

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITE M'HAMED BOUGARA-BOUMERDES



Faculté de Technologie
Département Ingénierie des Systèmes Electriques
Mémoire de Master

Présenté par

Madani mehdi

Ali Allouache

Filière : Génie Biomédical
Spécialité : Instrumentation Biomédicale

Thème
Classification des images échocardiographie avec les
réseaux des neurones convolutionnels

Posé le 11 / 07 / 2021 devant le jury composé de :

Dr Messaoudi Nouredine	MCA	UMBB	Président
Dr Belkacem samia	MCA	UMBB	Examinatrice
Dr AMMAR Mohammed	MCA	UMBB	Encadreur

Année Universitaire : 2020/2021

Remerciements

Cette mémoire a été préparée au sein du département génie électrique

De l'université M'Hammed Bougera Boumerdes.

Tout d'abord nous tenant à exprimer notre grande reconnaissance pour notre encadreur Mr. Ammar Mohammed, professeur à l'université M'Hammed Bougera Boumerdes. Nous lui adressons un grand merci pour sa grande disponibilité, ses remarques et ses conseils tout au long de ce travail.

Nous adressons également notre profonde gratitude et notre reconnaissance à toute nos familles, nos très chers parents qu'Allah les bénisse, nos chers frères et sœurs a qui, sans leur soutien et encouragement, ce travail n'aurait pas vu le jour.

On remercie aussi nos amis et toutes les personnes qui nous sont très chères et qui ont contribué de près ou de loin à l'élaboration de ce travail

Sommaire

1	L'anatomie du cœur et ses pathologies	1
1.1	Anatomie.....	1
1.1.1	Tissus du cœur.....	3
1.2	Appareil cardiovasculaire	4
1.2.1	Circulation pulmonaire.....	4
1.2.2	Circulation systémique	4
1.3	Révolution cardiaque	6
1.4	Pathologies cardiaques.....	8
1.4.1	Cardiomyopathie dilatée	8
1.4.2	Cardiomyopathie hypertrophique	9
1.4.3	Myocardite.....	10
1.4.4	Infarctus du myocarde	11
2	<i>Imagerie par échographie</i>	13
2.1	Principes de base et fonctionnement.....	14
2.1.1	Ondes sonores	14
2.1.2	L'échographe	19
2.1.3	Génération de l'image	23
2.1.4	Rapport signal sur bruit	24
2.2	Échocardiographie	24
2.2.1	Caractéristiques	24
2.2.2	Résumé	27
3	<i>Convolutional neural network</i>	31
3.1	Introduction :.....	31
3.2	Définition de CNN :.....	31
3.3	Description des différents composants ou blocs de construction de base de CNN	33
3.3.1	Network Layers :	33
3.3.2	Couche convolutive :.....	33
3.3.3	Couche de mise en commun :.....	33
3.3.4	Fonctions d'activation (non linéaires) :.....	33

3.3.5	Couche entièrement connectée :	34
3.4	Fonctions de perte	35
3.5	Progrès récents dans les architectures CNN :	35
3.6	Classification des images :	35
3.6.1	LeNet-5 :	35
3.6.2	AlexNet :	36
3.6.3	ZFNet:	37
3.6.4	VGGNet :	37
3.6.5	GoogLeNet :	37
3.6.6	ResNet :	38
3.6.7	DenseNet :	39
3.7	Conclusion :	39
4	<i>Introduction</i>	41
4.1	Logiciels et librairies Utilisés dans l'implémentation	41
4.1.1	MATLAB	41
4.1.2	ResNet-50 :	41
4.1.3	Configuration Utilisé dans l'implémentation	42
4.1.4	Les base d'images.....	42
4.2	Plan de notre travail	43
4.1	Résultats obtenus et discussion.....	45
4.1.1	Résultats de la 1 ère expérience (avec un mini path de 32) :	45
4.1.2	Résultats de la 2 eme expérience (avec un mini path de 16) :	52
4.1.3	Résultats de la 3 eme expérience (avec un mini path de 64) :	59
5	Conclusion général.....	66

Listes des figures

<i>Figure I. 1 Anatomie du cœur humain coupé selon une vue en long axe. Image adaptée de [3].</i>	2
<i>Figure I. 2 Schéma des composantes de la paroi du cœur humain. Image tirée de [4].</i>	3
<i>Figure I. 3 Système cardiovasculaire chez l'humain. Image tirée du livre</i>	5
<i>Figure I. 4 Révolution cardiaque du cœur humain. Image tirée du livre [6].</i>	7
<i>Figure I. 5 Comparaison d'un cœur atteint de cardiomyopathie dilatée avec un cœur en santé. Image tirée du site [7].</i>	9
<i>Figure I. 6 Comparaison d'un cœur atteint de cardiomyopathie hypertrophique avec un cœur en santé. Image adaptée du site [8].</i>	10
<i>Figure I. 7 Comparaison d'un cœur atteint d'une myocardite avec un cœur en santé. Image adaptée du site [9].</i>	11
<i>Figure I. 8 Cause d'un infarctus du myocarde. Image tirée du site [10].</i>	12
<i>Figure I. 9 Comparaison d'un cœur après un infarctus avec un cœur en santé. Image tirée du site [11].</i>	12
<i>Figure I. 10 Trois acquisitions écho cardiographiques acquises dans des contextes différents (scanneur, opérateur et configuration).</i>	13
<i>Figure I. 11 Schéma illustrant la différence entre une onde transversale et une onde longitudinale.</i>	14
<i>Figure I. 12 Schéma illustrant une onde sonore émise par une source de son ainsi que son impact sur le milieu dans lequel elle se propage.</i>	15
<i>Figure I. 13 – Schéma illustrant la notion de période et longueur d'onde</i>	17
<i>Figure I. 14 Schéma illustrant le spectre des fréquences sonores. Image tirée du site [11].</i>	17
<i>Figure I. 15 Schéma illustrant le comportement d'une onde lors d'une rupture d'impédance acoustique.</i>	18
<i>Figure I. 16 Schéma illustrant les différentes composantes d'un échographe. Image tirée du site [12].</i>	19
<i>Figure I. 17 Schéma illustrant 3 formes de sondes écho cardiographiques ainsi que la forme des images générées par chacune d'elles. Figure adaptée de [14].</i>	21
<i>Figure I. 18 Schéma illustrant (a) une sonde échographique biplan commercialisée par la compagnie Vernon et (b) une sonde échographique biplan commercialisée par la compagnie Echoson.</i>	21
<i>Figure I. 19 Schéma illustrant une sonde 3D effectuant un balayage longitudinal.</i>	22
<i>Figure I. 20 Schéma illustrant trois autres types de balayages (axial, latéral et élévation) du transducteur 2D des sondes échographiques permettant la construction d'un volume 3D. Image tirée de [16].</i>	22
<i>Figure I. 21 Schéma illustrant la numérisation d'une acquisition échographique.</i>	23
<i>Figure I. 22 Schéma illustrant trois des fenêtres anatomiques permettant l'acquisition de différentes coupes du coeur humain avec une sonde échographique, soit la coupe parasternale, apicale et sous-costale. Dans cet exemple, les images proviennent de sondes 2D. Im.</i>	25
<i>Figure I. 23 Figure illustrant des coupes échocardiographie avec un schéma de l'anatomie visible par celle-ci. (a) Coupe apicale 4/5 cavités (4 ou 5 cavités du coeur sont visibles) (b) Coupe apicale 2/3 cavités (2 ou 3 cavités du coeur sont visibles) (c) Coupe par</i>	26
<i>Figure I. 24</i>	26
<i>Figure I. 25 – Figure illustrant un artefact d'ombrage sur une image échocardiographie. Image adaptée du site [19].</i>	28
<i>Figure I. 26 Figure illustrant un artefact de réverbération sur une image échocardiographie. Image adaptée du site [19].</i>	28
<i>Figure I. 27 Figure illustrant une acquisition échocardiographie où une importante partie du ventricule gauche est manquante suite à une mauvaise manipulation de la sonde par l'opérateur. La ligne jaune indique la région en question.</i>	29
<i>Figure I. 28 Exemples d'images écho cardiographiques ayant (a) beaucoup d'interférence à l'apex (pointe du coeur) et (b) un bruit ambiant si élevé que l'on voit à peine les parois du ventricule gauche.</i>	29
<i>Figure II. 1 Natural neurons.</i>	32
<i>Figure II. 2 artificiel neurones.</i>	32
<i>Figure II. 3 L'architecture des couches entièrement connectées[23]</i>	34
<i>Figure II. 4 L'architecture de LeNet-5[24].</i>	36
<i>Figure II. 5 The Architecture de AlexNet [25]</i>	36
<i>Figure II. 6 architecture de ZFNet [26].</i>	37

Figure II. 7 architecture de VGGNet[27].....	37
Figure II. 8 architecture de GOOGLE Net [28]	38
Figure II. 9 architecture de ResNet [29].....	39
Figure II. 10 architecture de DANSENet [30].....	39
Figure III. 1 base d'image utiliser pour l'apprentissage.....	43
Figure III. 2 coupes échocardiographie apical / parasternal grand / parasternal petit	44
Figure III. 3 tableau illustre le nombre de frame dans chaque axe de notre base de données	44
Figure III 4.1 distribution d'image dans la base de donnée (50%)	45
Figure III 4.2 confusion matrix de 50% de la base d'apprentissage	45
Figure III. 4 confusion matrix (0 ,5 mini path 32)	46
Figure III 5.1 distribution d'image dans la base de donnée (30%)	47
Figure III 5.2 confusion matrix de 30% de la base d'apprentissage	47
Figure III. 5 confusion matrix (0 ,3 mini path 32)	48
Figure III 6.1 distribution d'image dans la base de donnée (70%)	49
Figure III 6.2 confusion matrix de 70% de la base d'apprentissage	49
Figure III. 6 confusion matrix (0 ,7 mini path 32)	50
Figure III 7.1 distribution d'image dans la base de donnée (50%)	52
Figure III 7.2 confusion matrix de 50% de la base d'apprentissage	52
Figure III. 7 confusion matrix (0 ,3 mini path 16)	53
Figure III 8.1 distribution d'image dans la base de donnée (30%)	54
Figure III 8.2 confusion matrix de 30% de la base d'apprentissage	54
Figure III. 8 confusion matrix (0 ,5 mini path 16)	55
Figure III 9.1 distribution d'image dans la base de donnée (70%)	56
Figure III 9.2 confusion matrix de 70% de la base d'apprentissage	56
Figure III. 9 confusion matrix (0 ,7 mini path 16)	57
Figure III 10.1 distribution d'image dans la base de donnée (50%)	59
Figure III 10.2 confusion matrix de 50% de la base d'apprentissage	59
Figure III. 10 confusion matrix (0 ,3 mini path 64)	60
Figure III 11.1 distribution d'image dans la base de donnée (30%)	60
Figure III 11.2 confusion matrix de 30% de la base d'apprentissage	60
Figure III. 11 confusion matrix (0 ,5 mini path 64)	62
Figure III 12.1 distribution d'image dans la base de donnée (70%)	63
Figure III 12.2 confusion matrix de 70% de la base d'apprentissage	63
Figure III. 12 confusion matrix (0 ,7 mini path 64)	64

Liste des tableaux

tableau I 1 Impédances acoustiques de différents milieux et la vitesse de propagation de l'onde sonore correspondante [12]......	16
Tableau III. 1 résultat de sensibilité et taux de classification a 0 ,5 mini path 32.....	46
Tableau III. 2 résultat de sensibilité et taux de classification a 0 ,3 mini path 32.....	48
Tableau III. 3 résultat de sensibilité et taux de classification a 0 ,7 mini path 32.....	50
Tableau III. 4 résultat de sensibilité et taux de classification a 0 ,3 mini path 16.....	53
Tableau III. 5 résultat de sensibilité et taux de classification a 0 ,5 mini path 16.....	55
Tableau III. 6 résultat de sensibilité et taux de classification a 0 ,7 mini path 16.....	57
Tableau III. 7 résultat de sensibilité et taux de classification a 0 ,3 mini path 64.....	60
Tableau III. 8 résultat de sensibilité et taux de classification a 0 ,5 mini path 64.....	62
Tableau III. 9 résultat de sensibilité et taux de classification a 0 ,7 mini path 64.....	64

Abréviations

CDC: Centers for Disease Control and Prevention

FE : la fraction d'éjection

VES : le volume d'éjection systolique

VTD : volume télé diastolique

VTS : volume télé systolique

CMD : La cardiomyopathie dilatée

CMH : La cardiomyopathie hypertrophique

IRM : Imagerie par résonance magnétique

SNR : Le rapport signal sur bruit

CNN : Les réseaux de neurones network

ANN : Réseaux des neurones artificiels

MLP : perceptron multicouche traditionnel

2D/3D/4D : 2, 3 et 4 dimension

CPU : central processing unit

GPU: graphic processing unit

Résumé

Les réseaux de neurones convolutionnels sont des réseaux de neurones multicouches qui sont spécialisés dans des tâches de reconnaissance de forme. Notre travail consiste à reconnaître les différentes coupes écho cardiographiques en utilisant la classification par les réseaux de neurones convolutionnels à travers l'architecture de Resnet50 qui est constituée de nombreux blocs résiduels avec 3×3 couches de convolution. Les résultats obtenus ont montré que le choix du nombre d'époque et la taille de la base d'images ainsi que la profondeur du réseau ont une grande influence pour avoir des meilleurs résultats.

Mots-clés : réseaux de neurones convolutionnels, écho cardiographiques, reconnaissance, classification, Resnet50

ملخص

الشبكات العصبية التلافيفية هي شبكات عصبية متعددة الطبقات متخصصة في مهام التعرف على الأنماط. يتمثل عملنا في على شرائح صدى تخطيط القلب المختلفة باستخدام التصنيف حسب الشبكات العصبية التلافيفية من خلال بنيت Resnet50 التعرف التي تتكون من العديد من الكتل المتبقية ذات طبقات الالتفاف 3×3 . أظهرت النتائج التي تم الحصول عليها أن اختيار عدد الحقب وحجم قاعدة الصورة وكذلك عمق الحزوز له تأثير كبير للحصول على أفضل النتائج.

الكلمات المفتاحية : الشبكات العصبية التلافيفية، أصداء تخطيط القلب، والاعتراف، والتصنيف، Resnet50

Abstract

Convolutional Neural Networks are multi-layered neural networks that specialize in pattern recognition tasks. Our work consists in recognizing the different cardiographic echo slices using classification by convolutional neural networks through the architectures of Resnet50 which consists of many residual blocks with 3×3 convolution layers. The results obtained showed that the choice of the number of epochs and the size of the image base as well as the depth of the grating have a great influence to have the best results.

Keywords: convolutional neural networks, cardiographic echoes, recognition, classification, Resnet50

INTRODUCTION GENERALE

Nous vivons dans un monde numérique, où les informations sont stockées, traitées, indexées et recherchées par des systèmes informatiques, ce qui rend leur récupération une tâche rapide et pas cher. Au cours des dernières années, des progrès considérables ont été réalisés dans le domaine de classification d'images. Ce progrès est dû aux nombreux travaux dans ce domaine et à la disponibilité des bases d'images internationales qui ont permis aux chercheurs de signaler de manière crédible l'exécution de leurs approches dans ce domaine, avec la possibilité de les comparer à d'autres approches qu'ils utilisent les mêmes bases.

Dans notre projet on va étudier et appliquer les réseaux de neurones convolutionnels pour classifier des images échocardiographie. L'échocardiographie est l'examen de routine en cardiologie. L'apprentissage en profondeur a été utilisé dans l'analyse d'images médicales mais n'a pas encore été largement appliqué à l'échocardiographie, en partie en raison de la complexité des échocardiogrammes format multi-vues et multi-modalités. La première étape essentielle vers une solution complète assistée par ordinateur consiste à reconnaître les différentes coupes écho cardiographiques.

Dans notre travail nous avons utilisée le modèle resnet50 pour reconnaître trois plan de coupe : -coupe apical, -coupe parasternnale grande et parasternnale petite.

La classification des images a reconnu une avancée majeure en termes des performances grâce aux réseaux des neurones convolutifs (Convolutionals neural network CNN) qui a plusieurs modèles implémenter (LeNet-5, alexNet, ZFNet, VGGNet, GoogleNet, Resnet, DenseNet ...etc.)

Pour ce faire, nous avons structuré notre mémoire en trois chapitres :

- Pour le premier chapitre on va présenter l'anatomie de cœur et ces pathologies et l'imagerie par échographier
- Le deuxième chapitre est consacré à la description des réseaux de neurones convolution els ainsi que leurs l'intérêt dans le domaine de la classification des images.
- Dans le troisième chapitre, on va montrer la partie expérimentale de notre travail et on discute les différents résultats obtenus et à la fin on termine par une conclusion générale.

Chapitre 1

L'anatomie du cœur et ses
pathologies

1 L'anatomie du cœur et ses pathologies

Pendant des décennies, les maladies cardiaques ont reçu une grande attention de la médecine et de la recherche. En fait, c'est l'une des principales causes de décès en Amérique du Nord. Selon Statistique Canada [1], les maladies cardiaques sont la deuxième cause de décès au Canada depuis 2010. En 2011, les maladies cardiaques représentaient 19,7 % des décès au Canada. Aux États-Unis, selon un article du CDC (Centers for Disease Control and Prevention), elle est même devenue la première cause de décès depuis 1958. [2]

Dans ce chapitre notre objectif est d'apprendre et de comprendre les structures et les fonctions des principaux systèmes corporels et les structures du corps et leur fonctionnement

1.1 Anatomie

Comme nous le savons tous, le cœur est un organe utilisé pour faire circuler le sang dans le corps pour fournir de l'oxygène. Mais comment ça fonctionne ? Afin de bien comprendre l'impact des projets de recherche couverts dans cet article, il faut avoir une compréhension de base de la structure anatomique, du rôle et de la fonction du cœur. La figure 1.1 illustre l'anatomie du cœur humain. Il est situé sous le sternum, à peu près de la taille d'un poing fermé, et est principalement composé de tissu musculaire provenant du cœur. Bien que cela puisse paraître compliqué à première vue, la structure du cœur peut se résumer ainsi : C'est un organe composé de quatre chambres (deux oreillettes et deux ventricules), chacune avec une valve. Il est relié à d'autres parties du corps par différentes artères,.

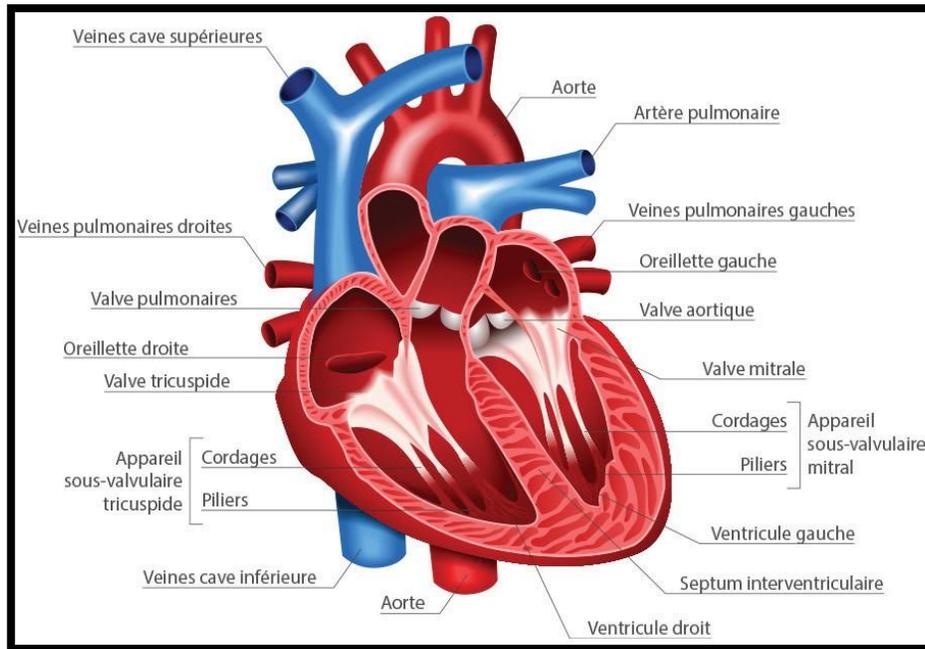


Figure I. 1 Anatomie du cœur humain coupé selon une vue en long axe. Image adaptée de [3].

La figure 1.2 offre une perspective plus détaillée sur la structure de la paroi du cœur. Elle est essentiellement composée de trois couches, soit l'endocarde à l'intérieur, le myocarde au milieu et l'épicarde à l'extérieur (à ne pas confondre avec le péricarde, l'enveloppe qui contient le cœur). L'endocarde est une mince membrane qui couvre l'intérieur des cavités et des valves du cœur. Le myocarde, considérablement plus épais, est le muscle cardiaque qui se contracte pour pomper le sang. L'épicarde (aussi une mince membrane) recouvre l'extérieur du cœur [3].

1.1.1 Tissus du cœur

Le cœur est composé d'un mélange de différents tissus, notamment du tissu épithélial, du tissu conjonctif, du tissu musculaire et du tissu nerveux. Le tissu épithélial, également appelé tissu de revêtement, agit comme une membrane recouvrant les surfaces interne (endocarde) et externe (épicarde) de la paroi cardiaque. L'endocarde est en contact avec le sang, tandis que l'épicarde est en contact avec le liquide péricardique. Leurs cellules sont étroitement regroupées, ce qui rend ce tissu résistant aux dommages mécaniques, aux microbes et à la perte de liquide. Le rôle du tissu conjonctif est de fixer et de soutenir les autres tissus. Contrairement au tissu épithélial, les cellules du tissu conjonctif sont peu nombreuses et éloignées les unes des autres. Le tissu musculaire a la capacité de se contracter sous la stimulation de l'influx nerveux. Ce type de tissu est le plus abondant dans le cœur et forme la partie de contraction : le muscle cardiaque. Enfin, le tissu nerveux perçoit et transmet au tissu musculaire des directives sous forme d'influx nerveux. En d'autres mots, c'est le tissu responsable du contrôle du bon fonctionnement du cœur. [4].

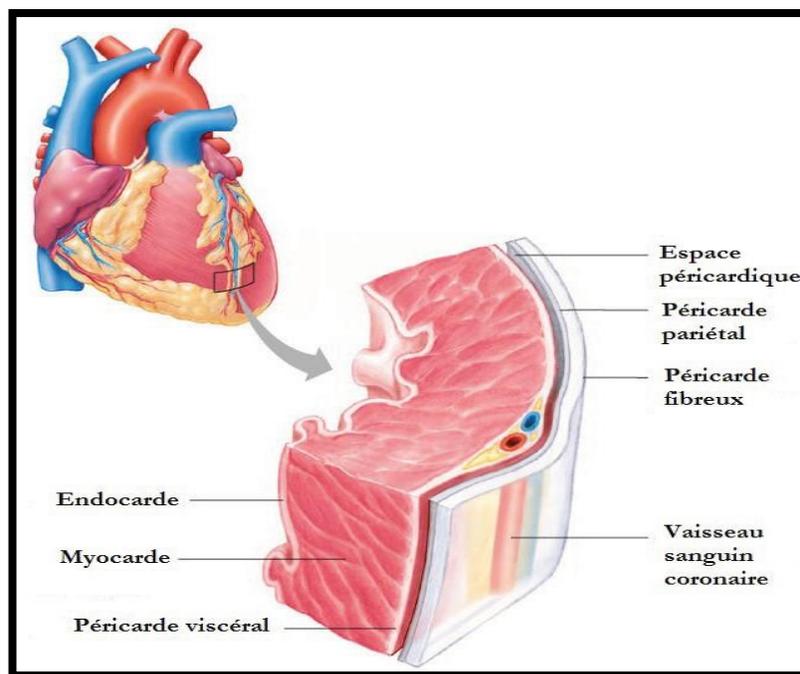


Figure I. 2 Schéma des composantes de la paroi du cœur humain. Image tirée de [4].

1.2 Appareil cardiovasculaire

Le cœur humain fait partie d'un système circulatoire complexe nommée appareil cardiovasculaire ou système sanguin. Ce système est responsable de transporter le sang du cœur vers les extrémités du corps et les divers organes, puis de le réachemine vers le cœur. Ce système fermé, illustré à la figure 1.3, est divisé en 2 parties : la circulation pulmonaire (en bleu) et la circulation systémique (en rouge).

1.2.1 Circulation pulmonaire

La fonction de la circulation pulmonaire, aussi appelée petite circulation, est d'acheminer le sang pauvre en dioxygène et riche en gaz carbonique aux poumons, puis de renvoyer le sang réoxygéné au cœur. Pour y arriver, le ventricule droit pompe d'abord le sang vers les poumons par l'intermédiaire du tronc pulmonaire. Celui-ci se subdivise en deux artères pulmonaires droite et gauche se connectant chacun à un poumon. En s'écoulant dans les poumons, le sang capte le dioxygène (O₂) et libère du dioxyde de carbone (CO₂). Enrichi d'oxygène, le sang est acheminé à l'oreillette gauche du cœur via les veines pulmonaires droite et gauche respectives au poumon. Puis, l'oreillette gauche se contracte de façon à acheminer le sang jusqu'au ventricule gauche.

1.2.2 Circulation systémique

La fonction de la circulation systémique, aussi appelée grande circulation, est d'acheminer le sang riche en dioxygène à tous les organes ainsi qu'aux extrémités du corps, puis de renvoyer le sang pauvre en dioxygène au cœur.

Pour y parvenir, le ventricule gauche expulse le sang riche en dioxygène par l'aorte qui elle est connectée aux déférentes artères parcourant le corps. Les premières connectées à l'aorte, nommées artères coronaires (non visibles dans la FIGURE 1.1), servent à alimenter le muscle cardiaque en dioxygène. Les artères suivantes débouchent sur les lits capillaires des membres supérieurs, soit la tête et les bras. Puis, l'aorte se prolonge et débouche sur les artérioles et les lits capillaires des membres inférieurs, de façon à approvisionner les organes abdominaux et les jambes. Dans les lits capillaires, le sang cède la majorité de son dioxygène aux tissus et reçoit le dioxyde de carbone produit par la

respiration cellulaire. Les capillaires s'agglomèrent en veinules, puis en veines. Ainsi, le sang désoxygéné provenant des extrémités du corps se canalise dans la veine cave supérieure pour le haut du corps et dans la veine cave inférieure pour le sang provenant du tronc et du bas du corps. Les deux veines caves acheminent leur sang

Dans l'oreillette droite qui à son tour le déverse dans le ventricule droit avant d'être réacheminé aux poumons pour s'enrichir à nouveau de dioxygène

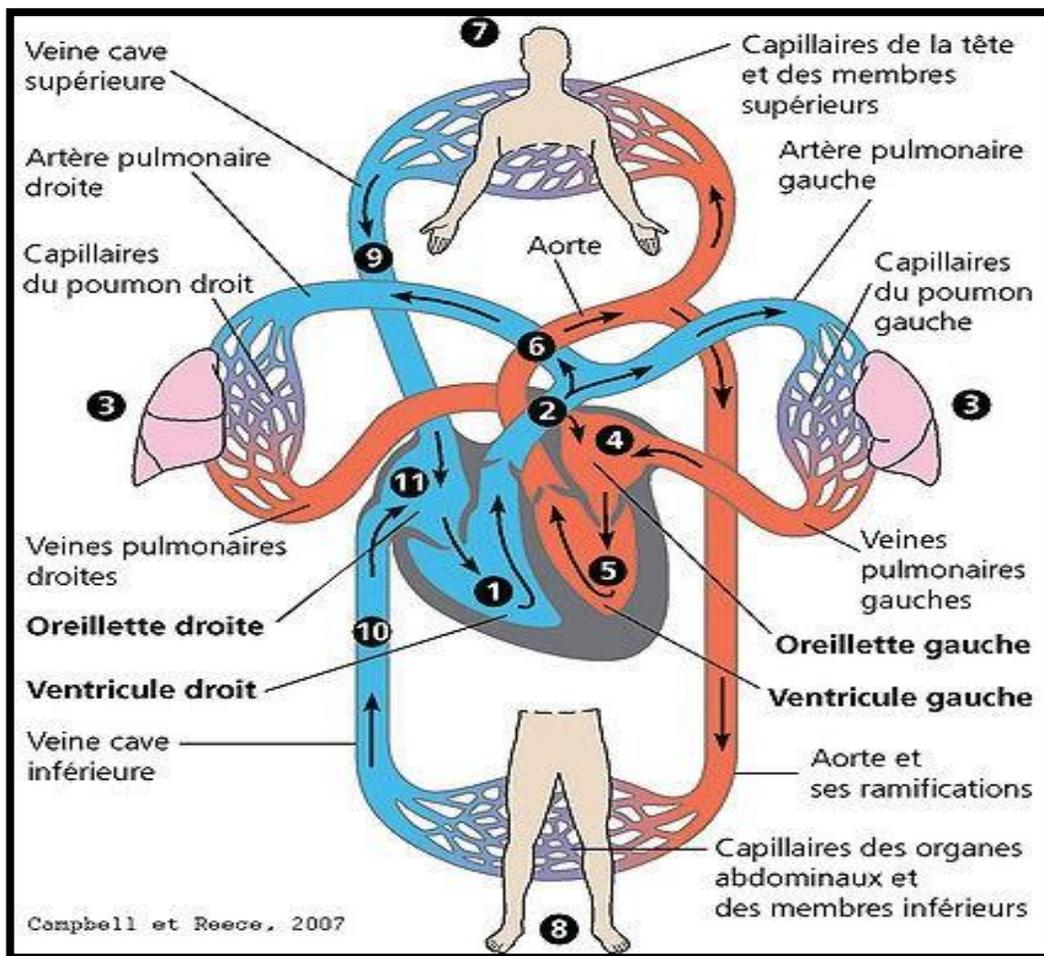


Figure I. 3 Système cardiovasculaire chez l'humain. Image tirée du livre

1.3 Révolution cardiaque

Le cœur pompe le sang pour fournir de l'oxygène au corps en battant, ce qui est aussi appelé « rotation du cœur » dans le domaine médical. Comme le montre la figure 1.4, celui-ci est divisé en trois étapes principales du cycle : la diastole, la systole auriculaire et la systole ventriculaire. La diastole (période de relaxation) est la période pendant laquelle le sang de la veine cave et des veines pulmonaires s'écoule dans l'oreillette puis dans le ventricule. En d'autres termes, le stade où le cœur est plein de sang. C'est à ce moment que le cœur est le plus dilaté.

Ensuite, il y a une brève période de contraction auriculaire, où tout le sang restant dans l'oreillette est transporté vers le ventricule. Une fois que le sang est drainé de l'oreillette, la valve auriculo-ventriculaire se ferme, empêchant le sang de retourner dans l'oreillette. Enfin, dans la dernière phase du cycle, lorsque les ventricules se contractent, le sang est éjecté dans le tronc des poumons et l'aorte.

C'est à cette étape que le ventricule gauche est le plus contracté.

Dans une révolution cardiaque, les phases (diastolique, systolique auriculaire et systolique ventriculaire) durent chacune 0.4, 0.1 et 0.3 seconde respectivement pour un total de 0.7 seconde. Au repos, le cœur d'un humain adulte a une fréquence cardiaque moyenne de 75 battements avec un débit de sang évacué de 4.75 litres de sang par minute. De plus, il peut battre plus de 2 milliards de fois en une vie.

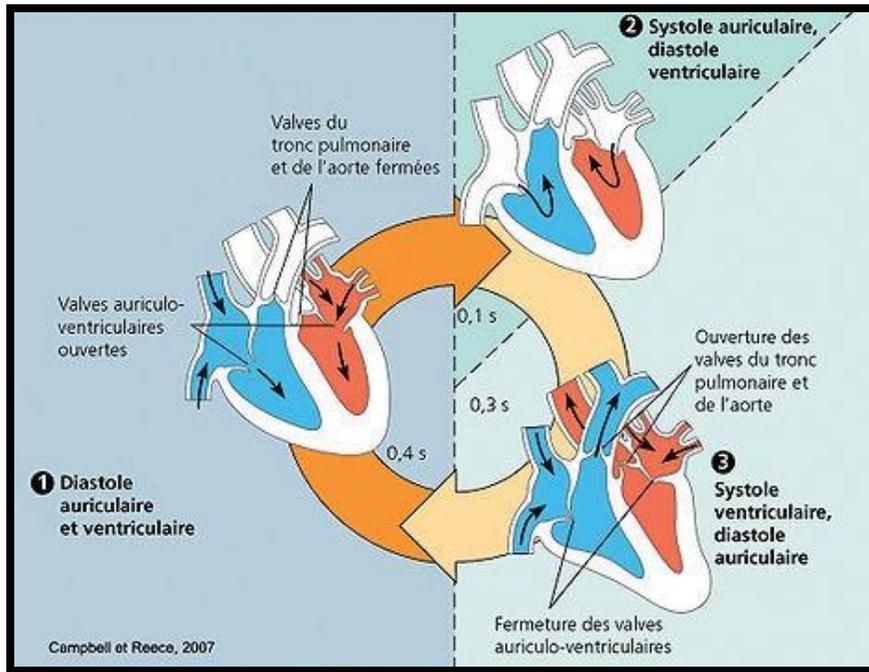


Figure I. 4 Révolution cardiaque du cœur humain. Image tirée du livre [6].

Une métrique importante utilisée pour quantifier de façon relative le sang évacué à chaque battement est la fraction d'éjection (FE). La FE représente le ratio de sang éjecté du ventricule gauche lors d'un battement et elle est régie par l'équation suivante :

$$FE(\%) = \frac{VES}{VTD} * 100 \text{ où } VES = VTD - VTS. \quad (1.1)$$

Ainsi, nous calculons la FE en divisant le volume d'éjection systolique (VES) par le volume télédiastolique (VTD), puis en multipliant par 100 pour obtenir un pourcentage. Le VES est le volume de sang éjecté à chaque contraction tandis que le VTD est le volume de sang présent dans le ventricule gauche avant la contraction, soit en phase de diastole. Le VES se calcule simplement en soustrayant le VTD avec le volume télésystolique (VTS). Le VTS est le volume de sang restant dans le ventricule gauche après la contraction, soit en phase de systole. Un cœur humain en bonne santé aura une FE de l'ordre de 50 à 70%. Toutefois, un cœur possédant une FE dans cette gamme n'est pas nécessairement en santé ; bien qu'elle soit utile et importante au diagnostic de maladies cardiaques, c'est une information complémentaire à d'autres observations.

1.4 Pathologies cardiaques

L'appareil circulatoire peut être atteint de plusieurs conditions pathologiques affectant ses fonctions et son comportement de différentes façons. Parmi les plus courantes, les examens sélectionnés pour cette étude permettent d'observer 4 pathologies cardiaques différentes, soit la cardiomyopathie dilatée, la cardiomyopathie hypertrophiée, l'infarctus du myocarde ainsi que la myocardite. Étant donné des fonctions vitales du cœur, il est important de comprendre l'impact qu'ont ces conditions sur le système cardiovasculaire.

1.4.1 Cardiomyopathie dilatée

La cardiomyopathie dilatée (CMD) est une maladie du muscle cardiaque où le muscle du ventricule (gauche ou droit) est affaibli. Ainsi, la cavité cardiaque est constamment dilatée à différents niveaux telle qu'illustrée à la figure 1.5, que ce soit pendant la phase de diastole ou de systole. Cela provoque une diminution non négligeable de la force de contraction du muscle cardiaque et de ce fait, la FE en est abaissée. Une FE diminuée signifie que la quantité de sang évacué proportionnellement au volume diastolique est moindre et donc que généralement moins de sang oxygéné est acheminé vers les organes vitaux et les membres externes du corps. Si la FE est trop basse, autrement dit, si le cœur ne réussit pas à pomper suffisamment de sang pour subvenir aux besoins du corps, nous nous retrouvons avec un cas d'insuffisance cardiaque. S'il y a pénurie d'oxygène, les tissus peuvent se dégrader de façon irréversible, voire létale. À noter qu'il est possible que le volume de sang éjecté à chaque battement soit constant malgré une FE abaissée si le volume diastolique est augmenté par une déformation du VG. Dans un contexte de CMD, ce phénomène se nomme remodelage du ventricule gauche.

1.4.2 Cardiomyopathie hypertrophique

La cardiomyopathie hypertrophique (CMH) est une maladie du muscle cardiaque qui atteint la structure même des cellules. Cette anomalie fait en sorte que le muscle cardiaque s'épaissit à différents endroits dépendamment du cas. La figure 1.6 en démontre un exemple. D'un patient à l'autre, le muscle cardiaque peut être hypertrophié de façon symétrique, c'est-à-dire que les parois sont atteintes uniformément, ou bien asymétrique, où certaines parois sont plus affectées que d'autres. Ainsi, les ventricules gauches de cœurs atteints de CMH peuvent se manifester sous une multitude de formes différentes. Ceci représente un défi additionnel à sa segmentation. Dans certains cas, le muscle peut être épaissi à la sortie du cœur, pouvant potentiellement créer un blocage. Un trait commun des cas de CMH est une FE généralement élevée. En effet, l'état du muscle cardiaque réduit considérablement la capacité du ventricule gauche à se remplir de sang et cause donc de l'insuffisance cardiaque.

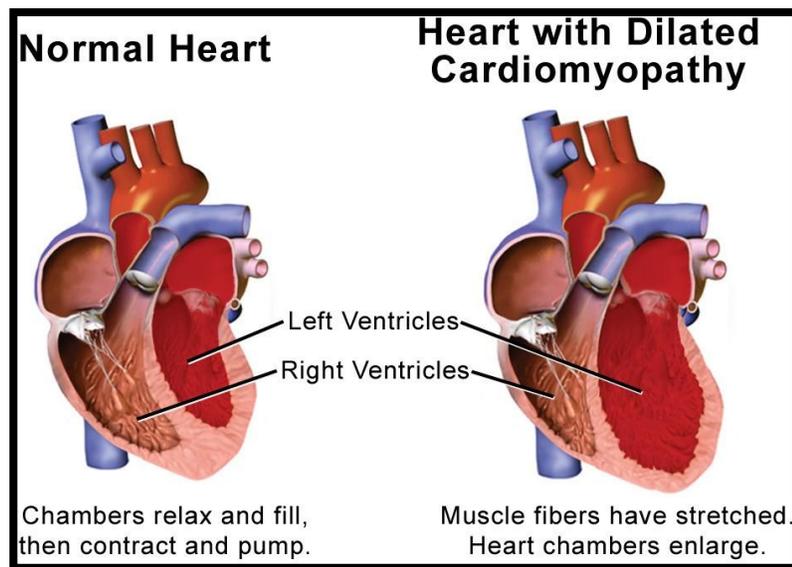


Figure 1.5 Comparaison d'un cœur atteint de cardiomyopathie dilatée avec un cœur en santé. Image tirée du site [7].

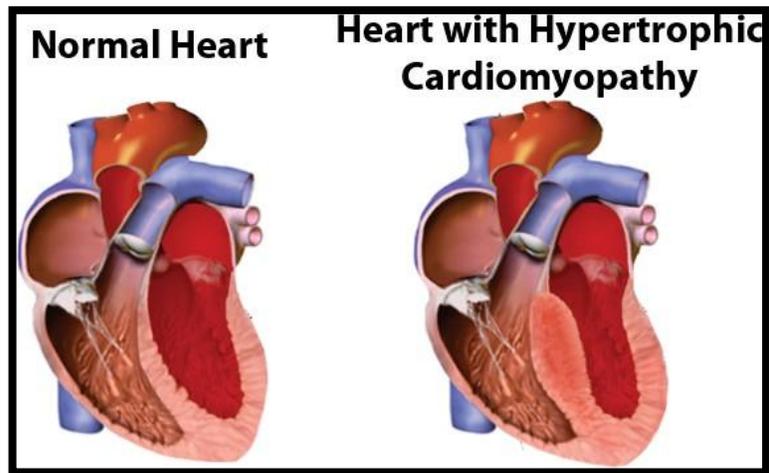


Figure 1.6 Comparaison d'un cœur atteint de cardiomyopathie hypertrophique avec un cœur en santé. Image adaptée du site [8]

1.4.3 Myocardite

La myocardite est une inflammation du muscle cardiaque causée principalement par une infection virale pouvant potentiellement entraîner une destruction de ses cellules. L'inflammation du myocarde diminue la force de contraction du cœur, donc sa capacité d'alimenter la totalité des organes du corps en dioxygène. Contrairement aux pathologies cardiaques présentées précédemment, cette maladie peut être traitée relativement facilement. Souvent, les patients atteints d'une myocardite se rétablissent avec le temps sans aide externe. Toutefois, il se peut que la condition s'aggrave. Dans le cas échéant, il existe des traitements antiviraux pour accélérer le rétablissement. Comme le démontre la figure 1.7, l'inflammation du muscle cardiaque peut être inégal, donnant à la surface de la cavité une allure gondolée

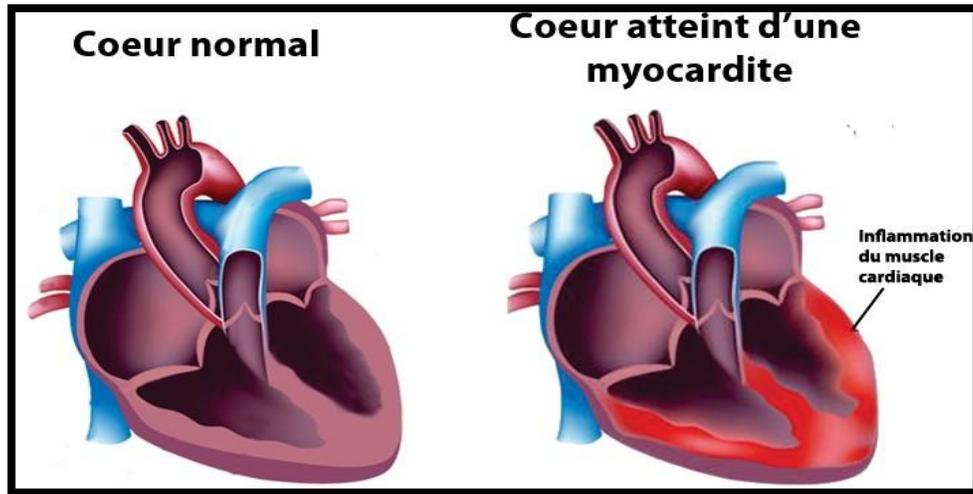


Figure I. 7 Comparaison d'un cœur atteint d'une myocardite avec un cœur en santé. Image adaptée du site [9]

1.4.4 Infarctus du myocarde

L'infarctus du myocarde est une condition cardiaque définie par la mort cellulaire (nécrose) d'une région plus ou moins importante du muscle cardiaque. Il survient lorsque le myocarde ne reçoit pas assez de sang pour s'alimenter en dioxygène. Comme illustré à la figure 1.8, une telle situation se produit lorsqu'une (ou plusieurs) artère coronaire, chargée d'acheminer l'oxygène au myocarde, est obstruée par des plaques de cholestérol. L'artère ainsi bouchée, le débit de sang devient insuffisant pour garder le muscle cardiaque en vie. Lorsqu'il y a rupture de ces plaques, un caillot sanguin se forme à l'intérieur de l'artère coronaire, obstruant ainsi complètement celle-ci. Le terme « crise cardiaque » est utilisé pour caractériser les cas aigus d'infarctus du myocarde. Les dommages sont irréversibles et comme illustré à la figure 1.9, la paroi cicatrisée du myocarde déforme considérablement les cavités. Puisque les régions nécrosées peuvent être localisées à différents endroits sur la surface du cœur, il n'y a pas de forme générale de l'allure du ventricule gauche atteint. Les parties nécrosées perdent leur capacité à se contracter et la capacité du cœur à se contracter pour faire circuler le sang en est affectée. En effet, si le ventricule gauche peine à se contracter, la quantité de sang pompé à chaque battement est moindre. Conséquemment, la FE de cœurs atteints d'infarctus du myocarde est très faible.

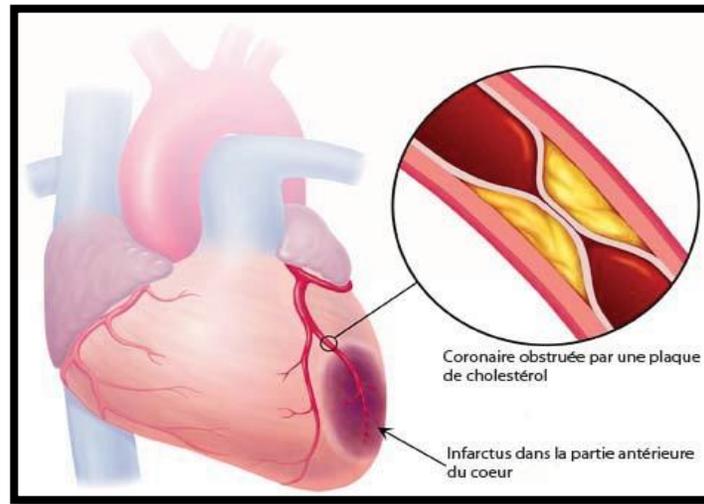


Figure I. 8 Cause d'un infarctus du myocarde. Image tirée du site [10].

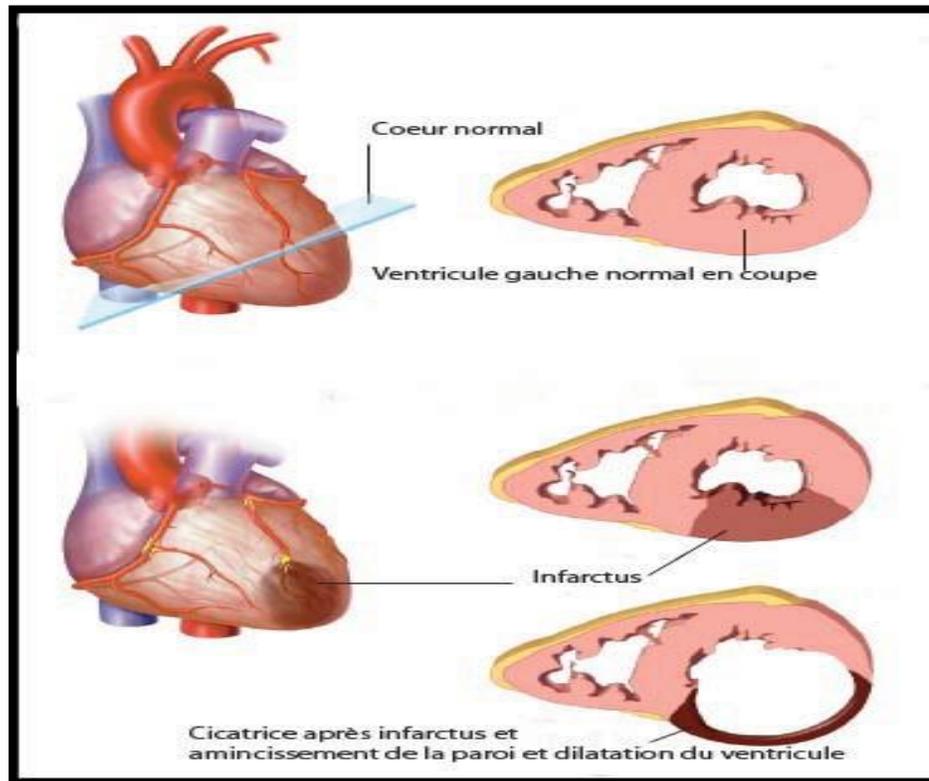


Figure I. 9 Comparaison d'un cœur après un infarctus avec un cœur en santé. Image tirée du site [11].

2 Imagerie par échographie

L'échographie est une technique d'imagerie basée sur la réflexion des ultrasons utilisés en médecine depuis 1952. Cette modalité d'imagerie possède de nombreux avantages : elle est non invasive, peu dispendieuse et les résultats sont disponibles en temps réel. De plus, un échographe est simple d'utilisation par rapport à d'autres technologies comme les scanners d'imagerie par résonance magnétique (IRM). Cependant, cette technique possède des défauts non négligeables. Tout d'abord, les images d'examen échographiques manquent généralement de netteté, sont souvent bruitées et peuvent contenir des régions souffrant de baisse de signal. De plus, le résultat visuel est très dépendant de l'équipement, de la calibration de l'expert ainsi que son aptitude à effectuer l'examen tel qu'on peut observer avec la figure 2.1. Enfin, puisque l'air (présent dans les poumons) et les os dévient et bloquent la propagation des ultrasons destinés à être transmis par les tissus mous, cette technologie n'est pas utile pour visualiser l'entièreté du corps humain.

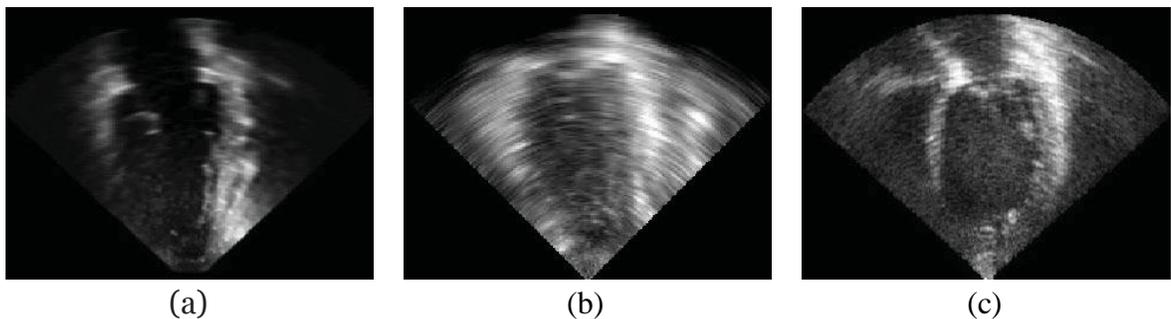


Figure I. 10 Trois acquisitions écho cardiographiques acquises dans des contextes différents (scanneur, opérateur et configuration).

2.1 Principes de base et fonctionnement

Dans cette section, des concepts de bases nécessaires à la compréhension de la technologie derrière les images échographiques sont couverts.

2.1.1 Ondes sonores

Par définition, une onde sonore (ou onde acoustique), est une perturbation mécanique longitudinale qui se propage dans un milieu physique (comme l'eau ou l'air). En d'autres mots, c'est une vibration ayant la forme d'une onde longitudinale qui traverse un milieu. Une onde longitudinale est un type d'onde où la déformation du milieu s'effectue à une certaine fréquence dans la même direction que la propagation, contrairement à une onde transversale où la déformation s'effectue dans la direction perpendiculaire à la propagation. Ce phénomène est illustré à la figure 2.2.

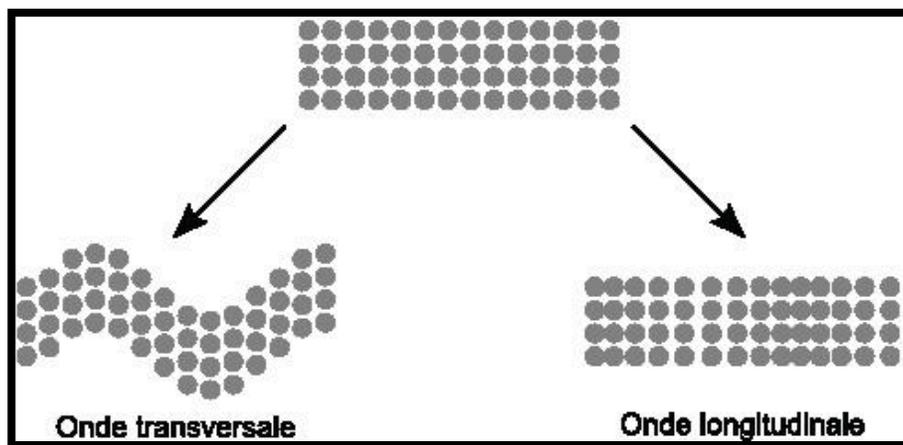


Figure I. 11 Schéma illustrant la différence entre une onde transversale et une onde longitudinale.

Les ronds gris représentent la matière d'un milieu comme l'air ou l'eau. En imaginant une propagation allant de la gauche vers la droite, nous observons une perturbation perpendiculaire à celle-ci pour l'onde transversale et parallèle pour l'onde longitudinale.

L'onde sonore produit des zones de compression et de raréfaction des molécules du milieu dans laquelle elle se propage. C'est ce qu'on appelle la vibration sonore et ce phénomène est illustré à la figure 2.3. L'onde sonore va se comporter différemment d'un milieu à l'autre en raison de la structure

des molécules. Par exemple, si les molécules sont très compactes, il sera plus difficile de les faire vibrer, donc pour l'onde de traverser le milieu. Chaque milieu possède sa propre résistance quant à la propagation des ondes acoustique. Autrement dit, les ultrasons peuvent se propager plus facilement dans certains milieux par rapport à d'autres.

Ce concept se nomme l'impédance acoustique. Les impédances acoustiques des principaux tissus ainsi que leur vitesse de propagation c correspondante sont présentées dans le tableau 2.1.

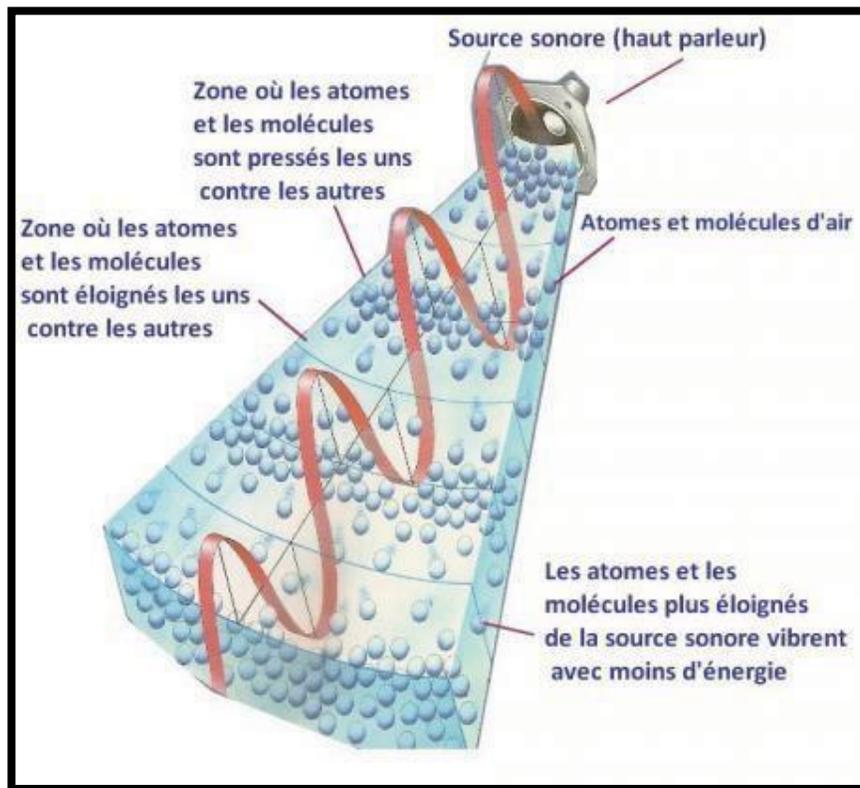


Figure I. 12 Schéma illustrant une onde sonore émise par une source de son ainsi que son impact sur le milieu dans lequel elle se propage.

Dans cet exemple, le milieu est de l'air. Image tirée du site [8].

Tissus	Impédance acoustique (105kg/m ² /s)	Vitesse de propagation (m/s)
Air	0.0004	343
Sang	1.66	1566
Foie	1.66	1566
Rein	1.62	1567
Graisse	1.33	1446
Os	3.75 - 7.38	2070 - 5350

tableau I. 1 Impédances acoustiques de différents milieux et la vitesse de propagation de l'onde sonore correspondante [12].

Afin de quantifier et de décrire les ondes sonores, il faut connaître les définitions et concepts suivants. D'abord, l'amplitude d'une onde sonore correspond aux variations de pression du milieu dans lequel elle se propage. Comme qu'illustrée à la figure 2.3, l'amplitude est élevée lorsque les molécules du milieu sont comprimées (pression forte) et basse lorsque les molécules du milieu sont plus éloignées (pression basse). Puis, on nomme « la période » (notée T) l'intervalle de temps qui sépare deux états d'amplitude (de pression) successifs et identiques. La longueur d'onde est la distance physique parcourue par l'onde pendant une période, ou en d'autres mots, la distance qui sépare deux molécules successives dans le même état vibratoire. Ces deux derniers concepts sont illustrés à la figure 2.4.

La fréquence d'une onde, dont l'unité est le Hertz (Hz), est le nombre de périodes complétées par secondes. Ainsi, nous calculons la fréquence avec l'inverse de la période :

$$f = 1/T \text{ où } T \text{ est en secondes.} \quad (2.1)$$

Par exemple, la fréquence d'une onde sonore possédant une période T de 0.2 seconde est de 5Hz. Grâce à l'ouïe, l'humain a la capacité d'entendre les sons dont la fréquence est entre 20 Hz et 20 kHz. Les ondes sonores possédant une fréquence plus basse ou plus élevée que cet intervalle sont inaudibles

par l'humain et sont communément appelées infrasons et ultrasons respectivement. Le spectre des ondes sonores est illustré à la figure 2.5

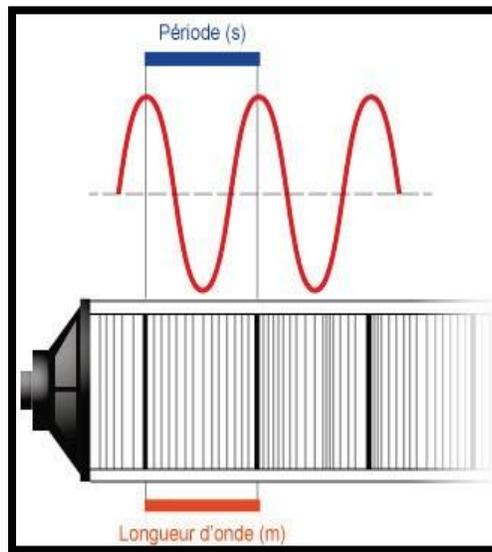


Figure I. 13 – Schéma illustrant la notion de période et longueur d'onde

Le haut de la figure représente la variation de l'amplitude de l'onde sonore émise par le haut-parleur au bas de la figure. L'espacement entre les lignes de la partie du bas de la figure illustre la compression des molécules du milieu dans lequel le son se propage. Les lignes noires représentent une pression élevée et les zones où les lignes sont plus espacées représentent une pression moindre. Image tirée du site [13].

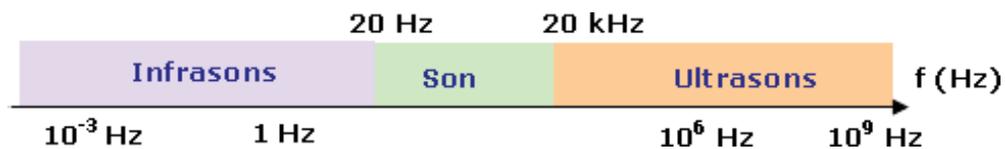


Figure I. 14 Schéma illustrant le spectre des fréquences sonores. Image tirée du site [11].

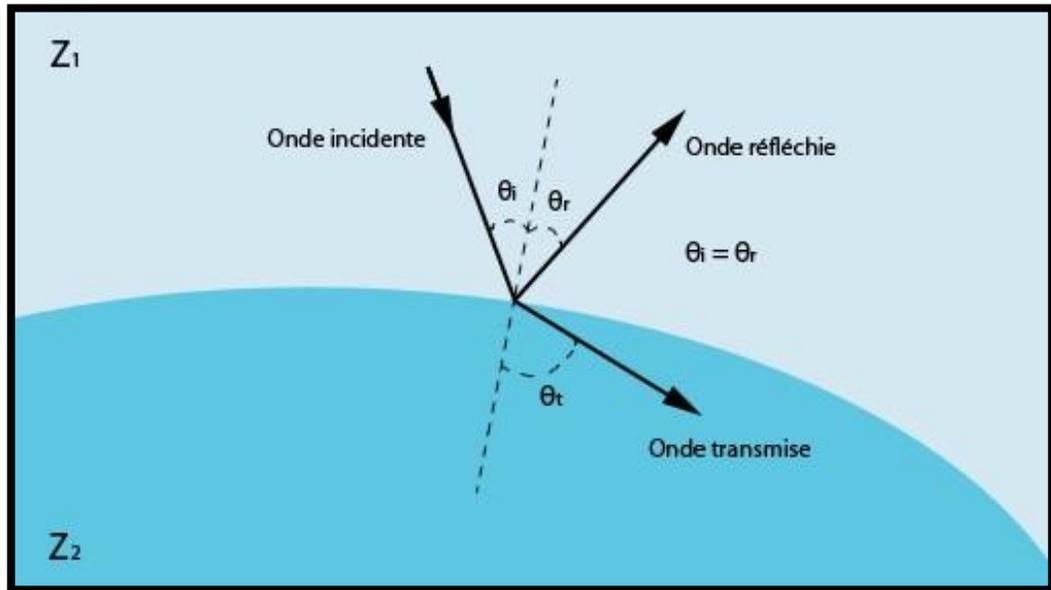


Figure I. 15 Schéma illustrant le comportement d'une onde lors d'une rupture d'impédance acoustique.

Enfin, la vitesse de propagation de l'onde sonore c dépend de l'impédance acoustique du milieu Z ainsi que de sa masse volumique ρ selon l'équation suivante :

$$c = Z/\rho. \quad (2.2)$$

Lorsqu'une onde passe d'un milieu à un autre (concept nommé rupture d'impédance), une partie de celle-ci est transmise en profondeur avec une nouvelle vitesse et une autre partie est réfléchie suivant les lois de l'optique géométrique. Ce phénomène est illustré à la figure 2.6.

Le coefficient de réflexion de l'onde R dépend directement des impédances acoustiques des deux milieux de propagation et est défini par l'équation suivante :

$$R = (Z_2 - Z_1) / (Z_2 + Z_1)^2. \quad (2.3)$$

Lors d'une rupture d'impédance entre l'air et un tissu mou, le coefficient de réflexion R peut atteindre presque 100%. Par contre, entre deux tissus mous d'impédance acoustique similaire, R est généralement inférieur à 10% et plus de 90% de l'onde est transmise au deuxième milieu. Ainsi, le signal émis par la sonde peut atteindre une plus grande distance, permettant de visualiser des organes situés plus loin de l'origine de la sonde.

2.1.2 L'échographe

Un examen échographique est effectué à l'aide d'un échographe tel qu'illustré à la figure 2.7. Il est constitué d'au moins une sonde (servant à l'émission et la réception d'ultrasons), d'un système de visualisation (moniteur), d'un système informatique (servant à interpréter le délai entre la réception et l'émission des ultrasons en image), d'une console de commande (servant à la saisie des données du patient et à la calibration de l'équipement) ainsi que d'un système d'enregistrement des données. Généralement, il est possible d'enregistrer l'information de façon numérique (image DICOM) ou analogique (vidéo, impression).

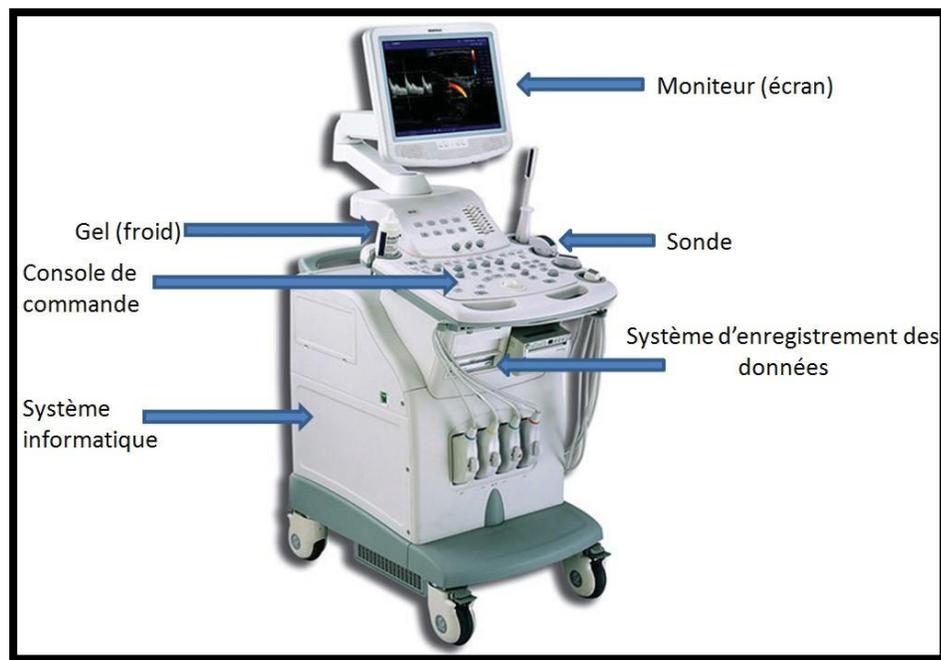


Figure I. 16 Schéma illustrant les différentes composantes d'un échographe. Image tirée du site [12]

La sonde est placée directement sur le corps du patient, puis émet des ultrasons qui se propagent dans les tissus pour ensuite récupérer les ondes réfléchies par ceux-ci. Bien que la sonde soit appliquée directement sur la peau, il existe une fine couche d'air entre la sonde et celle-ci, créant une rupture d'impédance avec un coefficient de réflexion R si élevé de sorte que seule une insignifiante proportion est transmise dans le corps pour sa visualisation. Pour pallier à ce problème, un gel avec une impédance acoustique similaire à celle de la peau est appliqué pour assurer un contact complet. Ainsi, la majorité des ultrasons sont transmis aux tissus mous.

Il existe plusieurs catégories de sondes permettant de répondre aux besoins des centres hospitaliers. Il y a les transducteurs 2D, les transducteurs biplans et les transducteurs 3D/4D. Les transducteurs 2D sont les plus couramment utilisés en clinique et permettent, comme le nom l'indique, de générer une image échographique en 2D. Elles existent sous plusieurs formes de façon à s'adapter aux différentes contraintes anatomiques à considérer lors de l'acquisition des données. La figure 2.8 illustre trois formes de sondes échographiques. À noter qu'il existe plusieurs dizaines de types de sondes différentes.

Les sondes biplans servent à générer simultanément deux images échographiques 2D perpendiculaires et sont majoritairement utilisées sur les sondes que l'on insère dans des cavités du corps humain. Les modèles et les plans choisis varient d'un fabricant à l'autre. La figure 2.9 illustre deux sondes de cette catégorie.

Les sondes échographiques 3D/4D sont essentiellement des transducteurs 2D motorisés, permettant d'effectuer un balayage mécanique dans l'espace. La figure 2.10 illustre une telle sonde effectuant un balayage longitudinal. Ainsi, une multitude d'images 2D sont acquises pour former un éventail d'ultrason. L'image 3D est créée en interpolant les images 2D. Dans le cas des images 4D, c'est le même principe, mais à répétition à chaque temps t . Il existe des types de balayages du transducteur 2D autres que longitudinal permettant de reconstruire un volume 3D : le mode axial, latéral et élévation. Ceux-ci sont illustrés à la figure 2.11

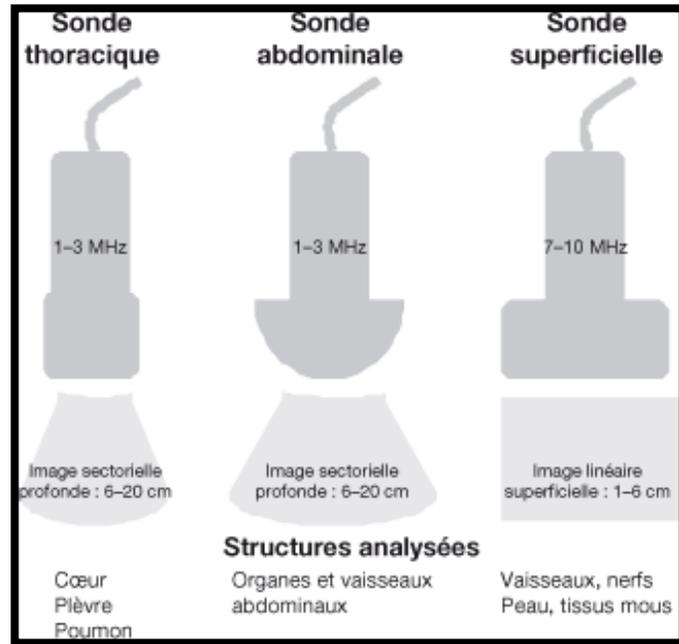


Figure I. 17 Schéma illustrant 3 formes de sondes écho cardiographiques ainsi que la forme des images générées par chacune d'elles. Figure adaptée de [14].

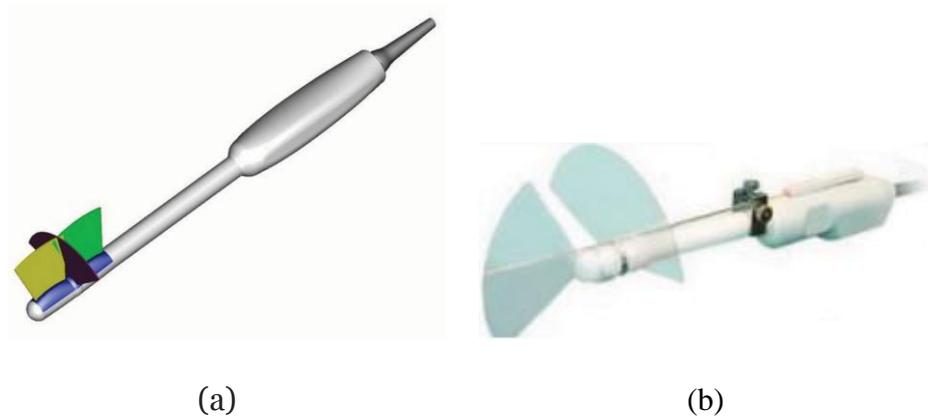


Figure I. 18 Schéma illustrant (a) une sonde échographique biplan commercialisée par la compagnie Vermon et (b) une sonde échographique biplan commercialisée par la compagnie Echoson.

Toutes deux sont conçues pour être insérées dans une cavité. (a) Sonde qui acquiert selon un plan latéral et horizontal, et (b) selon un plan horizontal et vertical. Figure tirée de [14] pour (a) et [15] pour (b).

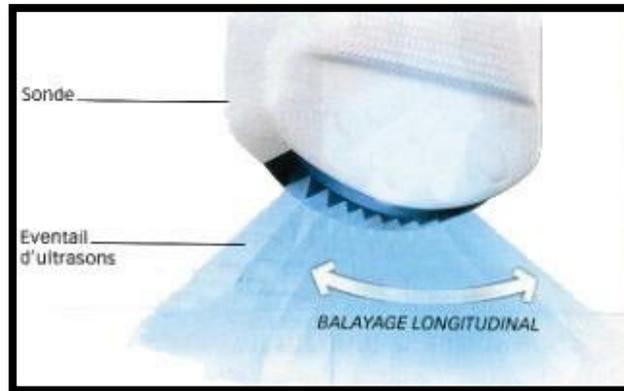


Figure I. 19 Schéma illustrant une sonde 3D effectuant un balayage longitudinal.

L'éventail d'ultrasons est constitué d'une collection d'images 2D utilisée pour reconstruire une image 3D à l'aide d'interpolations. Image adaptée du site [15].

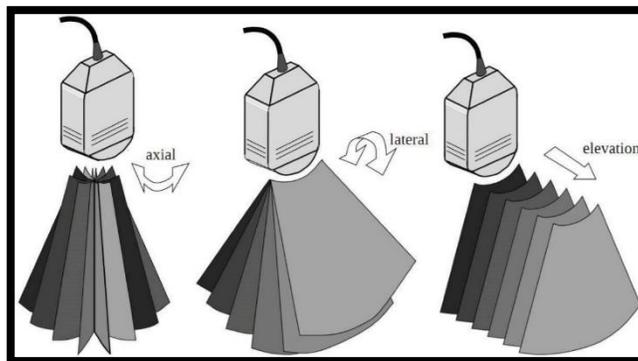


Figure I. 20 Schéma illustrant trois autres types de balayages (axial, latéral et élévation) du transducteur 2D des sondes échographiques permettant la construction d'un volume 3D. Image tirée de [16].

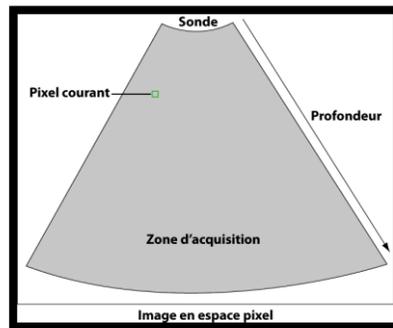


Figure I. 21 Schéma illustrant la numérisation d'une acquisition échographique.

Considérant une sonde produisant une image sectorielle, un pixel donné aura une intensité proportionnelle avec l'intensité de l'écho émis à l'emplacement physique correspondant. Plus le délai entre l'émission et la réception de l'écho est long, plus le pixel donné sera profond relativement à l'emplacement de la sonde. L'intensité des pixels à l'extérieur de la zone d'acquisition est zéro.

2.1.3 Génération de l'image

La sonde échographique est un transducteur, c'est-à-dire un dispositif convertissant un signal physique en un autre. Dans le cas de la sonde échographique, des vibrations sonores (échocs captés) sont transformées en signal électrique. Suite à l'émission, les ultrasons sont réfléchis partout où il y a un changement d'impédance acoustique c.à-d. à chaque fois que l'onde traverse un tissu différent. Les ondes sonores réfléchies (ou « échocs ») sont captées par la sonde. Les vibrations des échocs sont transformées en impulsions électriques et cette information est utilisée afin de générer une image numérique (à noter que chaque pixel de l'image résultante correspond à un écho émis à un emplacement donné). Pour ce faire, le scanneur considère deux paramètres avec chaque écho capté : son intensité ainsi que le délai entre l'émission de l'ultrason et sa réception. Le délai permet de déduire une information de profondeur à partir du lieu d'émission de l'ultrason (la sonde). Autrement dit, il permet de trouver le pixel dans l'image correspondant à l'origine de l'écho (l'information spatiale). La figure 2.12 illustre cette notion de profondeur. La force de l'écho est utilisée afin de déterminer l'intensité du pixel. Un écho fort résultant d'une réflexion R élevée aura une intensité forte (pixel clair) tandis qu'un écho faible résultant d'une réflexion R basse) aura une intensité faible (pixel foncé).

2.1.4 Rapport signal sur bruit

Le rapport signal sur bruit (SNR en anglais pour Signal-to-noise ratio) est une métrique permettant de décrire la qualité d'un signal désiré par rapport à son bruit de fond. Le SNR est défini par la relation suivante entre la puissance moyenne P du signal et celle du bruit :

$$\text{SNR} = P \text{ signal} / P \text{ noise} \quad (2.4)$$

La définition de la puissance d'un signal et de son bruit dépend du contexte et de sa nature. Par exemple, la définition de la puissance d'un signal sonore (dans un contexte d'un scanner échographique) n'est pas la même définition que la puissance d'un signal de réponse magnétique avec un scanner par résonance magnétique. Cependant, peu importe son contexte d'utilisation, un SNR élevé indique toujours une bonne qualité d'image tandis qu'un SNR moindre indique toujours une qualité inférieure. Un SNR de 3/1 est généralement un requis minimal en échographie pour que l'image soit considérée de bonne qualité [15]. En perspective, un SNR de 3/1 implique que le quart du signal total capté (signal + bruit) est du bruit, ce qui est tout de même non négligeable.

2.2 Échocardiographie

L'échocardiographie (ou échographie cardiaque) est une échographie du cœur et l'un des examens diagnostiques les plus utilisés en clinique. La simplicité d'utilisation et le faible coût de ce type d'imagerie sont des attraits particuliers de cette technologie. De plus, l'examen est relativement rapide et dure entre dix minutes à une demi-heure.

2.2.1 Caractéristiques

Le cœur est un organe intrathoracique. Ainsi, il est entouré d'os (les côtes) et d'air (les poumons). À cause de leur impédance acoustique fort différente de celui du tissu du coeur, ces milieux agissent comme des boucliers en bloquant et déviant les ultrasons normalement destinés à être transmis au coeur. Afin de pallier à ce problème, il faut se servir de « fenêtres » anatomiques, c'est-à-dire d'appliquer la

sonde à des endroits précis sur le torse du patient où il y aura peu d'interférences osseuses ou aériennes. Ces ouvertures sont en nombre limité et les trois principales sont illustrées à la figure 2.13. La figure 2.14 illustre côte à côte différentes coupes écho cardiographiques avec leur schéma anatomique respectif afin de bien saisir à quoi correspond l'information contenue dans les images générées.

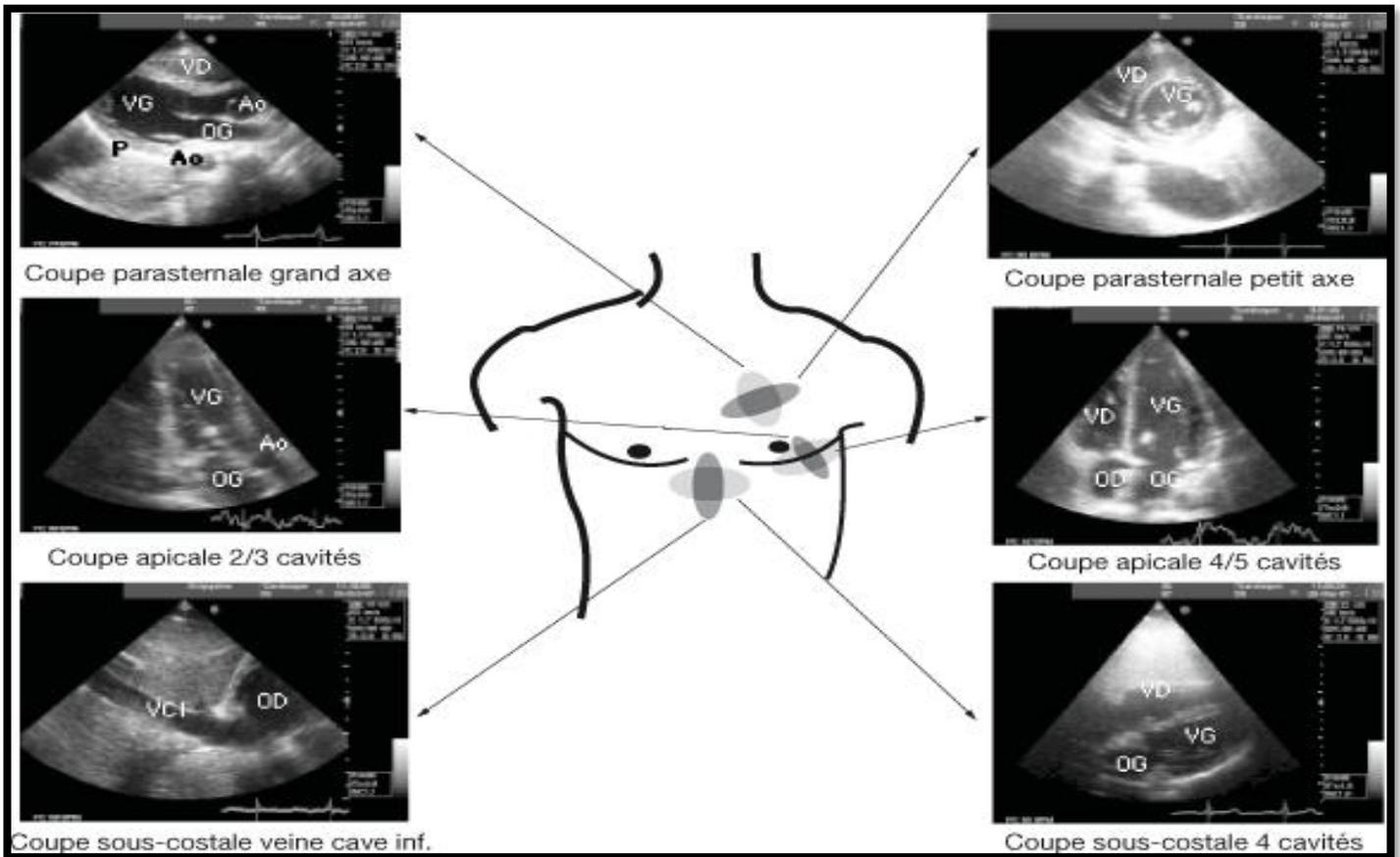


Figure 1. 22 Schéma illustrant trois des fenêtres anatomiques permettant l'acquisition de différentes coupes du cœur humain avec une sonde échographique, soit la coupe parasternale, apicale et sous-costale. Dans cet exemple, les images proviennent de sondes 2D. Im

Chez certains patients, il est possible que ces fenêtres anatomiques soient extrêmement réduites, voire inexistantes. Par exemple, un patient trop maigre où l'espace entre les côtes sont des creux (donc des trous d'air) complique grandement l'application de la sonde. À l'autre extrême, il est aussi difficile d'effectuer cet examen

avec les patients trop obèses en raison des épaisses couches de matière grasse sous-cutanées. Dans ces situations, une autre modalité d'imagerie est préférable.

Les images échographiques utilisées dans ce projet pour valider la méthode sont

Toutes acquises en 3D avec la fenêtre « coupe apicale » de façon à inclure le ventricule gauche dans son entièreté (dans la mesure du possible).

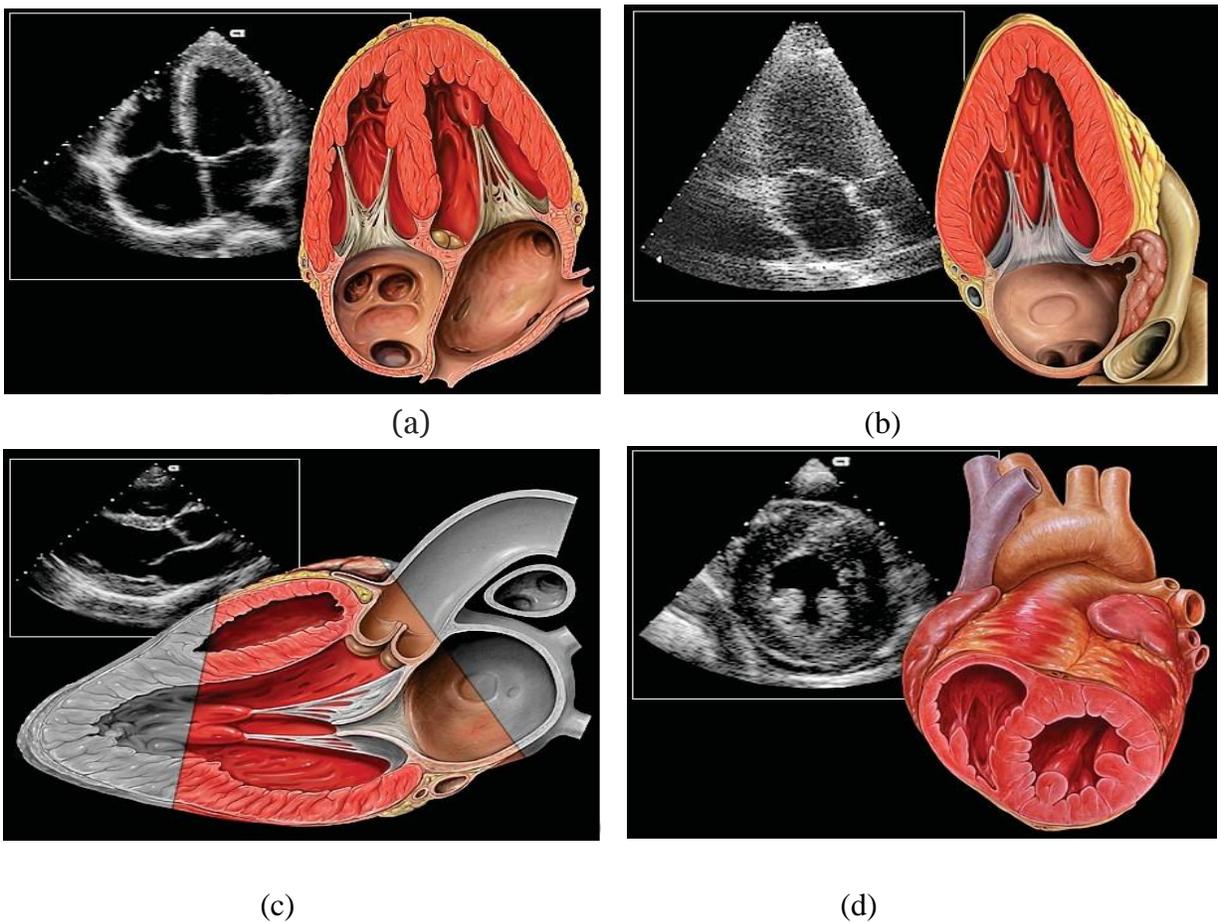


Figure I. 23 Figure illustrant des coupes échocardiographie avec un schéma de l'anatomie visible par celle-ci. (a) Coupe apicale 4/5 cavités (4 ou 5 cavités du coeur sont visibles) (b) Coupe apicale 2/3 cavités (2 ou 3 cavités du coeur sont visibles) (c) Coupe par Figure I. 24

2.2.2 Résumé

Les images écho cardiographiques peuvent contenir différents types d'artefacts visuels nuisant à la qualité ou pouvant même complètement obstruer de l'information importante. Le premier est un artefact d'ombrage et est considéré comme le talon d'Achille de l'échocardiographie. Il se manifeste lorsque l'ultrason transmis rencontre un milieu avec une impédance acoustique significativement différente de celui du tissu du cœur (comme de l'os ou de l'air). Ainsi, presque la totalité du signal sonore est réfléchi vers le transducteur et trop peu est transmis aux tissus en arrière. Visuellement, cet artefact se traduit par des zones sans information causées par des baisses de signal. La figure 2.15 illustre ce phénomène.

Un autre type d'artefact propre à l'échocardiographie est la réverbération. Il est causé par la réflexion continue du signal sonore entre deux interfaces très réfléchissantes. Ainsi, le signal «rebondi» d'un milieu à l'autre. Le résultat visuel de cet artefact ressemble à un effet de « god's ray » (rayons de lumière qui percent un nuage). La figure 2.16 illustre ce phénomène.

Il existe d'autres types d'artefacts chez les images échographiques, mais les deux décrits plus haut sont les principaux qui affectent les acquisitions cardiaques. Outre les artefacts d'origine physique, la qualité de l'image peut également être affectée par l'opérateur. En effet, l'aptitude de l'expert à manipuler la sonde sur le patient est primordiale afin de générer des images de bonne qualité. Il arrive parfois qu'une partie du ventricule gauche soit exclue de l'acquisition. Un tel exemple est démontré à la figure 2.17. Ceci se produit lorsque l'angle ou le positionnement de la sonde n'est pas adéquat. De plus, il est possible que les paramètres de l'échographe ne soient pas configurés optimalement pour la tâche à effectuer. Ainsi, comme nous observons à la figure 2.18, il peut y avoir dans l'acquisition des zones d'interférences (a), ou du bruit ambiant très persistant (b) rendant difficile le repérage du myocarde.

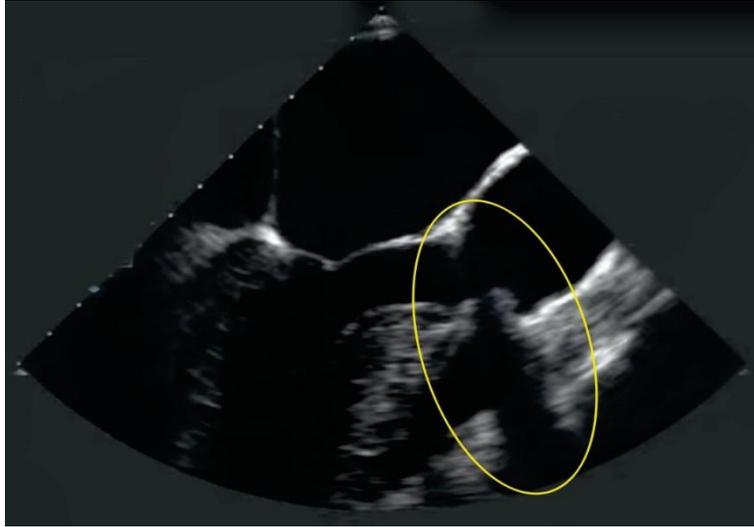


Figure I. 25 – Figure illustrant un artefact d'ombrage sur une image échocardiographie. Image adaptée du site [19].



Figure I. 26 Figure illustrant un artefact de réverbération sur une image échocardiographie. Image adaptée du site [19].

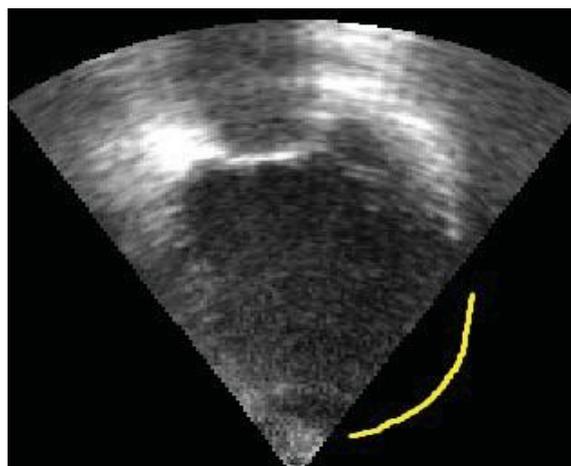
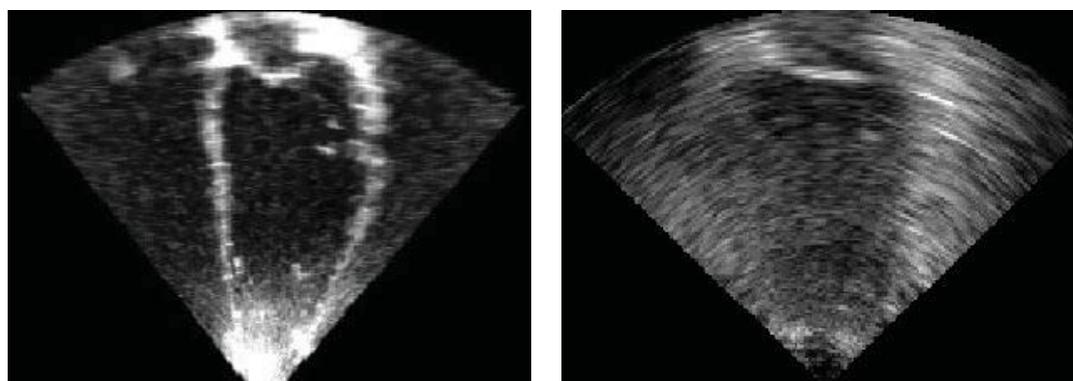


Figure I. 27 Figure illustrant une acquisition échocardiographie où une importante partie du ventricule gauche est manquante suite à une mauvaise manipulation de la sonde par l'opérateur. La ligne jaune indique la région en question.



(a)

(b)

Figure I. 28 Exemples d'images écho cardiographiques ayant (a) beaucoup d'interférence à l'apex (pointe du coeur) et (b) un bruit ambiant si élevé que l'on voit à peine les parois du ventricule gauche.

Chapitre 2

Les réseaux de neurones convolutionnels

3 Convolutional neural network

3.1 Introduction

L'apprentissage profond est un ensemble de méthodes d'apprentissage qui tentent de modéliser des données avec architectures complexes combinant différentes transformations non linéaires. Les briques élémentaires de l'apprentissage en profondeur sont les réseaux de neurones, qui sont combinés pour former les réseaux de neurones profonds.

Ces techniques ont permis d'importants progrès dans les domaines du son et le traitement d'images, y compris la reconnaissance faciale, la reconnaissance vocale, la vision par ordinateur, le traitement automatisé du langage, la classification de texte (par exemple reconnaissance des spams). Les applications potentielles sont très nombreuses. Un spectaculaire exemple est le programme Alpha GO, qui a appris à jouer au jeu de go par la méthode d'apprentissage profond, et a battu le champion du monde en 2016.

Notre objective dans ce chapitre est de comprendre c'est quoi les réseaux de neurones, comment il fonctionne et quelle sont ces architectures

Parmi les types des architectures les réseaux de neurones :

- Les Convolution Neural Networks (CNN), particulièrement adaptés au traitement d'images

3.2 Définition de CNN

Les réseaux de neurones network [16] (CNN) est un type spécifique de réseaux de neurones artificiels (ANN), qui ont démontré des performances élevées sur diverses tâches visuelles, y compris la classification d'images, segmentation [17], récupération d'images [18], détection d'objets, sous-titrage d'images, reconnaissance de visage [19], estimation de pose [20], reconnaissance de panneaux de signalisation, traitement de la parole, transfert de style neuronal [21]

Les CNN sont conçus pour traiter les données qui entrent dans la forme de plusieurs tableaux, par exemple, une image couleur composée de trois tableaux 2D contenant des intensités de pixels dans les canaux tricolores. Ils utilisent leur convolution des filtres pour extraire des informations des images, les calques antérieurs détectent les contours, les calques ultérieurs peuvent détecter une partie des objets, alors même les couches ultérieures peuvent détecter des objets complets, tels que comme des visages, ou d'autres formes géométriques complexes [22].

Le CNN composé d'un ensemble de couches pouvant être regroupées par leurs fonctionnalités, trois principaux types de couches sont : la couche convolutive, la couche de mise en commun et la couche entièrement connectée

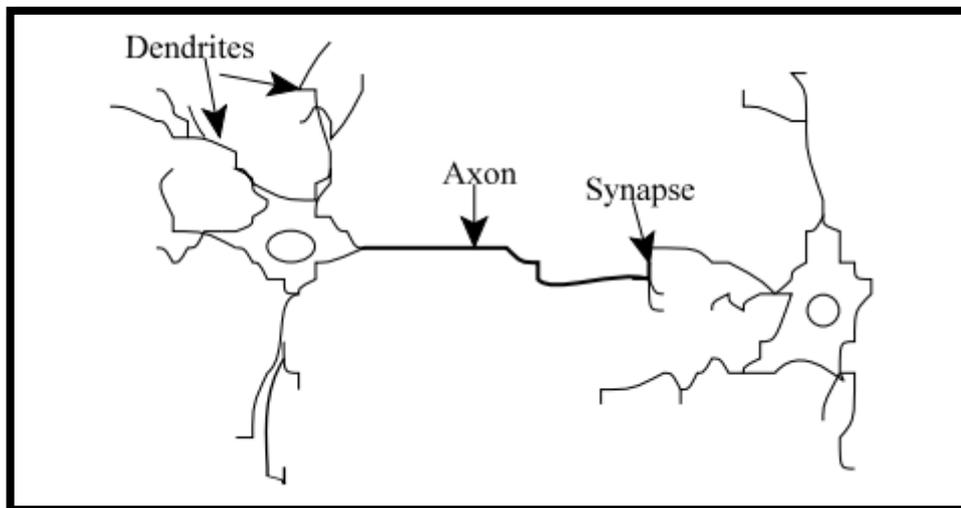


Figure II. 1 Natural neurons

La complexité des vrais neurones est très abstraite lors de la modélisation artificielle neurones. Ceux-ci se composent essentiellement d'entrées (comme les synapses), qui sont multipliées par des poids (force des signaux respectifs), puis calculé par une fonction mathématique qui détermine l'activation du neurone. Une autre fonction (qui peut être l'identité) calcule la sortie du neurone artificiel (parfois en fonction d'un certain au seuil). Les ANN combinent des neurones artificiels afin de traiter l'information

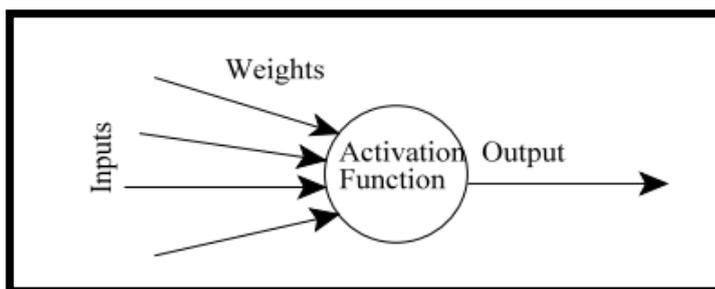


Figure II. 2 artificiel neurones

3.3 Description des différents composants ou blocs de construction de base de CNN

3.3.1 Couche réseau

Comme nous l'avons mentionné précédemment, qu'un CNN est composé de plusieurs blocs de construction (appelés couches de l'architecture), dans cette sous-section, nous avons décrit en détail certains de ces éléments constitutifs avec leur rôle dans L'architecture CNN.

3.3.2 Couche convolutive

Couche convolutive est le composant le plus important de toute architecture CNN. Il contient un ensemble de noyaux convolutifs (également appelés filtres), qui sont convolués avec l'image d'entrée (métriques à N dimensions) pour générer une carte des caractéristiques

3.3.3 Couche de mise en commun

La mutualisation les couches sont utilisées pour sous-échantillonner les cartes de caractéristiques (produites après des opérations de convolution), c'est-à-dire qu'il prend les cartes de caractéristiques de plus grande taille et le réduit à des cartes de caractéristiques de plus petite taille.

Tout en réduisant les cartes de caractéristiques, il conserve toujours les caractéristiques (ou informations) les plus dominantes dans chaque étape du pool. L'opération de regroupement est effectuée en spécifiant la région regroupée taille et la foulée de l'opération, similaire à l'opération de convolution. Il existe différents types des techniques de pooling sont utilisées dans différentes couches de pooling telles que max pooling, min pooling, pooling moyen, pooling fermé, pooling arborescent, etc. Le pooling max est le plus populaire et surtout technique de mise en commun utilisée

3.3.4 Fonctions d'activation (non linéaires)

La tâche principale de toute fonction d'activation dans tout modèle basé sur un réseau neuronal est de mapper l'entrée à la sortie, où la valeur d'entrée est obtenue en calculant la somme pondérée des neurones entrée et en ajoutant un biais avec lui (s'il y a un biais). En d'autres termes, la fonction d'activation décide si un neurone se déclenchera ou non pour une

entrée donnée en produisant la production.

Dans l'architecture CNN, après chaque couche apprenable (couches avec poids, c'est-à-dire convolutives et FC couches) des couches d'activation non linéaires sont utilisées. Le comportement de non-linéarité de ces couches permet le modèle CNN pour apprendre des choses plus complexes et réussir à mapper les entrées aux sorties de manière non linéaire. La caractéristique importante d'une fonction d'activation est qu'elle doit être différentiable en afin d'activer la rétropropagation des erreurs pour entraîner le modèle. L'activation la plus couramment utilisée les fonctions dans les réseaux de neurones profonds sont décrites.

3.3.5 Couche entièrement connectée

Habituellement, la dernière partie (ou couches) de chaque architecture CNN (utilisée pour la classification) est constituée de couches entièrement connectées, où chaque neurone à l'intérieur d'une couche est connecté à chaque neurone de sa couche précédente. La dernière couche des couches entièrement connectées est utilisée comme couche de sortie (classificateur) de l'architecture CNN.

Les couches entièrement connectées sont un type de réseau de neurones artificiels (ANN) à anticipation et il suit le principe du réseau neuronal perceptron multicouche traditionnel (MLP). Les couches FC prendre l'entrée de la couche convolutive ou de mise en commun finale, qui se présente sous la forme d'un ensemble de métriques

(Cartes de caractéristiques) et ces métriques sont aplaties pour créer un vecteur et ce vecteur est ensuite alimenté dans la couche FC pour générer la sortie finale de CNN comme le montre la Fig.1

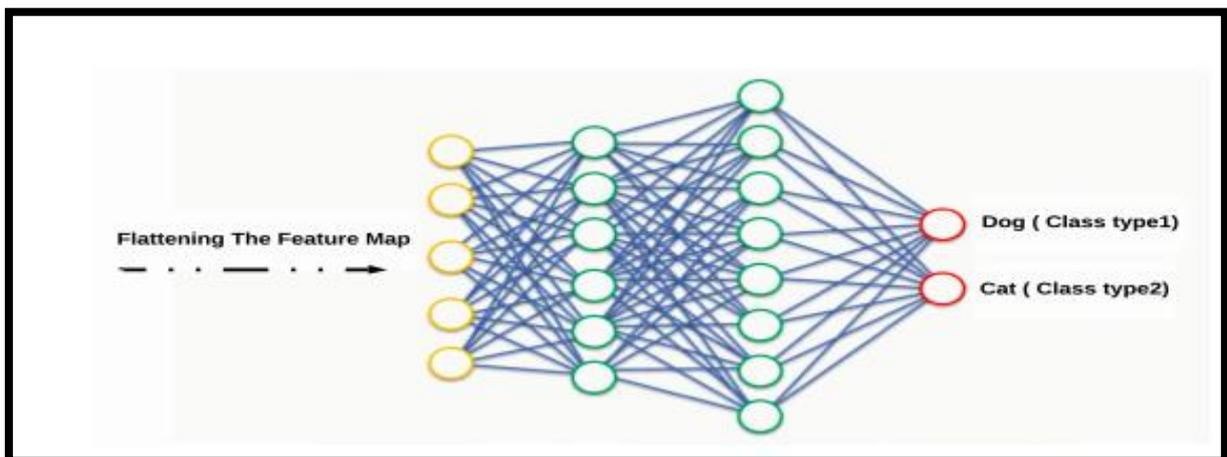


Figure II. 3 L'architecture des couches entièrement connectées[23]

3.4 Fonctions de perte

Nous savons que la dernière couche de chaque architecture CNN (basée sur la classification) est la couche de sortie, où le classement final a lieu. Dans cette couche de sortie, nous calculons l'erreur de prédiction générée par le modèle CNN sur les échantillons d'apprentissage en utilisant une fonction de perte. Cette erreur de prédiction indique au réseau à quel point sa prédiction est éloignée de la sortie réelle, puis cette erreur sera optimisée pendant le processus d'apprentissage du modèle CNN.

La fonction de perte utilise deux paramètres pour calculer l'erreur, le premier paramètre est l'estimation sortie du modèle CNN (également appelée prédiction) et la seconde est la sortie réelle (également connu sous le nom d'étiquette). Il existe différents types de fonctions de perte utilisées dans différents types de problème

3.5 Progrès récents dans les architectures CNN

Jusqu'à présent, il a couvert les concepts de base de CNN ainsi que différents composants de base ou blocs de construction de CNN, nous discutons du processus d'apprentissage de CNN avec plusieurs algorithmes d'apprentissage avec des lignes directrices afin d'améliorer l'efficacité (y compris prétraitement, initialisation des paramètres et régularisation vers CNN). Dans cette section, nous essayons d'expliquer d'architecture CNN réussie qui montre la classification des images dans le domaine de la vision par ordinateur. La vision par ordinateur a trois sous-domaines principaux où plusieurs architectures CNN (modèles) contribuent un rôle essentiel pour obtenir d'excellents résultats.

3.6 Classification des images

Dans la classification d'images, nous supposons que l'image d'entrée contient un seul objet, puis nous doivent classer l'image dans l'une des classes cibles présélectionnées en utilisant des modèles CNN. Certaines des principales architectures (modèles) de CNN conçues pour la classification d'images sont brièvement décrites comme suit :

3.6.1 LeNet-5

LeNet-5[12] est l'une des premières architectures de CNN, conçue pour classer les chiffres manuscrits. Il a été introduit par LeCun et al. En 1998. Le LeNet-5 à 5 couches (entraînables), c'est-à-dire trois couches convolutives et deux couches FC. Parmi eux, chacun des deux premières couches de convolution sont suivies d'une couche de mise en commun maximale (pour sous-échantillonner les cartes de caractéristiques) et ensuite, la dernière couche de convolution est suivie de deux couches entièrement connectées. La dernière couche de ces

3.6.3 ZFNet

FNet [14] a été présenté par Zeiler et Fergus dans ECCV-2014, il a une architecture presque similaire comme AlexNet sauf qu'ici, ils ont utilisé un filtre 7×7 avec une foulée 2 dans la 1ère couche convolutive. Dans cas d'AlexNet, Krizhevky et al. Utilisez un filtre 11×11 avec foulée 4 dans la 1ère couche convolutive. Comme un résultat ZFNet devient plus efficace qu'AlexNet et devient le gagnant de l'ILSVRC-2013. L'architecture de ZFNet est illustrée à la Fig.4

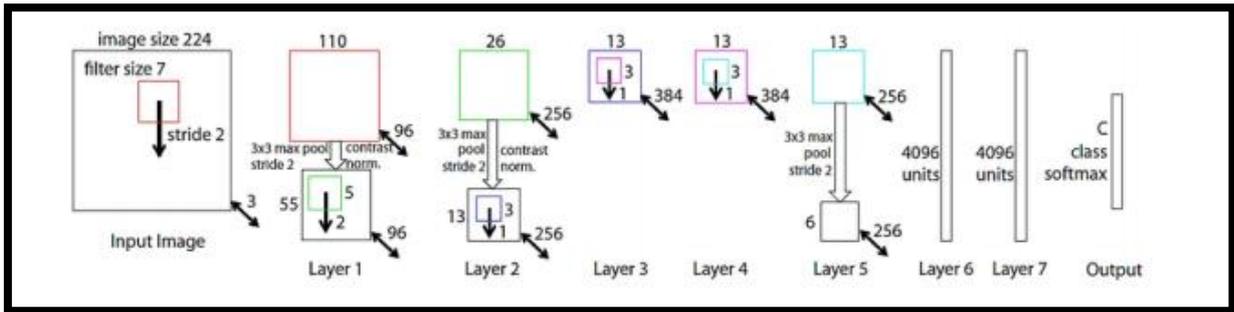


Figure II. 6 architecture de ZFNet [26]

3.6.4 VGGNet

VGGNet [15] est l'une des architectures CNN les plus populaires, introduite par Simonyan et Zisserman en 2014. Les auteurs ont introduit au total 6 configurations CNN différentes, parmi lesquelles le VGGNet-16 (configuration D) et le VGGNet-19 (configuration E) sont les plus réussis ceux. L'architecture de VGGNet-16 est illustrée à la Fig.5

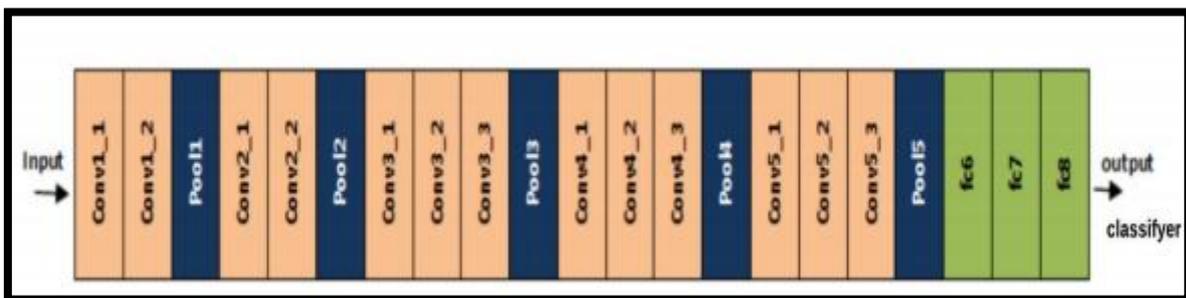


Figure II. 7 architecture de VGGNet[27]

3.6.5 GoogLeNet

L'architecture GoogleNet [16] est différente de toutes les CNN conventionnelles évoquées précédemment modèles, il utilise des branches de réseau au lieu d'utiliser une architecture séquentielle à une seule ligne. Le GoogleNet a été proposé par Szegedy et al.

En 2014. Le GoogleNet a 22 couches pondérées (apprenables),

il a utilisé le « module de lancement » comme élément de base du réseau. Le traitement de ce module s'effectue en parallèle dans le réseau, et chaque module (un simple module de base) se compose de 1×1 , 3×3 et 5×5 couches de convolutions filtrées en parallèle, puis il combine leurs cartes de caractéristiques de sortie, ce qui peut entraîner une sortie de caractéristiques de très grande dimension. Résoudre ce problème, ils ont utilisé le module de création avec réduction de la dimensionnalité (comme indiqué sur la Fig.35-2) dans leur architecture de réseau au lieu de la version naïve (de base) du module de création (comme indiqué dans la Fig.6).

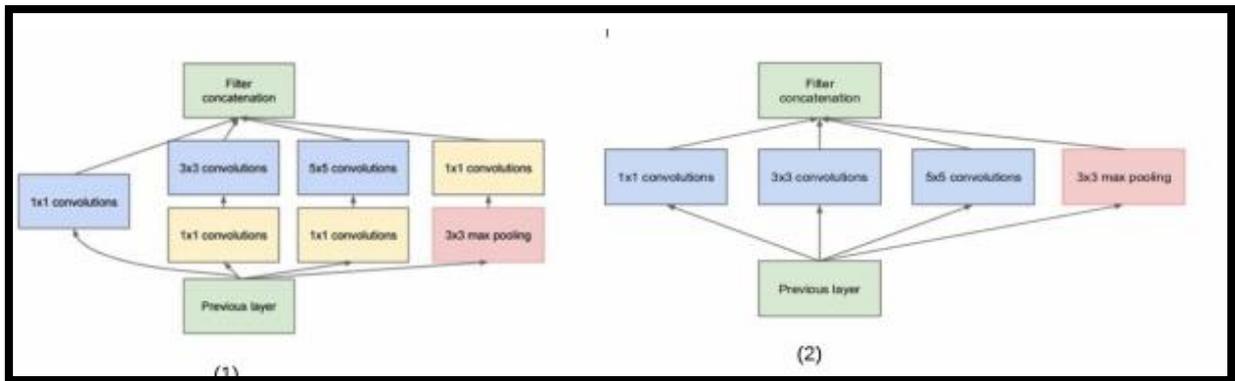


Figure II. 8 architecture de GOOGLE Net [28]

3.6.6 ResNet

Étant donné qu'un modèle CNN profond souffre de problèmes de gradient de disparition comme nous l'avons vu précédemment, He et Al. De Microsoft, a introduit l'idée de « connexion de saut d'identité » pour résoudre le gradient de disparition problème en proposant le modèle ResNet [50]. L'architecture du ResNet utilise une cartographie résiduelle ($H(x) = F(x) + x$) au lieu d'apprendre une application directe ($H(x) = F(x)$) et ces blocs sont appelés bocks résiduels. L'architecture ResNet complète est constituée de nombreux blocs résiduels avec

3×3 couches de convolution. La Fig.7 illustre la différence entre la cartographie directe et la cartographie résiduelle.

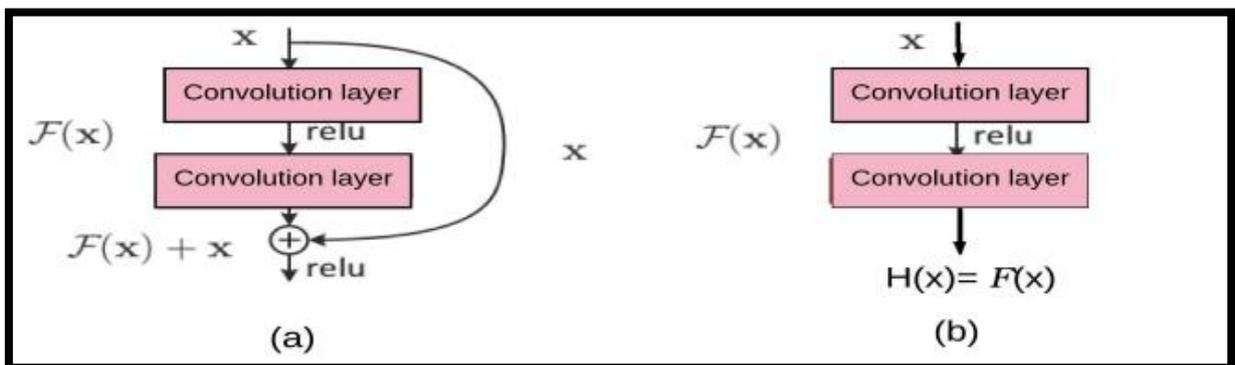


Figure II. 9 architecture de ResNet [29]

3.6.7 DenseNet

DenseNet [16] étend l'idée de mappage résiduel en propageant la sortie de chaque bloc à tous les blocs suivants à l'intérieur de chaque bloc dense du réseau, comme illustré à la Figure 35. Par propager l'information dans les deux sens vers l'avant et vers l'arrière pendant la formation du modèle renforce la capacité de propagation des caractéristiques et résout le problème du gradient de fuite. DenseNet a été introduit par Huang et al. En 2016 et il devient le lauréat de l'ILSVRC-2016. La figure 37 montre un modèle basé sur DenseNet.

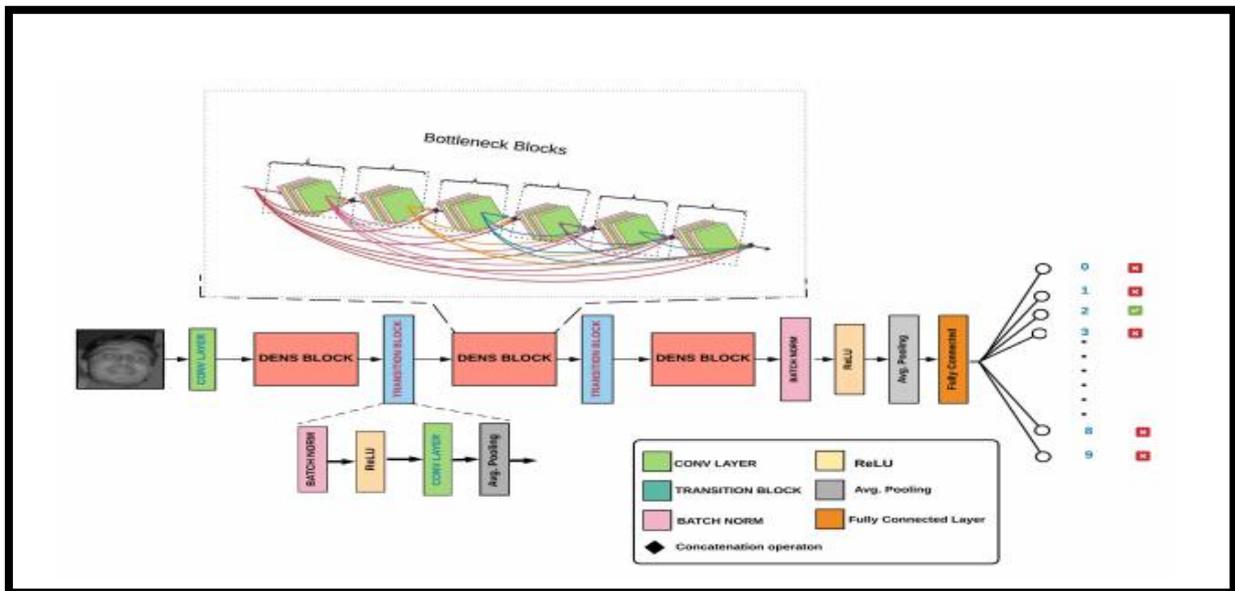


Figure II. 10 architecture de DANSENet [30]

3.7 Conclusion

Les réseaux de neurones convolutifs (CNN) sont devenus un algorithme de pointe pour les problèmes de vision par ordinateur, de traitement du langage naturel et de reconnaissance de formes. Ce CNN a été utilisé pour construire de nombreux modèles de cas d'utilisation, de la simple reconnaissance de chiffres à l'analyse d'images médicales complexes. Ce chapitre a essayé d'expliquer chaque composant d'un CNN, comment il fonctionne pour l'analyse d'images, et d'autres choses pertinentes. Ce chapitre donne également un aperçu de la fondation de CNN aux derniers modèles et mentionne quelques domaines d'application.

Chapitre 3

Implémentation

4 Introduction

En traitement d'image, La classification d'images est un problème d'apprentissage supervisé : dans lequel en devra définir un ensemble de classes cibles (objets à identifier dans des images), puis entraîner un modèle à les reconnaître à l'aide d'exemples de photos étiquetées.

Notre travail repose essentiellement sur, la classification d'image médicale acquise par une imageur écho cardiographique. Pour valider notre travail nous avons implémenté la méthode de réseaux de neurones convolutifs (en anglais Convolutional Neural Network, CNN). L'objectif dans ce chapitre est d'avoir classifié les images des coupes de cœur par CNN.

4.1 Logiciels et bibliothèques Utilisés dans l'implémentation

4.1.1 MATLAB

MATLAB (« matrix laboratory ») est un langage de script² émulé par un environnement de développement du même nom ; il est utilisé à des fins de calcul numérique. Développé par la société The MathWorks, MATLAB permet de manipuler des matrices, d'afficher des courbes et des données, de mettre en œuvre des algorithmes, de créer des interfaces utilisateurs, et peut s'interfacer avec d'autres langages comme le C, C++, Java, et Fortran. Les utilisateurs de MATLAB (environ 4 millions en 2019³) sont de milieux très différents comme l'ingénierie, les sciences et l'économie dans un contexte aussi bien industriel que pour la recherche. Matlab peut s'utiliser seul ou bien avec des toolboxes « boîte à outils ». [31]

4.1.2 ResNet-50

ResNet-50 est un modèle pré-entraîné qui a été entraîné sur un sous-ensemble de la base de données ImageNet et qui a remporté le concours ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) en 2015. Le modèle est entraîné sur plus d'un million d'images, a 177 couches au total, correspondant à un réseau résiduel de 50 couches, et peut classer les images en 1000 catégories d'objets (par exemple, clavier, souris, crayon et de nombreux animaux). L'ouverture du fichier `resnet50.mlpkginstall` depuis votre système d'exploitation ou depuis MATLAB lancera le processus d'installation pour la version que vous possédez.

4.1.3 Configuration Utilisé dans l'implémentation

La configuration du matériel utilisé dans notre implémentation est :

- Un PC portable Dell i5 CPU 2.40 GHZ
- Carte graphique INTEL HD
- RAM de taille 4 GO
- Disque dur de taille 500 GO
- Système d'exploitation Windows 10 pro 64

4.1.4 Les base d'images

Dans notre base de données on a utilisé 927 frame que on a extrait de trois séquence échocardiographie de 3 axe du cœur (coupe apicale, parasternale petite, parasternal grand)



Figure III. 1 base d'image utiliser pour l'apprentissage

4.2 Plan de notre travail

Notre travail expérimental est d'effectuer une classification d'images à l'aide d'un réseau de neurones convolutifs (CNN) dans MATLAB. Le réseau neuronal convolutif (CNN) est une puissante technique d'apprentissage automatique du domaine de l'apprentissage en profondeur. Une collection d'images divergentes est nécessaire pour former les CNN. Plus la collection est grande, plus les fonctionnalités apprises par CNN sont riches. Ces fonctionnalités surpassent souvent des fonctionnalités telles que HOG, LBP ou SURF. Former un CNN avec une grande collection d'images diverses n'est pas une tâche facile. Cependant, il existe un moyen facile. Nous pouvons utiliser CNN concerné pour tirer parti de la puissance de CNN. Cela permet d'économiser énormément de temps et d'efforts lorsque nous utilisons CNN pré-entraîné comme extracteur de caractéristiques. Dans cette leçon, on a utilisé « ResNet-50 » comme ensemble de données d'images CNN et des frame pré-entraînées. La classification d'images à l'aide d'un réseau de neurones convolutifs est un sujet très passionnant. Cependant, la stratégie utilisée que on a utilisée a rendu les choses simples à l'aide d'un réseau de neurones convolutifs dans MATLAB.

Nous avons opté pour la programmation sous Matlab pour les raisons suivantes :

- La syntaxe de Matlab, est simple et très parlante, elle est facile à assimiler.
- Matlab possède de nombreuses fonctions prédéfinies utiles en mathématiques que nous avons étudié au cours de nous travaux pratiques.
- Matlab facilite la manipulation des images médicale tout format.

Matlab propose des commandes pour créer certaines matrices particulières très simplement.

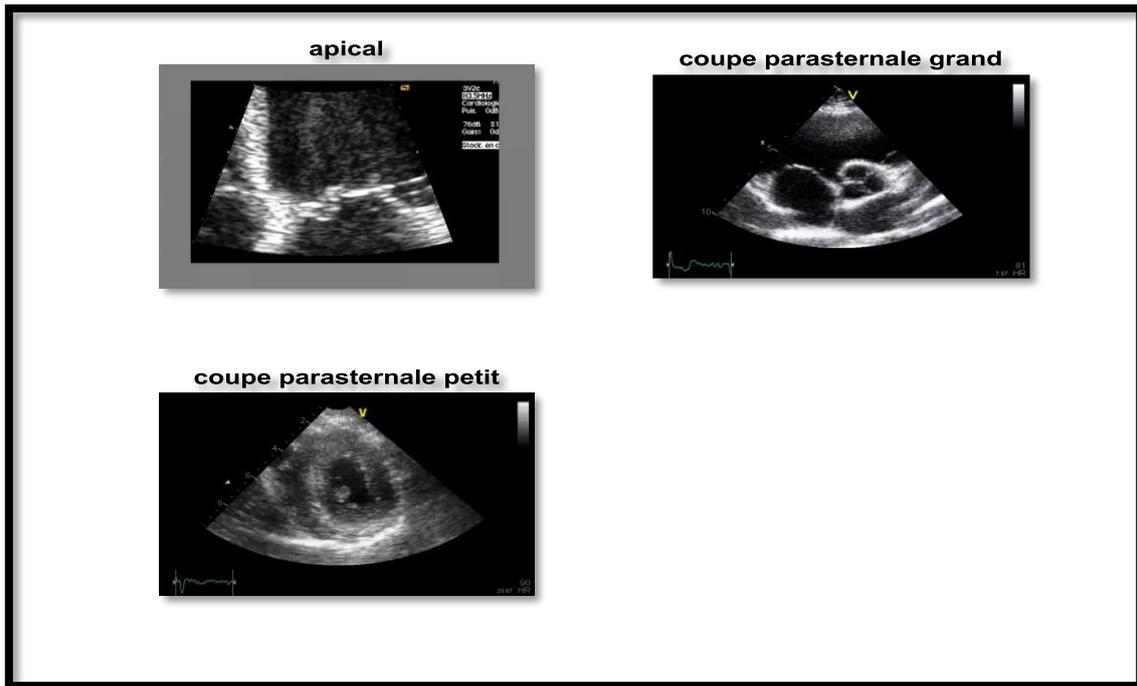


Figure III. 2 coupes échocardiographie apical / parasternal grand / parasternal petit

Dans notre base de données on a utilisé 927 frame que on a extrait de trois séquence échocardiographie de 3 axe du cœur (apicale, parasternale petite, parasternal grand)

Label	Count
Coupe_parasternale_petit_axe	258
apicale	433
coupe_parasternale_grand_axe	236

Figure III. 3 tableau illustre le nombre de frame dans chaque axe de notre base de donnée

4.1 Résultats obtenus et discussion

4.1.1 Résultats de la 1^{ère} expérience (avec un mini path de 32)

- On utilise 50 % dans la base de d'apprentissage et 50 % dans la base de test

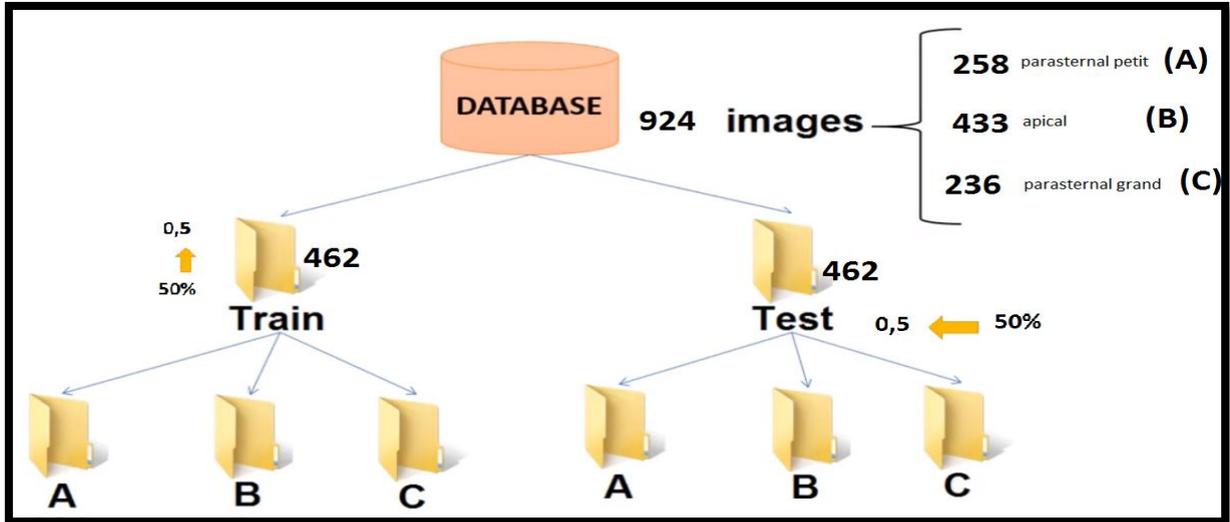


Figure III 4.1 distribution d'image dans la base de donnée (50%)

Confusion Matrix: resnet50

Output Class	Coupe _p arasternale _p etit _a xe	180 27.7%	0 0.0%	30 4.6%	85.7% 14.3%
	apicale	0 0.0%	303 46.7%	0 0.0%	100% 0.0%
	coupe _p arasternale _g rand _a xe	1 0.2%	0 0.0%	135 20.8%	99.3% 0.7%
		99.4% 0.6%	100% 0.0%	81.8% 18.2%	95.2% 4.8%
		Target Class			
		Coupe _p arasternale _p etit _a xe	apicale	coupe _p arasternale _g rand _a xe	

Figure III 4.2 confusion matrix de 50% de la base d'apprentissage

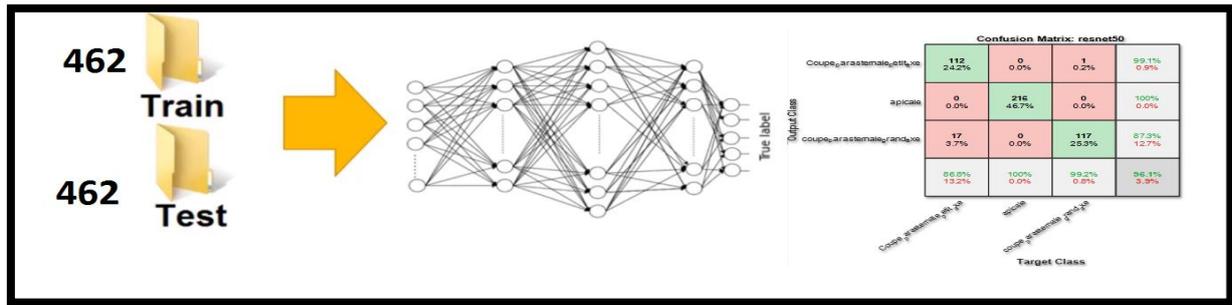


Figure III. 4 confusion matrix (0,5 mini path 32)

	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	99%	100%	81,8%
Taux de classification	95,2%		

Tableau III. 1résultat de sensibilité et taux de classification a 0,5 mini path 32

A . On utilise 50 % dans la basse de d'apprentissage et 50 % dans la basse de test

- | | | |
|----------------------------|-----------------------|------------------------|
| (tp)(A>true positive = 180 | (B>true positive =303 | (c>true positive = 135 |
| (tn)(A>true negative=438 | (B>true negative =346 | (c>true negative=483 |
| (fp)(A>false positive =1 | (B>false positive =0 | (c>false positive =30 |
| (fn)(A>false negative =30 | (B>false negative =0 | (c>false negative =1 |

Taux de classification = $(180+303+135)/(180+30+303+1+135)=0.962$

Taux de classification =96.2%

La sensitivity(A)=(Tp)/(TP+ FN)= $180/(180+30)=0.857$

La sensitivity(A)=85.7%

La specificity(A)=(TN)/(TN+ FP)= $438/(438+30)=0.994$

La specificity(A)=99.4%

La sensitivity(B)=(Tp)/(TP+ FN)= $303/(303+0)=1$

La sensitivity(B)=100%

La specificity(B)=(TN)/(TN+ FP)= $346/(346+0)=1$

La specificity(B)=100%

La sensitivity(C)=(Tp)/ (TP+ FN) = $135/(135+1)=0.99$

La sensitivity(C)=99.3%

La specificity(C)=(TN) / (TN+ FP) = $483/(483+30)=0.818$

La specificity(C)=81.8%

- On utilise 30 % dans la base de d'apprentissage et 70 % dans la base de test

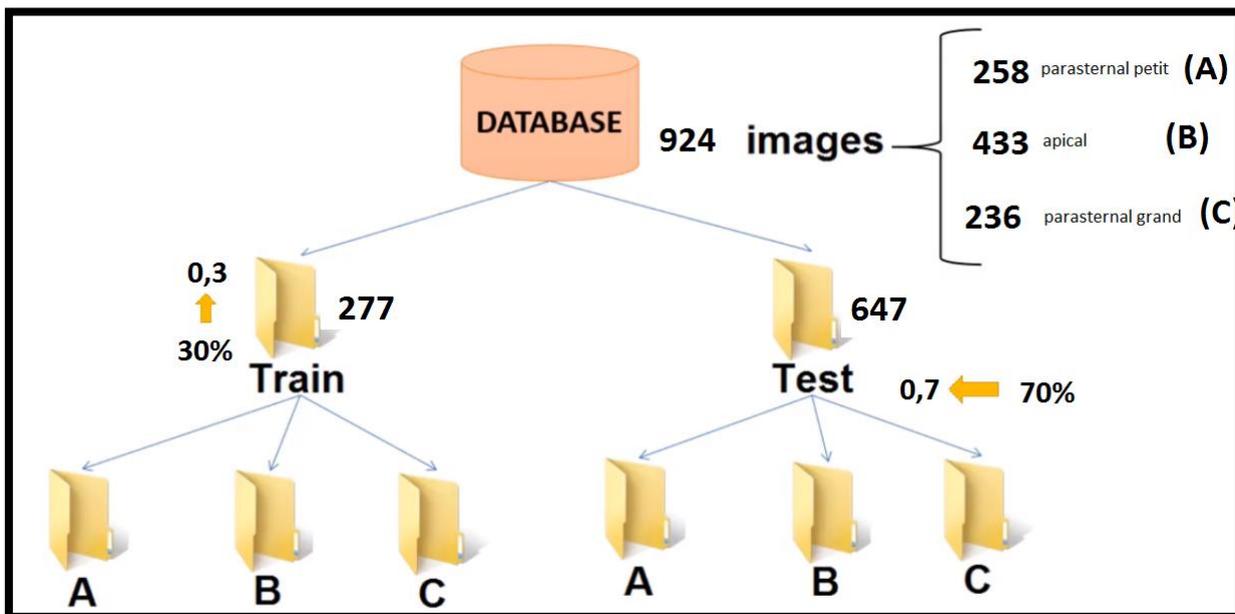


Figure III 5.1 distribution d'image dans la base de donnée (30%)

Confusion Matrix: resnet50

Output Class	Coupe_p_arasternale_p_etit_a_xe	apicale	coupe_p_arasternale_grand_a_xe	
Coupe_p_arasternale_p_etit_a_xe	112 24.2%	0 0.0%	1 0.2%	99.1% 0.9%
apicale	0 0.0%	216 46.7%	0 0.0%	100% 0.0%
coupe_p_arasternale_grand_a_xe	17 3.7%	0 0.0%	117 25.3%	87.3% 12.7%
	86.8% 13.2%	100% 0.0%	99.2% 0.8%	96.1% 3.9%
	Coupe_p_arasternale_p_etit_a_xe	apicale	coupe_p_arasternale_grand_a_xe	
	Target Class			

Figure III 5.2 confusion matrix de 30% de la base d'apprentissage

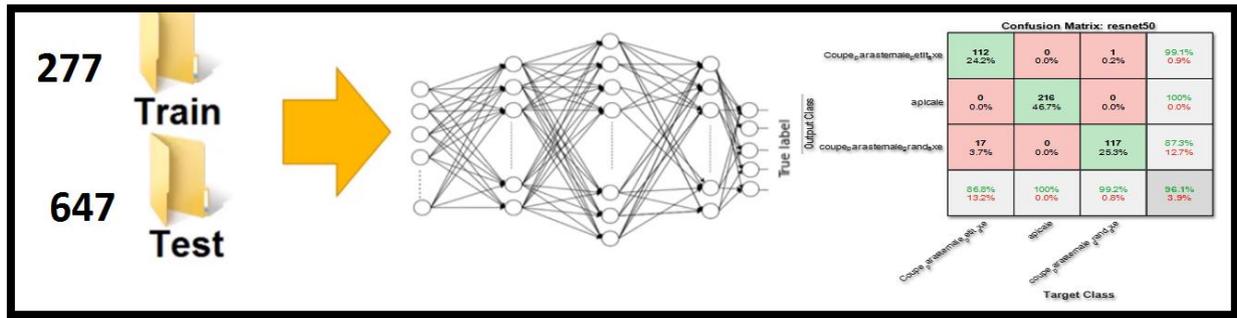


Figure III. 5 confusion matrix (0, 3 mini path 32)

	axe		
	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	86,8%	100%	99,2%
Taux de classification	96,1%		

Tableau III. 2 résultat de sensibilité et taux de classification a 0,3 mini path 32

B. On utilise 30 % dans la base de d'apprentissage et 70 % dans la base de test

(TP):(A) true positive = 112 (B)true positive =216 (c)true positive = 117
 (TN):(A) true negative=333 (B)true negative =247 (c)true negative=328
 (FP):(A) false positive =17 (B>false positive =0 (c>false positive =1
 (FN):(A) false negative =1 (B>false negative =0 (c>false negative =17

Taux de classification=(112+216+117)/(112+1+216+17+117)=0.952

Taux de classification =95.2%

La sensitivity(A)=(TP)/(TP+FN)=112/(112+1)=0.991

La sensitivity(A)=99.1%

La specificity(A)=(TN)/(TN+FP)=333/(333+17)=0.868

La specificity(A)=86.8%

La sensitivity(B)=(TP)/(TP+FN)=216/(216+0)=1

La sensitivity(B)=100%

La specificity(B)=(TN)/(TN+FP)=247/(247+0)=1

La specificity(B)=100%

La sensitivity(C)=(Tp)/(TP+FN)=117/(117+17)=0.873

La sensitivity(C)=87.3%

La specificity(C)=(TN)/(TN+FP)=328/(328+1)=0.992

La specificity(C)=99.2%

- On utilise 70 % dans la base de d'apprentissage et 30 % dans la base de test

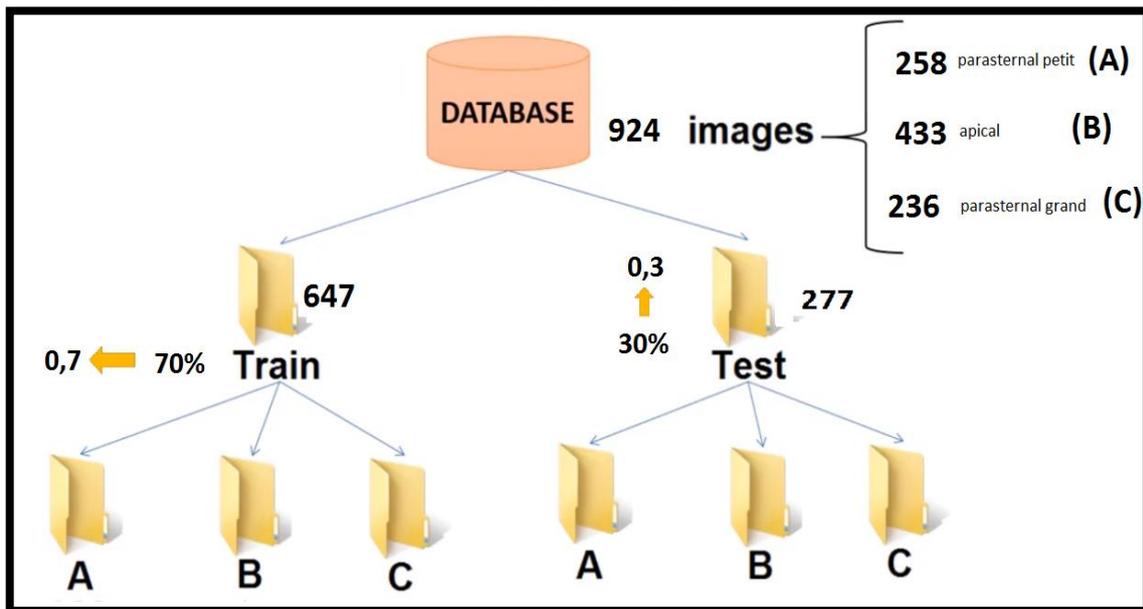


Figure III 6.1 distribution d'image dans la base de donnée (70%)

Confusion Matrix: resnet50

Output Class	Coupe_p_arasternale_p_etit_a_xe	73 26.3%	0 0.0%	5 1.8%	93.6% 6.4%
	apicale	0 0.0%	130 46.8%	0 0.0%	100% 0.0%
	coupe_p_arasternale_g_rand_a_xe	4 1.4%	0 0.0%	66 23.7%	94.3% 5.7%
		94.8% 5.2%	100% 0.0%	93.0% 7.0%	96.8% 3.2%
		Coupe_p_arasternale_p_etit_a_xe	apicale	coupe_p_arasternale_g_rand_a_xe	
		Target Class			

Figure III 6.2 confusion matrix de 70% de la base d'apprentissage

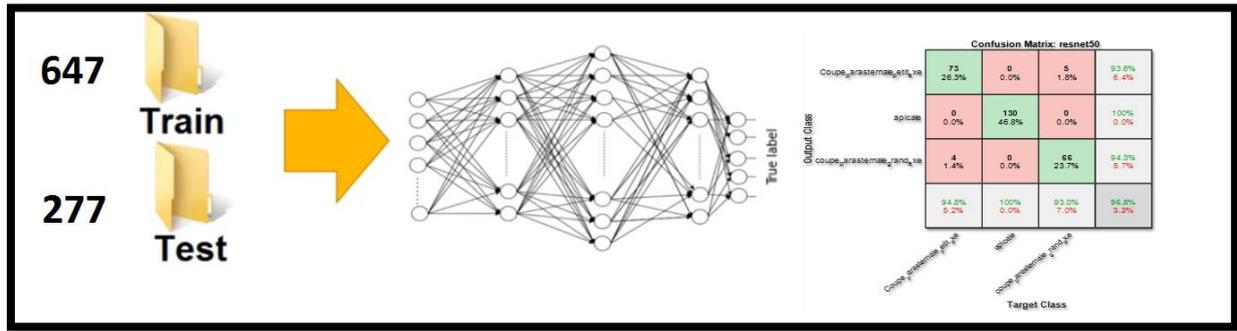


Figure III. 6 confusion matrix (0,7 mini path 32)

	axe		
	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	94,8%	100%	93%
Taux de classification	96,8%		

Tableau III. 3 résultat de sensibilité et taux de classification a 0,7 mini path 32

C . On utilise 70 % dans la base de d'apprentissage et 30 % dans la base de test

- (tp)(A>true positive = 73 (B>true positive =130 (c>true positive = 66
- (tn)(A>true negative=196 (B>true negative =148 (c>true negative=203
- (fp)(A>false positive =4 (B>false positive =0 (c>false positive =5
- (fn)(A>false negative =5 (B>false negative =0 (c>false negative =5

Taux de classification= $(73+130+66)/(73+5+130+4+66)=0.968$

Taux de classification =96.8%

La sensitivity(para-petite)=(TP)/(TP+FN)= $73/(73+5)=0.936$

La sensitivity(para-petite)=93.6%

La specificity(para-petite)=(TN)/(TN+FP)= $196/(196+4)=0.948$

La specificity(para-petite)=94.8%

La sensitivity(apical)=(TP)/(TP+FN)= $130/(130+0)=1$

La sensitivity(apical)=100%

La specificity(apical)=(TN)/(TN+FP)= $148/(148+0)=1$

La specificity(apical)=100%

La sensitivity(para-grand)=(TP)/(TP+FN)= $66/(66+5)=0.873$

La sensitivity(para-grand)=94.3%

La specificity(para-grand)=(TN)/(TN+FP)= $203/(203+5)=0.992$

La specificity(para-grand)=93%

Discussion

Après l'analyse des résultats obtenus, On constate les remarques suivantes :

D'après la Figures 39,40,41 La précision de l'apprentissage et de test augmente avec le nombre d'image ceci reflète qu'à chaque image le modèle apprend plus d'informations. Si la précision est diminuée alors on aura besoin de plus d'information pour faire apprendre notre modèle et par conséquent on doit augmenter le nombre d'époque et vice versa

D'après l'augmentation d'apprentissage de 30% vers 50% et la diminution dans la base de teste de 70% vers 50% en analyse que :

- * la sensitivity(para-petite) diminuer de 99.1% vers 85.7% et La specificity(para-petite) augmentée de 86.8% vers 99.4%

- * La sensitivity(apical) et la specificity(apical) reste toujours 100%

- * La sensitivity(para-grand) augmenter de 87.3% vers 93.3% et La specificity(para-grand) diminuer de 99.2% vers 81.8%

D'après la deuxième augmentation de la base d'apprentissage de 50% vers 70% et la diminution de la base de teste de 50% vers 30% en analyse que :

- * la sensitivity(para-petite) augmente de 87.7% vers 93.6% et La specificity(para-petite) diminuer de 99.4% vers 94.8%

- * La sensitivity(apical) et la specificity(apical) reste toujours 100%

- * La sensitivity(para-grand) augmenter de 93.3% vers 94.3% et La specificity(para-grand) augmente de 81.8 vers 93%

D'après les trois expériences en conclure que le taux de classification augmente lorsque on augmente la base d'apprentissage et diminuer la base de teste

4.1.2 Résultats de la 2 eme expérience (avec un mini path de 16)

- On utilise 30 % dans la base de d'apprentissage et 70 % dans la base de test

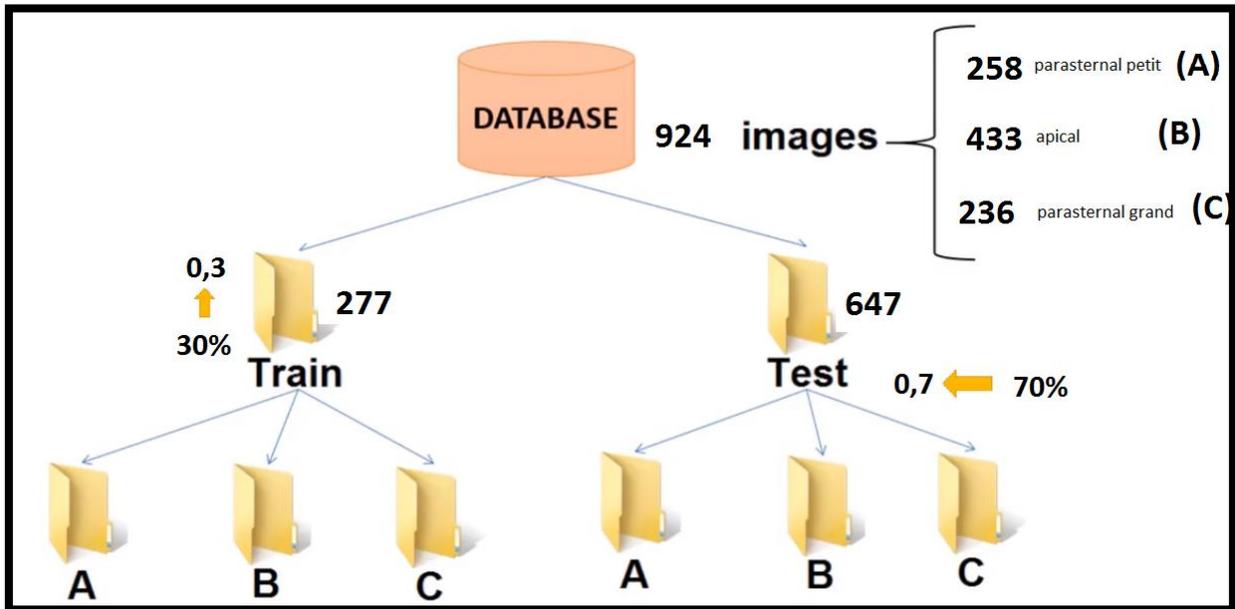


Figure III 7.1 distribution d'image dans la base de donnée (50%)

Confusion Matrix: resnet50

Output Class	Coupe_p_arasternale_p_etit_a_xe	169 26.0%	0 0.0%	19 2.9%	89.9% 10.1%
	apicale	0 0.0%	303 46.7%	0 0.0%	100% 0.0%
	coupe_p_arasternale_g_rand_a_xe	12 1.8%	0 0.0%	146 22.5%	92.4% 7.6%
		93.4% 6.6%	100% 0.0%	88.5% 11.5%	95.2% 4.8%
		Coupe_p_arasternale_p_etit_a_xe	apicale	coupe_p_arasternale_g_rand_a_xe	
		Target Class			

Figure III 7.2 confusion matrix de 50% de la base d'apprentissage

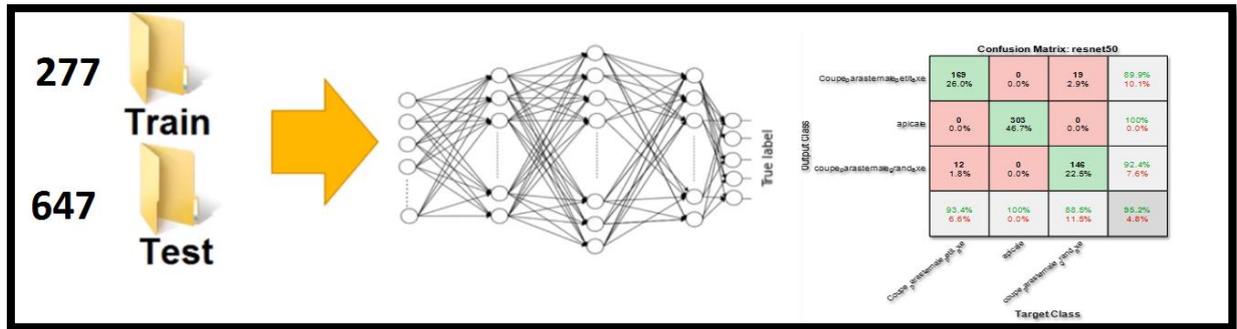


Figure III. 7 confusion matrix (0,3 mini path 16)

	Axe		
	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	93,4%	100%	88,5%
Taux de classification	95,2%		

Tableau III. 4 résultat de sensibilité et taux de classification a 0,3 mini path 16

On utilise 30 % dans la base de d'apprentissage et 70 % dans la base de test

(tp)(A)true positive = 169 (B)true positive =303 (c)true positive = 146
 (tn)(A)true negative=449 (B)true negative =346 (c)true negative=472
 (fp)(A>false positive =12 (B>false positive =0 (c>false negative =12
 (fn)(A>false negative =19 (B>false negative =0 (c>false positive =19

Taux de classification=(169+303+146)/(169+19+303++12+146)=0.952
 Taux de classification =95.2%

La sensitivity(A)=(Tp)/(TP+FN)=169/(169+19)=0.899
 La sensitivity(A)=89.9%

La specificity(A)=(TN)/(TN+FP)=449/(449+12)=0.934
 La specificity(A)=93.4%

La sensitivity(B)=(Tp)/(TP+FN)=303/(303+0)=1
 La sensitivity(B)=100%

La specificity(B)=(TN)/(TN+FP)=346/(346+0)=1
 La specificity(B)=100%

La sensitivity(C)=(Tp)/(TP+FN)=146/(146+12)=0.924
 La sensitivity(C)=92.4%

La specificity(C)=(TN)/(TN+FP)=472/(472+19)=0.885
 La specificity(C)=88.5%

- On utilise 50 % dans la base de d'apprentissage et 50 % dans la base de test

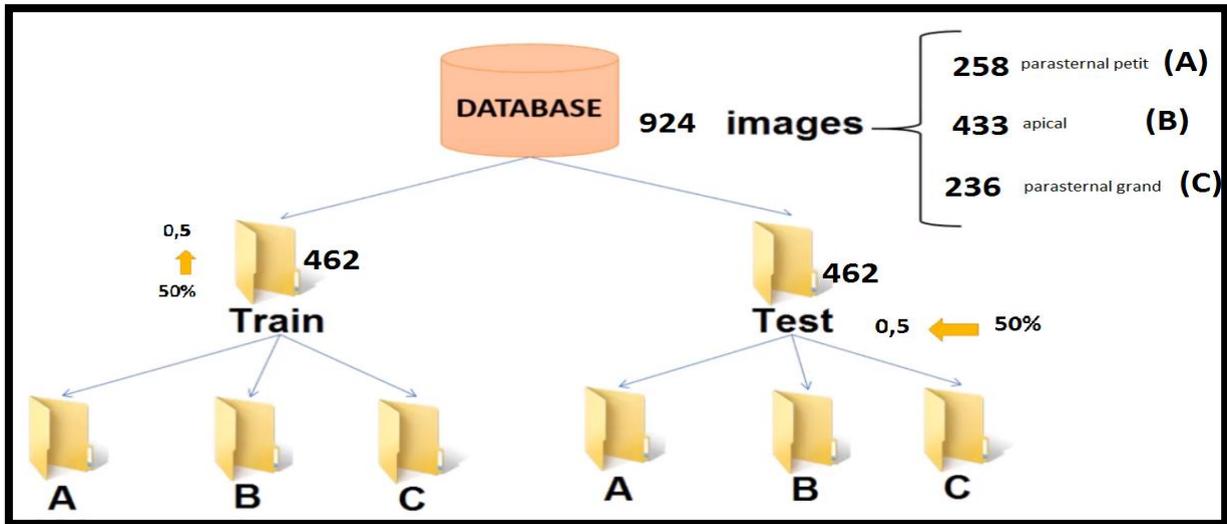


Figure III 8.1 distribution d'image dans la base de donnée (30%)

Confusion Matrix: resnet50

Output Class	Coupe_p_arasternale_p_etit_a_xe	126 27.2%	0 0.0%	25 5.4%	83.4% 16.6%
	apicale	0 0.0%	216 46.7%	0 0.0%	100% 0.0%
	coupe_p_arasternale_g_rand_a_xe	3 0.6%	0 0.0%	93 20.1%	96.9% 3.1%
		97.7% 2.3%	100% 0.0%	78.8% 21.2%	94.0% 6.0%
		Coupe_p_arasternale_p_etit_a_xe	apicale	coupe_p_arasternale_g_rand_a_xe	
		Target Class			

Figure III 8.2 confusion matrix de 30% de la base d'apprentissage

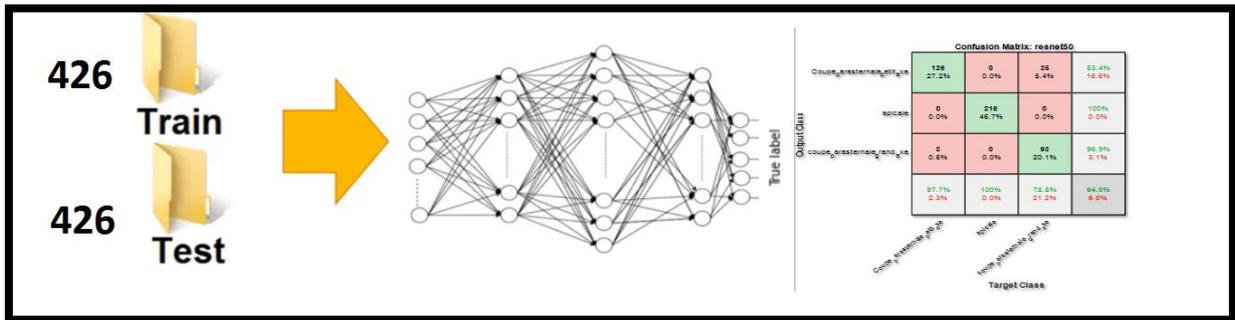


Figure III. 8 confusion matrix (0,5 mini path 16)

	axe		
	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	97,7%	100%	78,8%
Taux de classification	94%		

Tableau III. 5 résultat de sensibilité et taux de classification a 0,5 mini path 16

On utilise 50 % dans la base de d'apprentissage et 50 % dans la base de test

(tp)(A>true positive = 126 (B>true positive =216 (c>true positive = 93
 (tn)(A>true negative=309 (B>true negative =247 (c>true negative=342
 (fp)(A>false positive =3 (B>false positive =0 (c>false positive =3
 (fn)(A>false negative =25 (B>false negative =0 (c>false negative =25

$$\text{Taux de classification} = (126+216+93)/(126+25+216+93+3) = 0.94$$

Taux de classification = 94%

La sensitivity(A) = $(Tp)/(TP+FN) = 126/(126+25) = 0.834$
 La sensitivity(A) = 83.4%
 La specificity(A) = $(TN)/(TN+FP) = 309/(309+3) = 0.977$
 La specificity(A) = 97.7%

La sensitivity(B) = $(Tp)/(TP+FN) = 216/(216+0) = 1$
 La sensitivity(B) = 100%
 La specificity(B) = $(TN)/(TN+FP) = 247/(247+0) = 1$
 La specificity(B) = 100%

La sensitivity(C) = $(Tp)/(TP+FN) = 93/(93+25) = 0.969$
 La sensitivity(C) = 96.9%
 La specificity(C) = $(TN)/(TN+FP) = 342/(342+3) = 0.788$
 La specificity(C) = 78.8%

- On utilise 70 % dans la base de d'apprentissage et 30 % dans la base de test

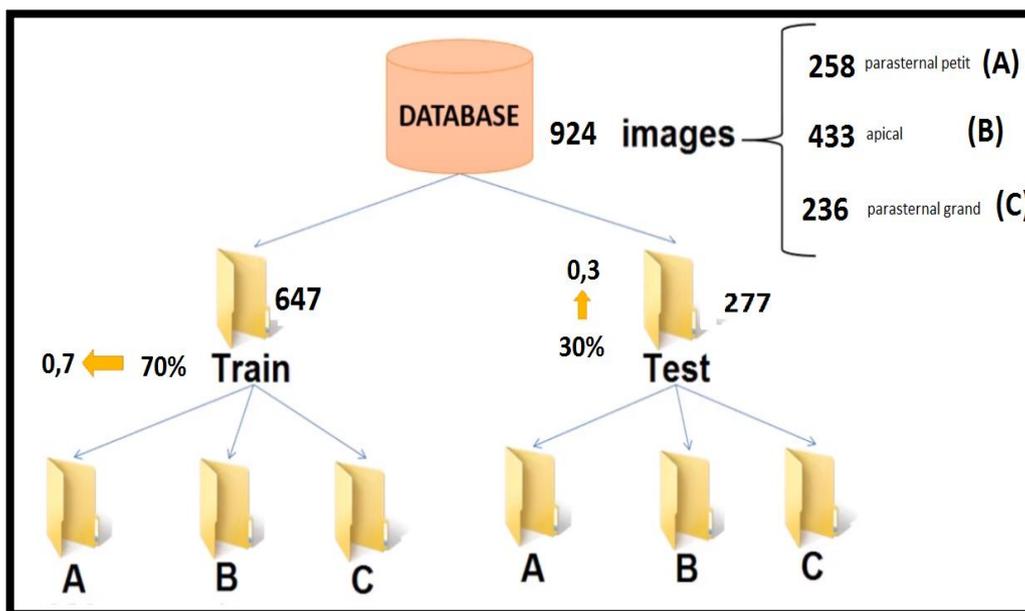


Figure III 9.1 distribution d'image dans la base de donnée (70%)

Confusion Matrix: resnet50

Output Class	Coupe _p arasternale _p etit _a xe	58 20.9%	0 0.0%	0 0.0%	100% 0.0%
	apicale	0 0.0%	130 46.8%	0 0.0%	100% 0.0%
	coupe _p arasternale _g rand _a xe	19 6.8%	0 0.0%	71 25.5%	78.9% 21.1%
		75.3% 24.7%	100% 0.0%	100% 0.0%	93.2% 6.8%
		Target Class			
		Coupe _p arasternale _p etit _a xe	apicale	coupe _p arasternale _g rand _a xe	

Figure III 9.2 confusion matrix de 70% de la base d'apprentissage

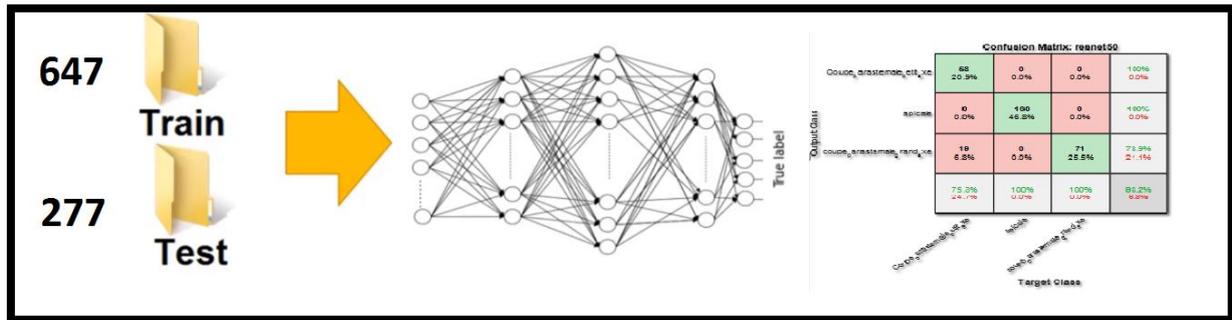


Figure III. 9 confusion matrix (0,7 mini path 16)

	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	75,3%	100%	100%
Taux de classification	93,2%		

Tableau III. 6 résultat de sensibilité et taux de classification a 0,7 mini path 16

On utilise 70 % dans la base de d'apprentissage et 30 % dans la base de test

- | | | |
|---------------------------|-----------------------|-----------------------|
| (tp)(A>true positive = 58 | (B>true positive =130 | (c>true positive = 71 |
| (tn)(A>true negative=201 | (B>true negative =148 | (c>true negative=188 |
| (fp)(A>false positive =19 | (B>false positive =0 | (c>false positive =0 |
| (fn)(A>false negative =0 | (B>false negative =0 | (c>false negative =19 |

Taux de classification=(58+130+71)/(58+130+17+19)=0.932
 Taux de classification =93.2%

La sensitivity(A)=(Tp)/(TP+FN)=58/(58+0)=1
 La sensitivity(A)=100%
 La specificity(A)=(TN)/(TN+FP)=201/(201+19)=0.753
 La specificity(A)=75.3%

La sensitivity(B)=(Tp)/(TP+FN)=130/(130+0)=1
 La sensitivity(B)=100%
 La specificity(B)=(TN)/(TN+FP)=148/(148+0)=1
 La specificity(B)=100%

La sensitivity(C)=(Tp)/(TP+FN)=71/(71+19)=0.969
 La sensitivity(C)=78.9%
 La specificity(C)=(TN)/(TN+FP)=188/(188+0)=1
 La specificity(C)=100%

Discussion

Après l'analyse des résultats obtenus, On constate les remarques suivantes :

D'après la Figures 42, 43,44 La précision de l'apprentissage et de test augmente avec le nombre d'époque, ceci reflète qu'à chaque époque le modèle apprend plus d'informations. Si la précision est diminuée alors on aura besoin de plus d'information pour faire apprendre notre modèle et par conséquent on doit augmenter le nombre d'époque et vice versa

D'après l'augmentation d'apprentissage de 30%vers 50% et la diminution dans la base de teste de70% vers 50% en analyse que :

- *la sensitivity(para-petite) diminuer de89.9% vers 83.4% et La specificity(para-petite) augmentée de 88.5% vers 97.7%

- * La sensitivity(apical) et la specificity(apical) reste toujours 100%

- * La sensitivity(para-grand) augmenter de 92.4% vers 96.6% et La specificity(para-grand) diminuer de88.5% vers 78.8%

D'après la deuxième augmentation de la base d'apprentissage de50% vers 70% et la diminution de la base de teste de 50% vers 30% en analyse que :

- *la sensitivity(para-petite) augmente de83.4% vers 100% et La specificity(para-petite) diminuer de 97.7% vers 75.3%

- * La sensitivity(apical) et la specificity(apical) reste toujours 100%

- * La sensitivity(para-grand) diminuer de 96.6% vers 78.9% et La specificity(para-grand) augmente de78.8 vers 100%

D'après les trois expériences en conclure que le taux de classification augmente lorsque on augmente la base d'apprentissage et diminuer la base de teste

4.1.3 Résultats de la 3 eme expérience (avec un mini path de 64)

- On utilise 30 % dans la base de d'apprentissage et 70 % dans la base de test

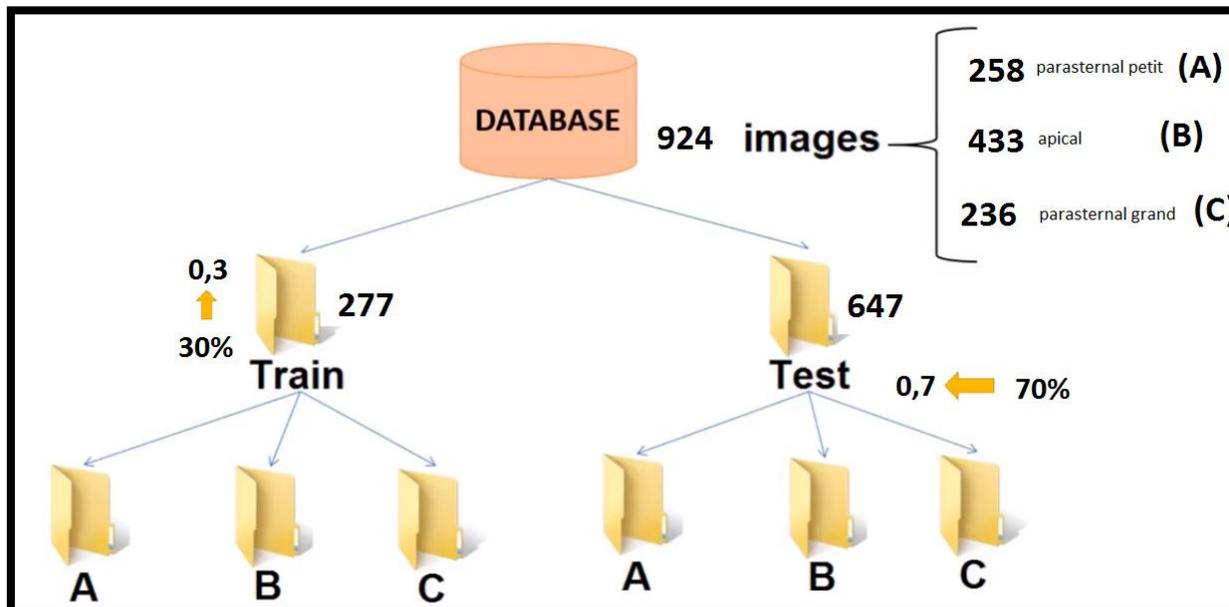


Figure III 10.1 distribution d'image dans la base de donnée (50%)

Confusion Matrix: resnet50

Output Class	Coupe_p_arasternale_petit_axe	apicale	coupe_p_arasternale_grand_axe	
Coupe_p_arasternale_petit_axe	173 26.7%	0 0.0%	27 4.2%	86.5% 13.5%
apicale	0 0.0%	303 46.7%	0 0.0%	100% 0.0%
coupe_p_arasternale_grand_axe	8 1.2%	0 0.0%	138 21.3%	94.5% 5.5%
	95.6% 4.4%	100% 0.0%	83.6% 16.4%	94.6% 5.4%
	Coupe_p_arasternale_petit_axe	apicale	coupe_p_arasternale_grand_axe	
	Target Class			

Figure III 10.2 confusion matrix de 50% de la base d'apprentissage

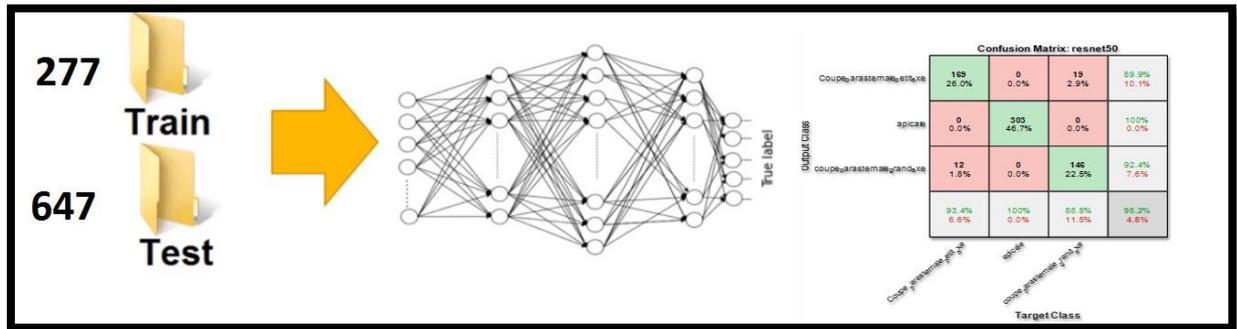


Figure III. 10 confusion matrix (0,3 mini path 64)

	axe		
	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	95,6%	100%	83,6%
Taux de classification	94,6%		

Tableau III. 7 résultat de sensibilité et taux de classification a 0,3 mini path 64

On utilise 30 % dans la base de d'apprentissage et 70 % dans la base de test

tp)(A>true positive = 173 (B>true positive =303 (c>true positive = 138
 (tn)(A>true negative=441 (B>true negative =346 (c>true negative=476
 (fp)(A>false positive =8 (B>false positive =0 (c>false positive =27
 (fn)(A>false negative =27 (B>false negative =0 (c>false negative =8

Taux de classification=(173+303+138)/(173+303+138+27+8)=0.946
 Taux de classification =94.6%

La sensitivity(A)=(Tp)/(TP+FN)=173/(173+27)=0.865
 La sensitivity(A)=86.5%
 La specificity(A)=(TN)/(TN+FP)=441/(441+8)=0.956
 La specificity(A)=95.6%

La sensitivity(B)=(Tp)/(TP+FN)=303/(303+0)=1
 La sensitivity(B)=100%
 La specificity(B)=(TN)/(TN+FP)=346/(346+0)=1
 La specificity(B)=100%

La sensitivity(C)=(Tp)/(TP+FN)=138/(138+8)=0.945
 La sensitivity(C)=94.5%
 La specificity(C)=(TN)/(TN+FP)=476/(476+27)=0.836
 La specificity(C)=83.6%

- On utilise 50 % dans la base de d'apprentissage et 50 % dans la base de test

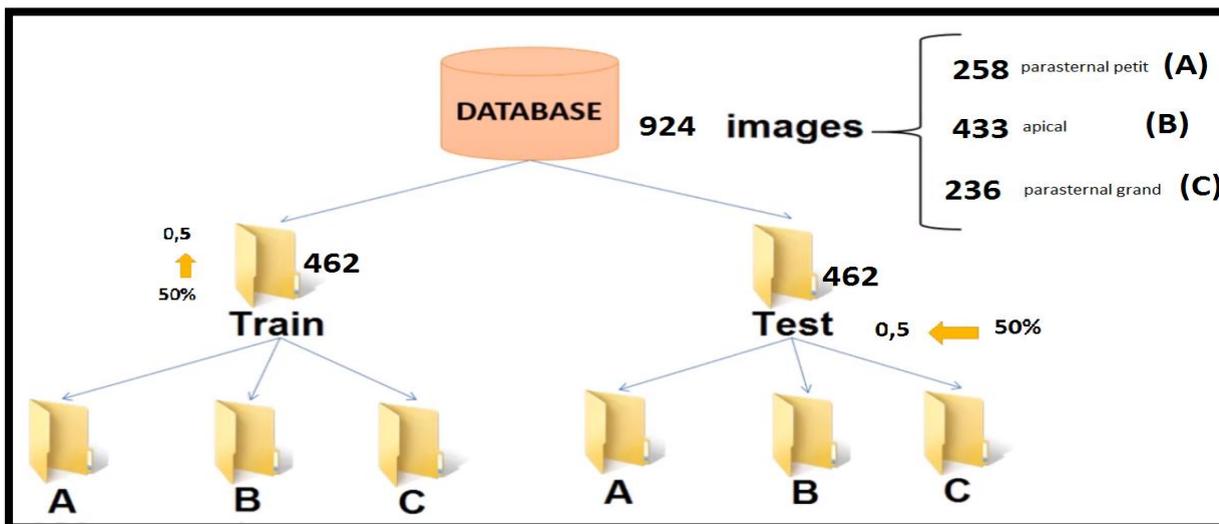


Figure III 11.1 distribution d'image dans la base de donnée (30%)

Confusion Matrix: resnet50

Output Class	Coupe _p arasternale _p etit _a xe	118 25.5%	0 0.0%	6 1.3%	95.2% 4.8%
	apicale	0 0.0%	216 46.7%	0 0.0%	100% 0.0%
	coupe _p arasternale _g rand _a xe	11 2.4%	0 0.0%	112 24.2%	91.1% 8.9%
		91.5% 8.5%	100% 0.0%	94.9% 5.1%	96.3% 3.7%
		Target Class			
		Coupe _p arasternale _p etit _a xe	apicale	coupe _p arasternale _g rand _a xe	

Figure III 11.2 confusion matrix de 30% de la base d'apprentissage

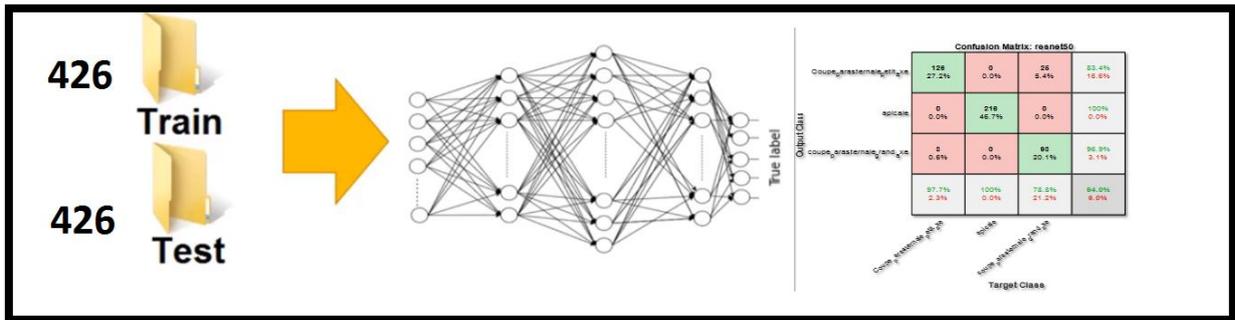


Figure III. 11 confusion matrix (0,5 mini path 64)

	axe		
	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	91,5%	100%	94,9%
Taux de classification	96,3%		

Tableau III. 8 résultat de sensibilité et taux de classification a 0,5 mini path 64

On utilise 50 % dans la base de d'apprentissage et 50 % dans la base de test

tp)(A>true positive = 118	(B>true positive =216	(c>false negative =11
(tn)(A>true negative=328	(B>true negative =246	(c>true positive = 112
(fp)(A>false positive =11	(B>false positive =0	(c>true negative=334
(fn)(A>false negative =6	(B>false negative =0	(c>false positive =6

Taux de classification=(118+216+112)/(118+216+112+6+11)=0.963
 Taux de classification =96.3%

La sensitivity(A)=(Tp)/(TP+FN)=118/(118+6)=0.952
 La sensitivity(A)=95.2%

La specificity(A)=(TN)/(TN+FP)=328/(328+11)=0.915
 La specificity(A)=91.5%

La sensitivity(B)=(Tp)/(TP+FN)=216/(216+0)=1
 La sensitivity(B)=100%

La specificity(B)=(TN)/(TN+FP)=246/(246+0)=1
 La specificity(B)=100%

La sensitivity(C)=(Tp)/(TP+FN)=112/(112+11)=0.911
 La sensitivity(C)=91.1%

La specificity(C)=(TN)/(TN+FP)=334/(334+6)=0.949
 La specificity(C)=94.9%

- On utilise 70 % dans la base de d'apprentissage et 30 % dans la base de test

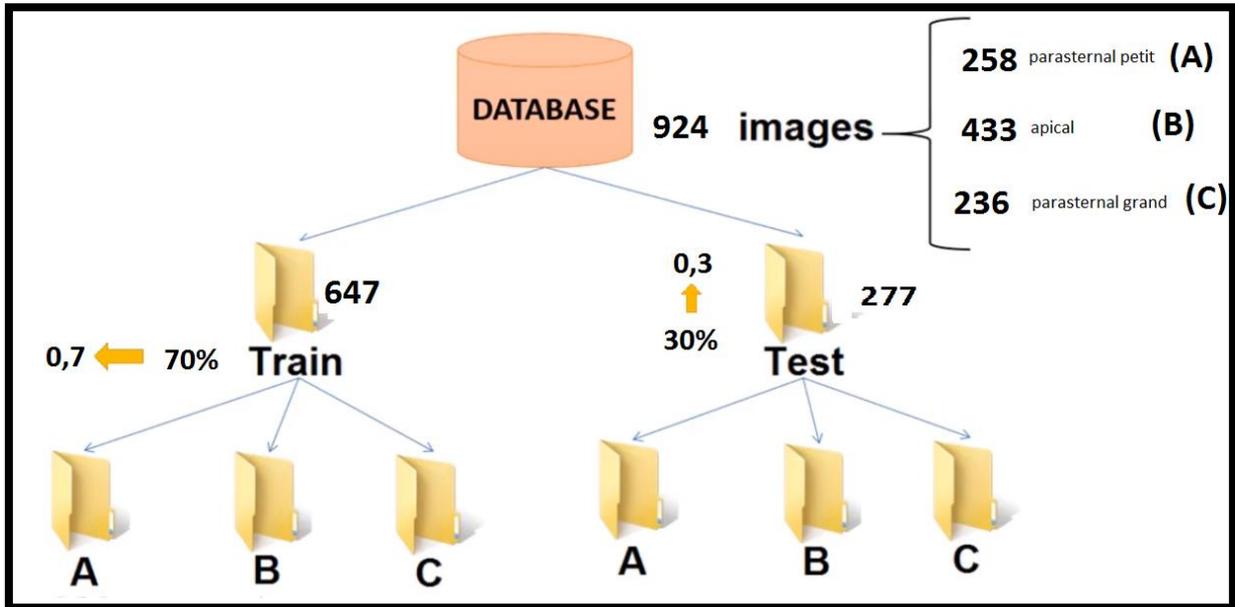


Figure III 12.1 distribution d'image dans la base de donnée (70%)

Confusion Matrix: resnet50

Output Class	Coupe_p_arasternale_p_etit_a_xe	71 25.5%	0 0.0%	0 0.0%	100% 0.0%
	apicale	0 0.0%	130 46.8%	0 0.0%	100% 0.0%
	coupe_p_arasternale_g_rand_a_xe	6 2.2%	0 0.0%	71 25.5%	92.2% 7.8%
		92.2% 7.8%	100% 0.0%	100% 0.0%	97.8% 2.2%
		Coupe_p_arasternale_etit_a_xe	apicale	coupe_p_arasternale_g_rand_a_xe	
		Target Class			

Figure III 12.2 confusion matrix de 70% de la base d'apprentissage

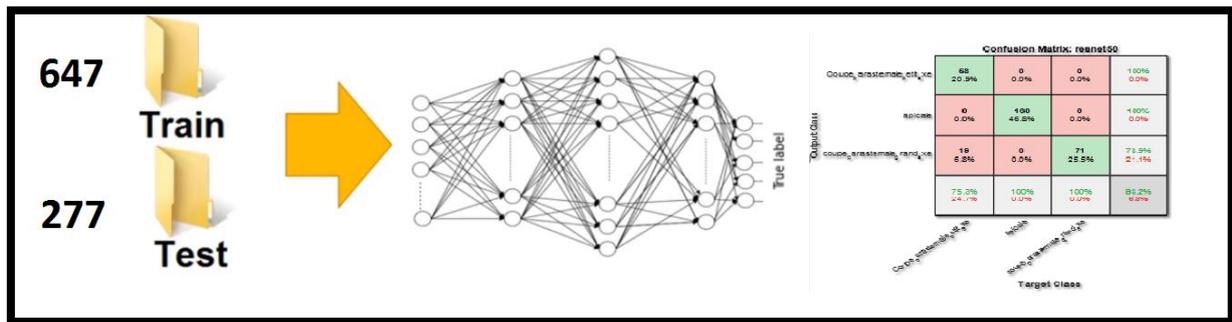


Figure III. 12 confusion matrix (0,7 mini path 64)

	axe		
	Coupe parasternal petit axe	Coupe apical	Coupe parasternal grand axe
Sensibilité	92,2%	100%	100%
Taux de classification	97,8%		

Tableau III. 9 résultat de sensibilité et taux de classification a 0,7 mini path 64

On utilise 70 % dans la base de d'apprentissage et 30 % dans la base de test

tp)(A>true positive = 71	(B>true positive =130	(c>true positive = 71
(tn)(A>true negative=201	(B>true negative =148	(c>true negative=201
(fp)(A>false positive =6	(B>false positive =0	(c>false positive =0
(fn)(A>false negative =0	(B>false negative =0	(c>false negative =6

$$\text{Taux de classification} = (71+130+71)/(71+130+71+6) = 0.978$$

$$\text{Taux de classification} = 97.8\%$$

$$\text{La sensitivity(A)} = (Tp)/(TP+FN) = 71/(71+0) = 1$$

$$\text{La sensitivity(A)} = 100\%$$

$$\text{La specificity(A)} = (TN)/(TN+FP) = 201/(201+6) = 0.922$$

$$\text{La specificity(A)} = 92.2\%$$

$$\text{La sensitivity(B)} = (Tp)/(TP+FN) = 130/(130+0) = 1$$

$$\text{La sensitivity(B)} = 100\%$$

$$\text{La specificity(B)} = (TN)/(TN+FP) = 148/(148+0) = 1$$

$$\text{La specificity(B)} = 100\%$$

$$\text{La sensitivity(C)} = (Tp)/(TP+FN) = 71/(71+6) = 0.922$$

$$\text{La sensitivity(C)} = 92.2\%$$

$$\text{La specificity(C)} = (TN)/(TN+FP) = 201/(201+0) = 1$$

$$\text{La specificity(C)} = 100\%$$

Discussion

Après l'analyse des résultats obtenus, On constate les remarques suivantes :

D'après la Figures 45, 46,47 La précision de l'apprentissage et de test augmente avec le nombre d'époque, ceci reflète qu'à chaque époque le modèle apprend plus d'informations. Si la précision est diminuée alors on aura besoin de plus d'information pour faire apprendre notre modèle et par conséquent on doit augmenter le nombre d'époque et vice versa

D'après l'augmentation d'apprentissage de 30%vers 50% et la diminution dans la base de teste de70% vers 50% en analyse que :

*la sensitivity(para-petite) augmente de85.5% vers 95.2% et La specificity(para-petite) diminuer de 95.6% vers 91.5%

* La sensitivity(apical) et la specificity(apical) reste toujours 100%

* La sensitivity(para-grand) diminuer de 94.5% vers 91.1% et La specificity(para-grand) augmente de 83.6% vers 94.9%

D'après la deuxième augmentation de la base d'apprentissage de50% vers 70% et la diminution de la base de teste de 50% vers 30% en analyse que :

*la sensitivity(para-petite) augmente de95.2% vers 100% et La specificity(para-petite) augmente de 91.5% vers 92.2%

* La sensitivity(apical) et la specificity(apical) reste toujours 100%

* La sensitivity(para-grand) augmente de 91.1% vers 92.2% et La specificity(para-grand) augmente de94.4 vers 100%

D'après les trois expériences en conclure que le taux de classification augmente lorsque on augmente la base d'apprentissage et diminuer la base de teste

4.2 Conclusion

Les résultats obtenus se sont améliorés à mesure que nous avons approfondie notre réseau et augmenté le nombre d'époque. La base d'apprentissage est également un élément déterminant dans les réseaux de neurones convolutionnels, il faut avoir une base d'apprentissage de grande taille pour aboutir à des meilleurs résultats

5 Conclusion générale

Dans ce projet nous avons discuté des notions fondamentales des réseaux de neurones en générale et des réseaux de neurones convolutionnels en particulier.

Les paramètres du réseau sont difficiles à définir a priori. Avec l'apprentissage profond et le CNN l'application de Resnet50 à simplifier notre classification de manier que l'augmentation de la base d'apprentissage et la diminution de la base de teste augmente le taux de classification et vise ver sa, par contre de taux d'erreur est en relation avec l'augmentation du nombre de mini path notamment nous avons rencontré quelques problèmes dans la phase d'implémentation, l'utilisation d'un CPU a fait que le temps d'exécution était trop couteux. Afin de régler ce problème on doit utiliser des réseaux de neurones convolutionnels plus profonds déployé sur un GPU au lieu d'un CPU sur des bases plus importantes.

Bibliographie

- [1] : Statistique Canada, « Principales causes de décès 2011 », Dernière consultation le 14 septembre 2016. <http://www.statcan.gc.ca/pub/82-625-x/2014001/article/11896-fra.htm>.
- [2]: J. Xu, S. Murphy, K.D. Kochanek et B.A. Bastian, « Deaths: Final Data for 2013 », National Vital Statistics Reports, 64(2):1–119, 2016.
- [3] : Société Française de cardiologie, « Généralités sur le cœur », Dernière consultation le 14 septembre 2016. <http://www.sante-sur-le-net.com/fiches-info/generalites-sur-le-coeur/>.
- [4] : Nandry, S., « Le cœur humain : organisation générale », Dernière consultation le 14 septembre 2016. <http://biofaculte.blogspot.ca/2015/06/le-coeur-humain-schema-animation.html>
- [5], [6]: N.A. Campbell et J. B. Reece, Biologie Campbell Reece. ERPI, 3 édition, 2007
- [7], [8] : Wikiversity, « Maladies cardiaques - Myocardite », Dernière consultation le 14 septembre 2016. https://en.wikiversity.org/wiki/Wikiversity_Journal_of_Medicine/Medical_gallery_of_Blausen_Medical_2014.
- [9] : Docteurcllic, « Medical gallery of Blausen Medical 2014 », Dernière consultation le 14 septembre 2016. <http://www.docteurcllic.com/maladie/myocardite.aspx>.
- [10], [11] : Doctorette.info, « Infarctus du myocarde causes et traitements », Dernière consultation le 14 septembre 2016. <http://www.doctorette.info/maladies-cardiovasculaires/infarctus-myocarde-causes-traitements>.
- [12]: Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, 86(11):2278–2324, 1998, pp. 1-46
- [13]: C. Farabet, C. Couprie, L. Najman and Y. LeCun, "Learning Hierarchical Features for Scene Labeling," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, 2013. pp. 1915- 1929
- [14]: Alex Krizhevsky, Geoffrey E Hinton, “Using very deep autoencoders for content-based image retrieval,” ESANN, 2011, pp. 1-7
- [15]: S. Ren, K. He, R. Girshick and J. Sun, "Faster RCNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, 2017, pp. 1137-1149

- [16]: O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, “Show and tell: A neural image caption generator,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3156-3164
- [17]: Y. Taigman, M. Yang, M. Ranzato and L. Wolf, “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 1701-1708
- [18]: A. Toshev, Ch. Szegedy, “DeepPose: Human Pose Estimation via Deep Neural Networks,” Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1653-1660
- [19]: P. Sermanet and Y. LeCun, “Traffic sign recognition with multi-scale Convolutional Networks,” The 2011 International Joint Conference on Neural Networks, San Jose, CA, 2011, pp. 2809-2813
- [20]: Y. LeCun and Y. Bengio, “Convolutional networks for images, speech, and time series,” in The Handbook of Brain Theory and Neural Networks, M. A. Arbib, Ed. Cambridge, MA: MIT Press, 1995, pp. 255–258
- [21]: Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. “A neural algorithm of artistic style.” arXiv preprint arXiv:1508.06576, 2015, pp. 1-16
- [22]: S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. CoRR, abs/1502.03167, 2015
- [23]: Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, Nov 1998.
- [24]: A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 25, pages 1097–1105. Curran Associates, Inc., 2012.
- [25]: Smistad, E. et Lindseth, F., « Real-time Tracking of the Left Ventricle in 3D Ultrasound Using Kalman Filter and Mean Value Coordinates », Dans MIDAS Journal, 2014.
- [26]: M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. In D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, editors, Computer Vision – ECCV 2014, pages 818–833, Cham, 2014. Springer International Publishing.
- [27]: K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556, 2014
- [28]: C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In The IEEE Conference on Computer

Vision and Pattern Recognition (CVPR), June 2015.

[29]: K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.

[30]: G. Huang, Z. Liu, and K. Q. Weinberger. Densely connected convolutional networks. CoRR, abs/1608.06993, 2016

[31]: <https://fr.wikipedia.org/wiki/MATLAB>