Chapitre 1

Contexte Médical

1. Introduction

L'évolution des techniques d'imagerie médicale, spécialement l'imagerie numérique, a récemment changé et d'une manière profonde, la pratique médicale. Cette évolution a donné lieu à une nouvelle dimension liée aux techniques de traitement des images médicales et à l'information extraite par le moyen de ces techniques.

Toutefois, et en revanche de la technologie avancée dans le secteur médical, l'analyse du cancer du sein reste un réel problème de santé public et un sujet de recherche d'actualité très délicat à aborder.

L'imagerie mammographique est l'une des modalités la plus couramment utilisée. Cet outil à quelle nous nous sommes intéressés dans ce travail est devenu un outil indispensable pour tout examen clinique lié au cancer du sein.

Dans ce chapitre, nous présenterons brièvement l'anatomie du sein chez la femme, ce qui va nous permettre d'aborder la notion du cancer du sein et son impact sur la population féminine. Ensuite, nous exposerons la procédure du dépistage du cancer du sein, puis nous continuerons avec quelques bases de mammographie.

2. L'anatomie du sein chez la femme

Un sein normal comprend des éléments mammaires glandulaires entourés de stroma adipeux et mammaire, qui sont entourés d'une structure fibreuse en nid d'abeille à ligaments de Cooper minces. Les éléments glandulaires sont composés de canaux lactifères menant du mamelon et de la branche aux canaux excréteurs, aux canaux inter-lobulaires et aux canaux terminaux conduisant aux acini producteurs de lait (figure 1.1)[1]. Les conduits sont tapissés d'épithélium composé d'une couche myoépithéliale cellulaire externe et d'une couche cellulaire sécrétoire interne. Les canaux et le tissu glandulaire s'étendent vers l'arrière selon une distribution en éventail composée de 15 à 20 lobes drainant chacun des canaux lactifères, la majeure partie du tissu glandulaire se trouvant dans la partie supérieure de la poitrine près de l'aisselle. Le tissu adipeux entoure le tissu glandulaire. La graisse rétro-glandulaire est située en arrière du tissu glandulaire et est décrite par le Dr Laszlo Tabar comme un « no man's land » dans lequel aucun tissu glandulaire ne doit être observé. Le muscle pectoral se trouve derrière la graisse au-dessus de la paroi thoracique[2].

La taille et le poids du sein varient durant la vie d'une femme. Ils sont principalement influences par la quantité de tissu adipeux présent dans le sein. Le développement du sein chez la femme se produit en deux étapes sous l'influence des hormones : à l'approche de la puberté, le volume du sein augmente en réponse à l'action de l'œstrogène sur les canaux galactophores, la deuxième phase survient sous l'influence de la production de progestérone quand débute l'ovulation. La croissance du sein est essentiellement terminée environ un an et demi après la fin de la période des premières menstruations, tandis que le développement des alvéoles continue en réponse à une stimulation hormonale[3].

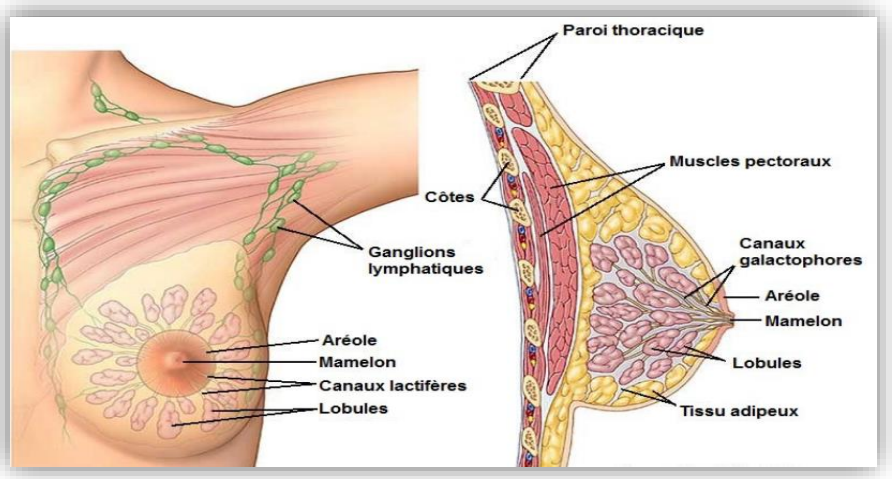


FIGURE 1.1 : L'ANATOMIE DU SEIN CHEZ LA FEMME[1].

3. Cancer de sein

Le cancer du sein est le cancer dont le diagnostic est le plus fréquent chez les femmes dans le monde[4]. Sa présence augmente avec l'âge, mais il peut aussi toucher des femmes jeunes puisqu'un quart des cas sont diagnostiqués avant l'âge de 50 ans. Le cancer du sein prend naissance dans les cellules du sein. La tumeur cancéreuse (maligne) est un groupe de cellules cancéreuses qui peuvent envahir et détruire le tissu voisin. Elle peut aussi se propager (métastases) à d'autres parties du corps. Le premier symptôme du cancer du sein est la présence d'une boule au niveau du sein, correspondant à la tumeur. Elle peut également s'accompagner de ganglions durs au niveau de l'aisselle (ganglions axillaires) correspondant à une propagation du cancer, ainsi qu'à des modifications cutanées au niveau du sein et du mamelon (peau d'aspect capitonné et un mamelon qui entre au lieu de sortir). Le sein peut progressivement se déformer et s'ulcérer, ce qui se traduit parfois par un écoulement du mamelon, d'un seul côté. Si le cancer est diagnostiqué tardivement, la tumeur peut se propager et déclencher d'autres symptômes tels que des nausées, des vomissements, une perte de poids, une jaunisse, des douleurs osseuses ou encore des céphalées, un essoufflement ou une toux.

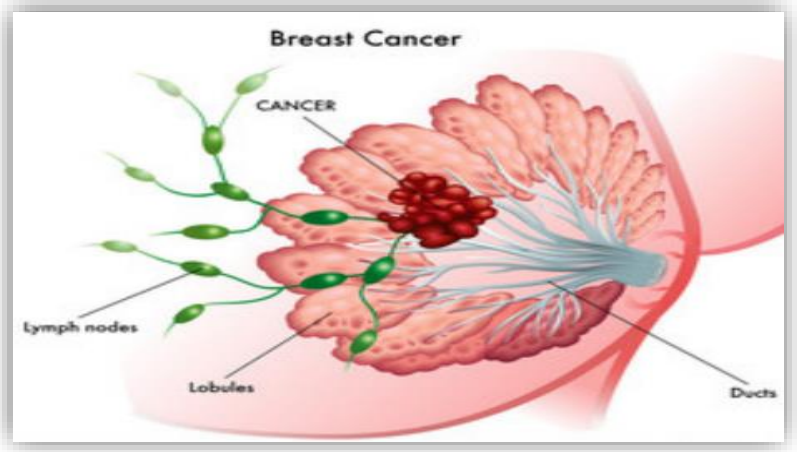


FIGURE 1.2 : EXEMPLE D'UN CANCER DE SEIN[4].

En bref, le cancer du sein est une tumeur maligne se développant à partir des cellules constituant la glande mammaire. Les cellules malignes se multiplient de manière anarchique et forment la tumeur. Lorsqu'un cancer du sein n'est pas traité, les cellules tumorales se propagent localement et envahissent les organes de voisinage. Elles peuvent également se propager par voie sanguine ou lymphatique pour atteindre les organes situés à distance (métastases). Les organes les plus fréquemment atteints par les métastases sont les os, les poumons, le foie et le cerveau. Il existe deux types de tumeurs du sein : les tumeurs non cancéreuses ou ”bénigne” et les tumeurs cancéreuses ”maligne”[5].

3.1. Tumeurs bénignes

Les tumeurs bénignes ont des contours bien délimités. Elles ont une croissance lente et restent localisées dans le tissu ou l'organe dans lequel elles sont apparues. Elles n'entraînent pas de métastases dans d'autres endroits du corps. Les tumeurs bénignes sont composées de cellules qui ressemblent aux cellules normales du tissu concerné. La tumeur bénigne qui se développe le plus souvent dans le sein est le fibroadénome.

Lorsqu’une tumeur est diagnostiquée comme étant bénigne, les médecins la laissent généralement seule plutôt que de la retirer. Même si ces tumeurs ne sont généralement pas agressives envers les tissus environnants, elles peuvent parfois continuer à se développer, en appuyant sur les organes et en provoquant des douleurs ou d’autres problèmes. Dans ces situations, la tumeur est retirée, permettant à la douleur ou aux complications de s’apaiser.[6]

3.2. Tumeurs malignes

Les tumeurs malignes sont cancéreuses et agressives car elles envahissent et endommagent les tissus environnants. Lorsqu’une tumeur est soupçonnée d’être maligne, le médecin procédera à une biopsie afin de déterminer la gravité ou l’agressivité de la tumeur.[6]

3.2.1. Les symptômes

Les symptômes listés ci-dessous ne signifient pas nécessairement qu’il s’agit d’un cancer du sein. Mais si c’est le cas, il est important de le détecter le plus tôt possible[5].

a) Une boule dans un sein : Une boule ou une masse dans un sein est le signe d’un cancer du sein le plus couramment observé. Cette masse, en général non douloureuse, est le plus souvent de consistance dure et présente des contours irréguliers. Elle apparaît par ailleurs comme « fixée » dans le sein.

b) Des ganglions durs au niveau de l’aisselle : Une ou plusieurs masse(s) dures à l'aisselle signifient parfois qu’un cancer du sein s’est propagé aux ganglions axillaires. Les ganglions restent toutefois indolores.

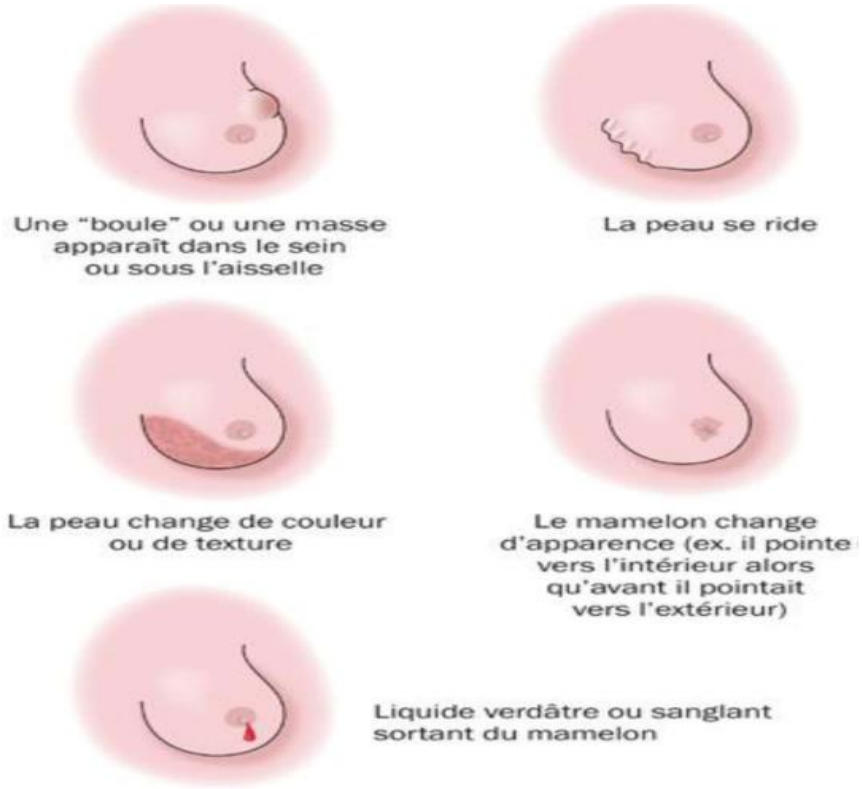


FIGURE 1.3 : SYMTÔMES DU CANCER DE SEIN[4].

c) Des modifications de la peau du sein et du mamelon : La peau du sein peut devenir capitonnée (et prendre l'aspect d'une peau d'orange) ou plissée.

d) Le mamelon peut pointer vers l'intérieur, alors qu'habituellement, il est dirigé vers l'extérieur.

e) Le sein peut se déformer et perdre de son galbe, des rides peuvent apparaître.

f) La peau du sein peut être rouge, écorchée (ulcérée), recouverte de croûtes et celle du mamelon peut se mettre à peler.

g) Un écoulement d'un seul mamelon : peut être le signe de cancer du sein, en particulier s'il se manifeste sans qu'il n'y ait de compression du mamelon et s'il contient du sang ou s'il est verdâtre.

h) Un changement de la taille ou de la forme du sein : Une rougeur, un œdème et une chaleur importante au niveau du sein peuvent être le signe d'un cancer du sein inflammatoire (le cancer inflammatoire du sein est un cancer rare ; il représente 1 à 4 % de tous les cas de cancer du sein. Les cellules cancéreuses ont la caractéristique de se déplacer rapidement dans les vaisseaux lymphatiques de la peau du sein qu'elles finissent par bloquer. C'est ce qui provoque l'inflammation locale du sein). La constatation d'un de ces signes impose une consultation médicale sans retard. Le médecin jugera des examens complémentaires nécessaires.

3.2.2. Les autres symptômes

Si le cancer n'est pas diagnostiqué dès l'apparition des premiers symptômes, la tumeur peut grossir et se propager vers d'autres parties du corps, entraînant ainsi d'autres symptômes dits plus tardifs, tels que des douleurs osseuses, des nausées, une perte d'appétit, une perte de poids et une jaunisse, une essoufflement, une toux et une accumulation de liquide autour des poumons (épanchement pleural), des maux de tête, une vision double et une faiblesse musculaire[5].

4. Dépistage du cancer du sein

Le dépistage du cancer consiste à réaliser des examens de surveillance d'une personne supposée en bonne santé, dans le but d'arriver à détecter des anomalies pouvant être des signes avant-coureur d'un cancer, bien avant l'apparition des premiers symptômes. C'est une démarche qui s'avère très efficace en ce sens qu'elle permet de réduire de façon significative le taux de mortalité lié au cancer du sein en améliorant les chances de guérison. En effet, s'il est détecté tôt, le cancer du sein peut être guéri dans neuf cas sur dix, ce qui participe à une réduction de 15 à 21% du taux de mortalité lié à ce cancer[7][8].

4.1. L'imagerie par résonance magnétique (IRM)

L'imagerie par résonance magnétique, est un moyen d'investigation dans le dépistage et le suivi thérapeutique des cancers. Proposant une visualisation de coupes anatomiques extrêmement fines, elle permet de différencier les tissus pathologiques des tissus sains et de réaliser une "cartographie" très précise et en trois dimensions des tumeurs. L'IRM utilise les modifications provoquées par un champ magnétique, envoyé par la machine, dans les noyaux des atomes d'hydrogène présents dans l'organisme. Cet examen est indolore et très fiable. L'imagerie par résonance magnétique n'utilise pas les rayons X[9]. Parmi toutes les modalités radiologiques, l'IRM est certainement la technique la plus performante dans la recherche des lésions du sein. Rarement, une prise de contraste diffuse du tissu fibroglandulaire peut limiter l'interprétation de l'examen, voire même masquer une lésion. La spécificité de l'IRM est un peu moins bonne. Des petits foyers de prise de contraste sont difficilement caractérisables. Certaines lésions bénignes comme les fibroadénomes peuvent avoir un comportement identique à celui des tumeurs malignes. Cette situation requiert très souvent une biopsie ou une exégèse. On peut rencontrer également des situations où l'on trouve à l'IRM une prise de contraste dont l'aspect est inquiétant, et dans le même temps, ne voir aucune lésion sur les mammographies et à l'échographie (même après réexamen). Il existe alors une indication à effectuer une biopsie sous IRM[6][3].

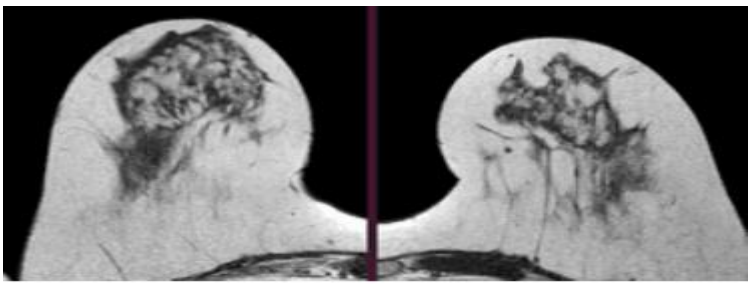


FIGURE 1.4: EXEMPLE D'IMAGERIE IRM MAMMAIRE[8].

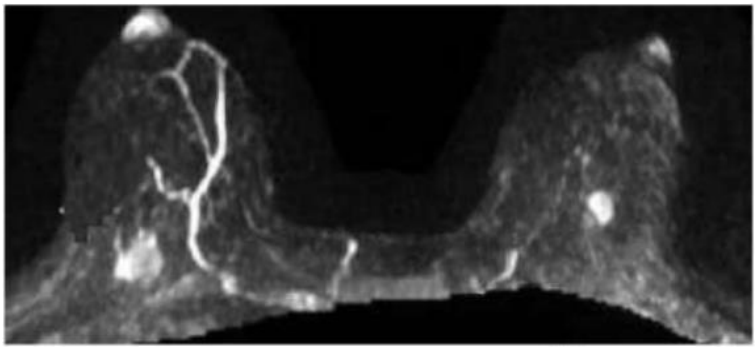


FIGURE 1.5: EXEMPLE D'IRM MAMMAIRE BI-SEIN[5].

4.2. L'échographie

L'échographie est une technique d'exploration de l'intérieur du corps basée sur les ultra-sons. Une sonde envoie un faisceau d'ultrasons dans la zone du corps à explorer. Selon la nature des tissus, ces ondes sonores sont réfléchies avec plus ou moins de puissance. Le traitement de ces échos permet une visualisation des organes observés[10].



FIGURE 1.6: EXEMPLE D'ECHOGRAPHIE MAMMAIRE[9].

Une échographie complète certaines mammographies difficiles à interpréter, en cas de seins denses rendant son analyse très difficile car pouvant masquer de petites lésions ayant la même densité que le tissu mammaire environnant (figure 1.6)[11].

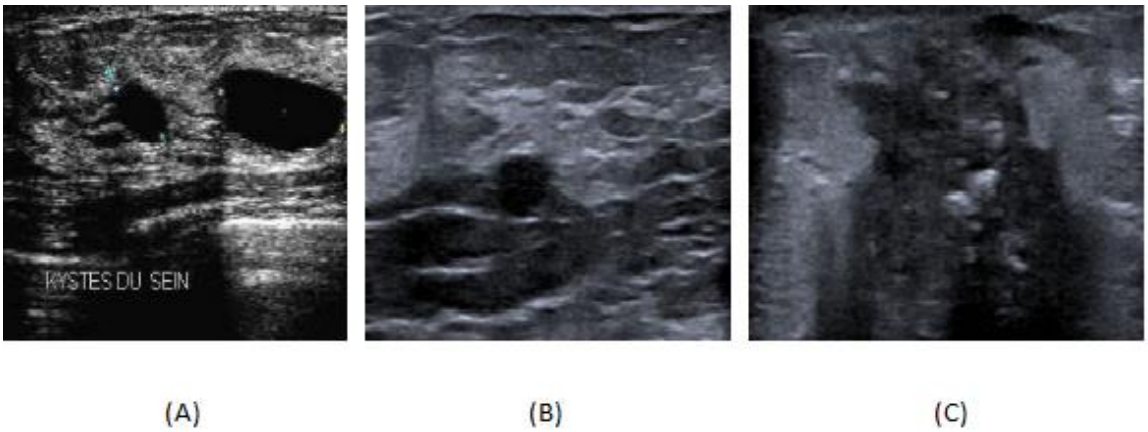


FIGURE 1.7: EXEMPLES D'IMAGES ECHOGRAPHIQUES DU SEIN (A) KYSTE DU SEIN, (B) LESION NON VISUALISE SUR LA MAMMOGRAPHIE, (C) PRESENCE D'UNE MASSE DE 5 CM, CONTENANT DES CALCIFICATIONS[10].

4.3. La mammographie

La mammographie est la technique de référence incontournable pour l'exploration du sein, la plus performante en matière de surveillance et de dépistage précoce du cancer du sein[12]. La qualité de l'image mammographique est caractérisée par trois paramètres physiques : la résolution, le contraste et le bruit. Elle dépend de la chaine radiologique (figure 1.8) dont la défaillance d'un seul maillon de cette chaine peut fausser le diagnostic. Selon Bremond, près de 40 % des cancers ratés à la mammographie sont dus à une mauvaise qualité des clichés. Il s'agit le plus souvent d'un défaut de positionnement du sein[11].

Un mammographe à le même principe qu'une radiographie standard, la quantité de rayon X produites dans le tube est de faible dose.

Comme on peut le voir sur la figure 1.8, Le sein à radiographier est comprimé entre le support patient, sous lequel se trouve le détecteur, et la pelote de compression. Actuellement, cette compression est assurée par un système pneumatique avec commande au pied ce qui permettre de libérer les mains du manipulateur pour donner un bon étalement et un positionnement correct au sein.

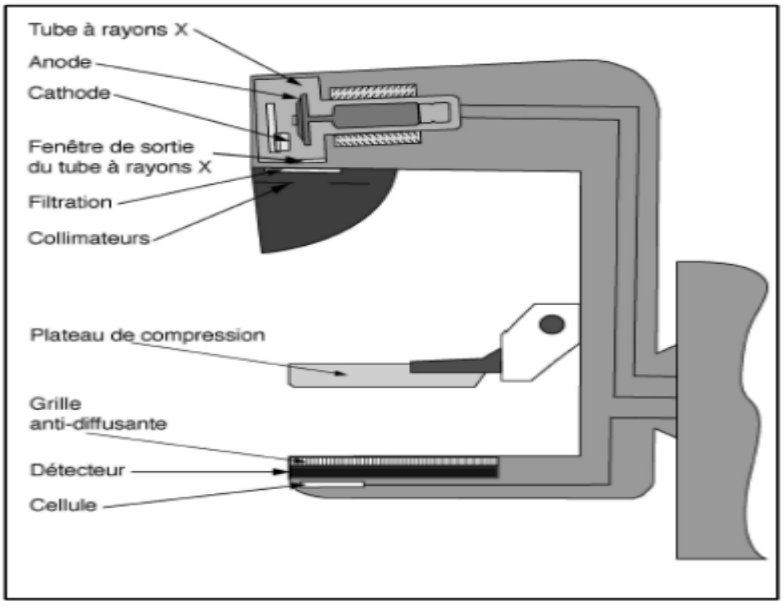


FIGURE 1.8: LES COMPOSANTS D'UN MAMMOGRAPHE.

4.3.1. Acquisition des images mammographiques

Le sein a une structure tridimensionnelle ; l'image mammographique est le résultat de la superposition des différents éléments anatomiques qui ont des coefficients d'atténuation aux rayons X presque identiques. Les densités des structures mammographiques sont basses et il y a un recouvrement entre les densités d'éléments normaux et des tissus pathologiques. La réalisation d'une mammographie passe par plusieurs étapes. D'abord les rayons X sont générés par un tube à rayons X. Ce faisceau est filtré puis collimaté pour éviter les fuites de radiations hors de la zone de l'examen. Il traverse ensuite le sein sous compression, pour atteindre le récepteur d'image placé immédiatement sous le sein. Une Grille anti-diffusante est le plus souvent interposée entre le sein et le récepteur pour atténuer les rayons X diffusés

par le sein qui peuvent causer une diminution de la lisibilité de l'image en introduisant un flou indésirable. Différentes vues sont obtenues selon l'angle de traversée de ces rayons. La figure (1.9) montre les deux vues les plus utilisées dans l'acquisition d'images mammographiques:[13]

La vue Médio Latérale Oblique (MLO) son acquisition est réalisée avec une position oblique par rapport au sein. Lors de la capture de ce type de vue, une région appartenant au muscle pectoral peut être présente dans la mammographie. Cette région se trouve au coin supérieur gauche ou au coin supérieur droit de l'image (figure 1.9.A).

La vue Carnio Caudale (CC) la position du sein lors de l'acquisition de la vue CC doit être horizontale (figure 1.9.B). Notons que cette vue montre autant que possible les tissus glandulaires (les canaux et les lobules) ainsi que les tissus gras qui entourent les tissus glandulaires.

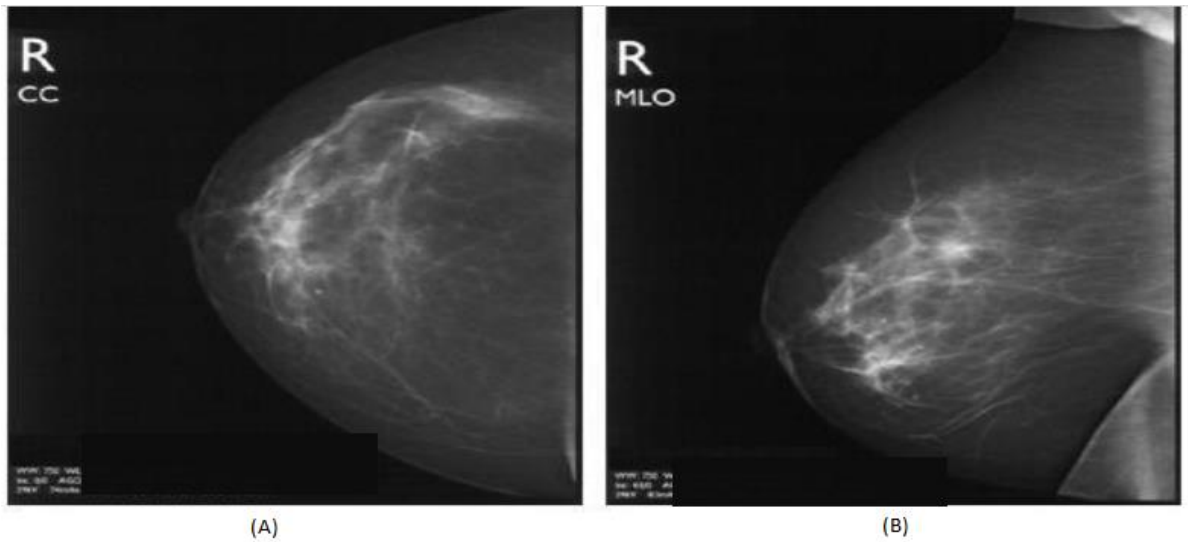


FIGURE 1.9 : DEUX VUES DE MAMMOGRAPHIE : (A) LA VUE MLO ET (B) LA VUE CC.

4.3.2. Corrélation entre la structure du sein et les images mammographiques

Une coupe longitudinale passant par le mamelon (figure 1.10), laisse entrevoir l'architecture de la glande mammaire : elle se divise en 15 à 20 unités glandulaires indépendantes, lobes mammaires, divisés en lobules par des cloisons conjonctives, qui s'achèvent en douzaines de minuscules bulbes, acini, base de tout raisonnement physiopathologique du sein. Les lobes, lobules et bulbes sont reliés entre eux par les canaux galactophores, qui se terminent par un canal unique s'abouchant au niveau du mamelon, situé au centre d'une zone pigmentée, l'aréole. La plus grande partie du sein est constituée de tissu adipeux qui comble l'espace situé entre les différentes structures du sein. Les tissus mammaires baignent dans des liquides lymphatiques, qui permettent d'éliminer les déchets, les cellules mortes et autres débris.

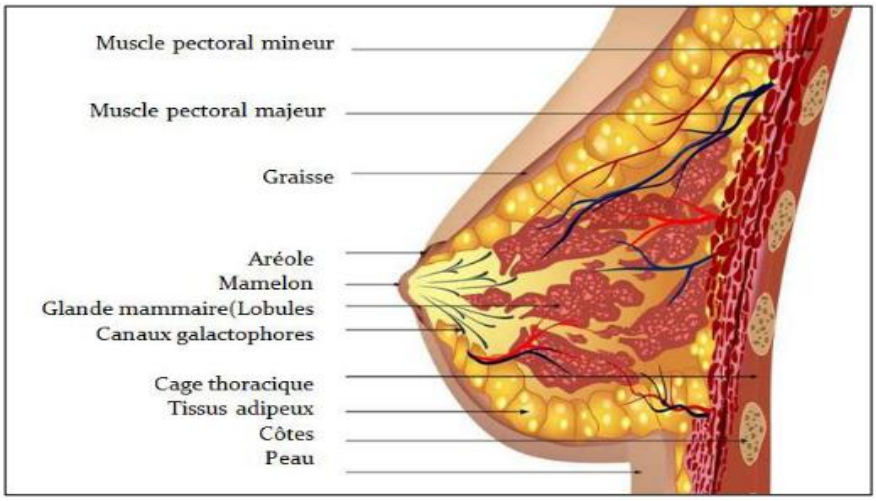


FIGURE 1.10 : COUPE DU SEIN EN PROFIL.

Radiologiquement, l'image mammographique générée est le résultat de différences atténuation aux rayons X par les différents types de tissu. L'aspect de la glande mammaire est en effet extrêmement variable selon la patiente, son âge, et par rapport au cycle menstruel. En effet, la graisse est considérée comme une zone radio transparente, elle apparait très sombre sur un cliché mammographique (figure 1.11).

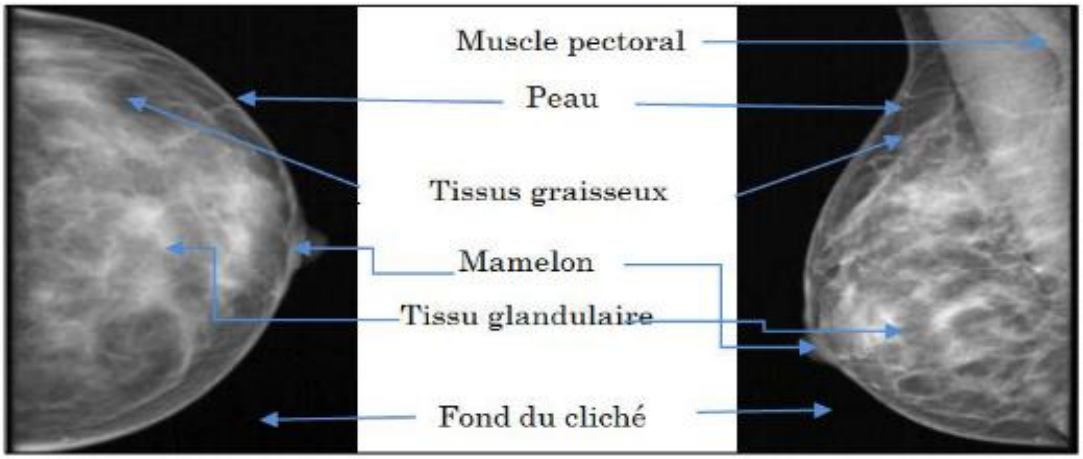


FIGURE 1.11 : ASPECT RADIOLOGIQUE DU SEIN NORMAL SUR UNE INCIDENCE FACE/OBLIQUE.

En revanche, les zones radio opaques apparaissent claires et correspondent au tissu fibro-glandulaire et au calcium qui est le composant essentiel des MCS[1] mammaires.

4.3.3. Signes d'appel radiologique d'un cancer du sein

Le dépistage du cancer du sein amène à découvrir quarts types d'anomalies purement radiologiques : les masses, les MCS, les asymétries de densité, distorsions architecturales. Dans le cadre de cette étude on s'intéresse à l'analyse des deux premiers signes d'un cancer du sein à savoir les masses et les MCS.

[1] microcalcifications

A. L'analyse des masses mammaires

Une masse est une lésion occupant l'espace visible sous deux incidences différentes. Si elle est visible sur une seule incidence, il s'agit d'une asymétrie. En excluant les « opacités physiologiques », les kystes, il reste que le problème des masses qui doivent être décrit par des critères radiologiques : Taille, forme, contour, densité.

- La taille : elle est variable de quelques millimètres à plusieurs centimètres. Cependant la taille ne prédit pas le caractère malin, sauf sur des clichés successifs lorsque l'on voit la taille augmenter régulièrement.

- La densité : les masses sont en général denses (plus blanches) que le tissu mammaire adjacent.

- La forme : on peut envisager quatre formes de masse (tableau 1.1).

- Le contour : c'est le meilleur critère d'étude des masses. Le tableau 1.1 montre cinq types de contours masses.

B. L'analyse des calcifications mammaires

Les calcifications mammaires sont des petits dépôts calciques, composés des substances chimiques $Ca_3(PO4)_2$, $CaCO_3$ et $Mg_3(PO\ 4)_2$, opaques aux rayons X et visibles sur une mammographie. Selon leur origine, on distingue les microcalcifications intra-canalaires qui se situe soit au niveau des canaux intermédiaires responsables de l'acheminement (figure 1.12.A) soit dans les lobules d'où son nom : les microcalcifications lobulaires (figure 1.12.B). Leur morphologie est déterminée par l'endroit de leur formation[14].

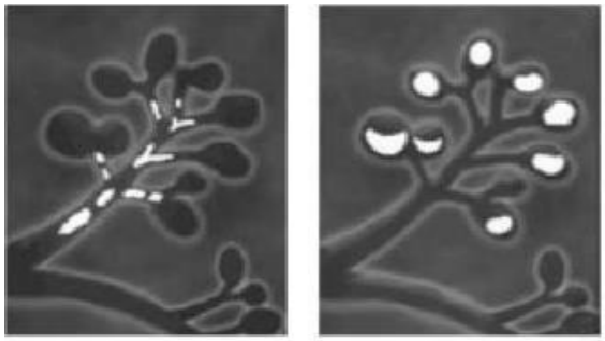


FIGURE 1.12 : EXEMPLE DE : (A) MCS CANALAIRES ET (B) MCS LOBULAIRES.

L'analyse des calcifications permet de distinguer deux grandes catégories, Les macrocalcifications qui sont des ponctuations blanches larges, bien visibles en mammographie, de taille dépasse 1mm et elles sont identifiées toujours comme bénignes, et les microcalcifications qui sont définies comme des calcifications de petite taille inférieures à 0.5 mm (en général comprises entre 200 et 500 µm), parfois à la limite de la visibilité et parfois on peut facilement les confondues avec le bruit présent dans les images de mammographiques. Pour cela l'étude des calcifications nécessite des clichés radiologiques d'une qualité irréprochable, en excluant les corps étrangers (fausses calcifications) et les artéfacts (traces de doigts, pommades opaques, un mamelon mal positionné,..). L'analyse

morphologique, et l'étude de leur mode de distribution et leur nombre restent les éléments majeurs de diagnostic, après une confirmation par des micro ou macro-biopsies.

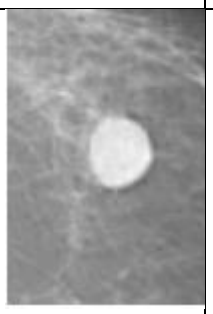
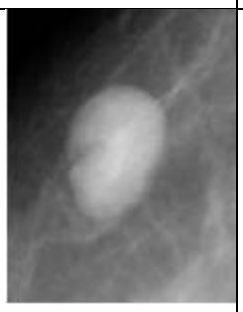
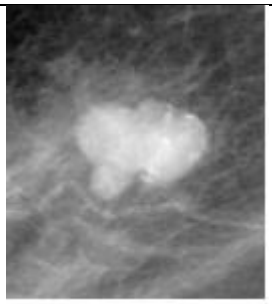
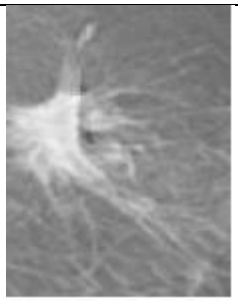
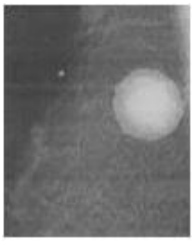
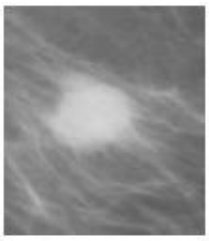
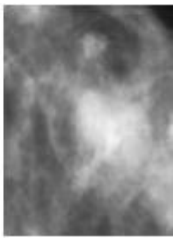
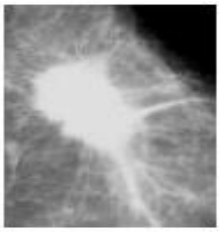
Forme					
	Ronde	Ovale	Lobulée		Irrégulière
Cliché correspondant					
Contour					
	Net	Masqué	Microlobulé	Flou	Spiculé
Cliché correspondant					
Interprétations	Il s'agit d'une transition brusque entre la lésion et le tissu environnant. Le contour est alors net et bien défini.	Le contour est caché par le tissu normal adjacent. Ce terme est employé pour caractériser une masse circonscrite dont une partie du contour est cachée.	De courtes dentelures du contour créent de petites ondulations.	Le contour est mal défini. Ce caractère peut correspondre à une infiltration.	La masse est caractérisée par des lignes radiaires (spicules.) prenant naissance sur le contour de la masse.

TABLEAU 1.1 : TABLEAU RÉCAPITULATIF DE DIFFÉRENTS FORMES ET CONTOUR DES MASSES.

-La Morphologie : permet le plus souvent de séparer MCS bénignes et malignes. Les microcalcifications arrondies ou ovales, uniformes dans leurs tailles et leurs formes, sont souvent bénignes. A l'inverse, plus elles sont irrégulières, plus elles ont de risque d'être malignes. Le tableau 1.3 donne des exemples explicatifs de chaque cas.

-Nombre de MCS : lorsque ce paramètre augmente, le risque de carcinome croit. L'existence d'un certain nombre de µcas groupées en une zone sous forme d'un foyer est en soi un signe qui mérite d'être pris en considération. Le terme foyer est utilisé lorsque de multiples calcifications (au moins 3) occupent un petit volume (>1 cm^2). Cette distribution reflète une lésion bénigne ou maligne.

- **Répartition :** la distribution des microcalcifications sur le sein décrit leur degré de malignité. Les déférentes distributions possibles des MCS sont récapitulées dans le tableau 1.2.

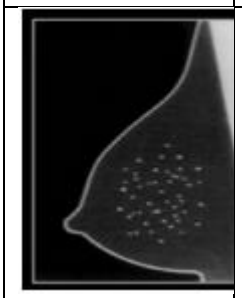
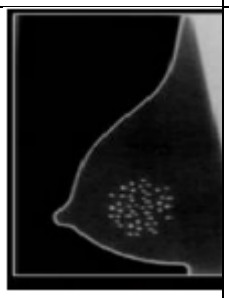
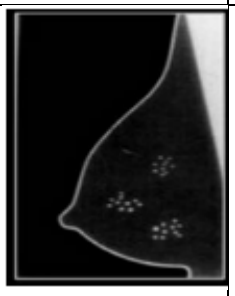
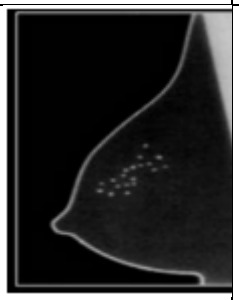
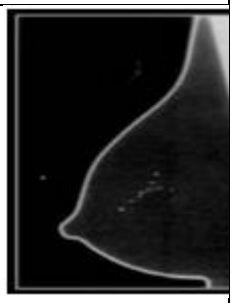
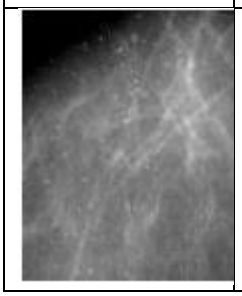
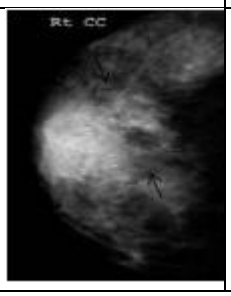
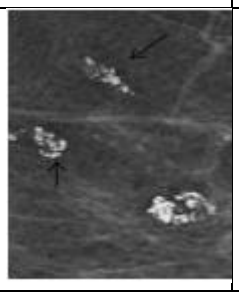
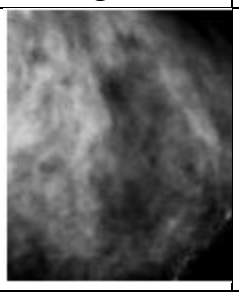
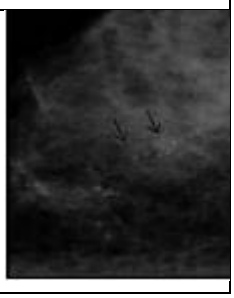
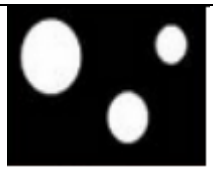
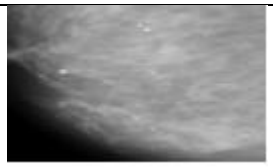
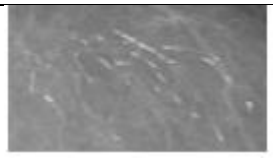
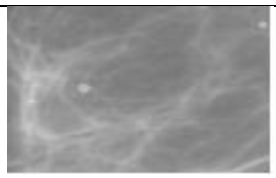
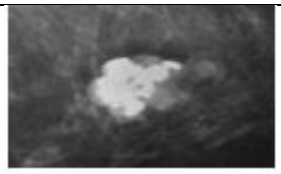
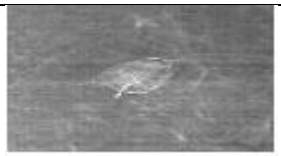
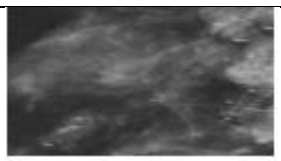
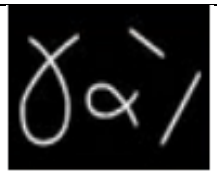
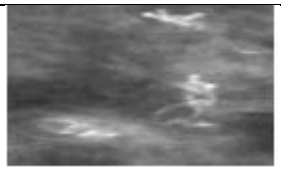
Eparses	Régional	En foyer	Segmentaire	Linéaire
				
Les MCS sont distribuées de façon aléatoire dans l'ensemble du sein.	Les MCS sont dispersées dans un large volume du tissu mammaire.	Lorsque de multiples MCS (au moins cinq) occupent un petit volume.	Les MCS sont disposées les unes derrières les autres sous forme d'une ligne.	Dépôts calciques dans des canaux galactophores et leurs branches.
				

TABLEAU 1.2 : TABLEAU DE DIFFÉRENTES DISTRIBUTIONS DES MCS.

Types	Forme	Cliché	Critère de forme et de taille
rondes			✓ Forme ronde. ✓ Tailles variables.
à centres clairs			✓ Taille de 1 mm à plus de 1 cm. rondes ou ovales, à surface lisse et à centre clair. ✓ La paroi qui les entoure est plus épaisse que celle des (Mcs) en coquille d'œuf.
en bâtonnets			✓ Habituellement associées à une dilatation du canal galactophore orientées vers le mamelon.
vasculaires			✓ Forme en rails ou linéaires associées à des structures tubulaires.
cutanées			✓ Ont typiquement un centre clair. ✓ Des clichés en incidence tangentielle confirment leurs localisation cutanée.
coralli-formes			✓ De grande taille (≥ 2-3 mm de diamètre).
coquille d'œuf			✓ Très fines apparaissent comme des dépôts calciques sur la surface d'une sphère. ✓ mesure moins de 1 mm.
lait calcique			✓ Sédimentées dans le fond de kystes. ✓ L'incidence de profil permet de démontrer leurs formes caractéristiques : semi-lunaires, en croissants, curvilignes ou linéaires.
de suture			✓ Correspondent à des dépôts calciques sur du matériel de suture. ✓ Typiquement linéaires ou tubulaires avec des nœuds.

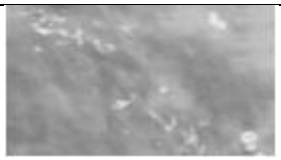
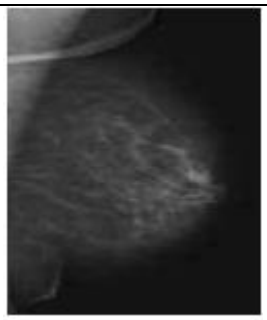
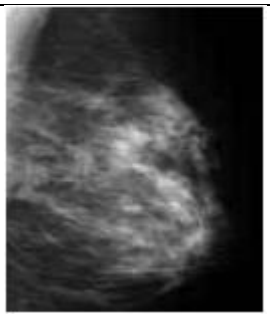
dystro-phiques			✓ Plus de 0.5 mm de diamètre. ✓ De formes irrégulières. ✓ Souvent rencontrées dans un sein irradié.

TABLEAU 1.3 : TABLEAU RÉCAPITULATIF DE DIFFÉRENTS TYPES DE MICROCALCIFICATION.

4.3.4. Densité mammaire

La détermination de la densité mammaire fait partie intégrante de l'analyse des clichés radiologiques, soulignant ainsi la difficulté potentielle à séparer et détecter une lésion qui possède la même densité que le tissu mammaire normal environnant. Le sein d'adulte jeune est riche en tissu glandulaire et conjonctif et apparaît en mammographie comme une plage très dense occupant la presque totalité du sein (tableau 1.4).

Type de densité mammaire	Cliché correspondant	Interprétations
1		Le sein est presque entièrement graisseux : seins clairs. (moins de 25 % de la glande mammaire).
2		Il y a des opacités fibroglandulaires dispersées. (approximativement 25 à 50 % de la glande mammaire).
3		Le tissu mammaire est dense et hétérogène. (approximativement 51 à 75 % de la glande mammaire).

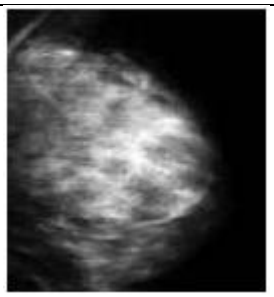
4		Le tissu mammaire est extrêmement dense. (plus de 75 % de la glande mammaire).

TABLEAU 1.4 : LA CLASSIFICATION BI-RADS DE LA DENSITÉ MAMMAIRE.

À la ménopause, les seins sont souvent « déshabités » et donc tout à fait radiotransparents. Les images tumorales s'y inscrivent alors avec une très particulière netteté. La classification BI-RADS de *l'American College of Radiology* permet de classer les seins en quatre groupes pour lesquelles le risque de rater un cancer en raison du caractère hétérogène du tissu fibro-glandulaire augmente lorsque l'on passe de la catégorie 1 à la catégorie 4 (tableau 1.4).

4.3.5. Classification des pathologies mammaires

Le grand avantage d'un lexique de classification des lésions mammaires est de standardiser les comptes-rendus mammographiques en employant des termes identiques, clairs et concis. Cette standardisation évite l'utilisation les mots, et les expressions inconvenants, aberrants et inadaptés, comme « seins normalement meublés et charpentés », « lacune échographique », « nodule plein », etc. De nos jours, deux systèmes d'aide à la rédaction des comptes rendus sont pratiqués, on note : la classification de Le Gal et celle de BIRADS.

La classification de Le Gal ; conçue en 1976, traite juste le cas des MCS selon leurs morphologie en 5 types de valeur prédictive de malignité croissante (tableau 1.5). Cette donnée morphologique a été associée par le même auteur à d'autres critères tels que l'âge, le nombre de foyers, et le nombre de MCS par foyer afin d'établir un score permettant d'aider à la décision chirurgicale[11].

Type I	Type II	Type III	Type IV	Type V
Rondes	Régulières	Grains de sel	Points irréguliers	Vermiculaires
				
0% de cancers	39% de cancers	39% de cancers	59% de cancers	96% de cancers

TABLEAU 1.5 : CLASSIFICATION DES MICROCALCIFICATIONS SELON LE GAL.

Bien que cette classification à l'avantage d'être simple. Seulement, son défaut est qu'elle traite seulement la forme des MCS sans passer par leurs distributions. Le cas des masses est absolument non envisagé. Pour pallier aux lacunes de cette classification, *l'American College of Radiology* l'ACR (novembre 1998) a établi le système BIRADS (*Breast Imaging Reporting and Data System*).

Ce système de classement des images radiologiques, actuellement utilisé et recommandé pour le dépistage du cancer du sein, fournit au médecin traitant une indication de la conduite à tenir. L'échelle varie de 0 à 5. Chaque degré correspond à un niveau de suspicion de l'image mammographique (tableau 1.6).

Le radiologue effectue un examen clinique, réalise une échographie (si nécessaire) et compare avec des anciens clichés (si disponible). Il rédige une conclusion de la forme :

Pas de lésion suspecte décelée ce jour. Pas d'anomalie péjorative constatée par rapport à l'examen du ..... ACR 2 bilatéral

Cette conclusion indique clairement le degré de suspicion de cancer et indique la conduite à tenir. Cependant, arriver à une telle conclusion reste une tache extrêmement délicate et essentielle pour la survie ou non d'une femme. La difficulté se résume au niveau de l'étape de détection des (contours, formes, distribution, sein dense, ...), de l'étape de décision (mammographie bénigne ou maligne) et, d'autre part du volume important des mammographies analysées, ce qui nécessite un effort et un temps énorme, d'où la nécessité de l'automatisation de la lecture des mammographies et de la prise de décision[15].

Catégorie	Lésion	Description	Conduite à tenir
ACR0		- Cliché insuffisant dans l'attente d'un second avis, pour une classification définitive.	Nécessité d'investigations complémentaires
ACR1	Masse	-Aucune anomalie.	Négative : Mammographie normale
	MCS		
ACR2	Masse	-Opacité ronde avec macrocalcifications. -Ganglion intramammaire. -Opacité ronde correspondant à un/des kystes en échographie.	Anomalies bénignes ne nécessitant ni surveillance ni examen complémentaire
	MCS	-Calcifications sur matériel de suture (Cicatrice connue). -MCS sans opacité. - MCS annulaires ou arciformes, semi-lunaires, sédimentées. -Calcifications cutanées et calcifications punctiformes régulières diffuses.	

ACR3	Masse	-Opacité bien circonscrite(s), ronde(s), ovale(s) sans microlobulation, non calcifiée(s), non liquidiennes en échographie.	Anomalies probablement bénigne ≤ 2 % une surveillance à court terme est conseillée (de 6 mois et un à 1 an).
	MCS	- MCS rondes ou punctiformes régulières ou pulvérulentes, peu nombreuses, en petit amas rond isolé. -Petit(s) amas rond(s) ou ovale(s) de calcifications amorphes, peu nombreuses.	
ACR4	Masse	-Masse spéculée sans centre dense. -Opacité non liquidienne ronde ou ovale aux contours lobulés, ou masqués, ou ayant augmenté de volume. -Distorsion architecturale en dehors d'une cicatrice connue et stable.	Une anomalie indé-terminée ou suspecte qui nécessite une vérification histologique.
	MCS	-MCS punctiformes régulières nombreuses et/ou groupées en amas aux contours ni ronds, ni ovales. -MCS pulvérulentes groupées et nombreuses. -MCS irrégulières, polymorphes ou granulaires, peu nombreuses.	
ACR5	Masse	-Opacité mal circonscrite aux contours flous et irréguliers. -Opacité spiculée à centre dense.	Lésions très évocatrices de cancer.
	MCS	-MCS vermiculaires, arborescentes irrégulières, polymorphes ou granulaires, nombreuses et groupées. -Groupement de MCS quelle que soit leur morphologie, dont la topographie est galactophorique. -MCS associées à une anomalie architecturale ou à une opacité. -MCS groupées ayant augmenté en nombre ou MCS dont la morphologie et la distribution sont devenues plus suspectes.	

TABLEAU 1.6 : CLASSIFICATION BIRADS DE L'AMERICAN COLLEGE OF RADIOLOGY (ACR).

4.4. L'examen de confirmation diagnostic : l'examen histologique

Certaines anomalies indéterminées ou suspectes nécessitent des examens plus poussés. Pour faire face aux mammographies douteuses ou aux échographies ambiguës un examen

histologique est recommandé. Ce dernier est pratiqué sur les tissus prélevés par ponction cytologique ou biopsie. L'aspiration ou la ponction à l'aiguille utilise une fine aiguille pour aspirer du liquide ou des cellules provenant de la lésion suspecte.[16]

L'intervention est rapide, mais peut être inconfortable à cause de la sensibilité des seins. Cependant les microbiopsies sont indiquées pour explorer les anomalies de type masses ou opacités, visibles. Elles sont faites sous échographie à l'aide d'un pistolet automatique muni d'une aiguille.

Les macrobiopsies par mammotome (figure 1.13) sont indiquées pour explorer les foyers de MCS. Après une anesthésie locale et une petite incision de 3-4 mm, la sonde (aiguille creuse) est introduite dans la zone choisie (le tissu suspect), creusée par un petit couteau cylindrique rotatif, puis aspirer.[17]

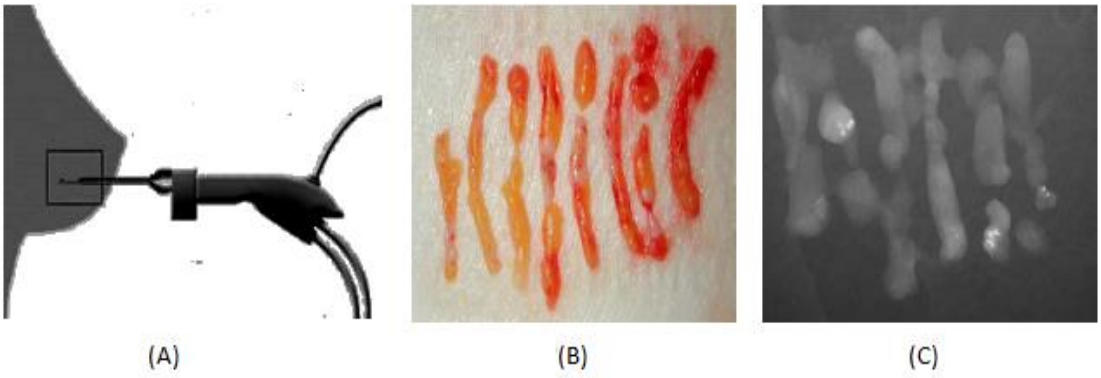


FIGURE 1.13 : LA MACROBIOPSIE PAR MAMMOTOME : A) LA MAMMOTOME B) LES PIÈCES PRÉLEVÉES C) L'IMAGE RADIOGRAPHIQUE.

4.5. Nouveauté en imagerie du sein

L'angiographie et la tomosynthèse mammaire sont deux nouvelles applications ou extension de mammographie numérique actuelle (figure 1.14.A). La différence entre le mammographe et l'appareil à tomosynthèse se résume dans le tube à rayons X qui reste statique dans le mammographe, tandis que dans la tomosynthèse un arc de cercle est effectué. Cette technique récente de mammographie 3D permet, en association avec la mammographie 2D conventionnelle d'augmenter le taux de détection des cancers (+ 30% en comparaison à la mammographie 2D seule) tout en réduisant le nombre de faux positifs[18]. Les premières expériences cliniques ont montrées que ce procédé est particulièrement efficace pour éliminer l'effet de superposition glandulaire, causant ainsi des diagnostics faussement négatifs (figure 1.14.A). L'avantage principal de la tomosynthèse est de limiter la superposition de tissus (figure 1.14.B). Ainsi, on peut espérer rendre les signes radiologiques associés à des lésions suspectes plus facilement discernables[14].

L'angiographie est une mammographie numérique plein champ avec injection d'un produit de contraste iodé (figure 1.14.C). Les expériences cliniques ont montrées qu'elle permet de visualiser l'angiogenèse des cancers du sein.

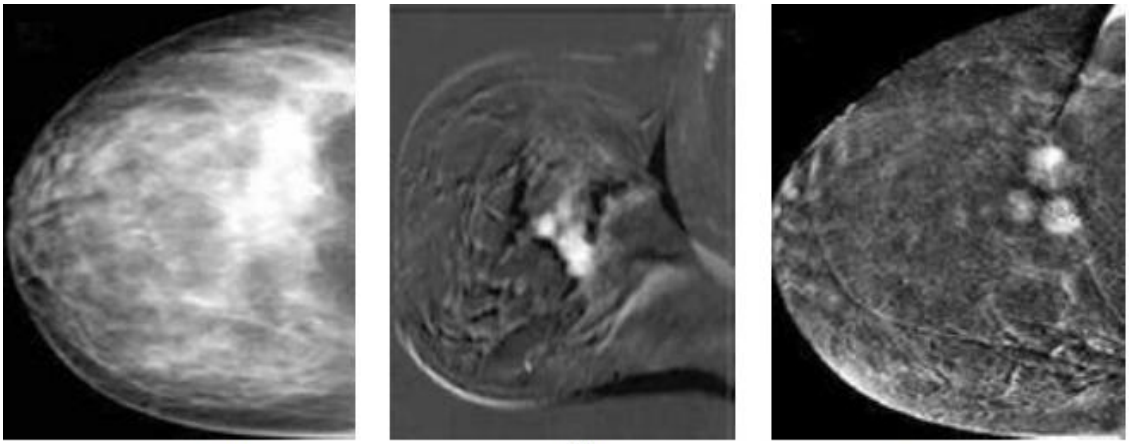


FIGURE 1.14 : (A) MAMMOGRAPHIE SEIN DROIT : OPACITÉ DE 3 CM ASSOCIÉE À DES MICROCALCIFICATIONS, (B) BILAN TOMOSYNTHÈSE CORRESPONDANT ET LE (C) BILAN PAR ANGIOMAMMOGRAPHIE.

Une autre exploration du parenchyme mammaire qui permet d'évaluer la dureté des tissus et d'augmenter la spécificité en échographie mammaire est l'échographie avec élastographie par ondes de cisaillement.

4.6. Système de détection assistée par ordinateur (DAOx)

En a vu que la mammographie de dépistage contribue largement à la détection précoce du cancer du sein. Cependant cette interprétation dépend fortement à l'expertise du radiologue (l'apprentissage reçu) et de sa connaissance du problème. Cependant, plusieurs études ont montré que les radiologues manquent entre 4% et 38% de détection de cancer[19].

Toutefois, deux grands types d'erreurs peuvent subvenir : le faux positif et le faux négatif. Le cas du faux positif est le cas qui présente des conséquences négatives, puisque les zones bénignes sont identifiées comme zones cancéreuses. Plus grave, le cas de faux négatifs qui mettent en danger la vie du patient et se produisent quand une anomalie existe et elle n'est pas détectée par le radiologue. C'est à partir de ce constat que l'idée d'utiliser les systèmes de Détection Assisté par Ordinateur (DAOx) est née afin d'offrir une seconde interprétation aux radiologues, qui ne font que résoudre un problème de vision[20].

Plusieurs études cliniques portant sur des logiciels commercialisés ont démontré la nécessité d'inclure un système de DAOx pour l'interprétation de mammographies[21].

En Algérie, ces systèmes sont actuellement appliqués en mammographie diagnostique, toutefois leur place en mammographie de dépistage n'est pas encore assurée : le futur dépistage précoce en Algérie se doit d'en bénéficier, souligne le Pr. *Abdelkrim BENDIB*, chef du service central d'Imagerie Médicale et de Radiologie du CHU Mustapha[22].

5. Les bases de mammographie

Il existe plusieurs bases d'images utilisées par les chercheurs en imagerie du sein dans le but de concevoir des systèmes de DAOx, dont certaines sont publiques et d'autres sont privées. Toutefois, toutes ne respectent pas les exigences que doivent satisfaire une base dédiée à la conception d'un système de DAOx. Dans un premier temps, nous allons évoquer

brièvement les exigences relatives à la création d'une base de mammographies, puis nous finirons en détaillant quelques-unes des bases disponibles tout en mettant l'accent sur leurs forces et leurs faiblesses. Dans des études menées entre 1996 et 1998, a énuméré certaines exigences qui doivent être respectées dans la création de base de mammographies numériques. Certaines d'entre elles sont abordées ci-dessous[20].

A. La sélection des cas : Toute base de mammographies doit contenir divers cas avec des images de sein normal, de même que celles de toutes les pathologies rencontrées en imagerie du sein et quelque soit la densité mammaire. Les images de sein normal mais contenant des structures qui peuvent prêter à confusion doivent être maintenues dans la base, car elles sont importantes pour rendre les classifieurs plus robustes. Les cas doivent être collectés par des radiologues ayant une réelle expérience en mammographie ; chaque cas doit contenir quatre images obtenues à l'aide des incidences standards (MLO medio latérale oblique et CC cranio-caudale), à moins que la femme n'ait subit auparavant une mastectomie du sein. Quant à l'acquisition des images, elle doit être réalisée dans les bonnes conditions, en veillant à ce que le positionnement de la femme et l'exposition aux rayons-X soient conformes à la procédure standard, de manière à éviter tout flou cinétique sur les images.

B. La vérité terrain : La preuve de la réalisation de la biopsie pour tous les cas évoqués doit être disponible. Par ailleurs, l'annotation doit inclure la vérité terrain concernant la cytologie/histologie de tous les cas, et les limites du contour de la lésion obtenue par la segmentation manuelle, réalisée par les radiologues experts en imagerie du sein.

C. Les informations assistées à chaque cas : Les informations relatives à l'âge, aux antécédents personnels et familiaux de la patiente peuvent être incluses dans la mesure où elles peuvent s'avérer utiles pour améliorer les systèmes de DAOx. Quant à la classification de la densité du sein selon l'ACR et le BI-RADS, elles sont obligatoires.

D. L'organisme de la base et sa distribution : Il n'y a pas de format de fichier spécifique, toutefois, la plupart des images médicales sont sauvegardées au format DICOM (*Digital Imaging and Communications in Medicine).* Une partition de ces images en sous base d'apprentissage et de test est également conseillée en vue de faciliter la comparaison des différents systèmes DAOx. Enfin, les bases doivent être accessibles à tous et de préférence via internet, sans oublier qu'une maintenance permanente doit être obligatoirement assurée par les propriétaires de la base[20].

5.1. Mammographie Image Analysis Society Digital Mammogram Database (MIAS)

Mammographic Image Analysis Society (MIAS) [Suckling et al., 1994] est la plus ancienne base d'images publiquement disponible ; elle a été conçue au Royaume-Uni en 1994 et est encore largement utilisée dans l'état-de-l'art. La MIAS contient 161 cas pour un total de 322 images numérisées obtenues en incidence MLO, sur lesquelles sont présentes toutes les pathologies à savoir les lésions bénignes et malignes, mais également les images normales. Cette base dispose d'un nombre important de masses spiculées, ainsi que d'une information relative à la densité du sein ; toutefois, la classification des masses réalisée par les radiologues ayant évalué cette base ne respecte pas les standards de l'ACR. Cependant, avec

l'augmentation de l'usage de la classification BI-RADS de l'ACR, certains auteurs se sont essayés à une classification des masses de la MIAS afin qu'elle puisse correspondre aux standards en vigueur en imagerie du sein. Un autre défaut de la MIAS est l'annotation de ces images, qui consiste à indiquer le centre et le rayon de la région d'intérêt, c'est-à-dire la région où se situe la pathologie. Ce genre d'annotation est considérée comme insuffisant pour certaines études comme la segmentation des masses, où l'on souhaite que toutes les lésions circonscrites ou spiculées soient manuellement segmentées afin de faire une comparaison très précise avec les méthodes automatiques. Pour finir, l'autre inconvénient de la MIAS, c'est la résolution à laquelle les images sont numérisées et qui fait que cette base ne convient pas pour des expériences liées à la détection des microcalcification[20].

5.2. Digital Database for Screening Mammography (DDSM)

Constituée en 1999 aux Etats-Unis (USA), la *Digital Database for Screening Mammography* (DDSM) est la plus grande base publique de mammographies. Elle contient 2620 cas incluant deux images (incidences MLO et CC) de chaque sein, pour un total de 10480 images où se retrouvent toutes les pathologies, des images normales à celles contenant des masses bénignes et malignes. Pour chacune de ces images, les informations relatives à l'âge de la patiente, la densité du sein et une classification respectant le BI-RADS sont fournies.

À la différence de la MIAS, les annotations concernant les pathologies sont faites manuellement par les radiologues, qui encerclent de façon un peu grossière la région suspecte. Toutefois, des travaux de recherche ont souligné que la DDSM n'était pas adéquate pour la validation des algorithmes de segmentation, en raison du manque de précision de la segmentation manuelle[6][20].

5.3. Banco Web LAPIMO Plus

Plus récente que les deux précédentes bases, la BancoWeb LAPIMO est une base créée en 2010 au Brésil. Elle dispose de 320 cas avec 1473 images (MLO et CC) contenant toutes les pathologies. Les antécédents de la patiente, de même qu'une classification BI-RADS sont fournis. En ce qui concerne les masses, seules certaines images sont annotées sous forme de ROI ; mais toutes les pathologies disposent d'une description textuelle[20].

5.4. INbreast

L'ensemble de données INbreast contient des annotations de vérité terrain au niveau des pixels « *pixel-level ground truth annotations* » et des preuves histologiques des cancers. Un total de 115 cas ont été collectés, dont 90 ont deux images (MLO et CC) de chaque sein et les 25 autres cas proviennent de femmes ayant subi une mastectomie et deux vues d'un seul sein ont été incluses. Cela fait un total de 410 images. Huit des 91 cas avec 2 images par sein ont également des images acquises à des moments différents (suivi). La base de données comprend des exemples de mammographies normales, mammographies avec masses, mammographies avec calcifications, distorsions architecturales, asymétries et images avec résultats multiples[23].

6. Conclusion

Dans ce chapitre ont été abordées des notions de généralité en rapport avec le cancer du sein. Après avoir présenté de façon concise l'anatomie du sein, les types des tumeurs et leurs symptômes en passant par le dépistage du cancer du sein ainsi que quelques bases de mammographie. L'objectif qui sous-tend la présentation de ces éléments, c'est de mettre en lumière les difficultés rencontrées par les radiologues dans l'analyse des mammographies. Notre travail va s'orienter vers la détection des tumeurs dans le sein en utilisant le DL qui est le sujet de notre suivant chapitre.

Chapitre 2

Deep Learning

1. Introduction

Depuis son point de départ dans les années 1950 avec les travaux du mathématicien Alan TURING, l'intelligence artificielle (IA), discipline mathématique et technique destinée à reproduire l'intelligence humaine, s'est développée par cycles successifs parallèlement à la croissance de la puissance de calcul informatique disponible.

Dans les années 1980, le concept d'apprentissage automatique « *Machine Learning* » se développe permettant à une machine de déduire une « règle à suivre » uniquement à partir de l'analyse de données. Cette période voit apparaître la majorité des algorithmes « apprenants » utilisés aujourd'hui (réseau de neurones, apprentissage par renforcement, machines à vecteurs de support, etc…). Ces avancées se concrétisent notamment par le succès de l'ordinateur Deep Blue face au grand maître d'échecs, Gary KASPAROV en 1992.

À partir des années 2000, un nouveau cycle d'innovations en IA se met en place avec le formidable développement d'internet et des très grandes infrastructures de calcul, offrant un accès à un volume de données encore jamais atteint dans l'histoire humaine. Avec cette capacité nouvelle, le développement des techniques d'apprentissage profond « *Deep Learning* » permet aux machines de commencer à surpasser les performances des meilleurs experts humains dans des domaines comme la reconnaissance visuelle, l'analyse documentaire ou la traduction.

La technologie d'apprentissage en profondeur est une famille de méthodes de calcul qui permettent à un algorithme de se programmer lui-même en apprenant à partir d'un large éventail d'exemples qui démontrent le comportement souhaité. Les applications de la technologie d'apprentissage en profondeur à l'imagerie du cancer peuvent aider les pathologistes à détecter et à classer le cancer aux premiers stades de son développement pour permettre aux patients d'avoir des traitements appropriés qui peuvent augmenter la survie ou la récupération des patients. Les analyses statistiques et autres approches analytiques, basées sur les données de ScienceDirect (une source de recherche scientifique), suggèrent que, depuis la fin des années 1990, la forte augmentation des études sur les technologies d'apprentissage en profondeur dans l'imagerie du cancer semble être due à des taux de mortalité élevés de certains types de cancer (par exemple, le poumon et le sein) afin de résoudre les problèmes consécutifs d'une détection et d'une caractérisation plus précises des types de cancer pour appliquer des thérapies anticancéreuses efficaces[24].

2. L'apprentissage automatique « Machine Learning »

L'apprentissage automatique est un processus qui crée un modèle à partir d'un jeu de données d'entraînement, dans le but, par exemple, de classifier, mesurer ou prendre des décisions sur de nouvelles données (figure 2.1)[25].

Les algorithmes d'apprentissage peuvent se catégoriser selon le mode d'apprentissage qu'ils emploient :

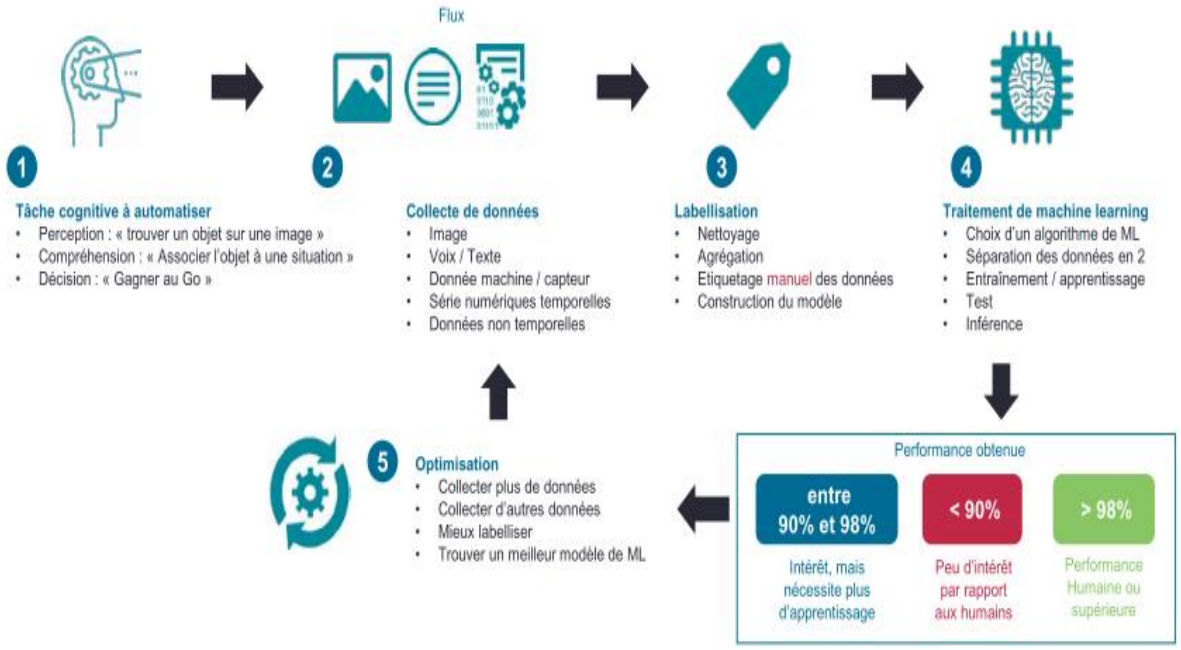


FIGURE 2.1 : PROCESSUS DE MACHINE LEARNING.

2.1. L'apprentissage supervisé

Consiste à analyser un jeu de données empiriques, appelée base de données d'entraînement. Ces données sont étiquetées ou labellisées, c'est à dire associées à une description ou information, afin de pouvoir caractériser un nouveau jeu de données inconnu. L'algorithme apprend ainsi à partir de milliers ou de millions d'exemples étiquetés : il cherche la relation qui permet de relier les données aux labels[26].

2.2. L'apprentissage non supervisé

Il s'agit, pour un logiciel, de trouver des structures sous-jacentes à partir de données non étiquetées. Puisque les données ne sont pas étiquetées, l'apprentissage non supervisé est beaucoup plus complexe puisqu'ici le système va devoir détecter les similarités dans les données qu'il reçoit et les organiser en fonction de ces dernières. Cette façon de travailler présente un avantage indéniable en ce sens que la phase de catégorisation de l'apprentissage supervisé est un processus gourmand en ressources humaines. Son élimination, ou tout du moins sa réduction retire un frein à l'implémentation de la technologie[26].

Apprentissage supervisé	Apprentissage non supervisé
Données d'entrée sont étiquetées.	Données d'entrée sont non étiquetées.
Utilise le jeu de données d'apprentissage.	Utilise tout le jeu de données en entrée.
Utilisé pour la prédiction.	Utilisé pour l'analyse.
Classification et régression.	Regroupement, estimation de la densité, et réduction de la dimensionnalité.

TABLEAU 2.1 : DIFFÉRENCES ENTRE L'APPRENTISSAGE SUPERVISÉ ET NON SUPERVISÉ.

2.3. Apprentissage semi-supervisé

Les systèmes à apprentissage semi-supervisé combinent à la fois le principe d'apprentissage supervisé et celui non supervisé. L'intérêt de cette combinaison provient du fait que l'étiquetage de données nécessite l'intervention d'un utilisateur humain. Lorsque les jeux de données deviennent très grands, cette opération peut s'avérer fastidieuse. Dans ce cas, l'apprentissage semi-supervisé, qui ne nécessite que quelques étiquettes, revêt un intérêt pratique évident[6].

2.4. Apprentissage par renforcement

L'apprentissage par renforcement consiste, pour un agent autonome, à apprendre les actions à prendre, à partir d'expériences, de façon à optimiser une récompense quantitative au cours du temps. L'agent est plongé au sein d'un environnement, et prend ses décisions en fonction de son état courant. En retour, l'environnement procure à l'agent une récompense, qui peut être positive ou négative. L'agent cherche, au travers d'expériences itérées, un comportement décisionnel (appelé stratégie ou politique, et qui est une fonction associant à l'état courant l'action à exécuter) optimal, en ce sens qu'il maximise la somme des récompenses au cours du temps.

Dans la littérature, il existe une grande diversité d'approches en machine learning. Le choix d'une méthode dépend du problème adressé, notamment du type de modèle d'inférence à déterminer. On distingue notamment les approches selon la complexité du modèle : spécifique ou généraliste. Un autre critère important est la taille de la base d'apprentissage. Trop petite, elle ne permet pas par exemple d'envisager les méthodes de Deep Learning[6].

3. L'apprentissage profond « *Deep Learning* »

3.1. Principes

D'après les travaux de *Deng L* et *al*[27], le « *Deep Learning* » est une classe de techniques d'apprentissage automatique appartient au domaine de « *Machine Learning* » (voir figure 3.1) dans lesquelles de multiples couches de traitement de calcul itératif dans des architectures hiérarchisées supervisées sont exploitées pour les algorithmes d'apprentissage non supervisé pour des tâches d'analyse et de classification. L'apprentissage profond consiste essentiellement à calculer des caractéristiques hiérarchiques des paramètres des réseaux de neurones artificiels pour les représentations vectorielles des données d'observation ou d'entrées. La famille des méthodes d'apprentissage en profondeur s'enrichit de plus en plus, englobant celles des réseaux de neurones, des modèles probabilistes hiérarchiques, ainsi que de nombreux algorithmes d'apprentissage des fonctionnalités supervisés et non supervisés.

Comme tous les réseaux de neurones l'apprentissage profond a besoin à des paramètres d'entrée *X* pour des observations de sorties et résultats *Y*. L'apprentissage automatique est une méthode d'apprentissage statistique dans laquelle chaque instance d'un

jeu de données est décrite par un ensemble de caractéristiques ou d'attributs. En revanche, le terme « *Deep Learning* » est une méthode d'apprentissage statistique qui extrait des caractéristiques ou des attributs à partir de données brutes. Il utilise pour cela des réseaux de neurones comportant de nombreuses couches cachées, des données volumineuses et de puissantes ressources de calcul. Les termes semblent quelque peu interchangeables, cependant, avec les méthodes d'apprentissage profond, l'algorithme construit automatiquement des représentations des données. En revanche, les représentations de données sont codées en dur comme un ensemble de fonctionnalités dans des algorithmes d'apprentissage automatique, nécessitant des processus supplémentaires tels que la sélection et l'extraction des paramètres.

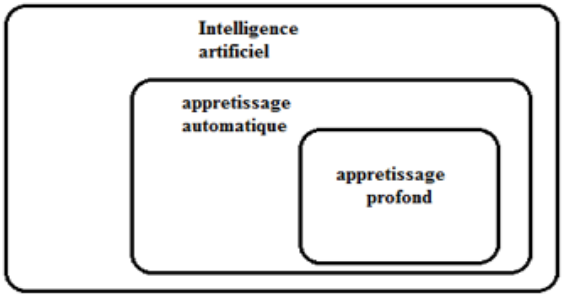


FIGURE 2.2 : LE CONTEXTE DE DEEP LEARNING.

3.2. Réseau de Neurones et l'architecture « Deep Learning »

Un réseau neuronal artificiel tire sa structure et son concept de travail du système nerveux humain. Le premier réseau neuronal artificiel appelé perceptron est basé sur un neurone biologique. Il a une couche d'entrée et une couche de sortie, la couche d'entrée est directement connectée à la couche de sortie. Il peut être utilisé pour des tâches très simples comme la classification de modèles séparables linéairement. Afin de résoudre les problèmes ayant des couches plus complexes de neurones entre les couches d'entrée et de sortie ont été introduites appelées couche cachée. Ainsi, le travail de base d'un réseau neuronal consiste à prendre l'entrée, à la traiter et à la transmettre en sortie à la couche suivante. La figure 2.3 donne l'architecture d'un réseau neuronal. Chaque neurone du réseau prend d'abord l'entrée, effectue la sommation de l'entrée pondérée, applique la fonction d'activation, calcule la sortie et la transmet à la couche suivante[28].

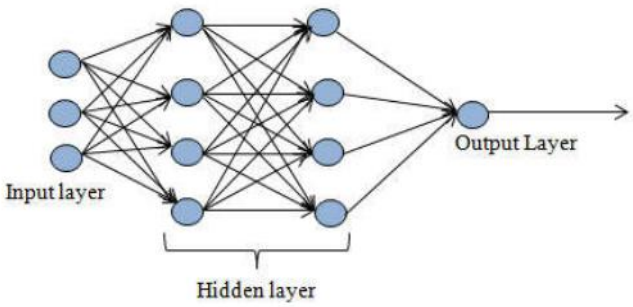


FIGURE 2.3 : ARCHITECTURE DE RÉSEAU DE NEURONES.

3.3. Deep Learning vs Machine Learning?

Le moyen le plus simple de comprendre la différence entre l'apprentissage automatique et l'apprentissage profond est de savoir que l'apprentissage profond est un apprentissage automatique. Plus précisément, l'apprentissage en profondeur est considéré comme une évolution de l'apprentissage automatique. Il utilise un réseau neuronal programmable qui permet aux machines de prendre des décisions précises sans l'aide des humains.

En fait, l'apprentissage profond est techniquement un apprentissage automatique et fonctionne de manière similaire (d'où la raison pour laquelle les termes sont parfois mal échangés). Cependant, ses capacités sont différentes.

Bien que les modèles d'apprentissage automatique de base s'améliorent progressivement quelle que soit leur fonction, ils ont encore besoin de conseils. Si un algorithme d'intelligence artificielle renvoie une prédiction inexacte, un ingénieur doit intervenir et effectuer des ajustements. Avec un modèle d'apprentissage en profondeur, un algorithme peut déterminer par lui-même si une prédiction est précise ou non via son propre réseau de neurones.

Un modèle d'apprentissage en profondeur est capable d'apprendre grâce à sa propre méthode de calcul, une technique qui donne l'impression d'avoir son propre cerveau.

Donc l'apprentissage automatique utilise des algorithmes pour analyser les données, apprendre de ces données et prendre des décisions éclairées en fonction de ce qu'il a appris tandis que le D.L structure des algorithmes en couches pour créer un « réseau de neurones artificiels » qui peut apprendre et prendre des décisions intelligentes par lui-même[29].

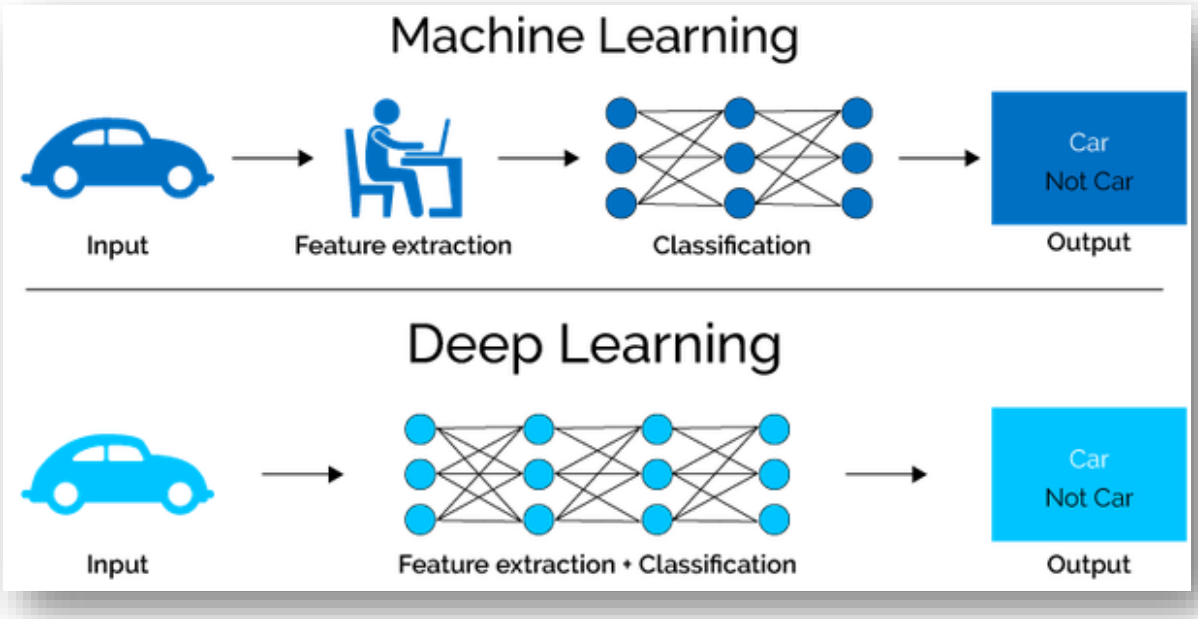


FIGURE 2.4 : ML VERS DL.

Facteurs	DL	ML
Exigence de données	Nécessite de grandes données	Peut s'entraîner sur des données moindres
Exactitude	Fournit une grande précision	Donne moins de précision
Temps d'entraînement	Prend plus de temps à s'entraîner.	Prend moins de temps pour s'entraîner.
Dépendance matérielle	Nécessite un GPU pour s'entraîner correctement.	S'entraîne sur CPU.
Réglage hyperparamètre	Peut être réglé de différentes manières.	Capacités de réglage limitées.

TABLEAU 2.2 : ML VERS DL.

3.4. Domaines d'application de l'apprentissage profond

Ces techniques se développent dans le domaine de l'informatique appliquée à la robotique, à la bio-informatique, la reconnaissance ou comparaison de formes, la sécurité, la santé, la pédagogie assistée par l'informatique, et plus généralement à l'intelligence artificielle. L'apprentissage profond peut par exemple permettre à un ordinateur de mieux reconnaitre des objets hautement déformables et/ou analyser par exemple les émotions révélées par un visage photographié ou filmé, ou analyser les mouvements et position des doigts d'une main, ce qui peut être utile pour traduire le langage des signes, améliorer le positionnement automatique d'une caméra, etc… Elles sont utilisées pour certaines formes d'aide au diagnostic médical (ex. : reconnaissance automatique d'un cancer en imagerie médicale), ou de prospective ou de prédiction (ex. : prédiction des propriétés d'un sol filmé par un robot)[30].

3.4.1. Entreprises adoptant le « *Deep Learning* »

La révolution de l'apprentissage en profondeur est dérivée de trois technologies principales, la première est la disponibilité des mégadonnées (big data), la seconde est les algorithmes actuels d'apprentissage en profondeur basés sur le cerveau humain et la troisième la puissance de traitement rapide. Le potentiel de l'apprentissage en profondeur est assez efficace, tout comme le coût et les efforts initiaux. De nombreuses entreprises comme IBM Watson, Google DeepMind, etc. s'efforcent de trouver la solution optimale pour les gros problèmes d'imagerie médicale. Des entreprises comme Hitachi, Philips, Siemen, etc. ont déjà commencé à investir dans ce domaine et les laboratoires de recherche d'IBM, Google travaillent au développement d'applications d'imagerie efficaces. IBM Watson et DeepMind collaborent avec différents prestataires de soins de santé afin d'apprendre comment

l'apprentissage en profondeur peut être appliqué à l'imagerie médicale et comment il peut résoudre les défis de l'imagerie médicale[31].

3.5. Les techniques de « *Deep Learning* »

En général, un réseau de neurones artificiels est une technique conçue pour simuler l'activité du cerveau humain et ces neurones biologiques. En particulier, il est utilisé pour les tâches de reconnaissances automatiques des formes et des systèmes informatiques intelligents. L'apprentissage profond est une approche spécifique utilisée pour la construction d'une topologie des réseaux de neurones, pour le traitement des grandes masses des données sur des tâches d'apprentissage automatique, chaque réseaux de neurones fournit des bons résultats pour des types de données spécifiques tels que (vidéo, document, audio …). L'avantage de ses techniques de « *Deep Learning* » se réside dans l'identification manuelle des caractéristiques des données d'une façon automatique dans la partie des couches cachées de processus d'apprentissage. Il facilite ainsi et accélère le temps d'apprentissage du réseau de neurones et permets d'obtenir de meilleurs résultats. Parmi les techniques de « *Deep Learning* » utilisé nous trouvons les réseaux de neurones convolutifs et récurrent défini comme suite :

3.5.1. Les réseaux de neurones convolutifs (CNN) les plus connus

D'après les travaux de *M.Zeiler et R.Fergus*[32] , les réseaux de neurones convolutifs « *CNN* » sont des types spécifiques de réseaux de neurones artificiels qui utilisent des perceptrons et un algorithme d'unité d'apprentissage automatique pour l'apprentissage supervisé, permettant d'analyser des données. Les CNN s'appliquent beaucoup plus au traitement d'images, aussi au traitement du langage naturel « NLP » et à d'autres types de tâches cognitives. Comme d'autres types de réseaux de neurones artificiels, un réseau de neurones convolutionnel comporte une couche d'entrée, une couche de sortie et une ou plusieurs couches cachées (voir figure 2.5). Certaines de ces couches sont convolutionnelles et utilisent un modèle mathématique et algébrique pour transmettre les résultats aux couches successives.

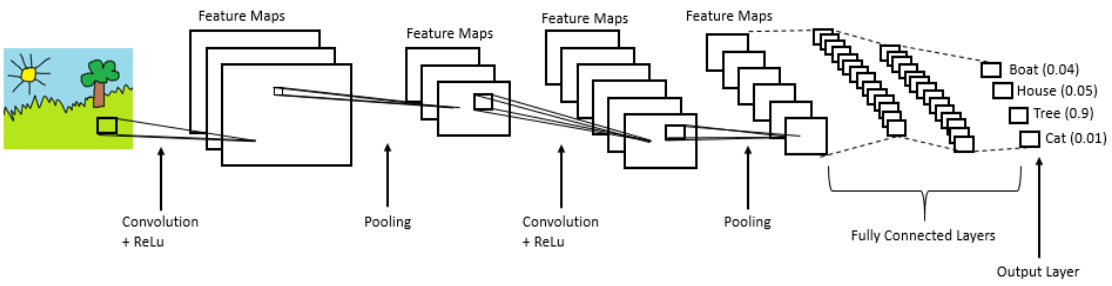


FIGURE 2.5 : L'ARCHITECTURE DES RÉSEAUX DE NEURONES CONVOLUTIFS.

Les réseaux « CNN » sont des exemples typiques de la technique d'apprentissage profond, où un modèle plus sophistiqué accélère l'évolution de l'intelligence artificielle en proposant des systèmes simulant différents types d'activités biologiques du cerveau humain. Généralement les réseaux « CNN » sont composés de quatre principales couches :

- La couche de convolution « *CONV* » qui traite les données d'entrées ou les résultats des couches intermédiaires ou cachées.
- La couche de *Pooling* « *POOL* », qui permet de compresser l'information en réduisant la taille de l'image intermédiaire.
- La couche d'activation, c'est des fonctions de rectification linéaire souvent appelée par « *ReLU* » qui permet la mise à jour des paramètres de la couche actuelle ainsi que les couches précédentes.
- La couche entièrement connectée « *Fully Connected* », qui est une couche de type perceptron.

Pour bien comprendre les réseaux de neurones convolutifs, il est important de connaître les bases des réseaux de neurones. Les principaux éléments à retenir sont :

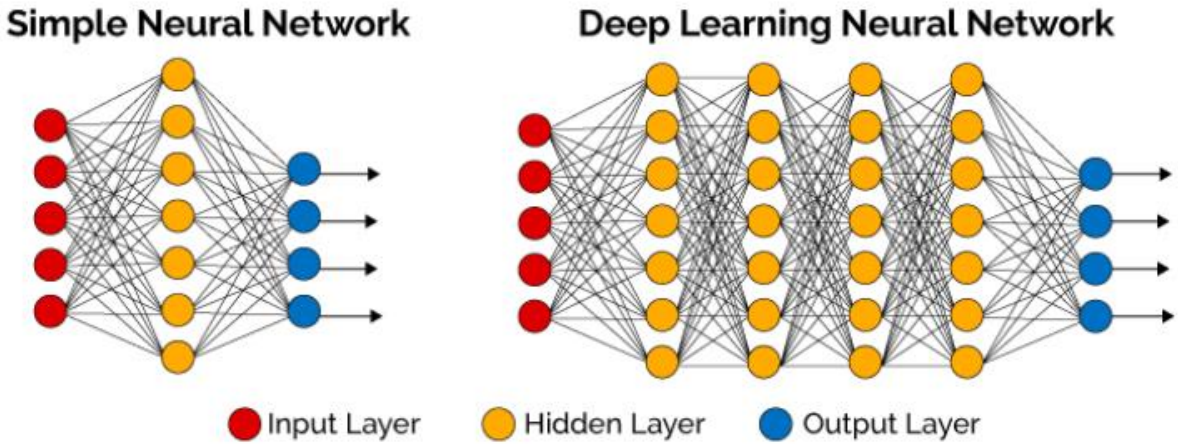


FIGURE 2.6 : RÉSEAU DE NEURONES DU DEEP LEARNING.

- Un réseau de neurones est un système composé de neurones, généralement répartis en plusieurs couches connectées entre elles [33];
- Un tel système s'utilise pour résoudre divers problèmes statistiques, le réseau calcule à partir de l'entrée un score (ou probabilité) pour chaque classe. La classe attribuée à l'objet en entrée correspond à celle de score le plus élevé [33];
- Chaque couche reçoit en entrée des données et les renvoie transformées. Pour cela, elle calcule une combinaison linéaire puis applique éventuellement une fonction non-linéaire, appelée fonction d'activation. Les coefficients de la combinaison linéaire définissent les paramètres (ou poids) de la couche [33];
- Un réseau de neurones est construit en empilant les couches : la sortie d'une couche correspond à l'entrée de la suivante [33];

- Cet empilement de couches définit la sortie finale du réseau comme le résultat d'une fonction différentiable de l'entrée [33];
- La dernière couche calcule les probabilités finales en utilisant, en général, pour fonction d'activation la fonction logistique (classification binaire) ou la fonction softmax (classification multi-classes) [33];
- La dernière couche calcule les probabilités finales en utilisant, en général, pour fonction d'activation la fonction logistique (classification binaire) ou la fonction softmax (classification multi-classes) [33];
- Les valeurs des poids des couches sont apprises par rétropropagation du gradient : on calcule progressivement (pour chaque couche, en partant de la fin du réseau) les paramètres qui minimisent la fonction de perte régularisée. L'optimisation se fait avec une descente du gradient stochastique[33].

Les réseaux de neurones convolutifs (CNN) ont connu un énorme succès dans de nombreux domaines de la vision par ordinateur et de l'analyse d'images médicales. Cependant, il existe toujours un immense potentiel d'amélioration des performances mammographie des systèmes de détection du cancer du sein en intégrant toutes les informations que le radiologue utilise.

Il existe plusieurs architectures dans le domaine des réseaux convolutifs qui ont un nom. Les plus courants sont :

3.5.1.1 AlexNet

Le premier travail qui a popularisé les réseaux convolutifs en vision par ordinateur a été AlexNet, développé par Alex Krizhevsky, Ilya Sutskever et Geoff Hinton. L'AlexNet a été soumis au défi ImageNet ILSVRC en 2012 et a largement surpassé le deuxième finaliste (top 5 erreur de 16% par rapport au finaliste avec 26% d'erreur). Le réseau avait une architecture très similaire à LeNet, mais était plus profond, plus grand et comportait des couches convolutionnelles empilées les unes sur les autres (auparavant, il était courant de n'avoir qu'une seule couche CONV toujours immédiatement suivie d'une couche POOL).

3.5.1.2 GoogLeNet

Le gagnant de l'ILSVRC 2014 « ImageNet Large-Scale Visual Recognition Challenge » était un réseau convolutif de Szegedy et al. de Google. Sa principale contribution a été le développement d'un Inception Module qui a considérablement réduit le nombre de paramètres dans le réseau (4M, comparé à AlexNet avec 60M). De plus, utilisant la mise en commun moyenne au lieu des couches entièrement connectées en haut du ConvNet, éliminant une grande quantité de paramètres qui ne semblent pas avoir beaucoup d'importance. Il existe également plusieurs versions de suivi de GoogLeNet, la plus récente Inception-v4[34].

3.5.1.3 VGGNet

VGGNet est inventé par VGG (Visual Geometry Group) de l'Université d'Oxford, bien qu'il soit le finaliste de l'ILSVRC 2014. Sa principale contribution a été de montrer que la profondeur du réseau est un élément critique pour de bonnes performances. Leur meilleur réseau final contient 16 couches CONV / FC et, de manière attrayante, présente une architecture extrêmement homogène qui n'effectue que des convolutions 3x3 et une mise en commun 2x2 du début à la fin. Un inconvénient du VGGNet est qu'il est plus coûteux à évaluer et utilise beaucoup plus de mémoire et de paramètres (140M). La plupart de ces paramètres se trouvent dans la première couche entièrement connectée, et il a été constaté depuis que ces couches FC peuvent être supprimées sans dégradation des performances, ce qui réduit considérablement le nombre de paramètres nécessaires[34].

3.5.1.4 ResNet

Residual Network développé par Kaiming He et al. a été le gagnant de l'ILSVRC 2015. Il comporte des connexions de saut « *skip connections* » spéciales et une utilisation intensive de la normalisation par lots « *batch normalization* ». L'architecture manque également de couches entièrement connectées à la fin du réseau[34].

3.5.1.5 Inception-V3, V4 et Inception-ResNet

Inception-V3, V4 et Inception-ResNet sont des versions améliorées d'Inception-V1 et V2 (Szegedy et al.2015, 2016a, b). L'idée d'Inception-V3 était de réduire le coût de calcul des réseaux profonds sans affecter la généralisation. À cette fin, Szegedy et al. remplacé les filtres de grande taille (5x5 et 7x7) par des filtres petits et asymétriques (1x7 et 1x5) et utilisé la convolution 1x1 comme goulot d'étranglement avant les grands filtres (Szegedy et al.2016b). Le placement simultané de la convolution 1x1 avec un filtre de grande taille fait que l'opération de convolution traditionnelle ressemble davantage à une « *cross-channel correlation* ».

Dans Inception-V3, une opération convolutive 1x1 a été utilisée, qui mappe les données d'entrée dans 3 ou 4 espaces séparés qui sont plus petits que l'espace d'entrée d'origine, puis mappe toutes les corrélations dans ces espaces 3D plus petits, via des convolutions régulières (3x3 ou 5x5). Dans Inception-ResNet, Szegedy et al. combiné la puissance de l'apprentissage résiduel et du blocage initial (He et al.2015a; Szegedy et al.2016a). Ce faisant, la concaténation de filtre a été remplacée par la connexion résiduelle.

De plus, Szegedy et al. a montré expérimentalement qu'Inception-V4 avec des connexions résiduelles (Inception-ResNet) a le même pouvoir de généralisation que InceptionV4 ordinaire mais avec une profondeur et une largeur accrues. Cependant, ils ont observé qu'Inception-ResNet converge plus rapidement qu'Inception-V4, qui montre que la formation avec des connexions résiduelles accélère considérablement la formation des réseaux Inception[34].

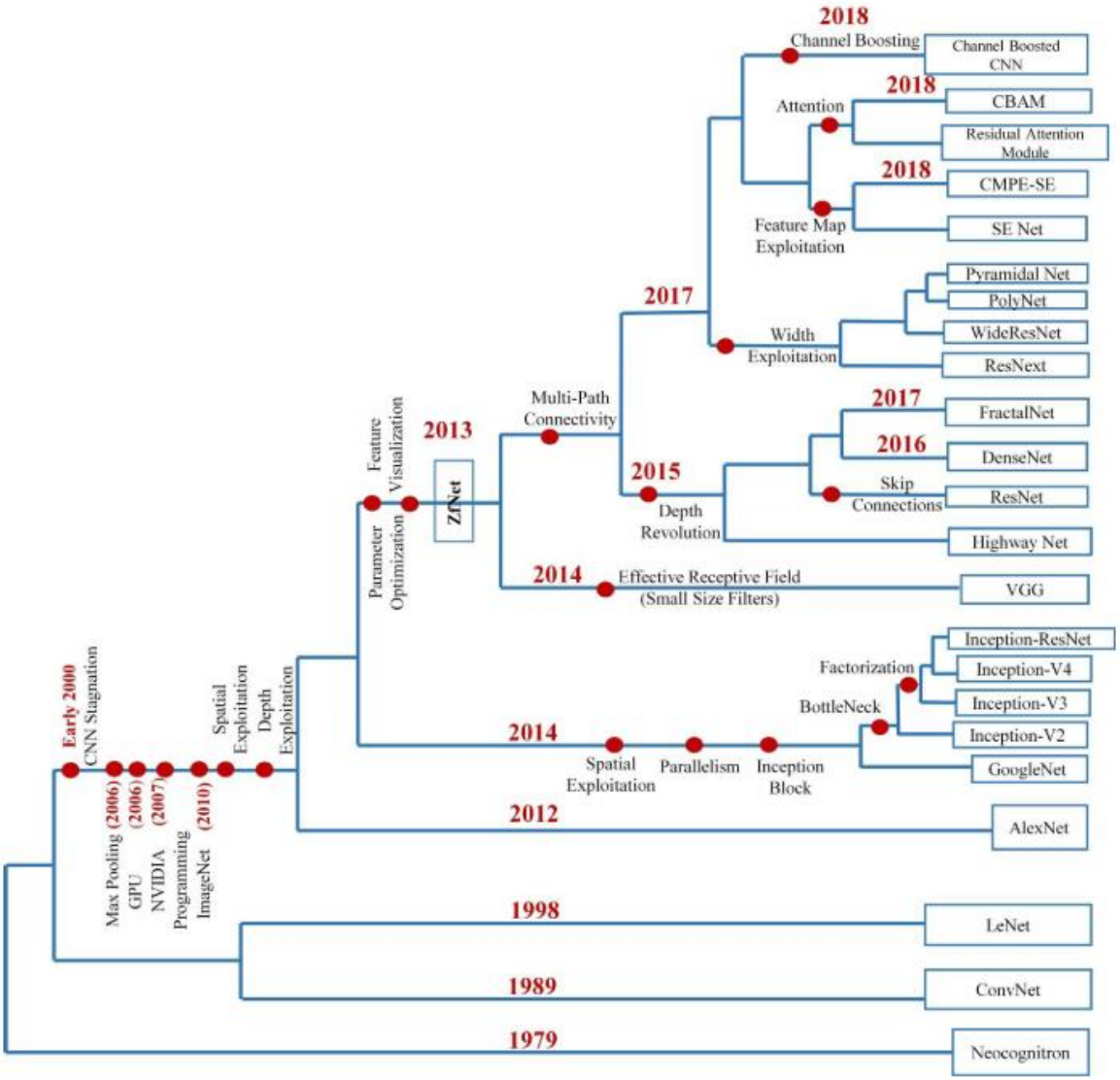


FIGURE 2.7 : HISTOIRE ÉVOLUTIVE DES CNN PROFONDS MONTRANT LES INNOVATIONS ARCHITECTURALES DE CONVNET AUX ARCHITECTURES À CE JOUR[34].

3.5.2. Les réseaux de neurones récurrent « *RNN* »

Les réseaux neurones récurrents *« RNN »* sont des types de réseaux de neurones artificiels avancés « ANN ». Ils intègrent les notions de continuités temporelles dans les entrées de ces couches de réseaux de neurones ces réseaux de neurones récurrents sont adaptés pour des données d'entrée de taille variable (voir figure 2.8). Ils conviennent en particulier pour l'analyse de séries temporelles. Ils sont utilisés en reconnaissance automatique de la parole ou de l'écriture manuscrite, plus en général en reconnaissance de formes ou encore en traduction automatique.

De point de vue dans l'application pratique, les RNN ont été un domaine de concentration actif pour de nombreux professionnels, notamment pour le traitement d'images[35].

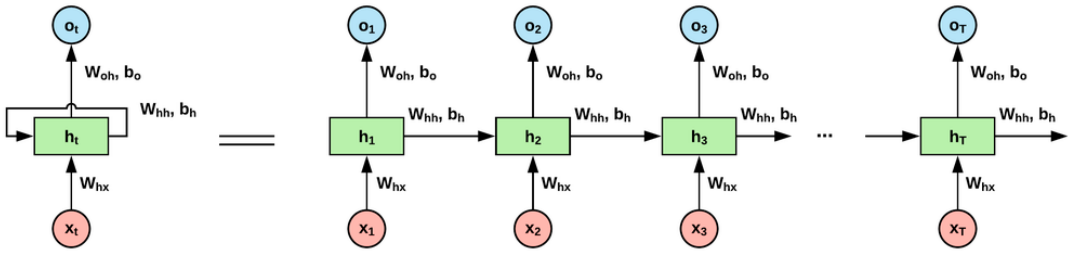


FIGURE 2.8 : L'ARCHITECTURE DES RÉSEAUX DE NEURONES RÉCURRENTS.

De point de vue technique. Les réseaux de neurones récurrents utilisent les mêmes techniques d'apprentissage utilisées dans les réseaux de neurones classiques tels que l'algorithme de rétropropagation du gradient. Cependant, ces réseaux sont limités par le problème de disparition du gradient pour le traitement de grandes séquences pour mémoriser les évènements passés. Nous trouvons dans les littératures des architectures particulières des « *RNN* » qui répondent à ce problème tel que les réseaux « *LSTM* » « *Long Short-Term Memory* ». Ces techniques permettent d'étudier les comportements des réseaux de neurones récurrents avec la théorie des bifurcations, mais la complexité de cette architecture augmente très rapidement avec le nombre de neurones.

3.6. Les techniques avancées de « *Deep Learning* »

Récemment, des techniques avancées d'apprentissage en profondeur, en particulier le region-based CNN (RCNN), ont obtenu des succès remarquables dans une diversité de tâches d'apprentissage automatique, de statistiques et de vision par ordinateur, par exemple, la détection d'objets, la catégorisation d'images, la segmentation d'images, etc. est parce que CNN identifie uniquement la classe de l'objet, pas l'emplacement de l'objet dans une image. Surtout lorsque plusieurs objets sont dans l'image, CNN ne peut pas fonctionner correctement en raison d'interférences. R-CNN identifie la région d'intérêt (RoI) en utilisant la recherche sélective et en utilisant un CNN, il les classe comme premier plan et arrière-plan. De nos jours, R-CNN est le meilleur détecteur basé sur CNN car il peut identifier l'emplacement de l'objet et fonctionne également bien pour une image contenant plusieurs objets.

En raison de la formation coûteuse et de la détection lente des objets, R-CNN est modifié et fast R-CNN, faster R-CNN, region-based fully convolutional networks (R-FCN), single shot detector (SSD), you only look once (YOLO), grid-CNN (G-CNN), spatial pyramid pooling network (SPP-Net), R-CNN minus R, propose network (Pro-Net), mask R-CNN, etc., évoluent un après un autre. Ces nouveaux algorithmes rendent la détection des objets et sa classification plus précises et plus rapides. Pour appliquer ces techniques dans des applications du monde réel telles que la détection de visage, la détection de personnes, la détection de panneaux de signalisation et la surveillance du trafic, etc., le problème est de choisir le modèle ou la technique adapté à leur application. Un compromis entre vitesse, précision et simplicité de ces techniques est essentiel pour obtenir le résultat souhaité[36].

Ces techniques aident également au diagnostic médical tel que la détection de tumeurs, la segmentation du noyau, et bien plus encore. Pour la détection d'objets, de nombreux

ensembles de données tels que les ensembles de données PASCAL VOC, les ensembles de données Microsoft COCO, les ensembles de données ImageNet, etc. sont disponibles[36].

3.6.1. R-CNN

L'architecture R-CNN comprend trois étapes comme le montre la figure 2.9. La première étape consiste à prendre l'image d'entrée et à rechercher des objets probables en utilisant un algorithme appelé recherche sélective qui génère environ 2000 propositions de région. Dans la deuxième étape, en haut de chaque proposition de région, un CNN est exécuté. La troisième étape consiste à prendre chaque sortie CNN et à fournir cette sortie dans (i) une machine à vecteurs de support (SVM) et (ii) un régresseur linéaire. SVM classe la région de l'objet dans une image et le régresseur linéaire resserre la boîte englobante « bounding box » de l'objet détecté.

R-CNN présente quelques inconvénients notables. (i) Pour classer ~ 2000 propositions de régions par image, il faut énormément de temps pour former le réseau, (ii) La détection d'objets en temps réel n'est pas possible car elle prend environ 47 s pour chaque image de test, (iii) R-CNN devrait entraîner 3 modèles différents séparément (CNN qui génère une fonction d'image, un classificateur qui prédise la classe et le modèle de régresseur qui aide à resserrer les cadres de délimitation). Par conséquent, le pipeline est vraiment difficile à former. En raison de ces limitations, le RCNN doit être amélioré. S'inspirant de R-CNN, différents modèles évoluent année après année[36].

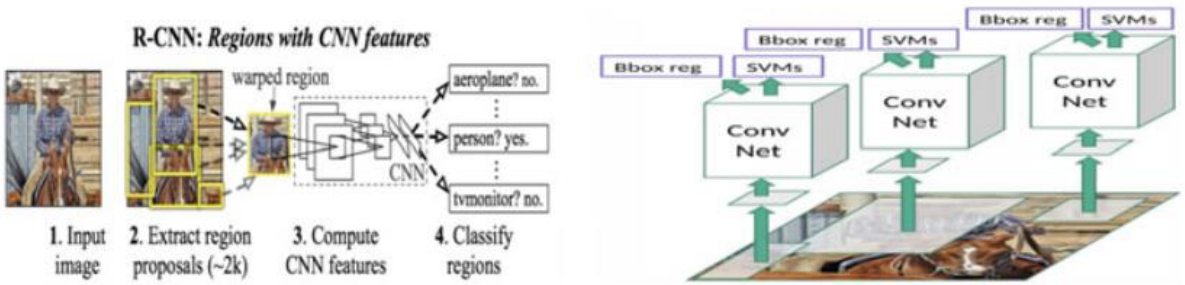


FIGURE 2.9 : L'ARCHITECTURE R-CNN.

3.6.2. Fast R-CNN

L'approche Fast R-CNN est similaire à celle du R-CNN. La différence est au lieu de fournir les propositions régionales au CNN ; l'image d'entrée est directement envoyée au CNN pour générer une carte de caractéristiques convolutives. À partir de là, les propositions de région sont identifiées et déformées en carrés. Pour alimenter une couche entièrement connectée FC, la couche de pooling RoI les remodèle en une taille fixe. Pour prédire la classe de la région proposée, la couche softmax est utilisée à la place de SVM comme le montre la figure 2.10.

Cet algorithme est plus rapide que R-CNN car ~ 2 K propositions de régions ne sont pas obligées d'être transmises au CNN à chaque fois. Il n'est alimenté qu'une seule fois pour

générer la carte des caractéristiques. L'apprentissage du CNN, du classifieur et du régresseur de boîte englobante du R-CNN rapide se fait simultanément dans un seul modèle.

Cependant avec tous ces avantages, la proposition de région est le principal problème de cette technique. Ces propositions ont été faites en utilisant la recherche sélective qui était un processus lent et est devenu un goulot d'étranglement du processus global[36].

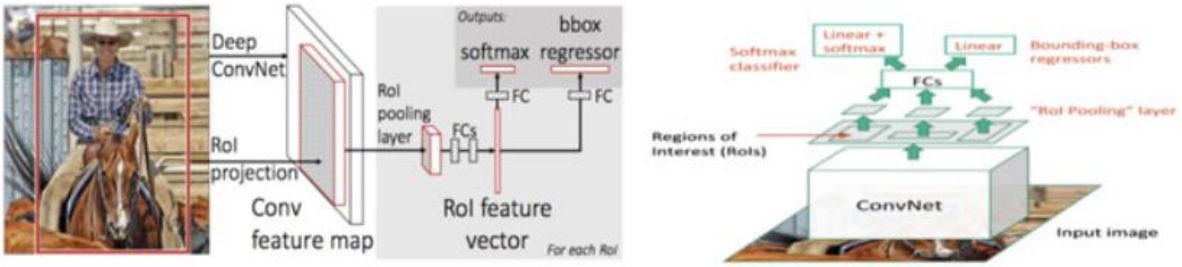


FIGURE 2.10 : L'ARCHITECTURE FAST R-CNN[37].

3.6.3. Faster R-CNN

Les deux algorithmes ci-dessus (R-CNN et Fast R-CNN) utilisent la recherche sélective pour découvrir les propositions de région. La recherche sélective est un processus lent et chronophage qui affecte les performances du réseau. Par conséquent, un algorithme de détection d'objets a été proposé ici qui élimine l'algorithme de recherche sélective et permet au réseau d'apprendre les propositions de région.

Semblable à Fast R-CNN, l'image est fournie en tant qu'entrée à un réseau convolutif qui fournit une carte de caractéristiques convolutives. Au lieu d'utiliser un algorithme de recherche sélective sur la carte des caractéristiques pour identifier les propositions de région, un réseau séparé est utilisé pour prédire les propositions de région. Les propositions de région prédite sont ensuite remodelées en utilisant une couche de regroupement RoI qui est ensuite utilisée pour classer l'image dans la région proposée et prédire les valeurs de décalage pour les cadres de délimitation[36].

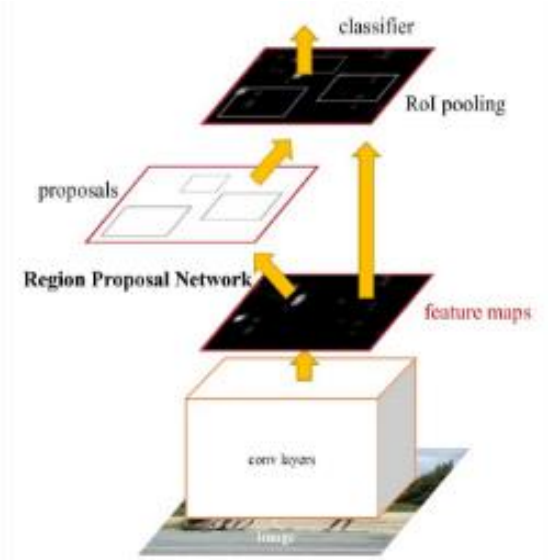


FIGURE 2.11 : L'ARCHITECTURE FASTER R-CNN.

3.6.4. YOLO

Tous les algorithmes de détection d'objets précédents utilisent des régions pour localiser l'objet dans l'image. Le réseau ne regarde pas l'image complète. Au lieu de cela, les parties de l'image qui ont de fortes probabilités de contenir l'objet. YOLO ou You Only Look Once est un algorithme de détection d'objets très différent des algorithmes basés sur la région vus ci-dessus. Dans YOLO, un seul réseau convolutif prédit les boîtes englobantes et les probabilités de classe pour ces boîtes.

Le fonctionnement de YOLO est que nous prenons une image et la divisons en une grille SxS, dans chacune des grilles, nous prenons m boîtes englobantes. Pour chacune des boîtes englobantes, le réseau génère une probabilité de classe et des valeurs de décalage pour la boîte englobante. Les cadres de délimitation ayant la probabilité de classe supérieure à une valeur de seuil sont sélectionnés et utilisés pour localiser l'objet dans l'image.

YOLO est plus rapide que les autres algorithmes (45 images par seconde). La limitation de l'algorithme YOLO est qu'il se débat avec de petits objets dans l'image, par exemple il pourrait avoir des difficultés à détecter un troupeau d'oiseaux. Cela est dû aux contraintes spatiales de l'algorithme[36].

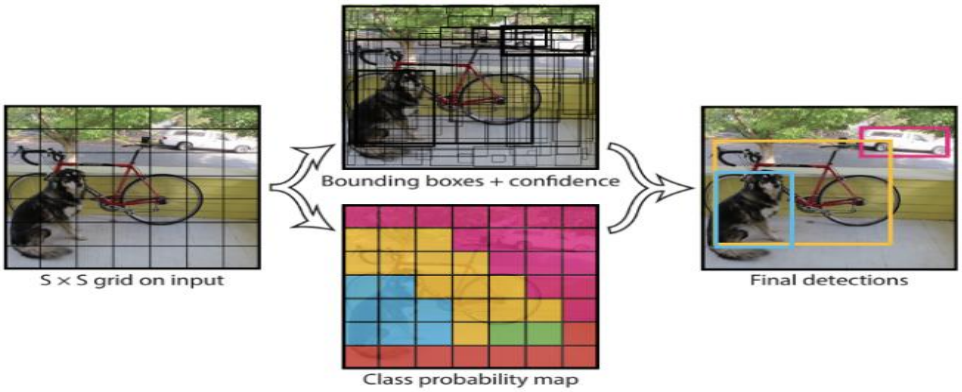


FIGURE 2.12 : YOLO.

3.6.5. Mask R-CNN

Faster R-CNN classe les objets mais il ne peut pas trouver quel pixel fait partie d'un objet dans une image. Par conséquent, le masque R-CNN est développé qui est utilisé pour la segmentation par exemple. La segmentation d'instance comporte deux parties qui sont la détection d'objets et la segmentation sémantique, également appelée segmentation au niveau des pixels. Mask R-CNN a été construit par l'équipe de recherche Facebook AI (FAIR) en avril 2017.

Le modèle peut être divisé en deux parties : un réseau de proposition de région (RPN) et un classificateur de masque binaire. La première étape consiste à obtenir un ensemble de boîtes englobantes qui pourraient éventuellement contenir un objet pertinent et le réseau RoI Align fonctionne sur les principes de la détection d'objets, mais il génère plusieurs boîtes de délimitation possibles plutôt qu'une seule définie. Ces boîtes sont affinées à l'aide d'un autre modèle de régression[36], [38].

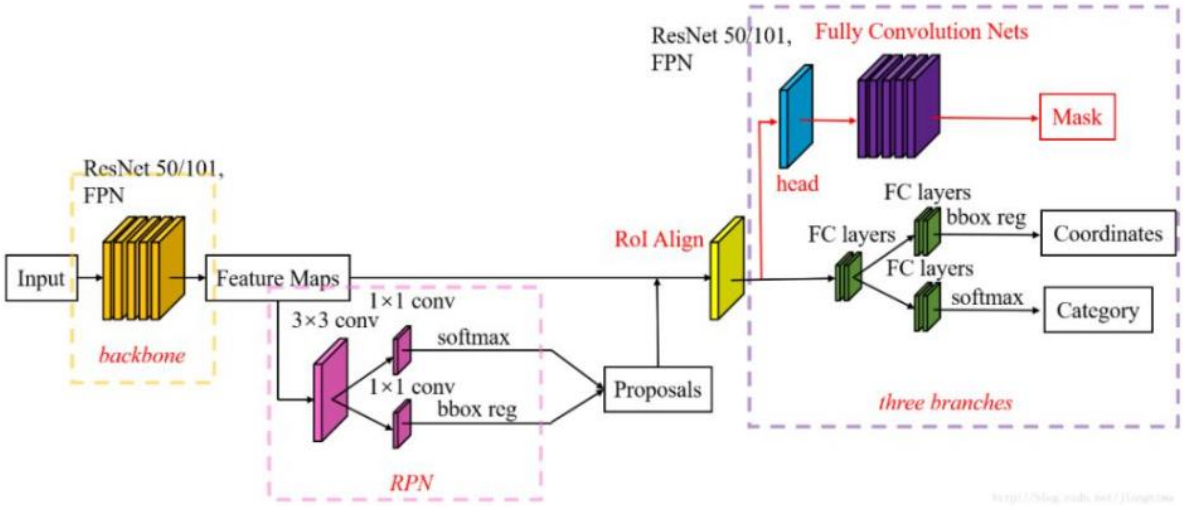


FIGURE 2.13 : L'ARCHITECTURE MASK R-CNN.

Étapes suivies dans Mask R-CNN pour détecter des objets :

1. Premièrement, nous utilisons un extracteur de caractéristiques (un CNN) pour extraire les caractéristiques de l'image entière.

2. Le réseau de proposition de région est appliqué à ces cartes de caractéristiques. Cela renvoie les propositions d'objet, c'est-à-dire les ROI ainsi que leur score d'objectivité.

3. Ensuite, les ROI et les cartes d'entités correspondantes sont transmises via la couche ROI Align. Dans Mask R-CNN, la couche ROI Align est utilisée à la place du ROI Pooling. La couche RoI Align est conçue pour corriger le désalignement d'emplacement causé par la quantification dans le regroupement RoI. Une région d'intérêt est mappée avec précision de l'image d'origine à la carte des caractéristiques sans arrondir aux nombres entiers.

4. Une couche softmax est utilisée au-dessus du réseau entièrement connecté pour prédire les classes. En plus de la couche softmax, une couche de régression linéaire est également utilisée parallèlement à la sortie des coordonnées de la boîte englobante pour les classes prédites.

5. La sortie de la couche ROI Align va également séparément à la couche convolutionnelle pour prédire le masque.

Fonction de perte dans le masque R-CNN :

La perte de masque R-CNN est constituée de la perte due au RPN (Regional Proposed Network) et de la perte due à la classification, à la localisation et au masque de segmentation.

1. Loss(RPN)= RPN_Class Loss + RPN_BBox Loss

2. Loss(Mask R-CNN)= Loss(class labels prediction) + Loss(Bounding Box prediction) + Loss (Mask Prediction)

Total Loss= Loss(RPN) + Loss(Mask R-CNN)

Donc, le problème d'optimisation est de minimiser la perte totale.[38]

3.7. Deep transfert learning

L'apprentissage en profondeur a récemment reçu une attention croissante de la part des chercheurs et a été appliqué avec succès dans de nombreux domaines. Dans certains domaines, comme la bio-informatique et la robotique, il est très difficile de construire un ensemble de données à grande échelle bien annoté en raison des coûts d'acquisition des données et des annotations coûteuses, ce qui limite son développement. L'apprentissage par transfert assouplit l'hypothèse selon laquelle les données d'entraînement doivent être indépendantes et distribuées de manière identique avec les données de test, ce qui motive à utiliser l'apprentissage par transfert pour résoudre le problème des données d'entraînement insuffisantes[39].

L'apprentissage par transfert peut être appliqué à travers plusieurs stratégies différentes à la fois dans l'espace d'apprentissage profond et d'apprentissage automatique. Il existe 3 stratégies principales pour effectuer un apprentissage par transfert sur des modèles d'apprentissage profond :

3.7.1. Utilisation directe de modèles pré-entraînés

La stratégie la plus simple consiste à résoudre une tâche cible en appliquant directement un modèle à partir d'une tâche source. De tels modèles sont généralement de grands réseaux de neurones (des millions de paramètres), entraînés pendant des jours, voire des semaines, sur des machines à la pointe de la technologie. Les grandes entreprises (entreprises, universités, etc.) ont tendance à diffuser de tels modèles au public, dans le but d'améliorer le développement de ce domaine. Certains modèles pré-entraînés utilisés directement incluent le BERT ainsi que YOLO (You Only Look Once), GloVe, UnsupervisedMT et plus encore[39].

3.7.2. Extraction de caractéristiques à partir de modèles pré-entraînés

Au lieu d'utiliser le modèle de bout en bout comme dans l'exemple précédent, nous pouvons traiter le réseau de neurones pré-entraîné comme un extracteur de caractéristiques en supprimant la dernière couche de sortie entièrement connectée.

Cette approche nous permet d'appliquer directement un nouvel ensemble de données pour résoudre un problème entièrement différent.

Il apporte 2 avantages principaux :

- Permet de spécifier les dimensions de la dernière couche entièrement connectée.
- Permet d'utiliser un modèle linéaire léger (par exemple SVM linéaire, régression logistique).

La stratégie « feature extraction » est la mieux adaptée aux situations où le jeu de données de la tâche cible est très petit[39].

3.7.3. Fine-tuning les dernières couches des modèles pré-entraînés

Nous pouvons aller plus loin en entraînant non seulement le classificateur de sortie, mais également en affinant les pondérations dans certaines couches du modèle pré-entraîné. En règle générale, les premières couches du réseau (en particulier pour CNN) sont gelées, tandis que les dernières sont libérées pour le réglage.

Cela nous permet d'effectuer un entraînement complet sur le modèle existant et de modifier les paramètres aux toutes dernières couches[39].

4. Le « Deep Learning » pour le diagnostic du cancer du sein

De nombreux chercheurs ont utilisé des approches DL dans l'analyse d'images médicales. Le succès de DL dépend en grande partie de la disponibilité d'un grand nombre d'échantillons d'apprentissage pour apprendre les mappages de caractéristiques des images, ce qui donne des résultats très précis dans la classification. Par exemple, la tâche de classification d'images, le réseau est formé sur plus d'un million d'images avec plus de 1000 données de classe. Cependant, dans le cas des images médicales, la quantité de données d'entraînement disponibles n'est pas si grande. En outre, il est également difficile d'acquérir un grand nombre d'images étiquetées, car l'annotation elle-même est une tâche coûteuse et pour certaines maladies (par exemple, les lésions) sont rares dans les ensembles de données. De plus, l'annotation de ces échantillons de données, s'il en existe, dans différentes classes souffre de variations intra observateurs, car l'annotation est très subjective et repose sur les connaissances et l'expérience de l'expert[40]. Pour surmonter le défi de l'insuffisance de données, de nombreux groupes de recherche ont mis au point différentes stratégies :

(1) utiliser des correctifs 2D ou des cubes 3D au lieu d'utiliser l'image entière comme entrée, ce qui réduit également les paramètres du modèle et atténue le surajustement ;

(2) en introduisant l'augmentation des données à l'aide de certaines transformations affines (translation, rotation et retournement) et en entraînant le réseau sur les données augmentées ;

(3) en transférant l'approche d'apprentissage en utilisant des poids préformés et en remplaçant simplement les dernières couches par la nouvelle classe ciblée à la place ;

(4) en utilisant des modèles formés avec de petites tailles d'entrée, puis en transformant les poids dans les couches entièrement connectées en noyaux convolutifs.

Dans les sous-sections suivantes on va voir les études récentes pertinentes sur le diagnostic du cancer du sein[41].

4.1. CNN pour l'estimation de la densité mammaire

Le résumé des méthodes est présenté dans le tableau 2.3 avec les mesures de performance pour chaque méthode et les ensembles de données utilisés dans ces études.

Auteur	Méthode	Dataset/Nombre	Tâche	Valeur/s
Mohamed et al	CNN (AlexNet; transfer learning)	Private, University of Pittsburgh/200,00 DM[2] (multiview)	Breast density estimation	AUC[3] (0.9882)
Ahn et al	CNN (transfer learning)	Private, Seoul University Hospital/397 DM (multiview)	Breast density estimation	Correlation coefficient (0.96)
Xu et al	CNN (scratch based)	Public, INbreast dataset/410 DM (multiview)	Breast density estimation	Accuracy (92.63%)
Wu et al	CNN (transfer learning)	Private, New York University School of Medicine/201,179 cases (multiview)	Breast density estimation	Mean AUC (0.934)
Kallenberg et al	Conventional sparse autoencoder, ie, CNN+stacked autoencoder	Private, Dutch Breast Cancer Screening Program and Mayo Mammography, Minnesota/493+668 images (multiview)	Breast density estimation and risk scoring	Mammographic texture (0.91) and AUC (0.61)
Ionescu et al	CNN	Private dataset/67,520 DM (multiview)	Breast density estimation and risk scoring	Average match concordance index of (0.6)
Geras et al	Multiview deep neural network	Private, New York University/886,000 image (multiview)	Breast density estimation and risk score	Mean AUC (0.735)

TABLEAU 2.3 : RÉSUMÉ DES METHODES CONVOLUTIVES BASÉES SUR LES RESEAUX DE NEURONES POUR L'ESTIMATION DE LA DENSITÉ MAMMAIRE.

4.2. CNN pour la détection de la masse mammaire

Nous résumons les méthodes de détection et de classification des lésions en détail dans le tableau 2.4 et illustrons les ensembles de données utilisés, les tâches et les mesures de performance.

Auteur	Méthode	Dataset/Nombre	Tâche	Valeur/s
Dhungel et al	Hybrid CNN+level set	Public, INbreast dataset/410 images (multiview)	Mass detection, classification of benign, and malignant	Accuracy (0.9) and sensitivity (0.98)
Dhungel et al	CRF[4]+CNN	Public, INbreast and DDSM/116 and 158 images (multiview)	Lesion detection and segmentation	Dice score (0.89)

[2] DM: digital mammogram.
[3] AUC: area under the curve.
[4] CRF: conditional random field.

Zhu et al	Fully convolutional network+ CRF	Public, INbreast and DDSM/116 and 158 images (multiview)	Lesion segmentation	Dice score (0.97)
Wang et al	Stacked autoencoder (transfer learning)	Private, Sun Yat-Sen University/1000 Digital mammogram	Detection and classification of calcifications and masses	Accuracy (0.87)
Riddli et al	Faster R-CNN (transfer learning)	Public, DDSM (2620), INbreast (115), and private dataset by Semmelweis University Budapest/847 images	Detection and classification	AUC (0.95)
Singh et al	Conditional generative adversarial network and CNN	Public and private, DDSM and Reus Hospital Spain dataset/567+194 images	Lesion segmentation and shape classification	Dice score (0.94) and Jaccard Index (0.89)
Agarwal et Carson	CNN (scratch based)	Public, DDSM/8750 images (multiview)	Classification of mass and calcifications	Accuracy (0.90)
Gao et al	Shallow-deep convolutional neural network, ie, 4 layers CNN+ResNet	Private, Mayo Clinic Arizona (49 subjects) and public, INbreast dataset (89 subjects) (multiview)	Lesion detection and classification	Accuracy (0.9) and AUC (0.92)
Hagos et al	Multi-input CNN	Private (General Electric, Hologic, Siemens) dataset/28,294 images/(multiview)	Lesion detection and classification	AUC (0.93) and CPM (0.733)
Tuwen et al	Fast R-CNN and Mask R-CNN with ResNet variants as backbone	Private (General Electric, Hologic, Siemens) dataset/23,405 images (multiview)	Lesion detection and classification	Sensitivity (0.97) with 3.56 FP[5] per image
Jung et al	RetinaNet model	Public and private, INbreast and GURO dataset by Korea University Guro Hospital/410+222 images (multiview)	Mass detection and classification	Accuracy (0.98) with 1.3 FP per image

[5] FP: false positive.

Shen et al	CNN end-to-end (transfer learning through visual geometry group 16 and ResNet)	Public, DDSM and INbreast/2584 +410 (multiview)	Classification of masses	AUC (0.96)

TABLEAU 2.4 : RÉSUMÉ DES METHODES CONVOLUTIVES BASÉES SUR LES RESEAUX DE NEURONES POUR LA DÉTECTION DE LA MASSE MAMMAIRE.

4.3. « CNN features » pour la classification de masse

Nous résumons les méthodes de détection et de classification des lésions en détail dans le tableau 2.5 et illustrons les ensembles de données utilisés ainsi que les mesures de performance des tâches.

Auteur	Méthode	Dataset/Nombre	Tâche	Valeur/s
Levy et Jain	AlexNet and GoogleNet (transfer learning)	Public, DDSM dataset/1820 images (multiview)	Breast mass classification	Accuracy (0.924), precision (0.924), and recall (0.934)
Samala et al	Multistage fine-tuned CNN (transfer learning)	Private+public, University of Michigan and DDSM/4039 ROIs (multiview)	Classification performance on varying sample sizes	AUC (0.91)
Jadoon et al	CNN- Discrete wavelet and CNN-curvelet transform	Public, image retrieval in medical applications dataset/2796 ROI patches	Classification	Accuracy (81.83 and 83.74) and receiver operating characteristic curve (0.831 and 0.836) for both methods
Huynh et al	CNN (transfer learning)	Private, University of Chicago/219 images (multiview)	Classification of benign and malignant tumor	AUC (0.86)
Domingues et Cardoso	Autoencoder	Public, INbreast/116 ROIs	Classification of mass vs normal	Accuracy (0.99)
Wu et al	GAN[6] and ResNet50	Public, DDSM dataset/10,480 images (multiview)	Detection and classification of benign and malignant calcifications and masses	AUC (0.896)

[6] GAN: generative adversarial network.

Sarah et al	CNN (transfer learning)	Public, Full-field digital mammography and DDSM/14,860 images (multiview)	Classification	AUC (0.91)
Wang et al	CNN and long short-term memory	Public, Breast Cancer Digital Repository (BCDR-F03)/763 images (multiview)	Classification of breast masses using contextual information	AUC (0.89)
Shams et al	CNN and GAN	Public, INbreast and DDSM (multiview)	Classification	AUC (0.925)
Gastounioti et al	Texture feature+CNN	Private/106 cases (mediolateral oblique view only)	Classification	AUC (0.9)
Dhungel et al	Multi-ResNet	Public, INbreast (multiview)	Classification	AUC (0.8)

TABLEAU 2.5 : RÉSUMÉ DES METHODES CONVOLUTIVES BASÉES SUR LES RESEAUX DE NEURONES POUR LA CLASSIFICATION DE LA MASSE MAMMAIRE.

5. Conclusion

Nous vivons dans une époque où les possibilités sont infinies et la technologie de « *Deep Learning* » peut nous aider à réaliser de nouvelles avancées technologiques. Il est très important de souligner que la hiérarchie des couches et la supervision de l'apprentissage sont les principaux facteurs clés pour développer une application réussie avec un apprentissage en profondeur. La raison en est que la hiérarchie est essentielle pour une classification appropriée des données et la supervision estime qu'il est important de maintenir la base de données. Le « *Deep Learning* » repose sur l'optimisation des applications existantes en « *Machine Learning* » et son caractère innovant sur le traitement de la couche hiérarchique. Ce chapitre à dresser un état de l'art exhaustif sur la technologie du *« Deep Learning »* en général et son utilisation dans la détection du cancer du sein objet de notre travail dans ce mémoire de Master. Le chapitre suivant présente notre contribution dans ce domaine en présentant différents modèles de *« Deep Learning »* pour la détection et la clacification du tumeurs mammaires, les moèles utilisés sont tester et évaluer sur le jeu de données CBIS-DDSM[42].

Chapitre 3

Une approche basée Deep Learning pour la détection du cancer du sein :

1. Introduction

Après avoir présenté dans les parties précédentes un état de l'art sur l'utilisation des techniques du « *Deep Learning* » dans le domaine du cancer du sein, ou on présenter les approches, les techniques et les outils. Nous allons consacrer ce chapitre pour présenter en détail les modèles qu'on propose pour la détection d'anomalies et leur diagnostique dans les images de mammographie à l'aide de « *Deep Convolutional Neural Networks* ».

On a testé plusieurs modèles et nous joignons que ceux avec meilleurs résultats pour chaque catégorie. La première dite de 2 classes (2 sorties soit masse/calcification, soit bénigne/maligne). La deuxième catégorie dite de modèle à 4 classes (4 sorties : masse bénigne/masse maligne/calcification bénigne/calcification maligne). Ainsi qu'on va être les premiers à utiliser un modèle de pointe pour la segmentation d'instances « *mask R-CNN* ». Il étend « *Faster R-CNN* », le modèle utilisé pour la détection d'objets, en ajoutant une branche parallèle pour la prédiction des masques de segmentation sur chaque région d'intérêt (RoI) afin d'améliorer les performances de détection. Tout en adoptant CBIS-DDSM comme base de donnée.

2. Ressources utilisée

Avant de présenter notre approche, nous présentons le jeux de données (dataset) utilisé pour tester et évaluer notre démarche. La section suivante sera dédiée à cette tâche.

2.1. Dataset

Le dataset de référence est La DDSM (*Digital Database for Screening Mammograph*) qui contient des mammographies de film numérisés dans un format lossless-JPEG, cette version est maintenant obsolète. Nous avons utilisé une version ultérieure de DDSM appelée CBIS-DDSM[42]« *Curated Breast Imaging Subset of Digital Database for Screening Mammograph* » qui contient des images qui sont converties au format DICOM standard, où toutes les images ont été segmentées et étiquetées. L'ensemble de données a été téléchargé à partir du site Web[7] du CBIS-DDSM, et comprenait à la fois des vues craniocaudales (CC) et médiolatérales obliques (MLO) pour la plupart des examens. Chaque vue a été traitée comme une image distincte dans cette étude.

2.2. Prétraitement des données (Data pre-processing)

Modèle 1-4 : Toutes les mammographies ont été converties au format PNG (extraction des RoIs de la mammographie dont tous les « abnormality patches » ont été extraits de chaque image CBIS-DDSM en appliquant son masque binaire, qui est également fourni avec elle. Le RoI résultant a ensuite été redimensionné à 150x150) et l'ensemble de données relatives est fourni sous la forme des fichiers CSV on les transforme en un ensemble de tableaux numpy, contenant les images et les étiquettes à utiliser pour la formation (training) et les tests.

[7] https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM

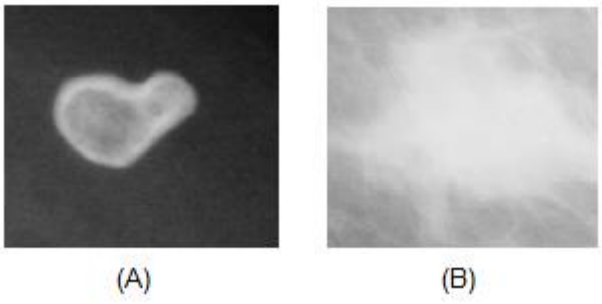


FIGURE 3.1 : EXEMPLE DE (A) : CALCIFICATION (B) : MASSE.

Avant que ces données puissent être réellement utilisées comme entrée pour les modèles NN, quelques étapes de prétraitement sont nécessaires. En fonction de la tâche de classification spécifique (par exemple masse/calcification, bénigne/maligne, ...), les actions à effectuer peuvent être légèrement différentes. La liste suivante décrit toute la séquence de préparation des données :

1. Nous créons un tableau numpy à partir de chaque CSV file data (fourni avec CBIS-DDSM).

2. Importez les données d'entraînement et de test sous forme de tableaux numpy à partir de fichiers .npy.

3. Remappez les étiquettes en fonction du nombre et des classes impliquées dans la classification spécifique. Si la tâche consiste simplement à faire la distinction entre les masses et la calcification, alors seulement deux étiquettes (0-1) sont nécessaires. Inversement, quatre étiquettes (0-3) sont nécessaires lorsqu'il est également important de distinguer les anomalies bénignes des anomalies malignes.

4. Normaliser les valeurs de pixels pour qu'elles soient dans une plage compatible avec le modèle choisi. Les modèles Scratch CNN utilisent une entrée dans la plage (0,1) virgule flottante, tandis que VGGNet et d'autres modèles pré-entraînés sont conçus pour fonctionner avec des images en (0,255) qui sont ensuite prétraitées avec des transformations personnalisées.

5. Mélangez l'ensemble d'apprentissage (et les étiquettes en conséquence, bien sûr).

6. Divisez les données de formation en sous-ensembles « *training* » et « *validation* ». Le premier sera utilisé pour calculer la fonction de perte, tandis que l'ensemble de validation n'est utilisé que pour surveiller les performances réelles sur un ensemble indépendant pendant l'entraînement.

7. Instanciez les générateurs « *Keras* » en tant que sources de données pour le réseau. Les paramètres d'augmentation des données peuvent être spécifiés à ce stade.

À la fin du processus, un ou plusieurs des échantillons résultants sont effectivement visualisés pour vérifier que :

- Les données sont formatées comme prévu (size, range, ...).
- Le contenu des images est toujours significatif et n'a pas été accidentellement corrompu pendant le processus.

Modèle 5 :

1. Extraire les métadonnées CBIS-DDSM du fichier Excel fourni et enregistrer les annotations dont nous avons besoin dans un fichier JSON.

2. Toutes les mammographies ont été converties au format PNG (full image).

3. Présentation des outils utilisés

3.1. Le software

Plusieurs Framework et bibliothèques open sources sont disponibles dans la littérature, la grande majorité supporte le langage Python. Nous avons utilisé le langage *python 3.7* et une listes de quelques bibliothèques préparées avec des paramètres :

- La bibliothèque Keras.
- La bibliothèque Tensorflow.
- La bibliothèque Numpy.
- La bibliothèque OpenCV-python.
- La bibliothèque Scikit-image.
- La bibliothèque Pillow.

3.2. Le hardware

Le « *Deep Learning* » est un domaine avec des exigences des calculs intenses et la disponibilité des ressources (surtout GPU) qui vont fondamentalement influencer sur l'expérience de l'utilisateur car sans ses ressources, il faudra trop de temps pour la construction des modèles ainsi la détection des erreurs de programmations adjacentes, ce qui rend la tâche de programmation très lourdes.

Les expérimentations ont tous été effectuées sur une machine qui offre des performances acceptables dont voici les caractéristiques utiliser dans notre travaille :

CPU	Intel® Core™ i3-2330M (2.20 GHz)
GPU	NONE
RAM	4 GO

TABLEAU 3.1 : LES PERFORMANCES DE NOTRE MACHINE.

Cependant, vue les fortes exigences de calculs et espaces nécessaires pour l'implémentation des solutions à base de « *TensorFlow* » et « *Keras* », nous avons opté pour l'utilisation de calculs en *Cloud* via la plateforme « *Google Colab* ». Cette dernière offre la possibilité de calcul *GPU*, *TPU* et même *CPU*. C'est une plateforme gratuite pour des sessions de temps limitées sur le web. Elle utilise l'environnement « *Jupyter Notebook* » pour les différentes routines de programmation en langage Python.

4. Solution proposée

L'ensemble de données d'intérêt est l'ensemble de données CBIS-DDSM dont les images mammographiques présentent deux types d'anomalies mammaires : la masse et la calcification, qui peuvent être bénignes ou malignes.

Notre travail consiste à proposer une solution basé *« Deep Learning »* pour la détection du cancer de sein. Pour atteindre cet objectif on a proposé cinq réseaux deep afin

d'avoir un taux de détection élevé par rapport aux travaux existants. On a classé les cinq modèles utilisés en deux grandes catégories : la première dite de 2 classes (2 sorties soit masse/calcification, soit bénigne/maligne). La deuxième catégorie dite de modèle à 4 classes (4 sorties : masse bénigne, masse maligne, calcification bénigne, ou bien calcification maligne). Dans les sections suivantes on va présenter en détail chaque modèle.

5. Architecture des modèles proposés et résultats

En ceux qui suivent des modèles de classification des mammographies de dépistage pour détecter les lésions malignes et le type de lésion (masse vs calcification) à l'aide de CNN. Pour ce qui concerne le dernier modèle en va utiliser Mask R-CNN qui est un modèle de pointe pour la segmentation d'instances et combine la détection d'objets (lesions dans notre cas), où le but est de classer les objets individuels et de les localiser à l'aide d'un cadre de sélection « *bounding box* ». Mask R-CNN est construit sur Faster R-CNN utilisant l'architecture ResNet 101 pour l'extraction de caractéristiques de l'image ainsi que le réseau de proposition de région (RPN) pour générer une région d'intérêts (RoI). Il a une branche supplémentaire pour prédire les masques de segmentation sur chaque région d'intérêt (RoI) d'une manière pixel à pixel.

5.1. Modèle 1 (VGG 2-classes Masse/Calcification)

L'exploration des méthodes d'apprentissage en profondeur dans le domaine de l'imagerie mammaire a été limitée, car seul un petit nombre d'ensembles de données publics sont disponibles (par exemple DDSM, INbreast) c'est pour cela on a opté pour le transfert learning.

VGG16 est un modèle de réseau de neurones convolutif proposé par Simonyan et al., avec plusieurs couches convolutives 3x3 en cascade parfois entrelacées avec 2x2 maxpooling layers (figure 3.2). Développé pour le défi ILSVRC2014, il a pu atteindre une précision top-5 de 92,7% sur ImageNet dataset. Sa taille énorme fait de la formation un processus extrêmement lent ; néanmoins, VGG16 est souvent utilisé pour l'apprentissage par transfert, grâce à sa flexibilité[43].

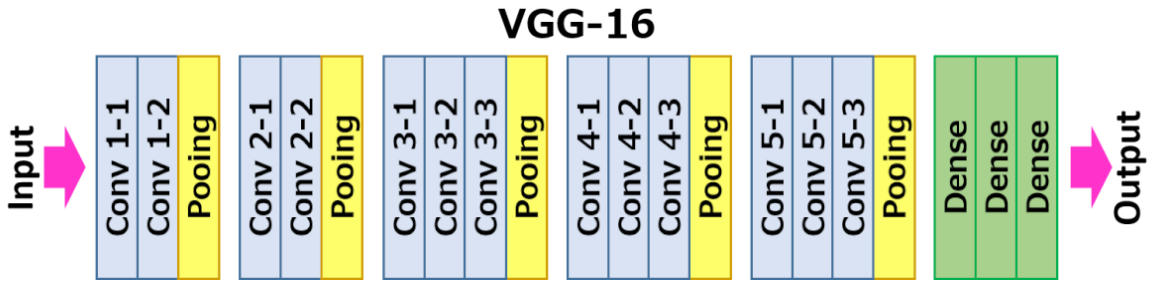


FIGURE 3.2 : ARCHITECTURE VGG-16.

Ce modèle consiste à distinguer les masses et les calcifications dans les mammographies.

5.1.1. Modèle 1.1 « VGG avec feature extraction »

L'extraction de caractéristiques comme on l'a présenté dans le chapitre 2 (voir section 3.7.2.) est une approche d'apprentissage par transfert qui consiste à prendre un modèle déjà formé sur un autre jeu de données (éventuellement similaire), à découper la couche finale « *fully-connected* » et à la remplacer par une nouvelle (c.à.d. avec de nouveaux poids « *weights* »). La base convolutive est figée, c.à.d. qu'elle ne peut pas être entraînée, car elle sert d'extracteur de caractéristiques pour les nouvelles images et ne doit pas être modifiée par le processus d'apprentissage.

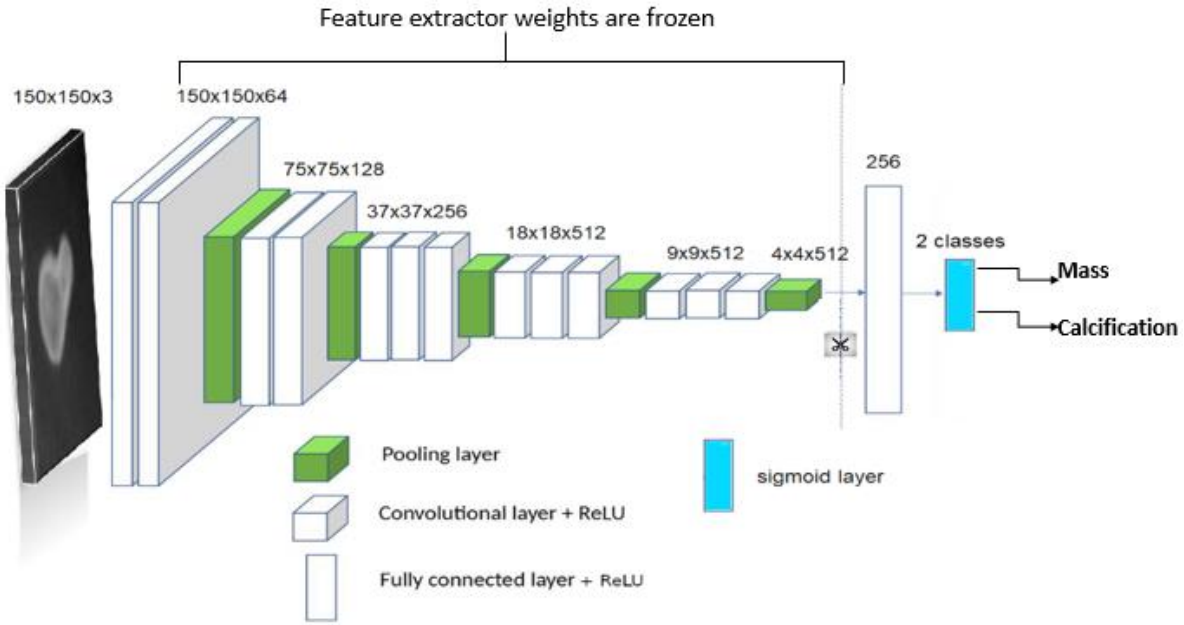


FIGURE 3.3 : ARCHITECTURE VGG16 FEATURE EXTRACTION.

5.1.2. Les résultats du modèle 1.1

On ce qui suit les résultats de l'utilisation du modèle VGG 16 Vanilla ainsi que les résultats après l'application des méthodes de régularisation.

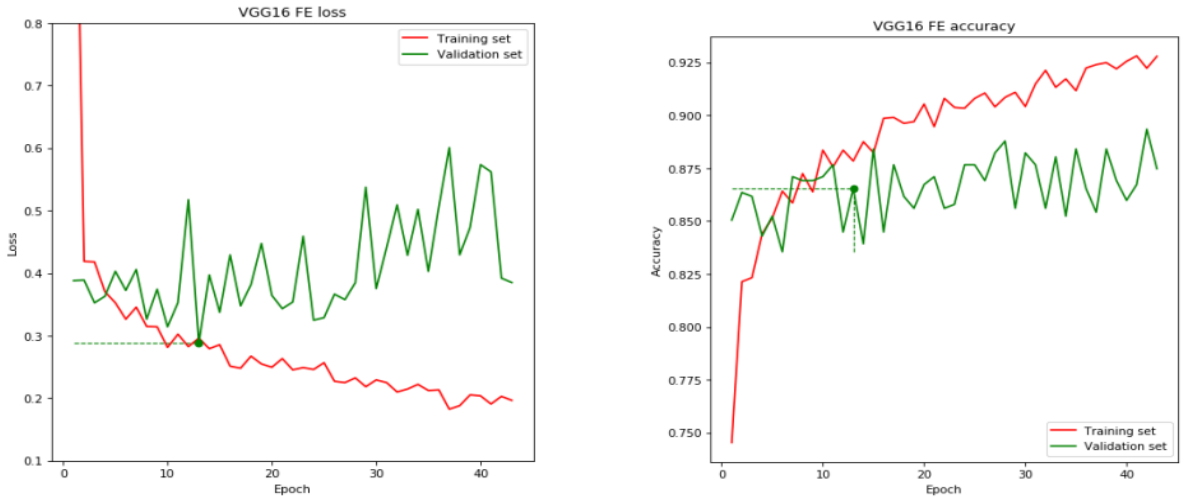
i. **Vanilla :** Le VGG16 d'origine a 4096 couches de FC suivies de 1000 neurones softmax, ce qui est bien pour ImageNet mais certainement surdimensionné pour notre objectif. Par conséquent, la base convolutive est laissée telle quelle, et le bloc entièrement connecté est réduit en une seule couche avec 256 neurones alimentant une unité de sortie sigmoïdale (figure 3.3).

Feature extraction (vanilla)				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
13	**86.54**	**86.90**	**28.86**	**43.20**

TABLEAU 3.2: RÉSULTATS VGG16 FEATURE EXTRACTION (VANILLA).

Le réseau apprend très vite à faire la distinction entre les masses et les calcifications, avec plus de 86% de précision. D'un autre côté, en regardant le graphique des pertes,

il commence très vite à se surajuster « *overfitting* » : une méthode de régularisation est très nécessaire.



(a) Loss (b) Accuracy

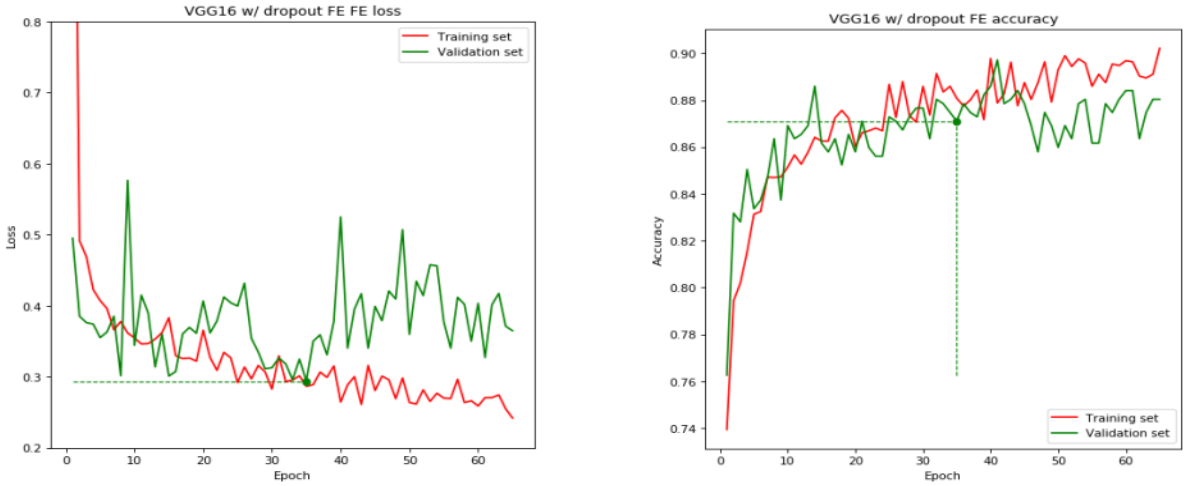
FIGURE 3.4 : VGG16 FEATURE EXTRACTION (VANILLA).

ii. Avec dropout : Un « dropout » est ajouté à la fin du réseau pour réduire le surajustement « overfitting ».

Feature Extraction (w/ dropout)				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
35	**87.10**	**87.50**	**0.2940**	**0.4637**

TABLEAU 3.3: RÉSULTATS VGG16 FEATURE EXTRACTION AVEC DROPOUT.

Comme prévu, le dropout a atténué l'ampleur du sur ajustement, mais il n'a pas été suffisant pour l'éliminer. Dans l'ensemble, le temps de formation a augmenté et la précision des tests s'est améliorée à 87,5%.



(a) Loss (b) Accuracy

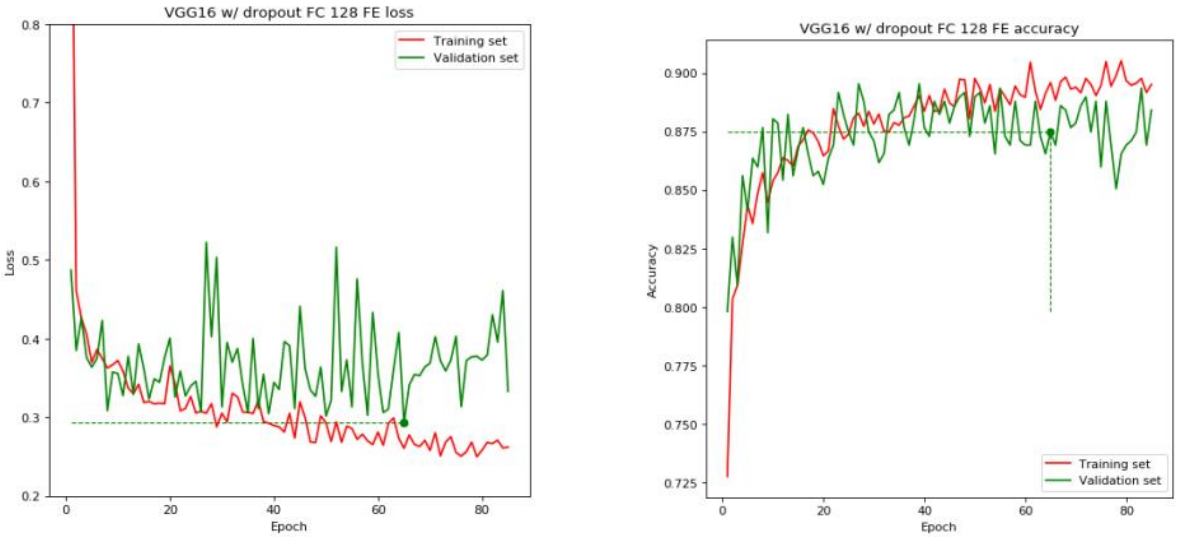
FIGURE 3.5 : VGG16 FEATURE EXTRACTION AVEC DROPOUT.

iii. **Avec dropout et FC plus petite :** Une autre façon de réduire le sur ajustement consiste à réduire la taille du réseau, par ex. la couche FC. Notez que cette approche est déconseillée lorsque le nombre de neurones est déjà relativement faible par rapport à la complexité du problème : cela entraînera probablement une baisse de précision. La couche entièrement connectée compte désormais 128 neurones au lieu de 256.

Feature Extraction (w/ dropout and smaller FC)				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
65	**87.48**	**86.61**	**0.2930**	**0.4261**

TABLEAU 3.4: RÉSULTATS VGG16 FE AVEC DROPOUT ET PLUS PETITE FC LAYER.

En réduisant la dernière couche FC et en utilisant le « dropout », le sur ajustement a diminué encore plus. Cependant, le redimensionnement du FC a légèrement nui aux performances, en termes de temps d'entraînement et de précision finale.



(a) Loss (b) Accuracy

FIGURE 3.6 : VGG16 FEATURE EXTRACTION AVEC DROPOUT ET PLUS PETITE FC LAYER.

iv. **Data augmentation :** l'ensemble d'entraînement est déjà augmenté, mais son intensité peut être augmentée avec des transformations plus radicales. On étend la plage de cisaillement à 15 ° (à partir de 10 °) et introduisant des décalages jusqu'à 20% de la taille. L'objectif est d'avoir un ensemble d'entraînement plus varié afin de réduire le sur ajustement.

Feature Extraction (w/ dropout and more augmentation)				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
73	**87.48**	**85.71**	**0.2910**	**0.3940**

TABLEAU 3.5: RÉSULTATS VGG16 FE AVEC DROPOUT ET STRONGER DATA AUGMENTATION.

L'utilisation d'une augmentation de données plus forte a considérablement réduit le sur ajustement. La précision du test est restée plus ou moins inchangée, ce qui indique que les performances du réseau ont déjà été poussées à la limite.

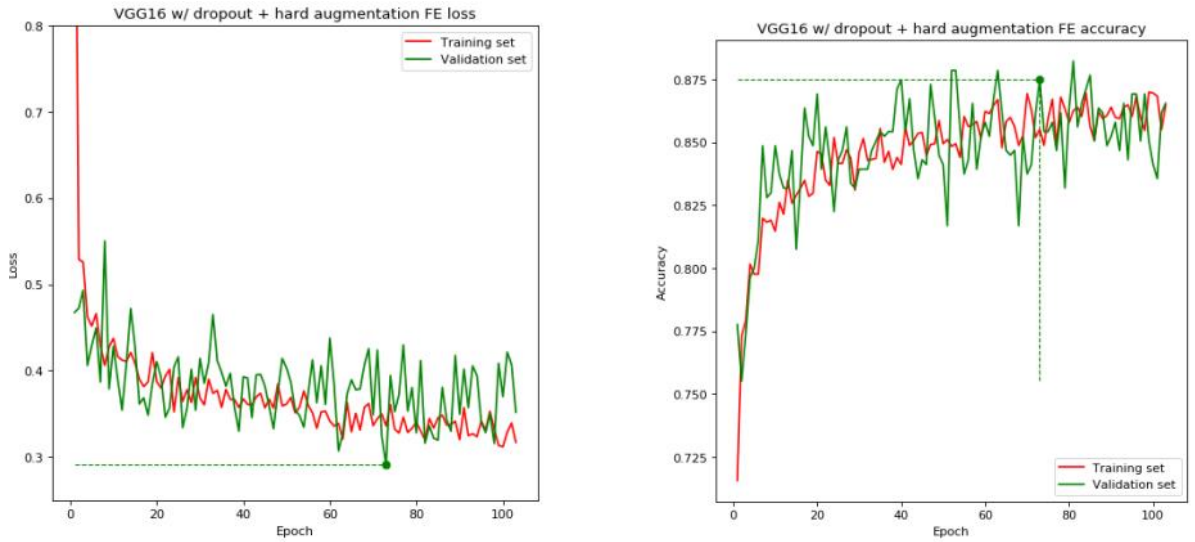


FIGURE 3.7 : VGG16 FE AVEC DROPOUT ET STRONGER DATA AUGMENTATION.

5.1.3. Modèle 1.2 « VGG avec fine tuning »

Le « fine tuning » étend l'extraction des caractéristiques comme on l'a présenté dans le chapitre 2 (voir section 3.7.3.). L'architecture de ce modèle est décrite dans la figure 3.8; le « unfreezing » (c.à.d le fait de rendre entraînable) une ou plusieurs des dernières couches convolutives après que le nouveau bloc FC ait été entrainé. S'entraînant pendant quelques époques supplémentaires, éventuellement avec un taux d'apprentissage très faible, ces couches convolutives obtiennent un changement légèrement ajusté (fine-tuned) pour améliorer les performances.

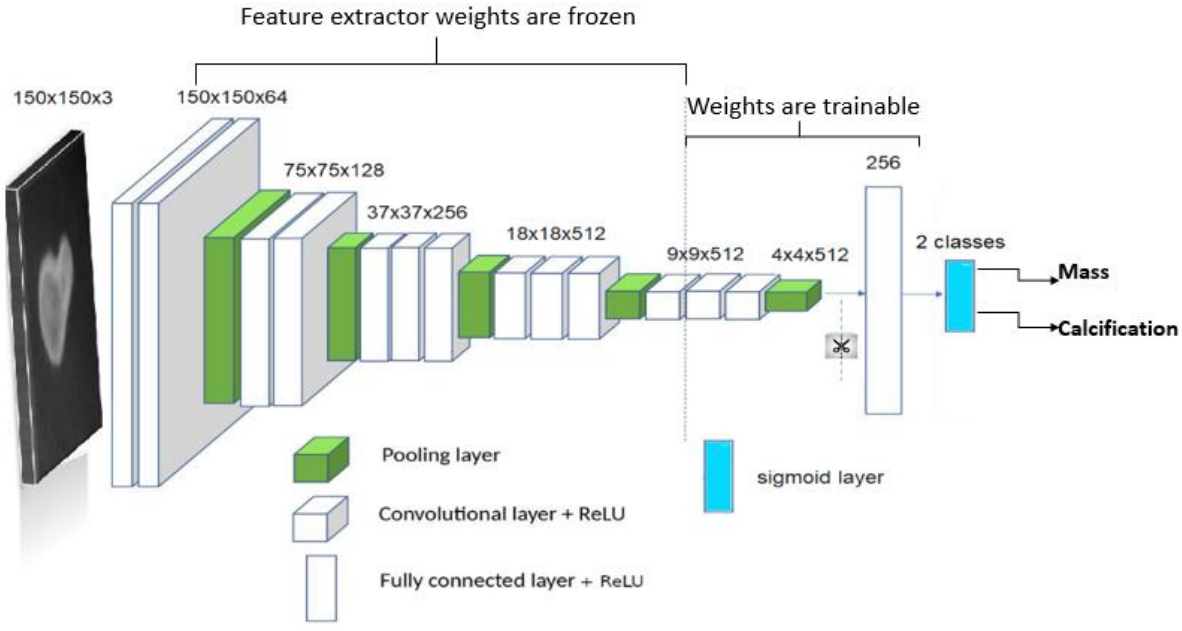


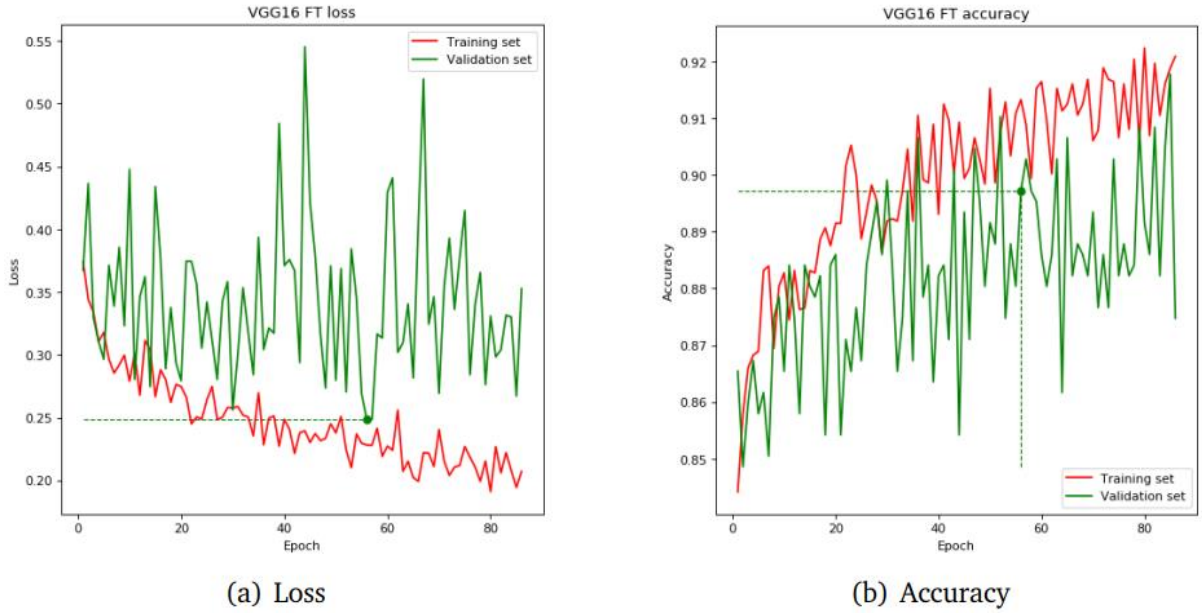
FIGURE 3.8 : ARCHITECTURE VGG16 2-CLASS FINE TUNING (2 LAYER).

5.1.4. Les résultats du modèle 1.2

i. **Une couche :** à partir du dernier modèle (extraction de caractéristiques avec augmentation), la 3ème couche Conv2D du 5ème bloc est dégelée. Après un certain temps, un nouveau minimum de « validation loss » est atteint, comme le montre la figure 3.9 (a).

Fine tuning (1 layer)				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
56	**89.72**	**89.58**	**0.2484**	**0.3228**

TABLEAU 3.6 : RÉSULTATS VGG16 2-CLASS FINE TUNING (1 LAYER).



(a) Loss (b) Accuracy

FIGURE 3.9 : VGG16 2-CLASS FINE TUNING (1 LAYER).

Grâce à « fine tuning », le NN a obtenu une précision de 89,58% sur l'ensemble de test, avec un gain de 3,87% par rapport au précédent feature extraction VGG16 modèle.

ii. **Deux couches :** Le même processus peut être appliqué de manière itérative, cette fois à partir du 1-layer finetuned et en « unfreezing » l'avant-dernière couche convolutive.

Fine tuning (2 layers)				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
9	**91.03**	**91.37**	**0.1935**	**0.3312**

TABLEAU 3.7: RÉSULTATS VGG16 2-CLASS FINE TUNING (2 LAYER).

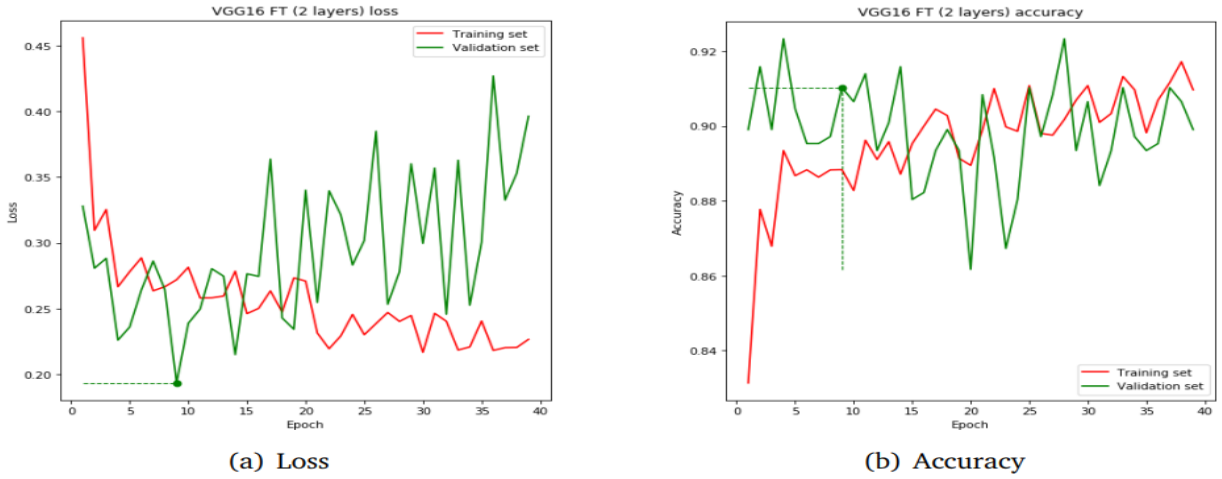


FIGURE 3.10 : VGG16 2-CLASS FINE TUNING (2 LAYER).

Après très peu d'époques, « validation loss » est tombée trop et la précision a atteint la valeur exceptionnelle de 91,4% sur l'ensemble de test. Le réseau a démarré immédiatement après. En effet, le « unfreezing » de la deuxième couche laissait trop de paramètres réseau libres pour le réglage, donc le résultat que nous avons obtenu est fondamentalement le meilleur que nous puissions en tirer. Pour cette raison, affiner (fine-tuning) une troisième couche ne serait pas du tout utile.

5.2. Modèle 2 « scratch CNN 2-classes Benign/Malignant »

Cette tâche consiste à construire un modèle pouvant distinguer entre les cas bénins et ceux malins.

On ajuste les étiquettes pour le problème de classification binaire, de sorte que

- 0 correspond à " Benign " ;
- 1 mappe à " Malignant ".

Son architecture est représentée dans la figure 3.11.

Conv2D (64 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Conv2D (128 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Conv2D (256 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Flatten ( )

Dense (64 , relu)

Dropout (0.5)

Dense (1 , sigmoid)

Modèle CNN B-M

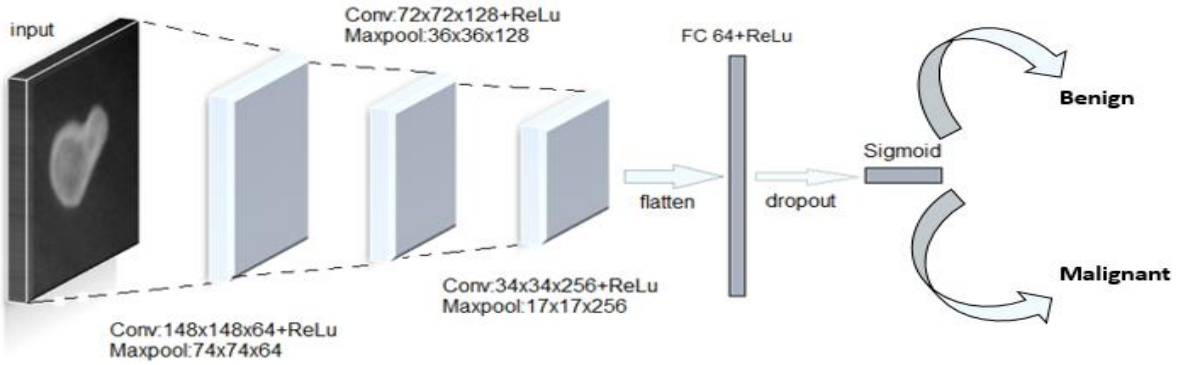


FIGURE 3.11 : ARCHITECTURE CNN B/M (2-CLASSES).

5.2.1. Les résultats du modèle 2

Voici les résultats de la classification Bénigne/Maligne du modèle de scratch CNN :

CNN 2-class Benign-Malignant				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
419	**68.97**	**69.05**	**0.5248**	**0.5918**

TABLEAU 3.8 : RÉSULTATS CNN B/M (2-CLASSES).

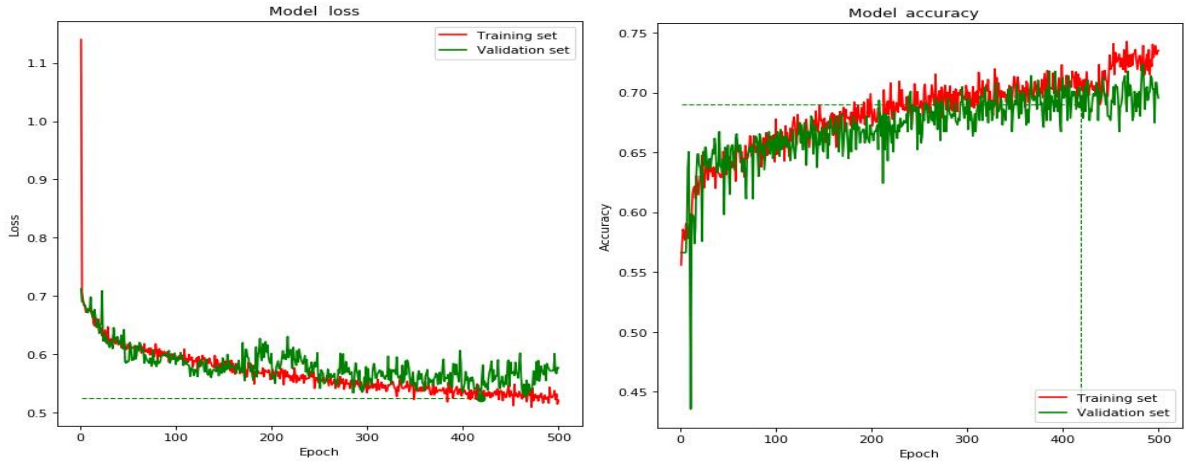


FIGURE 3.12 : CNN B/M (2-CLASSES).

Après plus de 400 époques notant que la précision a atteint la valeur de 69,05% sur l'ensemble de test. Il faut signaler que nous avons traité les 2 cas de type (masse ou bien calcification) en même temps tout ça c'est une préface au 4ème modèle que nous allons voir (section 5.4).

5.3. Modèle 3 « scratch CNN 4-classes »

Dans cette tâche au lieu de distinguer uniquement les masses et les calcifications, il y a un défi supplémentaire : prédire si l'anomalie est bénigne ou maligne. Ajustez les étiquettes pour le problème de classification, de sorte que :

- 0 corresponde à " masse bénigne " ;
- 1 corresponde à " masse maligne " ;
- 2 corresponde à " calcification bénigne " ;
- 3 corresponde à " calcification maligne ".

5.3.1. Modèle 3.1

Le modèle utilisé dans la classification à deux classes est un bon point de départ, car les deux tâches se chevauchent partiellement et fonctionnent sur le même ensemble de données.

Bien entendu, certaines modifications sont nécessaires, notamment dans la couche de sortie, car le problème est désormais une classification multiclasse: 4 neurones softmax remplacent l'ancienne unité sigmoïde.

De même, la fonction de perte est désormais la crossentropie catégorique, une version généralisée de la crossentropie binaire.

Conv2D (32 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Conv2D (64 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Conv2D (128 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Conv2D (256 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Flatten ( )

Dense (48 , relu)

Dropout (0.5)

Dense (4 , softmax)

Modèle 3.1

5.3.2. Résultats du modèle 3.1

Voici les résultats de la classification de la 2ème catégorie 4 sorties du modèle 3.1 de scratch CNN :

1er Modèle CNN 4-classes				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
208	**63.55**	**57.74**	**0.7709**	**0.9402**

TABLEAU 3.9 : RÉSULTATS SCRATCH CNN 1 (4-CLASSES).

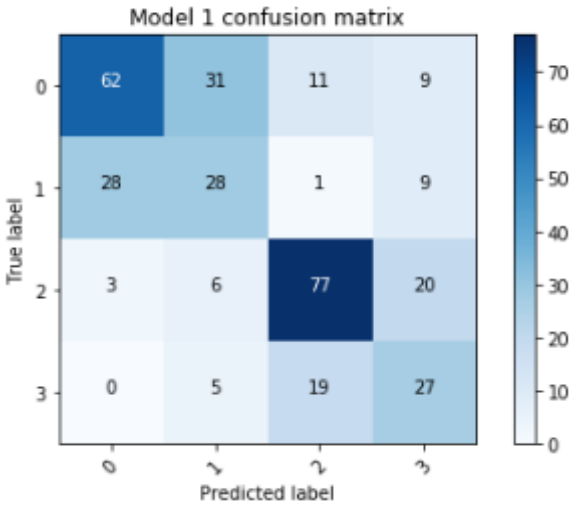


FIGURE 3.13 : 4-CLASSES CNN 1 CONFUSION MATRIX.

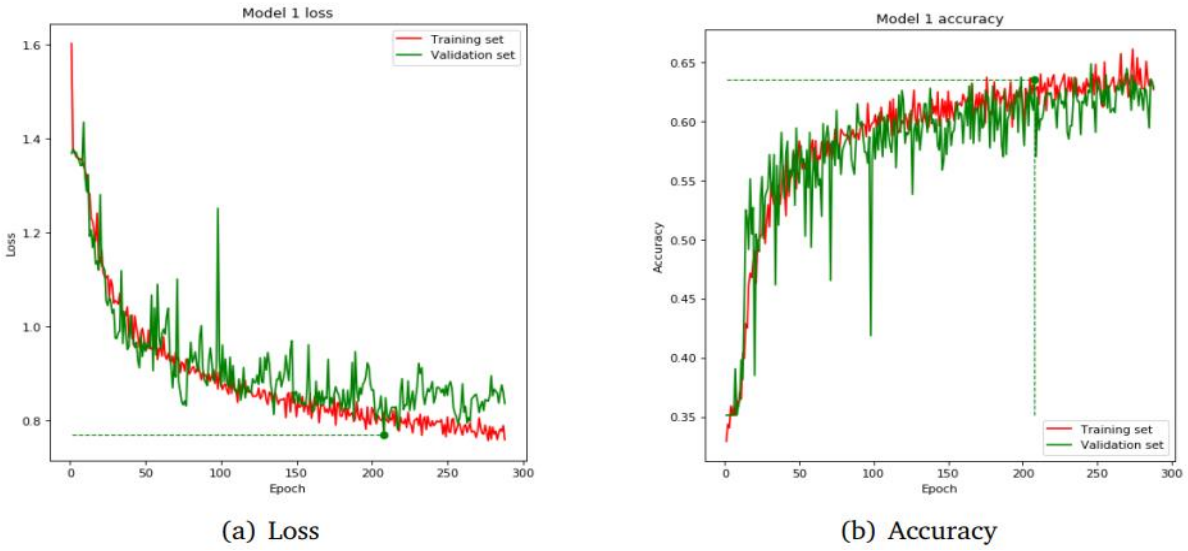


FIGURE 3.14 : SCRATCH CNN 1 (4-CLASSES).

Le réseau a obtenu une précision finale de 57,7% sur l'ensemble de test. Cette valeur est nettement basse, mais c'est une conséquence évidente de l'augmentation du nombre de classes. Maintenant, il suffit de mal prévoir le type ou la malignité pour considérer que l'échantillon est mal classé.

La matrice de confusion fournit ici (figure 3.13) des informations très utiles. En se souvenant de la façon dont les étiquettes ont été mappées aux classes, on peut observer que 44 images ont été affectées du mauvais type (masse/calc), soit 13% des échantillons. Il convient de souligner que la plupart des erreurs sont en effet dues à une mauvaise classification bénigne/maligne, plutôt qu'à un échange masse/calcification. Apparemment, la tâche de prédire la nocivité est une tâche plus difficile que d'identifier le type d'anomalie. Enfin, un sur-ajustement modéré est perceptible à la fin de la formation.

5.3.3. Modèle 3.2 « more filters »

Le deuxième modèle est moins profond, mais avec deux fois plus de filtres qu'auparavant dans les couches correspondantes. Alternativement, il peut être vu comme un clone du précédent où le premier bloc conv a été ignoré. L'hypothèse à vérifier est que plus de filtres dans les premiers stades convolutifs peuvent aider à obtenir une meilleure précision. Plus précisément :

• Trois couches conv2D avec 64, 128, 256 filtres.

• La dernière couche FC a un peu plus de neurones (64, auparavant 48) pour faire face à l'augmentation des filtres.

Conv2D (64 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Conv2D (128 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Conv2D (256 , (3 , 3) , relu)

MaxPooling2D ((2 , 2))

Flatten ( )

Dense (64 , relu)

Dropout (0.5)

Dense (4 , softmax)

Modèle 3.2 More Filters

5.3.4. Les résultats du modèle 3.2

Voici les résultats de la classification de la 2ème catégorie 4 sorties du modèle 3.2 de scratch CNN :

2ème Modèle CNN 4-classes				
Epoch stop	**Validation accuracy**	**Testing accuracy**	**Validation loss**	**Testing loss**
483	**65.42**	**61.31**	**0.7440**	**0.9047**

TABLEAU 3.10 : RÉSULTATS SCRATCH CNN 2 (4-CLASSES).

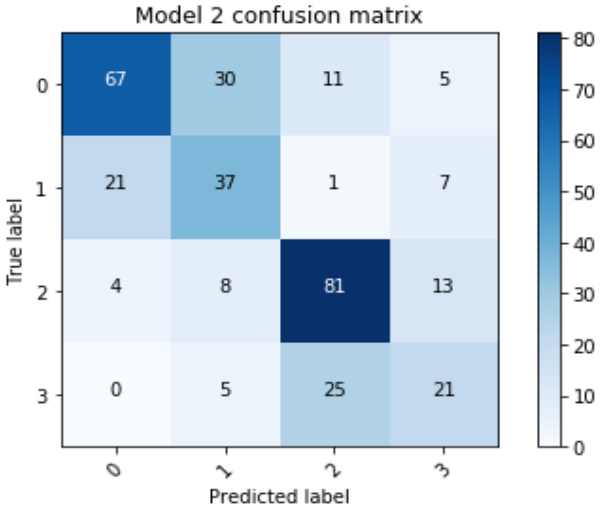
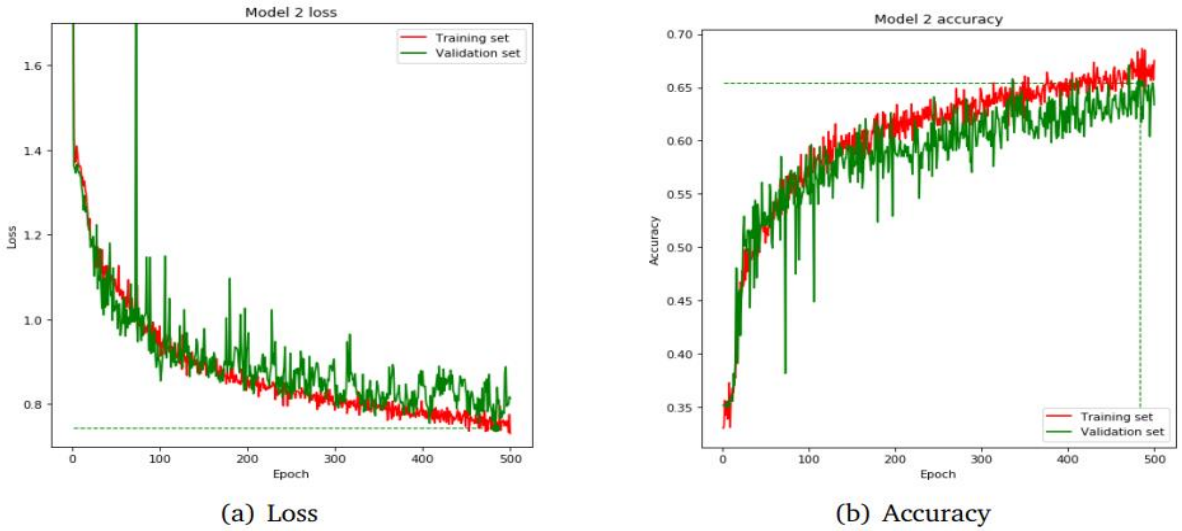


FIGURE 3.15 : 4-CLASSES CNN 2 CONFUSION MATRIX.



(a) Loss (b) Accuracy

FIGURE 3.16 : SCRATCH CNN 2 (4-CLASSES).

La précision globale est meilleure que dans le modèle précédent (61,3% contre 57,7%). Même après un très long entraînement, la perte diminuait encore ; augmenter le taux d'apprentissage serait ici un geste risqué, car l'évolution est déjà assez bruyante.

Après plus de 480 époques, le NN a atteint une précision de classification de 61,3%. Il est intéressant de noter comment le modèle est plus enclin à classer les anomalies bénignes

comme malignes plutôt que l'inverse. Dans un contexte médical, cela pourrait être une caractéristique souhaitable, car les faux négatifs sont potentiellement plus nocifs que les faux positifs.

5.4. Modèle 4 « Composite de modèle 1 et modèle 2 »

La tâche en 4 classes peut être décomposée en deux sous-tâches : classification de masse/calcification et diagnostic bénigne/maligne. L'idée est de former deux modèles indépendamment, chacun se concentrant sur une seule sous-tâche ; les deux réseaux fonctionneront alors en parallèle, chacun produisant la moitié de la prédiction complète à 4 classes. La logique derrière cette approche de « division et impera » est que chaque réseau se spécialise pour résoudre un sous-problème plus petit et potentiellement plus facile, obtenant ainsi de meilleurs résultats que de traiter l'ensemble du problème d'un seul coup (figure 3.17).

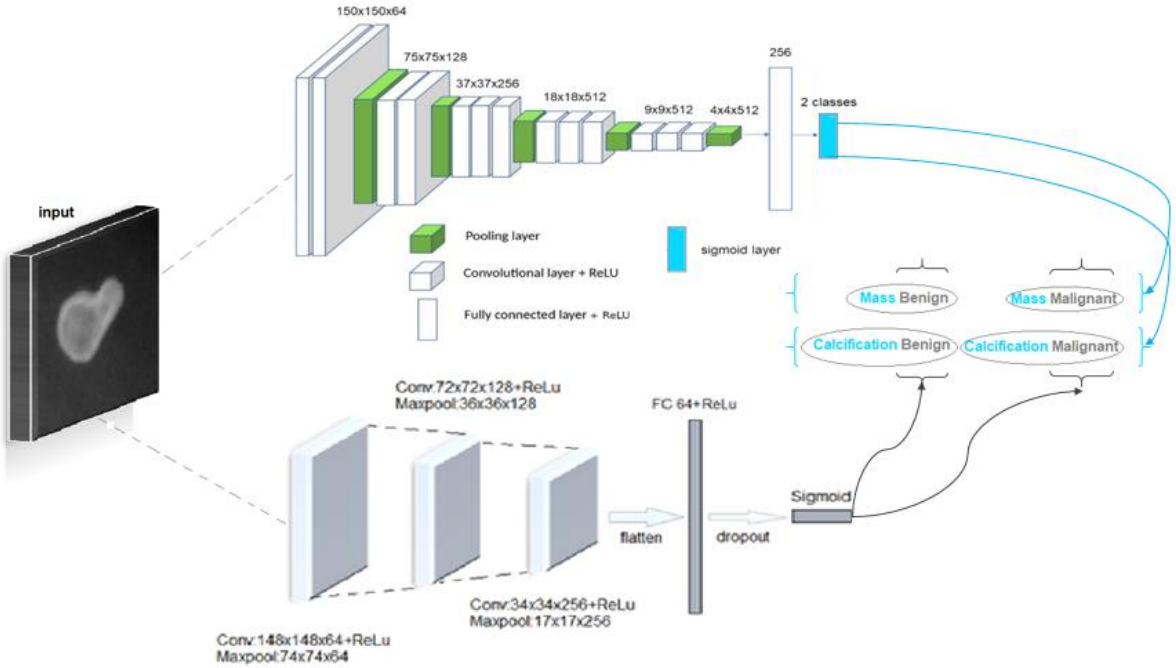


FIGURE 3.17 : ARCHITECRTURE DU CLASSIFIEUR COMPOSITE 4-CLASSES.

Le modèle de classification masse/calcification est le réseau VGG16 fine-tuned 2 layers de la section 5.1.3., avec une précision (accuracy) de 91,37% sur l'ensemble de test.

Le modèle de classification bénigne-maligne est plutôt un CNN avec 3 blocs convolutifs (64, 128, 256 canaux respectivement), 64 « fully-connected » et un dropout de 0.5 de la section 5.2. Formé depuis plus de 400 époques, il a une précision de 69% pour distinguer les anomalies bénignes et malignes sur l'ensemble de test.

5.4.1. Les résultats du modèle 4

La prédiction finale est simplement la concaténation des réponses des deux réseaux. Par exemple, si le premier classificateur produit « mass » et le second dit « benign », l'étiquette finale sera « mass benign ».

En pratique, l'adoption de cette stratégie se traduit par une précision de **62,5%** sur l'ensemble de test, ce qui représente une amélioration par rapport aux modèles tout-en-un à 4 classes (était de 61,3% comme décrit dans la section 5.3.3).

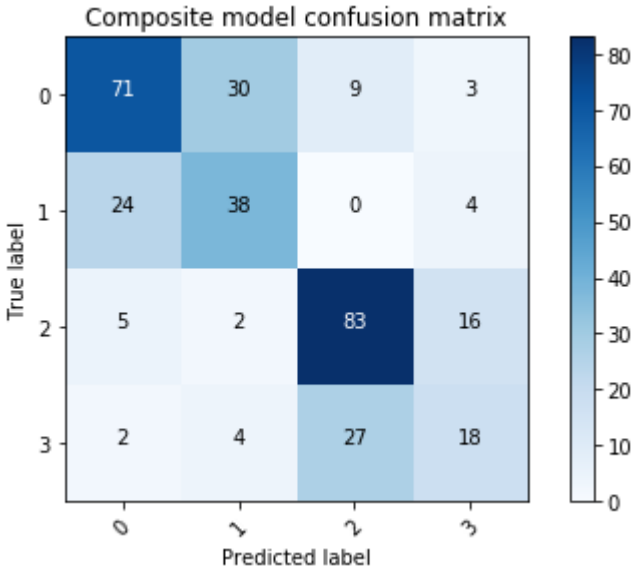


FIGURE 3.18 : COMPOSITE 4-CLASSES CLASSIFIEUR CONFUSION MATRIX.

5.5. Modèle 5 « Mask R-CNN »

Pour une image donnée le masque R-CNN, en plus de l'étiquette de classe et des coordonnées de la boîte englobante pour chaque objet, renverra également le masque d'objet. Donc il a trois sorties : pour chaque objet candidat, une étiquette de classe et la boîte englobante « bounding-box » ; la troisième sortie est le masque d'objet comme on l'a présenté dans le chapitre 2 (voir section 3.6.5.). L'architecture de ce modèle proposé est décrite dans la figure 3.19.

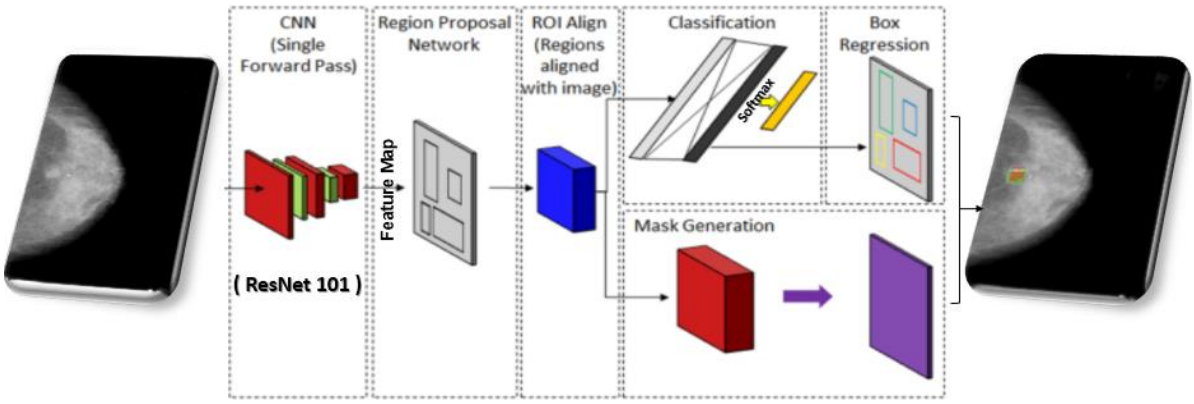


FIGURE 3.19 : ARCHITECTURE MASK R-CNN PROPOSÉE.

5.5.1. Backbone

Une « *Deep Convolutional Network* » architecture pour extraire les « *features* » de l'image. L'image d'entrée est introduite dans un CNN, souvent appelé backbone, qui est un réseau pré-entraîné dans notre architecture est ResNet101. Les couches de classification (entièrement connectées FC) du réseau backbone sont supprimées afin de l'utiliser comme

extracteur de caractéristiques. Cela rend également le réseau entièrement convolutif « *fully convolutional* », il peut donc prendre n'importe quelle image de taille d'entrée. Ces « *features*» agissent comme une entrée pour la couche suivante.

5.5.2. RPN

Réseau de proposition de région (RPN) pour proposer des boîtes de délimitation d'objets candidats. Autrement dit « *Region Proposal Network* » propose des régions d'intérêt (RoI) à partir des cartes de caractéristiques « *feature maps* » produites par le réseau de neurones convolutifs, les rectangles en pointillés ci-dessous (figure 3.20) sont ces propositions.

Il utilise un CNN pour générer plusieurs régions d'intérêt (RoI) à l'aide d'un classificateur binaire léger. Il le fait en utilisant 9 boîtes d'ancrage « *anchors boxes* » sur l'image. Le classificateur renvoie les scores d'objet / non-objet. La « *Non-Max suppression* » est appliquée aux « *Anchors* » avec un score d'objectivité élevé. Pour prédire plusieurs objets ou plusieurs instances d'objets dans une image, « *Mask R-CNN* » fait des milliers de prédictions. La détection finale des objets est effectuée en supprimant les boîtes d'ancrage qui appartiennent à la classe d'arrière-plan et les autres sont filtrées par leur score de confiance.

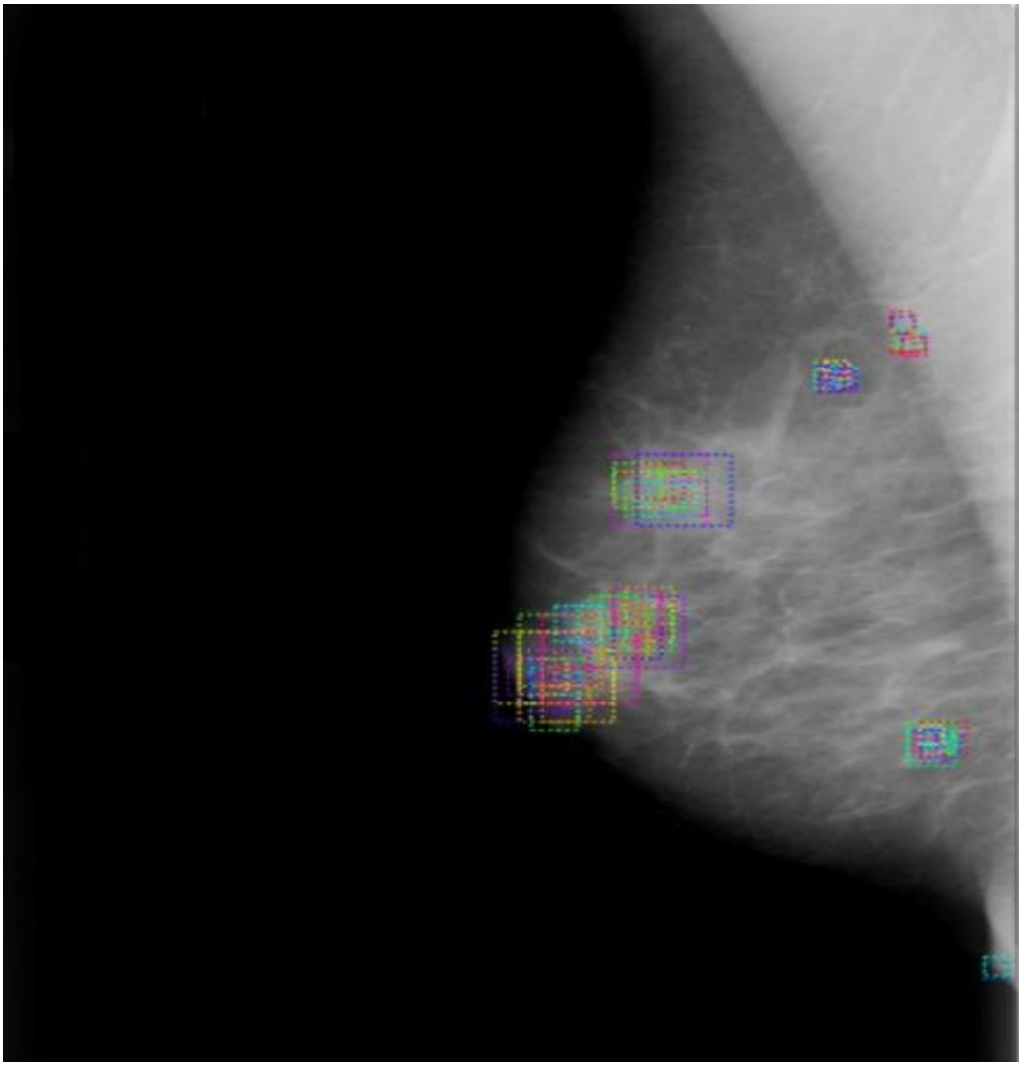


FIGURE 3.20 : RÉGIONS D'INTÉRETS (ROI) GÉNERÉE PAR LE (RPN) AVEC PLUSIEURS « ANCHOR BOXES ».

5.5.3. RoI Align et Classification

La suppression non maximale supprimera toutes les boîtes englobantes où IoU est inférieur ou égal à 0,5. Sélectionnez la zone de délimitation avec la valeur la plus élevée pour IoU et supprimez les autres zones de délimitation pour identifier le même objet. IoU (Intersection over Union) calcule l'intersection sur l'union des deux boîtes englobantes, « *ground truth bounding box* » et « *predicted bounding box* » par l'algorithme.

Le réseau RoI Align génère plusieurs « *bounding boxes* » au lieu d'un seul défini et les déforme dans une dimension fixe.

Les « *warped features* » sont ensuite introduites dans des couches entièrement connectées « *FC layers* » pour effectuer une classification à l'aide de softmax et la prédiction du « *boundary box* » est encore affinée à l'aide du modèle de régression.

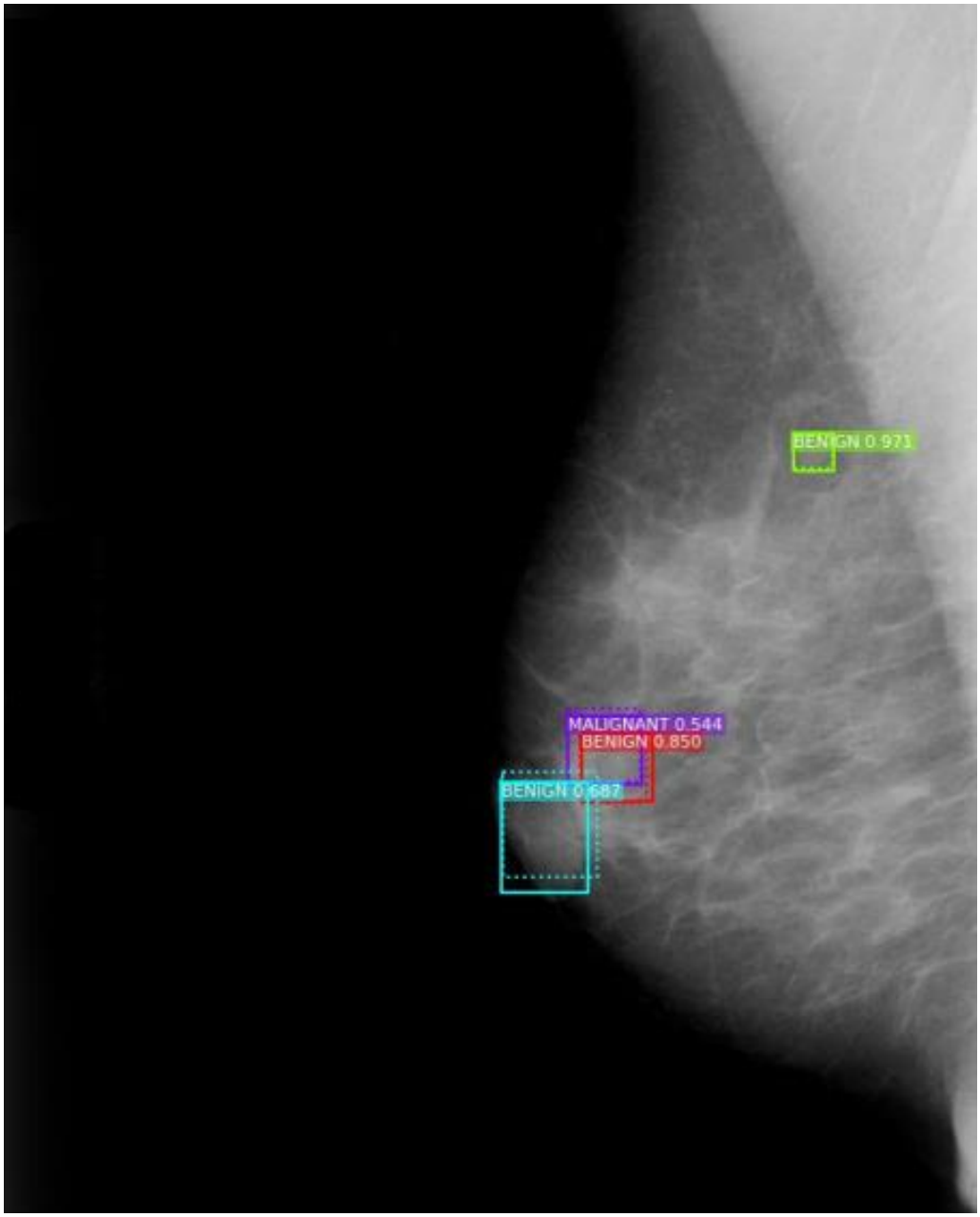


FIGURE 3.21 : RÉSULTATS DE LA DÉTECTION ET DE LA CLASSIFICATION.

5.5.4. Mask

Les « *warped features* » sont également introduites dans le classificateur de masque « *Binary mask classifier* » , qui se compose de deux CNN pour générer un masque binaire pour chaque RoI. « *Mask Classifier* » permet au réseau de générer des masques pour chaque classe sans concurrence entre les classes (figure 3.22).



FIGURE 3.22 : RÉSULTAT DE SEGMENTATION DE MASSE.

5.5.5. Évaluation du modèle mask R-CNN

Les « *ground truth* » sont représentées par la couleur verte, tandis que les prédictions sont représentées par la couleur rouge. Ceux sont quelques exemples sur les prédictions finales de « *Mask R-CNN* » :

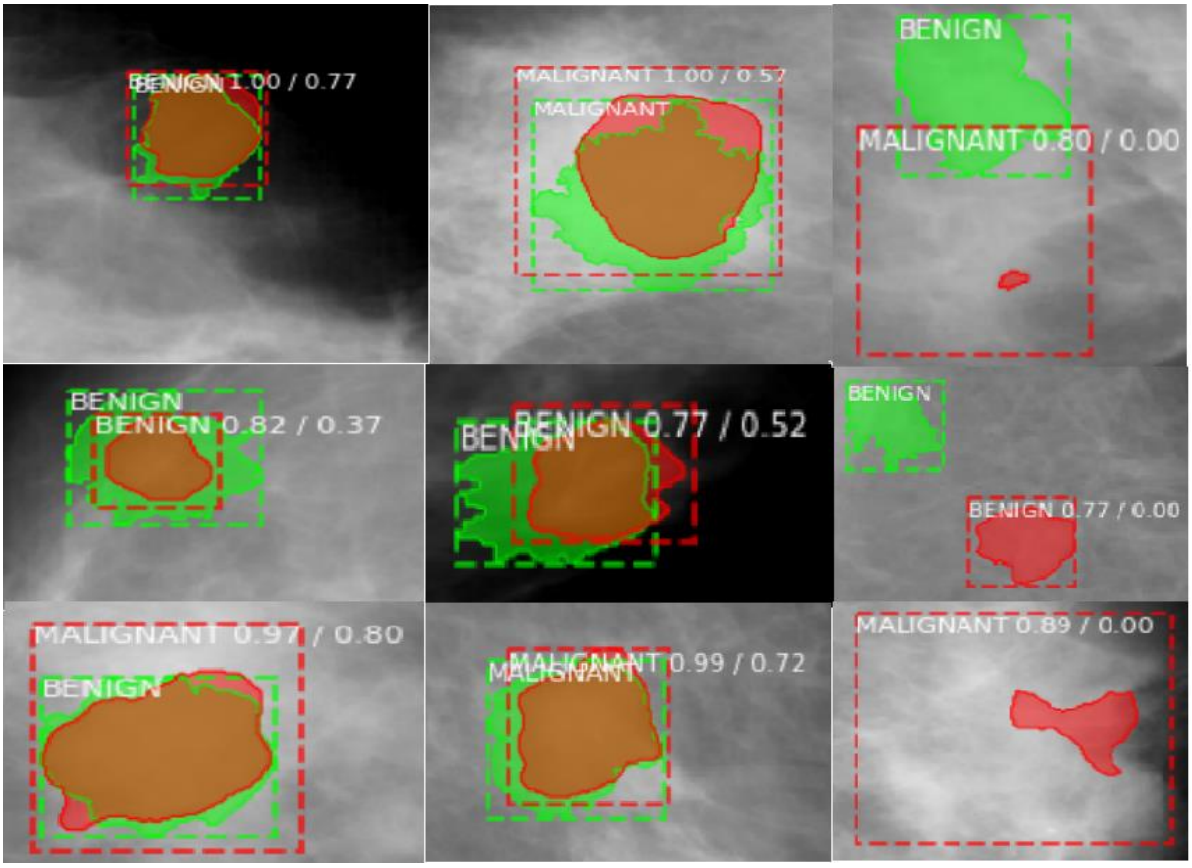


FIGURE 3.23 : GT ET PRÉDICTIONS FINALES DES MASSES AVEC MASK R-CNN, GT=GREEN, PRED= RED, CAPTIONS: SCORE/IOU.

i. Perte de formation (Training loss)

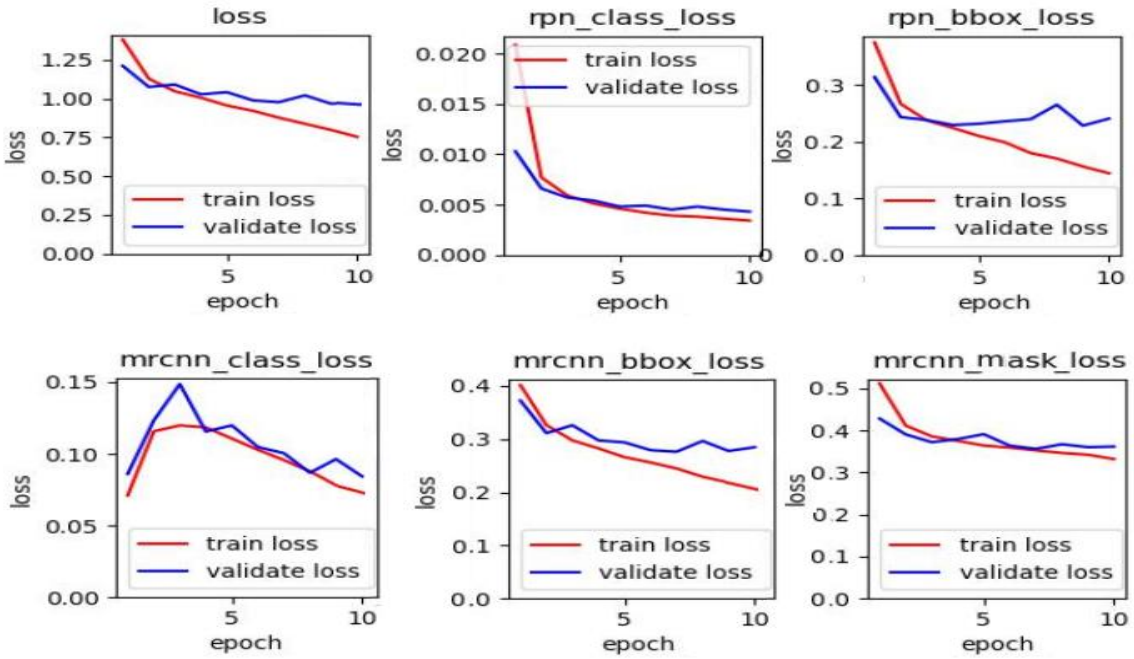


FIGURE 3.24 : TRAINING LOSS.

ii. Résultat du test

La matrice de confusion de classification de masses a donnée :

(TP=95, FP=66, FN=23, TN=72)

On va choisir « *Dice coefficient* » comme métrique pour le problème de segmentation d'image car Dice ainsi que Jaccard aussi sont les meilleures mesures pour découvrir la similitude entre la « *ground truth* » et l'image segmentée. Ce coefficient a été développé à l'origine pour les données binaires. La métrique Intersection over Union (IoU), également appelée indice Jaccard, est essentiellement une méthode pour quantifier le pourcentage de chevauchement entre le masque cible et notre sortie de prédiction. Cette métrique est étroitement liée au « *Dice coefficient* ». D'autres métriques d'évaluation du modèle *« Mask R-CNN* » sont aussi calculées.

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)} \quad IoU = \frac{Dice}{2 - Dice}$$

Modèle Mask R-CNN					
Dice Similarity Coefficient	**Mass segmentation IoU**	**Sensitivity (TPR)**	**Specificity**	**F1 Score**	**Test Accuracy**
68.10%	51.63%	80.51%	52.17%	68.09%	65.23%

TABLEAU 3.11: RÉSULTATS MASK R-CNN.

6. Discussion des résultats

Dans cette section nous discutons les résultats obtenus par notre travail, une approche basée Deep Learning pour la détection du cancer du sein.

Le meilleur modèle pour la 1ère catégorie à 2 classes (Mass/Calcification) a obtenu une précision de **91,37%** sur l'ensemble de test.

Et celui à 2 classes (Benign/Malignant) a obtenu une précision de **69,05%** sur l'ensemble de test.

L'adoption de la stratégie composite de 2 modèles pour la 2ème catégorie à 4 classes se traduit par une précision de **62,5%** sur l'ensemble de test.

L'utilisation brûte de Mask R-CNN pour la segmentation et la détection des masses est nouvelle et en tant qu'évaluation des performances, la précision est de **65,23%** sur l'ensemble de test avec une sensitivité (TPR) de **80,51%**. Dans l'évaluation de la segmentation des régions du cancer du sein, l'indice Dice est de **68.10%.** Ces résultats indiquent que notre schéma proposé peut être utile pour la détection du cancer du sein.

En comparant les résultats avec ceux présentés dans de nombreux articles prenant CBIS-DDSM comme dataset, les modèles ont atteint une précision de pointe[44]–[46]. Par

exemple le travail 2019 de Nhung Le, Nithish Addepalli et Ravi Choudhary nommé « *Breast Cancer Detection Through High Resolution Screening Mammograms* » de Center For Data Science, New York University, a obtenu une meilleure précision de 60% avec 4 classes et 74% pour 2 classes[12].

7. Conclusion

Ce chapitre décrit en détail notre contribution dans le domaine de la détection du cancer du sein en utilisant de différents modèles d'apprentissage profond « *Deep Learning* ». Afin de chercher un taux de détection acceptable ; nous avons testé cinq architecture qu'on a classé en deux catégorie de classes. La première catégorie contient les solutions ayant deux classes de sorties, quant à la deuxième elle contient les solutions offrant quatre classes pour détecter le type d'anomalies masses ou bien calcifications ainsi que leur diagnostique bénigne ou maligne.

Les résultats obtenus sont encourageants par rapport aux travaux existants car on a obtenu un taux de précision très encourageant dans les deux catégories. Dans la première catégorie on a obtenus un taux de 91,37% et un taux de 62,5% dans la deuxième catégorie qui est meilleure par rapport à un travail de Nhung Le, Nithish Addepalli et Ravi Choudhary[12].

Les solutions proposées sont testées et évaluées en utilisant une dataset de référence celle de CBIS-DDSM « *Curated Breast Imaging Subset of Digital Database for Screening Mammograph* » ; cette base a subi un processus de prétraitement consistant.

Les résultats obtenus nous encouragent à préparer un article scientifique pour le soumettre à un journal indexé.