



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Mouloud Mammeri de Tizi-Ouzou
Faculté De Génie Electrique et d'Informatique Département de
Génie Biomédical

Mémoire de fin d'études
En vue de l'obtention du diplôme de MASTER en
Génie Biomédical

Spécialité : Instrumentation Biomédicale

Présenté par :

Boudjemai Melissa

**Utilisation de l'apprentissage profond pour identifier des anomalies dans
les images médicales : Application à la mammographie**

Soutenu publiquement, le 01 /10/ 2025, devant les jurys composés de :

Mr MOUAS L'Hocine	MCB UMMTO	PRESIDENT
Mr MEDDOUR Cherif	MCB UMMTO	EXAMINATEUR
Mr BOURKACHE Nouredine	MCB UMMTO	PROMOTEUR

Dfca chꞑcb : 2024-2025

Remerciements

Je tiens à remercier en premier, ALLAH le tout puissant de m'avoir donné la volonté et le courage de percer et d'aboutir dans mes études.

Nous tenons à remercier toutes les personnes qui nous ont aidés lors de la rédaction de ce mémoire.

J'exprime mes vifs remerciements à mon directeur de mémoire Mr BOURKACHE Nouredine qui a suivi attentivement et activement mon travail, aussi pour sa patience, sa disponibilité, ses encouragements et ses conseils précieux. Je m'estime chanceuse d'avoir l'occasion de travailler avec vous dans le cadre de mon mémoire de fin d'étude, je vous remercie infiniment.

Je remercie également les membres du jury pour avoir accepté d'examiner ce travail, et j'adresse mes sincères remerciements à tout le personnel qui m'ont aidé à la réalisation de ce travail.

Et enfin, mes sincères remerciements s'adressent à tous les enseignants du département de génie biomédical de la faculté de génie électrique et d'informatique MOULOUD MAMMERI TIZI OUZOU, qui m'ont suivi le long de mon parcours d'études ainsi que tous les étudiants de ma promotion 2024/2025.

Dédicaces

Avant tout, je rends grâce à Dieu Tout-Puissant, pour Sa protection, Sa sagesse et Sa lumière qui m'ont guidée à chaque étape de ce parcours.

Je dédie également ce travail à moi-même, pour ma force, ma persévérance et ma détermination à surmonter les épreuves.

À mes chers parents, qui sont mon refuge, ma plus grande source d'amour et de sacrifices. Rien de tout cela n'aurait été possible sans votre soutien inconditionnel, vos prières et vos encouragements constants. Ce mémoire est avant tout le vôtre.

À toute ma famille, qui représente pour moi un cocon de tendresse, de force et de valeurs. Merci pour votre présence, vos conseils, vos prières et votre confiance en moi. Vous êtes ma base solide et mon inspiration de chaque jour.

À mes amis, pour leurs encouragements, leur patience et les moments de bonheur partagés.

*En particulier, à **Thamila et Ikram**, pour leur soutien précieux, leur amitié sincère et leur place unique dans mon cœur.*

À ma binôme Lamia, je te souhaite plein de courage, de réussite et de belles victoires dans l'avenir.

SOMMAIRE

Liste des figures

Liste des tableaux

Liste des abréviations

Introduction générale.....2

Chapitre I : L'imagerie médicale

I.1. Introduction.....3

I.2. Historique de l'imagerie médicale.....3

I.3. Principe de l'imagerie médicale.....4

I.4. Bases physiques de l'imagerie médicale.....5

I.5. Architecture générale d'un système d'imagerie.....5

- I.5.1. Source d'énergie.....6
- I.5.2. Détecteurs et capteurs.....10
- I.5.3. Chaîne d'acquisition, traitement et affichage.....10

I.6. Les principales techniques d'imagerie médicale.....11

- I.6.1. Imagerie par résonance magnétique (IRM)11
- I.6.2. Échographie.....12
- I.6.3. Fluoroscopie.....13
- I.6.4. Tomodensitométrie (TDM)13
- I.6.5. Panoramique dentaire.....14
- I.6.6. Mammographie.....16

I.7. Conclusion.....17

Chapitre II : La Mammographie

II.1. Introduction.....18

II.2. Fonctionnement de la mammographie.....19

- II.2.1. Processus et étapes de la mammographie.....19
- II.2.2. Principe de fonctionnement d'un mammographe.....20

II.3. Interprétation et classification des images mammographiques.....21

- II.3.1. Classification des mammographies en six catégories (BI-RADS)22

II.4. Types de lésions mammaires.....	23
• II.4.1. Kystes du sein.....	23
• II.4.2. Tumeurs non cancéreuses.....	23
• II.4.3. Microcalcifications.....	24
• II.4.4. Masse solide.....	24
• II.4.5. Tumeurs cancéreuses.....	24
II.5. Le cancer du sein.....	25
• II.5.1. Statistiques épidémiologiques du cancer du sein : données mondiales et régionales.....	25
• II.5.2. Anatomie du sein.....	27
• II.5.3. Types de cancer du sein.....	28
• II.5.4. Dépistage et diagnostic du cancer du sein.....	29
II.6. Technologies avancées en mammographie.....	29
• II.6.1. Imagerie améliorée par IA.....	29
• II.6.2. Mammographie 3D.....	30
• II.6.3. Biopsie liquide.....	30
• II.6.4. Dépistage par thermographie.....	30
• II.6.5. Échographie sans fil.....	31
II.7. Conclusion.....	32

Chapitre III : L'Apprentissage profond dans l'Analyse des Images Mammographiques

III.1. Introduction.....	33
III.2. L'intelligence artificielle.....	34
• III.2.1. Définition.....	34
• III.2.2. Principe de l'intelligence artificielle.....	34
• III.2.3. Applications de l'IA dans le domaine de la santé.....	34
III.3. L'apprentissage automatique et l'apprentissage profond.....	35
• III.3.1. Le Deep Learning.....	35
• III.3.2. Domaines d'applications.....	36
III.4. Les types d'apprentissage automatique.....	36
• III.4.1. Apprentissage supervisé.....	36
• III.4.2. Apprentissage non supervisé.....	36
• III.4.3. Apprentissage semi-supervisé (hybride).....	37

III.5. Les réseaux de neurones.....	37
• III.5.1. Différents types de réseaux.....	39
○ III.5.1.1. Réseaux neuronaux convolutifs (CNN)	39
○ III.5.1.2. Réseaux neuronaux récurrents (RNN)	40
• III.5.2. Apprentissage des réseaux de neurones.....	40
• III.5.3. Domaines d'application des réseaux de neurones.....	41
III.6. Applications en mammographie.....	42
• III.6.1. Algorithmes et modèles profonds utilisés.....	43
○ III.6.1.1. CNN classiques	43
○ III.6.1.2. Réseaux pré-entraînés	43
○ III.6.1.3. Modèles récents	43
• III.6.2. Architecture U-Net.....	44
○ III.6.2.1. Méthode U-Net.....	44
○ III.6.2.2. Applications biomédicales de U-Net.....	44
○ III.6.2.3. Architecture détaillée.....	45
III.7. Bases de données et études de référence en mammographie.....	45
III.8. Défis actuels et limites liés aux données et à l'IA.....	47
III.9. Perspectives Futures.....	48
III.9. Conclusion.....	50
 Chapitre IV : Implémentations et résultats	
IV.1. Introduction.....	51
IV.2. Les outils et environnements de travail.....	51
• IV.2.1. Python sur PC local.....	51
• IV.2.2. Visual Studio Code (VS Code)	52
• IV.2.3. Google Colab.....	53
• IV.2.4. Kaggle Notebooks.....	54
IV.3. Bases de données utilisées.....	55
• IV.3.1. MIAS.....	55
• IV.3.2. DDSM.....	55
• IV.3.3. INBreast.....	56
• IV.3.4. Bases fusionnées.....	56

IV.4. Modèles et méthodes.....	56
• IV.4.1. Présentation des modèles testés.....	56
• IV.4.2. Adaptation du modèle DenseNet121.....	58
• IV.4.3. Paramètres d'entraînement.....	59
IV.5. Évaluation et métriques.....	59
IV.6. Résultats expérimentaux.....	61
• IV.6.1. Résultats par base.....	61
• IV.6.2. Comparaison des environnements (Colab vs Kaggle vs PC local).....	64
IV.7. Tests de performances sur un ensemble de mammographies.....	65
IV.8. Conclusion.....	69
Conclusion générale.....	70
Résumé	

Liste d'abréviations

Chapitre I : Imagerie médicale

- **2D** : *Two-Dimensional* (imagerie bidimensionnelle)
- **3D** : *Three-Dimensional* (imagerie tridimensionnelle)
- **DICOM** : *Digital Imaging and Communications in Medicine* (norme internationale pour le stockage et la transmission d'images médicales)
- **NFS 99-300** : Norme Française de Santé 99-300 (norme relative aux systèmes d'imagerie médicale en France)

Chapitre II : La Mammographie

- **BRCA1** : *Breast Cancer Gene 1* (gène de prédisposition au cancer du sein et de l'ovaire)
- **BRCA2** : *Breast Cancer Gene 2* (gène de prédisposition au cancer du sein et de l'ovaire)

Chapitre III: L'Apprentissage profond dans l'Analyse des Images Mammographiques

- **ADN** : Acide Désoxyribonucléique
- **AlexNet** : Réseau de neurones convolutifs développé par Alex Krizhevsky, lauréat d'ImageNet 2012
- **CBAM** : *Convolutional Block Attention Module* (module d'attention pour réseaux convolutifs)
- **CBIS-DDSM** : *Curated Breast Imaging Subset of the Digital Database for Screening Mammography* (sous-ensemble annoté de la base DDSM)
- **CNN** : Convolutional Neural Network (Réseau de neurones convolutif)
- **GPU** : *Graphics Processing Unit* (processeur graphique utilisé pour l'accélération des calculs)
- **GRU** : *Gated Recurrent Unit* (variante simplifiée du LSTM)
- **INBreast** : Base de données de mammographies numériques annotées
- **JPG** : Joint Photographic Experts Group (Format d'image compressée)
- **LeNet** : Réseau de neurones convolutionnel développé par Yann LeCun pour la reconnaissance de chiffres manuscrits
- **LJPEG** : *Lossless JPEG* (format d'image compressé sans perte utilisé dans certaines bases médicales)
- **LSTM** : *Long Short-Term Memory* (type de réseau de neurones récurrents)
- **Mask R-CNN** : *Mask Region-Based Convolutional Neural Network* (réseau de neurones convolutifs pour la segmentation d'images)
- **MIAS** : *Mammographic Image Analysis Society* (base de données d'images mammographiques)
- **PNG** : *Portable Network Graphics* (format d'image compressé sans perte)
- **SENet** : *Squeeze-and-Excitation Network* (réseau intégrant des blocs d'attention par recalibrage de canaux)
- **VGGNet** : *Visual Geometry Group Network* (réseau convolutif profond développé à l'université d'Oxford)

- **VinDr-Mammo** : *Vietnamese Digital Mammography Dataset* (base de données vietnamienne de mammographies)

Chapitre IV : Implémentations et résultats

- **AUC** : Area Under the Curve (Aire sous la courbe)
- **FN** : False negatives (Faux négatifs)
- **FP** : False positives (Faux positifs)
- **IDE** : Integrated Development Environment (Environnement de développement intégré)
- **RAM** : Random Access Memory (Mémoire vive)
- **ROC** : Receiver Operating Characteristic (Caractéristique de performance de réception)
- **SGD** : Stochastic Gradient Descent (Descente de gradient stochastique)
- **SSD** : Solid State Drive (Disque à état solide)
- **Test**: Test set (Ensemble de test)
- **TN** : True negatives (Vrais négatifs)
- **TP** : True positives (Vrais positifs)
- **Train** : Training set (Ensemble d'entraînement)
- **UHD** : Ultra High Definition (Ultra haute définition)
- **Val** : Validation set (Ensemble de validation)

Listes des figures :

Chapitre 1 : Imagerie médicale

FIGURE I.1: TUBE DE COOLIDGE.....	6
FIGURE I.2: LE SCHEMA DE TUBE DE COOLIDGE	7
FIGURE I.3: SPECTRE GLOBAL D'EMISSION DES RAYONS X.....	8
FIGURE I.4: ELECTRON DIFFUSE	8
FIGURE I.5: : TRANSITIONS ELECTRONIQUES ET EMISSION DE RAYONS X CARACTERISTIQUES SUITE A UNE INTERACTION ELECTRON-ELECTRON.....	10
FIGURE I.6: REPARTITION DES TECHNIQUES D'IMAGERIE SELON LA FREQUENCE DU RAYONNEMENT SELON LA FREQUENCE.....	11
FIGURE I.7: APPAREIL D'IRM	12
FIGURE I.8: APPAREIL D'ECHOGRAPHIE	12
FIGURE I.9:APPAREIL DE FLUOROSCOPIE.....	13
FIGURE I.10: APPAREIL DE TDM.....	14
FIGURE I.11: APPAREIL PANORAMIQUE DENTAIRE	14
FIGURE I.12: APPAREIL MAMMOGRAPHE.....	16

Chapitre 2 : La mammographie

FIGURE II.1:VUE FRONTALE DE L'APPAREIL MAMMOGRAPHIQUE	19
FIGURE II.2:VUE OBLIQUE DE L'APPAREIL MAMMOGRAPHIQUE.....	19
FIGURE II.3:DETAILS PRECIS DES COMPOSANTS D'UN COMPRESSEUR DE MAMMOGRAPHE POUR DES EXAMENS MEDICAUX DE HAUTE PRECISION.....	20
FIGURE II.4:SCHEMA ANNOTE D'UN MAMMOGRAPHE NUMERIQUE ET DE SES PRINCIPALES COMPOSANTES	20
FIGURE II.5: UN KYSTE DU SEIN	23
FIGURE II.6: LA TUMEUR NON CANCEREUSE	23
FIGURE II.7: UN MICRO-CLACIFICATION	24
FIGURE II.8:UNE MASSE SOLIDE	24
FIGURE II.9: TUMEURS CANCEREUSES.....	24
FIGURE II.10:MAMMOGRAPHIE MONTRANT UN CANCER DU SEIN A DROITE.	25
FIGURE II.11: LA STRUCTURE DU SEIN	27
FIGURE II.12 : ECHOGRAPHIE SANS FIL.....	31

Chapitre III : L'Apprentissage profond dans l'Analyse des Images Mammographiques

FIGUREIII.1 : APPRENTISSAGE PROFOND ET IA.....	33
FIGUREIII.2 : STRUCTURE GENERALE D'UN RESEAU DE NEURONES.....	37
FIGUREIII.3 : STRUCTURE D'UN RESEAU DE NEURONES ARTIFICIELS.....	38
FIGUREIII.4 : FONCTIONNEMENT D'UN RESEAU NEURONAL EN PHASE D'APPRENTISSAGE.....	40

Chapitre IV : Implémentations et résultats

<i>Figure IV.1: architecture de model densnet121</i>	57
FIGURE IV.2 LA MATRICE DE CONFUSION	60
FIGURE IV.3COURBE ROC - INBREAST	62
FIGURE IV.4 : MATRICE DE CONFUSION SUR DDSM.....	64
FIGURE IV.5 : COURBE ROC SUR DDSM	65
FIGURE IV.6 : INTERFACE GRAPHIQUE DE CLASSIFICATION DE MAMMOGRAPHIES.....	66
FIGURE IV.7 : EXEMPLE DE PREDICTION AVEC 100% DE CONFIANCE.	66
FIGURE IV.8 : EXEMPLE DE PREDICTION AVEC 88% DE CONFIANCE.	67
FIGURE IV.9. EXEMPLE DE FAUX POSITIF (54% DE CONFIANCE)	67
FIGURE IV.10. EXEMPLE DE VRAI POSITIF (100% DE CONFIANCE)	68

LISTES DES TABLEAUX

Chapitre II : la Mammographie

Tableau II.1 les six catégories de mammographies selon la classification <i>Bi-Rads</i> de l'ACR.	22
--	-----------

Chapitre III : L'Apprentissage profond dans l'Analyse des Images Mammographiques

Tableau III.1 : comparatif des principales bases de données dédiées à l'imagerie mammographique.....	46
--	-----------

Chapitre IV : Implémentations et résultats

Tableau IV.1 : Résultats par base de données.....	61
---	-----------

Tableau IV .2 – Comparaison des environnements d'exécution (avec MIAS Dataset)	65
---	-----------

Les avancées technologiques dans le domaine médical ont profondément transformé la manière dont les maladies sont détectées et prises en charge. Parmi elles, le cancer du sein occupe une place particulière : il représente aujourd'hui l'une des principales causes de mortalité chez les femmes à travers le monde. Selon l'Organisation mondiale de la Santé (OMS, 2023), il s'agit du cancer le plus fréquemment diagnostiqué dans cette population, ce qui en fait un enjeu majeur de santé publique. Dans ce contexte, la détection précoce et l'établissement d'un diagnostic fiable constituent des étapes essentielles pour améliorer les chances de survie et optimiser les stratégies thérapeutiques.

L'imagerie médicale joue un rôle central dans ce processus. Parmi les outils disponibles, la mammographie demeure la méthode de référence pour le dépistage et le diagnostic du cancer du sein. Elle permet d'explorer la structure interne du tissu mammaire et de détecter d'éventuelles anomalies, telles que des tumeurs ou des microcalcifications. Toutefois, l'interprétation de ces images reste une tâche complexe : elle peut varier selon les radiologues, voire chez un même observateur, ce qui entraîne parfois des erreurs de diagnostic.

Pour pallier ces limites, les récents progrès de l'intelligence artificielle (IA), et plus particulièrement de l'apprentissage profond, offrent de nouvelles perspectives. Ces approches permettent de concevoir des modèles capables d'extraire automatiquement des caractéristiques pertinentes des images médicales, apportant ainsi une aide précieuse aux radiologues dans leur prise de décision. L'intégration de ces techniques dans le dépistage par mammographie apparaît ainsi comme une voie prometteuse pour accroître la précision, la rapidité et la fiabilité des diagnostics.

Ce mémoire s'intéresse à l'apport de l'apprentissage profond dans l'analyse des mammographies. Il s'organise en quatre chapitres : une présentation des bases de l'imagerie médicale, une description des principes et limites de la mammographie, une étude des concepts et applications du deep learning appliqués aux images, et enfin les expérimentations réalisées sur plusieurs bases de données, accompagnées d'une analyse des résultats obtenus.

À travers cette approche, ce mémoire vise à mettre en lumière les apports et les limites de l'apprentissage profond dans l'aide au diagnostic du cancer du sein, tout en ouvrant une réflexion sur les perspectives futures de l'imagerie médicale assistée par ordinateur.

Chapitre I

L'imagerie Médicale

I.1. Introduction

L'imagerie médicale regroupe un ensemble de techniques permettant de visualiser le fonctionnement du corps humain, que ce soit au niveau de la physiologie ou du métabolisme. Elle offre ainsi des images claires, compréhensibles et porteuses d'informations essentielles pour la médecine. Son objectif principal est d'aider à établir un diagnostic, à détecter d'éventuelles anomalies et à assurer le suivi de l'efficacité des traitements [1].

L'imagerie biomédicale, quant à elle, se situe à l'intersection de la biologie, de la médecine et de l'ingénierie. Elle met à profit des technologies comme la tomodensitométrie (TDM) ou l'imagerie par résonance magnétique (IRM) afin de diagnostiquer et de surveiller différentes pathologies, le tout de manière non invasive.

Ces méthodes, dont certaines existent depuis le début du XX^e siècle, ont profondément transformé la pratique médicale. Grâce aux avancées de l'informatique, elles permettent aujourd'hui d'explorer de façon indirecte mais précise l'anatomie, la physiologie et le métabolisme du corps humain. D'abord pensées comme des outils de diagnostic, elles sont désormais également utilisées en recherche biomédicale pour approfondir notre compréhension du fonctionnement de l'organisme.

Dans ce premier chapitre, nous proposerons un aperçu global de l'imagerie médicale, en retraçant son évolution historique depuis la découverte des rayons X jusqu'aux technologies les plus récentes. Nous présenterons aussi les principes généraux qui sous-tendent les différentes méthodes d'imagerie, en mettant en avant leurs bases physiques et les étapes essentielles allant de l'acquisition des images à leur affichage. Enfin, nous introduirons brièvement les principales techniques utilisées en pratique médicale, parmi lesquelles l'IRM, l'échographie, la fluoroscopie, la tomodensitométrie et la mammographie.

I.2. Historique de l'imagerie médicale :

L'histoire de la radiologie a commencé avec la découverte des rayons X par Wilhelm Roentgen en 1895. Au cours de sa carrière, ce physicien allemand a centré ses recherches sur des travaux comme la piézoélectricité, l'absorption et les chaleurs spécifiques des gaz, ainsi que les propriétés des électrons (tube de Crookes). Ces études ont conduit de manière inattendue à la découverte en 1895 de rayons invisibles capables de traverser la plupart des substances, laissant des

ombres qui pourraient être enregistrées sur des plaques photographiques. En raison de la nature inconnue de ces rayons à l'époque, Roentgen les a alors qualifiés de « Rayons X ».

Il a partagé ses découvertes extraordinaires dans son rapport intitulé "On a New Kind of Rays", qui a été publié le 28 décembre 1895. La nouvelle de sa découverte s'est rapidement répandue, en particulier dans le milieu médical. En février 1896, les applications cliniques ont commencé à être mises en œuvre. Le travail exceptionnel de Roentgen dans la découverte des fondements de ce qui allait devenir la radiologie moderne. Ses autres contributions scientifiques lui ont valu le premier prix Nobel de physique en 1901.

La radiographie la plus ancienne consistait à enregistrer des images sur des plaques photographiques en verre. Puis à partir de 1918 la radiographie fait un saut technologique avec l'introduction du film par le fondateur de la Eastman Kodak Company, George Eastman. Aujourd'hui, la radiographie utilise une technologie d'imagerie numérique et d'archivage pour enregistrer et stocker électroniquement des images radiographiques.

Les progrès technologiques au fil des années ont contribué à l'évolution continue de l'imagerie médicale. La médecine de la Seconde Guerre mondiale jusqu'aux années 1950 a propagé la technologie des ultrasons. Par la suite, les années 1970 ont vu l'arrivée d'autres méthodes d'imagerie, largement utilisées de nos jours, telles que la tomodensitométrie et l'imagerie par résonance magnétique (IRM) .

I.3. Principe d'imagerie médicale :

Le but de l'imagerie médicale est de créer une représentation visuelle intelligible d'une information à caractère médical. Cette problématique s'inscrit plus globalement dans un cadre technique : l'objectif est en effet de pouvoir représenter sous un format relativement simple une grande quantité d'informations issues d'une multitude de mesures acquises selon un mode bien défini .

L'image obtenue peut être traitée informatiquement pour obtenir par exemple :

- Une reconstruction tridimensionnelle d'un organe ou d'un tissu ;
- Un film ou une animation montrant l'évolution ou les mouvements d'un organe au cours du temps ;
- Une imagerie quantitative qui représente les valeurs mesurées pour certains paramètres biologiques dans un volume donné [2].

Dans un sens plus large, le domaine de l'imagerie médicale englobe toutes les techniques permettant de stocker et de manipuler ces informations. Ainsi, il existe des normes conçues pour assurer la qualité et la sécurité des examens d'imagerie, tout en protégeant les patients et les professionnels de santé. Comme par exemple : La norme NFS 99-300 (Démarche Qualité en Imagerie Médicale) et la norme DICOM qui un standard international pour la transmission, le stockage et le traitement des images médicales numériques[3].

I.4. Bases physiques de l'imagerie médicale :

L'imagerie médicale repose sur des principes physiques et technologiques permettant de produire des images du corps humain à des fins diagnostiques ou thérapeutiques. Elles s'appuient sur l'interaction entre différentes formes d'énergie telles que les rayons X, les ondes ultrasonores, les champs magnétiques ou encore les émissions radioactives et les tissus biologiques. Ces interactions, comme l'absorption, l'atténuation, la réflexion, la diffusion ou encore la résonance, fournissent des informations essentielles sur les structures anatomiques, les fonctions physiologiques ou la présence d'anomalies pathologiques. Pour cela, une énergie spécifique est envoyée dans le corps : elle doit être capable de traverser les tissus et d'interagir avec eux. Contrairement à la lumière visible, qui pénètre peu en profondeur et est réservée à l'observation externe (dermatologie, endoscopie), des rayonnements non visibles comme les rayons X, les ondes radio ou les ultrasons sont utilisés pour explorer l'intérieur du corps. Toutefois, le simple passage de l'énergie à travers les tissus ne suffit pas : C'est plutôt l'interaction entre l'énergie et les tissus qui rend possible la création d'une image exploitable. Enfin, bien que la qualité de l'image soit essentielle pour le diagnostic, elle doit toujours être obtenue dans le respect des limites de sécurité pour le patient. Par exemple, améliorer une image radiographique peut nécessiter une dose plus élevée de rayons X, une IRM plus précise peut demander un temps d'acquisition plus long, et une échographie plus nette peut exiger une intensité accrue des ultrasons. Ces choix techniques doivent donc toujours prendre en compte l'impact sur la santé et le confort du patient.

I.5. Architecture générale d'un système d'imagerie :

I.5.1. Source d'Énergie :

- **Ultrasons** : L'ultrason représente une onde mécanique et élastique, se déplaçant à travers des milieux fluides, solides, gazeux ou liquides . Les ultrasons sont utilisés en échographie pour envoyer des ondes ultrasonores à travers les tissus du corps et détecter les échos pour former des images en temps réel.
- **Champs Magnétiques** : Ces champs magnétiques, souvent générés par des aimants supra-conducteurs, sont indispensables pour créer des images de haute qualité en médecine. Les champs magnétiques sont utilisés en IRM pour polariser les noyaux atomiques et générer des images détaillées en 2D et 3D [4].
- **Les rayons X** : Les rayons X sont des radiations invisibles qui peuvent pénétrer dans le corps humain et rentrent en interaction avec celui-ci. L'image en radiologie est créée grâce à la différence d'atténuation du rayonnement X imposée par les divers éléments du corps humain (os, graisse, muscles, eau, air, vaisseaux...). Les rayons X sont une forme de rayonnement électromagnétique à haute fréquence constitué de photons dont l'énergie varie d'une centaine d'eV, à plusieurs MeV. La principale propriété des rayons X est de traverser la matière en étant partiellement absorbés en fonction de la densité de celle-ci et de l'énergie du rayonnement, ce qui permet d'avoir une information sur l'intérieur des objets qu'ils traversent. Les rayons X sont parmi les principales modalités de l'imagerie médicale et du contrôle non destructif des tissus [5].

➤ Production des rayons X

Les rayons X (RX) sont produits dans des tubes à RX également appelés tubes de Coolidge. Le principe est le suivant : des électrons émis par une cathode (un filament, le plus souvent en tungstène, chauffé par le passage d'un courant électrique) sont accélérés par une différence de potentiel élevée en direction d'une cible constituée d'une anode en métal (en tungstène également) [5]. Cela crée ensuite des RX, qui sont émis par la cible selon deux mécanismes :

-Le freinage des électrons par les atomes de la cible crée un rayonnement continu (Rayonnement de freinage ou Bremsstrahlung en allemand) dont une partie appartient au domaine des RX [5].



Figure I.1: Tube de Coolidge

- Les électrons accélérés ont une énergie suffisante pour exciter certains des atomes de la cible, en perturbant leurs couches électroniques internes. Ces atomes excités émettent des RX en retournant à leur état fondamental (Rayonnement de fluorescence) [5].

- La structure de base d'une source de RX : La figure 2 montre le principe d'émission des rayons X dans le tube dit de Coolidge

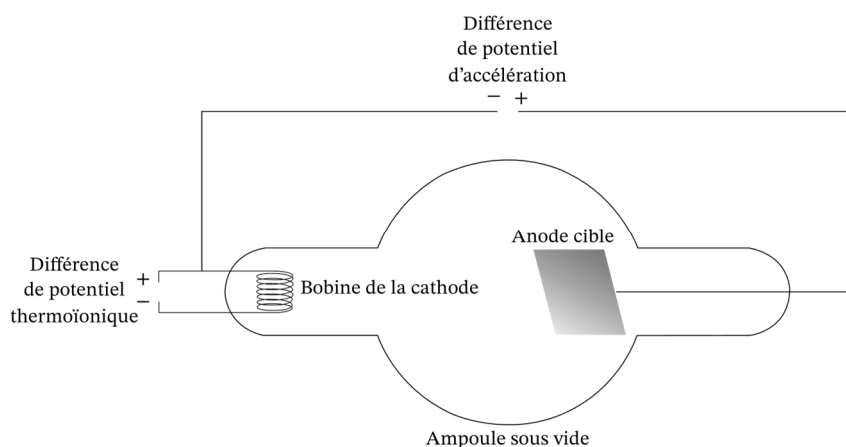


Figure I.2: Le schéma de tube de Coolidge

<https://www.nagwa.com/fr/explainers/725153535794/>

Le tube de Coolidge est composé des éléments essentiels suivants :

- Tube en verre ou en céramique : il constitue l'enceinte qui contient tous les autres éléments. Il doit être sous vide (l'air doit être éliminé pour qu'il ne constitue pas d'obstacle sur le chemin des électrons accélérés.)
- Une cathode qui constitue la source des électrons : Elle est sous la forme d'un filament qui sera chauffé par un courant de chauffage et qui va émettre les électrons par effet thermoïonique.

- Une coupelle de focalisation : Dont le rôle est d'empêcher les électrons émis par la cathode de se disperser dans toutes les directions.
- Une anticathode (anode) : Elle constitue la cible pour les électrons accélérés et où seront produits les RX.
- Un support pour la cible (Cuivre en général) dans lequel est incorporé un système de refroidissement qui va empêcher l'anode de fusionner sous l'effet de la chaleur dégagée par les électrons incidents.
- Une fenêtre en feuilles minces de béryllium : Elle a le rôle de laisser sortir les RX tout en gardant le tube étanche à l'air extérieur [6].

La figure 3 représente le spectre du rayonnement émis par le tube de Coolidge. En général c'est un spectre (global) composé de deux types de spectres, le spectre continu et le spectre de raies. Dans le cas où l'énergie cinétique des électrons accélérés n'est pas assez importante pour libérer des électrons des couches internes des atomes de la cible, alors on n'obtient qu'un spectre continu.

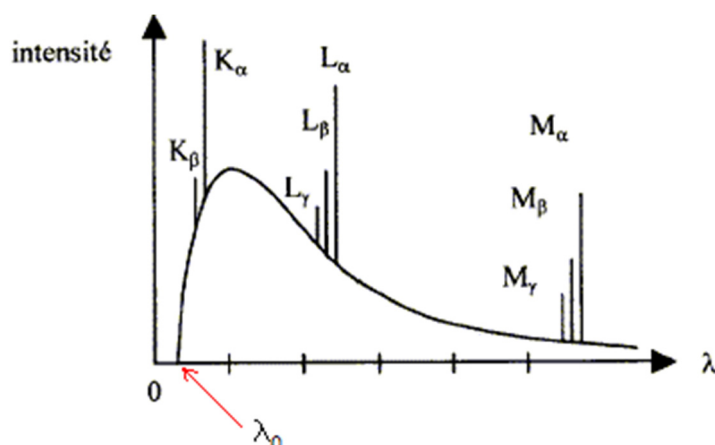


Figure 1.3: Spectre global d'émission des rayons X

<https://www.google.com/url?sa=i&url=http%3A%2F%2Fwww.ch>

Le spectre continu de rayon X : Lorsqu'un électron ayant une énergie cinétique initiale E_0 heurte (interagit avec) un des atomes de la cible, l'électron peut alors perdre une énergie E , qui peut se traduire par l'émission d'un photon de RX qui rayonne à partir du site de la collision. (Il y a très peu d'énergie transférée au recul de l'atome en raison de la masse relativement élevée de ce dernier, on peut donc la négliger)

L'électron diffusé de la figure 4, ayant maintenant une énergie inférieure à E_0 , peut heurter un autre atome de la cible, générant un deuxième photon, dont l'énergie sera généralement différente de celle du photon produit dans la première collision .

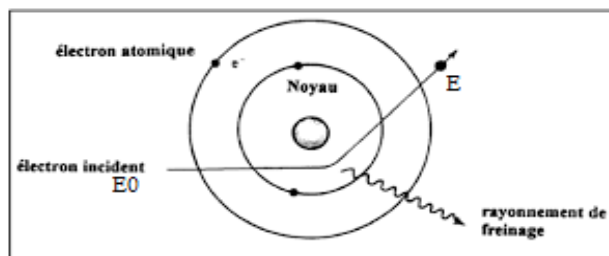


Figure I.4: interaction électron-noyau et Rayonnement de freinage

https://l2bichat2017-2018.weebly.com/uploads/1/1/2/5/112587633/p2_ue2_cours6_image-rie_par_rayons_x_ron%C3%A9o.pdf

Ce processus de diffusion peut continuer jusqu'à ce que l'électron perde toute son énergie. Chacun des photons générés par ces collisions forme une partie du spectre continu de RX.

Dans le schéma du spectre de la figure 3, on voit bien la longueur d'onde de seuil nettement définie $\lambda_0 = \lambda_{min}$ sous laquelle le spectre continu n'existe pas. Cette longueur d'onde minimale correspond à une unique collision frontale avec un atome cible, collision dans laquelle un électron incident perd toute son énergie cinétique initiale E_0 . Pratiquement toute l'énergie de l'électron est transférée d'un photon unique, dont la longueur d'onde associée (la plus petite longueur d'onde possible du RX) est déterminée par :

$$E_0 = E_{max} = h.f = h.c/\lambda_{min} \Rightarrow \lambda_{min} = h.c/E_{max}$$

La longueur d'onde de seuil minimal est totalement indépendante du matériau de la cible .

Le spectre de Raies : Si un faisceau d'électrons est dirigé vers une cible, les électrons sont susceptibles d'avoir une énergie suffisante pour arracher les électrons les plus liés au noyau des atomes constituant la cible. Si un électron d'un niveau inférieur est expulsé, les électrons des niveaux les plus élevés vont occuper la place laissée vacante en libérant l'excès d'énergie sous forme de photons de RX .

Concrètement, si des électrons de la couche K correspondant au niveau quantique $n = 1$ sont arrachés, les électrons d'état d'énergie plus élevée occupent la place laissée vacante conduisant à la production de raies caractéristiques $K\alpha$, $K\beta$, $K\gamma$

De même, si la place vacante est sur la couche L ($n = 2$), il y a production de raies caractéristiques $L\alpha$, $L\beta$, $L\gamma$... Le processus se répète pour toutes les couches M, N etc. Chacune de ces raies est en fait constituée de plusieurs raies très rapprochées correspondant à la structure des niveaux d'énergie atomique .

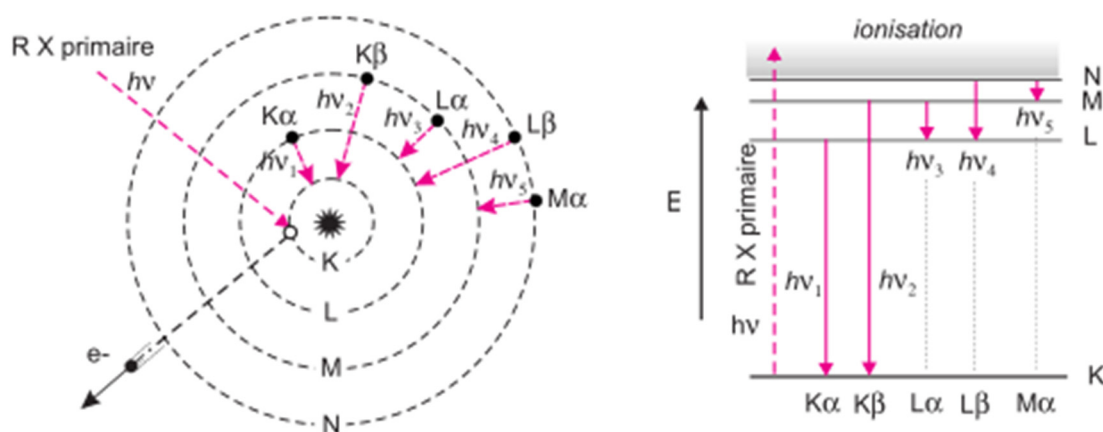


Figure I.5: : Transitions électroniques et émission de rayons X caractéristiques suite à une interaction électron-électron

<https://www.analyticaltoxicology.com/spectrometrie-de-fluorescence-x/>

I.5.2. Détecteurs et capteurs :

Les détecteurs et capteurs utilisés sont des composants essentiels qui captent les signaux émis par la source d'énergie (Rayons X, ultrasons, champs magnétiques) lorsqu'ils interagissent avec les tissus du corps. Ces détecteurs convertissent les signaux analogiques en signaux électriques qui peuvent être traités pour former des images médicales.

I.5.3. Chaîne d'acquisition, traitement et affichage :

Une fois que les signaux sont captés par les détecteurs, ils passent par une chaîne d'acquisition qui les convertit en données numériques. Ces données sont ensuite traitées par des algorithmes complexes pour améliorer la qualité de l'image, réduire le bruit et corriger les artefacts. Enfin, les images traitées sont affichées sur des moniteurs pour permettre aux médecins et aux radiologues d'analyser les structures anatomiques et de poser des diagnostics.

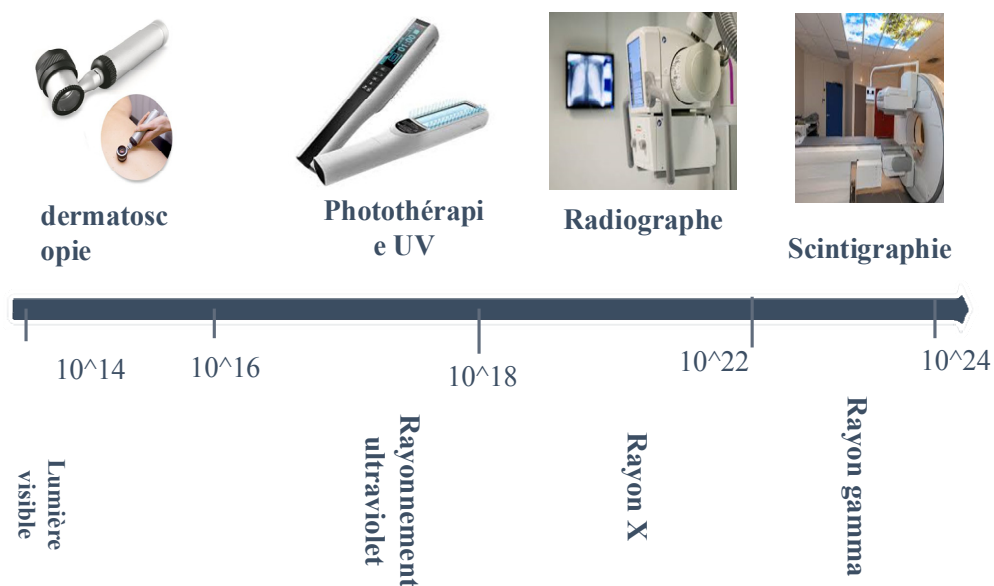


Figure I.6: répartition des techniques d'imagerie en fonction de la fréquence

I.6. Les principales techniques d'imagerie médicales :

I.6.1. Imagerie par résonance magnétique (IRM) :

IRM est une technique d'imagerie médicale permettant d'obtenir des vues en 2 et 3 dimensions de l'intérieur du corps humain de façon non-invasive avec une résolution en contraste relativement élevée. L'IRM repose sur le principe de la résonance magnétique nucléaire (RMN) qui utilise les propriétés quantiques des noyaux atomiques.

Dans l'RMN on distingue 4 applications différentes sont Angio-IRM, Imagerie du tenseur de diffusion, Spectroscopie RMN et Cholangio-IRM.

Depuis sa création les domaines d'application en imagerie médicale de L'IRM se sont considérablement étendus, particulièrement utilisé pour l'exploration des pathologies osseuse et articulation, cérébrales, ORL, gynécologique, vasculaires et cardiaques.

L'appareil d'IRM se compose d'un aimant principal produisant un champ magnétique intense, de bobines de gradient qui déterminent la position des signaux dans l'espace, et de bobines radiofréquence qui diffusent et captent les ondes requises pour créer l'image. Le patient est placé sur une table d'examen motorisée qui se déplace à travers le tunnel de l'appareil [7].



Figure I.7: Système IRM

I.6.2. Echographie :

L'échographie est une technique d'imagerie médicale non invasive permettant de visualiser en temps réel les structures internes du corps grâce à l'utilisation d'ultrasons.

Elle fonctionne par l'émission d'ondes sonores à haute fréquence, qui traversent les tissus et sont réfléchies différemment selon leur composition, permettant ainsi la formation d'images.

Parmi les principales indications figurent l'échographie obstétricale, abdominale, cardiaque (échocardiographie) et vasculaire (doppler).

Couramment utilisée en pratique clinique, l'échographie est devenue un outil de diagnostic de première ligne, notamment pour l'examen des tissus mous, du système cardiovasculaire, des organes abdominaux et pour le suivi des grossesses.

L'échographe est constitué des éléments suivants :

- Une sonde, permettant l'émission et la réception d'ultrasons ;
- Un système informatique, transformant le délai entre l'émission et la réception de l'ultrason en image ;
- Une console de commande, permettant la saisie des données du patient et les différents réglages ;
- Un système de visualisation : le moniteur ;
- un système d'enregistrement des données, soit de manière analogique (cassette vidéo, impression papier), soit de manière numérique (format DICOM) [8].

Le tout est disposé sur un chariot mobile, permettant d'effectuer l'examen au chevet même du patient.



Figure I.8: Appareil d'échographie

I.6.3. Fluoroscopie :

La fluoroscopie, ou radioscopie, est une modalité de la radiologie médicale qui consiste à acquérir en instantané des images dynamiques de l'intérieur des structures [9].



Figure I.9: Appareil de fluoroscopie

<https://fr.wikipedia.org/wiki/Fluoroscopie>

Les principaux composants comprennent un tube à rayons X, des filtres de mise en forme spectrale, un dispositif de restriction de champ (collimateur), une grille anti-diffusion, un récepteur d'images, un ordinateur de traitement d'images et un dispositif d'affichage. Les composants

auxiliaires, mais indispensables, comprennent un générateur haute tension, un support patient (table ou divan) et le matériel permettant de positionner la source de rayons X et le récepteur d'images par rapport au patient [10].

Les applications les plus courants sont : explorations digestives, Guidage de procédures interventionnelles, Arthrographie, Pédiatrie, Neuroradiologie interventionnelle, Applications dentaires et maxillo-faciales.

I.6.4. Tomodensitométrie (TDM) :

La TDM, plus communément appelée scanner, est une technique d'imagerie médicale qui permet d'obtenir des images en coupes fines de l'anatomie interne à l'aide de rayons X. Elle repose sur l'absorption variable des rayons X par les différents tissus du corps, associée à un traitement informatique permettant de reconstruire des images en 2 ou 3 dimensions.

Les principales applications du scanner incluent le scanner cérébral, thoracique, abdominal, osseux, ainsi que l'angioscanner destiné à l'exploration des vaisseaux sanguins.

Un appareil TDM se compose, avant tout, d'un tube rayonneur X, d'un détecteur, d'un anneau rotatif (Statif), d'une table d'examen motrice et d'un ordinateur réassembleur d'image. Au cours de l'examen, le tube X émet un rayon transversal et le faisceau parallèle à l'axe de rotation du scanner pendant que le statif tourne autour du patient. Les détecteurs reçoivent les rayons X affaiblis par les tissus irradiés et les transmettent au système informatique et les dispositifs numériques se chargent de composer des balances. Les coupes peuvent se faire de deux façons : 3D et 2D et cela facilite l'élaboration d'un angle intérieur du patient réaliste [11].



Figure I.10: Appareil de TDM

<https://www.sfendocrino.org/cushing-infos/examen-scanner.php>

I.6.5. Panoramique dentaire :

Le panoramique dentaire est une technique d'imagerie radiographique permettant d'obtenir une vue d'ensemble en deux dimensions des mâchoires, des dents, des articulations temporo-mandibulaires et des structures avoisinantes [12].

Elle fonctionne grâce à la rotation simultanée d'un émetteur de rayons X et d'un capteur autour de la tête du patient, ce qui permet de capturer l'ensemble de l'arcade dentaire en une seule prise [12].

Cette technique est couramment utilisée en dentisterie pour diagnostiquer les infections, les caries, les dents incluses, les anomalies de position ou encore pour évaluer la structure osseuse avant des traitements orthodontiques ou implantaires.

L'appareil de scanners panoramiques dentaires est composé de nombreux éléments technologiques fonctionnant ensemble de manière harmonieuse. Au cœur du système se trouve un tube émetteur de rayons X puissants et précis ainsi qu'un capteur numérique ultraperformant. Un bras motorisé fait pivoter cet attelage de manière orbitale autour de la tête du patient, qui est stabilisée confortablement grâce à un repose-menton ergonomique [13]. Pendant toute la durée de l'examen tomodensitométrie, des images panoramiques de haute définition des mâchoires, dents et sinus sont capturées sous tous les angles possibles. Ces clichés dentaires de pointe, aussi complets que précis, sont ensuite transmis à un logiciel médical permettant aux praticiens d'analyser avec acuité l'état de santé buccodentaire du patient.



Figure I.11: Appareil panoramique dentaire

I.6.6. Mammographie : La mammographie est une radiographie des seins qui utilise des rayons X à faible dose. Elle est réalisée grâce à un appareil appelé mammographe. Elle permet de visualiser en deux dimensions les structures mammaires afin de détecter d'éventuelles anomalies telles que des masses, des microcalcifications ou des lésions suspectes [14].

Principalement utilisée dans le cadre du dépistage du cancer du sein, notamment chez les femmes de plus de 40 ans, la mammographie peut également être prescrite à visée diagnostique en présence de signes cliniques.

La mammographie joue un rôle central dans le dépistage précoce et le suivi des pathologies mammaires, souvent complétée par une échographie pour affiner l'analyse.



Figure I.12: Le mammographe

Il est constitué de plusieurs éléments essentiels permettant la production, la détection et le traitement des images mammaires :

- Tube à rayons X.
- Générateur haute tension.
- Système de compression du sein.
- Dispositif de détection (détecteur d'image).
- Système d'acquisition et de traitement d'image.
- Support du sein (table ou plateau).

- Console de commande.
- Bras en C (ou bras articulé) [15].

I.7. Conclusion :

L'imagerie médicale est caractérisée par une diversité de techniques qui concernent l'exploration du corps humain, qu'il s'agisse de structures anatomiques ou de fonctions physiologiques, du fonctionnement en temps réel au fonctionnement métabolique. Qu'elles utilisent les rayons X, les ultrasons, les champs magnétiques, toutes les modalités présentent une architecture analogue : une source d'énergie, des détecteurs, et une chaîne d'acquisition, de traitement et d'affichage. Les systèmes présentés ont pu être développés pour remplir des besoins diagnostiques concrets, dans une résolution croissante, avec une sécurité accrue pour le patient, pour des analyses plus performantes.

Certaines de ces techniques ont pu se spécialiser dans des domaines cliniques particuliers. C'est le cas de la mammographie, modalité de radiologie du sein qui utilise les rayons X à faible dose, l'un des outils les plus utilisés pour le dépistage précoce du cancer du sein. Le chapitre suivant s'intéressera donc précisément à cette technique, ses principes de fonctionnement, ses éléments constitutifs et les évolutions récentes vers le numérique ou la 3D.

Chapitre II

La mammographie

II.1. Introduction :

La mammographie est une méthode d'imagerie médicale qui permet d'examiner les tissus internes des seins à l'aide de rayons X à faible dose. Cette analyse aide à repérer d'éventuelles anomalies ou tumeurs, même avant qu'elles ne soient détectables par un examen manuel. Son importance réside dans sa capacité à repérer rapidement les signes de cancer du sein, ce qui est crucial pour un diagnostic précoce et des traitements plus efficaces. En effet, la mammographie est reconnue comme l'un des moyens les plus performants pour le dépistage précoce du cancer du sein, ce qui peut considérablement améliorer les chances de rétablissement et sauver des vies. Son rôle essentiel dans la détection précoce du cancer du sein en fait un élément déterminant des programmes de santé publique visant à diminuer la mortalité liée à cette maladie.

Il existe deux types de mammographie : la mammographie de dépistage et la mammographie de diagnostique. La mammographie de dépistage peut être proposée lors du dépistage organisé du cancer du sein à partir de 40 ans. La mammographie diagnostique peut être proposée par votre médecin, pour un dépistage individuel plus rapproché, ou après la découverte d'une anomalie pendant un examen clinique des seins [16].

Dans ce chapitre dédié à la mammographie, nous nous introduirons dans le vif de l'imagerie médicale pour examiner en détail le rôle essentiel de la mammographie dans la détection précoce d'éventuelles anomalies mammaires. En explorant les principes fondamentaux de la mammographie, les différentes techniques d'imagerie utilisées, ainsi que l'acquisition et l'interprétation des images radiographiques du sein, nous chercherons à approfondir notre compréhension de cet outil précieux dans la prévention et le traitement du cancer du sein. En outre, nous aborderons la diversité des lésions mammaires détectables, l'importance capitale de la détection précoce du cancer du sein, et les nouvelles technologies émergentes qui façonnent l'avenir de la mammographie. Ce chapitre vise à offrir une perspective complète et éclairante sur l'impact crucial de la mammographie sur la santé des femmes et la lutte contre le cancer du sein

II.2. Fonctionnement de la mammographie :

II.2.1. Processus et étapes de la mammographie :

Dès que le patient arrivé, le manipulateur pose des questions sur ses antécédents médicaux et se renseigne également sur son état, comme si elle est ménopausée ou si elle suit un traitement hormonal. Si elle est enceinte ou avait la possibilité de l'être, il est crucial de l'indiquer.

En effet, des mesures spécifiques doivent être mises en place pour garantir la protection du fœtus des rayons X pendant l'examen.

Ensuite, le patient se déshabille, se positionner près du mammographe, généralement en position debout, bien que cela puisse occasionnellement se faire en position assise.



Figure II.1: Vue frontale du mammographe

L'examen proprement dit dure environ quinze minutes et se déroule en plusieurs phases :

Un de ses seins est placé sur une plaque intégrée au mammographe. Une seconde plaque vient s'appuyer sur ce sein et l'aplatit de manière verticale. Cette compression vise à garantir une image de bonne qualité tout en réduisant l'exposition aux radiations. Bien qu'elle ne soit pas douloureuse, elle peut provoquer une certaine gêne. Cette étape dure moins d'une minute.

Le technicien se positionne derrière un écran de protection. Il demande au patient de rester immobile lors de la prise de l'image, afin d'assurer sa clarté. Il actionne ensuite à distance le mammographe pour capturer la première image.

Dès que l'image est obtenue, la compression se relâche automatiquement.

L'appareil pivote de 45°. Le sein est alors positionné sur le côté, toujours entre les deux plaques, pour permettre la capture d'une seconde image en oblique.

Le technicien effectue les mêmes étapes pour l'autre sein. La mammographie peut être complétée par un cliché de profil ou focalisé sur une zone précise du sein.



Figure II.2 : Vue du mammographe en prise oblique

Après la prise d'images, le radiologue se rend dans une autre pièce pour les développer et vérifier leur clarté. Cela peut prendre quelques minutes. Pendant ce temps, vous attendez dans la salle de radiologie au cas où il serait nécessaire de refaire certaines images, ou pour faire un éventuel examen complémentaire comme l'échographie mammaire.

Si le mammographe est un appareil de radiologie numérique, le processus de l'examen reste le même (standard). Les images sont visibles immédiatement sur un écran et sont également imprimées sur des clichés.



Figure II.3: Détails des composants relatifs au compresseur de mammographe pour des examens médicaux de haute précision

II.2.2. Le principe de fonctionnement d'un mammographe :

Le principe de fonctionnement d'un mammographe repose sur l'utilisation de rayons X pour capturer des images détaillées des tissus mammaires. Lors d'une mammographie, le sein est placé entre deux plaques et comprimé pour obtenir une image radiographique claire. Les rayons X sont émis à travers le sein et capturés par un détecteur qui convertit les informations en une image visible à l'écran. Le processus de compression permet de réduire l'épaisseur du sein, minimisant ainsi la diffraction des rayons X et assurant une image de meilleure qualité [17].

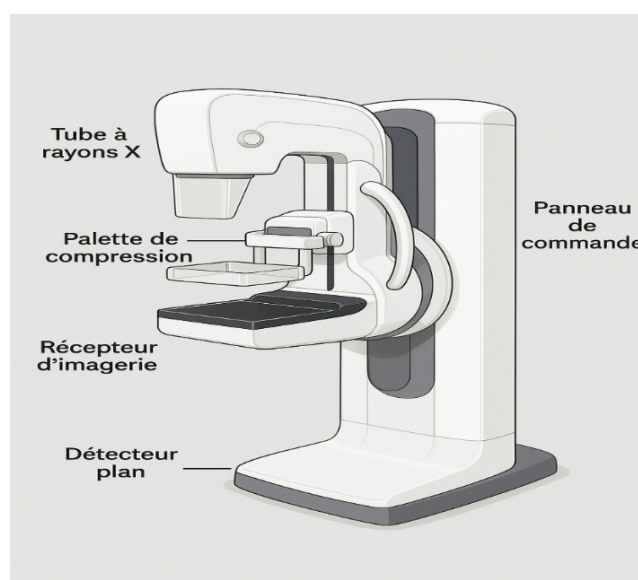


Figure II.4 : Schéma annoté d'un mammographe numérique et de ses principaux composants

Les images obtenues sont ensuite interprétées par des radiologues pour détecter d'éventuelles anomalies ou signes de cancer du sein.

L'appareil de mammographie se compose de plusieurs éléments fondamentaux, chacun ayant une fonction précise dans le processus d'imagerie. En haut de l'appareil se trouve le tube à rayons X, qui représente la source principale de rayonnement. Celui-ci émet des rayons X qui pénètrent le sein pour créer une image. Juste en dessous, la plaque de compression est utilisée pour aplatir le sein de manière uniforme, ce qui aide à réduire son épaisseur, améliore la qualité de l'image, diminue la quantité de rayons X nécessaires et limite les mouvements, ce qui évite le flou. En face de cela, le récepteur d'imagerie capture les rayons X après qu'ils aient traversé les tissus mammaires afin de former l'image. Il peut s'agir d'un film radiographique ou d'un capteur numérique, en fonction du type de système utilisé. Le détecteur plan, quant à lui, est un capteur numérique avancé utilisé dans les systèmes de mammographie numériques directs ; il


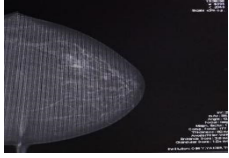
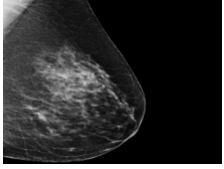
convertit les rayons X en signaux électriques pour produire une image de haute résolution. Finalement, le panneau de commande permet à l'opérateur (radiologue ou manipulateur en électroradiologie médicale) de gérer les paramètres de l'examen, tels que la dose de rayons X, la capture d'image ou encore la sélection des vues à prendre [18].

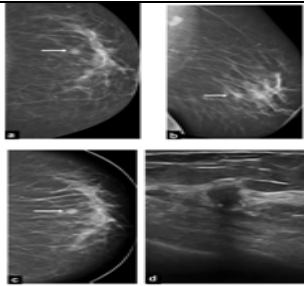
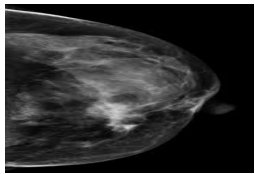
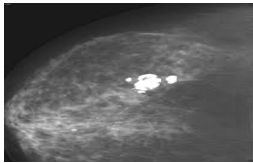
II.3. Interprétation et classification des images mammographiques :

II.3.1. Classification des mammographies en six catégories :

En fonction du degré de suspicions dans les mammographies, *L'American College OF Radiology (ACR)* a développé le système *BI-RADS (Breast Imaging Reporting and Data System)* pour standardiser les comptes rendus de mammographie. [19] le tableau I.1 résume la *classification Bi-Rads* de l'ACR correspondant au système de classement des mammographies.

Tableau I.1 les six catégories de mammographies selon la classification *Bi-Rads* de l'ACR.

CATÉGORIES DES IMAGES MAMMO- GRA- PHIQUES	Interprétation	Image radiographique
ACR 0 ACR 1	Mammographie normale	
ACR 2	Il existe des anomalies bénignes ne nécessitant ni surveillance ni examen complémentaire	
ACR 3	Il existe une anomalie probablement bénigne pour laquelle une surveillance à court terme est conseillée	

ACR 4	Il existe une anomalie indéterminée ou suspecte qui indique une vérification histologique.	
ACR 5	Il existe une anomalie évocatrice d'un cancer	
ACR 6	Cancer prouvé histologiquement	

II.4. Types de lésions mammaires :

Il existe plusieurs catégories de lésions du sein qui peuvent être détectées lors d'une mammographie.

II.4.1. Kystes du sein : Les kystes du sein sont des poches remplies de fluide qui peuvent se développer dans le tissu mammaire. En général, ils sont sans danger, mais ils peuvent parfois entraîner des douleurs et de l'inconfort [20].

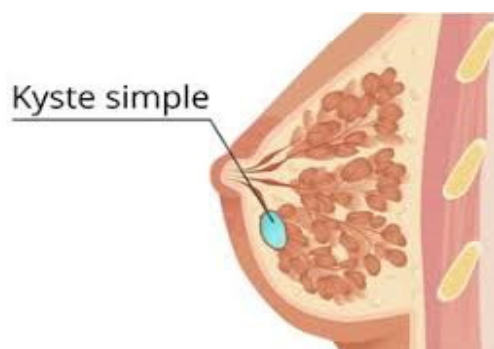


Figure II.5 : Un kyste du sein

<https://www.sante-sur-le-net.com/sante-femme/gynecologie/kyste-du-sein/>

II.4.2. Tumeurs non cancéreuses : Les tumeurs non cancéreuses du sein, comme les fibroadénomes, sont des formations qui ne sont pas malignes et peuvent se présenter sous forme de nodules dans une mammographie. Bien qu'elles ne soient pas cancéreuses, un suivi ou une prise en charge peut s'avérer nécessaire [21].



Figure II.6: la tumeur non cancéreuse

<https://medtube.fr/tags/fibroadenome/62746>

II.4.3. Microcalcifications : Les microcalcifications sont de petites accumulations de calcium qui peuvent être mises en évidence sur une mammographie. Elles peuvent parfois signaler des conditions précoces de cancer ou des risques accrus, surtout lorsqu'elles apparaissent regroupées de façon atypique [22].

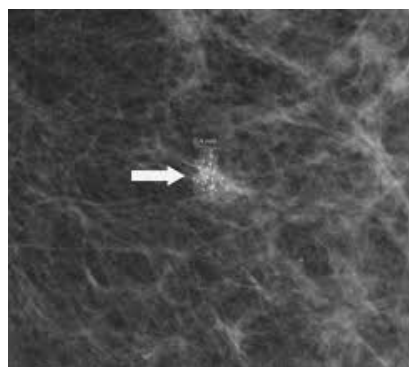
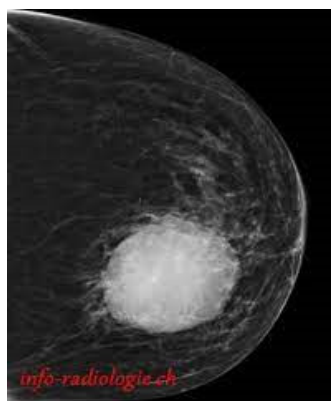


Figure II.7 : exemple de microcalcification

<https://www.depistagesein.ca/microcalcifications/>

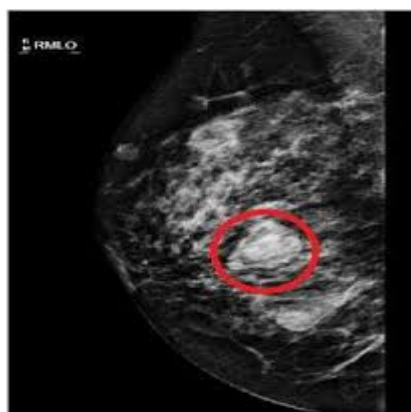
II.4.4. Masse solide : Les masses solides dans le sein peuvent être soit bénignes soit malignes. Elles peuvent se manifester comme des zones denses ou des boules sur une mammographie et nécessitent souvent des examens complémentaires pour préciser leur type.



FigureII.8 : une masse solide

<https://info-radiologie.ch/palpation-masse-sein.php>

II.4.5. Tumeurs cancéreuses : Les tumeurs cancéreuses du sein, dont les carcinomes, sont des formations qui peuvent apparaître sous différentes formes dans une mammographie. Une évaluation détaillée et un traitement immédiat sont essentiels [23].

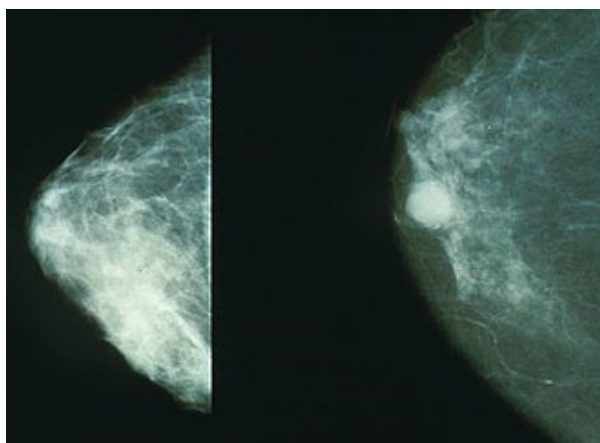


FigureII.9: Tumeurs cancéreuses

https://www.researchgate.net/figure/Mammographie-ayant-une-tumeur-benigne-circons-crite_fig12_338103710

II.5. Cancer du sein :

Le cancer du sein est une pathologie maligne qui se développe au niveau de la glande mammaire, c'est un cancer qui naît dans les unités cellulaires dont la fonction est de sécréter le lait, les unités ductolobulaires du sein, essentiellement chez la femme [24]. Huit cancers du sein sur dix se déclarent après 50 ans. Il touche environ 1,8 million de femmes chaque année dans le monde. Les projections indiquent qu'une femme sur huit sera concernée par un diagnostic de cancer du sein au cours de sa vie. Grâce aux progrès thérapeutiques réalisés au fil des années, les taux de rémission ont significativement augmenté.



FigureII.10 : Mammographie montrant un cancer du sein à droite.

https://fr.wikipedia.org/wiki/Cancer_du_sein

Entre 5 à 10 % de ces cancers ont une origine génétique héréditaire ; 85 à 90 % des cas ont des origines environnementales ou inconnues. Les facteurs de risque sont induite par la prise d'hormones, œstrogènes et la progestérone contenues dans les contraceptifs ou les traitements contre la ménopause, la consommation d'alcool, de sucre, de produits laitiers, de graisses hydrogénées, l'obésité, le manque d'activité physique. Une première grossesse tardive et l'absence d'allaitement favoriseraient aussi ce cancer [25].

II.5.1. Statistiques épidémiologiques du cancer du sein : données mondiales et régionales

En 2022, on estimait à 670 000 le nombre de décès dus au cancer du sein dans le monde.

Environ la moitié des cancers du sein apparaissent chez des femmes qui ne présentent aucun facteur de risque spécifique autre que le sexe et l'âge.

Dans 157 pays sur 185, le cancer du sein est déclaré comme étant la première cause de cancer chez les femmes en 2022.

Le cancer du sein est pratiquement présent dans tous les pays du monde.

Les hommes représentent environ 0,5 % à 1 % des personnes touchées par ce type de cancer.

Des interventions efficaces pour un diagnostic précoce et en temps utile, associées à un traitement complet, à une réadaptation et à des soins palliatifs, sont essentielles pour réduire la charge du cancer du sein et pour atteindre et maintenir un fonctionnement et un bien-être optimaux [26].

Le cancer du sein (CS) est le cancer féminin le plus fréquent, à la fois dans les pays développés et dans les pays en développement, avec cependant une incidence deux à trois fois plus basse dans ces derniers. En Algérie l'incidence élevée des CS chez la femme jeune représente une réalité épidémiologique mais les données sont en réalité insuffisantes.

Une étude descriptive réalisée en 2019 à Ouargla (Algérie) basée sur les cas incidents de CS invasif enregistrés chez les femmes de cette localité (Wilaya de Ouargla), par le registre du cancer durant la période allant du 01/01/2015 jusqu'au 31/12/2019. Le nombre de cas de CS invasifs enregistrés entre 2015 et 2019 était de 424, ce qui représente 45 % des cancers féminins diagnostiqués durant cette période, avec une incidence standardisée moyenne de 41,62 pour 10 000 femmes. [27] L'âge moyen est de 48,03±12,34 ans et l'âge médian est de 47 ans. Les décès par CS représentent 41,3 % des décès par cancer chez la femme. Le carcinome canalaire infiltrant représente plus de 79 % des différents types histologiques du CS. Les fréquences des tumeurs de grades II et III étaient de 42,9 % et 50 % respectivement. Le stade métastatique est retrouvé chez 20,8 % des cas.

Le taux d'incidence standardisée moyenne enregistré dans cette localité est largement inférieur aux taux des différentes wilayas du nord. L'âge moyen de diagnostic est proche de celui enregistré sur le territoire national, ce qui affirme la particularité de survenue du CS chez la femme jeune dans notre pays. L'augmentation des taux d'incidence standardisée de 2015 à 2019 semble due essentiellement à une amélioration du recueil et la diversification des sources de données comme c'est le cas pour tout nouveau registre. La hausse du taux de la mortalité et du stade métastatique peut être expliquée par l'absence d'une stratégie de dépistage, dont la détection précoce associée à un traitement adapté a permis l'amélioration de la survie des malades dans les pays développés [28].

II.5.2. L'anatomie du sein :

Chez la femme, le sein est un organe destiné à produire du lait pour nourrir l'enfant, qui se développe à la puberté sous l'influence des hormones.

Chez l'homme, le sein est un reliquat d'organe mammaire, semblable au sein féminin pré-pubère, qui présente une morphologie relativement similaire à celle des femmes, mais dans une dimension toutefois atrophiée.

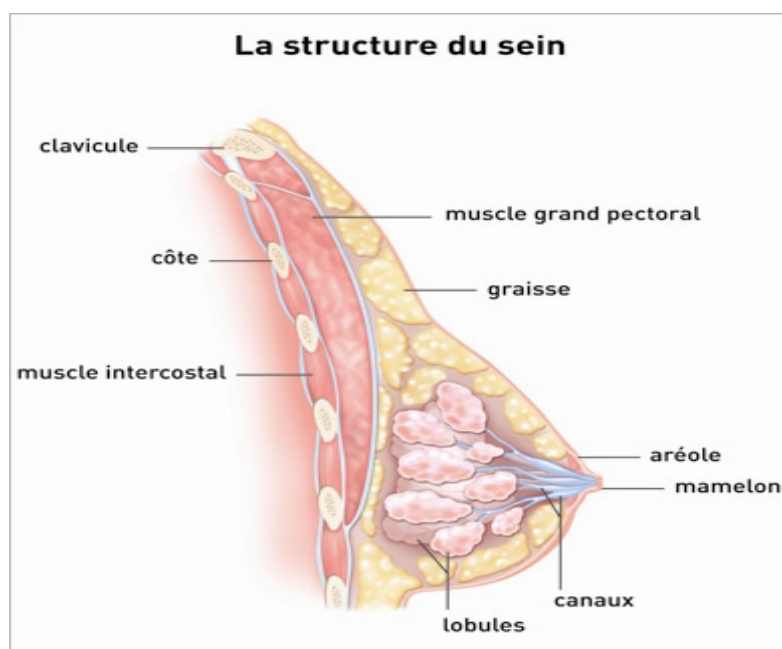


Figure II.11: la structure du sein

<https://www.europadonna.fr/le-cancer-du-sein/quest-ce-qu-un-sein/>

Le sein est composé d'une glande mammaire, de fibres de soutien (ligaments de Cooper) et de graisse (tissu adipeux) ; le tout est recouvert par la peau. La quantité de chacune de ses composantes peut varier d'une femme à l'autre. Chez l'homme, le sein se compose aussi de tissu graisseux, de canaux et de lobules, mais qui sont en moins grand nombre que chez la femme. Le sein est situé par-dessus le muscle pectoral. On trouve également dans le sein, des nerfs, des vaisseaux sanguins et lymphatiques. La glande mammaire est divisée en 15 à 20 sections qu'on appelle lobes, composés de lobules. Ceux-ci sont reliés à des canaux qui se rendent sous le mamelon (situé au centre du sein). On peut également observer des chaînes de ganglions lymphatiques qui filtrent les microbes et protègent le corps contre l'infection et la maladie. Le cancer du sein peut se développer tant au niveau d'un canal galactophore que d'un lobule et il peut également se retrouver au niveau des ganglions lymphatiques. Chez l'homme, le cancer est principalement présent dans les canaux et beaucoup plus rarement dans les lobules [29].

II.5.3. Type de cancer du sein :

- **Cancers du sein non invasifs (ou in situ)**

-Carcinome canalaire in situ (CCIS) : Se développe dans les canaux galactophores sans envahir les tissus adjacents. C'est un précurseur du cancer invasif.

-Carcinome lobulaire in situ (CLIS) : Affecte les lobules mammaires et indique un risque accru de développer un cancer invasif, mais il n'est pas considéré comme un cancer à part entière [30].

- **Cancers du sein invasifs**

-Carcinome canalaire invasif (CCI) : Le plus fréquent (~80 % des cas), il commence dans les canaux lactifères et envahit les tissus voisins.

-Carcinome lobulaire invasif (CLI) : Moins fréquent (~10-15 % des cas), il débute dans les lobules et peut être plus diffus, rendant la détection plus difficile [31].

- **Cancers du sein rares**

-Maladie de Paget du sein : Se manifeste au niveau du mamelon et de l'aréole sous forme d'eczéma, souvent associé à un carcinome canalaire in situ ou invasif.

-Carcinome inflammatoire du sein : Forme agressive où les cellules cancéreuses bloquent les vaisseaux lymphatiques de la peau, provoquant rougeur, gonflement et chaleur au sein.

-Carcinome médullaire, mucineux, tubuleux, papillaire, adénoïde-kystique, etc. : Types rares ayant des caractéristiques histologiques spécifiques et des pronostics variés [32].

- Cancers du sein liés à des mutations génétiques

-Certains cancers sont liés à des mutations des gènes BRCA1 et BRCA2, augmentant le risque de cancer du sein et de l'ovaire [33].

II.5.4. Dépistage et diagnostic de cancer du sein :

On rappelle qu'il existe deux types de mammographies :

- La mammographie de dépistage : La mammographie de dépistage est utilisée lorsqu'il n'y a pas de signes de cancer du sein [34]. Pour identifier le cancer du sein, la mammographie consiste à effectuer une radiographie. C'est la méthode la plus efficace pour le dépistage du cancer du sein. La mammographie est le seul examen de dépistage qui permet de réduire le nombre de décès attribuables au cancer du sein. Cet examen est offert à toutes les femmes admissibles de 50 à 74 ans.
- La mammographie diagnostique : La mammographie diagnostique est une radiographie qui emploie des radiations de faible dose pour produire des images du sein. La mammographie de diagnostic aide à poser un diagnostic lorsqu'il y a des signes de cancer du sein [34]. Elle vise à diagnostiquer un trouble du sein déjà connu, comme une masse ou un changement du mamelon ou de la peau ou à la suite d'un dépistage suspect.

II.6. Technologies Avancées en Mammographie :

II.6.1. L'imagerie améliorée par IA :

La mammographie connaît une percée importante avec l'aide de l'intelligence artificielle (IA), ce qui améliore la précision des méthodes de dépistage et réduit les faux positifs. Grâce à des algorithmes puissants, elle aide les radiologues à prioriser les cas complexes et urgents, tout en simplifiant les flux de travail. En utilisant l'IA, les images peuvent être analysées plus rapide-

ment et efficacement pour un diagnostic précoce. En utilisant des techniques d'imagerie avancées et une compression améliorée, elle contribue à l'optimisation de la qualité d'image. En utilisant des techniques d'imagerie avancées et une compression améliorée, elle contribue à l'optimisation de la qualité d'image [35].

II.6.2. Mammographie 3D :

La mammographie en 3D, aussi connue sous le nom de tomosynthèse numérique des seins, représente une avancée significative dans la détection du cancer du sein. En produisant des images tridimensionnelles du tissu mammaire, elle permet de pallier les inconvénients de la mammographie traditionnelle en 2D, en particulier la confusion due à la superposition des tissus. Cette méthode peut accroître le taux de détection des cancers jusqu'à 40 %, tout en limitant les rappels superflus, ce qui aide à réduire l'anxiété chez les patientes. De plus en plus intégrée dans les établissements de santé, surtout aux États-Unis, la mammographie en 3D se révèle être une technologie à la fois fiable, efficace et de plus en plus accessible, améliorant ainsi la qualité du dépistage [35].

II.6.3. Biopsie liquide :

La biopsie liquide représente une technique non invasive qui permet d'identifier des marqueurs associés au cancer, tels que l'ADN tumoral circulant (ADNtc), à partir d'un simple prélèvement sanguin. En complément des méthodes d'imagerie classiques, elle présente un potentiel significatif pour la détection précoce du cancer du sein, en révélant des modifications moléculaires avant qu'elles ne deviennent visibles via les examens d'imagerie. Cette méthode pourrait donc renforcer les chances de détection précoce et faciliter un traitement rapide. En outre, elle permet d'ajuster les traitements en temps réel, en fonction des mutations génétiques observées. Moins invasive qu'une biopsie de tissu, la biopsie liquide est en train de se développer en tant qu'outil prometteur pour un dépistage du cancer du sein plus précis et personnalisé [35].

II.6.4. Dépistage par thermographie :

La thermographie du sein utilise la technologie infrarouge pour repérer les changements de température à la surface du sein, ce qui peut indiquer des problèmes potentiellement liés à des maladies, y compris le cancer. Cette méthode est particulièrement pertinente pour les femmes ayant un tissu mammaire dense, où la mammographie traditionnelle peut parfois être moins efficace. Elle peut alors servir d'outil supplémentaire en identifiant des zones nécessitant une

attention particulière. Cependant, la thermographie n'est pas considérée comme une méthode de dépistage indépendante. Son manque de précision et le risque de résultats faussement positifs restreignent son utilisation dans la pratique médicale. En résumé, bien qu'elle fournisse des informations précieuses, elle doit être employée en complément d'examen reconnus, et non en remplacement de la mammographie [35].

II.6.5. Echographie sans fil :

L'échographie sans fil offre une solution portable et accessible pour le dépistage du cancer du sein. Sa facilité d'utilisation et le peu de formation nécessaire permettent aux personnels de santé de réaliser des examens dans les cliniques, dans des centres mobiles ou même au domicile des patientes.



Figure II.12 : Echographie sans fil

<https://www.yesenmed.com/fr/collections/wireless--usb-ultrasound-probe>

Cette technologie facilite un dépistage rapide, avec des résultats souvent fournis le jour même, diminuant ainsi l'angoisse des femmes et accélérant le processus de prise en charge. De nombreuses recherches ont montré que son efficacité est similaire à celle des échographes classiques pour identifier des tumeurs. En améliorant l'accès aux soins, l'échographie sans fil apparaît comme un outil prometteur pour renforcer les initiatives de dépistage du cancer du sein [35].

II.7. Conclusion :

La mammographie est devenue la méthode clé pour le dépistage précoce du cancer du sein. Grâce à sa capacité à identifier les lésions préoccupantes, elle a nettement augmenté les chances de détection à un stade initial, contribuant ainsi à réduire la mortalité liée à cette pathologie. Dans ce chapitre, nous avons abordé les principes fondamentaux de la mammographie, les diverses techniques d'imagerie, le processus de capture des images, ainsi que les types de lésions mammaires que l'on peut observer. Nous avons aussi examiné les avancées récentes, telles que la mammographie en 2D, la thermographie, l'échographie sans fil et les méthodes non invasives comme la biopsie liquide.

Néanmoins, malgré ces progrès technologiques, la mammographie présente toujours certaines limites, notamment pour les seins denses, les faux positifs ou négatifs, et les variabilités d'interprétation entre radiologues. C'est dans cette optique que l'intelligence artificielle (IA), et plus particulièrement l'apprentissage profond (deep learning), joue un rôle de plus en plus crucial. En effet, ces technologies novatrices offrent de nouvelles perspectives pour automatiser et perfectionner l'analyse des images mammographiques, améliorer la précision du diagnostic et diminuer la charge de travail des professionnels de santé.

Le chapitre suivant se concentrera ainsi sur l'application du *Deep learning* au traitement des images mammographiques. Nous explorerons les principes de l'apprentissage profond, les types d'algorithmes employés (comme les réseaux de neurones convolutifs), ainsi que leurs utilisations concrètes dans la détection des anomalies mammaires. Ces nouvelles approches marquent une nouvelle ère dans la lutte contre le cancer du sein, en combinant la richesse des données médicales avec la puissance des ressources informatiques et les capacités d'apprentissage automatique de l'AI.

Chapitre III

*L'Apprentissage profond
dans l'Analyse des Images
Mammographiques*

III.1. Introduction :

L'intelligence artificielle (IA) connaît un essor remarquable dans le domaine médical, notamment en imagerie. Le rôle de l'IA en imagerie ne se limite pas à l'automatisation des processus ; elle modifie fondamentalement l'approche du diagnostic des maladies, le rendant plus précis et plus efficace. L'apprentissage profond (en anglais : *deep learning*) est un sous-domaine de l'intelligence artificielle qui utilise des réseaux neuronaux ayant de nombreuses couches pour résoudre des tâches complexes. [36] Ces techniques ont permis des progrès importants et rapides dans les domaines de l'analyse du signal sonore ou visuel et notamment de la reconnaissance faciale, de la reconnaissance vocale, de la vision par ordinateur, du traitement automatisé du langage. Durant les années 2000, ces progrès ont suscité des investissements privés, universitaires et publics importants, notamment de la part des GAFAM (Google, Apple, Facebook, Amazon, Microsoft). [37]. L'apprentissage profond permet d'extraire les principales caractéristiques des mammographies numériques pour distinguer les populations à risque de cancer. Il est une véritable révolution dans l'analyse automatique des images. Il est particulièrement adapté aux méga données médicales (Big Data) et permet d'en extraire des connaissances utiles. Cette nouvelle technologie d'IA permet de détecter automatiquement les lésions, de proposer des diagnostics différentiels et de rédiger des rapports radiologiques préliminaires. Ce chapitre examine les principes, les techniques, les usages et les défis de l'IA dans le domaine de l'imagerie mammaire. L'intention est de mettre en lumière les méthodes d'apprentissage automatique les plus courantes et de démontrer leur performance dans l'identification et la catégorisation des anomalies mammaires.

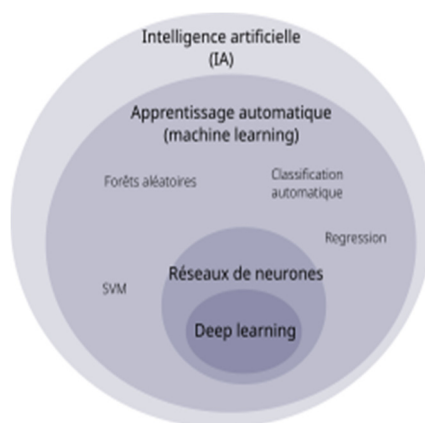


Figure III.1 : Apprentissage profond et IA

https://fr.wikipedia.org/wiki/Apprentissage_profond

III.2. L'intelligence artificielle :

III.2.1. Définition :

On peut la définir ainsi : l'Intelligence Artificielle est un domaine de l'informatique dont le but est de recréer un équivalent technologique à l'intelligence humaine. L'IA n'est pas une technologie à part entière mais un ensemble de technologies et d'outils [38].

En fait, l'IA est une discipline scientifique inventée en 1955 par deux mathématiciens, John MacCathy et Marvin Lee Minsky.

« L'IA est la science de programmer les ordinateurs pour qu'ils réalisent des tâches qui nécessitent de l'intelligence lorsqu'elles sont réalisées par des êtres humains. » [39].

-Marvin Lee Minsky-

L'intelligence artificielle est donc une technologie qui permet aux machines, qu'il s'agisse d'ordinateurs ou de programmes informatiques, d'exécuter des tâches généralement réservées à l'intelligence humaine. Aujourd'hui, les ordinateurs ont dépassé les capacités du cerveau humain en matière de puissance de calcul, de compréhension, d'adaptation, de communication et d'apprentissage profond. L'intelligence artificielle englobe également des dispositifs informatiques ou robotiques ainsi qu'un ensemble d'algorithmes qui permettent de prendre des décisions ou de résoudre des problèmes [40].

III.2.2. Le principe de l'intelligence artificielle :

L'objectif fondamental de l'intelligence artificielle est de reproduire le fonctionnement du cerveau humain au sein des machines, en s'appuyant sur un ensemble d'algorithmes d'apprentissage. Ces derniers permettent à ces machines de développer une forme d'intelligence propre [41].

III.2.3. Les applications de l'IA dans le domaine de la santé :

- L'intelligence artificielle est utilisée pour identifier automatiquement des anomalies (tumeurs, lésions, kystes) dans des images comme les radiographies, IRM, scanners et mammographies
- L'IA permet de délimiter automatiquement les contours d'organes ou de lésions dans une image médicale.

- L'IA classe des images ou des structures (ex. bénin vs malin) selon des critères appris automatiquement.
- L'IA aide à fusionner plusieurs types d'images pour une meilleure visualisation.
- L'IA peut modéliser la progression de maladies à partir de séries d'images.
- Les systèmes IA génèrent des rapports à partir d'images, combinant vision par ordinateur et traitement du langage naturel [42].

III.3. L'apprentissage automatique et l'apprentissage profond :

L'apprentissage automatique classique est généralement effectué sur des représentations des données obtenues par extraction de caractéristiques, cette dernière étant effectuée au moyen d'un algorithme [43].

III.3.1. Deep Learning :

Le deep learning constitue une méthode d'intelligence artificielle permettant à un ordinateur d'apprendre de manière autonome à identifier des motifs dans les images, tels que les anomalies sur les mammographies, sans nécessiter une programmation explicite [51]. Pour entraîner un modèle capable de détecter les premiers signes d'un cancer du sein, il est nécessaire de lui fournir un grand nombre d'images annotées, en distinguant les zones normales et anormales. Après cet apprentissage, le système peut analyser de nouvelles images et détecter automatiquement des anomalies potentielles, même inédites [44].

L'apprentissage profond repose sur l'idée d'apprendre directement les représentations des données plutôt que de les définir manuellement. Les couches cachées des réseaux de neurones permettent ainsi de construire une hiérarchie de représentations, allant de motifs simples (par exemple les contours) à des structures plus complexes. Ce processus hiérarchique justifie la qualification de « profond » [45].

Une image peut être représentée par ses intensités de pixels, ses contours ou encore ses formes. Les techniques de deep learning apprennent ces représentations automatiquement, à partir de données brutes, sans intervention humaine explicite [46].

Dans le cas des réseaux de neurones convolutifs (CNN), les premières couches extraient des motifs élémentaires, comme les bords et les textures, tandis que les couches plus profondes détectent des structures plus complexes et pertinentes pour la tâche de classification [47].

Le deep learning nécessite des ressources matérielles importantes : le CPU pour coordonner le traitement des données et le GPU pour exécuter rapidement les calculs intensifs liés aux images médicales [48].

III.3.2. Domaines d'applications :

L'apprentissage profond connaît actuellement une croissance impressionnante dans divers secteurs, en particulier dans le domaine de la santé, où ses aptitudes en matière de traitement de grandes quantités de données et de reconnaissance de motifs complexes sont pleinement mises à profit. Dans le secteur médical, il permet notamment d'aider les professionnels de santé à poser des diagnostics en utilisant des données d'imagerie. Des systèmes fondés sur le deep learning peuvent automatiquement identifier certaines maladies, comme les cancers (y compris le mélanome cutané ou le cancer du sein grâce aux mammographies) et la rétinopathie diabétique [49].

Ces technologies offrent un gain significatif en rapidité et en précision lors des dépistages, tout en diminuant la subjectivité des analyses médicales. De plus, une application remarquable dans le domaine de la santé publique est le projet *Horus*, un dispositif portable basé sur la plateforme NVidia Jetson. Cet appareil innovant, destiné aux personnes malvoyantes ou aveugles, vocalise les scènes vues par une caméra intégrée, leur permettant ainsi de percevoir leur environnement et les individus qui les entourent [50].

L'apprentissage profond est également utilisé en bio-informatique, en particulier pour l'analyse des séquences d'ADN, l'identification des segments non codants du génome et dans les technologies de cytométrie, ce qui favorise une meilleure compréhension des mécanismes biologiques [51].

III.4 Les types d'apprentissage automatique :

III.4.2.1. Apprentissage supervisé :

L'apprentissage supervisé est une méthode d'apprentissage automatique dans laquelle un modèle est entraîné à partir de données d'entrée et de sortie étiquetées. Le modèle utilise ces données pour apprendre à faire des prédictions précises sur de nouvelles données. Par exemple dans la classification d'images, des images étiquetées avec des étiquettes telles que "chien" ou "chat" sont utilisées pour entraîner un modèle à identifier les animaux dans de nouvelles images [52].

III.4.2.2. Apprentissage non supervisé :

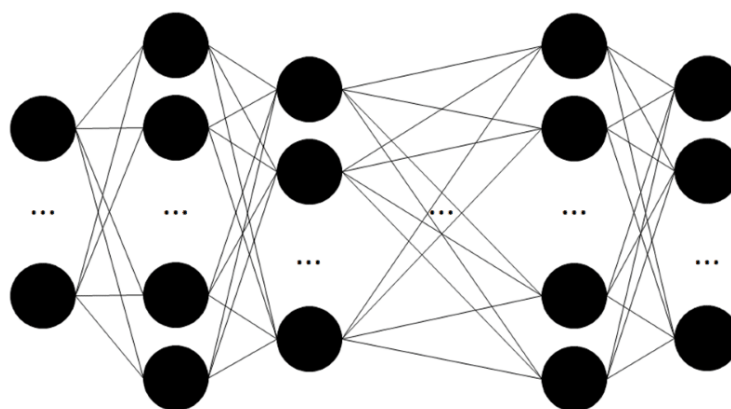
L'apprentissage Non Supervisé est une méthode d'apprentissage automatique où un algorithme est entraîné à trouver des motifs ou des structures dans un ensemble de données sans l'aide d'étiquettes ou de réponses prédéfinies. Les algorithmes d'apprentissage non supervisé sont souvent utilisés pour explorer et analyser des données, détecter des anomalies, regrouper des données similaires et réduire la dimensionnalité des données [53].

III.4.2.3. Apprentissage semi-supervisé (hybride) :

L'apprentissage Semi-Supervisé dit « hybride » est une méthode d'apprentissage automatique qui combine à la fois des données étiquetées et non étiquetées pour améliorer les performances de la classification ou de la prédiction. Cette approche permet de tirer parti des données non étiquetées pour renforcer la qualité de la modélisation et de l'apprentissage, tout en limitant la nécessité d'étiqueter manuellement toutes les données d'entrée. Les modèles hybrides peuvent inclure des techniques telles que la classification par ensembles, la régularisation, la propagation de labels, etc [54].

III.5. Les réseaux de neurones :

C'est un ensemble de « neurones artificiels » organisés de la manière illustrée dans la figure 2.



FigureIII.2 : Structure générale d'un Réseau de neurones

<https://www.google.com/url?sa=i&url=https%3A%2F%2Fperspective.orange-business.com%2Ffr%2Ftutoriel-machine-learning-comprendre-ce-quest-un-reseau-de-neurones->

Cette structure comporte plusieurs couches distinctes comme le montre la figure 3.

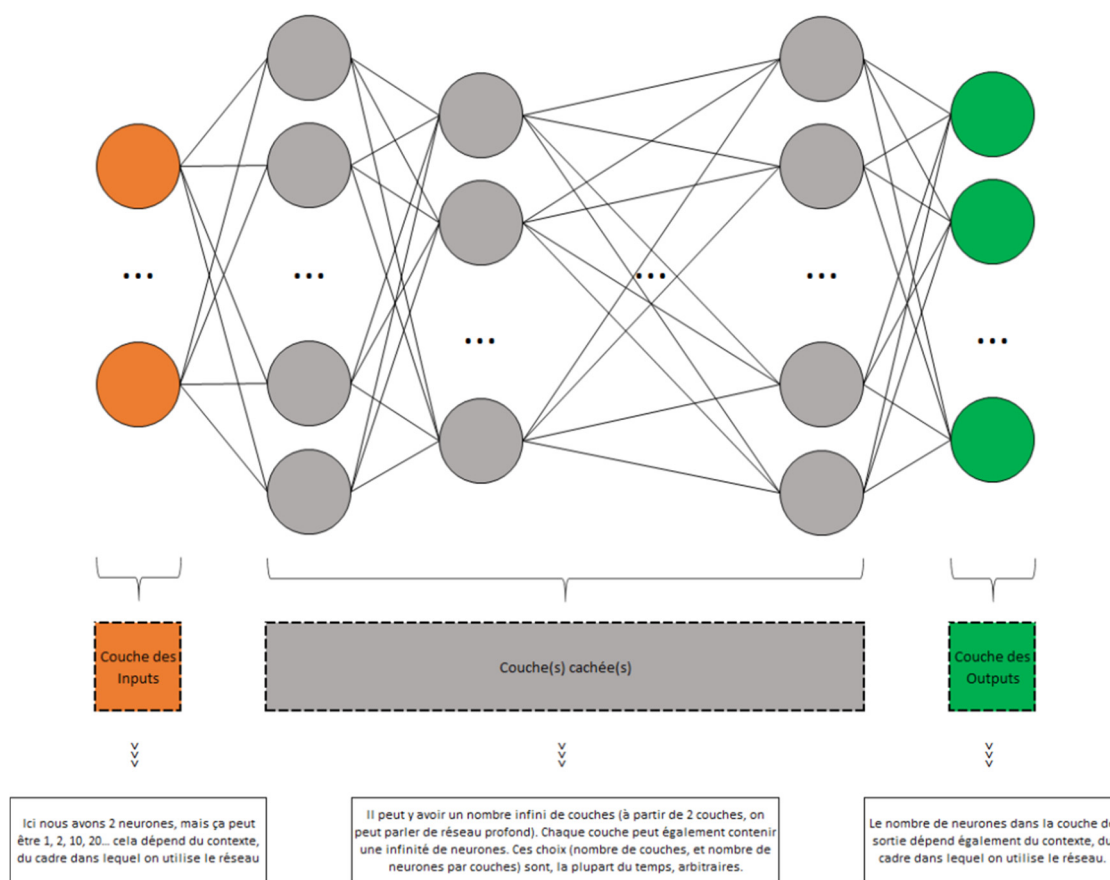


Figure III.3 : Structure d'un réseau de neurones artificiels

<https://perspective.orange-business.com/fr/tutoriel-machine-learning-comprendre-ce-quest-un-reseau-de-neurones-et-en-creer-un/>

Chaque « couche » est connectée, reliée à l'ensemble des neurones de la couche précédente.

Dans le domaine de l'intelligence artificielle, un réseau de neurones artificiels est un ensemble organisé de neurones interconnectés permettant la résolution de problèmes complexes tels que la vision par ordinateur ou le traitement du langage naturel [55].

Il s'agit d'un type particulier d'algorithmes d'apprentissage automatique (comme les machines à vecteur de support ,arbres de décision, K plus proches voisins, etc.) caractérisés par un grand nombre de couches de neurones, dont les coefficients de pondération sont ajustés au cours d'une phase d'entraînement [55].

III.5.1. Les différents types de réseaux de neurones :

Il existe deux principaux types de systèmes de deep Learning dotés d'architectures différentes :

III.5.1.1. Les réseaux neuronaux convolutifs (CNN) :

Les réseaux de neurones convolutifs désignent une sous-catégorie de réseaux de neurones : ils présentent donc toutes les caractéristiques listées ci-dessus. Cependant, les CNN sont spécialement conçus pour traiter des images en entrée. Leur architecture est alors plus spécifique : elle est composée de deux blocs principaux :

Le premier bloc fait la particularité de ce type de réseaux de neurones, puisqu'il fonctionne comme un extracteur de caractéristiques. Pour cela, il effectue de la machine à modèle en appliquant des opérations de filtrage par convolution. La première couche filtre l'image avec plusieurs noyaux de convolution, et renvoie des "cartes de caractéristiques", qui sont ensuite normalisées (avec une fonction d'activation) et/ou redimensionnées [56].

Ce procédé peut être réitéré plusieurs fois : on filtre les cartes de caractéristiques obtenues avec de nouveaux noyaux, ce qui nous donne de nouvelles cartes de caractéristiques à normaliser et redimensionner, et qu'on peut filtrer à nouveau, et ainsi de suite. Finalement, les valeurs des dernières cartes de caractéristiques sont concaténées dans un vecteur. Ce vecteur définit la sortie du premier bloc, et l'entrée du second [56].

Le second bloc n'est pas caractéristique d'un CNN : il se retrouve en fait à la fin de tous les réseaux de neurones utilisés pour la classification. Les valeurs du vecteur en entrée sont transformées (avec plusieurs combinaisons linéaires et fonctions d'activation) pour renvoyer un nouveau vecteur en sortie. Ce dernier vecteur contient autant d'éléments qu'il y a de classes : l'élément i représente la probabilité que l'image appartienne à la classe i . Chaque élément est donc compris entre 0 et 1, et la somme de tous vaut 1. Ces probabilités sont calculées par la dernière couche de ce bloc (et donc du réseau), qui utilise une fonction logistique (classification binaire) ou une classification multi-classe comme fonction d'activation [57].

III.5.1.2. Les réseaux neuronaux récurrents (RNN) :

Un réseau neuronal récurrent (« Recurrent Neural Network » ou RNN), est un réseau de neurones profond entraîné sur des données séquentielles ou des données de séries temporelles pour créer un modèle de machine learning (ML) capable de tirer des prédictions séquentielles ou de tirer des conclusions sur la base d'entrées séquentielles [58].

Un RNN pourrait être utilisé pour prédire les niveaux d'inondation quotidiens sur la base des données météorologiques, des marées et des inondations quotidiennes passées. Les RNN peuvent également être utilisés pour résoudre des problèmes ordinaux ou temporels tels que la traduction linguistique, le traitement automatique du langage naturel (NLP), l'analyse des sentiments, la reconnaissance vocale et le sous-titrage [59].

III.5.2. Apprentissage des réseaux de neurones :

Au départ, un réseau de neurones est un système "ignorant". Cependant, grâce à un processus d'apprentissage basé sur les exemples qui lui sont fournis, il évolue progressivement pour devenir un modèle capable de représenter avec précision le comportement observé en relation avec les variables descriptives [60].

L'apprentissage des réseaux de neurones désigne la procédure par laquelle les poids synaptiques sont mis à jour, permettant ainsi au réseau de mieux accomplir la tâche qui lui est assignée [61]. On peut distinguer deux types d'apprentissages : l'apprentissage « supervisé » et l'apprentissage « non supervisé ».

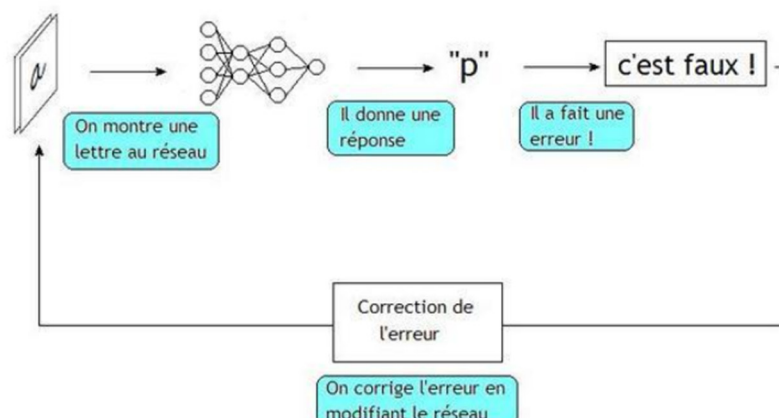


Figure III.4 : Fonctionnement d'un réseau neuronal en phase d'apprentissage

III.5.3. Les domaines d'applications des réseaux de neurones :

- ✓ Traitement d'image : compression d'images, reconnaissance de formes et de motifs, classification, reconnaissance de caractères et de signatures, chiffrement...
- ✓ Traitement du signal : traitement de la parole, identification de sources, filtrage, classification, ...
- ✓ Contrôle : diagnostic de pannes, commande de processus, contrôle qualité, robotique...
- ✓ Classification d'espèces animales étant donnée une analyse ADN
- ✓ Modélisation de l'apprentissage et perfectionnement des méthodes de l'enseignement
- ✓ Approximation d'une fonction inconnue ou modélisation d'une fonction connue mais complexe à calculer avec précision. [62]

C'est quoi un neurone dans un réseau de deep learning appliqué aux mammographies ?

Un neurone artificiel constitue une petite unité de traitement qui reçoit des données, les analyse et génère une nouvelle information. Concernant les mammographies, ces données peuvent inclure des valeurs de pixels d'une image, ou des attributs extraits (comme des textures ou des contours). Chaque neurone reçoit un signal d'entrée, effectue un calcul (dans la plupart des cas une somme pondérée), applique une fonction d'activation, et transmet le résultat aux neurones suivants. Des milliers de neurones collaborent au sein d'un réseau afin de déchiffrer des motifs visuels présents dans les mammographies [63].

Le Poids : Dans un réseau de neurones, chaque lien entre neurones est attribué à un poids. Ce poids reflète le niveau d'importance qu'un neurone attribue aux données qu'il reçoit. Concernant une mammographie, un poids important dans une région spécifique de l'image peut signaler que cette zone est particulièrement pertinente pour identifier une anomalie. Les poids sont modifiés de manière automatique durant l'apprentissage par une technique nommée rétro propagation, afin de perfectionner les performances du modèle [63].

Le Biais : Le biais constitue un élément additionnel intégré au calcul d'un neurone, permettant d'ajuster la sortie sans tenir compte des entrées. Cela confère au réseau une plus grande flexibilité pour représenter des relations complexes présentes dans les données.

Cependant, le terme "biais" peut également faire référence à un biais d'apprentissage : si le modèle a été formé uniquement avec des mammographies de jeunes femmes, son efficacité

pourrait diminuer sur des images de femmes plus âgées. C'est la raison pour laquelle il est essentiel d'utiliser un ensemble de données diversifié et équilibré durant le processus d'apprentissage [63].

La Perte : La fonction de perte évalue l'écart entre la prédiction du modèle (par exemple, "absence de tumeur") et la vérité (par exemple, "Présence de tumeur"). Plus l'écart est considérable, plus la valeur de la perte augmente. L'objectif lors de l'entraînement est de réduire cette perte. Cela aide le modèle à fournir des prédictions de plus en plus précises sur de nouvelles images de mammographies [63].

III.6 Applications en mammographie :

Classification : L'intelligence artificielle permet de distinguer de manière autonome les lésions bénignes des lésions malignes dans les images de mammographie. En s'appuyant sur des réseaux de neurones convolutifs formés sur d'importantes bases de données d'images annotées, ces systèmes peuvent atteindre des niveaux de performance comparables à ceux des radiologues chevronnés [64].

Détection et segmentation : L'intelligence artificielle localise et identifie les anomalies telles que les masses, les microcalcifications et les déformations architecturales. Des modèles tels que U-Net ou Mask R-CNN sont appliqués pour cerner avec précision les zones suspectes, facilitant ainsi l'analyse et les décisions cliniques [64].

Prédiction du risque : En associant les images à des données cliniques, certains algorithmes sont capables d'évaluer la probabilité de développer un cancer du sein, même avant l'apparition de signes visibles. Ces outils de prévision peuvent contribuer à adapter les protocoles de dépistage en fonction des profils de risque individuels [64].

Triage automatique : Les systèmes d'intelligence artificielle permettent de classer les cas selon leur niveau de suspicion. Cela optimise la gestion du flux de travail en milieu clinique en permettant aux radiologues de traiter en priorité les examens les plus urgents, ce qui réduit les délais de diagnostic pour les patientes présentant un risque élevé [64].

III.6.1. Algorithmes et modèles profonds souvent utilisés :

III.6.1.1. CNN classiques : [65]

LeNet : Considéré comme l'un des premiers réseaux de neurones convolutifs, il a été créé pour la reconnaissance des chiffres manuscrits. Il sert de référence éducative pour les réseaux contemporains, présentant une architecture simple et peu profonde .

AlexNet : A transformé le domaine du deep learning en 2012 grâce à une architecture plus complexe et à l'utilisation des GPU pour accélérer l'apprentissage. Il a également introduit les fonctions ReLU et la régularisation par dropout .

VGGNet : Reconnu pour sa structure uniforme constituée de petites convolutions (3x3), il facilite une plus grande profondeur tout en demeurant relativement simple, ce qui le rend très prisé pour des tâches de classification d'images .

III.6.1.2. Réseaux pré-entraînés : [67]

ResNet (Réseau Résiduel) : A intégré des connexions résiduelles permettant de construire des réseaux très profonds sans dégradation des performances. ResNet est largement adopté en imagerie médicale en raison de sa robustesse et de sa capacité à saisir des détails complexes .

EfficientNet : Propose une méthode d'optimisation intégrale de la taille, de la profondeur et de la largeur du réseau pour maximiser la précision tout en réduisant les coûts computationnels. C'est un excellent compromis entre performance et efficacité, idéal pour des applications médicales à grande échelle .

III.6.1.3. Modèles récents : [68]

Vision Transformers (ViT) : En opposition aux CNN, ils emploient un système d'attention pour examiner des sections d'image. Ils ont démontré des performances compétitives, voire meilleures que celles des CNN, dans certaines applications d'imagerie médicale, notamment grâce à leur capacité à appréhender les relations globales dans les images .

Modèles hybrides : Ces structures allient les atouts des CNN avec d'autres éléments tels que :

Des modules d'attention (par exemple : SENet, CBAM) pour renforcer la capacité du réseau à se concentrer sur des zones d'intérêt.

Des réseaux récurrents (comme LSTM ou GRU) pour traiter des séquences d'images ou des antécédents médicaux.

Des combinaisons CNN-ViT, tirant parti à la fois de la puissance locale des convolutions et de l'attention globale des Transformers [69].

III.6.2. Architecture U-NET :

III.6.2. 1. La méthode U-NET :

U-Net est un réseau neuronal convolutif développé pour la segmentation d'images . Ce réseau repose sur un réseau neuronal entièrement convolutif dont l'architecture a été modifiée et étendue pour fonctionner avec moins d'images d'entraînement et permettre une segmentation plus précise . La segmentation d'une image 512×512 prend moins d'une seconde sur un GPU moderne (2015) utilisant l'architecture U-Net [70].

L'architecture U-Net a également été utilisée dans les modèles de diffusion pour la débruitage d'images itératives. Cette technologie sous-tend de nombreux modèles modernes de génération d'images, tels que DALLÉ , Midjourney et Stable Diffusion [71].

- **À quoi sert U-net ?**

U-Net est spécialement conçue pour effectuer des tâches de segmentation sémantique. Également appelée prédiction dense, la segmentation sémantique est un processus de vision par ordinateur consistant à attribuer une étiquette sémantique à chaque pixel d'une image pour en comprendre la signification [72].

Au lieu de simplement détecter des objets dans une image, la segmentation sémantique vise à diviser l'image en plusieurs régions significatives. Chacune va se voir assigner des étiquettes représentant son contenu sémantique. Si l'on prend l'exemple d'une image de rue, la segmentation sémantique va identifier et étiqueter les pixels correspondant à la route, aux voitures, aux piétons, aux bâtiments ou encore aux arbres [72].

III.6.2.2. Application de U-net dans le domaine biomédical :

La segmentation sémantique est une tâche précise de la vision par ordinateur. U-Net est utilisée dans des domaines dans lesquels la segmentation précise des images sont essentiels pour l'analyse et la compréhension des données visuelles.

À l'origine, c'est dans ce but que U-Net a été développé. Il est donc largement utilisé pour la segmentation d'images médicales. Il intervient dans la détection de cellules, la segmentation d'organes, la localisation de tumeurs, l'ophtalmologie, ou encore dans l'analyse d'images de tomographie par ordinateur (TDM) ou d'imagerie par résonance magnétique (IRM) [72].

III.6.2. 3. L'architecture de U-Net :

U-Net présente une architecture symétrique en forme de U, qui permet une segmentation précise des images tout en préservant les détails contextuels importants. La partie supérieure du U représente la section de l'encodeur. Elle se compose de couches de convolution et de pooling. Ces couches réduisent progressivement la résolution spatiale de l'image en extrayant des caractéristiques pertinentes.

La partie centrale du U est le point crucial de l'architecture. C'est une sorte de passerelle entre l'encodeur et le décodeur. Elle permet de conserver les informations contextuelles à différentes échelles. Enfin, la partie inférieure du U représente la section du décodeur. Elle se compose de couches de déconvolution et de concaténation. Celles-ci vont contribuer à la reconstruction d'une carte de segmentation de la même taille que l'image d'entrée [72].

III.7. Les bases de données et les études de références dans le domaine de la mammographie :

Les informations et les études de référence en mammographie sont des éléments indispensables pour la recherche, la mise au point d'algorithmes et l'amélioration des techniques médicales liées au dépistage et au diagnostic du cancer du sein.

Les bases de données de mammographie comme DDSM, MIAS, CBIS-DDSM, INBREAST et VINDR-MAMMO offrent des ensembles d'images mammographiques destinés à la recherche et à l'avancement des technologies médicales [73].

Les données ainsi que les études de référence jouent un rôle crucial dans la validation des modèles d'intelligence artificielle, des algorithmes de détection de lésions et d'autres outils employés dans l'imagerie médicale.

Les études de référence permettent d'évaluer les performances des divers algorithmes de traitement d'images pour la détection précoce du cancer du sein [74].

Tableau III.1 : comparatif des principales bases de données dédiées à l'imagerie mammographique

Nom de la DB	Nombre d'images	Format	Caractéristiques de la DB	Applications
DDSM	26602	LJPEG, PNG	Ancienne base, images numérisées, annotations de masses et calcifications	Détection de lésions, entraînement IA, classification[74]
CBIS-DDSM	3 100	DICOM, (et PNG)	Version améliorée de DDSM, annotations mieux structurées, BIRADS inclus	Apprentissage profond, segmentation, classification[75]
INbreast	2084	DICOM (et PNG)	Images haute qualité, annotations précises (contours, types), base très propre	Segmentation fine, analyse radiologique avancée[76]
MIAS	322	PGM (et PNG)	Petites images en gris, annotations simples (type, position, sévérité)	Études de base, détection automatique simple[77]
VINDr-Mammo	20 000	DICOM	Réaliste, annotée par radiologues, BIRADS, densité mammaire et types d'anomalies indiqués	Détection, diagnostic assisté par IA, projets cliniques[78]

Une recherche récente a mis en œuvre des méthodes de deep learning pour améliorer la détection des problèmes liés aux seins lors des mammographies. Les scientifiques ont utilisé des réseaux de neurones convolutifs afin d'examiner les images mammographiques, obtenant ainsi des résultats encourageants en matière de sensibilité et de spécificité. Cette méthode a permis de constater une réduction significative des diagnostics tardifs du cancer du sein et une meilleure identification des lésions inquiétantes, proposant ainsi une technique de dépistage plus

précise et efficace. Ces résultats ouvrent de nouvelles perspectives pour le diagnostic et le traitement des patients souffrant de cancer du sein [79].

L'évaluation des performances des modèles de deep learning dans le domaine de la mammographie s'effectue généralement à travers des indicateurs tels que la sensibilité, la spécificité, la précision, le rappel et la courbe ROC (Receiver Operating Characteristic). Ces indicateurs aident à apprécier l'efficacité du modèle dans la détection des anomalies mammaires.

La sensibilité reflète la capacité du modèle à identifier correctement les cas positifs existants (anomalies réelles), tandis que la spécificité évalue sa capacité à exclure avec succès les cas négatifs (absence d'anomalies). La précision révèle le pourcentage de résultats positifs valides parmi tous les résultats prédit par le modèle, tandis que le rappel indique la proportion de cas positifs effectivement identifiés par rapport à tous les cas positifs présents dans les données.

La courbe ROC est un diagramme qui montre la performance d'un modèle selon divers seuils de classification. Elle aide à juger de la capacité du modèle à distinguer entre les catégories positives et négatives. Plus la surface sous la courbe ROC se rapproche de 1, meilleures seront les performances du modèle [80].

III.8. Défis actuels et limites liés aux données et à l'intelligence artificielle

-Problèmes associés aux données de mammographie [81]

- **Déséquilibre des classes :** Les images saines sont largement prédominantes comparées aux images montrant des anomalies. Cela complique l'apprentissage des cas pathologiques.

Exemple : un système peut apprendre à toujours indiquer « sein sain » car cela est correct dans 95 % des situations, mais cela signifie qu'il peut rater les cancers !

- **Qualité inégale des images :** Certaines bases de données incluent des images anciennes, à faible résolution ou numérisées (comme MIAS et DDSM). Cela peut restreindre l'efficacité des modèles modernes, qui exigent des images d'une excellente qualité.
- **Annotations incomplètes ou inexactes :** Les zones préoccupantes sont parfois mal délimitées ou incorrectement classées (bénin/malin). Cela affecte la précision et la fiabilité de l'apprentissage.

- **Manque de diversité :** Les données proviennent souvent d'un seul pays, d'un hôpital ou d'un groupe ethnique. Par conséquent, les modèles peuvent échouer à s'appliquer à d'autres populations.

- **Accès limité aux données réelles**[82]

- **Limitations concernant l'interprétabilité des modèles IA :** Les algorithmes d'apprentissage profond sont efficaces, mais leur fonctionnement demeure souvent mystérieux.
- **Boîtes noires :** Les modèles, notamment les réseaux de neurones profonds (CNN), prennent des décisions sans qu'on puisse comprendre précisément comment cela se produit. Dans le domaine médical, où chaque choix doit être justifié, cela peut poser problème.

Exemple : un système conclut qu'une image représente un cancer, mais il est impossible de déterminer quels aspects spécifiques de l'image ont conduit à cette conclusion.

- **Difficulté à expliquer les résultats aux médecins :** Si un médecin ne comprend pas la raison pour laquelle l'IA propose un diagnostic, il pourrait ne pas lui faire confiance, même si le résultat est correct.
- **Risque d'erreurs non détectées :** Sans interprétation claire, les erreurs dans les modèles peuvent passer inaperçues, représentant ainsi un danger pour les patients.

III.9. Perspectives Futures :

- **Incorporation de l'intelligence artificielle dans les processus cliniques :**

L'intégration des technologies d'intelligence artificielle dans les processus cliniques pourrait renforcer l'efficacité des radiologues en offrant une aide intelligente pour l'analyse des images mammographiques. Des modèles d'apprentissage profond peuvent être employés pour analyser préalablement les images et mettre en avant les zones d'intérêt, ce qui permet aux radiologues de mieux gérer leur temps et d'améliorer leur précision dans le diagnostic [83].

- **Vers une intelligence artificielle compréhensible (XAI) :**

La progression vers des modèles d'intelligence artificielle compréhensible s'avère cruciale pour instaurer la confiance parmi les professionnels de la santé concernant les choix faits par les systèmes d'IA en matière de mammographie. Les chercheurs développent des méthodes pour clarifier les décisions des modèles d'apprentissage profond, permettant ainsi aux radiologues de saisir comment les résultats ont été établis et de les interpréter de façon plus confiante [83].

- **Possibilités offertes par les approches multimodales et la coopération humain-machine :**

Les approches multimodales, intégrant des données de différentes sources telles que la mammographie et l'échographie, présentent un potentiel important pour améliorer la précision du diagnostic. Par ailleurs, la coopération entre humains et machines, où les radiologues collaborent avec des systèmes d'intelligence artificielle pour analyser les images et faire des choix, peut conduire à des diagnostics plus exacts et à une meilleure gestion des patients [83].

III.10. Conclusion :

L'apprentissage profond a apporté des avancées majeures dans l'examen des mammographies, augmentant la précision, la rapidité et la fiabilité des méthodes de détection des anomalies au niveau des seins. En étudiant d'importants volumes de données d'images mammographiques, ces modèles de deep learning sont capables de repérer des structures et des détails subtils qui pourraient passer inaperçus pour l'œil humain, ce qui permet une identification précoce des lésions et une prise en charge optimale des patients.

Les réseaux neuronaux profonds ont prouvé leur aptitude à acquérir des représentations très distinctives des images mammographiques, leur permettant de différencier efficacement les tissus sains des tissus suspects. En intégrant des architectures contemporaines telles que les réseaux neuronaux convolutifs (CNN) et les réseaux antagonistes génératifs (GAN), les scientifiques ont réussi à rehausser la sensibilité, la spécificité ainsi que la précision des modèles de deep learning dédiés à la détection des anomalies.

En outre, le deep learning a conduit à des usages novateurs comme la segmentation automatique des tissus, le classement des lésions et l'évaluation du risque de cancer, fournissant aux professionnels de la santé des outils robustes pour optimiser les diagnostics et les traitements. En intégrant ces innovations dans les pratiques cliniques, le deep learning pourrait transformer la façon dont les radiologues analysent les mammographies et prennent des décisions médicales.

Dans la suite de ce mémoire, il serait intéressant de se pencher sur des aspects expérimentaux ou techniques tels que l'optimisation des structures de modèles, la gestion des ensembles de données déséquilibrés, la compréhension des modèles de deep learning et les défis liés à leur validation clinique. En fin, il serait envisageable d'exploiter pleinement le potentiel du deep learning pour l'analyse des mammographies et d'ouvrir de nouvelles voies pour l'amélioration des soins en oncologie mammaire.

Chapitre IV

*Implémentation et
résultats*

IV.1. Introduction

Dans ce chapitre, nous allons présenter en détail les différentes expérimentations que nous avons réalisées dans le cadre de notre projet de classification d'images médicales grâce à des entraînements adaptés des modèles issus de l'apprentissage profond. Divers tests ont été effectués en utilisant des environnements de développement différents, tels que notre propre ordinateur (sous Python/VS-Code), Google Colab et Kaggle Notebooks, et en exploitant des grandes bases de données médicales (mammographies), à savoir MIAS, DDSM et INBreast.

Ces expériences visent à évaluer la performance du modèle que nous avons utilisé, à savoir : DenseNet121. L'un des modèles basés sur les réseaux de neurones convolutifs suivant l'architecture communément appelée réseaux de neurones densément connectés (*Densely connected convolutional Network* ou simplement : *DenseNet*).

Ce chapitre fait suite aux sections précédentes, dans lesquelles nous avons présenté les principes théoriques du Deep Learning et expliqué comment peut-t-il être appliqué au diagnostic assisté du cancer du sein.

Notre analyse se termine par une comparaison des environnements de développement, ainsi que sur la manière dont le choix de la base de données affecte les résultats obtenus par le modèle.

IV.2. Les outils et environnements de travail :

Dans cette partie nous allons présenter les outils et les bibliothèques utilisées afin de réaliser notre système.

IV.2.1. Python sur PC local :

- Version utilisée : Python 3.10

Bibliothèques principales :

- **PyTorch** (Paszke et al., 2019) – Framework de Deep Learning permettant l'implémentation des réseaux de neurones.
- **Torchvision** – pour le traitement des images et le chargement des datasets standards.
- **Scikit-learn** (Pedregosa et al., 2011) – pour le calcul des métriques de classification (accuracy, précision, rappel, F1-score).

- **NumPy** – opérations matricielles et manipulation des tableaux numériques.
- **Matplotlib & Seaborn** – visualisation des résultats expérimentaux (courbes d'apprentissage, matrices de confusion).

Configuration matérielle testée :

- Processeur : Intel Core i5-8350U (8^e génération, fréquence de base 1.70 GHz, jusqu'à 1.90 GHz en mode turbo).
- Mémoire vive (RAM) : 8 Go (7,85 Go utilisables).
- Carte graphique : Intel UHD Graphics 620 intégrée (128 Mo de mémoire dédiée).
- Stockage : SSD 256 Go (environ 80 Go libres au moment des tests).
- Système d'exploitation : Windows 11 Professionnel 64 bits.

Avantages :

- Contrôle complet sur l'environnement Python et les dépendances installées (bibliothèques, versions, compatibilités).
- Possibilité de travailler hors ligne sans dépendre d'une connexion Internet.
- Adapté pour des tests unitaires, du prétraitement d'images et l'exécution de petits scripts.

Limites :

- Absence de GPU dédiée : l'exécution des modèles de Deep Learning est extrêmement lente sur CPU seule ou GPUs faibles.
- Mémoire RAM limitée (8 Go) : Peu adaptée pour des bases de données médicales volumineuses et minimale pour les moyennes bases comme DDSM ou INBreast.
- Stockage restreint (80 Go libre/238 Go exploitables) : limite l'intégration de plusieurs Datasets lourds.
- Risque de surchauffe et ralentissements lors de l'entraînement des grands modèles.

IV.2.2. Visual Studio Code (VS Code) est un environnement de développement intégré (IDE) gratuit et open-source développé par Microsoft, largement utilisé pour la programmation en Python et l'apprentissage profond [84].

Avantages :

- Léger et rapide à installer.
- Large choix d'extensions (ex. Python, Git, Jupyter).
- Interface simple et personnalisable [84].

Limites :

- Moins interactif que Jupyter Notebook pour la visualisation.
- Nécessite parfois une configuration manuelle (extensions, environnements virtuels).
- Les calculs lourds restent lents sur CPU en local [84].

IV.2.3. Google Colab :

- **Présentation** : Google Colab est un service de notebooks Jupyter en ligne qui permet d'écrire et d'exécuter du code en Python depuis son navigateur web sans la nécessité d'installer python sur le PC client. Gratuit, accessible avec un compte Google, Colab est particulièrement indiqué pour l'analyse de données et la data science avec un codage Python. Ses cellules de code et ses blocs de texte (pour commenter ou expliquer le code) permettent à la fois de structurer et commenter le code. Il fournit un environnement préconfiguré avec les bibliothèques couramment utilisées pour le Machine Learning (TensorFlow, PyTorch, Keras...) et offre un accès à des GPU et TPU (Gratuites ou payantes) [85].
- **Configuration matérielle disponible** :
 - ✓ GPU Tesla T4 (16 Go) ou P100 (Gratuites mais limitées en temps d'utilisation)
 - ✓ TPU (expérimentale, mais payante généralement),
 - ✓ RAM 12–16 Go. (Gratuites mais limitées en temps d'utilisation)
- **Gestion des données** :
 - ✓ Intégration avec Google Drive (synchronisé avec le Drive personnel),
 - ✓ Chargement direct (Upload) depuis l'ordinateur.

- **Avantages :**
 - ✓ Accès gratuit à des GPU/TPU,
 - ✓ Environnement déjà configuré (PyTorch, TensorFlow, etc.),
 - ✓ Exécution plus rapide que sur un PC classique.

- **Limites :**
 - ✓ Temps de session limité (12h maximum et si disponibles),
 - ✓ Déconnexion possible après un certain temps d'inactivité ou saturation des ressources hôtes.

IV.2.4. Kaggle Notebooks

Il s'agit d'une plateforme en ligne destinée aux data scientists, permettant de coder en Python et en langage R directement dans des notebooks avec accès à une communauté et à une large collection de Datasets sans être obligé d'installer Python sur le PC personnel. [86]

- **Configuration matérielle :**
 - ✓ GPU Tesla P100 (16 Go),
 - ✓ 12 à 13 Go de RAM,
 - ✓ 20 Go de stockage temporaire.

- **Gestion des données :**
 - ✓ Intégration native avec Kaggle datasets download,
 - ✓ Possibilité de travailler directement sur les jeux de données hébergés par Kaggle (sans être obligé de télécharger les bases d'images...)

- **Avantages :**
 - ✓ Accès simplifié aux datasets médicaux publics,
 - ✓ Environnement stable et gratuit,
 - ✓ Résultats reproductibles.

- **Comparaison avec Colab :**
 - ✓ Kaggle facilite l'accès aux bases de données mais offre moins de flexibilité avec les fichiers personnels (configuration fixe...)
 - ✓ Colab est mieux adapté pour sauvegarder des modèles et des checkpoints via Google Drive. Et permet aussi plus de flexibilité concernant les Datasets (organisation des jeux de données, nombre de classes...etc.) [86].

IV.3. Bases de données utilisées :

IV.3.1. MIAS (Mammographic Image Analysis Society)

- **Taille** (originale) : 322 images.
- **Taille** (avec data augmentation) : 3816 images.
- **Résolution** (originale) : 1024×1024 pixels.
- **Classes** : normales, bénignes, malignes.
- **Format** : PNG.
- **Prétraitements appliqués :**
 - ✓ Redimensionnement à 224×224 pixels,
 - ✓ Normalisation des valeurs de pixels,
 - ✓ Conversion en tenseurs PyTorch. [87]

IV.3.2. DDSM (Digital Database for Screening Mammography)

- **Taille** : 26 620 images.
- **Résolution** (originale) : jusqu'à 4000×5000 pixels.
- **Annotations** : contours des lésions, diagnostic.
- **Prétraitements appliqués :**
 - ✓ Réduction de la résolution (224×224 pixels),
 - ✓ Équilibrage des classes (sous-échantillonnage),
 - ✓ Normalisation. [88]

IV.3.3. INBreast

- **Taille** (originale) : 2084 images.
- **Résolution** (originale) : 3328×4084 pixels.
- **Format** (originale) : DICOM. ou JPG (transformée)
- **Particularité** : haute qualité, annotations précises, largement utilisé pour la validation. [89]

IV.3.4. Bases fusionnées

- **Méthodologie** :
 - ✓ Intégration des trois bases (MIAS, DDSM, INBreast),
 - ✓ Harmonisation (taille 224×224 , normalisation identique),
 - ✓ Équilibrage des classes (sous-échantillonnage ou sur-échantillonnage).
- **Avantages attendus** :
 - ✓ Enrichissement de la variabilité (résolutions et sources différentes),
 - ✓ Meilleure généralisation du modèle,
 - ✓ Réduction du surapprentissage (*Overfitting*) [90].

IV.4. Modèles et méthodes :

IV.4.1. Présentation des modèles testés :

Dans le cadre de ce travail, plusieurs modèles de réseaux de neurones convolutionnels ont été considérés. Le modèle principal est **DenseNet121**, reconnu pour son efficacité en classification d'images médicales. **DenseNet-121** est une architecture de *deep learning* conçue principalement pour la classification d'images, mais également applicable à d'autres tâches telles que la segmentation. Elle appartient à la famille des réseaux de neurones convolutionnels (CNN) et se distingue par son schéma de connectivité original.

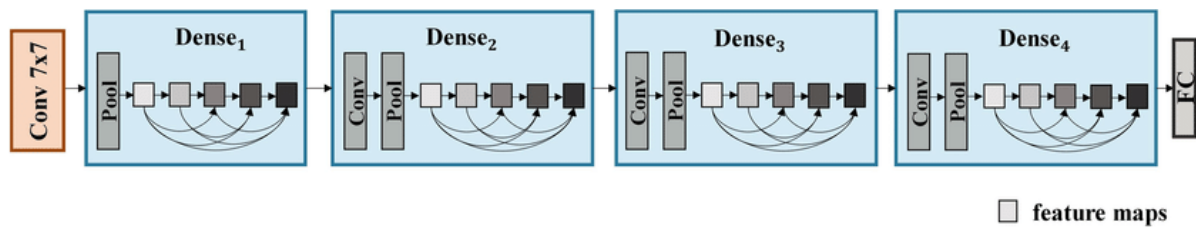


Figure IV.1: architecture de model densnet121

https://www.researchgate.net/figure/The-structure-of-DenseNet-121_fig2_342368106

➤ Caractéristiques principales

• **Connectivité dense :**

Contrairement aux CNN traditionnels, où chaque couche n'est connectée qu'à la couche suivante, DenseNet établit une connexion entre chaque couche et toutes les couches suivantes de manière directe. Ainsi, la sortie de chaque couche est transmise à toutes les couches ultérieures.

Cette connectivité dense favorise une meilleure réutilisation des caractéristiques et améliore la propagation du gradient, permettant d'entraîner des réseaux plus profonds sans souffrir du problème de gradient qui disparaît.

• **Réduction du nombre de paramètres :**

Les architectures DenseNet nécessitent moins de paramètres que les CNN classiques, car elles évitent la redondance dans l'apprentissage des caractéristiques. Chaque couche recevant déjà les informations des couches précédentes, le réseau n'a pas besoin de réapprendre plusieurs fois les mêmes représentations [91].

À titre comparatif, deux autres modèles ont également été étudiés :

ResNet (Resnet18/50)

ResNet18 fait partie de la famille des **ResNet (Residual Networks)**, introduite pour résoudre le problème de dégradation de performance dans les réseaux très profonds. L'idée clé est l'utilisation de **connexions résiduelles** (skip connections), qui permettent de propager directement l'information d'une couche à une autre, sans passer par toutes les transformations intermédiaires.

ResNet18, avec ses 18 couches, constitue une architecture relativement légère qui permet un

bon compromis entre performance et temps d'entraînement. Il est souvent utilisé comme point de départ ou pour des applications nécessitant une rapidité d'exécution.

ResNet50 est une version plus profonde de la famille ResNet, composée de 50 couches. Grâce à ses connexions résiduelles, il est capable d'apprendre des représentations plus riches sans perte de précision liée à la profondeur. Son utilisation est particulièrement adaptée aux tâches complexes nécessitant une discrimination fine entre classes, comme la classification des images médicales.

EfficientNet-B0 et EfficientNet-B1

EfficientNet introduit une nouvelle approche d'optimisation appelée *compound scaling*, qui ajuste de manière équilibrée la profondeur (nombre de couches), la largeur (nombre de filtres par couche) et la résolution d'entrée des images. Contrairement aux approches traditionnelles qui augmentent seulement la profondeur ou la largeur, EfficientNet maintient un compromis optimal entre performance et efficacité.

- **EfficientNet-B0** est la version de base, plus légère et rapide à entraîner.
- **EfficientNet-B1** est une version légèrement plus grande, offrant de meilleures performances avec un coût computationnel modéré à élever [91]

Le choix de DenseNet121 comme modèle de référence repose sur un compromis entre profondeur du réseau, performances rapportées dans la littérature et faisabilité expérimentale dans les environnements de calcul disponibles.

IV.4.2. Adaptation du modèle DenseNet121

Pour ce projet, nous avons retenu DenseNet121 comme modèle principal, en raison de sa capacité à favoriser une réutilisation efficace des caractéristiques extraites grâce à ses connexions denses entre les couches. Cependant, une adaptation a été nécessaire afin de l'appliquer à la tâche de classification des mammographies en trois classes : *malignant*, *bénigne* et *normal*.

Organisation des données

Les bases de données utilisées (**DDSM**, **INbreast**, et **MIAS**) ont été prétraitées et organisées de manière homogène, afin de faciliter l'entraînement du modèle. Chaque base a été divisée en trois sous-dossiers distincts :

- **train** : 70 % des images utilisées pour l'apprentissage.
- **val** : 15 % des images utilisées pour la validation intermédiaire.
- **test** : 15 % des images réservées à l'évaluation finale.

À l'intérieur de chacun de ces dossiers, les images ont été réparties en trois sous-classes :

- *malignant* : correspondant aux cas de cancers confirmés,
- *bénigne* : regroupant les anomalies non cancéreuses,
- *normal* : représentant les cas sains.

Cette organisation hiérarchique a permis de structurer correctement les DataLoaders de PyTorch pour l'entraînement, la validation et le test [92].

Adaptation du modèle DenseNet121

Le modèle DenseNet121 pré-entraîné sur ImageNet a été importé depuis la bibliothèque Torchvision.models. Afin d'adapter le réseau à notre problématique :

- La dernière couche fully-connected a été remplacée par une couche linéaire avec 3 neurones de sortie (correspondant aux classes *malignant*, *benign* et *normal*).
- Les images issues des différentes bases (DDSM, INbreast et MIAS) ont été redimensionnées en 244×244 pixels, afin d'être compatibles avec les couches convolutives initiales du réseau.
- Les données ont été normalisées selon les valeurs moyennes et écarts-types de l'ensemble ImageNet, pour maintenir la compatibilité avec les poids pré-entraînés.

Ainsi, DenseNet121 a pu être entraîné et évalué sur les trois bases de données, tout en conservant une cohérence entre les jeux de données et l'architecture du modèle.

IV.4.3 Paramètres d'entraînement

- Batch size = 16 (Nombre limité d'échantillons traités avant que le modèle change)
- Nombre d'époques = 16 (Nombre d'itérations sur le jeu de données d'entraînement).
- Taille des images = 244×244 pixels.
- Fonction de perte : CrossEntropyLoss
- Optimiseur (par défaut Adam ou SGD)
- Environnement d'exécution : CPU (lent, mais faisable)

IV.5 Évaluation et métriques

Pour évaluer les performances du modèle DenseNet121 et des autres architectures comparatives, plusieurs métriques standards en classification médicale ont été utilisées :

- **L'exactitude (Accuracy)** : Mesure du pourcentage global de prédictions correctes par rapport à l'ensemble des échantillons testés (et pré-diagnostiqués).

Telle que : l'Exactitude = $(TP + TN) / (TP + TN + FP + FN)$ où : TP = vrai positif ; FP = faux positif ; TN = vrai négatif ; FN = faux négatif.

- **La précision (Precision)** : Évalue la reproductibilité des mesures, (rapprochement des mesures les unes par rapport aux autres pour les différentes classes : bénignes, malignes, normales)

Tel que : Précision = $TP / (TP + FP)$.

- **Le Rappel (recall)** : Une métrique qui mesure la capacité d'un modèle à identifier correctement toutes les instances positives dans un Dataset.

Tel que : Recall = $TP / (TP + FN)$

- **F1-Score** : Représente la moyenne harmonique entre la précision (précision) et le rappel (Recall).

$$F1\ score = 2 * (Precision \times Recall) / (Precision + Recall)$$

- **AUC (Area Under the Curve : l'aire sous la courbe ROC)** : mesure la capacité du modèle à distinguer correctement entre les différentes classes.

Il s'agit d'un graphique qui représente les performances d'un modèle de classification d'images. Elle est tracée en portant le Taux de Vrais Positifs (TVP) sur l'axe des ordonnées et le Taux de Faux Positifs (TFP) sur l'axe des abscisses. Un modèle idéal se situe en haut à gauche du graphique, et une courbe plus proche du coin supérieur gauche indique une meilleure performance

- **Matrice de confusion** : fournit une vue détaillée des prédictions du modèle en indiquant le nombre de vrais positifs (*True positif TF*), vrais négatifs (*True négatif TN*), faux positifs (*False positif FP*), et faux négatifs (*False négatif FN*).

Elle permet de mieux identifier les types d'erreurs commises.

		True Class	
		Positive	Negative
Predicted Class	Positive	TP	FP
	Negative	FN	TN

Figure IV.2 La matrice de confusion

<https://www.picsellia.fr/post/matrice-confusion-computer-vision>

- **Temps d'exécution selon l'environnement** : les expériences ont été réalisées en local (Laptop personnel) sous Python et à l'aide de *Visual Studio Code*. L'entraînement a été effectué principalement sur **CPU**, ce qui a considérablement rallongé le temps d'exécution par rapport à une exécution avec une **GPU** (généralement plus adaptée aux modèles de *deep learning*). Cette limitation matérielle a été prise en compte dans l'interprétation des résultats [93].

IV.6 Résultats expérimentaux

IV.6.1 Résultats par base

Tableau 4.1 Résultats par base de données

Base de données	Exactitude (Accuracy %)	Précision	Rappel (Recall)	F1-Score	AUC (macro)	Temps d'exécution (CPU)
MIAS	0.98571	0.98571	0.98571	0.98571	0.999138	6h
IN-BREAST	0.99136	0.99136	0.99136	0.99135	0.99981	3h
DDSM	0.99928	0.99928	0.99928	0.99928	0.99998	36h

IV.6.1.1. INBreast Dataset:

La Figure IV.2 représente la **courbe ROC (Receiver Operating Characteristic)** illustrant la performance du modèle de *deep learning* entraîné sur le dataset **INbreast** en traçant le **taux de vrais positifs (sensibilité)** en fonction du **taux de faux positifs (1-spécificité)** pour différents seuils de décision. Dans notre analyse, les trois graphes de ROC (correspondant au 3 classes : bénigne, maligne, normal) convergent tous vers **1**, ce qui indique que le modèle atteint presque une **sensibilité et une spécificité maximales** simultanément. Cette convergence traduit une excellente capacité discriminative du modèle : c'est-à-dire qu'il distingue de manière quasi parfaite les lésions bénignes des lésions malignes. L'**AUC (Area Under the Curve)**, qui mesure l'aire sous la courbe ROC, approche la valeur 1, confirmant que le modèle présente une performance optimale sur ce dataset, avec un risque très faible de confusion entre les classes.

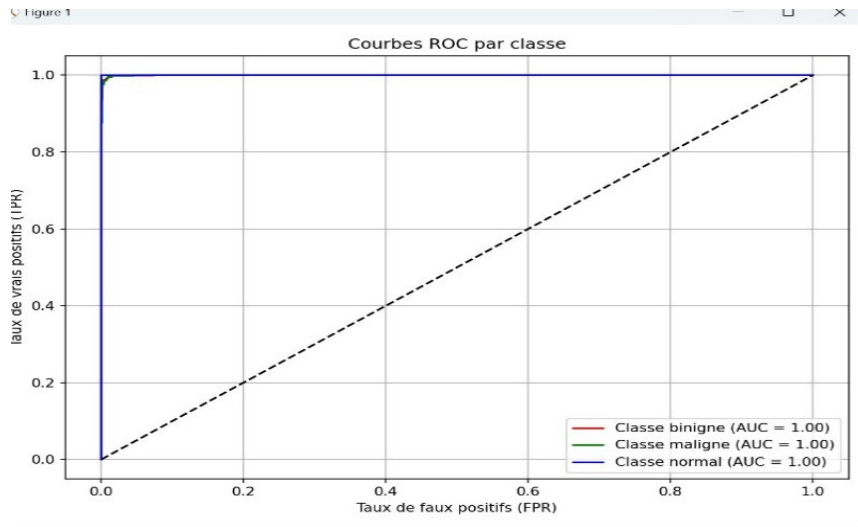


Figure IV.3.Courbe ROC - INbreast

IV.6.1.2. DDSM DATASET :

La matrice de confusion issue de la classification des mammographies du dataset **DDSM** met en évidence l'excellente performance du modèle deep learning sur les trois classes considérées (**bénin, malin et normal**).

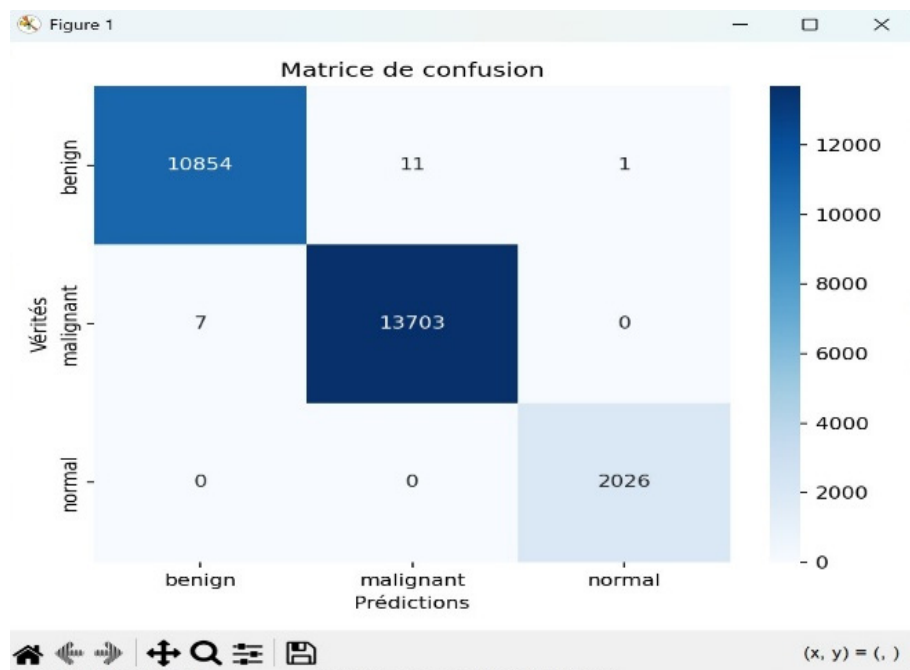


Figure IV.4 : Matrice de confusion sur DDSM

Les valeurs situées sur la diagonale principale (10 854 bénins correctement prédits, 13 703 malins correctement identifiés et 2 026 normaux correctement classés) témoignent d'une capacité de prédiction très fiable. Les erreurs de classification sont extrêmement limitées : seulement quelques cas de lésions bénignes prédites comme malignes (11), un bénin classé à tort comme normal (1), ainsi que 7 malins prédits comme bénins. Aucune confusion n'apparaît pour la classe normale. Ces résultats traduisent un modèle à la fois **robuste et précis**, avec un taux d'erreur négligeable, confirmant sa pertinence comme outil d'aide au diagnostic pour la détection des anomalies mammaires dans le dataset DDSM.

Les métriques dérivées de la matrice de confusion confirment la qualité exceptionnelle du modèle. L'**exactitude globale (accuracy)** atteint **99,93 %**, traduisant une quasi-absence d'erreurs de classification. Pour la classe **bénigne**, la précision est de **99,94 %**, le rappel de **99,89 %** et la spécificité de **99,96 %**, indiquant une très bonne capacité à identifier correctement les cas bénins tout en minimisant les faux positifs. La classe **maligne** présente également des performances remarquables, avec une précision de **99,92 %**, un rappel de **99,95 %** et une spécificité de **99,91 %**, ce qui montre que les lésions malignes sont détectées avec une fiabilité élevée. Enfin, pour la classe **normale**, les résultats sont quasi parfaits, avec une précision de **99,95 %**, un rappel de **100 %** (aucun cas normal n'a été mal classé) et une spécificité de **99,99 %**. Ces résultats démontrent que le modèle atteint une **performance optimale sur le dataset DDSM**, offrant une excellente capacité de discrimination entre les différentes catégories de mammographies.

La Figure IV.5 représentant la **courbe ROC** sur **la base DDSM** illustre aussi de très hautes performances du modèle. On voit que les trois graphes (correspondant au 3 classes : bénigne, maligne, normal) convergent tous vers **1** pour ce deuxième dataset, ce qui indique que le modèle atteint une **sensibilité et une spécificité maximales** simultanément. L'**AUC** se rapproche également de 1 pour ce dataset, avec un risque très faible de confusion entre les classes. Ça veut dire que le modèle est capable de distinguer de manière parfaite les lésions bénignes des lésions malignes.

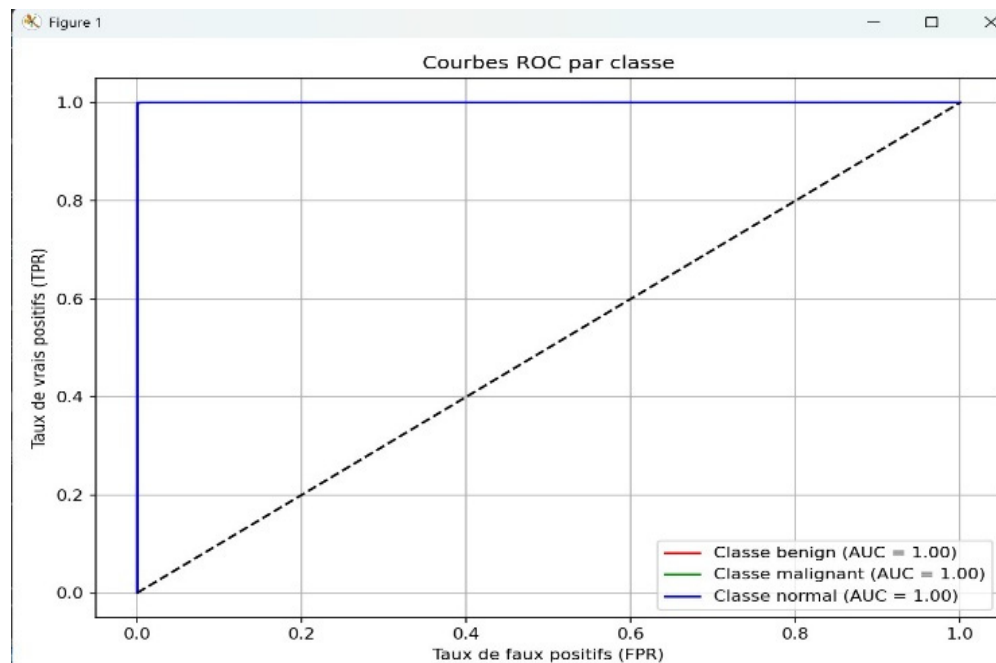


Figure IV.5 : Courbe ROC sur DDSM

IV.6.1.3. MIAS dataset

Les mêmes performances que les deux précédentes bases ont été obtenues : la **courbe ROC** sur **la base MIAS** illustre aussi de très hautes performances du modèle. L'**AUC pour cette troisième base** se rapproche aussi de 1 pour ce dataset, Le modèle entraîné sur cette base est sensé distinguer de manière parfaite les lésions bénignes des lésions malignes des mammographies.

IV.6.2 Comparaison des environnement Colab vs Kaggle vs PC local

Afin d'évaluer l'impact de l'environnement d'exécution sur les performances expérimentales, nous avons comparé trois plateformes : Google Colab, Kaggle et un PC local sur la base MIAS. Les critères étudiés sont le temps d'entraînement, l'utilisation mémoire GPU/CPU et la vitesse d'inférence.

Tableau 4.2 – Comparaison des environnements d'exécution (avec MIAS Dataset)

Environnement	GPU/CPU disponibles	Temps d'entraînement (16 époques)	Utilisation mémoire GPU/CPU	Vitesse d'inférence
Google Colab	GPU Colab (Tesla T4 ou P100), CPU (2,3 GHz, 12,7 Go de RAM)	sans GPU : 4h Avec GPU : 35-40 min	8 Go de RAM	533 images/s
Kaggle	GPU P100 (16 Go VRAM), CPU 2,2 GHz, RAM 13 Go	Avec GPU (1 h 7 min)	6 à 8 Go de RAM	200 images/s
PC local	Cœurs physiques : 4 Cœurs logiques : 8 Fréquence CPU : 1896.00 MHz	6h	RAM ~6,99 Go	8.75 images/s

IV.7. Tests de performances sur un ensemble de mammographies

Les performances du modèle entraîné pour la classification des mammographies sur un ensemble d'images mammographiques présentant des cas bénins, malins et des cas sans anomalies (sain / normal) ont été évaluées à l'aide du **modèle sauvegardé** sur le disque dur du pc local et sur google drive (pour les tests sur Colab).

Nous présenterons dans ce paragraphe les résultats des tests effectués via une interface graphique (créée avec python, voire Figure IV.6) pour la prédiction des lésions mammaires. Pour cela nous chargeons des images mammographiques pré-diagnostiquées par des radiologues (le diagnostic est connu à l'avance) puis nous estimons la capacité du modèle à détecter correctement les différents cas de la mammographie analysée.

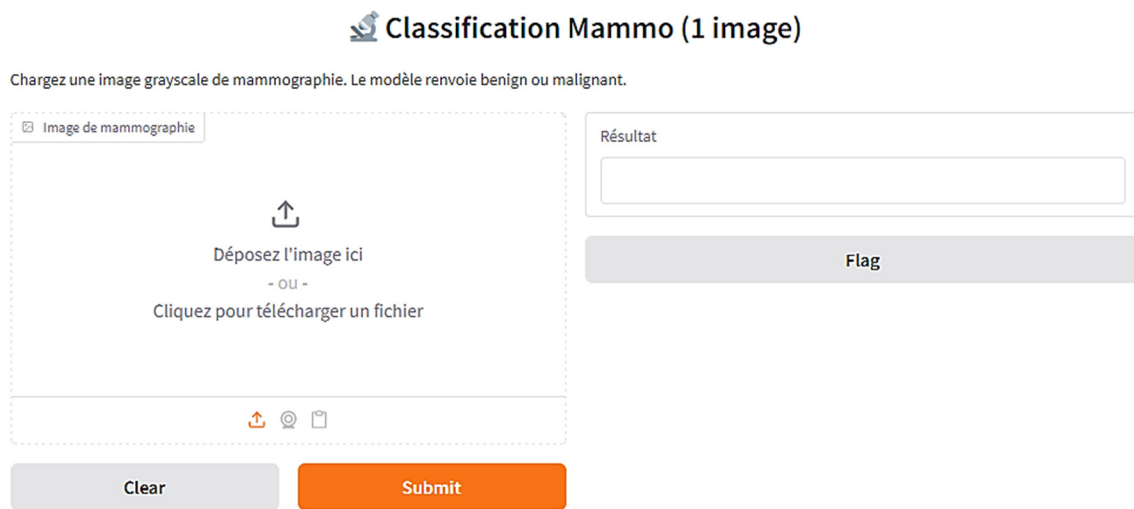


Figure IV.6: interface graphique de classification de mammographies

L'image de la figure IV.7 illustre un exemple de prédiction obtenue par notre modèle de prédiction de la mammographie. Il s'agit là, d'un cas avec une lésion bénigne d'après le rapport des radiologues et prédite correctement (bénin) par notre modèle.



Figure IV.7 : exemple de prédiction avec 100% de confiance.

Dans ce premier test on peut lire : **Classe : bénigne, confiance : 1** et **score [1.00, 0.00]** ce qui signifie que ce cas est idéalement prédit avec 100% de confiance (0% de probabilité qu'elle présente une pathologie quelconque).

Dans un deuxième exemple, l'analyse de l'image mammographique (figure figure IV.8) révèle aussi une probabilité importante d'un cas sain mais un peu loin de la certitude absolue (88% au lieu de 100%).



Figure IV.8 : exemple de prédiction avec 88% de confiance.

En fait, dans certains cas le modèle réalise des prédictions encore plus basses pour quelques cas de la base de test. Parfois le modèle fait des erreurs de classification et renvoie des fausses prédictions comme c'est le cas de la figure IV.9. :



Figure IV.9. Exemple de faux positif (54% de confiance)

Dans ce cas précis, la figure montre une mammographie avec une analyse du modèle qui suggère une possible lésion maligne : c'est-à-dire que l'analyse effectuée par le modèle indique que la probabilité d'un cancer du sein est de 54%, avec une probabilité de 46% pour un état bénin ou normal. Donc selon ces pourcentages, la probabilité d'une masse maligne est plus élevée que celle d'un état normal ou bénin. Ce qui veut dire que la décision finale du modèle est un cancer alors que dans la réalité (selon l'annotation des radiologues de cette mammographie) cette image est saine et ne présente aucune pathologie ! ce qui suggère un faux positif : Cette mammographie n'est pas correctement identifiée !

Visuellement l'erreur du modèle peut s'expliquer par une confusion dans la phase d'apprentissage très probablement causée par la présence d'une masse dense à bordures bien définies sur la mammographie (présence d'une petite masse claire au centre de la partie inférieure de cette mammographie). Un autre cas visuellement semblable à ce cas (avec une masse suspecte sur la mammographie, Figure IV.10) est prédit maligne par le modèle et avec 100% de confiance. Ce qui est correct selon le diagnostic des radiologues accompagnant cette image).



Figure IV.10. Exemple de vrai positif (100% de confiance)

On voit bien dans cet exemple, que la distribution des probabilités confirme bien le diagnostic des radiologues : la classe *malignant* atteint une probabilité de 100 %, tandis que la classe *non-maligne* est réduite à 0 %. Cette confiance très élevée indique que le modèle a reconnu les caractéristiques d'une tumeur cancéreuse avec une certitude très élevée comparativement au cas précédent sur lequel la modèle a hésité (offrant seulement 54%, de confiance). Ces résultats

traduisent donc la capacité du modèle à détecter d'une manière acceptable des cas pathologiques manifestes, ce qui est particulièrement pertinent pour une utilisation dans l'aide au diagnostic du cancer du sein.

IV.8. Conclusion

Les résultats obtenus grâce à notre modèle confirme sa grande capacité à bien classifier des images mammographiques. Les performances élevées obtenues, notamment une exactitude de validation pouvant atteindre 100% pour plusieurs cas, témoigne de la robustesse de l'apprentissage de notre modèle et ce, malgré la taille limitée des bases de données. D'ailleurs, certaines limites s'imposent et préconisent la nécessité d'un enrichissement du dataset mais avec moins d'images de faible qualité ou celles présentant des anomalies très confuses afin de faciliter l'apprentissage du modèle et améliorer sa robustesse. Or, l'analyse des performances montre que la qualité et la taille des datasets influencent fortement la capacité de généralisation des modèles et réduisent leurs performances lors des détections des anomalies.

Les contraintes matérielles constituent également un facteur déterminant. L'utilisation de GPU avec une VRAM suffisante permet d'augmenter la taille des lots (batches size) et d'accélérer l'entraînement en réduisant considérablement le temps de calcul même avec un nombre d'époques élevé. Tandis qu'un CPU seul entraîne des temps d'exécution très longs et impose une réduction du nombre de lots et le nombre d'époques. La RAM et la VRAM doivent donc être optimisées pour éviter les lenteurs et permettre un traitement efficace des images.

Conclusion générale

La lutte contre le cancer du sein demeure un défi majeur de santé publique, et les avancées technologiques récentes, en particulier celles issues de l'intelligence artificielle, offrent des perspectives déterminantes. Ce mémoire a montré que l'imagerie médicale, et plus spécifiquement la mammographie, reste l'outil de référence pour le dépistage et le diagnostic du cancer du sein, malgré certaines limites liées à la subjectivité et à la variabilité de l'interprétation humaine.

L'intégration de l'apprentissage profond dans ce domaine apparaît comme une approche prometteuse. Les expérimentations réalisées sur des bases de données reconnues (MIAS, DDSM, INBreast) ont mis en évidence la capacité des modèles à extraire automatiquement des caractéristiques pertinentes et à accompagner les radiologues dans leur prise de décision. Bien que des contraintes techniques et méthodologiques persistent, les résultats obtenus confirment l'efficacité et le potentiel des réseaux de neurones profonds pour l'analyse des mammographies.

Pratiquement, l'application de l'intelligence artificielle à l'imagerie médicale ouvre la voie vers des diagnostics plus rapides, précis et fiables. Néanmoins, des efforts restent nécessaires pour améliorer la qualité des données disponibles, optimiser les architectures de modèles et assurer une intégration clinique sécurisée et éthique.

À l'avenir, la synergie entre l'expertise médicale et les outils d'IA pourrait transformer en profondeur les pratiques de dépistage et de traitement, contribuant ainsi à sauver davantage de vies et à renforcer la performance des systèmes de santé.

Les références :

1. BOUKERROU, Zohra. **Imagerie médicale et détection des pathologies thoraciques**. Mémoire de Master. Université d'Alger 1, Faculté des Sciences, 2020.
2. DOUBROVINE, Tania. **Méthodes avancées pour la reconstruction tridimensionnelle en imagerie médicale**. Thèse de doctorat. Université de Rennes 1, 2016.
3. NATIONAL ELECTRICAL MANUFACTURERS ASSOCIATION (NEMA). **Digital Imaging and Communications in Medicine (DICOM) Standard**. Rosslyn (VA) : NEMA, 2020.
4. HAACKE, E. Mark, BROWN, Robert W., THOMPSON, Michael R. et VENKATESAN, Ramesh. **Magnetic Resonance Imaging: Physical Principles and Sequence Design**. 2^e éd. New York : Wiley-Blackwell, 2014
5. DELANNOY, Yvon et PEIX, Jean-Louis. **Physique des rayonnements : des rayons X à la tomographie par émission de positons**. Paris : Ellipses, 2006.
6. KHAN, Faiz M. **The Physics of Radiation Therapy**. 5^e éd. Philadelphia : Lippincott Williams & Wilkins, 2014.
7. 14.Gibaud,B.&Frandon,J. **Imagerie médicale : principes et applications**. Springer, 2016.
8. SZABO, Thomas L. **Diagnostic ultrasound imaging: inside out**. 2^e éd. Amsterdam : Academic Press, 2013.
9. BOUKHARI, Fatima Zohra. **Apport de la fluoroscopie dans la prise en charge des malformations congénitales cardiaques**. Thèse de doctorat en médecine. Université Mohammed V de Rabat, Faculté de Médecine et de Pharmacie, 2019..
10. EL KAMOUN, Salma. **Radiologie interventionnelle : place de la fluoroscopie dans le diagnostic et le traitement des pathologies digestives**. Thèse de doctorat en médecine. Université Cadi Ayyad, Faculté de Médecine de Marrakech, 2017.
11. BENSALD, Youssef. **Applications cliniques et principes physiques du scanner en radiologie médicale**. Thèse de doctorat en médecine. Université Mohammed V, Faculté de Médecine et de Pharmacie, Rabat, 2018.
12. RANGANATHAN, S. et KUMAR, V. **Radiology for Dentistry**. 2^e éd. New Delhi : Jaypee Brothers, 2015.
13. CARLSON, David et FRANK, Allan. **Essentials of Dental Radiography**. London : Quintessence Publishing, 2016.

14. MILIONI DE CARVALHO, Pablo. **Imagerie vasculaire du sein par rayons X, tridimensionnelle, quantitative et à faible dose.** Thèse de doctorat, Université Paris-Sud, École doctorale Sciences et Technologies de l'Information, des Télécommunications et des Systèmes, 2014.
15. BENABADJI, R. **Techniques et systèmes de mammographie numérique : principes et applications.** Mémoire de Master en Imagerie Médicale. Université de Tlemcen, 2017.
16. **DUFOUR, Bruno et DENIS-BACLET, Catherine.** *Imagerie diagnostique du sein : mammographie, échographie, IRM, techniques interventionnelles.* 2^e éd. Issy-les-Moulineaux : Elsevier Masson, 2009. 568 p. ISBN 978-2-294-01790-1.
17. **Bushberg, J.T. et al.** *The Essential Physics of Medical Imaging, 3^e édition.* , 2012.
18. **HUDA, Walter.** *Review of Radiologic Physics. 4^e éd. St. Louis : Mosby., 2010.*
19. **BENAYED, Manel.** *Apport du système BI-RADS dans l'interprétation des mammographies.* Thèse de doctorat en médecine. Alger : Université d'Alger, Faculté de Médecine, 2018.
20. **DUFOUR, Bruno et DENIS-BACLET, Catherine.** *Imagerie diagnostique du sein : mammographie, échographie, IRM, techniques interventionnelles.* 2^e éd. Issy-les-Moulineaux : Elsevier Masson, 2009.
21. **BENSLIMANE, Yasmine.** *Étude des lésions bénignes du sein diagnostiquées en imagerie médicale.* Thèse de doctorat en médecine. Alger : Université d'Alger, Faculté de Médecine, 2018.
22. **OUARET, Nadia.** *Étude radiologique des microcalcifications mammaires : corrélation anatomo-radiologique.* Thèse de doctorat en médecine. Oran : Université d'Oran 1, Faculté de Médecine, 2017.
23. **GUINDO, Ilias.** *Apport du couple échographie-mammographie dans le diagnostic des pathologies mammaires.* Mémoire de spécialité. Bamako : Université des Sciences, des Techniques et des Technologies de Bamako, 2022.
24. **HOLLAND, James F., FREI, Emil et al.** *Cancer Medicine.* 8^e éd. Shelton (CT) : People's Medical Publishing House, 2010.
25. **MOON, Greenberg (ou éd.), éd.** *Breast Cancer Prevention and Treatment.* Springer. 2016.
26. **Cancer du sein. Fiches d'information.** [En ligne] OMS (Organisation mondiale de la Santé), 14 août 2025. <https://www.who.int/fr/news-room/fact-sheets/detail/breast-cancer>.
27. **DJEMAI, Yamina.** *Étude descriptive des cas de cancer du sein invasif enregistrés dans la Wilaya de Ouargla de 2015 à 2019.* Mémoire de Master, Université Kasdi Merbah – Ouargla, Faculté des Sciences de la Nature et de la Vie, 2019.
28. **KOUADRI, Mouna.** *Étude anatomo-clinique et radiologique du cancer du sein.* Thèse de Doctorat en Médecine, Université d'Alger 1, Faculté de Médecine, 2017.
29. **BOUNAB, Lila.** *Étude anatomo-clinique et radiologique du cancer du sein : à propos de 100 cas.* Thèse de Doctorat en Médecine, Université d'Alger 1, Faculté de Médecine, 2016.
30. **DABAS, Nidhi et GUPTA, Rajesh.** *Breast Cancer: Biology, Prevention and Treatment.* Singapore : Springer, 2022.

31. **ABDELKADER, Nouria.** *Profil épidémiologique et anatomopathologique du cancer du sein : étude de 200 cas.* Thèse de Doctorat en Médecine, Université d'Oran 1, 2019.
32. **MERZOUK, Siham.** *Prédisposition génétique au cancer du sein : étude des gènes BRCA1 et BRCA2.* Mémoire de Master en Biologie, Université de Tlemcen, 2017.
33. **AMERICAN CANCER SOCIETY.** *Breast Cancer Facts & Figures 2021-2022.* Atlanta : American Cancer Society, 2021.
34. **Inview Imaging.** *Mammogram 2024 : Top 5 Breakthrough Technologie.* [En ligne] 8 octobre 2024. . <https://www.inviewimaging.com/2024/10/08/mammogram-2024-top-5-breakthrough-technologies/>.
35. **CHOLLET, François.** *Deep Learning with Python.* Shelter Island : Manning Publications, 2017. 384 p. ISBN 978-1617294433.
36. **FLAMENT, Marie-Pierre.** *L'intelligence artificielle au service de la recherche clinique : Quel est son impact dans la recherche clinique ?* Mémoire de master, Université de Lille, 2019. 82 p.
37. **MINSKY, Marvin.** *The Society of Mind.* New York : Simon & Schuster, 1986. 368 p. ISBN 978-0671657130.
38. **LEFEBVRE, Sophie.** *Applications de l'intelligence artificielle dans l'analyse d'images médicales.* Mémoire de master, Université de Strasbourg, 2021. 95 p.
39. **BENOIT, Laure.** *Intelligence artificielle pour la détection précoce de cancers dans les images mammographiques.* Mémoire de master, Université de Rennes, 2020. 92 p.
40. **BISHOP, Christopher M.** *Pattern Recognition and Machine Learning.* New York : Springer, 2006. ISBN 978-0-387-31073-2.
41. **LeCun, Yann; Bengio, Yoshua; Hinton, Geoffrey.** *Deep learning.* *Nature*, 2015, vol. 521, no 7553, p. 436-444.
42. **SHEN, Li, et al.** *Deep Learning to Improve Breast Cancer Early Detection on Screening Mammography.* *Scientific Reports*, 2019, vol. 9, no 12495. DOI : 10.1038/s41598-019-48995-4.
43. **Agarwal, Richa ; Yap, Moi Hoon ; Hasan, Md. Kamrul ; Zwiggelaar, Reyer ; Martí, Robert.** *Deep Learning in Mammography Breast Cancer Detection.* In: *Book: BreastCancer_5_*, 2022.
44. **FERRARI, Alessandro.** *Feature Extraction and Classification Methods for Medical Image Analysis.* Thèse de doctorat. Politecnico di Milano, 2012.
45. **ERHAN, Dumitru.** *Training Deep Architectures for AI.* Thèse de doctorat. Université de Montréal, 2009.
46. **LECUN, Yann, BENGIO, Yoshua et HINTON, Geoffrey.** *Deep learning.* *Nature*, 2015, vol. 521, n° 7553, p. 436-444. DOI: 10.1038/nature14539.

47. Gyawali, Dipesh, et al. *Comparative Analysis of CPU and GPU Profiling for Deep Learning Models*. 2023
48. *Deep Learning in Breast Cancer Imaging: A Decade of Progress and Future Directions* — Luo et al., 2023.
49. BENHAMOU, N. *Application de l'apprentissage profond à l'imagerie médicale : détection automatique de pathologies*. Thèse de doctorat. Université de Paris-Saclay, 2020.
50. DUBOIS, M. *Conception de systèmes portables intelligents pour l'assistance aux personnes malvoyantes*. Mémoire de Master. Université de Lille, 2018.
51. BENSLIMANE, S. *Application de l'apprentissage profond à l'analyse des séquences génomiques*. Mémoire de Master. Université de Constantine, 2021.
52. BOUKERCHE, A. *Apprentissage supervisé et non supervisé pour la reconnaissance d'images médicales*. Mémoire de Master. Université de Tizi-Ouzou, 2020.
53. BENZAADA, S. *Méthodes hybrides d'apprentissage semi-supervisé pour l'analyse de données médicales*. Thèse de doctorat. Université de Lyon 1, 2019.
54. LACROIX, A. *Applications des réseaux de neurones artificiels au traitement automatique du langage naturel*. Thèse de doctorat. Université de Toulouse, 2018.
55. HASTIE, Trevor, TIBSHIRANI, Robert et FRIEDMAN, Jerome. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2^e éd. New York : Springer, 2009. ISBN 978-0-387-84858-7.
56. **Goodfellow, Ian ; Bengio, Yoshua ; Courville, Aaron**. *Deep Learning*. Cambridge (Massachusetts) : MIT Press, 2016.
57. **Benali, Houda**. *Classification d'images médicales par réseaux de neurones convolutionnels*. Mémoire de Master, Université de Constantine 2, 2020.
58. **Benaouda, Abdelhakim**. *Utilisation des réseaux de neurones récurrents pour la prévision de données séquentielles*. Mémoire de Master, Université de Constantine 2, 2020.
59. **GOLDBERG, Yoav**. *Neural Network Methods for Natural Language Processing*. San Rafael (California) : Morgan & Claypool Publishers, 2017.
60. **HAYKIN, Simon**. *Neural Networks and Learning Machines*. 3^e éd. Upper Saddle River (New Jersey) : Pearson, 2009.
61. **Bounoua, Chahinez**. *Étude comparative entre l'apprentissage supervisé et non supervisé pour la classification de données*. Mémoire de Master, Université de Béjaïa, 2020.

62. **Giger, M.L. ; Suzuki, K.** *Computer-Aided Detection and Diagnosis in Medical Imaging*. Boca Raton: CRC Press, 2018.
63. PYTHON.DOCTOR. Deep learning & python : C'est quoi le deep learning ? [en ligne]. Paris : Python.doctor, s.d. Disponible à : <https://python.doctor/page-deep-learning-python>
64. De Moor, T., Rodriguez-Ruiz, A., Mérida, A. G., Mann, R., Teuwen, J., et al. *Automated soft tissue lesion detection and segmentation in digital mammography using a U-Net deep learning network*. arXiv preprint, 2018.
65. SCHOPF, C. M., et al. *Artificial Intelligence-Driven Mammography-Based Future Cancer Risk Prediction*. Journal of the American College of Radiology, 2024.
66. FRIEDEWALD, S. M., et al. *Triaging mammography with artificial intelligence*. PubMed / NCBI, 2025.
67. **Jean-Edmond DASSE.** *Étude des méthodes de reconnaissances d'empreinte digitale à l'aide du deep learning*. Mémoire de Master, option Bases de Données et Génie Logiciel, Université (non précisée), 2021.
68. *A multi-model machine learning framework for breast cancer risk stratification using clinical and imaging data*. Miao, Lu et al., 2025.
69. Ali, R., Zahoor, S., Riaz, H. et al., 2025. *Hybrid convolutional neural network and bi-LSTM model with EfficientNet-B0 for high-accuracy breast cancer detection and classification*. **Computers in Biology and Medicine**, vol. 184, 109944. DOI 10.1016/j.combiomed.2025.109944.
70. **Alotaibi, A.**, *ViT-DeiT : An Ensemble Model for Breast Cancer Histopathological Images Classification*. 2022.
71. **Mehmood, A.**, *A Novel Channel Boosted Residual CNN-Transformer with Regional-Boundary Learning for Breast Cancer Detection*. 2025.
72. **Intelligence Artificielle School**, 2025. *U-net : le réseau de neurones par vision par ordinateur*. [En ligne] Disponible sur : <https://www.intelligence-artificielle-school.com/ecole/technologies/u-net-le-reseau-de-neurones-par-vision-ordinateur/>
73. **Ravishankar, H., et al.**, 2017. *Deep Learning for Mammography: Dataset and Algorithm Evaluation*. *Journal of Digital Imaging*, 30(4), pp. 495–507.
74. **Heath, M., Bowyer, K., Kopans, D., Moore, R., & P. Kegelmeyer**, 2001. *The Digital Database for Screening Mammography (DDSM)*. Thèse technique, Université du Massachusetts, USA.

75. Kopans, D. B., 2007. *Breast Imaging: The Digital Era and Databases*. Thèse de doctorat, Harvard University, USA.
76. Moreira, I. C., et al., 2012. *INbreast: Toward a Full-Field Digital Mammographic Database*. Mémoire de master, Université de Porto, Portugal.
77. Suckling, J., Parker, J., Dance, D. R., Astley, S. M., Hutt, I., et al., 1994. *The Mammographic Image Analysis Society Digital Mammogram Database (MIAS)*. Thèse de master, Université de Nottingham, UK.
78. Vijayakumar, D., et al., 2020. *VINDR-MAMMO: Vietnamese Mammography Dataset for AI Research*. Thèse de doctorat, Université de Hanoi, Vietnam.
79. **Rodriguez-Ruiz, A., et al.**, 2019. *Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists*. *JAMA Oncology*, 5(8), pp. 1090–1098.
80. **Kooi, T., et al.**, 2017. *Large Scale Deep Learning for Computer-Aided Detection of Mammographic Lesions*. *Medical Image Analysis*, 35, pp. 303–312.
81. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I., 2017. *A Survey on Deep Learning in Medical Image Analysis*. *Medical Image Analysis*, 42, pp. 60–88.
82. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D., 2019. *Key Challenges for Delivering Clinical Impact with Artificial Intelligence*. *BMC Medicine*, 17(1), 195.
83. **DANG, Lan-Anh**. *Impact de l'intelligence artificielle en mammographie de dépistage du cancer du sein*. Thèse de médecine, Université de Lille, 2021
84. ZHANG, Y., WANG, X. et LI, J. *An Empirical Study on the Usage of Visual Studio Code in Machine Learning Projects*. In: 2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER). IEEE, 2021. p. 145–156. DOI: 10.1109/SANER50967.2021.00025.
85. **GOOGLE**. *Colaboratory – Frequently Asked Questions*. [en ligne]. Google Research, 2025. Disponible à l'adresse : <https://research.google.com/colaboratory/faq.html>
86. **KAGGLE**. *Kaggle Notebooks Documentation*. [en ligne]. Disponible à l'adresse : <https://www.kaggle.com/docs/notebooks>
87. **SUCKLING, J., et al.** *The Mammographic Image Analysis Society Digital Mammogram Database*. Extrait de: Y. Gale, éd. *Proceedings of the 2nd International Workshop on Digital Mammography*. York, Royaume-Uni: Excerpta Medica, 1994, p. 375-378.

88. **HEATH, M., BOWYER, K., KOPANS, D., MOORE, R. et KEGELMEYER, P.** The Digital Database for Screening Mammography. In: *Proceedings of the 5th International Workshop on Digital Mammography*. Madison, WI: Medical Physics Publishing, 2001, p. 212-218.
89. **MOREIRA, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J. et Cardoso, J. S.** INbreast: Toward a Full-field Digital Mammographic Database. *Academic Radiology*, 2012, vol. 19, no. 2, p. 236-248. DOI: 10.1016/j.acra.2011.09.014.
90. **CHEN, L., Margolies, L. R., Ding, J., et al.** Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Scientific Reports*, 2019, vol. 9, article 12495. DOI: 10.1038/s41598-019-48995-4.
91. **GERAS, K. J., Wolfson, S., Kim, S. G., Moy, L. et Cho, K.** High-resolution breast cancer screening with multi-view deep convolutional neural networks. *arXiv preprint arXiv:1703.07047*, 2017. Disponible à l'adresse : <https://arxiv.org/abs/1703.07047>
92. **Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S. (2019).** *PyTorch: An Imperative Style, High-Performance Deep Learning Library*. arXiv. <https://doi.org/10.48550/arXiv.1912.01703>
93. **Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011).** *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830.

Résumé

Les avancées technologiques ont profondément transformé le diagnostic et la prise en charge des maladies. Le cancer du sein, première cause de mortalité féminine et le cancer le plus fréquent chez la femme (OMS, 2023), nécessite un dépistage précoce et fiable pour améliorer les chances de survie. La mammographie reste l'examen de référence, mais son interprétation demeure complexe et sujette à des erreurs. L'intelligence artificielle, en particulier l'apprentissage profond, offre de nouvelles perspectives grâce à l'analyse automatisée des images. Ces approches permettent d'extraire des caractéristiques pertinentes et d'assister les radiologues dans leurs décisions, améliorant ainsi précision et rapidité. Ce travail applique l'apprentissage profond à l'analyse mammographique, avec des résultats prometteurs dépassant 98 % de précision sur trois bases de données : MIAS, INBreast et DDSM.

Mots clés : Apprentissage profond, intelligence artificielle, diagnostic, cancer du sein.

Abstract:

Technological advances have significantly transformed disease diagnosis and management. Breast cancer, the leading cause of female mortality and the most common cancer among women (WHO, 2023), requires early and accurate detection to improve survival rates. Mammography remains the reference screening tool, yet its interpretation is complex and prone to errors. Artificial intelligence, especially deep learning, offers new opportunities through automated image analysis. These methods extract relevant features and assist radiologists, enhancing diagnostic accuracy and speed. This study applies deep learning to mammographic analysis, achieving promising results with over 98% accuracy on three major databases: MIAS, INBreast, and DDSM.

Keywords: Deep learning, Artificial intelligence, Diagnosis, Breast cancer.

ملخص

شهدت التكنولوجيا الطبية تطوراً كبيراً غير طرق التشخيص والعلاج. ويُعد سرطان الثدي أكثر أنواع السرطان انتشاراً بين النساء وأحد أهم أسباب الوفيات (منظمة الصحة العالمية، 2023). لذلك، يصبح الكشف المبكر والدقة في التشخيص أمراً ضرورياً لرفع فرص الشفاء. تبقى الماموغرافيا الوسيلة الأساسية للفحص، غير أن تفسير صورها قد يكون صعباً ويؤدي أحياناً إلى أخطاء. هنا يبرز دور الذكاء الاصطناعي، خاصة التعلم العميق، الذي يوفر أدوات تحليل آلي قادرة على استخراج الخصائص المهمة ومساندة الأطباء في اتخاذ القرار، مما يزيد من سرعة ودقة التشخيص. يركز هذا البحث على تطبيق التعلم العميق في تحليل صور الماموغرافيا، حيث حقق دقة تتجاوز 98% على ثلاث قواعد بيانات رئيسية MIAS وINBreast وDDSM.

التعلم العميق، الذكاء الاصطناعي، التشخيص، سرطان الثدي :الكلمات المفتاحية

Taqbaylit

Tazmert n tiknulujiyin tebeddel atas n tarrayin n usnifel d usellek n yexfawn. **Almus n tammut** (cancer du sein) d amezwaru n yiberdan i yezwaren ar timyarin, yerna d sebba tameqqrant n wamut (OMS, 2023). Ihi, assay n yifassen d usmad n tidet d ayen ilaqen i wakken ad ttwarnu tudert. Mammugrafi d tarrayt tameqqrant n usnifel, maca tafsirt n tugniwin-is tezmer ad d-ttawi-d tucqdiwin. Dagi i d-yettban lhağ n tazrurt tamenttut (IA) s usuldaf amaynut (deep learning), i izmer ad yesleh tafsirt n tugniwin d ad imudd afus i yimediyazen n radioloji. Asay-agi isenfel usuldaf deg usnifel n mammugrafi, yerna yebya ad yesleh-d tasiregt s tiddawin uffiren (ugar n 98%) deg krađ n tyezwa imqqaren : MIAS, INBreast, d DDSM.

Awalen yufraren : Asuldaf, tazrurt tamenttut, usnifel, almus n tammut.