l'université de Tébessa, pour la patience, la disponibilité et surtout les judicieux conseils, qui ont contribué à alimenter ma réflexion.

Je remercie également tous les membres du jury pour l'intérêt qu'ils ont porté à mon travail :

- Dr. . Zeggari Ahmed,*
- Dr. Gharbi Abdelhakim*
- Dr.*
- Dr.*

.....

Enfin je désire aussi remercier tous les enseignants de l'Université Larbi Tebessi département de mathématiques et d'informatique, en me fournissant des données précises pour réussir à mes études universitaires.

Résumé

Aujourd'hui, l'imagerie médicale est une composante essentielle de l'ensemble de la communauté des soins de la santé et du bien-être, au processus du diagnostic, sélection du traitement et du suivi. La segmentation d'image est une tâche primordiale dans le domaine des traitements des images médicales qui consiste à classer chaque pixel d'une image dans une instance, où chaque instance correspond à une classe. Cette tâche fait partie du concept de la compréhension de la scène ou de mieux expliqué le contexte global d'une image. De ce fait, l'assurance d'une bonne qualité de segmentation rassure tout le reste du processus de traitement de ces images et influe considérablement sur des éventuelles décisions qui seront prises par les médecins sur l'état du patient en cours de traitement, et cela peut avoir des conséquences même sur la vie de ce dernier. Dans ce travail, nous allons explorer l'utilisation des réseaux de neurones profonds pour accomplir la tâche de segmentation des images médicales *Ultra-sonores* du cancer du sein. Nous allons adapter et comparer quatre modèles des réseaux de neurones de convolution : Alexnet, Resnet18, Resnet80, et U-net on utilisant l'apprentissage de transfert. Les résultats montrent l'efficacité des modèles utilisés, et confirment que la profondeur du réseau a une grande influence sur les résultats.

Mots-clés : imagerie médicale, segmentation sémantique d'image · Apprentissage en profondeur, apprentissage de transfert.

Abstract

Today, medical imaging is an essential component of the entire healthcare and wellness community, in the process of diagnosis, treatment selection and follow-up.

Image segmentation is a primordial task in the field of medical image processing which consists in classifying each pixel of an image in an instance, where each instance corresponds to a class. This task is part of the concept of understanding the scene or better explaining the overall context of an image. Therefore, the assurance of a good quality of segmentation reassures the rest of the process of processing these images and considerably influences any decisions that will be taken by the doctors on the patient's condition during treatment, and this can have consequences even in the life of the patient.

In this work, we will explore the use of deep neural networks to accomplish the task of segmenting of ultra sound medical images of breast cancer. We will adapt and compare four models of convolutional neural networks: Alexnet, Resnet18, Resnet80, et U-net using transfer learning. The results show the efficiency of the models used, and confirm that the depth of the network has a great influence on the results.

Mots-clés : imagerie médicale, segmentation sémantique d'image, apprentissage en profondeur, apprentissage de transfert.

يعد التصوير الطبي اليوم مكوناً أساسياً لمجتمع الرعاية الصحية والعافية بأكمله في عملية اختيار العلاج التشخيصي ومتابعة تجزئة الصورة مهمة أساسية في مجال معالجة الصور الطبية والتي تتكون من تصنيف كل بكسل من الصورة في حالة حيث كل مثيل يتوافق مع فئة. هذه المهمة هي جزء من مفهوم فهم المشهد أو شرح أفضل للسياق العام للصورة يطمئن بقية عملية معالجة هذه الصورة وتؤثر بشكل كبير على أي قرارات سوف يتخذها الأطباء على حالة المريض أثناء العلاج ويمكن أن يكون لذلك عواقب حتى في حياة المريض. في هذا العمل، سنستكشف استخدام الشبكات العصبية العميقة لإنجاز مهمة تقسيم الصور الطبية فائقة الصوت لسرطان الثدي، وسنعمل على تكييف أربعة نماذج من شبكات الشبكات العصبية : Alexnet, Resnet18, Resnet80 وU-net نحن نستخدم تعلم النقل. تظهر النتائج فعالية النماذج المستخدمة، وتأكيد أن عمق الشبكة له تأثير كبير على النتائج.

الكلمات المفتاحية: التصوير الطبي، تجزئة الصورة الدلالية • التعلم العمق، التعلم التحويلي.

Sommaire

Introduction générale : xi

Chapitre 1:L'imagerie médicale..... xiii

1 .1 Introduction :	3
1.2 L'image :	3
1.3 L'image numérique :	3
1.4. Les caractéristiques d'une image numérique :	4
1.4.1 Le pixel :	4
1.4.2 La résolution:	5
1.4. 3 Formats d'image :	5
1.5. L' imagerie médicale :	6
1.6 Les modalités d'images médicales :	7
1.6.1 Les rayons.....	7
1.6.2 La radiographie :	7
1.6.3 Le scanner:	7
1.6.4 Les Image Ultra-sonores :L'échographie.....	8
1.6. 5 Les images par résonance Magnétique :	8
1.6.6 L'image Nucléaire et scanners hybrides.....	9
1.7 Conclusion :	10

Chapitre 2:La segmentation des images..... 11

2.1 Introduction :	11
2.2 Définition de la segmentation :	11
2.2.1 Définition formelle:	11
2. 3 Les objectifs de la segmentation :	12
2.4 approches de segmentation :	12
2.4.1 Segmentation basé sur les régions (<i>en anglais : régions-based segmentation</i>) :	12
2.4.2Les méthodes de segmentation par détection de contours :	17
✓ Les méthodes analytiques :	22
2.4.3Approche basée classification de pixel.....	22
✓ La classification non supervisée :	24
2.4.4 Les méthodes de segmentation par seuillage.....	24
2 .5 Segmentation sémantique d'images.....	25
2.6 Segmentation d'images médicales :	26
2.6.1 Etat de l'art :	26

2.6 Conclusion :	28
Chapitre 3:Deep learning	xxix
3.1 Introduction :	32
3.2 L'apprentissage en profondeur (Deep Learning) :	33
3.2.1 Définition :	33
3.2.2 Pourquoi l'apprentissage profond (deep-learning) ?	33
3.3 Réseaux de neurones:	34
3.3.1 Réseaux de neurones (Modèle biologique) :	34
3.3.2 Le perceptron :	35
3.3.3 Les réseaux de neurones en couches :	37
3.4 Introduction du réseau de neurones convolutifs (CNN)	38
3.4.1 Architecture de réseaux de neurone convolutifs CNN	38
1) Couche de convolution :	39
2) Couche de pooling :	41
3) Couches entièrement connectées	42
3.4.2 Fonction d'activation	42
3.4.3 Fonction de perte :	43
3.4.3.1.Types de fonction de perte:	43
1) <i>Fonction</i> de perte de régression:	43
2) Fonctions de perte de classification binaire:	44
3) Fonctions de <i>perte</i> de classification multi-classes:	44
3.5 Quelques réseaux convolutifs célèbres :	44
3.5.1 GoogLeNet :	44
3.5.2 VGG16 :	45
3.5.3 RESNET :	46
3.5.4 AlexNet	47
3.5.5 <i>ReNet</i> :	47
3.5.6 <i>U-Net</i> :	48
3.6 Apprentissage de transfert(Transfer learning)	49
3.6.1 utilisation l'apprentissage par transfert:	51
3.6.2. apporche de l'apprentissage par transfert:	52
1) former un modèle pour réutiliser:	52
2) Utilisation d'un modele pré-formé:	52
3) Extraction des caractéristiques:	52
3.7 Conclusion :	53

Chapitre 4:Application	liv
4.1 Introduction :.....	52
4.3 Architectures de réseau de neurones profonds.....	52
4.4 Indices de performance.....	52
4.5 Dataset « BUT » (Brest Ultrasound Tumor » :.....	53
4.6. Environnement Matériel et logiciel : Dans cette section, nous présenterons les environnements matériel et logiciel de notre travail	54
4.6.1 Environnement matériel.....	54
4.6.2 Environnement logiciel :	54
1)MATLAB :.....	54
2) Google Colab :	55
4.7 Résultats et discussion :	55
4.8 Conclusions :	57
Conclusion générale.....	58

Liste des figures

Chapitre 1:

figure1- 1 Représentation d'une image numérique.....	4
figure1- 2 Image et pixels.....	5
figure1- 3 Résolution d'une image.....	5
figure1- 4 : Perfusion de résonance magnétique.	9

Chapitre 2:

figure 2- 1 segmentation d'une image à 4 classes.....	11
figure 2- 2 : segmentation basé sur les régions.....	12
figure 2- 3 : Différentes régions d'une image.....	13
figure 2- 4 :Division successive des blocs	16
figure 2- 5 :Gradient d'une image à 2 régions	19
figure 2- 6 :Principe des contours actifs	20
figure 2- 7 :Figure L'image de droite montre le histogramme	25
figure 2- 8 : Image et étiquette des pixels	26

Chapitre 3:

Figure 3- 1: La relation entre intelligence artificielle, Machine Learning et Deep Learning.....	33
Figure 3- 2:Neurones artificiels. la forme la plus générale de les neurones.....	35
Figure 3- 3: illustration de base du neurone:.....	36
Figure 3- 4:réseaux de Neurone.....	37
Figure 3- 5:Les couches d'un réseau de neurones convolutifs	39
Figure 3- 6: opération d'un filtre sur une image.....	40
Figure 3- 7 :Pooling avec un filtre 2x2 et un pas de 2	42
Figure 3- 8 : L'aplatissement de l'image.....	42
Figure 3- 9 :courbe sigmoïde et courbe ReLU	43
Figure 3- 10: Module de création avec réduction de la dimensionnalité de l'architecture Googlenet.	45
Figure 3- 11 : Architecture VGG-16 CNN. Figure extraite de La conversation de Matthieu Cordon avec sa permission.....	46
Figure 3- 12 : Bloc résiduel de l'architecture RESNET [49]	47
Figure 3- 13 : Architecture de réseau de neurones convolutionnelle Alexnet. Chiffre reproduit de [51].....	47
Figure 3- 14 : Une couche de modélisation d'architecture renouvelante verticale et Dépendances spatiales horizontales. Extrait de [52].....	48
Figure 3- 15 : L'architecture U-net[53].....	49
Figure 3- 16 : Processus d'apprentissage du transfert	51

Chapitre 4:

Figure 4 - 1: Tumeurs bénignes segmentées par les huit modèles CNN.....	57
---	----

Liste des tableaux

La liste des tableaux :

Chapitre 1:

Tableau 1 - 1 : Formats d'image et leurs propriétés.....6

Chapitre 4:

Tableau 4 - 1 : Matrice de confusion pour la segmentation binaire.....53

Introduction générale :

L'imagerie médicale est un domaine vaste guidé par les avancées en instrumentation, en techniques d'acquisition et en traitement d'images. Les progrès réalisés dans ces grandes disciplines concourent tous à l'amélioration de la compréhension de phénomènes physiologiques comme pathologiques. En parallèle, l'accès à des bases de données d'imagerie plus large, associé au développement de la puissance de calcul, a favorisé le développement de méthodologies par apprentissage machine pour le traitement automatique des images dont les approches basées sur des réseaux de neurones profonds. Parmi les applications où les réseaux de neurones profonds apportent des solutions, on trouve la segmentation d'images qui consiste à localiser et délimiter dans une image les régions avec des propriétés spécifiques qui seront associées à une même structure. La segmentation d'image fait référence à la division de l'image ou de la classification de pixel en ce qui concerne leurs caractéristiques locales ou / et leur caractéristiques texture qui sont très importantes pour comprendre le contenu de l'image.

Malgré de nombreux travaux récents en segmentation d'images par réseaux de neurones, l'apprentissage des paramètres d'un réseau de neurones reste guidé par des mesures de performances quantitatives n'incluant pas la connaissance de haut niveau de l'anatomie.

Dans Le domaine de traitements des images médicales la segmentation d'image est une tâche primordiale dans le domaine des traitements des images médicales, où l'étiquetage des images à besoin de soins élevés dans le but d'identifier la diagnostique de la maladie et de la recherche clinique. Cette tâche fait partie du concept de la compréhension de la scène ou de mieux expliqué le contexte global d'une image. De ce fait, l'assurance d'une bonne qualité de segmentation rassure tout le reste du processus de traitement de ces images et influe considérablement sur des éventuelles décisions qui seront prises par les médecins sur l'état du patient en cours de traitement, et cela peut avoir des conséquences même sur la vie de ce dernier.

L'objectif de ce travail est de développer des méthodes permettant d'intégrer des réseaux de neurones profonds, en ciblant la segmentation des images médicales. Nous allons explorer l'utilisation des réseaux de neurones profonds pour accomplir la tâche de segmentation des images médicales ultra Sound du cancer du sein. Nous allons adapter et comparer quatre modèles des réseaux de neurones de convolution :

Alexnet, Resnet18, Resnet80, et U-net on utilisant l'apprentissage de transfert. Les résultats montrent l'efficacité des modèles utilisés, et confirment que la profondeur du réseau a une grande influence sur les résultats .

Nous avons structuré notre mémoire en quatre chapitres :

Le chapitre 1 est dédié aux notions de base sur les images, les différents systèmes de l'imagerie médicale ,

Le chapitre 2, est un survol sur les différentes classes de méthodes de segmentation, il introduit ainsi la segmentation sémantique.

Le chapitre 3 est un chapitre théorique qui vise à assimiler les divers concepts de base de l'apprentissage profond (Deep Learning).

Le chapitre 4 il présente les logiciels les plus connus et les plus utilisés dans la segmentation des images médicales et Résultats », est consacré aux différents tests et évaluations de la solution proposée.

Enfin, nous achevons ce mémoire par une conclusion générale sur le travail que nous avons effectué.

Chapitre 1:L'imagerie médicale

1.1 Introduction :

L'imagerie médicale est un moyen qui permet l'acquisition et la restitution des images internes du corps humain. Elle regroupe les moyens d'acquisition et de restitution d'images du corps humain à partir de différents phénomènes physique tel que l'absorption de rayon X la résonance magnétique nucléaire , la réflexion d'ondes ultra-sons ou de la rétroactivité, auxquels on associe parfois les techniques d'imagerie optique comme l'endoscopie.

L'imagerie médicale est un élément essentiel à la recherche clinique, qui est en perpétuelle évolution et joue de plus en plus un rôle important dans la prise en charge des patients L'étude des maladies et la mise au point de nouveaux traitements, elle fournit des moyens de suivi des réponse tumorale et progression des patients après le traitement.

Cependant, le traitement et l'évaluation quantitative de ces images par des mesures manuelles est fastidieux, longue et sujette à inter-opérateur, car l'évaluation visuelle peut être non objective. L'évolution de l'intelligence artificielle (IA) a contribué massivement dans le domaine médical à travers divers applications, y compris la détection des maladies résiduelles sans intervention chirurgicale, ainsi que d'autres pratiques de médecine personnalisée. Elle a permis d'automatiser la quantification des caractéristiques radiographiques des hénotypes tumoraux ainsi que le suivi des changements dans les tumeurs avant, pendant, et après traitement de manière quantitative.

1.2 L'image :

L'image est un signal de $2D$ (bidimensionnel) ou $3D$ (tridimensionnel) en tenant compte de la profondeur de l'image ou du temps. L'image continue est associée à une fonction continue à 2 variables $f(x, y)$, x et y étant des variables d'espace, coordonnées d'un point. On s'intéresse aux images discrètes obtenues par numérisation, et représentée par une fonction $f(i, j)$ avec i et j variables d'espace, désignant des variables discrètes, suivant respectivement x et y [1] .

1.3 L'image numérique :

L'image numérique est une image dont la surface est divisée en éléments d'une taille fixe appelés cellules ou pixels, ayant chacune comme caractéristiques : le niveau de couleurs (ou de gris) prélevé et l'emplacement correspondant dans l'image réelle.

La numérisation d'une image est la conversion de celle-ci de son état analogique en une image numérique représentée par une matrice bidimensionnelle de valeurs numériques $f(x, y)$, tel que :

$f(x, y)$: niveau d'intensité

et

x, y : coordonnées cartésiennes d'un point de l'image.

Pour des raisons de commodité de représentation pour l'affichage et l'adressage, les données d'images sont généralement rangées sous forme de tableau I de n lignes et p colonnes où chaque élément $I(x, y)$ représente un pixel de l'image et sa valeur est associée à un niveau de gris codé sur m bits (2^m niveaux de gris, $0 = \text{noir}$, $2^m - 1 = \text{blanc}$).

La valeur en chaque point exprime la mesure d'intensité lumineuse perçue par le capteur. [2]

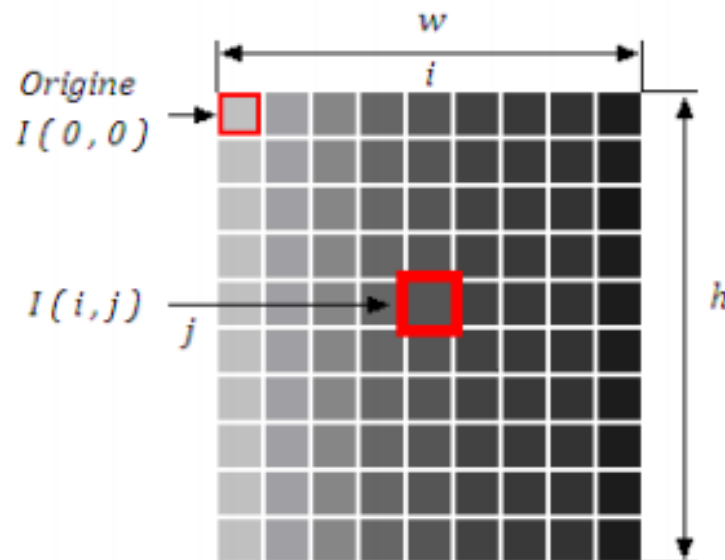


figure1- 1 Représentation d'une image numérique[2]

1.4. Les caractéristiques d'une image numérique :

1.4.1 Le pixel :

Le pixel est la taille du plus petit élément de l'image. Il désigne aussi un point de la matrice image. Le pixel peut avoir une dimension qui est associée à la résolution spatiale de l'image, on aperçoit en effet que celle-ci se compose d'une multitude de petits carrés comme illustré dans la figure 1.2. [3]

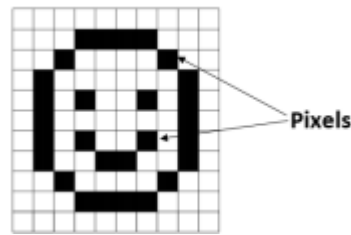


figure1- 2 Image et pixels[3]

1.4.2 La résolution:

La résolution est le nombre de points(pixel) constituant l'image c'est -à-dire sa « dimension informatique » (le nombre de colonnes de l'image que multiplie son nombre de lignes), Elle est exprimée en points par pouce par PPP, en anglais DPI , un pouce égale à 2.54 cm une résolution de 300 DPI signifie donc 300 colonnes et 300 rangées de pixels sur un pouce carré ce qui donne donc 90000 pixels sur un pouce carré, ce qui donne donc 90000 pixels sur un pouce carré .

La résolution d'une image numérique définit le degré de détail de l'image figure1.3

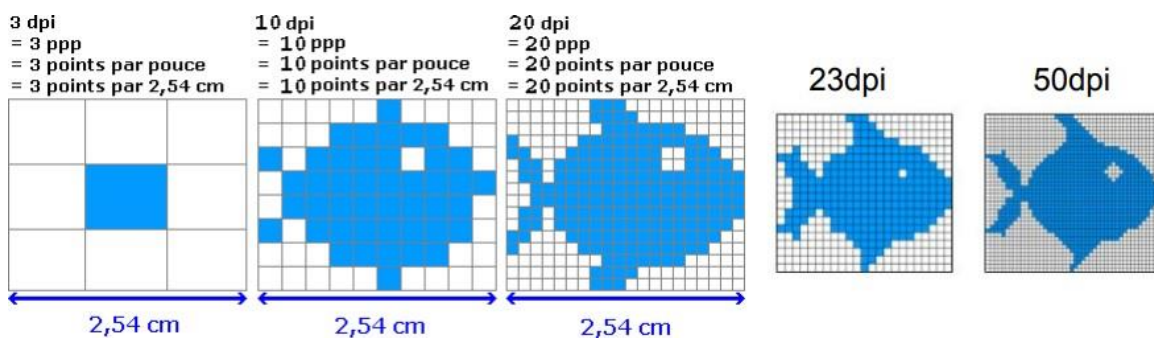


figure1- 3 Résolution d'une image[3]

1.4. 3 Formats d'image :

Différents formats d'image sont généralement adaptés à différentes applications. Par exemple, les images GIF présente un format de stockage d'image très basique limité à seulement 256 niveaux de gris ou couleurs.

Le tableau 1 1 : présente quelques formats les plus connues.

Tableau 1 1 : Formats d'image et leurs propriétés

Type	Name	Propretés
<i>GIF</i>	<i>Graphics interchange format</i>	Limité à seulement 256 couleurs (8 bits); sans perte compression
<i>JPEG</i>	<i>Joint Photographic Experts Group</i>	-Dans l'utilisation la plus courante aujourd'hui; -avec perte compression; -des variantes sans perte existent
<i>BMP</i>	<i>Bit map picture</i>	_Format d'image de base; -limité (généralement) compression sans perte; - des variantes avec perte existent
<i>PNG</i>	<i>Portable network graphics</i>	-Nouveau format de compression sans perte; -conçu pour remplacer GIF
<i>TIF/TIFF</i>	<i>Tagged image (file) format</i>	Très flexible, détaillé et adaptable format; -variantes compressées / non compressées exister

1.5. L'imagerie médicale :

Les images médicales est un ensemble d'images du corps humain à l'aide de différents phénomènes physique par exemple la résonance magnétique, et la représentation des variations spatiales et temporelles d'une ou de plusieurs propriétés physiques d'un sujet sous investigation. On parle d'images statiques lorsque la variation temporelle de la propriété n'est pas prise en considération, autrement on parle d'images dynamiques

Les propriétés physiques sont par exemple la densité électronique en radiographie et tomographie aux rayons X (en abrégé RX et XCT) ou la densité protonique et les temps de relaxation en résonance magnétique nucléaire (RMN) ou une distribution de radioactivité en tomographie par émission de photons simple (SPECT) ou par émission de positrons (PET).[4]

Les images médicales contiennent des valeurs proportionnelles aux caractéristiques d'absorption du tissu par rapport à un signal projeté à travers le corps.

Les types les plus courants sont la tomodensitométrie (TDM) et l'imagerie résonance magnétique. Les images CT, comme les rayons X conventionnels, représentent des valeurs qui sont directement Proportionnelles à la densité du tissu à travers lequel le signal est passé, par contre, les images par résonance magnétique présentent plus de détails. En CT et IRM, le format d'image 2D est généralement étendu à un volume 3-D; le volume 3-D est essentiellement juste une pile des images

2-D (plusieurs points de vus).

1.6 Les modalités d'images médicales :

Les appareils d'acquisition d'images médicales sont nombreux et permettent d'obtenir des informations différentes selon le procédé physique utilisé pour observer les tissus du corps humain et les principales modalités d'acquisition d'images médicale sont :

1.6.1 Les rayons

A. Historique :

Jusqu'au 8 novembre 1895, les médecins n'avaient pas accès à images de ce qui se passait à l'intérieur de corps d'un patient. Sur cette journée fatidique, Wilhelm Conrad Roentgen découvrir un "nouveau type de rayon" qu'il a appelé "Xrays" tout en expérimentant une cathode des escrocs Tube à rayons, Figure (1-4). Cette invention mis en mouvement une révolution dans le Diagnostic médical qui continue à ce jour.[5]

B. Définitions :

Le rayon X résulte d'une onde électromagnétique que est la propagation d'un champ électrique et d'un champ magnétique associés perpendiculaires entre eux et à la direction de propagation. Dans le vide Un rayonnement électromagnétique est donc associé à un faisceau de photons. Dans le cas d'un rayonnement monochromatique, les photons ont tous la même énergie. Au contraire, un rayonnement polychromatique est constitué de photons d'énergies différentes

1.6.2 La radiographie :

Le mot (radiographie) peut désigner l'ensemble des technique permettant de réaliser les cliché des structures internes d'une courante est radiographie médicale dans laquelle les clichés traduisent l'opacité plus ou moins marquée des tissus ou organes par une teinte plus ou moins claire Cependant, la qualité d'image n'est pas la meilleure. En effet, la radiologie fournie des images des organes dans toute leur épaisseur ce qui ne permet pas de voir toutes les liaisons.

1.6.3 Le scanner:

Au début du XXe siècle, Alessandro Vallebona, un radiologue italien, a inventé la tomographie conventionnelle. Cela est fait déplaçant simultanément la source de rayons X et détecteur de rayons X en tandem afin de garder un objet d'intérêt dans le centre du plan de

numérisation et flou de tous les autres objets. La tomographie conventionnelle a évolué, mais il était toujours inefficace pour l'imagerie des tissus mous et des grandes zones du corps [6].

Soixante-quinze ans après la naissance de l'imagerie médicale, la tomographie calculée (CT) a été développée en utilisant des projections radiographiques de plusieurs angles et puis construire une image bidimensionnelle avec un modèle mathématique qui incorpore toutes les données. En 1967, Sir Godfrey Hounsfield a inventé le premier scanner CT aux laboratoires de recherche EMI. Le premier patient vivant a été scanné le 1er octobre 1971. Alan M. Cormack (qui a créé le début de la mathématique des modèles utilisés dans le CT) et Hounsfield ont reçu conjointement un prix Nobel en 1979 pour le développement de "l'imagerie assistée par ordinateur." [7]

De nombreuses avancées technologiques ont été faites dans le CT, avec des améliorations de la vitesse de balayage, une tranche plus petite, une épaisseur, diminution de la dose de rayonnement et une meilleure qualité d'image. Il y a vingt-cinq ans, une étude CT typique pourrait prendre des dizaines de minutes en fonction du volume numérisé, alors qu'aujourd'hui, les études CT peuvent être effectuées dans une fraction d'une seconde, couvrant de grandes zones du corps, même en utilisant plus d'énergie de balayage. Les tranches de CT dans les machines modernes peuvent être aussi minces que des fractions d'un millimètre [6],

1.6.4 Les Images Ultra-sonores : L'échographie

Le diagnostic ultra-sonore a commencé avec l'étude des chauves-souris à la fin des années 1700 de Lazzaro Spallanzani. Spallanzani a essayé de comprendre comment les chauves-souris volent la nuit et a hypothéqué que les chauves-souris s'appuient sur le son pour naviguer [8].

L'échographie utilise des ondes sonores à haute fréquence au-dessus de la gamme d'audience humaine, qui sont transmises dans le corps. Les échos réfléchis sont mappés dans une image basée sur l'amplitude et le temps de retard des signaux retournés. Cette modalité a des avantages sur le CT. Parce que nous n'avons pas d'ionisation, le rayonnement est peu coûteux, elle est largement disponible et est effectuée face à face avec des patients.

1.6.5 Les images par résonance Magnétique :

L'histoire du développement de l'imagerie par résonance magnétique (IRM) est complexe et implique nombreuses contributions de cliniciens, scientifiques et ingénieurs techniques. Deux scientifiques, Felix Bloch et Edward Purcell, simultanément et indépendamment ont développé le concept de résonance magnétique nucléaire (RMN), pour laquelle ils ont été attribués conjointement

Le prix Nobel de la physique en 1952. En 1973, Paul Lauterbur a publié les premières images NMR, et plus tard a reçu le prix Nobel de 2003 en physiologie et médecine.[7]

L'imagerie par résonance magnétique est basée sur différents principes physiques que « CT ». Un aimant puissant est utilisé pour produire un très fort champ magnétique fixe autour du patient et des impulsions radiofréquences sont utilisées pour exciter les protons dans le corps, Quand les protons excités se détendent à un état de repos, les signaux de retour qui sont capturés et mappés dans une image. Aujourd'hui, l'IRM est utilisé pratiquement dans toutes les sous-spécialités de la médecine.

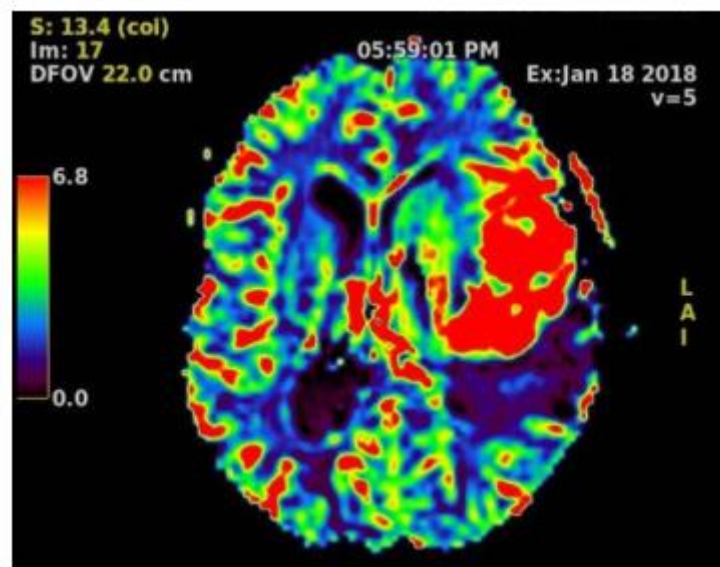


figure1- 4 : Perfusion de résonance magnétique.[9]

1.6.6 L'image Nucléaire et scanners hybrides

L'imagerie nucléaire et les scanners hybrides utilisent des traceurs radioactifs pour créer une imagerie fonctionnelle ; où des capteurs uniques détectent l'émission radioactive de ces traceurs.

Les scans des spectres mesurent des rayons gamma, et des détecteurs spéciaux enregistrent l'arrivée simultanée de la paire de rayons gamma et le nombre de ces arrivées. Les données sont utilisées pour mapper l'activité métabolique dans les organes du corps. James Robertson a construit le premier scanner pour animaux dans le laboratoire national de Brookhaven en 1961.

Le co-enregistrement fonctionnel et morphologique (c'est-à-dire le premier design hybride) a été proposé pour la première fois pour un Scanner PET-CT de David Townsend et Ronald Nutt. Un Scanner « prototype hybride Clinical » PET-CT a été utilisé pour la première fois cliniquement en 1998,[10].

1.7 Conclusion :

Dans ce chapitre, nous avons vus des concepts généraux et des caractéristiques des images, Nous avons défini l'imagerie médicale comme Toutes les technologies permettant de visualiser des parties du corps humain ou des organes, son but est de diagnostiquer et guider les gestes de traitement, tel que la ponction, ou surveillance de l'effet du traitement. Ensuite, nous avons répertorié quelques méthodes utilisées dans le domaine de l'imagerie, telles que: Radiologie, échographie.

Chapitre 2:La segmentation des images

2.1 Introduction :

La segmentation est le nom donné au processus générique par lequel une image est subdivisée en régions ou objets constitutifs. Elle occupe un rôle très important dans le traitement de l'image car c'est souvent la première étape vitale qui doit être réussie avant les tâches suivantes telles que l'extraction de caractéristiques, la classification, la description, etc. Après tout, si vous ne pouvez pas localiser les objets en premier lieu, comment pouvez-vous les classer ou les décrire ?

2.2 Définition de la segmentation :

Segmenter une image signifie, trouver ses régions homogènes et ses contours, ces deux derniers sont supposés être pertinents, c'est-à-dire que les régions doivent correspondre aux parties significatives des objets du monde réel, et les contours à leurs frontières apparentes.

Soit ' ' le domaine de l'image et ' ' la fonction qui associe à chaque pixel une valeur $f(x,y)$. [11]

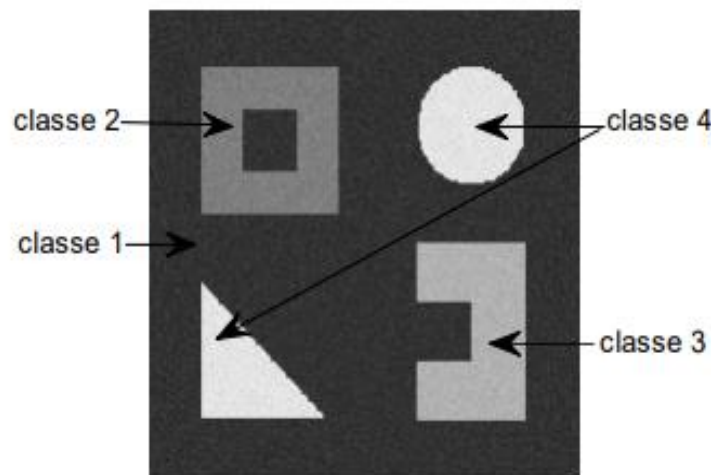


figure 2- 1 segmentation d'une image à 4 classes [11]

2.2.1 Définition formelle:

nous définissons un prédicat 'P' sur l'ensemble des parties de 'I', la segmentation de 'I' est définie comme une partition de 'I' en 'n' sous-ensemble $\{R_1, R_2, \dots, R_n\}$ tels que :

1. $I = \bigcup_{i=1}^n R_i$;
2. $R_i \neq \Phi$;
3. $\forall i, j \quad ; \quad i \neq j, R_i \cap R_j = \Phi$;

Ce qui se traduit par :

- L'union de régions R_i constituent les l'image initiale I

- les pixels doivent être traités et les appartenant à une même région R_i .
- Les régions R_i et R_j doivent être deux à deux disjointes[12]

2.3 Les objectifs de la segmentation :

L'objectif de la segmentation est découper l'image en plusieurs régions, dans lesquelles les pixels vérifient un certain critère d'homogénéité, tel que le niveau de gris ou la couleur, il existe de très nombreuses méthodes permettant ce découpage dont l'efficacité dépend avant tout de l'image. Elles visent à :

- Fournir des régions homogènes (selon un critère donné).
- Etudier et interpréter des structures anatomiques.
- Localiser de manière précise les contours des régions.
- Minimiser le bruit[13]

2.4 approches de segmentation :

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

2.4.1 Segmentation basé sur les régions (*en anglais : régions-based segmentation*) :

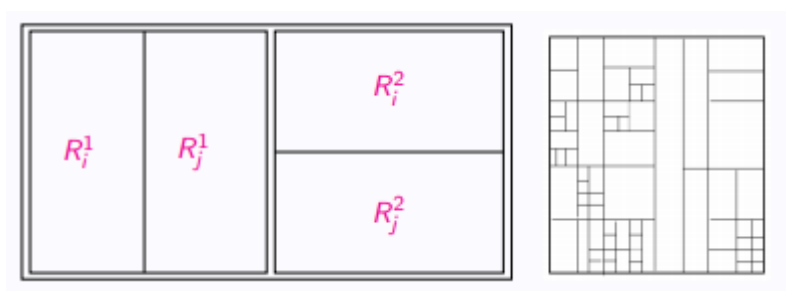


Image originale

Image partitionnée

figure 2- 2 : segmentation basé sur les régions[14]

A.Définition 1 :

Une **région** est un sous-domaine de l'ensemble de coordonnées E d'une image I . C'est un sous-ensemble de pixels de I lorsque $E \subset \mathbb{N}^2$ et un sous-ensemble de voxels de I lorsque $E \subset \mathbb{N}^3$. [14]

B.Définition 2 :

Une **frontière** est une structure marquant la séparation d'une région avec l'ensemble de ses régions connexes. Elle peut être un ensemble de pixels, de voxels, une courbe ou une surface. : [14]

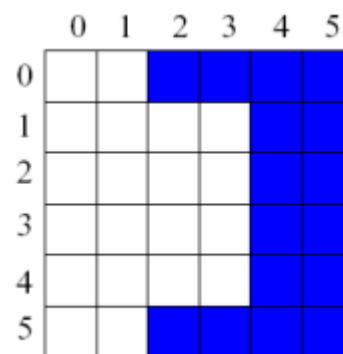


figure 2- 3 : Différentes régions d'une image [13]

L'approche segmentation par division de régions consiste à diviser l'image originale en régions homogènes au sens d'un critère donné. Ce processus est récursif et considère que la région initiale correspond à l'image à analyser. Si une région ne respecte pas un prédicat d'homogénéité, elle est divisée en quatre sous-régions de taille égale. Chaque sous-région est ensuite analysée. L'algorithme récursif s'arrête lorsque toutes les régions respectent le prédicat d'homogénéité [15]

✓ **Croissance de région (région growing) :**

Il a été suggéré d'utiliser un prédicat qui vérifie si l'erreur d'approximation de la touche de pixels par une fonction est inférieure à un seuil.

* L'image additionnelle est un ensemble d'échantillons quantifiés d'une fonction variable continue.

* L'idée est d'installer une surface de faible commande appropriée (par exemple, planaire ou biquadratique) sur les données d'image d'une région.

- 1) partitionner l'image dans les régions de graines initiales $R_i^{(0)}$ (par exemple, diviser l'image en 7x7 Régions).

- 2) adapter un modèle plan pour chaque région de la graine. Si $E(R_i^{(0)}, a, m)$ est assez petit, accepter $R_i^{(0)}$ et son modèle; sinon rejeter $R_i^{(0)}$
- 3) Pour chaque région, trouvez tous les points compatibles avec la région en considérant les voisins de la région (c'est-à-dire d'éviter les valeurs aberrantes).

$$C_i^{(k)} = [(x, y): (f(x, y, a, m))^2 < \varepsilon \text{ et } (x, y) \text{ est un } 4\text{-voisin de } R_i^{(0)}]$$

- 4) S'il n'y avait pas de points compatibles, alors $m = m + 1$. Si $M > M$, ne poussez $R_i^{(k)}$ plus loin; Sinon, passez à l'étape 3.
- 5) former la nouvelle région

$$R_i^{(k+1)} = R_i^{(k)} \cup C_i^{(k)}, \text{ Refaites le modèle à } R_i^{(k+1)} \text{ Et calcul } E(R_i^{(k+1)}, a, m)$$

- 6) Computer l'erreur de différence:

$$P^{(k)} = E(R_i^{(k+1)}, a, m) - E(R_i^{(k)}, a, m)$$

- 7) si $\rho^{(k)} < T_1$ allez à l'étape 3.
- 8) $m = m + 1$; Si $M > m$, ne poussez pas la région plus loin.
- 9) Remplacez la région au nouveau modèle $f(x, y, a, m)$ Si l'erreur d'ajustement diminue, Acceptez le nouveau modèle et passez à l'étape 3; Sinon, ne poussez pas la région plus loin [16]

Principe général :

- Approche ascendante
- Départ d'un pixel
- Analyse de ses pixels voisins et analyse du critère d'homogénéité
- Croissance de la région jusqu'au critère d'arrêt.

Avantage :

- Rapide et simple Cette technique
- Facile à exécuter

Inconvénients

- Cette technique très sensible au bruit
- Certains pixels ne peuvent pas être classés

✓ ***Segmentation par fusion de régions (Region Merging) :***

- les opérations de fusions de la région éliminent les fausses frontières et les régions parasites par fusionner des régions adjacentes appartenant au même objet.

- Les systèmes de fusion commencent par une partition satisfaisante (4) (par exemple, des régions produites à l'aide de seuils).

$$(4) P(R_i) = \text{Vrai}$$

- alors, ils procèdent à remplir l'état (5) en fusionnant progressivement une image adjacente Régions.

$$(5) P(R_i \cup R_j) = \text{False}$$

- 1) former des régions initiales dans l'image.
- 2) Construire un graphique d'adjacence des régions (RAG).
- 3) Pour chaque région faire:

A. Considérez sa région adjacente et son test pour voir s'ils sont similaires

B. Pour les régions similaires (c'est-à-dire $(5)P(R_i \cup R_j) = \text{Vrai}$) ?fusionner-les et modifier le RAG

- 4) Répétez l'étape 3 jusqu'à ce que aucune régions ne soit fusionnée.

Comment déterminer la similarité de la région?

- Basé sur les valeurs grises des régions
 - Comparez leurs intensités moyennes.
 - Utilisez le raccord de surface pour déterminer si les régions peuvent être approximées d'une surface.
 - Utilisez des tests d'hypothèse pour juger de la similitude des régions adjacentes (suppose que les valeurs d'intensité sont dessinées d'une distribution de probabilité).
- Sur la base de la faiblesse des frontières entre les régions.

✓ **Segmentation par division de régions (Split)**

- Les opérations de division de la région Ajoutez des frontières manquantes en divisant les régions contenant des parties d'objets différents, Fig.2.4

- Les schémas de fractionnement commencent par une partition satisfaisante (5), par exemple l'image entière.

$$P(R_i \cup R_j) = \text{False} \quad (5)$$

- Ensuite, ils procèdent à satisfaire la condition (4) en divisant progressivement les régions d'image.

$$P(R_i) = \text{True} \quad (4)$$

- Deux principales difficultés à mettre en œuvre cette approche:

* Décider quand diviser une région (par exemple, utiliser la variance, le raccord de surface).

* Décider de la division d'une région.

Décomposition régulière

- (1) si $P(R) = False$, divisé r en quatre quadrants
 (2) Si P est fausse sur n'importe quel quadrant, sous-remplissant

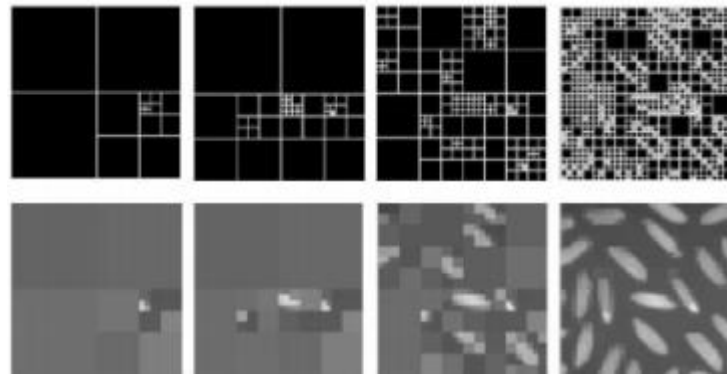


figure 2- 4 : Division successives des blocs [16]

✓ *Segmentation par division-fusion (Split and Merge)*

La scission ou la fusion pourrait ne pas produire de bons résultats lorsqu'il est appliqué séparément.

- De meilleurs résultats peuvent être obtenus en entrelacement de fusion et de scission.
- Cette stratégie prend une partition qui ne satisfait probablement aucune condition (4) ou (5) dans le but de produire une segmentation qui satisfait à la fois des conditions

1. divisé en quatre quadrants disjoints toute région R_i RI où $P(R_i) = False$
2. fusionner toutes les régions adjacentes R_j et R_k pour lequel

$$P(R_j \cup R_k) = Vrai$$

(3) Arrêtez-vous quand aucune autre fusion ou une fractionnement n'est possible[16]

Les inconvénients de cette méthode se situent à trois niveaux :

1. régions obtenues ne correspondent pas, dans tous les cas, aux objets représentés dans l'image.
2. Les limites des régions obtenues sont habituellement imprécises et ne coïncident pas exactement aux limites des objets de l'image.
3. La difficulté d'identifier les critères pour agréger les pixels ou pour fusionner et diviser les régions.
4. Les régions obtenues ne correspondent pas, dans tous les cas, aux objets représentés dans l'image.
5. Les limites des régions obtenues sont habituellement imprécises et ne coïncident pas exactement aux limites des objets de l'image.
6. La difficulté d'identifier les critères pour agréger les pixels ou pour fusionner et diviser les régions.[18]

2.4.2 Les méthodes de segmentation par détection de contours :

A. Définition 1 :

Un **contour** une liste d'arêtes ou de courbe mathématique qui modélise la liste des arêtes.

B. Définition 2 :

Un **contour** d'arêtes est un algorithme qui produit un ensemble d'arêtes {bords ou fragments de bord} à partir d'une image

Les bords physiques fournissent des informations visuelles importantes car elles correspondent aux discontinuités dans Les propriétés physiques, photométriques et géométriques des objets de scène. Le principe physique les bords correspondent à des variations significatives dans l'éclairage, l'orientation et la profondeur des surfaces de la scène. Étant donné que l'intensité de l'image est souvent proportionnelle au rayonnement des scènes, les bords physiques sont représentés dans l'image par des modifications de la fonction d'intensité.

Pour les détecter les contours, il existe deux types d'approches :

✓ **Les méthodes dérivatives :**

Les modèles dérivatifs consistent à modéliser les contours ou des zones d'images et supposent que l'image numérique provient de l'échantillonnage d'une onction scalaire a support borné et dérivable en tout point .un contour représente généralement une variation d'intensité entre deux régions connexes d'une image .la variation de la fonction scalaire que représente le contour correspond généralement aux extrema locaux du gradient ou aux zéro du laplacien de la fonction ainsi le gradient est une fonction vectorielle des pixels [i,j]

$$\nabla f[i, j] = \left(\frac{\delta f}{\delta x} [i, j], \frac{\delta f}{\delta y} [i, j] \right) \quad (2.1)$$

Alors que le Laplacien est une fonction scalaire de [i,j]

$$\nabla^2 f[i, j] = \left(\frac{\delta^2 f}{\delta x^2} [i, j], \frac{\delta^2 f}{\delta y^2} [i, j] \right) \quad (2-2)$$

Les méthodes dérivatives sont cependant moins adaptées au traitement d'images médicales que d'autres méthodes, car elles se basent principalement sur la présence de contraste entre les régions de l'image. Ces techniques ont aussi l'inconvénient d'être hautement sensibles au bruit présent dans l'image, ce qui est souvent cas dans les images médicales

Il ya deux approche à méthodes dérivatives

1. Approche laplacien : détermination des passages par zéro du laplacien
2. Approche gradient : détermination des extrema locaux dans la direction de gradient.

A. Détection de contours par l'approche gradient

Le gradient est un opérateur de commande de premier ordre de Ned comme vecteur $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ Le module et le La direction du gradient est défini par:

$$|\overrightarrow{\text{grad}}| = \sqrt{(\frac{\partial}{\partial x})^2 + (\frac{\partial}{\partial y})^2} \quad (2-3)$$

La direction de gradient est perpendiculaire à l'orientation des bords. Dans de nombreux systèmes proposés, La direction de gradient est utilisée pour localiser les bords. L'opérateur de module de dégradé est non linéaire et invariant à la rotation. Il est uniquement calculé en utilisant des dérivés dans x et y. Dans un bruyant image, l'utilisation de plusieurs dérivés directionnels peut être utile pour augmenter le signal à nois rapport. D'un point de vue informatique, selon le théorème de la direction[19]

Le gradient d'une image donne les taux de changement de niveau de gris par unité de distance dans les directions des axes de coordonnées. Il est définit comme un vecteur caractérisé par son amplitude et sa direction, tels que :

- L'amplitude est liée à la quantité de variation locale des pixels.
- La direction du gradient est orthogonale à la frontière qui passe au point considéré.

Le gradient d'une image dans un point p(x, y) est le vecteur I (x, y) qui est définit par :

$$I(x, y) = (\frac{\partial I(x,y)}{\partial x}, \frac{\partial I(x,y)}{\partial y})^t \dots\dots \quad (2-4)$$

$$\Delta I(x, y) = (\frac{\partial I(x,y)}{\partial x}, \frac{\partial I(x,y)}{\partial y})^t \dots\dots \quad (2.5)$$

Et pour calculer ce gradient, on effectue généralement, le produit de convolution de l'image avec un opérateur de dérivation fournissant deux masques M1 et M2 tels que :

$$\frac{\partial I(x,y)}{\partial x} = M1 \times I(x, y) \dots\dots \quad (2.6)$$

$$\frac{\partial I(x,y)}{\partial y} = M2 \times I(x, y) \dots\dots \quad (2.7)$$

Le module m de gradient est définit par :

$$m = \sqrt{(\frac{\partial I(x,y)}{\partial x})^2 + (\frac{\partial I(x,y)}{\partial y})^2} \dots\dots \quad (2.8)$$

Tandis que sa direction ϕ est définit par :

$$\phi = \arctan\left(\frac{\partial I(x,y)}{\partial y} \mid \frac{\partial I(x,y)}{\partial x}\right) \dots\dots \quad (2.9)$$

La direction du gradient maximise la dérivée directionnelle.

La dérivée de I(x,y) dans une direction donnée d s'écrit :

$$I(x, y) * d. \dots\dots\dots (2.10)$$

Voici une illustration montrant le gradient d'une image découpé en deux zones distinctes :

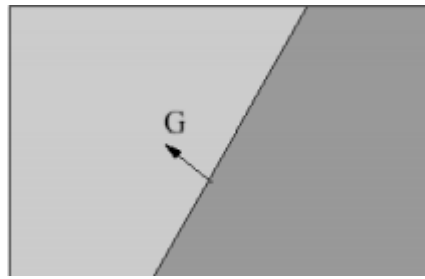


figure 2- 5 :Gradient d'une image à 2 régions [20]

Il existe plusieurs opérations permettant l'approximation du gradient, on peut citer [17] :

❖ Masque de Prewitt

$$M1=1/3 \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \text{ et } M2=1/3 \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \dots\dots(2.11)$$

❖ Masque de Sobel

$$M1=1/4 \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \text{ et } M2=1/4 \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \dots\dots(2.12)$$

B. Détection de contours par l'approche Laplacien :

Le laplacien d'une image d'intensité I(x, y) est défini par :

$$\nabla^2 I(x, y) = \frac{\partial^2 I(x, y)}{\partial x^2} + \frac{\partial^2 I(x, y)}{\partial y^2} \dots\dots\dots (2.13)$$

L'approximation des deux dérivées est donnée par la relation suivante :

$$\nabla^2 I(x, y) = I(x+1, y) + I(x-1, y) + I(x, y-1) + I(x, y+1) - 4I(x, y) \dots\dots\dots (2.14)$$

☞ Invariant aux rotations de l'image.

☞ Le laplacien est souvent utilisé en amélioration d'images pour accentuer l'effet de contour :

❖ *Les méthodes déformables :*

Une méthode déformable est une courbe (en 2D) ou une surface (en 3D) définie dans le domaine de l'image et qui se met sous l'influence de forces externes extraites de l'image et de forces internes

décrivant des propriétés du modèle déformable lui-même. Ces forces sont définies de façon à s'adapter au mieux aux contours de la structure à segmenter.

❖ *Contours actifs :*

Les contours actifs tirent leur origine des modèles élastiques [21], mais la communauté s'accorde à les attribuer à l'équipe Kass, Witkin et Terzopoulos [22] qui introduisirent les snakes ou courbes minimisant. Les snakes tiennent leur nom de leur aptitude à se déformer comme des serpents. Depuis la publication de cette équipe, les modèles déformables sont devenus un sujet très important pour la communauté du traitement d'images. De très nombreuses équipes s'y sont intéressées de manière plus ou moins approfondie

Les contours actifs sont définis par une courbe paramétrique pouvant être fermée ou non. Un snake consiste à placer aux alentours de la forme à détecter une ligne initiale de contour. Cette ligne va se déformer progressivement selon l'action de plusieurs forces qui vont la tirer ou la pousser vers la forme. Ces forces sont représentées par trois énergies associées au snake [23] :

1. Une énergie propre, due uniquement à la forme du contour, dite énergie interne.
2. Une énergie potentielle imposée par l'image dite énergie externe. C'est elle qui va attirer la ligne du snake vers les contours réels présents sur l'image
3. Une énergie de contexte qui exprime certaines contraintes supplémentaires qui peuvent être imposées par l'utilisateur vu le snake qu'il veut obtenir.

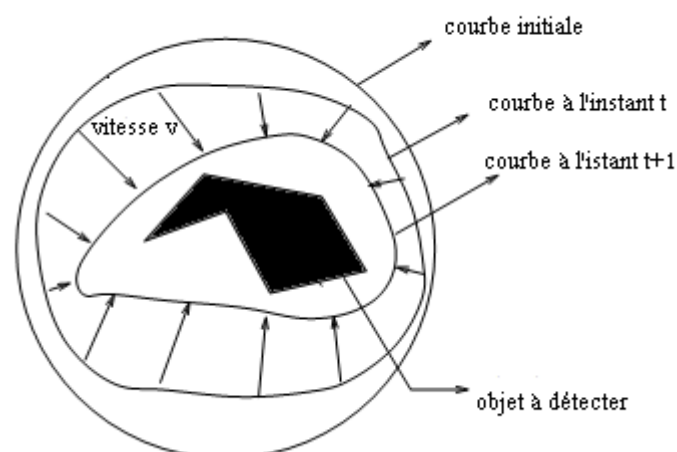


figure 2- 6 :Principe des contours actifs [23]

Soit $v(s,t)$ la position d'un point de la courbe à un instant t , et x, y les coordonnées Cartésiennes d'un point de l'image

$$v(s, t) = (x(s, t), y(s, t)) \quad \forall (s, t) \in [0, 1] * [0, +\infty[\quad (2.15)$$

Nous allons maintenant examiner les différentes énergies mises en jeu

▪ ***Les énergies :***

La fonctionnelle d'énergie attachée au contour actif est composée de trois énergies :

$$E(v): v = E_{interne}(v) + E_{externe}(v) + E_{contexte}(v) \quad (2.16)$$

Nous allons, par la suite, préciser ces trois types d'énergies, qui sont :

▪ ***L'énergie interne :***

Les énergies internes gèrent la cohérence de la courbe. Elles définissent la raideur de la courbe et la cohésion des points. Alors, elle est intrinsèque au snake. L'énergie interne est calculée à partir de deux forces appelées continuité et courbure

$$E_{interne} = E_{continuité} + E_{courbure} \quad (2.17)$$

$$E_{interne}(v) = \int_0^1 \left(\frac{\alpha}{2}(s) \|(s)\dot{v}\|^2 + \frac{\beta}{2}(s) \|(s)\ddot{v}\|^2 \right) ds \quad (2.18)$$

La force de continuité influe sur le rayon de courbure du contour en conduisant les points du contour à se positionner de manière à être équidistants. La deuxième force utilisée pour l'énergie interne est celle de courbure. Son but est d'éviter que le contour contienne des points isolés qui ne seraient pas cohérents avec la forme

▪ ***Energie externe :***

Les énergies externes prennent en compte les caractéristiques des images traitées. Parmi les énergies externes existantes, nous citons l'énergie de gradient (la dérivée première de l'image). Cette énergie externe est d'une importance première pour la détection du contour. En effet, un contour est généralement caractérisé par une forte différence entre les valeurs de plusieurs pixels.

$$F_{ballon}(v(s)) = k\vec{n}(s) \quad (2.19)$$

Où $\vec{n}(s)$ est un vecteur unitaire normal à la courbe en $v(s)$. L'intensité de l'énergie ballon est un scalaire généralement négatif (expansion du ballon) proportionnel à l'aire intérieure du contour.

L'image suivante illustre la progression du contour actif.

✓ Les méthodes analytiques :**A.Approche de Canny:**

Dans l'industrie, la technique de détection de bord Canny est l'une des détections de bord standard techniques. elle a été créé pour la première fois par John Canny pour sa thèse de maîtrise au MIT en 1983, et toujours surpasse la plupart des nouveaux algorithmes qui ont été développés. Pour rechercher des arêtes par séparer le bruit de l'image avant de trouver les bords de l'image, le Canny est un élément très important méthode. La méthode Canny est une meilleure méthode sans perturber les caractéristiques des arêtes dans le image ensuite en appliquant la tendance à trouver les bords et la valeur sérieuse du seuil .Les étapes algorithmiques sont les suivantes:

- Convolutionnez l'image $f(r, c)$ avec une fonction gaussienne pour obtenir une image .
- Appliquer le premier opérateur de gradient de différence pour calculer la force de l'arête puis la magnitude de l'arête et la direction est obtenue comme avant.
- Appliquer une suppression non maximale ou critique à la magnitude du gradient.
- Appliquer un seuil à l'image de suppression non maximale.
- Contrairement à Roberts et Sobel, l'opération Canny n'est pas très sensible au bruit. Si le Canny le détecteur fonctionnait bien ce serait supérieur[24].

B.Approche de Deriche:

Deriche a utilisé une méthode d'optimisation similaire à celle de Canny afin de définir un filtre de réponse impulsionnelle infinie [Der87, Der90]. La modification proposée, qui consiste à étendre à l'infini les bornes d'intégration dans l'expression des critères de performance, amène à une solution optimale similaire à celle de Canny.

2.4.3Approche basée classification de pixel

La classification d'images constitue une étape nécessaire et importante pour diverses applications telles que la quantification d'images et la segmentation d'images. Elle consiste à déterminer une classification des pixels dans l'espace des luminances, en utilisant les niveaux de gris présents dans l'image. Suite à la phase de classification, le niveau de gris moyen de chaque classe est affecté à tous les pixels de cette classe et chaque composante connexe d'une même classe constitue un objet La segmentation par classification de pixels implique la recherche des pixels d'image qui sont assez similaires pour être groupés dans une même classe. Cette approche utilise généralement les

algorithmes de la classification des données pour classifier les pixels d'une image en différentes classes tout en permettant l'optimisation d'un certain critère d'homogénéité.

✓ *Méthodes classification supervisées*

La classification supervisée repose sur des exemples de motifs ou de vecteurs de caractéristiques qui ont déjà été affecté à une classe définie. En utilisant un échantillon de vecteurs de caractéristiques comme notre donnée de formation, nous concevons un système de classification avec l'intention et l'espoir que de nouveaux les exemples de vecteurs de caractéristiques qui n'ont pas été utilisés dans la conception seront ensuite classés avec précision. En classification supervisée, il s'agit donc d'utiliser des exemples de formation pour concevoir un classificateur qui se généralise bien aux nouveaux exemples. En revanche, classification non supervisée ne repose pas sur la possession d'exemples existants d'une classe de modèles connue. Les exemples ne sont pas étiquetés et nous cherchons à identifier les groupes directement dans le corpus global de données et caractéristiques qui nous permettent de distinguer un groupe d'un autre. Les techniques de clustering sont un exemple de classification non supervisée a classification est automatisée, comme (k-means, c-means).

Dans la classification supervisée, le nombre de classes est connu et on dispose d'un ensemble de pixels déjà étiquetés, servant d'ensemble d'apprentissage. Il s'agit alors de pouvoir associer chaque nouveau pixel à la classe la plus adaptée en se servant des pixels déjà étiquetés [17].

A.Les réseaux de neurones :

Un « neurone formel » (ou simplement neurone est une fonction algébrique non linéaire et bornée, dont la valeur dépend de paramètres appelés coefficients ou poids .les variables de cette fonction sont habituellement appelées entrées du neurone ,et la valeur de la fonction est appelées sa sortie[25]la période d'apprentissage consiste à calculer les différents coefficients ,de sorte à ce que les sorties du réseaux soient égales aux valeurs souhaitées a un seuil d'erreur près pour les exemples utilisée Les réseaux de neurones sont habituellement utilisés en imagerie médicale comme classificateurs, ayant pour entrées les images médicales à segmenter et pour sortie les différentes classes de régions possibles sur ces images .on utilise un certain nombre d'exemples manuellement segmentés pour la phase d'apprentissage du réseaux

Cette période d'apprentissage est le principal inconvénient de la méthode ,car l'efficacité du système sera proportionnelle au nombre d'exemples manuellement traités, ainsi qu'a leur diversité

✓ La classification non supervisée :

Dans la classification non supervisée, aucune information sur le nombre et le contenu des classes possible n'est fourni. L'objectif est alors de pouvoir regrouper automatiquement des pixels considérés similaires dans une même classe. Dans ce cas il s'agira de définir une fonction de similarité entre pixels qui sera maximum entre les pixels d'une même classe et minimum avec ceux des autres classes. [26]

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

B. L'algorithme des K-moyennes

L'algorithme des K-moyennes (en anglais K-means) est une méthode de classification automatique, qui a pour objectif le partitionnement des données d'une image (pixels) en K classes (ou clusters), chaque pixel étant caractérisé par un vecteur de l attributs ,A partir d'une variance intra classe et en maximisant l'écartement entre les classes les étapes principale de l'algorithme k-moyennes sont :

1. Choix aléatoire de la position initiale des K classes;
2. (Ré-) affecter les objets a une classe suivant un critère de minimisation des distances;
3. Une fois tous les objets placés ,recalculer les K centroides
4. Réitérer les étapes 2 et 3 jusqu'a ce qu'aucune réaffectation ne soit faite.

La classification des éléments de l'image se fait manière itérative en alternant l'étape de classification et l'étape de mise à jour des centres ,jusqu'à la stabilisation de la segmentation[27]

2.4.4 Les méthodes de segmentation par seuillage

Le seuillage a pour objectif de segmenter une image en plusieurs classes en n'utilisant que l'histogramme. On suppose donc que l'information associée à l'image permet à elle seule la segmentation, i.e. qu'une classe est caractérisée par sa distribution de niveaux de gris. A chaque pic de l'histogramme est associée une classe. [28].

L'idée de base de l'utilisation du seuil d'intensité dans la segmentation est très simple. Nous choisissons une certaine valeur de seuil telle que les pixels possédant des valeurs supérieures au seuil soient affectées à une région tandis que celles qui tombent en dessous du seuil sont affectées à une autre (Adjointe) région.

Seuiller une image consiste à définir une valeur d'intensité $\sigma \in E$ qui va séparer l'ensemble des pixels en deux régions R_0 et R_1 telles que

$$b(x, y) = \begin{cases} 1 & \text{if } I(x, y) > T \\ 0 & \text{otherwise} \end{cases} \quad (2.20)$$

Cette définition peut bien entendu être étendue au multi-seuillage. Quoiqu'il en soit, cette définition binaire suffit à présenter les principes de seuillages existants. La manière la plus simple de définir le seuil σ est de procéder de façon empirique ou en se basant sur les informations a-priori de l'image. Cette utilisation est pertinente dans le cas où l'on cherche à segmenter des objets dont on connaît l'intensité et qui sont les seuls à avoir cette intensité. Cela peut par exemple être le cas pour des images tomodynamométriques où l'on cherche à isoler les objets d'une densité connue.

Nous utiliserons cette approche dans le cadre de la segmentation des nœuds.

Cependant, plusieurs techniques ont été mises au point afin de déterminer automatiquement un seuil pertinent.

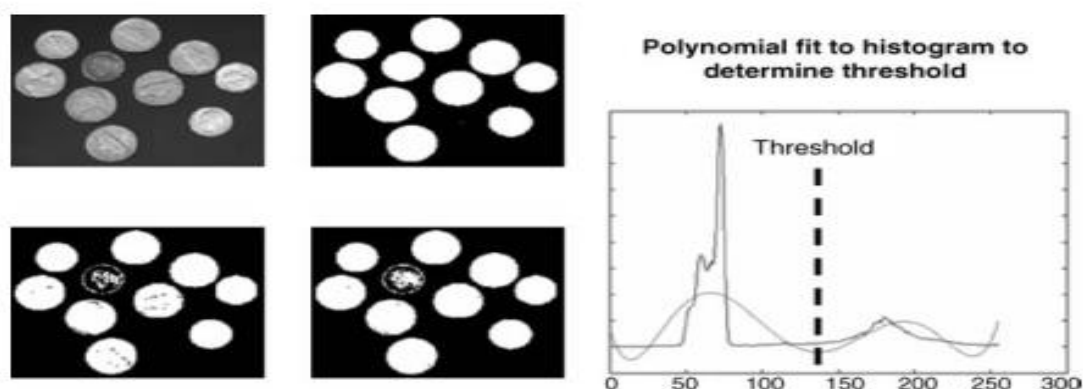


figure 2- 7 :Figure L'image de droite montre le histogramme[28]

2.5 Segmentation sémantique d'images

La tâche de segmentation sémantique d'image consiste à classer chaque pixel d'une image dans une instance, où chaque instance correspond à une classe. Cette tâche fait partie du concept de la compréhension de la scène ou de mieux expliquer le contexte global d'une image.

Dans le Domaine d'analyse d'images médicales, la segmentation d'images peut être utilisée pour les interventions guidées par image, radiothérapie ou diagnostic radiologique amélioré.



figure 2- 8 : Image et étiquette des pixels

2.6 Segmentation d'images médicales :

Le but de toute méthode de segmentation est l'extraction d'attributs caractérisant ces entités. Quel que soit l'origine de l'image. La diversité des images, la difficulté des problèmes, l'évolution de la puissance de calcul des ordinateurs ont conduit à l'introduction d'une multitude d'algorithmes de segmentation. La segmentation et l'annotation automatique de structures dans les images médicales sont des tâches essentielles à une multitude d'applications clés dont le diagnostic assisté, le suivi de pathologies et la recherche clinique. Bien qu'important, le processus de segmentation est très complexe et qui est dû notamment au faible contraste, à la superposition des régions d'intérêt et au bruit. Des approches efficaces de segmentation ont été proposées au fil des années, la plupart de ces approches requièrent l'intervention d'un humain pour étiqueter manuellement

les régions d'intérêt dans l'image ou le volume à segmenter.

Une segmentation des volumes anatomiques normaux ou pathologiques fiable et précise issus des systèmes d'imagerie est un objectif important en traitement de l'information médicale, elle constitue le premier maillon de la chaîne d'analyse qui aboutit à l'étude de la morphologie des structures internes, à la détection et quantification des lésions. De nombreuses méthodes de segmentation ont été proposées pour ce type d'images.

2.6.1 Etat de l'art :

La tâche de segmentation d'image sémantique consiste à classer chaque pixel d'une image dans une instance, où chaque instance correspond à une classe. Cette tâche fait partie du concept de la compréhension de la scène ou de mieux expliqué le contexte global d'une image. Dans le Domaine d'analyse d'image médical, la segmentation d'images peut être utilisée pour les interventions guidées par image, radiothérapie ou diagnostic radiologique amélioré.

On trouve dans la littérature plusieurs chercheurs qui ont été intéressés par ce domaine crucial :

Kumar et al.[29]. Proposé une méthode de segmentation basée sur un ensemble de dix réseaux U-net appelés Algorithme multi-u-net. Un projet de vote majoritaire s'agrandit les réponses de tous les modèles U-Net. Le but est de réduire l'incertitude de trouver les minima associés à l'initialisation aléatoire de chaque réseau. Cette méthode a atteint un score F1 (F1S) de 0,82.

Chiao et al.[30] Utilisé une extension du R-CNN plus rapide pour la segmentation tumorale appelée masque R-CNN. Ce CNN remplace la mise en commun de ROI de R-CNN plus rapide avec un alignement de retour sur investissement afin d'éviter de perdre des informations spatiales. Le modèle de masque R-CNN a obtenu une intersection sur l'union (IOU) de 0,75

Yap et al.[31]. Comparé trois modèles CNN pour la détection de la tumeur, nommément: Lenet, U-Net et FCNALEXnet. Ces réseaux ont ensuite été comparés contre quatre méthodes traditionnelles qui utilisent des fonctionnalités et des règles conçu manuellement. FCN-Alexnet a obtenu les meilleures performances avec F1S = 0:92 et tous les CNNs surperformé les méthodes traditionnelles.

Dans un autre travail, YAP et al.[31] Évalué quatre réseaux entièrement convolutionnels (FCNS) pour traiter le Problème de la segmentation de la tumeur. Les CNN évalués étaient FCN-Alexnet, FCN-8S, FCN-16S et FCN-32s. Les résultats expérimentaux montrent que les FCN-16 ont obtenu la meilleure performance des tumeurs bénignes avec F1S = 0:76, alors que les FCN-8 ont atteint F1S = 0:54, qui était le score le plus élevé dans les tumeurs malignes.

Hu et al. [49]présenté le réseau entièrement convolutionnel dilaté (DFCN), suivi d'une phase basée sur la phase Contour actif (PBAC) pour la segmentation tumorale. La convolution dilatée est introduite pour faire face à la mouche Bruit et ombres typiques des images de bus. PBAC est une méthode de post-traitement pour affiner la sortie DFCN.

Fang et al[31]. introduit une autre méthode de segmentation utilisant la méthode PBAC pour affiner la Sorties M-Net Model. Ce CNN comprend une entrée et une sortie multi-échelles, avec un convolutionnel en forme de V réseau. Les résultats expérimentaux montrent que M-Net + PBAC a surperformé à la fois des modèles U-Net et V-Net En atteignant F1S = 0:96.

Zhuang et al[31]. Proposé le porte-gate de dilatation résiduelle-Unut (RDAU-NET) pour la segmentation tumorale, qui est basé sur l'architecture U-net. RDAU-Filet intègre des éléments supplémentaires (par exemple, résiduel des unités, des convolutions dilatées et des modules de porte d'attention) pour améliorer ses capacités d'apprentissage. Rdau-net a été comparé à d'autres réseaux profonds, tels que FCN-8S, FCN-16, SEGNET et huit modèles basés sur U-Net.

Tao et al[32]. Proposé un réseau profond pour la détection de la tumeur, nommé modèle en une étape, qui comprend Les trois composants clés suivants: base-filet, seg-filet et CLS-Net basé sur

la boîte d'ancrage. Les deux premiers Les composants sont utilisés pour détecter des régions tumorales possibles, alors que le dernier composant est utilisé pour supprimer False régions tumorales.

Amiri et al. [33] a développé une méthode de segmentation en deux étapes basée sur l'architecture U-Net. La première étape détecte la région tumorale dans l'image de bus, alors que les segments de la deuxième étape sont détectés région pour obtenir la forme de la tumeur. Pour améliorer les résultats de la segmentation, augmentations de données de temps de test (changement d'image) avec la technique d'abandon scolaire est effectué. Ce processus est exécuté plusieurs fois, de sorte que le La règle suivante est définie: lorsque le coefficient de dés entre les différentes exécutions est faible, le réseau de détection est considéré comme valide; Sinon, il est considéré comme invalide et nécessite une validation supplémentaire.

Tableau2- 1 : Interrogé les approches basées sur le CNN pour la détection et la segmentation des tumeurs du sein

CNN model	Tâch	Performance	REF
Multi U-Net	Segmentation	F1s = 0.82	[29]
FCN- AlexNet	Detection	F1s = 0.92	[30]
FCN-16s	Segmentation	F1s = 0.76	[31]
DFCN	Segmentation	F1s = 0.89	[32]
M-Net	Segmentation	F1s = 0.96	[33]
SSD	Detection	F1s= 0:79	[34]
RDAU-Net	Segmentation	F1s = 0:84	[35]
VGG19	Detection	DR = 48%	[36]
Mask R-CNN	Segmentation	IoU = 0:75	[37]
Faster R-CNN	Detection	CL = 85%	[38]
One-step model	Detection	F1s = 0:90	[39]
Faster R-CNN	Detection	IoU = 0:83	[40]
WGAN-RDA-UNet	Segmentation	F1s = 0:89	[41]
U-Net	Segmentation	F1s = 0:86	[42]
DsUnet	Segmentation	F1s = 0:86	[43]

2.6 Conclusion :

Dans ce chapitre, nous avons vus des concepts généraux et de la segmentation des images, Nous avons faits un survol rapide sur les différentes techniques de la segmentation en générales. Ensuite, nous avons répertorié quelques méthodes utilisées dans le domaine de l'imagerie médicale .

Chapitre 3:Deep learning

3.1 Introduction :

L'apprentissage profond est dérivé de la recherche sur le réseau neuronal artificiel. L'une des structures d'apprentissage de profondeur communes est une perception multicouche contenant plusieurs cachées. Le mode de travail de base de l'apprentissage profond consiste à former une caractéristique plus abstraite de haut niveau en analysant et en combinant des fonctionnalités de bas niveau, dérivant et mappage, pour montrer la catégorie d'attributs, de sorte que les caractéristiques des données ayant une représentation de fonctionnalités distribuées se trouvent à avoir une identification caractéristique des données. L'apprentissage approfondi est devenu un nouveau domaine de la recherche sur l'intelligence artificielle et ses objectifs de recherche sont d'établir une représentation informatique d'un ordinateur similaire au cerveau humain pour analyser l'apprentissage, c'est-à-dire un réseau neuronal, qui imite la manière d'étudier les choses et Mécanismes pour expliquer les données, telles que le texte de la langue, l'audio, la photographie et plus encore. L'introduction de la technologie d'apprentissage profonde est une tentative dans le domaine de la segmentation de l'image médicale. Il a un grand nombre de données à l'avance pour améliorer la précision de la segmentation. Ce nouveau type de recherche et de mise en œuvre a des prospects généraux, qui peuvent être utilisés pour corriger les erreurs dans certains segments.

L'apprentissage en profondeur est un terme utilisé pour un sous-ensemble de modes d'apprentissage automatique qui utilisent plusieurs paramètres d'apprentissage de plusieurs couches, ce qui permet de créer une chaîne de fonctionnalités simples pouvant exprimer des concepts complexes. Les premières couches apprennent des fonctionnalités simples et la représentation devient peu à peu plus complexe. Au fur et à mesure que nous nous rapprochons de la couche de sortie, une chose qui rend le deep Learning intéressant est que nous ne disons pas au modèle quelles fonctionnalités extraire à chaque couche, tout ce que nous devons faire est de fournir des paires d'entrée et de sortie. Il est donc utile de savoir comment fonctionne l'apprentissage en profondeur et comment en tirer le meilleur parti, c'est ce que nous allons faire dans ce chapitre, nous allons présenter les principes de base de l'apprentissage automatique en mettant l'accent sur l'apprentissage en profondeur.

Dans le deep learning, les grands réseaux de neurones artificiels sont alimentés par des algorithmes d'apprentissage et des quantités croissantes de données, améliorant continuellement leur capacité à "penser" et à "apprendre au fur et à mesure qu'ils traitent les données. "Deep" fait référence aux nombreuses couches que le réseau neural accumule dans le temps afin d'améliorer sa performance à mesure qu'il progresse dans son apprentissage. Alors que l'essentiel du deep learning

s'effectue actuellement sous supervision humaine, l'objectif est de créer des réseaux neuraux capables de s'entraîner seuls et d'« apprendre » en toute autonomie.

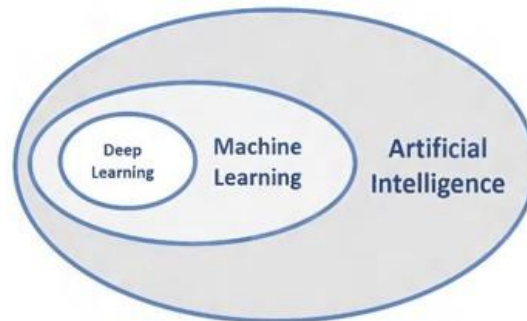


Figure 3- 1: La relation entre intelligence artificielle, Machine Learning et Deep Learning[34]

3.2 L'apprentissage en profondeur (Deep Learning) :

3.2.1 Définition :

L'apprentissage en profondeur (Deep Learning : DL) est une branche de l'apprentissage automatique qui enseigne aux ordinateurs à faire ce qui vient naturellement aux humains : apprendre de l'expérience. Les algorithmes d'apprentissage automatique utilisent des méthodes de calcul pour "apprendre" des informations directement à partir de données sans s'appuyer sur une équation prédéterminée comme modèle. DL est une classe de techniques d'apprentissage automatique qui exploitent de nombreuses couches de traitement d'informations non linéaires pour l'extraction et la transformation de fonctions supervisées ou non supervisées, et pour l'analyse et la classification de modèles. Les caractéristiques et les concepts de niveau supérieur sont donc définis en termes de niveaux inférieurs, et une telle hiérarchie de caractéristiques est appelée architecture profonde. La plupart de ces modèles sont basés sur un apprentissage non supervisé des représentations [35].

3.2.2 Pourquoi l'apprentissage profond (deep-learning) ?

L'apprentissage profond (deep-learning) avec leur capacité remarquable de dériver une signification des données compliquées ou imprécises, peut être utilisé pour extraire des modèles et détecter les tendances trop complexes pour être remarquées par l'homme ou techniques informatiques. Un réseau de neurones qualifiés peut être considéré comme un "expert" dans la catégorie des informations qu'il a été donné à analyser. Cet expert peut alors être utilisé pour fournir des projections nouvelles des situations d'intérêt et de réponse.

Les autres avantages incluent

1. Stockage des informations sur l'ensemble du réseau: informations tels que dans la programmation traditionnelle est stocké sur l'ensemble réseau, pas sur une base de données. [36].
2. Capacité à travailler avec des connaissances incomplètes: après formation, les données peuvent produire une sortie même avec des informations incomplètes.
3. Avoir la tolérance aux pannes : la corruption d'une ou plusieurs cellule.
4. Avoir une mémoire distribuée : pour que Ann soit capable d'apprendre, il est nécessaire de déterminer les exemples et d'enseigner au réseau en fonction de la sortie souhaitée en montrant ces exemples au réseau. Le succès du réseau est Directement proportionnelle aux instances sélectionnées et si l'événement ne peut pas être montré au réseau dans tous ses aspects, le réseau peut produire une fausse sortie
5. Capacité à faire apprendre à machine : les réseaux Neurones artificiels apprennent des événements et prennent des décisions en commentant les événements similaires.
6. Capacité de traitement parallèle: réseaux de neurones artificiels avoir une force numérique qui peut effectuer plus d'un emploi en même temps

3.3 Réseaux de neurones:

3.3.1 Réseaux de neurones (Modèle biologique) :

Les neurones reçoivent des signaux (impulsions électriques) par les dendrites et envoient l'information par les axones. Les contacts entre deux neurones (entre axone et dendrite) se font par l'intermédiaire des synapses. Les signaux n'opèrent pas de manière linéaire [31].

Schématiquement, on peut décomposer le neurone en 3 grandes entités :

- Un **corps cellulaire**, appelé péricaryon.
- Un **ensemble de dendrites** (de l'ordre de 7 000).
- Un **axone**.

Lorsque l'excitation du corps cellulaire dépasse un certain seuil, un afflux nerveux est émis le long de l'axone de façon centrifuge, c'est-à-dire du corps cellulaire à l'extrémité de l'axone.

La quantification de l'excitation se fait par **modulation de fréquence**. Plus le neurone est excité, plus rapprochées sont les impulsions dans l'axone [37].

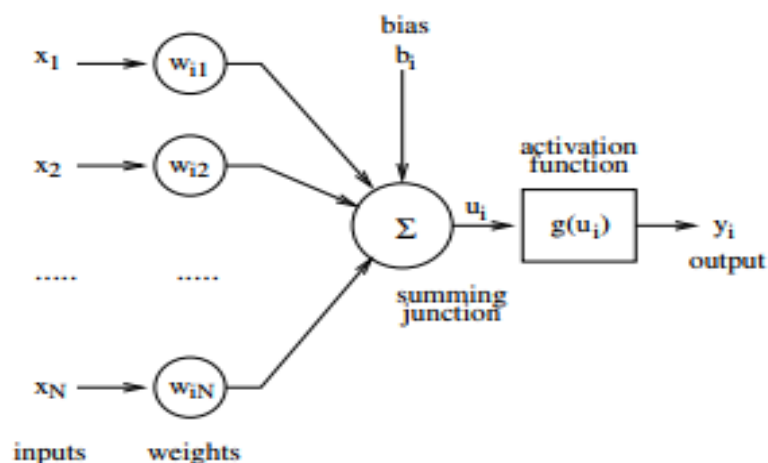


Figure 3- 2:Neurones artificiels. la forme la plus générale de les neurones[34]

3.3.2 Le perceptron :

Principes :

- pas de notion temporelle
- coefficient synaptique : coefficient réel
- sommation des signaux arrivant au neurone
- sortie obtenue après application d'une fonction de transfert[37]

Un neurone formel est composé :

- Des **entrées**, notées \mathbf{x} sous formes de vecteurs, représentant les dendrites
- Une **sortie**, notée \mathbf{y} , représentant l'axone
- Des **paramètres**, notés \mathbf{W} et \mathbf{b} , influençant le fonctionnement du neurone.

Le réseau neuronal est un modèle mathématique ou un modèle de calcul pour imiter la structure et la fonction du réseau neuronal biologique (système nerveux animal, en particulier cerveau), est utilisé pour estimer ou approcher la fonction .

les réseaux de neurones sont un grand ensemble d'unités différemment interconnectées Les entrées du neurone sont variables, tandis que les paramètres sont fixés lors de la construction du modèle, et la sortie est calculée en fonction des entrées et des paramètres. Équation d'un neurone formel

$$\hat{y} = f((W, x) + b) \quad (3.1)$$

Des **paramètres**, notés \mathbf{W} et \mathbf{b} , influençant le fonctionnement du neurone.

Il appliquera ensuite une fonction f à z , donnant la sortie \hat{y} .

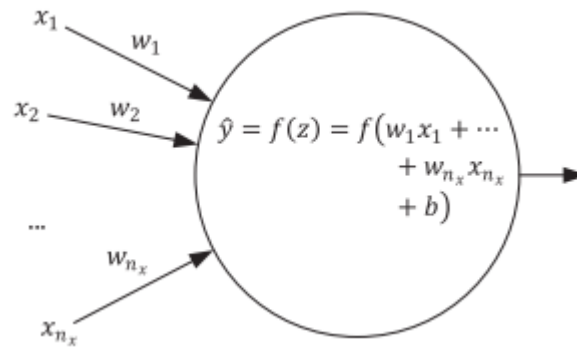


Figure 3- 3: illustration de base du neurone:

Les entrées ne sont pas mises dans une bulle. C'est simplement pour les distinguer à partir de nœuds qui effectuent un calcul réel.

- Les noms des poids sont écrits le long de la flèche. Cela signifie qu' avant de passer les entrées à la bulle centrale (ou nœud), l'entrée d'abord sera multiplié par le poids relatif, comme indiqué sur la flèche.

La première entrée, x_1 , sera multipliée par w_1 , x_2 , par w_2 , etc.

- La bulle centrale (ou nœud) effectuera plusieurs calculs en même temps. Tout d'abord, il additionnera les entrées (les $w_i x_i$ pour $i = 1, 2, \dots, n_x$), puis additionnez au résultat le biais b et, enfin, appliquez au résultat valoriser la fonction d'activation

les formules que nous utiliserons plus tard, pour multiplier x et w , nous utiliserons de multiplication matricielle, et, par conséquent, nous écrivons

$$w^T x = (w_1 \dots w_{n_x}) \begin{pmatrix} x_1 \\ \vdots \\ x_{n_x} \end{pmatrix} = w_1 x_1 + w_2 x_2 + \dots + w_{N_x} x_{n_x} \quad (3.2)$$

où w^T indique la transposition de w . z peut alors être écrit avec cette notation vectorielle comme :

$$z = w^T x + b \quad (3.3)$$

$$\text{et la sortie du neurone } \hat{y} \text{ comme : } \hat{y} = f(z) = f(w^T x + b) \quad (3.4)$$

$\hat{y} \rightarrow$ sortie neurone

$f(z)$ fonction d'activation (ou fonction de transfert) appliquée à z

$w \rightarrow$ poids (vecteur avec n_x composantes

- $b \rightarrow$ bias

Les réseaux de neurones artificiels (ANN) sont des machines d'apprentissage de la représentation introduites dans la fin des années 1950 pour apprendre des concepts complexes . ANN transforme une instance / exemple d'entrée) donnée (un vecteur D -dimensionnel des caractéristiques / covariats) dans une nouvelle représentation pouvant être utilisée dans le contexte de la classification ou de la régression ou d'autres paradigmes d'apprentissage. Ces transformations sont non linéaires et

éventuellement non convexes, et calculé en premier calcul une projection affine de l'entrée, suivie d'une fonction d'activation non linéaire monotone sur ces projections (qui peut ne pas être nécessairement point sage[38]). Les fonctionnalités résultantes correspondent à certaines représentations de niveau supérieur et abstraite des intrants.

3.3.3 Les réseaux de neurones en couches :

Dans la plupart des cas, le réseau neuronal artificiel peut modifier la structure interne sur la base de l'entrée ou de la stimulation des informations, qui est une structure système pouvant être adaptative . Le réseau neuronal moderne est un outil de modélisation de données basé sur des statistiques non linéaires et la figure 3.4 est un réseau de neurones de connexion complet schématique. Les réseaux de neurones typiques ont généralement les trois parties suivantes

1. Architecture: la structure spécifie les variables du réseau et de leurs topologies. Par exemple, le poids de la connexion neuronale dans le réseau neuronal et la valeur d'excitation des neurones peuvent être des variables du réseau neuronal.

2. Règle d'activité: la plupart des modèles de réseau neuronal ont une règle cinétique de courte durée appelée fonction d'incitation, également appelée fonction d'activation, utilisée pour convertir l'entrée d'une neurone à certaines valeurs spécifiques, telles que des types numériques continus sont convertis en $[0, 1]$ Sortie, c'est-à-dire définir la manière dont les neurones changent leurs valeurs d'excitation conformément à d'autres neurones. La fonction d'incitation générale dépend du poids du réseau (c'est-à-dire les paramètres du réseau).

3. Règle d'apprentissage: la règle d'étude spécifie comment les poids sont ajustés au fil du temps. Ceci est généralement considéré comme une règle cinétique à longue échelle. En général, les règles d'apprentissage dépendent de la valeur d'excitation des neurones. Il peut également s'appuyer sur la valeur cible fournie par le superviseur et la valeur du poids actuel.

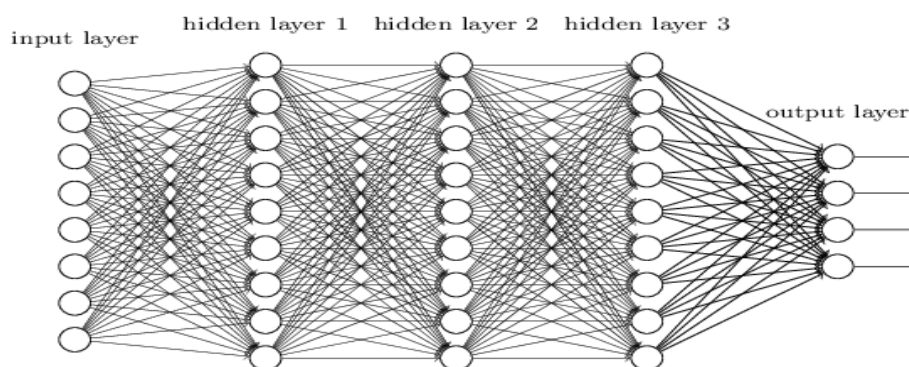


Figure 3- 4:réseaux de Neurone

Cette caractéristique de ce réseau neuronal est que chaque couche de neurone est supérieure connecté à chaque neurone de la couche suivante avec une entrée et une sortie.

La construction de réseaux de neurones est causée par l'inspiration de la fonction de la fonction du réseaux neurone biologique (humain ou d'un animal). Le réseau neuronal artificiel est généralement optimisé par une méthode d'apprentissage basée sur des types statistiques mathématiques et le réseau neuronal artificiel est également une application réelle de méthodes statistiques mathématiques, mais le réseau neuronal traditionnel a un poids excessif, une convergence lente.

3.4 Introduction du réseau de neurones convolutifs (CNN)

Un réseau de neurones convolutifs (CNN) est un réseau neuronal comportant une ou plusieurs couches utilisées principalement pour le traitement, la classification, la segmentation, la segmentation et les autres données corrélées automatiquement.

Une convolution est essentiellement glisser un filtre sur l'entrée une façon utile de penser à des convolutions est cette citation du Dr Prasad Samarakoon: "Une convolution peut être pensée comme" en regardant d'une fonction de prédiction de son résultat "

Plutôt que de chercher certaines fonctionnalités, il peut être plus efficace de regarder des portions plus petites de l'image.

3.4.1 Architecture de réseaux de neurone convolutifs CNN

Dans l'apprentissage profond, un réseau de neurones convolutionnels (CNN) est une classe de réseaux de neurones profonds, le plus souvent appliqué pour analyser les images visuelles. Maintenant, lorsque nous pensons à un réseau de neurones, nous pensons aux multiplications matricielles, Il utilise une technique spéciale appelée convolution. Maintenant, dans la convolution mathématique est une opération mathématique sur deux fonctions qui produisent une troisième fonction qui exprime la manière dont la forme d'une est modifiée par l'autre.

Il existe quatre types de couches: la couche de convolution, la couche de pooling, la couche de correction ReLU et la couche fully- connected.

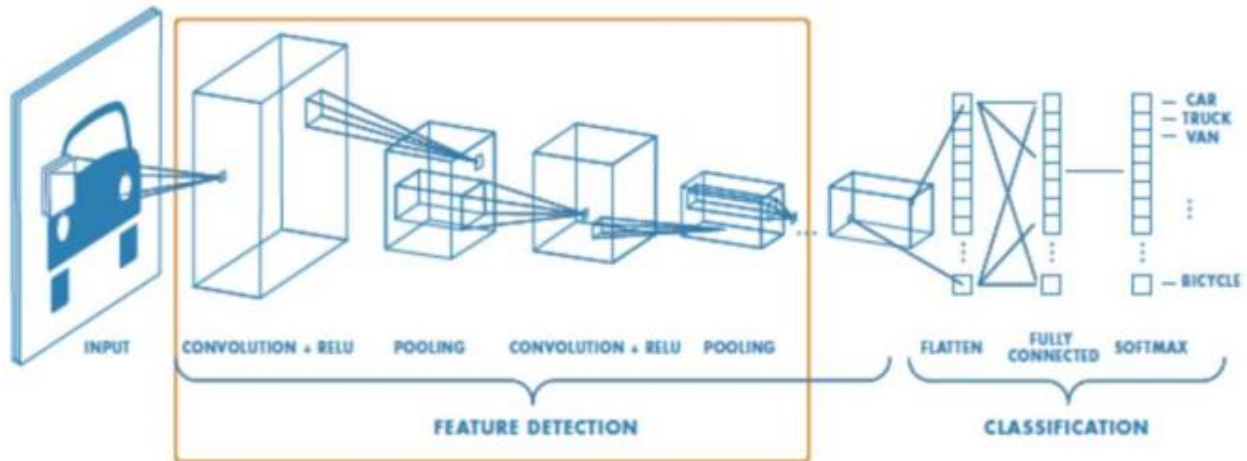


Figure 3- 5: Les couches d'un réseau de neurones convolutifs [41]

Une architecture CNN est formée de couches de traitement indépendantes :

1) Couche de convolution :

Dans l'analyse fonctionnelle, la convolution est un opérateur mathématique de la troisième fonction par deux fonctions F et G , la surface de la fonction symbolique F et le bord incurvé de la fonction de trajet de la g . G . Si une fonction de participation à la convolution est considérée comme une fonction d'indication de l'intervalle, la convolution peut également être considérée comme une promotion "moyenne mobile". Soit $F(x)$, $g(x)$ est deux fonctions de volume sur \mathbb{R} , avec des points:

$$\int_{-\infty}^{\infty} f(\tau)g(x - \tau)d\tau \quad (2.1)$$

Dans un réseau neuronal artificiel, qui a suggéré l'apprentissage des filtres convolutifs les opérations de convolution ne sont pas utilisées dans le sens conventionnel, où le filtre déplace à droite après chaque multiplication. Il convient de souligner que ce paradigme d'activité de convolution (détecteurs de fonctionnalités indépendants) n'est pas entièrement acceptable pour les images enregistrées. Les images du visage, par exemple, nécessitent un certain nombre de détecteurs de fonctionnalités dans divers endroits spatiaux. Certains modèles ont été créés pour tenir compte de cela.[42] Envisagez uniquement des réseaux liés localement.

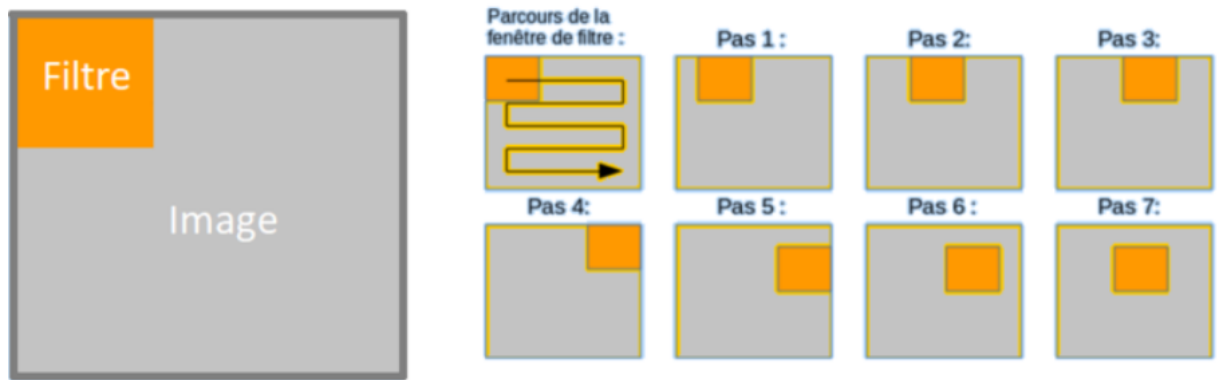


Figure 3- 6: opération d'un filtre sur une image[43]

La carte de caractéristiques correspond donc à toutes les sorties produites par les filtres et aura comme dimensions $W_{i+1}H_{i+1}C_{i+1}$, avec :

$$W_{i+1} = \frac{W_i - F + 2P}{S} + 1 \quad (3.5)$$

$$H_{i+1} = \frac{H_i - F + 2P}{S} + 1 \quad (3.6)$$

$$C_{i+1} = K \quad (3.7)$$

F est la taille du filtre, S est le pas de la convolution P est le *Padding* qui permet de contrôler la dimension spatiale de la carte en sortie,

K est le nombre de filtres utilisés.

Si les étapes et marges appliquées à l'image d'entrée permettent de contrôler le nombre de champs récepteurs (surfaces de traitement) à traiter, la profondeur peut donner le concept de volume de sortie, et elle peut avoir un volume identique à l'image Avec une profondeur de 3 pour les trois canaux RVB, la couche convolutive affichera également la profondeur à la sortie. C'est pourquoi nous ne parlons pas de «sortie» et d '«entrée», car l'entrée de la couche convolutive la sortie d'une autre couche convolutive.

L'opération de convolution peut être effectuée en utilisant l'équation suivante:

$$m_{ij} = b_i + \sum w_{ij} * x_j \quad (3.8)$$

L'organisation de la couche de convolution est basée sur le nombre de cartes de convolution et la taille des filtres. Par conséquent, la carte des caractéristiques est construite en ajoutant le biais avec la somme du produit de différentes caractéristiques de convolution bidimensionnelle.

2) Couche de pooling :

En général, après la couche de convolution, le réseau recevra une entrée représentative des fonctionnalités d'image, appelée carte de fonctionnalité, après avoir obtenu ces fonctionnalités, la prochaine étape consiste à poursuivre la classification. En théorie, les personnes peuvent utiliser toutes les caractéristiques de sortie convolutionnelles pour former notre classificateur, telles que l'utilisation du classificateur SoftMax, mais cela fera face à un grand nombre de paramètres incontrôlables. Par exemple, pour une image de 96×96 pixels, il est supposé que notre couche de convolution comporte 400 copies de volume, c'est-à-dire que la sortie est de 400 canaux, chaque volume du volume est 8×8 , puis chaque canal recevra une convolution Caractéristique de $(96-8 + 1) \times (96-8 + 1) = 89 \times 89$ -dimensionnelle, car il y a 400 canaux, chaque échantillon (exemple) obtient un vecteur de caractéristique convicenal de $89 \times 89 \times 400 = 3,168,400$ Vecteur de fonctionnalité de convicenal. Il est très difficile de former un classificateur comportant plus de 3 millions d'intrants de fonctionnalités, et en raison de paramètres excessifs, des paramètres sont sensibles à la classification, il est facile d'avoir un bon modèle pour des données spéciales. Pour effet de données générales, il est installé. (sur-adaptation).

Retour à l'image elle-même, nous savons qu'une paire d'images voit toujours la situation générale et, dans l'espace de l'espace, est associée, elle s'appelle des propriétés "statiques". C'est-à-dire que lorsque nous regardons une image paire, la fonctionnalité observée dans une zone d'image peut également apparaître dans une autre zone et, selon le contact, elle déduira naturellement le passé. Par conséquent, afin d'indiquer une grande caractéristique d'image, un naturellement souhaitable consiste à utiliser les statistiques de polymère à la raison et classer en fonction des caractéristiques de l'image. Par exemple, les personnes peuvent calculer une plage d'images d'une taille particulière pour obtenir une caractéristique spécifiée de la plage. Il est généralement utilisé pour calculer la valeur moyenne ou maximale de tous les points de pixels de la plage d'images. Ces statistiques récapitulatives ont non seulement perdu des caractéristiques d'image de la faible latitude, mais empêchent également les phénomènes ajustés d'améliorer les résultats de la classification. Ceci s'appelle la mise en commun basée sur l'opération basée sur des statistiques agrégées dans une région. Chihua est divisé en une image d'entrée dans plusieurs zones rectangulaires, sortie maximale ou moyennes de sortie pour chaque sous-région, généralement plus en utilisant le réservoir maximal, voir la figure 3.7.

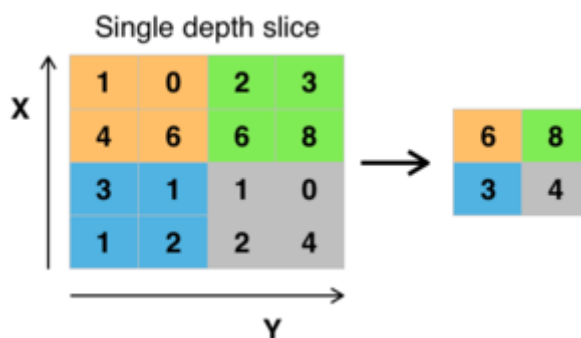


Figure 3- 7 :Pooling avec un filtre 2x2 et un pas de 2 [43]

3) Couches entièrement connectées

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. Leurs fonctions d'activations peuvent donc être calculées avec une multiplication matricielle suivie d'un décalage de polarisation.

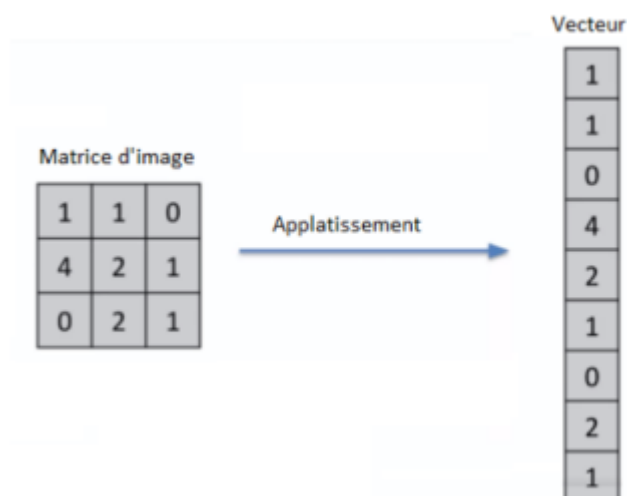


Figure 3- 8 : L'aplatissage de l'image[38]

3.4.2 Fonction d'activation

La compréhension du cerveau humain n'est pas linéaire, mais une cartographie non linéaire complexe. C'est-à-dire une des trois parties de base du réseaux de neurones mentionnées dans 2-1, l'utilisateur transforme l'entrée et la sortie entre le neurones. Par conséquent, les réseaux de neurones utilisent souvent des fonctions d'activation non linéaires pour simuler le comportement cognitif non linéaire du cerveau humain par le biais de fonctions d'activation non linéaire. Les fonctions d'activation non

linéaire couramment utilisées comprennent principalement les fonctions sigmoïdes, les fonctions tanh et les fonctions de rectification linéaire, définies comme suit:

$$\text{Sigmoid: } f(x) = \frac{1}{1+e^{-x}} \quad (3.9)$$

$$\text{Tanh: } f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.10)$$

$$\text{ReLU: } f(x) = \max(0, x) \quad (3.11)$$

Voir la figure 3.9 le côté gauche est une courbe de fonctions sigmoïde et la courbe de fonction relu est une pente près de 0 en raison de la fonction Sigmoid et l'état des non-lumières est généré, de sorte que cette expérience utilise la fonction relu.

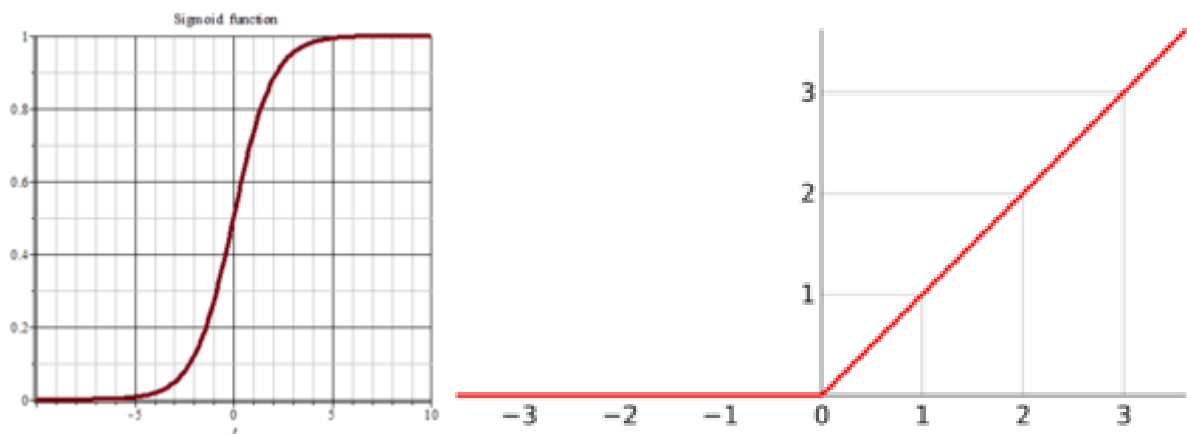


Figure 3- 9 :courbe sigmoïde et courbe ReLU[43]

3.4.3 Fonction de perte :

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 fonction Softmax permet de calculer la distribution de probabilités sur les classes de sortie.

3.4 .3.1.Types de fonction de perte:

1) Fonction de perte de régression:

Les modèles de régression permettent de prédire une valeur continue, par exemple la surface au sol, le nombre de pièces, la taille des pièces, prédire le prix de la pièce. La fonction de perte utilisée dans le problème de régression est appelée «fonction de perte de régression».

2) Fonctions de perte de classification binaire:

La classification binaire est un algorithme de prédiction dans lequel la sortie peut être l'un des deux éléments, indiqué par 0 ou 1. La sortie des algorithmes de classification binaire est un score de prédiction (principalement). Ainsi, la classification se produit en fonction du seuil de la valeur (la valeur par défaut est 0,5). Si le score de prédiction > seuil, 1 sinon 0.

3) Fonctions de *perte* de classification multi-classes:

Les classifications multi-classes sont les problèmes de modélisation prédictive où il y a plus de variables classes cibles. C'est juste l'extension du problème de classification binaire.

3.5 Quelques réseaux convolutifs célèbres :

Comme nous l'avons précédemment indiqué, certains réseaux profonds ont fait de telles contributions significatives sur le domaine qu'ils ont devenir des normes largement connues. C'est le cas d'Alexnet, VGG-16, Googlenet et ResNet. Telle était leur importance qu'ils soient actuellement utilisés comme blocs de construction pour de nombreuses architectures de segmentation. Pour cette raison, nous va consacrer cette section pour les examiner

3.5.1 GoogLeNet :

GoogLeNet est un réseau introduit par Szegedy et al.[48] qui a remporté le défi ILSVRC-2014 avec un test TOP-5 Précision de 93: 3%. Cette architecture CNN est caractérisée par sa complexité, souligné par le fait qu'il est composé par 22 couches et un bloc de construction nouvellement introduit appelé Module de création (voir la fig 3.10). Cette nouvelle approche s'est avérée que les couches CNN puissent être empilées d'une manière plus faite qu'une manière typique séquentielle. En fait, ces modules consistent en un réseau dans la couche réseau (NIN), une opération de mise en commun, un couche de convolution de grande taille et convolution de petite taille couche. Tous sont calculés en parallèle et suivis par 1×1 Opérations de convolution pour réduire la dimensionnalité.

Grâce à ces modules, ce réseau met une considération particulière sur la mémoire et le coût de calcul par de manière significative. Réduire le nombre de paramètres et d'opérations

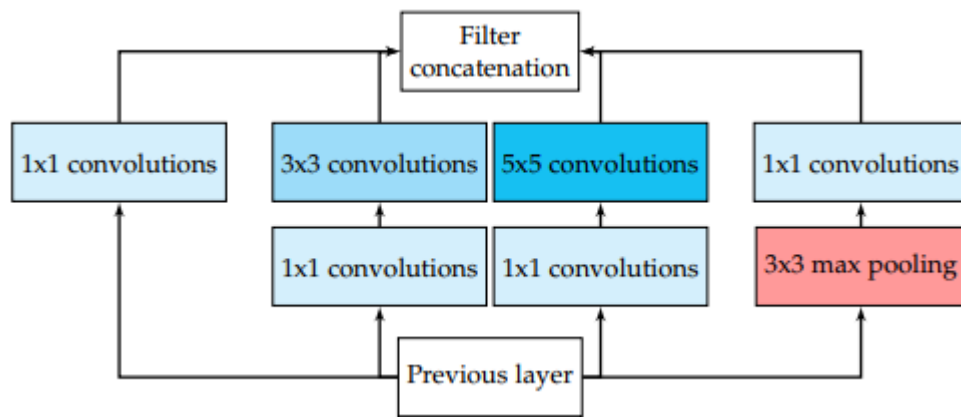


Figure 3- 10: Module de création avec réduction de la dimensionnalité de l'architecture GoogLeNet. [48]

3.5.2 VGG16 :

Le groupe Visual Geometry (VGG) est un modèle CNN introduit par le groupe Visual Geometry (VGG) de l'université d'Oxford. Ils ont proposé divers modèles et configurations de CNNs profonds, l'un d'entre eux a été soumis à Le défi de reconnaissance visuelle d'Imagenet à grande échelle (ILSVRC) -2013. Ce modèle, également appelé VGG-16 en raison de le fait qu'il soit composé de 16 couches de poids, est devenu Popular merci à sa réalisation de 92: 7% Top-5 Test précision Fig 11 montre la configuration de VGG-16. La principale différence entre VGG-16 et ses prédécesseurs est la Utilisation d'une pile de couches de convolution avec petit réceptif champs dans les premières couches au lieu de quelques couches avec grand champs réceptifs. Cela conduit à moins de paramètres et plus de non-linéarités entre les deux, faisant ainsi la fonction de décision plus discriminant et le modèle plus facile à former.

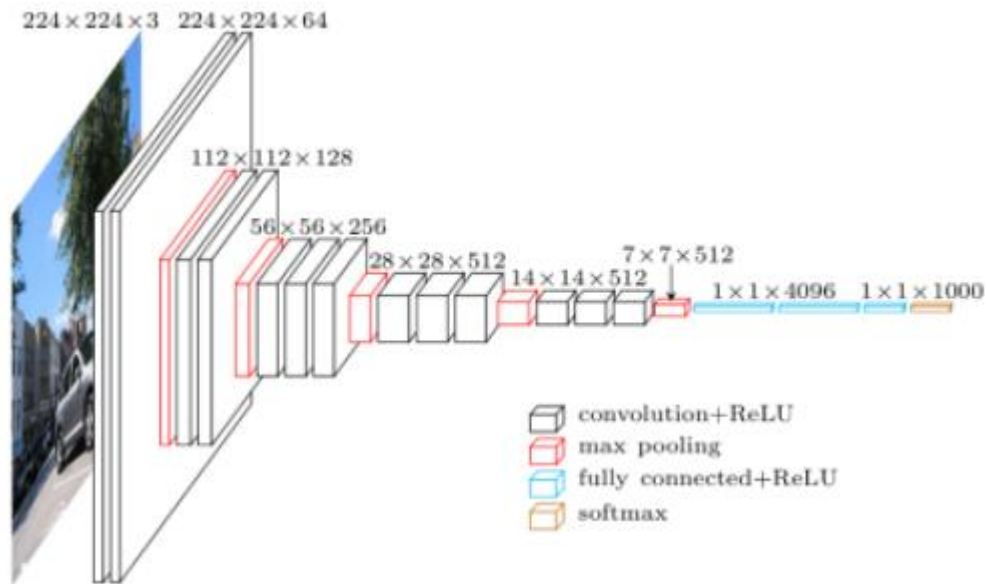


Figure 3- 11 : Architecture VGG-16 CNN. Figure extraite de La conversation de Matthieu Cordon avec sa permission.

3.5.3 RESNET :

Le RESNET de Microsoft[49] est spécialement remarquable grâce à GAGNANT ILSVRC-2016 avec une précision de 96: 4%. Mis à part cela Fait, le réseau est bien connu en raison de sa profondeur (152 couches) et l'introduction de blocs résiduels (voir la Fig3.12).

Les blocs résiduels traitent du problème de la formation un très profond architecture en introduisant l'identité Ignorer les connexions afin que Les couches peuvent copier leurs entrées sur la couche suivante. L'idée intuitive derrière cette approche est qu'elle assure que la couche suivante apprend quelque chose de nouveau et différent de ce que l'entrée a déjà codé (puisque'elle est fournie avec la sortie de la couche précédente et son entrée inchangée). De plus, ce type de connexion aide à surmonter le problème des gradients de disparition

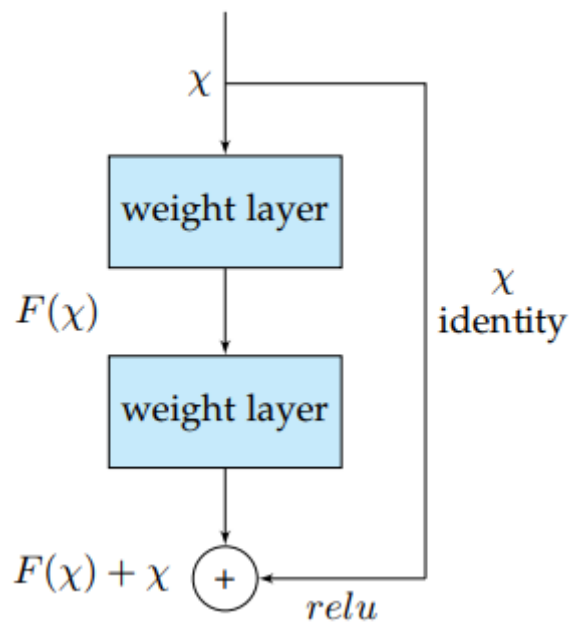


Figure 3- 12 : Bloc résiduel de l'architecture RESNET [49]

3.5.4 AlexNet

AlexNet était le pionnier profond CNN qui a remporté le ILSVRC-2012 avec une précision de test Top-5 de 84: 6% tandis que Le concurrent le plus proche, qui a utilisé des techniques traditionnelles au lieu d'architectures profondes, a atteint une précision de 73: 8% dans le même défi. L'architecture présentée par Krizhevsky et al.[50]était relativement simple. Cela consiste en Cinq couches convolutionnelles, max-piscine, des unités linéaires rectifiées (relatives) comme non-linéarités, trois entièrement connectées couches et décrochage. La fig3.13 montre que l'architecture CNN

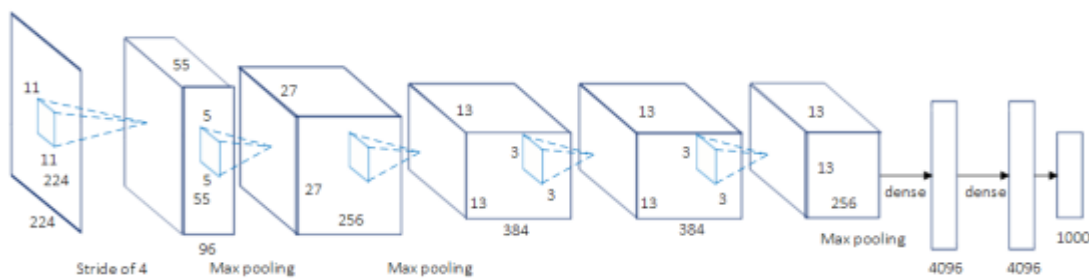


Figure 3- 13 : Architecture de réseau de neurones convolutionnelle Alexnet. Chiffre reproduit de [51]

3.5.5 ReNet :

Afin d'étendre des réseaux de neurones récurrents (RNNS) aux tâches multidimensionnelles, Graves et al. proposé un réseau de neurones récurrent multidimensionnels (MDNN) Architecture qui

remplace chaque récurrent Connexion des RNN standard avec des connexions D, où D est le nombre de dimensions de données spatio-temporelles. Basé Sur cette approche initiale, Visin el al. [52] proposé Renet architecture dans laquelle multidimensionnels, Ils utilisent des RNN de séquence habituelle. De cette façon, le Nombre de RNNS permet de réduire linéairement à chaque couche concernant au nombre de dimensions D de l'image d'entrée (2D). Dans cette approche, chaque couche convolution elle (convolution + la mise en commun) est remplacée par quatre RNNS balayant l'image verticalement et horizontalement dans les deux sens que nous pouvons voir à la fig 3.14

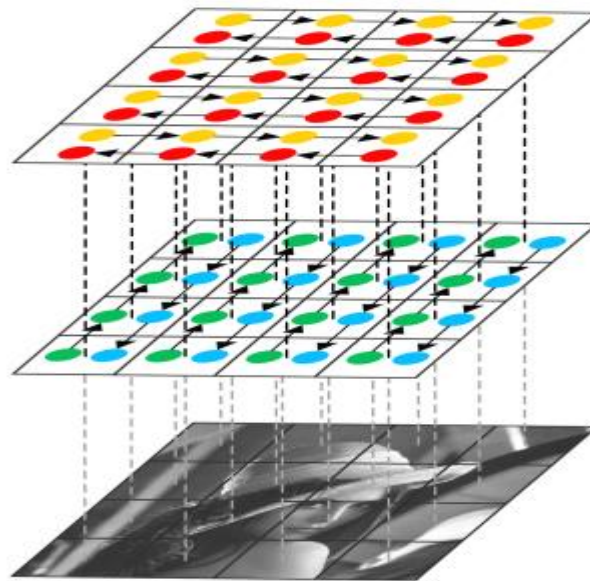


Figure 3- 14 : Une couche de modélisation d'architecture renouvelante verticale et Dépendances spatiales horizontales. Extrait de [52].

3.5.6 U-Net :

le U-Net a un Structure du décodeur de codeur, avec les connexions de saut entre les couches de codage et de décodage correspondantes qui permettent Le réseau de conserver des fonctionnalités de bas niveau pour la prédiction finale. 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 512×512 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 3 parties :

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

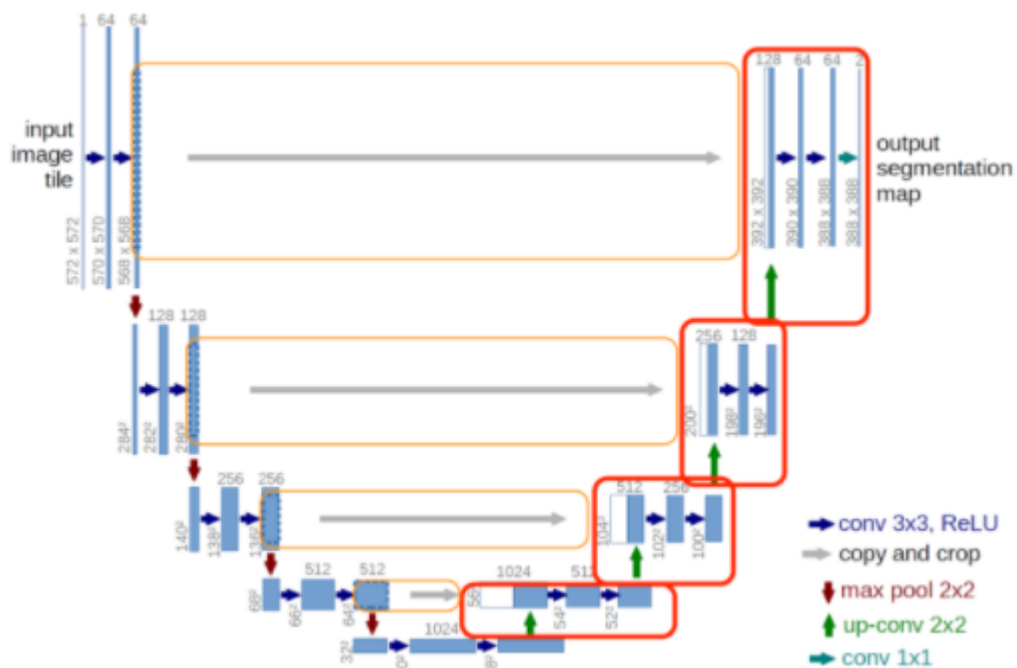


Figure 3- 15 : L'architecture U-net[53]

3.6 Apprentissage de transfert(Transfer learning)

Entraînement Un réseau neuronal profond de zéro n'est souvent pas réalisable en raison de diverses raisons: un jeu de données suffisant la taille est requise (et pas généralement disponible) et atteignant la convergence peut prendre trop de temps pour que les expériences soient valeur. Même si un jeu de données suffisamment grand est disponible et la convergence ne prend pas aussi longtemps, il est souvent utile de commencer avec des poids pré-entraînés au lieu d'aléatoires initialisées [20] [21]. Rajuster les poids d'un réseau pré-formé En continuant avec le processus de formation, c'est l'un des principaux Transférer des scénarios d'apprentissage.

Yosinski et al. [22] prouvé que les caractéristiques de transfert Même à partir de tâches distantes peut

être meilleure que d'utiliser aléatoire initialisation, en tenant compte de la transférabilité de caractéristiques diminue comme la différence entre le pré-entraîné tâche et la cible augmente.

Cependant, appliquer cette technique d'apprentissage de transfert est pas complètement simple. D'une part, là sont des contraintes architecturales qui doivent être remplies pour utiliser un réseau pré-préparé. Néanmoins, comme il n'est pas habituel de venir avec une toute nouvelle architecture, il est courant de réutiliser Les architectures de réseau déjà existantes (ou les composants) ainsi Activation de l'apprentissage de transfert. D'autre part, la formation processus diffère légèrement lorsque vous ajustiser la formation au lieu de la formation de zéro.

Il est important de choisir correctement les couches à l'aise - généralement la partie supérieure du réseau, Puisque le bas a tendance à contenir plus de fonctionnalités génériques - et choisissez également une politique appropriée pour le taux d'apprentissage, qui est généralement plus petit en raison du fait que le pré-entraîné les poids devraient être relativement bons donc il n'y a pas de besoin de les changer radicalement.

En raison de la difficulté inhérente à la collecte et à la création jeux de données de segmentation marqués par pixels, leur échelle n'est pas aussi Grand comme la taille des ensembles de données de classification tels que ImageNet [54].

Ce problème devient encore pire lorsqu'il s'agit de jeux de données RGB-D ou 3D, qui sont encore plus petits. Pour cette raison, transférer l'apprentissage et en particulier ajustement à partir des réseaux de classification pré-formés sont une tendance commune pour réseaux de segmentation et a été appliqué avec succès dans les méthodes que nous allons examiner dans les sections suivantes.

L'apprentissage de transfert est un outil d'identité dans l'apprentissage de la machine pour résoudre la base Problème de données d'entraînement insuffisantes.

.Le processus d'apprentissage de transfert illustre à la Fig 3.16. Certaines notations utilisées dans cette enquête ne doivent pas être définies clairement. Vol de tous, Nous donnons respectivement les définitions d'un domaine et une tâche: un domaine peut être Représenté par $D = \{\chi, p(x)\}$, qui contient deux parties: l'espace de fonctionnalité Et la distribution de probabilité de bord $p(x)$ où $x = \{x_1, \dots, x_n\} \in \chi$. Une tâche Peut être représenté par $t = \{y, f(x)\}$. Il consiste en deux parties: l'espace d'étiquette Y Et la fonction de prédiction cible $f(x)$. $F(x)$ peut également être considéré comme un conditionnel Fonction de probabilité $p(y | x)$. L'apprentissage de transfert peut être formel défini Comme suit:

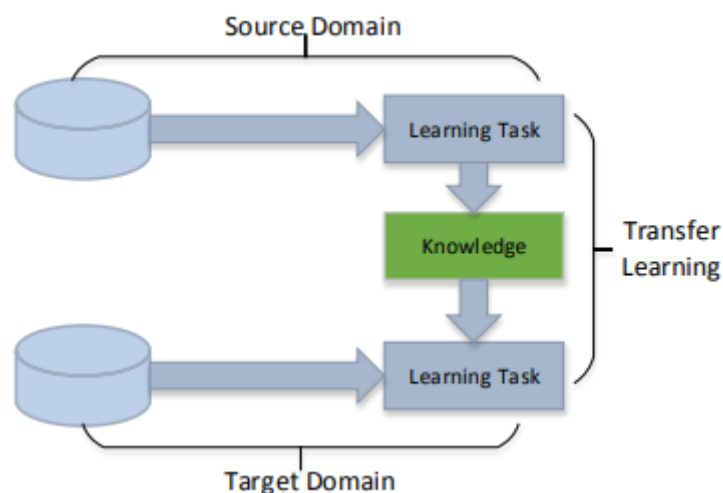


Figure 3- 16 : Processus d'apprentissage du transfert

Dans l'apprentissage par transfert, nous essayons de transférer autant de connaissances que possible de la tâche précédente sur laquelle le modèle a été formé à la nouvelle tâche à accomplir. Cette connaissance peut se présenter sous diverses formes en fonction du problème et des données. Par exemple, cela pourrait être la façon dont les modèles sont composés, ce qui nous permet d'identifier plus facilement de nouveaux objets.

3.6.1 utilisation l' apprentissage par transfert:

L'apprentissage par transfert présente plusieurs avantages, mais les principaux avantages sont le gain de temps d'entraînement, de meilleures performances des réseaux de neurones (dans la plupart des cas) et le fait de ne pas avoir besoin de beaucoup de données.

Habituellement, beaucoup de données sont nécessaires pour former un réseau de neurones à partir de zéro, mais l'accès à ces données n'est pas toujours disponible - c'est là que l'apprentissage par transfert est utile. Avec l'apprentissage par transfert, un modèle d'apprentissage automatique solide peut être construit avec relativement peu de données d'entraînement, car le modèle est déjà pré-entraîné. Ceci est particulièrement utile dans le traitement du langage naturel, car la plupart des connaissances d'experts sont nécessaires pour créer de grands ensembles de données étiquetés. De plus, le temps de formation est réduit car il peut parfois prendre des jours, voire des semaines, pour entraîner un réseau de neurones profond à partir de zéro sur une tâche complexe.

3.6.2. approche de l'apprentissage par transfert:

1) former un modèle pour réutiliser:

Imaginez que vous vouliez résoudre la tâche A mais que vous ne disposiez pas de suffisamment de données pour entraîner un réseau neuronal profond. Une façon de contourner ce problème est de trouver une tâche B connexe avec une abondance de données. Entraînez le réseau neuronal profond à la tâche B et utilisez le modèle comme point de départ pour résoudre la tâche A. Le fait que vous ayez besoin d'utiliser tout le modèle ou seulement quelques couches dépend fortement du problème que vous essayez de résoudre. Si vous avez la même entrée dans les deux tâches, il est possible de réutiliser le modèle et de faire des prédictions pour votre nouvelle entrée. Alternativement, changer et recycler différentes couches spécifiques à une tâche et la couche de sortie est une méthode à explorer.

2) Utilisation d'un modèle pré-formé:

La deuxième approche consiste à utiliser un modèle déjà pré-formé. Il existe un grand nombre de ces modèles, alors assurez-vous de faire un peu de recherche. Le nombre de couches à réutiliser et le nombre de couches à recycler dépendent du problème. Keras, par exemple, fournit neuf modèles pré-entraînés qui peuvent être utilisés pour l'apprentissage par transfert, la prédiction, l'extraction de caractéristiques et le réglage fin. Vous pouvez trouver ces modèles, ainsi que quelques brefs tutoriels sur la façon de les utiliser, [ici](#). Il existe également de nombreux instituts de recherche qui publient des modèles formés. Ce type d'apprentissage par transfert est le plus couramment utilisé tout au long de l'apprentissage profond

3) Extraction des caractéristiques:

Une autre approche consiste à utiliser l'apprentissage en profondeur pour découvrir la meilleure représentation de votre problème, ce qui signifie trouver les fonctionnalités les plus importantes. Cette approche est également connue sous le nom d'apprentissage de la représentation et peut souvent se traduire par une bien meilleure performance que celle obtenue avec une représentation conçue à la main

3.7 Conclusion :

À travers ce chapitre, nous avons présenté le Deep Learning qui utilise les réseaux de neurones convolutionnels les plus répandus.

Le Deep Learning est l'une des méthodes de développement le traitement d'images tel que la segmentation d'images. , nous nous sommes intéressés à la théorie de l'apprentissage profond et particulièrement au recours des CNNs Dans le prochain chapitre, nous exploiterons le Deep Learning pour réaliser une segmentation sémantique Ces réseaux sont capables d'extraire des caractéristiques d'images présentées en entrée et de classifier ces caractéristiques,

Chapitre 4:Application

4.1 Introduction :

La segmentation des images des tumeurs du sein est une tâche cruciale dans le traitement de l'image médicale. Dans cette partie, nous allons présenter notre choix qui consistant à appliquer quelques modèles d'apprentissage profond (DL) qui répondra au problème de la segmentation sémantique des images médicales pour la détection des tumeurs du sein.

Dans ce chapitre nous allons utiliser l'apprentissage de transfert pour réutiliser quelques modèles des réseaux de neurones profond pré- apprenait pour d'autre problématiques, nous allons faire faire un genre de comparaison entre ces différents modèles pour voire le plus réponsus a notre problématique de segmentation des images tumorales des sein

4.3 Architectures de réseau de neurones profonds

Quatre approches à la pointe de la technologie pour la construction de modèles de segmentation sémantique sont utilisées dans cette étude, nommément: RESNET18, RESNET50, U-NET et Alexnet [41,42, 48, 61].

La couche d'entrée des modèles CNN nécessite une image à 3 canaux car elles étaient préformées en utilisant la couleur images de la base de données Imagenet [50]. Par conséquent, l'image de bus d'origine grise d'origine est d'abord annulée à 128×128 pixels, comme suggéré dans [32]. Ensuite, l'image redimensionnée est reproduite deux fois pour former les trois couches d'une image couleur introduite à CNN. Par conséquent, les couches d'entrée de toute la segmentation sémantique Les modèles sont ajustés à $128 \times 128 \times 3$.

La couche de sortie est une couche softmax utilisée pour obtenir les probabilités de chaque pixel appartenant à un classer. Ensuite, un pixel d'entrée est classé comme pixel de tumeur si sa valeur de probabilité est supérieure à 0,5; autrement, Il est classé comme pixel de tissu normal.

4.4 Indices de performance

La performance de la segmentation est évaluée en comparant la similitude entre la segmentation de la trésorerie, effectuée manuellement par le radiologue et le résultat du CNN. Le tableau 4-1 montre les deux catégories Matrice de confusion, à partir de laquelle les indices de performance de la segmentation de base suivants sont calculés:

- le score F_1 (également appelé coefficient de similarité de dés) est la moyenne pondérée de la précision $\frac{TP}{TP+FP}$ et rappel, $\frac{TP}{TP+FN}$ qui considère le déséquilibre de classe:

$$F1s = \frac{2TP}{2TP+FP+FN} \quad (4.1)$$

L'intersection de l'union (également appelée l'indice Jaccard) quantifie le rapport de chevauchement entre Le masque de la vérité au sol et l'image segmentée:

$$IoU = \frac{TP}{TP+FP+FN} \quad (4.2)$$

Les deux indices renvoient des valeurs dans la plage [0, 1], où les valeurs vers l'unité indiquent une performance de segmentation adéquate.

Tableau 4 - 1 : Matrice de confusion pour la segmentation binaire.

		Classe prédite	
		Tumeur	Tissu Normal
Réal Classer	Tumeur	Vrai positif (TP)	faux négatif (FN)
	Tissu normal	Faux positif (FP)	Véritable négatif (TN)

Les indices $F1s$ et IoU comparent la similitude des résultats du CNNs avec segmentations de référence sur l'ensemble de tests. Chaque ensemble de formation est divisé au hasard en 80% pour Apprendre et 20% pour la validation.

4.5 Dataset « BUT » (Brest Ultrasound Tumor » :

Nous avons utilisé pour cette étude une partie des images ultra-soun du cancer du sein de la base ‘‘BUS’’ qui a été créé par Al-Debyani et all et publiée en 2020[55] Elle contient dans sa totalité 780 images avec une taille d'image moyenne de 500 * 500 pixels. Les images sont au format PNG. Les images de la vérité terrain sont présentées avec des images originales. Les images sont classées en trois classes, qui sont normales, bénignes et malignes. [55]

Vus la restriction matériels et logiciels, nous avons utilisé que 100 images de la bases.

4.6. Environnement Matériel et logiciel :

Dans cette section, nous présenterons les environnements matériel et logiciel de notre travail

4.6.1 Environnement matériel

Afin de mener à bien ce projet, il a été mis à notre disposition un ordinateur DEEL avec les caractéristiques suivantes :

Nom de l'appareil DESKTOP-E9L4TPI
 Processeur Intel(R) Core(TM) i5-4310U CPU @ 2.00GHz 2.60 GHz
 Mémoire RAM installée 12,0 Go
 ID de périphérique 4EB6E3B1-E64F-40B2-A321-17E8F2CA6CDD
 ID de produit 00330-80000-00000-AA648
 Type du système Système d'exploitation 64 bits, processeur x64
 Stylet et fonction tactile La fonctionnalité d'entrée tactile ou avec un stylet n'est pas disponible sur cet écran
 Édition Windows 10 Professionnel
 Version 20H2
 Installé le 01/11/2020
 Build du système d'exploitation 19042.985
 Expérience Windows Feature Experience Pack 120.2212.2020.0

4.6.2 Environnement logiciel :

Tout le programme était de développement à **MATLAB R2020A** et Google coalab :

1) MATLAB :

MATLAB est un logiciel commercial de calcul interactif. Il permet de réaliser des simulations numériques basées sur des algorithmes d'analyse numérique. Il peut donc être utilisée pour la résolution approchée d'équations différentielles, d'équations aux dérivées partielles ou de systèmes linéaires, etc... L'objectif de ces séances MATLAB est double : la connaissance de ce logiciel est en soi indispensable parce qu'il est de plus en plus utilisé dans l'industrie et les banques pour développer des prototypes de logiciels et tester de nouveaux algorithmes. Tous les ans entre vingt et trente pour cent des étudiants du DESS utilisent Matlab pendant leur stage. Ensuite son

apprentissage va passer par la mise en pratique des algorithmes d'analyse numérique 'étudiés plus théoriquement dans le reste du module

2) Google Colab :

Avant d'introduire Google Colaboratory, nous introduisons Jupyter Notebooks, la technologie qui est collaboratoire Basé sur. Jupyter est un outil à source ouverte et à base de navigateur Qui intègre des langues interprétées, des bibliothèques et des outils pour la visualisation [10]. Un ordinateur portable Jupyter peut fonctionner soit Localement ou sur le cloud. Chaque document est composé ou plusieurs cellules, où chaque cellule contient une langue de script ou un code de marquage, et la sortie est intégrée au document.

Les sorties typiques incluent du texte, des tables, des graphiques et des graphiques.

L'utilisation de cette technologie facilite la partage et la réplique des travaux scientifiques depuis les expériences et les résultats sont présentés. De manière autonome .

Google Colaboratory (A.K.A Colab) est un projet qui a la Objectif de diffuser l'enseignement de l'apprentissage de la machine et Recherche [6]. Les cahiers collaboratoires sont basés sur Jupyter et travailler comme un objet Google Docs: peut être partagé et Les utilisateurs peuvent collaborer sur le même ordinateur portable. Collaboratoire Fournit soit Python 2 et 3 exécutions préconfigurées est un produit de Google Research. Colab permet à n'importe qui d'écrire et d'exécuter le code MATLAB de son choix par le biais du navigateur Avec l'apprentissage de la machine essentielle et l'intelligence artificielle Bibliothèques, telles que Tensorflow, Matplotlib et Keras. La machine virtuelle sous l'exécution (VM) est désactivée après Une période de temps et toutes les données et configurations de l'utilisateur sont Perdu. Cependant, le cahier est préservé et il est également possible de transférer des fichiers du disque dur VM à l'utilisateur.

Compte de lecteur Google. Enfin, ce service Google fournit un Runtime accéléré GPU, également entièrement configuré avec le logiciel précédemment décrit. L'infrastructure de Google Colaboratory est hébergée sur la plateforme Google Cloud

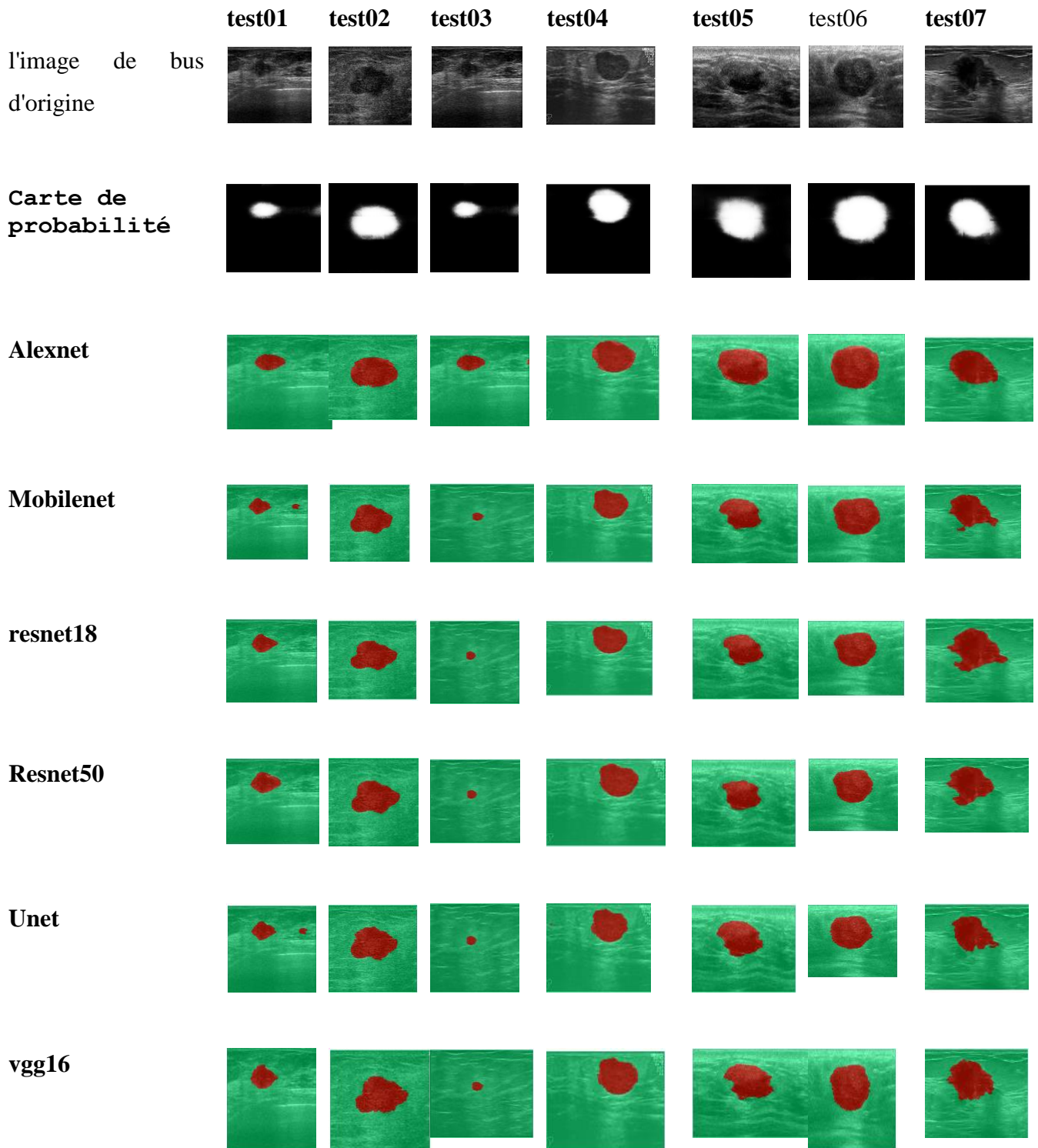
4.7 Résultats et discussion :

Les résultats de la segmentation des architectures du CNNs sont présentés dans le tableau ci-dessous qui montre la performance globale de la segmentation des architectures du CNN testées.

Notamment des modèles de segmentation Basé sur Resnet50 et Resnet18 obtenu F1S> 0,88 et IOU> 0,78, tandis que U-Net atteint légèrement Résultats plus bas. En général, FCN avec Alexnet a atteint les résultats de la segmentation les plus bas, tandis que le meilleur La performance de la segmentation

a été obtenue par RESNET50.

Ce résultat suggère que la profondeur du réseau est pertinente pour obtenir une meilleure qualité de segmentation, car des fonctionnalités plus subtiles sont extraites quand la profondeur du réseau grandit. Notez que RESNET50 est de 50 couches de profondeur, tandis que FCN-Alexnet dispose de 10 couches de convolution



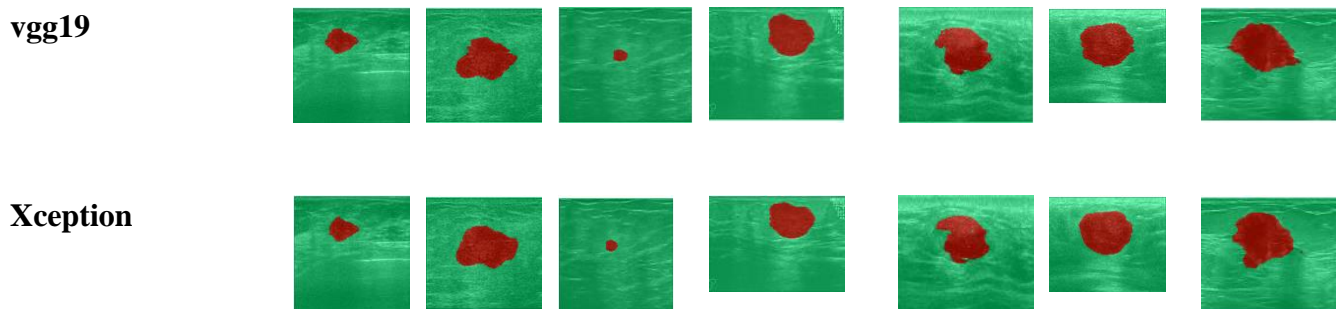


Figure 4 - 1: Tumeurs bénignes segmentées par les huit modèles CNN

4.8 Conclusions :

Dans la communauté de la vision de l'ordinateur, les architectures CNN sont régulièrement rendues publiques pour la reproductibilité des fins et de les réutiliser par le transfert d'apprentissage. Cependant, dans le cas de la segmentation basée sur le CNN proposées dans la littérature, un inconvénient notable est l'indisponibilité de des modèles pré-apprenais. En outre, plusieurs approches ont présenté des architectures CNN nouvelles ou modifiées qui sont tout à fait difficile de reproduire pour évaluer leurs performances sur de nouveaux jeux de données. Par conséquent, en raison de ces techniques Des difficultés, nous avons évalué les architectures publiques CNN pour la segmentation de l'image du cancer du sein de la base BUS. Quatre modèles de segmentation sémantiques basés au CNN ont été évalués: U-Net, Resnet18, Resnet50 et Alexnet. Des Modèles pré-apprenais ont été prises en compte dans ces modèles de segmentation sémantique pour évaluer la capacité de généralisation des modèles CNN pour la segmentation automatique des images « BUS ». Les résultats obtenus confirment l'efficacité de l'utilisation de l'apprentissage de transfert des modèles pré-apprenais dans d'autres taches pour la segmentation sémantique des images médicales ultra Sound, et renforcent l'hypothèse qui dit que : plus le réseau est profond, mieux la qualité de résultats augmentent .

Conclusion générale

Ce sujet a adopté la technologie d'apprentissage en profondeur, de diviser sémantiquement l'image médicale, a tenu une tentative d'introduction de technologies émergentes dans des domaines traditionnels, grâce à la visualisation de l'expérience et de précision des résultats, a complété l'étude sur l'algorithme d'apprentissage en profondeur. Dans le domaine du traitement des imageries médicales, et d'utiliser davantage les algorithmes d'apprentissage en profondeur pour aider le diagnostic de médecin, des soins médicaux intelligents.

Nous allons adapter et comparer quatre modèles des réseaux de neurones de convolution : Alexnet, Resnet18, Resnet80, et U-net. Par transfert d'apprentissage, ces CNN sont Finalisé pour segmenter les images de bus dans des pixels normaux et tumoraux.

L'objectif était Évaluer la capacité de généralisation des modèles CNN bien établis qui ont été développés par le Communauté de vision informatique pour la segmentation automatique des images de bus. Les modèles générés dans cette Des études, qui peut être testée sur de nouvelles images ou amélioration de l'apprentissage des transferts sur d'autres jeux de données de bus.

Peu de temps après une technologie d'apprentissage approfondie, dans le stade de développement rapide, diverses applications théoriques et pratiques ont encore un large domaine de recherche, en particulier dans le domaine traditionnel des soins médicaux, avec tous les aspects de la direction de la recherche. Pour la division sémantique, comment améliorer encore la précision, identifier les lésions est au cœur de ce problème. En raison de la particularité du domaine médical, les exigences de l'instrument et des logiciels ne sont pas strictes, les exigences en matière de précision et de précision seront donc très élevées. Le modèle proposé dans le domaine de l'apprentissage profond n'a pas de valeur caractéristique d'extraction optimisée pour les images médicales. Un modèle de réseau spécialement appliqué aux images médicales est une autre direction de recherche.

En outre, en raison des caractéristiques de l'incertitude des paramètres de la technologie d'apprentissage en profondeur, il n'existe aucune théorie complète pour fournir des recherches sur l'ajustement des paramètres. Il est donc nécessaire de coopérer avec des personnes dans le domaine professionnel, comment ajuster les paramètres dans le modèle réseau Pour améliorer ou optimal, la valeur est une autre direction qui doit être grandement étudiée.

Références Bibliographiques

Référence :

- [1] M. G. EL-HACHEMI, « Performances dans la Segmentation d'images médicales », École nationale Supérieure en Informatique, 2018.
- [2] F. HADJI, « Conception et réalisation d'un système de cryptage pour les images médicales », UNIVERSITE MOHAMED BOUDIAF-M'SILA FACULTE DES MATHEMATIQUES ET DE L ..., 2018.
- [3] P. Gravereau, « Introduction à la pratique de la diffraction des rayons X par les poudres », Sciences Chimiques, 2011.
- [4] A. Meziane, « Système d'information en imagerie médicale: moyen de diagnostic, de thérapie et d'économie », *RIST*, vol. 11, n° 2, 2001.
- [5] M. Bachaalany, « Utilisation de capteurs CMOS rapides pour l'imagerie X à très haute sensibilité », Strasbourg, 2012.
- [6] S. P. Wilk, « Axial transverse tomography of the chest », *Radiology*, vol. 72, n° 1, p. 42-50, 1959.
- [7] E. Bercovich et M. C. Javitt, « Medical imaging: from Roentgen to the digital revolution, and beyond », *Rambam Maimonides medical journal*, vol. 9, n° 4, 2018.
- [8] D. Kane, W. Grassi, R. Sturrock, et P. V. Balint, « A brief history of musculoskeletal ultrasound: 'From bats and ships to babies and hips' », *Rheumatology*, vol. 43, n° 7, p. 931-933, 2004.
- [9] G. Öz et al., « Clinical proton MR spectroscopy in central nervous system disorders », *Radiology*, vol. 270, n° 3, p. 658-679, 2014.
- [10] D. W. Townsend, « Combined positron emission tomography-computed tomography: the historical perspective », in *Seminars in Ultrasound, CT and MRI*, 2008, vol. 29, n° 4, p. 232-235.
- [11] L. Akli et A. M. Farid, « Segmentation d'images par colonies d'araignées artificielles », Université Mouloud Mammeri, 2014.
- [12] S. BENRAMDENE, « UTILISATION DES SYSTEMES D'INFORMATIONS GEOGRAPHIQUES POUR LA MESURE DE LA SURFACE ».
- [13] L. Lechlek, M. Ghougal, et F. Boutaouche, « La segmentation d'image médicale par la méthode d'ensembles de niveaux level sets », 2012.
- [14] A. Krähenbühl, « Segmentation et analyse géométrique: application aux images tomodensitométriques de bois », Université de Lorraine, 2014.
- [15] S. W. Zucker, « Region growing: Childhood and adolescence », *Computer graphics and image processing*, vol. 5, n° 3, p. 382-399, 1976.
- [16] P. J. Besl et R. C. Jain, « Segmentation through variable-order surface fitting », *IEEE Transactions on pattern analysis and machine intelligence*, vol. 10, n° 2, p. 167-192, 1988.
- [17] C. Solomon et T. Breckon, *Fundamentals of Digital Image Processing: A practical approach with examples in Matlab*. John Wiley & Sons, 2011.
- [18] K. C. Strasters et J. J. Gerbrands, « Three-dimensional image segmentation using a split, merge and group approach », *Pattern Recognition Letters*, vol. 12, n° 5, p. 307-325, 1991.
- [19] W. T. Freeman et E. H. Adelson, « The design and use of steerable filters », *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 13, n° 9, p. 891-906, 1991.
- [20] N. Laouadi, L. Triki, et I. Nini, « Segmentation d'image par région sur la base des contours des objets », 2017.
- [21] D. J. Burr, « Elastic matching of line drawings », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, n° 6, p. 708-713, 1981.
- [22] M. Kass, A. Witkin, et D. Terzopoulos, « Snakes: Active contour models », *International journal of computer vision*, vol. 1, n° 4, p. 321-331, 1988.
- [23] J.-J. Rousselle, « Les contours actifs, une méthode de ségmentation: application à l'imagerie médicale », Tours, 2003.
- [24] J. Canny, « A computational approach to edge detection », *IEEE Transactions on pattern*

analysis and machine intelligence, n° 6, p. 679-698, 1986.

- [25] I. Rivals, L. Personnaz, G. Dreyfus, et J. L. Ploix, « Modélisation, classification et commande par réseaux de neurones: principes fondamentaux, méthodologie de conception et illustrations industrielles », *Les réseaux de neurones pour la modélisation et la commande de procédés*, JP Corriou, ed.(Lavoisier Tec & Doc, 1995), 1995.
- [26] C. HOUASSINE, « Segmentation d'images par une approche biomimétique », Université de Boumerdès-M'hamed Bougara, 2012.
- [27] W. Aribi *et al.*, « Évaluation de Techniques Pyramidales de Fusion Multimodale (IRM/TEP) d'Images Cérébrales », in *5th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications*, 2009, vol. 19, p. 547-551.
- [28] S. Bres, J.-M. Jolion, et F. Lebourgeois, *Traitement et analyse des images numériques*. Hermes Science Publications, 2003.
- [29] V. Kumar *et al.*, « Automated and real-time segmentation of suspicious breast masses using convolutional neural network », *PloS one*, vol. 13, n° 5, p. e0195816, 2018.
- [30] Z. Cao, L. Duan, G. Yang, T. Yue, et Q. Chen, « An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures », *BMC medical imaging*, vol. 19, n° 1, p. 1-9, 2019.
- [31] M. H. Yap *et al.*, « Automated breast ultrasound lesions detection using convolutional neural networks », *IEEE journal of biomedical and health informatics*, vol. 22, n° 4, p. 1218-1226, 2017.
- [32] C. Tao *et al.*, « New one-step model of breast tumor locating based on deep learning », *Journal of X-ray Science and Technology*, vol. 27, n° 5, p. 839-856, 2019.
- [33] M. Amiri, R. Brooks, B. Behboodi, et H. Rivaz, « Two-stage ultrasound image segmentation using U-Net and test time augmentation », *International journal of computer assisted radiology and surgery*, vol. 15, n° 6, p. 981-988, 2020.
- [34] A. Holzinger, P. Kieseberg, E. Weippl, et A. M. Tjoa, « Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI », in *International Cross-Domain Conference for Machine Learning and Knowledge Extraction*, 2018, p. 1-8.
- [35] A. Rayhana et C. Radhia, « ÉVALUATION DE LA PERFORMANCE D'UN CAPTEUR LOGICIEL EN UTILISANT L'APPRENTISSAGE EN PROFONDEUR », UNIVERSITE MOHAMED BOUDIAF-M'SILA, 2019.
- [36] M. M. Mijwel, « Artificial neural networks advantages and disadvantages », *Retrieved from LinkedIn <https://www.linkedin.com/pulse/artificial-neural-net-work>*, 2018.
- [37] C. Chatelain, « Contributions à l'apprentissage profond, applications à la reconnaissance d'écriture manuscrite », Université de Rouen, 2019.
- [38] Y. Bengio, *Learning deep architectures for AI*. Now Publishers Inc, 2009.
- [39] Y. LeCun, P. Haffner, L. Bottou, et Y. Bengio, « Object recognition with gradient-based learning », in *Shape, contour and grouping in computer vision*, Springer, 1999, p. 319-345.
- [40] O. Russakovsky *et al.*, « Imagenet large scale visual recognition challenge », *International journal of computer vision*, vol. 115, n° 3, p. 211-252, 2015.
- [41] D. J. Hemanth et V. V. Estrela, *Deep learning for image processing applications*, vol. 31. IOS Press, 2017.
- [42] Y. Taigman, M. Yang, M. Ranzato, et L. Wolf, « Deepface: Closing the gap to human-level performance in face verification », in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, p. 1701-1708.
- [43] Z. MIMOUNE et A. OUAHAB, « Développement d'une Architecture Basée sur l'Apprentissage Profond (Deep Learning) pour la Détection d'Intrusion dans les Réseaux », Université Ahmed Draïa-Adrar, 2019.
- [44] K. Hornik, « Approximation capabilities of multilayer feedforward networks », *Neural networks*, vol. 4, n° 2, p. 251-257, 1991.
- [45] V. Nair et G. E. Hinton, « Rectified linear units improve restricted boltzmann machines », 2010.
- [46] A. L. Maas, A. Y. Hannun, et A. Y. Ng, « Rectifier nonlinearities improve neural network acoustic models », in *Proc. icml*, 2013, vol. 30, n° 1, p. 3.
- [47] K. He, X. Zhang, S. Ren, et J. Sun, « Delving deep into rectifiers: Surpassing human-level

- performance on imagenet classification », in *Proceedings of the IEEE international conference on computer vision*, 2015, p. 1026-1034.
- [48] C. Szegedy *et al.*, « Going deeper with convolutions », in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, p. 1-9.
- [49] K. He, X. Zhang, S. Ren, et J. Sun, « Deep residual learning for image recognition », in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, p. 770-778.
- [50] A. Krizhevsky, I. Sutskever, et G. E. Hinton, « ImageNet classification with deep convolutional neural networks », *Communications of the ACM*, vol. 60, n° 6, p. 84-90, 2017.
- [51] A. Krizhevsky, I. Sutskever, et G. E. Hinton, « Imagenet classification with deep convolutional neural networks », *Advances in neural information processing systems*, vol. 25, p. 1097-1105, 2012.
- [52] F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, et Y. Bengio, « Renet: A recurrent neural network based alternative to convolutional networks », *arXiv preprint arXiv:1505.00393*, 2015.
- [53] T. Brosch, Y. Yoo, L. Y. Tang, D. K. Li, A. Traboulsee, et R. Tam, « Deep convolutional encoder networks for multiple sclerosis lesion segmentation », in *International conference on medical image computing and computer-assisted intervention*, 2015, p. 3-11.
- [54] S. C. Wong, A. Gatt, V. Stamatescu, et M. D. McDonnell, « Understanding data augmentation for classification: when to warp? », in *2016 international conference on digital image computing: techniques and applications (DICTA)*, 2016, p. 1-6.
- [55] W. Al-Dhabyani, M. Gomaa, H. Khaled, et A. Fahmy, « Dataset of breast ultrasound images », *Data in brief*, vol. 28, p. 104863, 2020.

Liens internet

- [1] : <https://culturesciences.chimie.ens.fr/thematiques/chimie-physique/la-radiographie-i-histoire-de-la-decouverte-des-rayons-x-et-de-leur>