

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة سعد دحلب البليدة
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Master

Mention : Génie Biomédical
Spécialité : Instrumentation Biomédical

Présenté par :

Ferhah Hizia

&

Aroune Nesrine

Détection Automatique des Microcalcifications par les Images Mammographiques

Proposé par : Dr. L. Ait Mohamed

&

Pr. A. Cherfa

Année Universitaire : 2021/2022

Remerciements

Nous remercions Dieu de nous avoir accordé la force afin d'accomplir ce travail.

Nous remercions vivement nos encadreurs, madame **AIT MOHAMED** et madame **ASSIA CHERFA** d'avoir accepté d'encadrer ce travail. Nous les remercions pour la compréhension, l'encouragement, le soutien moral et scientifique indescriptibles accordé tout au long de ce travail. Ils ont su nous conseiller et nous renseigner lorsque nous avons besoin d'aide.

Leur écoute nous a permis de travailler efficacement. Nous leur adresse nos chaleureux remerciements pour m'avoir accordé de leur temps ainsi que leur disponibilité et la pertinence de leurs remarques. Qu'ils trouvent ici l'expression de notre profonde gratitude.

Mes remerciements vont également aux membres de jury, pour l'intérêt qu'ils ont porté à notre mémoire en acceptant d'examiner notre travail.

Aussi nous tiens à remercier le centre de développement des technologies avancées (CDTA), spécialement le directeur du centre, de m'avoir permis de passer notre stage de fin d'études.

Enfin nos remerciements vont à tous mes enseignants qui nous ont accompagnés durant tous notre cursus à travers leurs connaissances, soutiens, et conseils.

Dédicace

Nous tenons à dédier ce mémoire en particulier et en premier lieu à nos chers parents ;
Qui ont consacré leurs vies à l'éducation de leurs enfants,

Qui nous ont élevé avec amour et tendresse, et

Qui ont toujours cru en nous. A nos chers frères et sœurs, à tous nos proches.

A tous les membres de la famille Ferhah & Aroune ;

A toutes les personnes que nous portons dans notre cœur.

A tous nos amis(es) sans exception.

Aux étudiants de la spécialité génie biomédicale PROMO 2022.

Et à la fin, une dédicace spéciale à tous nos enseignants durant notre cursus scolaire, du
primaire à l'université.

Merci à toutes les personnes qui nous ont aidées de près ou de loin.

المخلص:

التصوير الإشعاعي للثدي هو أكثر تقنيات التصوير شيوعا للكشف عن الأورام في مرحلة مبكرة. انه حاليا التحقيق الرئيسي في فحص سرطان الثدي. يعد وجود التكلسات في صور التصوير الشعاعي للثدي أمرا مثيرا للاهتمام بشكل خاص للكشف المبكر عن سرطان الثدي. سنركز في هذا العمل على استخراج ثلاث معلومات مهمة للتكلسات الدقيقة: مواضعها وأرقامها وأشكالها تشمل هذه المذكرة تحليل صور إشعاعية للثدي بهدف استخراج حبيبات الكالسيوم غالبا ما تمثل المراحل الأولى لتكوين سرطان الثدي باستعمال طرق مختلفة للتركيبية الوسائل الرياضية والاندماج K.

الكلمات المفتاحية: الماموغرافية، التكلسات الدقيقة، الوسائل الرياضية، الاندماج K، التركيبية.

Résumé :

La mammographie est la technique d'imagerie la plus utilisée pour détecter des tumeurs à un stade précoce. Elle est actuellement la principale investigation dans le dépistage du cancer du sein. La présence de calcifications dans des images mammographiques sont particulièrement intéressante pour la détection précoce du cancer du sein. Dans ce travail nous nous intéresserons à l'extraction de trois paramètres importants des microcalcifications : leurs positions, leurs nombres et leurs formes, et cela par la segmentation des Mcs à partir des mammographies. Nous avons pour cela développé une approche basée sur les outils morphologiques, la classification par K-means et la fusion.

Mots clés : mammographie, microcalcification, morphologie mathématique, k-means, la fusion.

Abstarct:

Mammography is the most commonly used imaging technique to detect tumors at an early stage. It is currently the main investigation in breast cancer screening. The presence of calcifications in mammography images is particularly interesting for the early detection of breast cancer. In this work we will be interested in the extraction of three important parameters of microcalcifications: their positions, their numbers and their shapes, and this by the segmentation of Mcs from mammograms. For this, we have developed an approach based on morphological tools, classification by K-means and fusion.

Keywords: mammography, microcalcification, mathematical morphology, k-means, fusion.

Table de matière

Remerciements	I
Dédicace	II
Résumé	III
Table de matière	IV
Liste des figures	VII
Liste des tableaux	X
Liste des abréviations	XI
Introduction Générale	I

Chapitre 1 : Aspect médical

1.1. La mammographie	4
1.1.1. La mammographie analogique	4
1.1.2. La mammographie numérique	6
1.2. Les incidences en mammographie	7
1.2.1. Incidence cranio-caudale	7
1.2.2. Incidence de profil	7
1.2.3. Incidences complémentaires	8
1.3. L'anatomie du sein	8
1.4. Que voit-on sur une mammographie et à quoi est dû le contraste observé ?	10
1.5. Pathologies mammaires	11
1.5.1. Tumeurs bénignes	11
1.5.2. Tumeurs malignes (cancer du sein)	12
1.6. Anomalies mammaires radiologiques	13
1.6.1. Opacités mammaires	14
1.6.2. Calcifications mammaires	14
1.6.3. Classification de Le Gal des microcalcifications mammaires	15
1.7. Discussion	17

Chapitre 2 : Etat de l'art

2.1. Détection des microcalcifications : état de l'art	19
2.2. Tableau récapitulatif	27
2.3. Discussion	30

Chapitre 3 : Outils Mathématiques Utilisés

3.1. Apprentissage profond	32
3.1.1. Réseaux de neurones	33
3.1.2. Réseaux de Neurones Convolutifs	33
3.2. Morphologies mathématiques	38
3.2.1. Élément structurant	38
3.2.2. Dilatation	39
3.2.3. Erosion	39
3.2.4. Fermeture morphologique	40
3.2.5. Ouverture morphologique	40
3.2.6. Chapeau haut de forme (CHF).....	41
3.3. Fusion des données	42
3.3.1. Définition de la fusion d'images.....	43
3.4. Approche proposée	44
3.4.1. Détection automatique de la région d'intérêt	45
3.4.2. Détection automatique par apprentissage en profondeur : Approche 1	45
3.4.3. Détection automatique de la ROI par les outils morphologiques : Approche 2	48
3.4.4. Segmentation des microcalcifications.....	50
3.4.5. Segmentation par K-means.....	51
3.4.6. Fusion de plusieurs segmentations alternatives.....	53
3.5. Discussion	55

Chapitre 4 : Résultats et Interprétations

4.1. Base de données	57
4.2. Environnement et langage d'implémentation	58
4.2.1. Logiciel et matériel	58
4.2.2. Langage de programmation utilisé.....	58
4.3. Implémentation de l'algorithme	61
4.4. Résultats de la détection automatique de la région d'intérêt.....	62
4.4.1. Résultats de détection automatique de la ROI par apprentissage en profondeur : Approche 1	62
4.4.2. Résultats de détection automatique de la ROI par les outils morphologiques : Approche 2	65
4.4.3. Résultats de la segmentation des microcalcifications	71

4.4.4. Résultats de la labélisation par K-means.....	72
4.4.5. Résultats de la segmentation des microcalcifications par le CHF seuillé	73
4.4.6. Résultats de la segmentation des microcalcifications par la fusion	75
4.4.7. Résultats expérimentaux de la segmentation.....	76
4.5. Présentation de l'interface graphique	78
4.6. Application de l'interface.....	81
4.7. Discussion	84
Conclusion Générale	85
Références bibliographiques.....	88

Liste des figures

Figure 1.1: Mammographie analogique	5
Figure 1.2 : Mammographie numérique	6
Figure 1.3 : Mammographie ; incidence de face	7
Figure 1.4 : Mammographie ; incidence de profil	8
Figure 1.5 : Incidences complémentaires	8
Figure 1.6: Anatomie du sein	9
Figure 1.7: Mammographie d'un sein normal en vue de profil	10
Figure 1.8: Dystrophie fibrokystique	12
Figure 1.9 : Adénofibrome	12
Figure 1.10 : Tumeur maligne	13
Figure 1.11: Classification de Le Gal des microcalcifications mammaires.	16
Figure 1.12 : Mammographie montrant un foyer de microcalcifications	16
Figure 1.13 : Agrandissement de la région contenant l'amas de microcalcifications	17
Figure 3.1: Deep Learning sous-discipline du Machinelearning	33
Figure 3.2: Architecture standard d'un réseau de neurone convolutionnel	34
Figure 3.3: Convolution.....	35
Figure 3.4: Opérations pooling	36
Figure 3.5 : Fonction ReLU	37
Figure 3.6 : L'architecture de Fully-connected	38
Figure 3.7 :Exemples d'élément structurant.(a) Élément structurant linéaire, (b) élément structurant cercle, (c) élément structurant carré.....	38
Figure 3.8: Niveaux de traitements de la fusion d'images.	42
Figure 3.9: Organigramme général de la méthode proposée.....	45
Figure 3.10 : Division de l'image mammographique.	46
Figure 3.11: Organigramme global de l'étape de détection des Mcs.	47
Figure 3.12 : Architecture de ResNet 50	48
Figure 3.13 : Détection automatique de la ROI.	49
Figure 3.14:Organigramme de l'étape de segmentation des Mcs.....	51
Figure 3.15 : Organigramme du Clustering par K-means.	52

Figure 4.1: Exemples d'images de la base des données [2]. (a) Image CC_prior, (b) image CC_recent, (c) image MLO_prior, (d) imageMLO_recent.....	57
Figure 4.2: Interface d'Anaconda.	59
Figure 4.3: Description de l'interface Spyder.	61
Figure 4.4: Schéma des différentes étapes de l'approche proposée.	62
Figure 4.5: Exemples de classification des patches en classe Mcs et non Mcs. (a) Patch répertorié en classe Mcs, (b) patch répertorié en classe non Mcs.	63
Figure 4.6 : Matrice de confusion.	64
Figure 4.7 : Deux exemples d'images mammographiques.....	66
Figure 4.8: Masque du sein sans le mamelon, l'aréole et la peau. Colonne (a) image originale, colonne (b), (c), (d) image (a) érodée avec un élément structurant de taille (5x5), (9x9), et (19x19) respectivement.	66
Figure 4.9: Résultats du prétraitement du sein. Colonne (a) image originale, colonne (b) image prétraitée.....	67
Figure 4.10: Résultats de l'application du chapeau haut de forme. (a) image prétraitée, (b), (c), (d) et (e) application du CHF avec élément structurant de taille (3x3), (5x5), (7x7) et (9x9) respectivement.	68
Figure 4.11: Résultats de l'étape de segmentation par CHF seuillée. Colonne (a) image originale, colonne (b) image vérité de terrain, colonne (b) image après application du CHF et colonne (c) image CHF seuillée.....	69
Figure 4.12: Exemple des étapes de détection de la ROI.	70
Figure 4.13: Résultats de l'étape de détection des Mcs. Colonne (a) image vérité de terrain, Colonne (b) image originale, colonne (c) masque du foyer des Mcs détectés et colonne (d) foyer des Mcs détectés en NG.....	71
Figure 4.14: Résultats de la classification par K-means avec différents nombres de cluster k. Colonne (a) Image originale, colonnes (b), (c) et (d) k=5, k=10 et =15 respectivement.	72
Figure 4.15: Résultats de la segmentation par CHF avec différents éléments structurants. Colonnes (a) image de vérité de terrain, (b), (c), (d) et (e) application du CHF avec élément structurant (3x3), (5x5), (7x7) et (9x9) respectivement.	74
Figure 4.16: Exemple des différentes images seuillées.....	75
Figure 4.17: Résultats de l'étape de la fusion par l'union. La colonne (a) représente la vérité de terrain, La colonne (b) représente les résultats de la labélisation K-means, la	

colonne (c) représente les résultats du CHF seuillé, la colonne (d) représente les résultats de la fusion.	76
Figure 4.18: Résultats de la fusion. La colonne (a) représente l'image originale, colonne (b) image de vérité de terrain, colonne (c) résultats de la fusion par union, colonne (d) résultats de la fusion par intersection.	78
Figure 4.19: Résultats de l'approche proposée. La colonne (a) représente la ROI détectée automatiquement par CHF-, la colonne (b) représente la superposition des résultats de fusion avec GT, la colonne (c) représente la superposition des résultats de fusion avec la vérité terrain (GT) sur l'image en NG.	78
Figure 4.20: Fenêtre de système d'authentification.	79
Figure 4.21: Fenêtre de l'interface.	80
Figure 4.22: Lecture d'image.	81
Figure 4.23 : Détection automatique de la région d'intérêt.	82
Figure 4.24: Image à cropper.	83
Figure 4.25 : Détection semi-automatique de la région d'intérêt.	83
Figure 4.26 : Image segmentée.	84

Liste des tableaux

Tableau 1-1:Contraste dans une image mammographique.....	10
Tableau 1-2: Approche de Lamarque	11
Tableau1-3: Classification de le Gal	15
Tableau 2-1:Tableau récapitulatif de toutes les revues étudiées.....	30
Tableau 4-1 : Résultats de la précision en fonction du nombre d'époques.....	65
Tableau 4-2: Résultats des différentes étapes de la méthode proposée, en termes de dice et d'OF.....	77

Liste des abréviations

BMD: Breast Micro-Calcification Dataset
CC : Cranio-Caudal
CDTA : Centre de Développement des Technologies Avancées
CHF : Chapeau Haut de Forme
CHFB : Chapeau Haut de Forme Blanc
CNN: Convolutional Neural Network
DL: Deep Learning
ES : Élément Structurant
FN : Faux Négatifs
FP : Faux Positifs
GT : Grown Truth
GUI : Graphical User Interface
IA : Intelligence Artificiel
INT : Intersection
Mcs : Microcalcifications
MLO : Médio-Latéral-Oblique
NG : Niveau de gris
OF : Overlap Fraction
ReLU : Rectified Linear Units
Res-Net: Res-Network
ROI: Region Of Interest
SI : Indice de Similarité
UNI : Union
VN : Vrais Négatifs
VP : Vrais Positifs

Introduction Générale

Introduction Générale

Le traitement et l'analyse d'images médicales est un domaine en pleine expansion où le problème de la recherche automatique d'informations contenues dans les images médicales s'avère un besoin pressant. En effet, la grande diversité des dispositifs d'imagerie médicale, la difficulté d'interprétation de ces images ainsi que leur grand nombre, engendrent un travail fastidieux pour ceux qui doivent les interpréter. Pour traiter cette grande masse d'informations, les médecins s'orientent actuellement vers l'utilisation des systèmes d'aide à l'analyse et à l'interprétation de ces images. Cette analyse vise à faciliter le diagnostic posé par le praticien et à le rendre le plus précis et le plus fiable possible.

Le cancer du sein reste le cancer le plus fréquent chez la femme et l'une des premières causes de décès parmi les cancers gynécologiques. Beaucoup d'efforts ont été effectués sous forme de programmes de dépistage pour une prévention et une détection précoce. L'imagerie mammographique est l'une des modalités les plus utilisées. Ces images mammographiques fournissent des informations radiologiques sur la glande mammaire qui sont, malheureusement, extrêmement difficiles à exploiter par l'expert humain qui s'appuie sur sa vision et son expérience.

Suite à ce constat et au dernier développement des méthodes d'analyse d'image est venue l'idée de développer un système de détection des microcalcifications, afin d'offrir une aide complémentaire pour le radiologue et qui constitue un second lecteur potentiel.

Notre projet de fin d'études a eu pour but d'aider les radiologues à dépister le cancer du sein à un stade précoce. Ceci s'effectue à l'aide d'un processus de détection des microcalcifications (Mcs) à partir d'images mammographiques.

Ce projet de fin d'études est l'aboutissement d'une suite de travaux de recherche réalisés dans le cadre d'obtention du diplôme du Master et une collaboration avec le CDTA (Centre de Développement des Technologies Avancées) et qui porte sur l'analyse des images mammographiques pour l'aide à la détection du cancer du sein.

Pour plus de clarté, nous avons organisé l'ensemble de nos travaux en plusieurs chapitres. En plus de cette partie introductive, ce mémoire est constitué de quatre chapitres auxquels s'ajoute une conclusion qui résume les principaux résultats.

- Après cette introduction générale, **le chapitre 1** sera consacré à la présentation du problème du cancer du sein, son dépistage, son diagnostic et les lésions radiologiques qui l'affectent, nous nous intéressons aux foyers de Mcs et aux masses. Une brève

description anatomique du sein sera donnée, afin de mieux expliciter les différentes structures à localiser, et à les identifier sur les images mammographiques. Nous aborderons également l'appareil de mammographie et son fonctionnement, puisque c'est la modalité que nous avons utilisée.

- Dans le **chapitre 2**, qui s'intitule « État de l'Art », nous avons cité et donné un aperçu général des travaux déjà présentés par la communauté scientifique, liés aux méthodes de détection des Mcs, Nous exposerons-différentes architectures et techniques qui ont été proposées ces dernières années pour la détection des Mcs.
- Nous donnerons en détail dans le **chapitre 3**, l'approche automatique proposée pour le prétraitement des images, afin d'améliorer la détection des Mcs qui se divise en deux étapes. La première étape est la détection des Mcs, on ne donne pas le détail dans l'introduction. La deuxième étape consiste à segmenter des patches avec des Mcs on utilisant des morphologies mathématique tel que le k-means, top-hat et la fusion.
- Dans le dernier chapitre, nous présenterons et discuterons les résultats obtenus avec l'approche proposée, pour la détection automatique des Mcs. Différents outils d'évaluation des résultats, sont également présentés dans cette partie, ainsi qu'une brève présentation de l'interface réalisée.

En conclusion générale, nous présenterons un récapitulatif des étapes de notre travail, les difficultés rencontrées lors de l'implémentation de l'algorithme, et finalement, es perspectives que nous souhaitons proposer pour les travaux futurs.

Chapitre 1 :
Aspect médical

L'évolution des techniques d'imagerie médicale, spécialement l'imagerie numérique, a récemment changé et d'une manière profonde, la pratique médicale. Cette évolution a donné lieu à une nouvelle dimension liée aux techniques de traitement d'images médicales et à l'information extraite par le moyen de ces techniques [1].

Dans ce chapitre, nous allons présenter l'approche médicale de l'imagerie du sein, pour situer le contexte dans lequel l'application que nous cherchons à développer va être utilisée. La présentation de l'anatomie du sein et des changements bénins et malins qui peuvent se produire, n'est pas exhaustive, et est donnée uniquement dans le but de faciliter la compréhension de la problématique, pour en déduire les méthodes de détection. La littérature de l'aspect médical du cancer du sein est très vaste [1-17].

Concevoir un système de détection automatique des lésions en mammographie, ne doit pas être compris comme une tentative de remplacement du médecin. La question n'est pas de se substituer au médecin, mais de lui proposer des outils performants qui l'aident dans son travail d'analyse des clichés de mammographie.

Nous commençons, dans ce chapitre, par la présentation de la mammographie, le dispositif d'acquisition et le principe physique de la mammographie. Une partie sera accordée à l'anatomie du sein, les principales maladies du sein, les pathologies bénignes. Enfin nous on présente les pathologies du sein en imagerie médicale.

1.1. La mammographie

La mammographie est un examen radiologique consacré à la détection des pathologies du sein, la plus performante en matière de surveillance et de dépistage précoce de cancer de sein. La qualité de l'image mammographique est caractérisée par trois paramètres physiques : la résolution, le contraste et le bruit [2]. Il existe deux types de mammographie : la mammographie conventionnelle (analogique, ou « classique ») qui produit une image imprimée sur un film argentique (un cliché). La mammographie numérique qui construit l'image à partir de la valeur numérique de chacun des points.

1.1.1. La mammographie analogique

Elle utilise le même principe qu'une radiographie standard. La quantité de rayons X produite dans le tube est de faible dose, elle dépendra de la quantité de tissu adipeux dans le sein (densité) et du nombre d'images prises (**Figure 1.1**). Le sein radiographie étant positionné

sur une porte cassette et comprimé par une plaque transparente. Cette compression est assurée par un système pneumatique avec commande au pied, ce qui permet de libérer les mains du manipulateur, pour donner un positionnement correct de sein [3].

Le sein interposé entre un tube émetteur à rayons X et un couple « écran photo luminescent-film photosensible ». Les rayons X traversent le sein où ils sont absorbés selon la densité des tissus rencontrés et frappent l'écran, ce qui entraîne la production d'une lumière qui impressionne le film argentique. L'image obtenue traduit la densité des tissus traversés, elle prend en continu, tous les niveaux de gris entre le blanc (absorption importante des tissus denses glandulaires, fibreux...) et noir (faible absorption des tissus gras). Une forte compression permet de [4] :

- Immobiliser le sein, ce qui réduit le flou cinétique.
- Réduire la dose en diminuant l'épaisseur traversée, pour améliorer le contraste et réduire la dose absorbée.
- Rendre l'épaisseur homogène sur la plus grande partie du sein, donc d'homogénéiser la densité.
- Etendre le tissu pour que les petites anomalies ne soient pas cachées par le tissu voisin [5].



Figure 1.1: Mammographie analogique [6].

1.1.2. La mammographie numérique

Il existe deux techniques pour obtenir une mammographie numérique. La première où la numérisation est indirecte, consiste à transformer l'image du film argentique en image numérique, par l'intermédiaire d'un scanner ou d'une caméra CCD, qui affectent une valeur numérique déterminée à chaque NG [7]. La seconde où la numérisation est directe, utilise des mammographies numériques « plein champ » où la plaque et son film photographique sont remplacés par des détecteurs qui captent le rayonnement X en mode numérique et reconstruit l'image à partir d'un nombre fini de point dont chacun a une valeur [8] .

La mammographie numérique est une technique plus récente en imagerie mammaire (**Figure 1.2**). Cette technique permet, en association avec des technologies modernes RIS-PACS et par sa possibilité de travail en réseaux, les avantages suivants :

- Diminution de la dose de radiation.
- Accélère considérablement l'acquisition des images, et élimine le temps consacré au développement de la pellicule.
- Manipulation des images sur des écrans d'ordinateur de haute résolution, en fonctions des besoins (Zoom, inversion, contraste, fenêtrage).



Figure 1.2 : Mammographie numérique [9].

1.2. Les incidences en mammographie

Selon la partie du sein à laquelle s'intéresse l'examen, différentes incidences sont utilisées. Les incidences les plus courantes sont l'incidence oblique externe (ou médio-latéral-oblique), l'incidence de face (ou cranio caudal) et l'incidence de profil [10].

1.2.1. Incidence cranio-caudale

Sur cette incidence (**Figure 1.3**), les quadrants externes se superposent entre eux d'une part et les quadrants internes entre eux d'autre part. Cette incidence explore bien la région centrale et la région interne du sein.

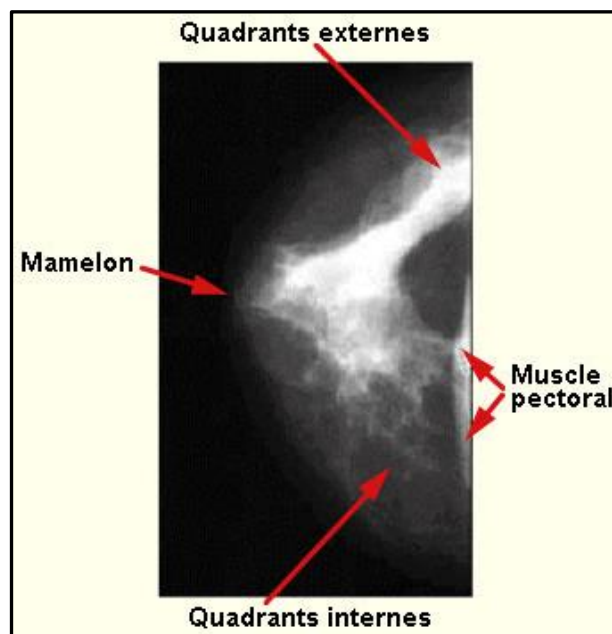


Figure 1.3 : Mammographie ; incidence de face [11].

1.2.2. Incidence de profil

Cette incidence (**Figure 1.4**) est généralement préférée au profil externe, car elle permet une analyse plus fine des lésions internes et une dissociation des structures externes.

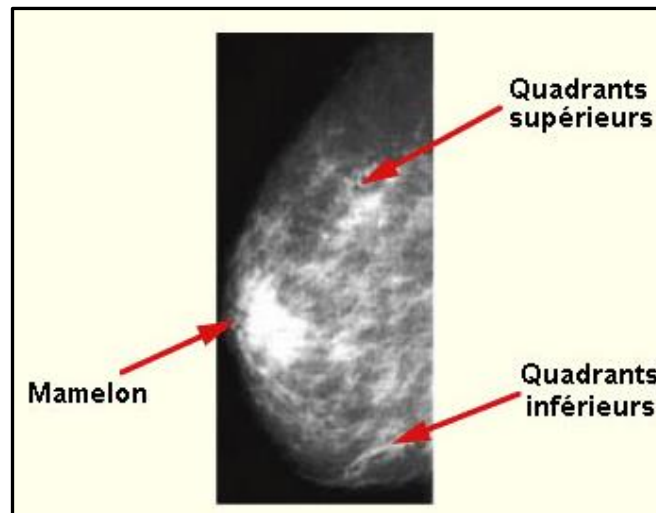


Figure 1.4 : Mammographie ; incidence de profil [12].

1.2.3. Incidences complémentaires

La figure 1.5 présente les trois vues possibles de la mammographie.

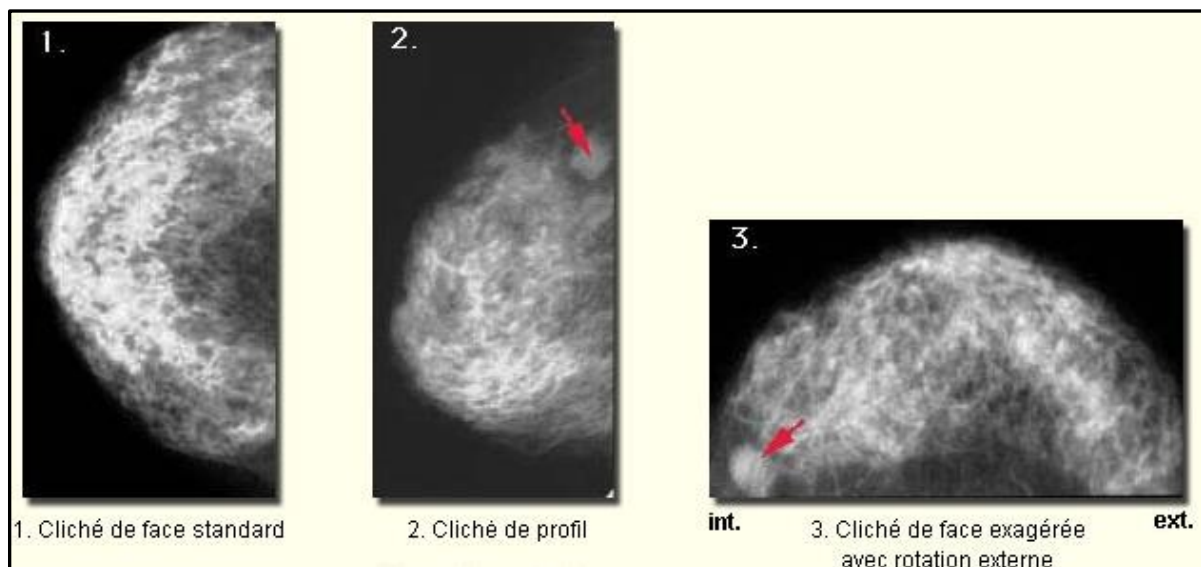


Figure 1.5 : Incidences complémentaires [13].

1.3. L'anatomie du sein

Le sein est un organe globuleux en nombre pair, situé en avant et en haut du thorax, entre la troisième et la septième côte [14], au-dessus du muscle grand pectoral, de structure complexe composée de graisse, de glandes et de canaux. Appelé aussi glande mammaire, sa fonction biologique est de produire du lait. Cette glande est constituée de 15 à 20 compartiments appelés

unités lobulaires ou lobes, reliés au mamelon par un réseau de canaux galactophores, servant au transport du lait (**Figure 1.6**). Les unités lobulaires sont séparées par des ligaments. Le sein contient aussi des vaisseaux sanguins, ainsi que des vaisseaux lymphatiques servant au transport d'un liquide clair appelé lymphes, et reliant des ganglions entre eux pour former des chaînes ganglionnaires, drainées en majorité dans les ganglions de l'aisselle.

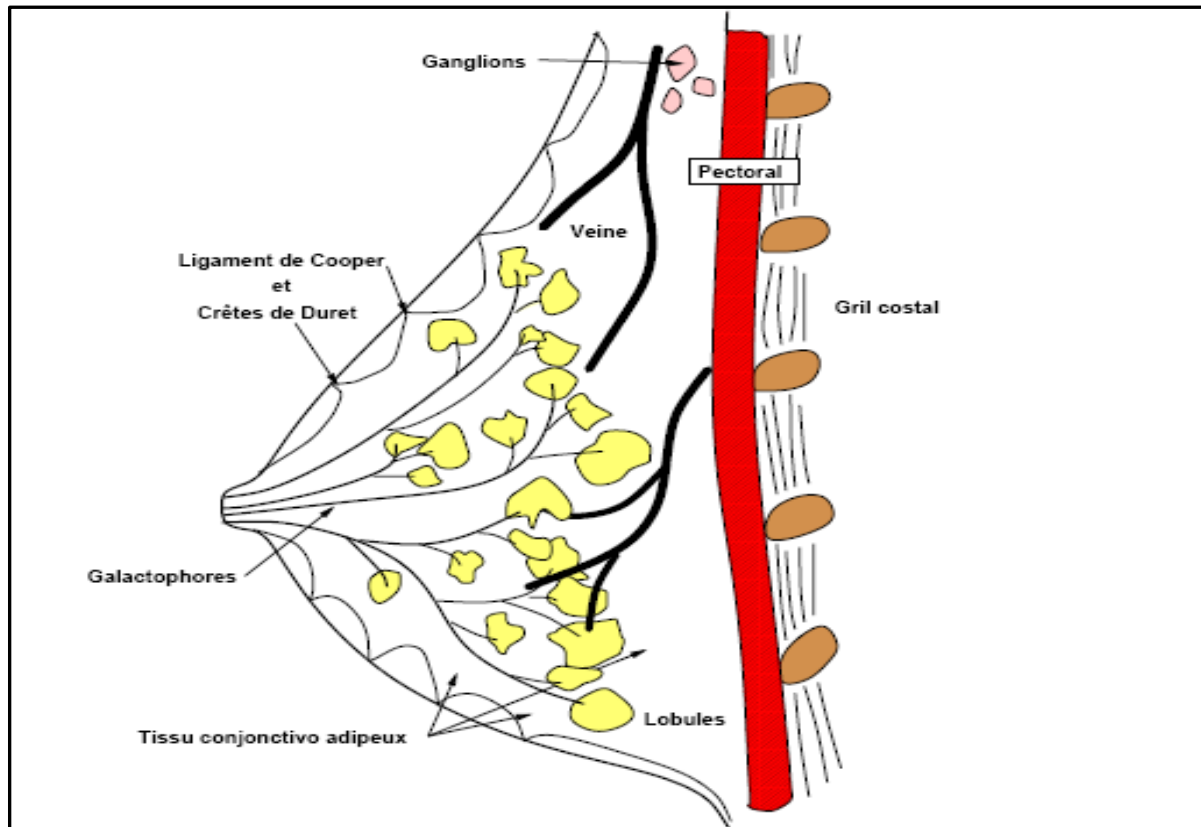


Figure 1.6: Anatomie du sein [15].

La **figure 1.7** est une mammographie d'un sein normal en vue de profil ; en le comparant avec le schéma de la **figure 1.6**, on peut reconnaître la structure en arches des crêtes de Duret, ainsi que le muscle pectoral en haut, à droite de la mammographie.



Figure 1.7: Mammographie d'un sein normal en vue de profil [16].

1.4. Que voit-on sur une mammographie et à quoi est dû le contraste observé ?

Dans le cas de la mammographie, le contraste est essentiel pour différencier entre deux types de tissus, ceux qui ont la densité de l'eau et ceux qui ont la densité de la graisse [17]. Les résultats sont présentés sur le tableau suivant (**Tableau 1-1**) :

Composant	Atténuation radiologique
Graisses	Radio transparent
Eau	Légèrement radio-opaque
Tissu conjonctif, fibre	Radio-opaque
Calcium	Très radio-opaque

Tableau 1-1:Contraste dans une image mammographique [18].

Une autre façon de faire le lien entre la réalité physique et son aspect radiologique, est l'approche expérimentale qu'utilise Lamarque, en comparant l'analyse histologique et les images obtenues par micro-radiographie, pour des exemples de tissus. Il donne le tableau de correspondance suivant (**Tableau 1-2**) [18] :

Structures histologiques	Traduction radiologique
Epithélium galactophorique	Peu perceptible
Tissu conjonctif	Opacité dense homogène
Fibrose	Opacité linéaire dense
Tissu graisseux	Image claire

Tableau 1-2: Approche de Lamarque [18].

1.5. Pathologies mammaires

Généralement, les pathologies mammaires sont des tumeurs qui se développent au niveau des lobules ou des canaux [18].

1.5.1. Tumeurs bénignes

Ce sont des masses tumorales présentant les caractères suivants : cellules régulières à croissance lente. Histologiquement identiques au tissu d'origine, elles refoulent le tissu voisin sans le détruire, ne donnent pas de métastases à distance, ne récidivent pas localement après exérèse, n'entraînent que rarement la mort du patient. Nous pouvons citer :

a) **Dystrophie fibrokystique : la maladie se manifeste par les lésions élémentaires bénignes :**

- Fibrose : la lésion la plus fréquente, atteint, à des degrés divers, le tissu conjonctif intra lobulaire.
- Adénose : c'est la multiplication des terminaisons canalaire avec leurs bases épithéliales.
- Hyperplasie épithéliale : multiplication du nombre de cellules tapissant la membrane basale.

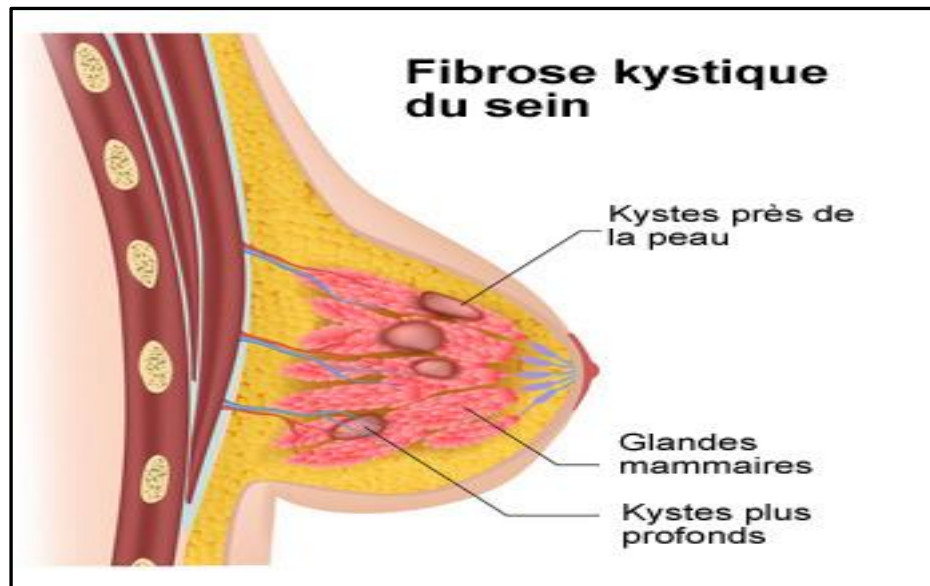


Figure 1.8: Dystrophie fibrokystique [19].

- b) **Adénofibrome** : tumeur bénigne du sein plus fréquemment observée chez la femme non ménopausée.

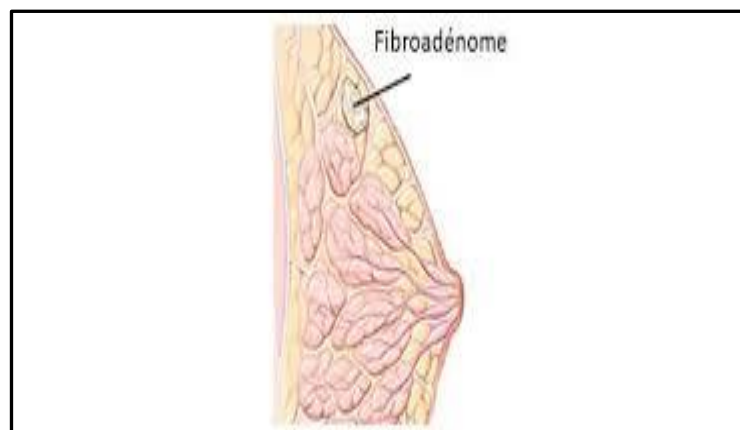


Figure 1.9 : Adénofibrome [20].

1.5.2. Tumeurs malignes (cancer du sein)

Le cancer du sein est le plus fréquent chez la femme, et représente l'une des premières causes de décès parmi les cancers gynécologiques.

Ce cancer est une prolifération anormale de cellules de l'organisme, au niveau de la glande mammaire. Ces cellules peuvent rester au niveau du sein, ou migrer dans l'organisme,

via les vaisseaux sanguins et lymphatiques. La majorité des cancers prennent naissance dans les canaux par lesquels coule le lait. Si la prolifération des cellules cancéreuses reste dans les canaux, on parle de « cancer in situ » (ou intra canalaire), en revanche, si les cellules sortent de la paroi des canaux, on parle de « cancer infiltrant ».

Cependant, malgré tous les progrès réalisés dans le domaine des recherches, les causes ne sont pas connues. Certains facteurs de risque ont été mis en évidence, favorisant l'apparition du cancer du sein, tels que l'âge, le sexe, la notion d'hérédité familiale, l'exposition prolongée aux rayons X, l'alcool et le tabagisme [21].

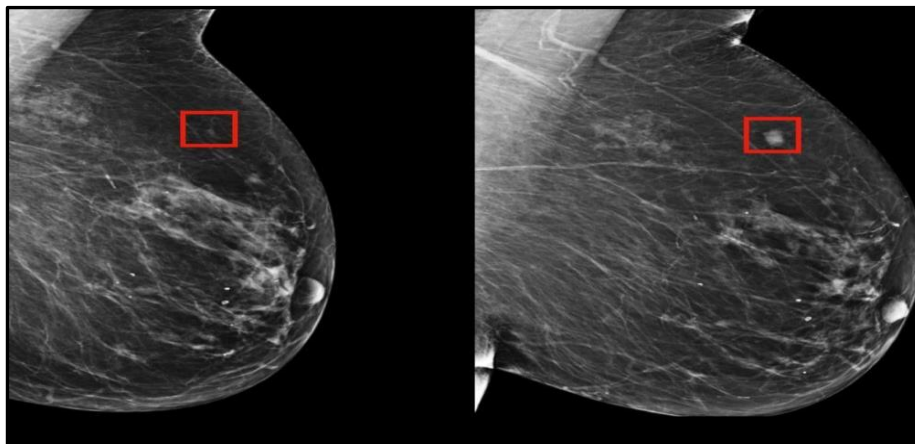


Figure 1.10 : Tumeur maligne [21].

1.6. Anomalies mammaires radiologiques

Dans la section précédente, nous avons essayé de caractériser les différents types de lésions. Dans [22], les auteurs orientent leurs analyses vers l'aspect radiologique, et décrivent les interprétations possibles. Ils y distinguent un certain groupe de signes radiologiques en relation avec le cancer.

Le dépistage du cancer du sein amène à découvrir quatre types d'anomalies purement radiologiques :

- Les opacités,
- Les microcalcifications,
- Les asymétries de densité,
- Les distorsions architecturales.

1.6.1. Opacités mammaires

Une opacité traduit une anomalie du tissu conjonctif ou épithélial. Elle correspond à une plage de surdensité anormale, qui ne se distingue pas d'une surdensité normale par un critère précis, mais par une combinaison de différentes caractéristiques : taille, densité, contour, forme, texture... Elle sera donc aisément visible au niveau d'une zone grasseuse, et beaucoup plus difficilement perceptible dans une zone dense de tissu conjonctif [23], c'est l'expérience qui permet au radiologue de distinguer une opacité sur une mammographie. Une surdensité sur plusieurs clichés effectués sous plusieurs incidences, implique une forte présomption en faveur de l'opacité.

1.6.2. Calcifications mammaires

Les calcifications sont des petits dépôts calciques, opaques aux rayons X, et visibles sur une mammographie. Leur origine n'est pas toujours connue, mais elles peuvent être liées à des sécrétions cellulaires. Leur morphologie est déterminée par l'endroit de leur formation [23]. L'analyse fine des calcifications permet de distinguer deux grandes catégories :

- a) **Macrocalcifications** : ce sont des dépôts grossiers de calcium dans le sein, avec un aspect très distinct à la mammographie. Elles sont toujours identifiées comme des affections bénignes.
- b) **Microcalcifications** : ce sont des dépôts calciques de petites tailles, comprises entre (200 μ m - 500 μ m), à la limite de visibilité. Elles peuvent être bénignes ou malignes, leur nature est précisée par des micro ou macro-biopsies. L'interprétation des microcalcifications (Mcs) dépend de plusieurs critères [23].
 - ✓ **la taille des microcalcifications** : les grosses Mcs sont le plus souvent bénignes et ont un plus fort contraste que les calcifications malignes ;
 - ✓ **la forme** : les Mcs rondes ou ovales ont plus de chances d'être bénignes que les Mcs irrégulières, en « pierres écrasées », ou linéaires ou branchées, qui ont plus de chance d'être malignes. Les calcifications trop petites pour être identifiées même sur des mammographies agrandies, dites pulvérulentes ont un risque intermédiaire.
 - ✓ **l'orientation**: les amas malins ont tendance à avoir une forme triangulaire ou losangique orientée vers le mamelon ;

- ✓ **le nombre** : Le nombre de Mcs est également un paramètre important à prendre en considération. Le cliché centré agrandi est important, pour préciser ce nombre. En effet certaines Mcs ne sont pas visibles en dehors du cliché centré agrandi. En pratique courante, il est difficile d'évoquer un foyer, s'il n'y a pas au moins 4 ou 5 Mcs groupées. Le foyer de Mcs est d'autant plus suspect que ces Mcs sont nombreuses et concentrées.
- ✓ **le polymorphisme** : si les Mcs dans un amas ont des formes variées, l'amas a plus de chance d'être malin.

Ces caractéristiques induisent aux nombreuses classifications des Mcs. La classification morphologique de Le Gal [23] a le mérite d'être claire et largement diffusée. Celle-ci inclus 5 types de Mcs (**Tableau 1-3**):

Type 1 :	Microcalcifications annulaires, arciformes ou polyédriques. Risque de cancer du sein quasi nul.
Type 2:	Microcalcifications rondes et de tailles variables. Risque de carcinome évalué à 22 %.
Type 3 :	Microcalcifications poussiéreuses pulvérulentes. Risque de cancer estimé à 36 %.
Type 4 :	Microcalcifications irrégulières associées à un risque de cancer estimé à 56 %.
Type 5 :	Microcalcifications vermiculaires ou branchées. Risque de carcinome estimé à 91 %.

Tableau 1-3: Classification de le Gal [23].

1.6.3. Classification de Le Gal des microcalcifications mammaires

Cette classification cherche à évaluer le caractère dangereux ou non des calcifications observées. On distingue cinq types de Mcs [23] :


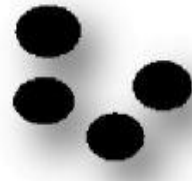



Type I	Type II	Type III	Type IV	Type V
Images rondes planes	Images régulières	Grains de sel	Points irréguliers	Vermiculaires
				
0% de cancers	39% de cancers	39% de cancers	59% de cancers	96% de cancers

Figure 1.11: Classification de Le Gal des microcalcifications mammaires [23].

Remarque :

Étant donné que notre travail se base sur la détection des Mcs, on va s'intéresser uniquement au signe radiologique des calcifications.

La figure qui suit (**Figure 1.12**) présente une mammographie d'un sein pathologique, en vue de profil, qui contient un foyer de Mcs, et on trouve dans la **figure 1.13** un agrandissement de la région contenant l'amas de Mcs.

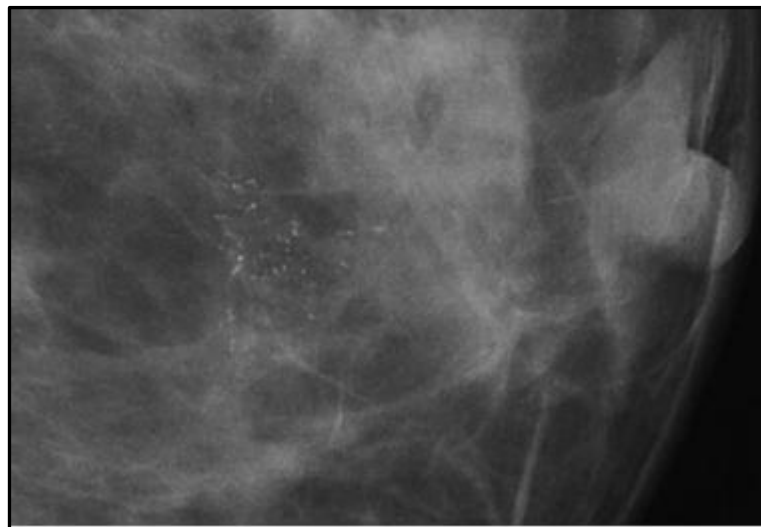


Figure 1.12 : Mammographie montrant un foyer de microcalcifications [22].

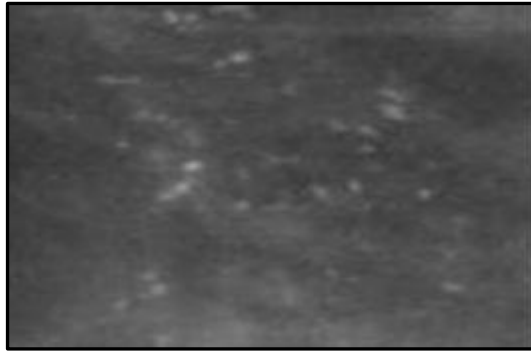


Figure 1.13 : Agrandissement de la région contenant l'amas de microcalcifications [22].

1.7. Discussion

Ce chapitre nous a permis de mettre en évidence l'intérêt de la mammographie, et les renseignements précieux qu'elle nous donne pour le diagnostic du cancer du sein. En raison de l'importance du cancer du sein, un grand nombre de groupes de recherche travaillent sur le problème de l'aide au diagnostic. Les premiers travaux datent de 1967. Les travaux se distinguent par les types des signes qui sont détectés (opacités, calcifications), le but recherché soit la détection et la classification. Dans le chapitre suivant, nous nous intéresserons au problème de la détection des Mcs. Nous citerons les différentes méthodes utilisées et les outils nécessaires.

Chapitre 2 :

Etat de l'art

Le traitement d'image peut être défini comme l'ensemble des techniques qui corrige les défauts de l'image et permet un confort de visualisation par l'augmentation du contraste, la correction des distorsions optiques et le filtrage de bruit. L'un des objectifs de base, désiré en introduisant ce domaine dans les différents secteurs de la vie, est d'aider le praticien dans sa prise de décision.

Les approches pour la détection de microcalcifications proposées dans la littérature sont très variées, et on retrouve des adaptations de quasiment toutes les méthodes du traitement d'images. Pour cela, dans ce chapitre, nous présentons un résumé des travaux de recherche, qui traitent de la classification des microcalcifications mammaires. Ensuite nous donnerons un récapitulatif, des méthodes utilisées dans ces travaux, et nous choisirons les méthodes sur lesquels notre travail sera construit.

2.1. Détection des microcalcifications : état de l'art

La littérature spécialisée dans le traitement automatique des images mammographiques est riche d'innombrables approches. Cependant, les méthodes qui traitent la segmentation et la détection des micro-classifications sont encore à un stade exploratoire, et tous les travaux publiés à ce jour ne sont pas infaillibles et souffrent pour certains d'entre eux, soit d'un manque de robustesse ou de précision. Nous présentons dans ce qui suit, un état de l'art sur les différentes approches en matière de segmentation et de détection des micro-classifications.

✚ **Boucher et al. (2009)** [24] utilisent une technique de segmentation du muscle pectoral basé sur la méthode des contours actifs pour délimiter le muscle pectoral du sein. Les résultats montrent l'efficacité de l'approche proposée avec une précision de 95%.

✚ **Balakumaran et Vennila (2011)** [25] réalisent un système de diagnostic assisté par ordinateur (CAD), pour la détection des microcalcifications dans les mammographies numérisées. Balakumaran and Vennila utilisent la décomposition en ondelettes sans opérateur d'échantillonnage, en différentes sous-bandes, en supprimant la sous-bande d'approximation la plus grossière, et la reconstruction à partir des sous-bandes ne contenant que des informations détaillées significatives. Les résultats obtenus sont meilleurs par rapport à la transformée en ondelettes 2D.

✚ **Kabbadj et al. (2012)** [26] présentent une nouvelle approche pour détecter les microcalcifications (MC) sur les mammographies numérisées, en utilisant la logique floue

et le SVM avec des caractéristiques géométriques et statistiques. La méthode a été testée sur 16 mammographies de la base de données MIAS et a obtenu une sensibilité égale à 99.66% et une spécificité égale à 99,11%.

✚ **Lewis et Aijuan Dong. (2012)** [27] développent la méthode (marker-controlled watershed), pour localiser la masse tumorale du sein. En premier, une sélection des marqueurs de premier plan et d'arrière-plan (foreground and background markers) est utilisée, ensuite une application de l'algorithme de SMCW (Segmentation Marker-Controlled Watershed) pour séparer la tumeur de son tissu environnant. L'approche proposée réussit assez bien à localiser les tumeurs. Les résultats obtenus ont été appliqués à l'ensemble de données MIAS, avec le taux de détection global de la masse tumorale égale à 90%.

✚ **Zhang et al. (2012)** [28] réalisent une série de tests, pour évaluer quantitativement l'impact de l'amélioration d'image, sur leur l'approche de segmentation. La méthode d'amélioration d'image combinatoire est basée sur la forme topographique d'une ROI (Regions Of Interest), pour réduire l'influence de l'arrière-plan dans la segmentation. D'après les résultats expérimentaux une AOM (Area Overlap Metric) égale à 0.70, l'étude démontre qu'en raison de l'utilisation de l'amélioration de l'image, la méthode DPA (Dynamic Programming Algorithm) présente une amélioration évidente dans la segmentation des régions d'intérêt suspectes (ROI).

✚ **Suapang et al. (2012)** [29] réalisent un nouvel algorithme mathématique de segmentation morphologique, en introduisant la théorie et les opérations morphologiques, afin de détecter les masses mammographiques. L'algorithme présente un meilleur compromis entre le lissage du bruit et l'orientation des bords, mais le calcul est plus complexe que les algorithmes de segmentation morphologique généraux. Les résultats expérimentaux montrent que l'algorithme est plus efficace pour le débruitage des images médicales et la détection des contours, que les algorithmes de segmentation basés sur des modèles habituellement utilisés, tels que l'opérateur Laplacien de gaussienne et le détecteur de contours de Sobel.

✚ **Malar et al. (2012)** [30] révèlent l'efficacité de l'analyse de la texture, sur les images mammographiques, pour détecter l'emplacement des microcalcifications. Les paramètres d'ondelettes, de Gabor et de la GLSDM (Gray Level Spatial Dependence Matrix) sont

calculés, et classés par la méthode Externe Learning Machine (ELM). Malar et al ont utilisé la base de données mini-MIAS et ont obtenu une précision égale à 94%.

✚ **Fathima et al. (2013)** [31] développent une approche de détection des tumeurs dans les mammographies, une détection de la région d'intérêt, est effectuée par rehaussement de contraste et les outils morphologiques. Par la suite une extraction des paramètres par la méthode de Gray Level Co-occurrence Matrix (GLCM), et Discret Wavelet Transform (DWT) sont appliqués, pour une classification par la méthode Support Vector Machine (SVM). La précision a été obtenue jusqu'à 95%.

✚ **Wajid et al. (2014)** [32] proposent le classificateur ESN (Echo State Network), pour détecter les cas de cancer malin. Les auteurs combinent l'ESN avec un LESH (Local Energy based Shape Histogram). Le système suggéré atteint une précision de classification élevée de 98%, ainsi qu'une sensibilité et une spécificité élevée. Les performances du classificateur d'ESN ont été testées par rapport à d'autres classificateurs, comme le SVM et affiche de meilleurs résultats.

✚ **Zouari et al. (2014)** [33] proposent une nouvelle approche NLS (Non Linear Stretching), pour améliorer le contraste des microcalcifications présents dans la mammographie numérisée, par une fonction exponentielle adaptative, qui effectue une modification locale du niveau de gris, pour mettre en évidence les microcalcifications. Zouari et al appliquent le NLS deux fois : localement et globalement. Les résultats de performance sont donnés en termes de mesure d'amélioration de type dérivée seconde SDME (Seconde-Derivative like Measure of enhancement).

✚ Dans les travaux de **Mustra et Grgic (2014)**. [34] , on y trouve une méthode de détection de la présence de microcalcifications, par l'amélioration du contraste des microcalcifications par rapport au tissu de fond, avec l'utilisation, de filtrage par ondelettes et de la morphologie mathématique. Les résultats obtenus de la détection automatique des zones suspectes des microcalcifications se sont avérés bons par rapport à d'autres méthodes qui n'utilisent pas la sélection manuelle des zones suspectes. Ils ont obtenu une amélioration significative de l'amélioration du contraste de microcalcifications.

✚ **Intharasombat et Piamsa-nga (2014)** [35] proposent une méthode AWAT (An Adaptive Window and Adaptive Threshold) de fenêtre adaptative et de seuil adaptatif, pour la détection de la microcalcification. Les maxima locaux sont identifiés à l'aide d'un seuil optimal adapté à chaque fenêtre, ou le seuil est égal à 2 fois l'écart-type (2SD). Ensuite, les petits objets sont éliminés à l'aide d'opérations morphologiques. Les maxima locaux restants sont classés en microcalcification ou tissu normal à l'aide de trois classificateurs : un perceptron multicouche (Multi Layer Perceptron), un réseau de neurones à fonction de base radiale (Radial Basis Function Neural Network) et un SVM. Les résultats expérimentaux démontrent que la méthode AWAT proposée fonctionnait mieux que les méthodes avec la taille de fenêtre fixe.

✚ **Chacon et al. (2015)** [36] proposent une méthode de diagnostic du cancer du sein, basée sur les conditions des tissus mammaires, en utilisant une analyse d'asymétrie thermographique du sein, par une comparaison des caractéristiques de la zone du sein droit et du sein gauche. Le modèle proposé se base sur un classificateur flou, utilisant des caractéristiques statistiques. Les résultats du système CAD ont été comparés à l'état pathologique des tissus des patients, et ont montré des résultats satisfaisants permettant de distinguer les tissus cancéreux des tissus sains, avec une sensibilité de 82,35% et une spécificité de 92,15%.

✚ **Esener et al. (2015)** [37] présentent un système automatique de diagnostic assisté par ordinateur (CAD), pour le diagnostic du cancer du sein. Dans la réalisation du système CAD, tous les patchs de mammographie sont d'abord prétraités par une égalisation d'histogramme, suivie d'un filtrage par moyennes non locales (NLM). Ensuite, une extraction de caractéristiques par l'algorithme « Local Configuration Pattern (LCP) » concaténé avec certaines caractéristiques statistiques fréquentielles est faite, puis une classification avec l'analyse discriminante linéaire de Fisher (FLDA) est réalisée. Une précision de reconnaissance maximale de 94,67% est atteinte en utilisant la nouvelle approche.

✚ **Hai Su et al. (2015)** [38] développent une méthode basée sur le SVM (Machine à Vecteur de Support) à l'aide de l'analyse conjointe des caractéristiques morphologiques et des informations génomiques fonctionnelles. Pour chaque patient, la région du sein a été détectée et segmentée pour identifier les régions tumorales. Le système proposé a été testé sur un ensemble de données contenant 86 patients, avec la validation croisée, qui nous montre que les caractéristiques morphologiques ainsi que les informations génomiques

fonctionnelles, produisent une meilleure performance de classification, pour différencier les patients de stade I, d'autres stades. La combinaison de l'image-omique et de la caractéristique génomique fonctionnelle peut améliorer la précision de la classification de 3%.

✚ **Tan et al. (2015)** [39] développent une nouvelle approche de détection assistée par ordinateur (CAD), basé sur l'analyse globale des caractéristiques des images, en utilisant la densité tissulaire mammographique. Les caractéristiques ont été extraites sur des images bilatérales, craniocaudales et médio latérale oblique, pour former deux réseaux de neurones artificiels ANNs. Les résultats ont ensuite été fusionnés par un troisième ANN, pour construire un classificateur de prédiction de l'examen de dépistage. La sensibilité obtenue est de 95% et la spécificité est de 80%.

✚ **Chen et al. (2015)** [40] présentent une approche de classification des grades du cancer de sein, par l'utilisation d'un classificateur SVM et d'un algorithme génétique (GA), pour l'optimisation des paramètres du modèle. Chen et al. Utilisent dans leurs approches les caractéristiques texturales, morphologiques, ellipsoïdales et acoustiques postérieures. Le système proposé classifie efficacement les tumeurs de haut grade et de bas grade avec une précision de 85,14%, une sensibilité de 79,31% et une spécificité de 86,55%.

✚ **Khalaf et Yassine (2015)** [41] proposent un système de diagnostic assisté par ordinateur (CAD) basé sur un ensemble de paramètres pour la détection des microcalcifications. Leur approche utilise une étape de prétraitement par DWT, une extraction des paramètres par des statistiques d'ordre supérieur (HOS (Higher Order Statistical)), par la transformée DWT (Discrete Wavelet Transform) et par la WPD (Wavelet Packet Decomposition) et leurs classifications par SVM. Les résultats obtenus sont validés à l'aide de la base de donnée MIAS et DDSM. Le système proposé atteint une sensibilité de 98,43%, une spécificité de 94,87% et une précision de 95,80%

✚ **Sainz de Cea et Yang (2015)** [42] utilisent dans leur approche (Bayes' risk) le risque de Bayes, pour la règle de décision dans la sortie du détecteur, pour laquelle le modèle de probabilité des faux positifs (FP) est déterminé directement à partir de l'image considérée. Les résultats de leurs méthodes pourrait réduire le taux de PF jusqu'à 44,16% et le taux de vrais positifs à 85%.

✚ **Zhang et al. (2016)** [43] développent une architecture d'apprentissage profond (DL), basée sur la méthode PGBM (Point-wise Gated Boltzmann Machine) pour la classification des tumeurs bénignes et des tumeurs malignes du sein. Li et al. Appliquent l'architecture DL à deux couches, et les résultats expérimentaux démontrent que les résultats obtenus surpassent les SF (Statical Features) en matière de précision de classification, sensibilité, spécificité et YI (youden's index). L'architecture DL pourrait être potentiellement utilisée dans CAD (computer aided diagnosis). Les résultats ont montré que les caractéristiques DL obtenaient de meilleures performances de classification avec une précision de 93,4%, une sensibilité de 88,6% et une spécificité de 97,1%.

✚ **Cascio et al. (2018)** [44] présentent une nouvelle méthode pour la détection des MC dans les mammographies, qui consiste en une segmentation des régions d'intérêt (ROI), basée sur un filtre spatial qui permet la détection de petites et grandes microcalcifications, le regroupement et la classification des MC se fait par des neurones artificiels. Les résultats démontrent que l'approche proposée atteint une sensibilité de 90%.

✚ **Karale et al. (2019)** [45] présentent un nouvel algorithme de détection des amas de microcalcifications dans une mammographie. Un opérateur énergétique non-linéaire 2D multi-échelle (Multiscale 2D No-Linear Energy Operator) est proposé pour renforcer le contraste entre les microcalcifications et le fond. Plusieurs types de paramètres sont extraits, de texture, de forme, d'intensité et les caractéristiques basées sur l'histogramme des gradients orientés (HOG) sont utilisés. Une nouvelle technique est proposée pour équilibrer les données non équilibrées.

✚ **Basile et al. (2019)** [46] proposent une approche en trois étapes, pour la détection de microcalcifications, une étape de pré-traitement par rehaussement de contraste et filtrage de Sobel, suivie d'une étape de détection de microcalcifications, basée sur la transformée de Hough, capable de saisir la forme caractéristique des structures d'intérêt. Enfin, une étape d'identification des clusters pour regrouper les microcalcifications, avec un algorithme de clustering capable de codifier des règles de domaine expert.

✚ **Azam et al. (2020)** [47] développent une méthode hybride de Canny Edge Detection et Otsu Thresholding et 2D Wavelet Transform, en exploitant la force de chacune des méthodes. La méthode proposée a surmonté le problème de l'algorithme de Canny en déterminant automatiquement les seuils haut et bas. La méthode hybride a été testée avec 30

images de mammographie. Ils obtiennent une précision égale à 92,80%, un F-score égale à 0.9280 et un taux d'erreur égale à 0.1375.

✚ **Touil et al. (2020)** [48] présentent une nouvelle méthode automatique, de détection des microcalcifications présentes dans la mammographie, basée sur la morphologie mathématique, cela à l'aide de divers types, et taille d'éléments structurants. Les résultats morphologiques obtenus, sont convertis en une carte, appelée indice de similarité structurelle (SSIM). Cette carte est utilisée, pour une analyse locale, à l'aide de super-pixels pour estimer automatiquement des valeurs seuils et enfin détecter d'éventuelles zones de microcalcifications. La méthode proposée a été évaluée à l'aide de l'ensemble de données INBreast accessible au public.

✚ **Mazumder et al. (2020)** [49] proposent une nouvelle technique basée sur la transformation stationnaire en ondelettes (SWT) (Stationary Wavelet Transform), pour la détection et la classification des microcalcifications mammaires. La SWE (Stationary Wavelet Energy) en question a été appliquée à plusieurs niveaux à des fins de décomposition, et l'énergie stationnaire des ondelettes (SWE) a ensuite été mise en œuvre, pour l'extraction des caractéristiques de chaque sous-bande de coefficient SWT. Ensuite, les paramètres sont classés comme bénins ou malins avec le classifieur SDE (Subspace Discriminant Ensemble).

✚ **Bulbul et al. (2020)** [50] réalisent une conception et une analyse numérique d'un biocapteur à base de «Photonic Crystal Fiber» (PCF), pour détecter les cellules cancéreuses dans le sein.

✚ **Desai et al. (2020)** [51] utilisent une méthode basée sur «Mul-TI-layer Perceptron Neural Network» (MLP) et le CNN «Convolutional Neural Net-work », pour la détection précoce de cancer de sein et la classification des cellules cancéreuses du sein.

✚ **Murakami et al. (2020)** [52] comparent entre la mammographie synthétique reconstruite à partir de tomosynthèse mammaire numérique (2DSM) et mammographie numérique plein champ (FFDM). Les images (2DSM) et (FFDM) sont analysées avec un « McNemar test ». L'évaluation BI-RADS entre 2DSM et FFDM a été évaluée à l'aide de «

kappa de Cohen ». Les taux moyens de détection du cancer avec les images 2DSM et FFDM sont de 84,6% et 87,8% respectivement.

✚ **Bai et al. (2021)** [53] appliquent le Deep Learning à la tomosynthèse mammaire numérique pour la détection du cancer du sein. Ils ont utilisé la technologie Digital Breast Tomosynthesis (DBT).

✚ **Gupta et al. (2021)** [54] détectent le cancer du sein à partir d'images histopathologiques, en utilisant des réseaux de neurones résiduels modifiés. Ils ont proposé une méthode basée sur un réseau neuronal résiduel pour la détection du cancer bénin ou malin. L'approche proposée offre de bonnes performances, le réseau obtient une précision de classification de 99,75% et une précision de 99,17% sur l'ensemble de données BreakHis.

✚ **Cao et al. (2021)** [55] présentent une méthode pour la détection de la masse mammaire en mammographie numérique basée sur «Anchor-Free Architecture» à l'aide de l'architecture de détection d'objets en une étape, appelée «Breast Mass Detection Network (BMassDNet)», introduisant une méthode de normalisation de troncature combinée avec l'égalisation adaptative de l'histogramme, pour améliorer le contraste entre la masse mammaire et le milieu environnant.

✚ **An et al. (2021)** [56] détectent et segmentent des masses mammaires, basées sur la fusion de caractéristiques multicouches, pour améliorer la précision de détection des masses mammaires. Ils ont proposé un modèle de détection D-Mask R-CNN basé sur Mask R-CNN et ils ont utilisé l'ensemble de donnée CBIS-DDSM. Les résultats ont montré que la valeur maP (Mean Average Precision) du modèle amélioré pour détecter les masses mammaires atteint 0.66 dans l'ensemble de test.

✚ **Sánchez-Cauce et al.(2021)** [57] présentent une méthode sur un réseau de neurones conventionnels multi-entrées pour la détection du cancer du sein, en utilisant comme entrée au CNN, des images thermiques et des données cliniques, puis construisant un modèle de classification multi-entrées qui exploite les avantages des réseaux de neurones convolutifs pour l'analyse d'images. Le modèle CNN atteint une précision de 97%, une spécificité de 100% et une sensibilité de 83%.

✚ Zhang et al. (2022) [58] présentent une méthode d'apprentissage en profondeur en utilisant l'algorithme R-CNN (Regional-Convolutional Neural Network), pour détecter les lésions anormales présentes dans le sein. L'algorithme a atteint la précision de 0,86.

2.2. Tableau récapitulatif

Le tableau 2.1 contient un résumé de (tous les articles étudiés), pour la réalisation d'une méthode de détection et de segmentation des microcalcifications.

Références	Année	Approches	Application	Validation	Base de données
Boucher et al. [24]	2009	Contours actifs.	Délimiter le muscle pectoral du sein.	/	/
Balakumaran et Vennila. [25]	2011	Ondelettes	Détection des microcalcifications dans les mammographies numérisées.	/	DDSM
Kabbadj et al. [26]	2012	SVM (Support Vector Machines) et logique floue	Détection des microcalcifications sur les mammographies numérisées.	Sensibilité = 99,66% spécificité = 99,11%.	MIAS
Lewis et Aijuan Dong. [27]	2012	Marked-controlled watershed	Localisation d'une masse tumorale du sein.	Détection = 90%	MIAS
Zhang et al. [28]	2012	DPA (Dynamic Programming Algorithm)	Segmentation des régions d'intérêt suspectes.	AOM (Area Overlap Metric) = 0.70	DDSM
Suapang et al. [29]	2012	Opérations morphologiques	Détection des masses mammographiques.	/	/
Malar et al. [30]	2012	ELM (Extreme Learning Machine)	Détection de l'emplacement des microcalcifications.	Précision = 94%	Mini-MIAS
Fathima et al. [31]	2013	SVM (Support Vector Machines)	Détection des tumeurs dans la mammographie.	Précision= 93,3% Sensibilité=90,92%. spécificité=97%.	MIAS
Wajid et al. [32]	2014	ESN (Echo State Network)	Détection du cancer malin.	Précision = 98%	MIAS
Zouari et al. [33]	2014	NLS (Non Linear Stretching)	Amélioration de contraste des microcalcifications présents dans la	SDME (Second-Derivative like Measure of	FDDSM

			mammographie numérisée	enhancement) = 118 SDME = 115 pour (NLS global)	
Musta et Grgic. [34]	2014	Amélioration de contraste	Détection de la présence des microcalcifications.	/	/
Intharasombat et Piamsa-nga. [35]	2014	AWAT (An Adaptive Window and Adaptive Threshold)	Détection des microcalcifications.	la méthode AWAT proposée fonctionnait mieux que les méthodes de taille de fenêtre fixe.	MIAS Nijmegen DDSM
Chacon et al. [35]	2015	Classificateur flou.	Détection du cancer du sein.	Sensibilité=82.35% Spécificité=92.15%	Lab at the Federal Fluminense University
Esener et al. [37]	2015	FLDA (Fisher's Linear Discriminant Analysis)	Détection du cancer du sein.	Précision=94.67%	IRMA avec 600 patches.
Hai Su et al. [38]	2015	SVM	Détection et segmentation des régions tumorales.	Les performances de classification obtenues par les caractéristiques combinée sont supérieures aux performances basées sur les caractéristiques d'image uniquement.	TCGA avec 68 patients.
Tan et al. [39]	2015	ANN (Artificial Neural Networks)	Détection automatique	Sensibilité =95% Spécificité =80 %.	IRB avec 1896 images
Chen et al. [40]	2015	SVM	Classification des grades de cancer du sein.	Précision= 85,14 % Sensibilité=79,31 % Spécificité=86,55 %	Changhua Christian Hôpital avec 148 images 3D
Khalaf et Yassine. [41]	2015	SVM	Détection des microcalcifications.	Sensibilité=98,43 % Spécificité=94,87 % Précision=95,80 %	DDSM / MIAS
Sainz de Cea et Yang. [42]	2015	Bayes' risk	Détection des microcalcifications.	Faux positif = 44,16 %. Vrai positif = 85%.	408 mammographies.
Zhang et al. [43]	2016	DL (Deep Learning) / PGBM	Classification des tumeurs bénignes et malignes de sein.	Précision = 93,4% Sensibilité = 88,6% Spécificité = 97,1%	IRB

		(Point-wise Gated Boltzmann Machine)			
Casio et al. [44]	2018	Réseaux de neurone.	Détection et classification des microcalcifications.	Sensibilité=90%	INbreast DB avec 410 images.
Karale et al. [45]	2019	A multiscale 2D non-linear energy operator.	Détection des microcalcifications dans une mammographie.	Sensibilité=100 %	DDSM /INbreast PGIMER-IITKGP mammogram
Basile et al. [46]	2019	Sobel, transformée de Hough et l'algorithme de clustering	Détection de microcalcifications.	Vrai positif=91,78%.	BCDR avec 364 images.
Azam et al. [47]	2019	Canny edge détection	Détection et segmentation des microcalcifications.	Précision =97.50% F-score =0.9280 taux d'erreur =0.1375	NCSM avec 30 images.
Touil et al. [48]	2020	Morphologie mathématique	Détection des microcalcifications présente dans la mammographie.	/	INBreast.
Mazumder et al. [49]	2020	SWT (Stationary Wavelet Energy)	Détection et classification des microcalcifications mammaires.	Sensibilité =94.12% Spécificité =92.48% Précision= 88.89%	MIAS 322 images mammographiques
Bulbul et al. [50]	2020	PCF (Photonic Crystal Fiber)	Détection des cellules cancéreuses dans le sein.	Sensibilité =92.2%	/
Desai et al. [51]	2020	CNN (Convolutional Neural network) /MLP (Multi-Layer Perceptron Neural network)	Détection précoce du cancer du sein.	/	/
Murakami et al. [52]	2020	McNemar test	Comparaison entre la mammographie synthétique reconstruite à partir de tomosynthèse mammaire numérique (2DSM) et mammographie numérique plein champ (FFDM).	Le taux moyen de détection du cancer de sein = 87.8% pour FFDM Le taux moyen de détection du cancer de sein = 84.2% pour 2DSM	Les données d'imagerie de 136 patients.

Bai et al. [53]	2020	DL (Deep Learning)	Détection de cancer du sein.	/	DBT
Gupta et al. [54]	2021	Réseau de neurone résiduel.	Détection de cancer du sein.	Précision de classification moyenne= 99,75 %. Précision de 99,18 %	BreakHis
Cao et al. [55]	2021	BMassDNet (Breast Mass Detection Network)	Détection de la masse mammaire en mammographie numérique.	INbbreast : FPPI (le faux positif par image) = 0.495 Taux de rappel = 0.930 DDSM : FFPI = 0.599 Taux de rappel = 0.943	INbbreast DDSM
An et al. [56]	2021	R-CNN (Regions with Convolutional Neural Networks)	Détection et segmentation des masses mammaires.	maP (Mean Average Precision) = 66%	CBIS-DDSM
Sánchez-Cauce et al. [57]	2021	CNN	Détection du cancer du sein.	Précision =97% Spécificité =100% Sensibilité =83%	DMR
Zhang et al. [58]	2022	R-CNN (Regions with Convolutional Neural Networks)	Détection des lésions anormales dans le sein.	Précision= 86%	Hôpital (Chine) 241 patients autre hôpital avec 98 patients

Tableau 2-1:Tableau récapitulatif de toutes les revues étudiées

2.3. Discussion

Notre travail consiste à détecter et à segmenter le foyer des microcalcifications pour faciliter leurs quantifications, car le nombre de microcalcifications présentes dans un groupement (foyer) est un indicateur précoce d'un cancer du sein. Des études sur la détection des microcalcifications montrent clairement qu'il n'existe pas de techniques universelles pour choisir une méthode de segmentation, mais que chaque approche donne un résultat plus ou moins bon selon le type d'application. Nous avons écarté un certain nombre de méthodes qui ne répondent pas au critère de facilité d'utilisation sur des applications mammographiques, et nous avons choisi une approche basée sur la fusion de plusieurs méthodes, comme le k-means et la morphologie mathématique, dont les notions de base et les transformations fondamentales seront présentées dans le chapitre suivant.

Chapitre 3 :
Outils Mathématiques Utilisés

Les mammographies sont des images qui n'offrent pas un bon contraste entre les tissus mammaires sain et malin ; ceci s'explique par le fait que l'atténuation des rayons X entre ces tissus varie peu, surtout chez les jeunes femmes qui ont un parenchyme mammaire très dense. L'analyse d'image se résume comme suit :

- Le prétraitement consiste en diverses opérations visant à améliorer la qualité de l'image et à faciliter la segmentation. Ces opérations sont principalement le rehaussement du contraste, la modification des histogrammes et la réduction du bruit.

- L'analyse a pour but d'extraire les caractéristiques permettant de classifier les pixels de l'image.

- La segmentation est un traitement de bas niveau qui consiste à créer une partition de l'image en sous-ensembles appelés régions. La segmentation est alors obtenue par extraction des composantes connexes des pixels appartenant à la même classe. Ainsi une région sera constituée de pixels d'une même classe.

Dans ce chapitre, nous allons passer en revue la méthode que nous avons utilisée pendant ce mémoire pour la segmentation des microcalcifications (Mcs). Pour cela, nous allons commencer par présenter l'approche globale et ses différents composants, ainsi que les outils mathématiques qui la décrivent. Nous allons commencer par présenter quelques notions sur les outils mathématiques utilisés dans notre approche, puis nous donnerons l'approche globale que nous avons adoptée. Nous allons détailler la chaîne de traitement qui consiste à détecter d'abord la zone des Mcs (Region Of Interest (ROI)), par la suite la région détectée sera segmentée par la fusion de deux approches différentes : le chapeau haut de forme seuillé (CHF) et la classification non supervisée (ou clustering).

3.1. Apprentissage profond

L'intelligence artificielle (IA) (**Figure 3.1**) va avoir un impact sur notre société, que l'on ne peut encore imaginer. Des algorithmes prétendent déjà pouvoir identifier les meilleurs candidats pour un emploi, assister les médecins, ou aider les avocats devant les tribunaux. Tout cela n'est pas totalement nouveau, puisque déjà, dans les années 1980, des systèmes experts assistaient les humains, avec un haut niveau d'expertise. Ce qui est nouveau aujourd'hui, c'est que les ordinateurs parviennent, de manière autonome, à exécuter des tâches, extrêmement

complexes, mais sans que leur concepteur ne comprenne parfois, ce qui s'est passé dans la « boîte noire » de l'apprentissage profond (deep-Learning) [59].

Dans l'apprentissage en profondeur, un modèle informatique apprend à effectuer des tâches de classification directement à partir d'images, de textes ou de sons. Les modèles d'apprentissage en profondeur peuvent atteindre une précision de pointe, dépassant parfois les performances d'un niveau humain. Les modèles sont entraînés en utilisant une grande masse de données étiquetées, et des architectures de réseau neuronal qui contiennent de nombreuses couches [60].

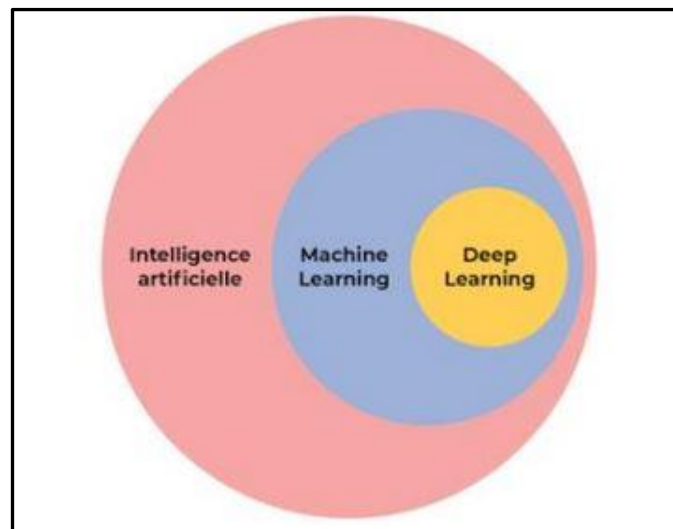


Figure 3.1: Deep Learning sous-discipline du Machine Learning [61].

3.1.1. Réseaux de neurones

Un réseau de neurones artificiels est une série d'algorithmes qui s'efforcent de reconnaître les relations sous-jacentes dans un ensemble de données, grâce à un processus qui imite le fonctionnement du cerveau humain. En ce sens, les réseaux de neurones font référence à des systèmes de neurones, de nature organique ou artificielle [62].

3.1.2. Réseaux de Neurones Convolutifs

Les réseaux de neurones convolutionnels ou convolutifs, sont à ce jour, les modèles les plus performants, pour classer des images. Désignés par l'acronyme CNN, de l'anglais Convolutional Neural Network, ils comportent deux parties bien distinctes. En entrée, une image est fournie, sous la forme d'une matrice de pixels, elle a 2 dimensions, pour une image

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

La première partie d'un CNN est la partie convolutive, à proprement parler. Elle fonctionne comme un extracteur de caractéristiques des images. Une image passe à travers une succession de filtres, ou noyaux de convolution, créant de nouvelles images, appelées cartes de convolutions (**Figure 3.2**). Certains filtres intermédiaires, réduisent la résolution de l'image, par une opération de maximum local. Au final, les cartes de convolutions sont mises à plat, et concaténées en un vecteur de caractéristiques, appelé code CNN [63].

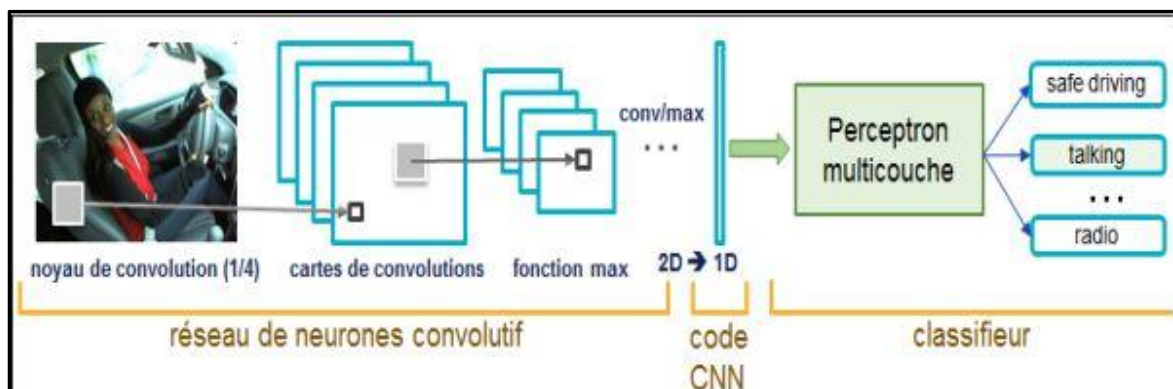


Figure 3.2: Architecture standard d'un réseau de neurone convolucif [64].

Il existe quatre types de couches pour un réseau de neurones convolutif : la couche de convolution, la couche de pooling, la couche de correction ReLU et la couche fully connected [65].

➤ Couche de convolution

La convolution est la partie la plus importante des CNN. Elle va opérer sur une image en 3 dimensions, qui a donc une longueur, une largeur et une profondeur. La profondeur peut être égale à 0, dans le cas, où l'image est en noir et blanc, et égale à 3, si elle est en couleur (R G B). En d'autres termes, notre image est, soit une matrice à 2 dimensions, si elle est en noir et blanc, soit un bloc à 3 dimensions, si elle est en couleurs [64].

La convolution consiste à faire déplacer un filtre, de taille donnée, sur l'image (**figure 3.3**). La question qui se pose maintenant est : qu'est-ce qu'un filtre ? Si on prend l'exemple d'une image en couleurs (profondeur =3), alors un filtre est lui aussi un bloc, de largeur et de hauteur, inférieures à celle de l'image, mais de même profondeur ; ce filtre, va être déplacé sur

le bloc de départ (image), il calcule pour chaque pixel la moyenne du pixel avec ses 8 proches voisins [64].

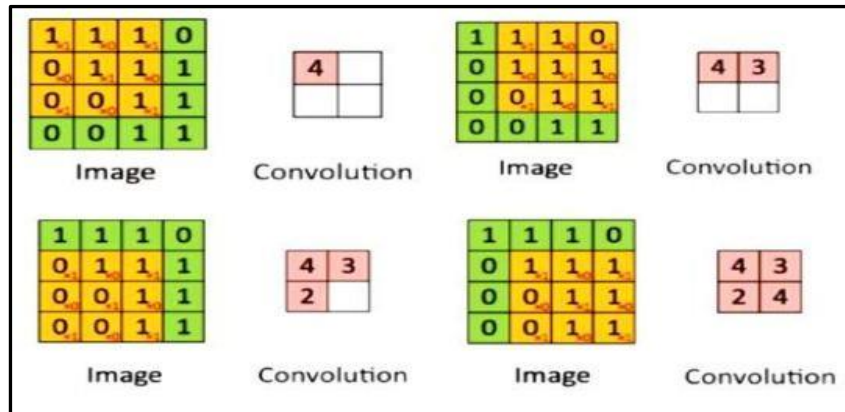


Figure 3.3: Convolution [66].

➤ Couche de Pooling (POOL)

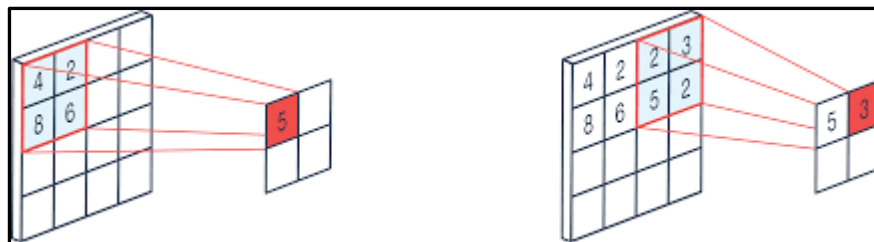
Ce type de couche est souvent placé entre deux couches de convolution : elle reçoit en entrée plusieurs cartes de caractéristiques (featuresmaps), et applique à chacune d'entre elles l'opération de pooling. L'opération de pooling consiste à réduire la taille des images, tout en préservant leurs caractéristiques importantes, cela permet de réduire le nombre de paramètres et de calculs dans le réseau, et d'éviter le sur-apprentissage [64].

Pour cela, on découpe l'image en cellules régulières, puis on remplace chaque cellule (par une certaine valeur). En pratique, on utilise souvent des cellules carrées de petite taille pour ne pas perdre trop d'informations. Les choix les plus communs sont des cellules adjacentes de taille 2*2 pixels qui ne se chevauchent pas, ou des cellules de taille 3*3 pixels, distantes les unes des autres d'un pas de 2 pixels (qui se chevauchent donc) [64]. On obtient en sortie le même nombre de feature maps qu'en entrée, mais celles-ci sont de tailles bien plus petites.

Parmi les méthodes de pooling les plus utilisés, on trouve le Max-Pooling qui consiste à remplacer chaque cellule par sa valeur maximale (Figure 3.4) ou encore l'Average Pooling (la figure 3.5) ne correspond pas à l'average pooling, vous pouvez le rajouter sur la figure 3.4 (a) max pooling et (b) average pooling) dont l'opération consiste à conserver à chaque pas, la valeur moyenne de la fenêtre du filtre [67].



(a)Maxpooling [71].



(b) Average pooling

Figure 3.4: Opérations pooling [67].

➤ **Couche de correction ReLU**

Pour améliorer l’efficacité du traitement, on intercale entre les couches de traitement, une couche qui va opérer une fonction mathématique (fonction d’activation) sur les signaux de sortie. Dans ce cadre, on trouve la fonction ReLU (Rectified Linear Units) qui désigne la fonction réelle non-linéaire définie par $ReLU(x) = \max(0,x)$ [72].

La couche de correction ReLU (**Figure 3.5**) remplace donc toutes les valeurs négatives reçues en entrées par des zéros. Elle joue le rôle de fonction d’activation [67].

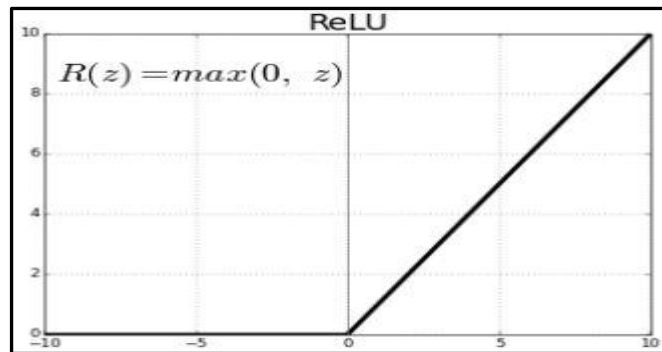


Figure 3.5 : Fonction ReLU [68].

➤ **Couche entièrement connectée (fully-connected)**

La couche fully-connected (**Figure 3.6**) constitue toujours la dernière couche d'un réseau de neurones, convolutif ou non : elle n'est donc pas caractéristique d'un CNN [69].

Ce type de couche reçoit un vecteur en entrée et produit un nouveau vecteur en sortie. Pour cela, elle applique une combinaison linéaire puis éventuellement une fonction d'activation aux valeurs reçues en entrée [69].

La dernière couche fully-connected permet de classifier l'image en entrée du réseau. Chaque élément du vecteur indique la probabilité pour l'image en entrée d'appartenir à une classe [69]. Elle détermine le lien entre la position des features dans l'image et une classe [69].

Cette couche peut être utilisée pour optimiser des objectifs tels que les scores de classe [69].

L'application de la fonction softmax aux sorties de la couche fully-connected, nous donne la probabilité d'une classe [69].

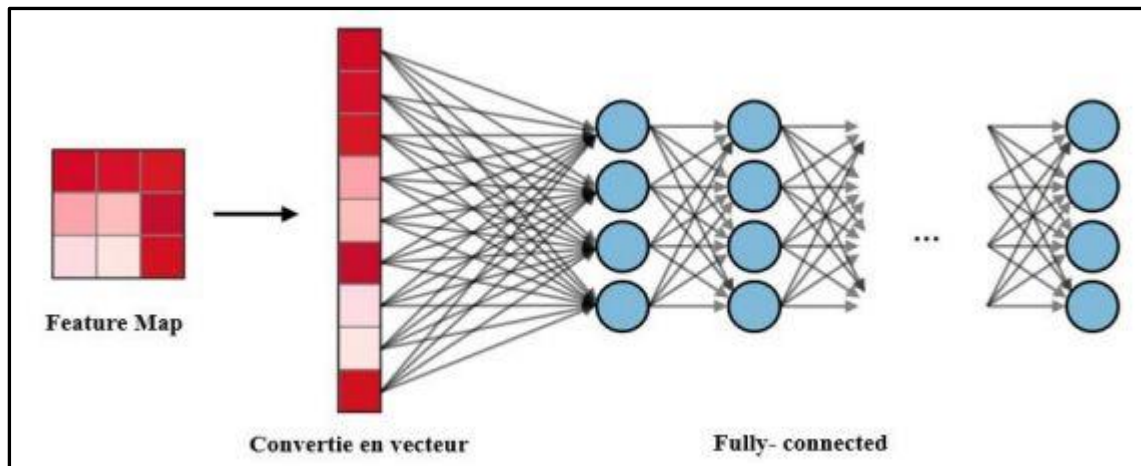


Figure 3.6 : L'architecture de Fully-connected [69].

3.2. Morphologies mathématiques

3.2.1. Élément structurant

Un élément structurant B est un ensemble de taille et de forme géométrique connues et repérées par son origine, appartenant ou non à l'élément. Cet élément est la configuration élémentaire de pixels à rechercher dans l'image. L'élément structurant choisi est déplacé de façon à ce que son origine passe par toutes les positions dans l'image à analyser. Pour chacune des positions de l'origine de l'élément (**Figure 3.7**), on se pose la question relative à l'union ou à l'intersection de l'élément structurant avec les objets de l'image. L'ensemble des points correspondant à une réponse positive permettant de construire une nouvelle image qui constitue l'image résultat. Pour la morphologie mathématique fonctionnelle l'élément structurant peut-être plat (élément plan) ou volumique [70].

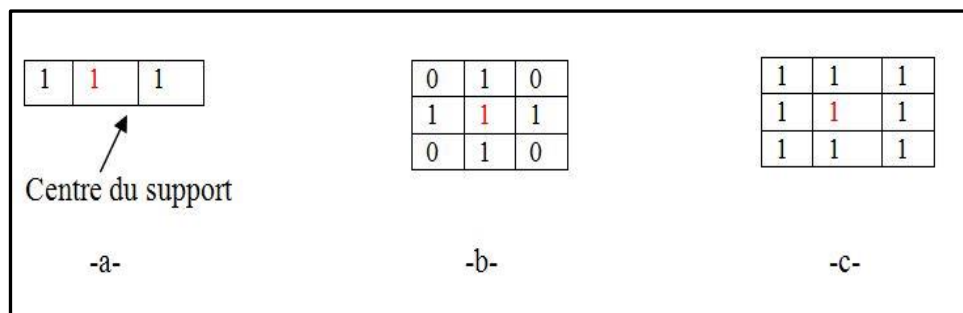


Figure 3.7 : Exemples d'élément structurant. (a) Élément structurant linéaire, (b) élément structurant cercle, (c) élément structurant carré [70].

3.2.2. Dilatation

Pour obtenir la fonction dilatée de $f(x)$, on attribue à $f(x)$ la valeur maximale qu'elle prend dans le domaine de B et ce, à chaque déplacement de B . Cette transformation comble les "vallées" et épaissit les "pics" : elle homogénéise l'image, l'éclaircit et tend à faire disparaître les objets sombres [71].

Algorithme de la dilatation :

```
-Ouvrir l'image à dilater.  
-Créer un nouveau tableau pour stocker les résultats de l'image de la  
taille de l'image originale.  
X=largeur de l'image.  
Y=hauteur de l'image.  
  
    Pour tous les i allant de 1 à X  
        Pour tous les j allant de 1 à Y  
            ImDil [i] [j]=max (voisinage compris dans l'élément structurant de ce  
pixel)  
        Fin de Pour.  
    Fin de Pour.
```

3.2.3. Erosion

Pour obtenir la fonction érodée de $f(x)$, on attribue à $f(x)$ la valeur minimale qu'elle prend dans le domaine de l'élément structurant B et ce, à chaque nouveau déplacement de B . Cette transformation a pour propriétés de réduire les "pics" de niveaux de gris et d'élargir les "vallées". Elle tend donc à homogénéiser l'image, à l'assombrir et à établir le bord des objets les plus sombres [71].

Algorithme de l'érosion :

-Ouvrir l'image à éroder.

-Créer une image érodée pour stocker les résultats de l'image de la taille de l'image originale

X=largeur de l'image.

Y=hauteur de l'image

Pour tous les **i** allant de **1** à **X**

Pour tous les **j** allant de **1** à **Y**

 ImEro [i-1] [j-1]=**min** (voisinage compris dans l'élément structurant de ce pixel)

Fin de Pour.

Fin de Pour.

-Reconstruire l'image érodée.

3.2.4. Fermeture morphologique

La fermeture d'un ensemble X , noté $X \bullet B$, est l'enchaînement d'une dilatation D_B suivie d'une érosion E_B par le même élément structurant B :

$X \bullet B = E_B(D_B(X))$	(3.1)
-----------------------------	-------

La fermeture est la duale de l'ouverture : la fermeture du complémentaire d'un ensemble est égale au complémentaire de l'ouverture de cet ensemble [72].

3.2.5. Ouverture morphologique

Une ouverture d'un ensemble X , noté $X \circ B$, est l'érosion par B suivie d'une dilatation avec B :

$X \circ B = D_B(E_B(X))$	(3.2)
---------------------------	-------

L'ouverture peut être caractérisée géométriquement, elle donne l'union de tous les B s inclus dans X . Ainsi, la forme de l'élément structurant permet de choisir les structures qui peuvent le contenir [78].

3.2.6. Chapeau haut de forme (CHF)

Le chapeau haut-de-forme est un opérateur issu de la morphologie mathématique permettant d'extraire des images à teintes de gris, des objets contrastés de petite taille, correspondant à des défauts de textures. Cette transformation est très populaire en ce qu'elle permet d'extraire d'une image les structures contrastées et de faible épaisseur, ceci indépendamment de la valeur absolue de l'éclairage ambiant.

La transformation chapeau haut-de-forme peut se diviser en « chapeau haut-de-forme noir » et en « chapeau haut-de-forme blanc ».

- Le chapeau haut-de-forme noir (Black Top-Hat) permet la détection des zones foncées des images. Cette transformation correspond à la différence, seuillée à une valeur "t", de l'image fermée et de l'image initiale. La fermeture agit telle un filtre de taille, dont la largeur est déterminée par la taille de la fermeture.
- Le chapeau haut-de-forme blanc (White Top-Hat) correspond à la différence seuillée (à une valeur z), de l'image initiale et de l'image ouverte.

En pratique, on commence par faire la différence entre la fonction $f(x)$ de départ et la fonction $f(x)$ "ouverte". Après, on choisit une valeur du seuil "t" qui retiendra ensuite uniquement des objets dont l'intensité en niveau de gris est supérieure à ce seuil.

L'algorithme du chapeau haut-de-forme :

Début

Image Originale

ImEro=Erosion de l'image originale.

ImDil=Dilatation de l'image érodée.

Pour i=1 jusqu'à largeur Faire

Pour j=1 jusqu'à hauteur Faire

{

WTH[i][j]=Image Original [i][j]-ImDil[i][j]

}

Fin de Pour.

Fin de Pour.

Fin.

Cette transformation seuillée doit son nom de transformation "chapeau haut-de-forme" du fait que l'on ne conserve de $f(x)$ que les portions entrant dans le chapeau (défini par la taille de l'élément structurant) et dépassant son sommet. Cette transformation est intéressante car elle est insensible aux faibles variations de niveaux de gris et elle nous permet de mettre en évidence les détails en extrayant les zones claires qui sont éliminées par ouverture.

3.3. Fusion des données

La fusion de données est un domaine en émergence, avec des applications variées. Historiquement, la plupart des techniques de fusion de données ont été développées pour des applications militaires : la surveillance des océans, la défense aérienne, le renseignement sur les champs de bataille et l'identification de cibles. La technologie est maintenant largement utilisée dans des applications non militaires, telles que l'acquisition de données, le contrôle automatisé, la surveillance d'équipements complexes et la robotique [73].

La fusion peut se pratiquer à trois moments différents du travail (Figure 3.8). Soit au niveau du pixel (fusion des pixels), soit au niveau des caractéristiques après une segmentation (fusion des objets extraits de l'image), soit au niveau décisionnel lors de la phase finale de la segmentation (fusion des objets extraits et identifiés) [73].

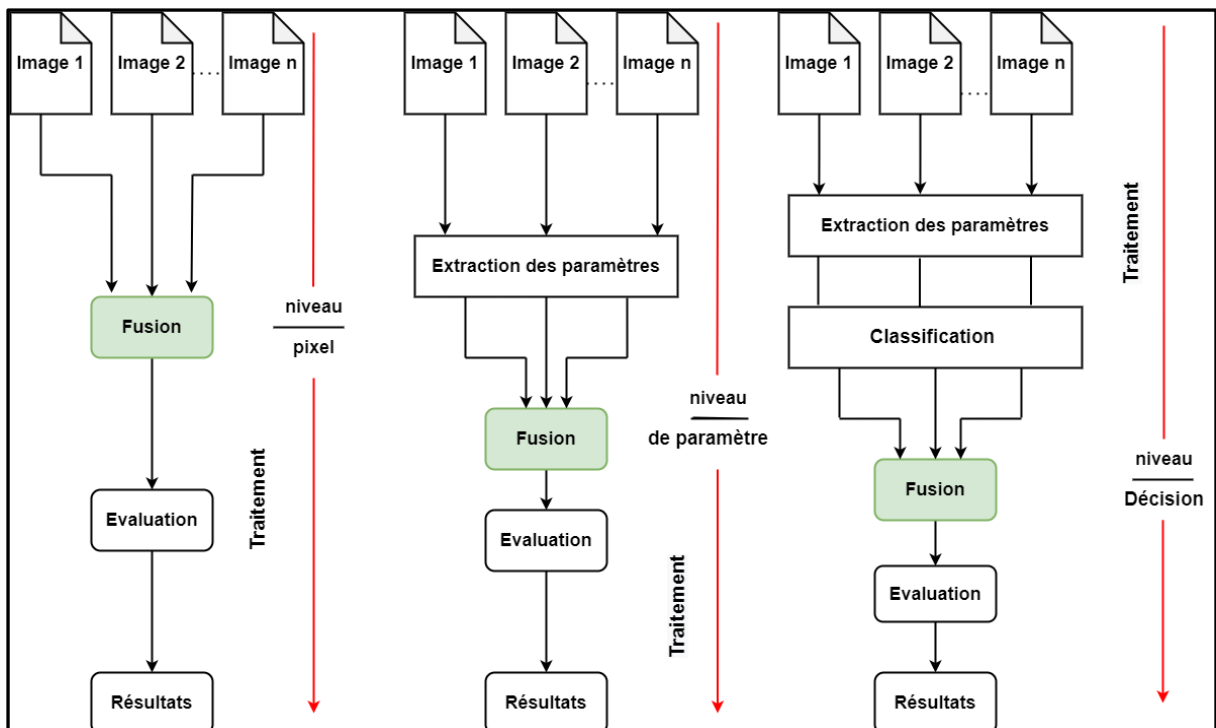


Figure 3.8: Niveaux de traitements de la fusion d'images [73].

3.3.1. Définition de la fusion d'images

Il existe plusieurs définitions de la fusion d'images dans le champ de la télédétection. On définit généralement la fusion d'images comme la combinaison de deux ou de plusieurs images différentes pour former à l'aide d'un algorithme une nouvelle image [73].

Le but de la fusion d'images est d'intégrer une information complémentaire à partir d'une multitude d'images pour créer une image fortement instructive, qui serait plus appropriée à la perception visuelle humaine et aux tâches de traitement par ordinateur. Cette technique est particulièrement utile lorsqu'on veut combiner des images de différentes modalités.

La procédure de fusion comprend généralement deux étapes. D'une part, les images doivent être superposables, ce qui implique leur géo référenciation par une correction géométrique et un rééchantillonnage approprié. La deuxième étape correspond à la fusion proprement dite [73].

a) Fusion au niveau du pixel

Pour effectuer la fusion au niveau du pixel, il doit y avoir plusieurs plans ou modalités différentes pour chaque image, afin de fusionner chaque pixel de chaque modalité [73].

b) Fusion au niveau des caractéristiques

Pour la fusion au niveau des caractéristiques, un vecteur d'attributs est extrait à partir de la sortie de chaque image. Ces vecteurs d'attributs sont ensuite combinés (fusionnés) et une déclaration d'identification est ensuite effectuée sur la base de ce vecteur conjoint. Les outils utilisés pour la déclaration d'identification comprennent les techniques statistiques (par ex., analyse de regroupements), les réseaux de neurones, les techniques structurelles et à base de connaissances [73].

c) Fusion au niveau des décisions

Pour la fusion au niveau des décisions (ou niveau des déclarations), chaque image effectue de façon indépendante un estimé ou une déclaration de la scène observée. Ces estimations sont ensuite combinées via un procédé de fusion. Les techniques de fusion de déclarations incluent les méthodes de votation, l'inférence bayésienne et la théorie de Dempster-Shafer (Théorie de l'évidence) [73].

Dans notre cas, on utilise la fusion au niveau des décisions pour but d'affiner la précision de la segmentation et de diminuer les faux positifs obtenus.

d) Applications de la fusion d'images

Le champ d'application de la fusion d'images est potentiellement vaste. Dans le contexte du rehaussement d'images, elle a pour tâche principale de réunir au sein d'une image, l'information provenant de différentes sources imagées (images multi sources). La procédure de fusion doit avant tout être considérée comme une technique d'assistance à l'interprétation visuelle. Pour des besoins de zonage ou de traitements numériques, la combinaison d'informations multisources procède d'une autre démarche. Pour des fins d'analyse visuelle, on souhaite habituellement une image la plus riche possible [73]. On retrouve parmi les domaines d'applications :

- L'aide à la navigation.
- Photos prises en dehors du focal (out-of-focus).
- L'imagerie Médicale.
- La télédétection.

3.4. Approche proposée

L'organigramme de la **figure 3.9** montre les différentes étapes, utilisées dans notre approche de détection et de segmentation des Mcs. L'approche proposée est divisée en deux étapes complémentaires, une première étape de détection de la région d'intérêt, et la deuxième pour la segmentation de cette dernière, afin de raffiner au mieux les résultats.

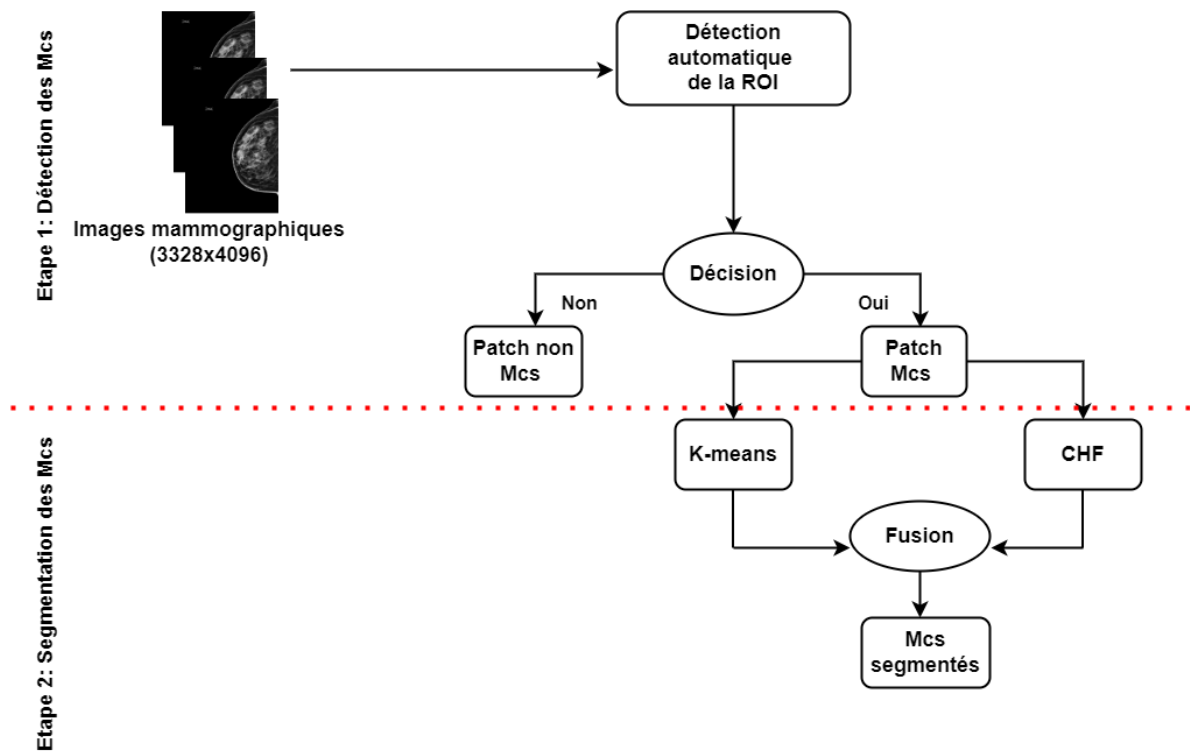


Figure 3.9: Organigramme général de la méthode proposée.

3.4.1. Détection automatique de la région d'intérêt

L'étape de détection de la région d'intérêt va nous permettre de réduire la zone de travail. Au lieu de travailler sur une image de taille (3328 x 4096), nous travaillerons sur une région d'intérêt de taille (100 x 100). L'avantage d'un tel traitement est double : d'une part, la zone de travail est considérablement réduite, pour des traitements ultérieurs, et donc le volume des données à traiter est réduit, d'autre part, travailler sur une région d'intérêt, rend le traitement plus performant, car plus ciblé, et moins perturbé par des informations, pouvant être considérées comme du bruit. Dans cette étape nous avons effectué deux approches différentes, pour la détection du ROI, puis nous avons choisi celle qui donne le meilleur résultat.

3.4.2. Détection automatique par apprentissage en profondeur : Approche 1

a) Division de l'image en patch

Dans cette étape nous avons effectué une division de l'image (figure 3.10), en patch de taille égale à 100 x 100. Cette taille est étudiée par rapport à la taille moyenne de Mcs qui varie entre 0.5mm et 5mm. Chaque patch de l'image sera utilisé comme entrée au classificateur, pour

un apprentissage en profondeur avec l'architecture du Resnet50. Cette étape permettra de faire la sélection des patches Mcs, parmi les autres patches qui ne contiennent pas de Mcs.

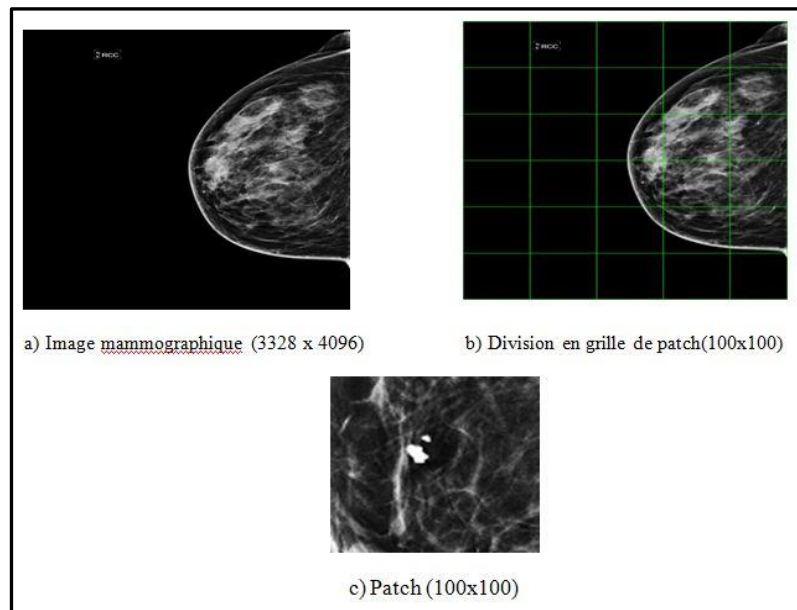


Figure 3.10 : Division de l'image mammographique.

L'approche 1 proposée est basée sur l'apprentissage en profondeur (Deep-learning) avec l'architecture du ResNet50. Après l'étape de division en grille de 100x100 de l'image, les différents patches sont répertoriés en deux classes différentes : patches Mcs et les patches non Mcs. Ainsi triée, les deux classes sont considérées comme entrée du classifieur pour la génération du modèle d'apprentissage.

Un système de classification automatique d'images par l'apprentissage en profondeur est composé de deux phases, la phase d'apprentissage qui consiste en la construction d'un modèle de décision et contient en général le 2/3 des données, et la phase de test, pour connaître les performances du modèle généré et contient le reste des données, soit 1/3 des données.

Le schéma global de l'approche 1 est illustré dans la **figure 3.11**.

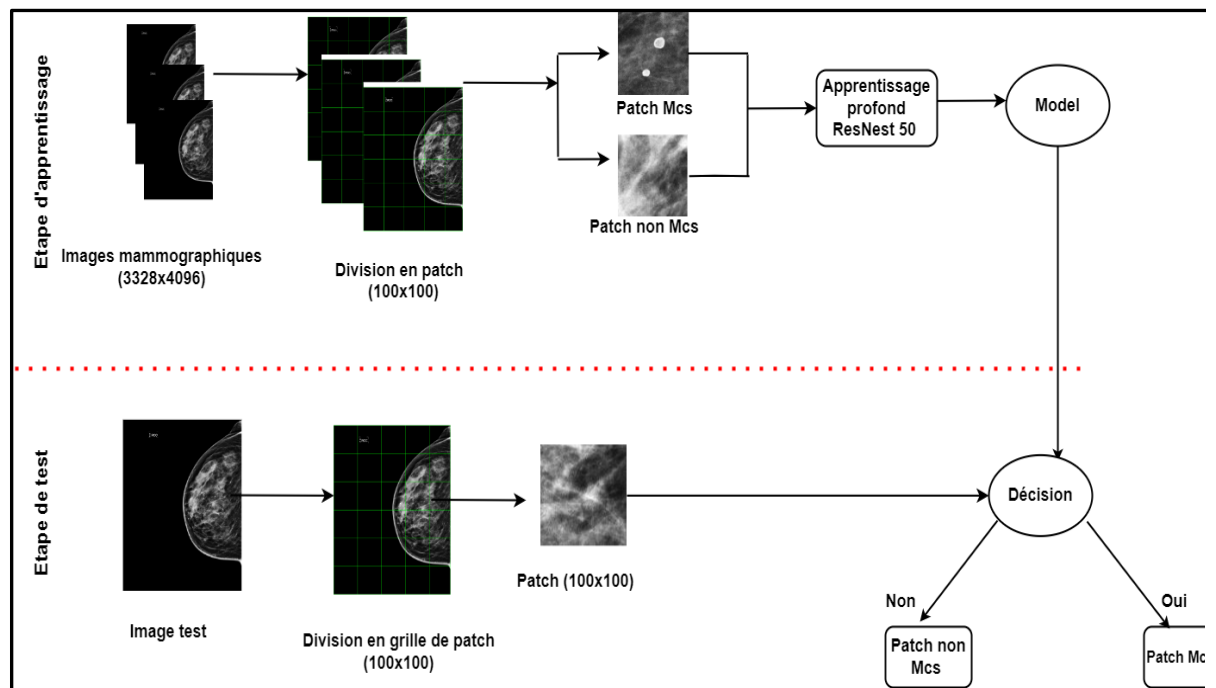


Figure 3.11: Organigramme global de l'étape de détection des Mcs.

Nous allons présenter dans ce qui suit, quelques définitions pour mieux comprendre l'apprentissage en profondeur.

b) Classification des régions d'intérêts par le ResNet50

La classification, est une étape que l'on retrouve, de façon récurrente dans le domaine biomédical. En effet, les biologistes et médecins sont régulièrement amenés, à établir un diagnostic, ce qui revient généralement, à attribuer une classe générale y , au phénomène x observé (pathologie, cellule, . . .). Cette étape va nous permettre, de classer les patchs de taille 100 x 100, en deux classes différentes, soit en un cas de microcalcification, ou bien en un cas de non microcalcification. L'étape de prédiction nécessite, un apprentissage important et crucial, pour atteindre un haut niveau de performance ; cela dépend, des types, qualitatif et quantitatif des données, pour cette étape de classification nous avons choisi l'architecture du ResNet50 [62].

Un réseau de neurones résiduels (ResNet) est un réseau de neurones artificiels (*artificial neuronal network*) d'un type qui s'appuie sur des constructions connues à partir de cellules pyramidales du cortex cérébral. Pour ce faire, les réseaux de neurones résiduels utilisent des connexions sautées ou des raccourcis pour sauter par-dessus certaines couches.

ResNet a de nombreuses variantes qui fonctionnent sur le même concept mais ont des nombres de couches différents (Figure 3.12). Resnet50 est utilisé pour désigner la variante qui peut fonctionner avec 50 couches de réseau de neurones [62].

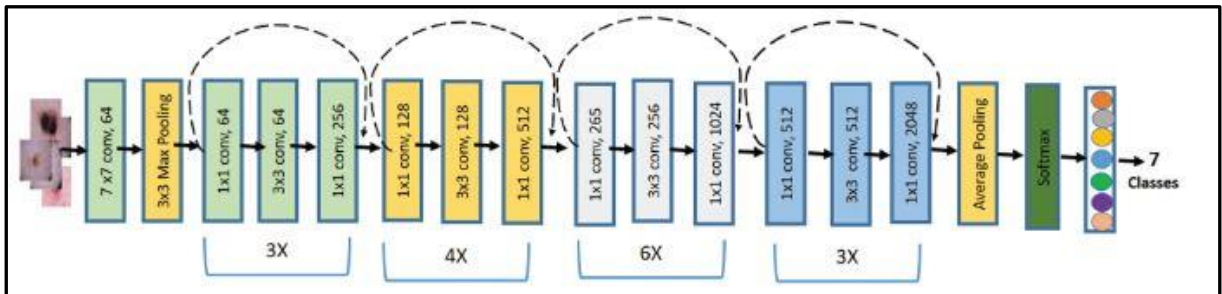


Figure 3.12 : Architecture de ResNet 50[62].

L'architecture de ResNet 50 contient quatre couches [62] :

- Une couche pooling avec une probabilité de 50 % - cela supprime 50 % des paramètres de manière aléatoire et réduit le sur ajustement.
- Une couche d'activation d'unité linéaire rectifiée (ReLU) avec régularisation L2/régression de crête. La valeur du paramètre de régularisation est 0,02. La régularisation est nécessaire pour éviter le sur ajustement du modèle, et cela se fait en ajoutant une pénalité avec la fonction de coût.
- Vient ensuite une autre couche avec une probabilité d'abandon de 50 %.
- Et enfin une couche d'activation soft max avec régularisation L2 (paramètre de régularisation = 0) [76].

3.4.3. Détection automatique de la ROI par les outils morphologiques : Approche 2

La deuxième approche proposée est basée sur l'utilisation des différents outils morphologiques :

- En premier, on effectue un prétraitement de l'image.
- En deuxième étape on applique le CHF
- En dernier, on divise l'image en patch de (100x100), la taille des patches est choisie de sorte qu'elle peut contenir un foyer de Mcs, et on applique une condition à chaque patch pour ne laisser que le patch qui vérifie la valeur maximum de tous les patches.

Les étapes de la détection automatique du ROI sont présentées dans la figure qui suit (Figure 3.13) :

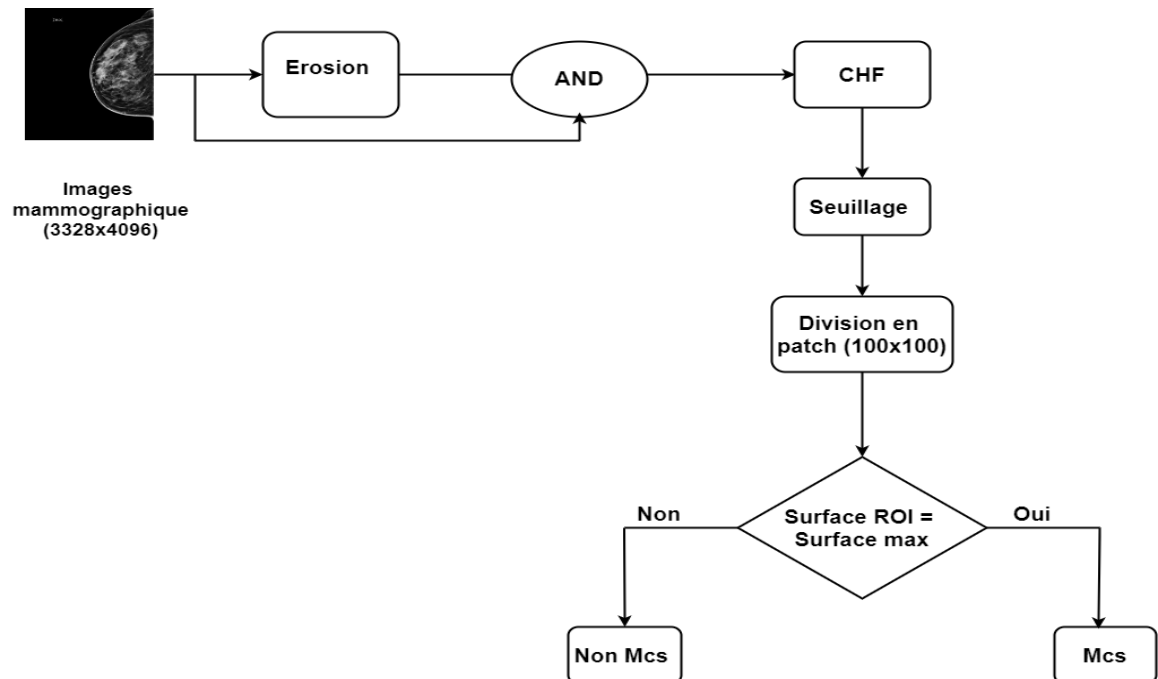


Figure 3.13 : Détection automatique de la ROI.

a) Prétraitement de l'image

Dans cette étape on supprime le mamelon, l'aréole et la peau qui entoure le sein, leurs intensités sont très élevées et peuvent compliquer la détection des Mcs. On applique l'érosion en deux itérations avec un élément structurant carré de taille (19x19), cela permet la suppression des particules plus petites que l'élément structurant utilisé. Cette taille d'élément structurant est suffisante pour supprimer le contour du sein qui présente une intensité proche des Mcs. Par la suite on multiplie le masque obtenu avec l'image originale, pour reconstruire notre image en niveaux de gris.

b) Segmentation par le chapeau haut de forme (CHF)

Les deux principales caractéristiques des Mcs sont leur petite taille et leur atténuation élevée. Une méthode de détection des Mcs basée sur l'opérateur chapeau haut-de-forme est présentée, elle combine ces deux caractéristiques.

L'opérateur chapeau haut-de-forme seuillé est composé de trois étapes :

- 1 - Ouverture de l'image avec un élément structurant.
- 2 - Soustraction de l'image résultant de l'ouverture à l'image originale.
- 3 - Application d'un seuil sur le résultat.

L'ouverture élimine toutes les structures plus petites que l'élément structurant et ne garde que les structures plus grandes. La soustraction de cette image à l'image originale contient donc uniquement les petites structures qui ont été effacées par l'ouverture [72].

Finalement, le seuillage ne prend en compte que les structures présentant un certain contraste. Nous avons utilisé le seuillage automatique d'Otsu [72].

Puisque la taille des Mcs n'est pas connue d'avance, plusieurs opérateurs chapeau haut-de-forme avec des éléments structurants de tailles différentes sont utilisés, pour choisir le meilleur résultat.

3.4.4. Segmentation des microcalcifications

La deuxième partie de notre approche est la segmentation des foyers de Mcs, après détection des régions d'intérêts par l'étape précédente. L'approche de segmentation est constituée par la fusion de deux méthodes différentes, l'une est basée sur la labélisation des différentes structures de la région d'intérêt, et la deuxième est basée sur les outils morphologique (CHF), l'étape de la segmentation est présentée dans la figure qui suit :

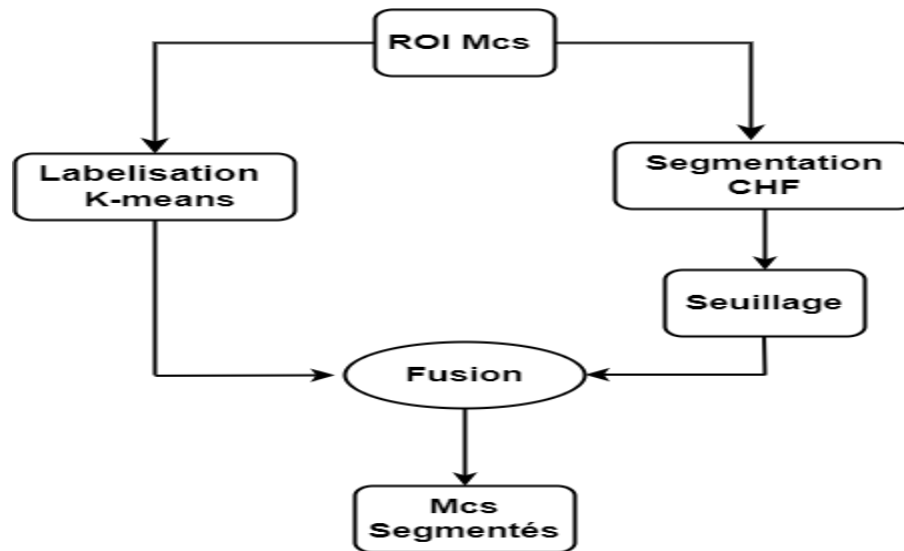


Figure 3.14: Organigramme de l'étape de segmentation des Mcs.

3.4.5. Segmentation par K-means

Le clustering est une méthode permettant de diviser un ensemble de données en un nombre spécifique de groupes. L'une des méthodes de clustering la plus populaire est le clustering par k-means. (Le clustering par k-means consiste à classer un ensemble de données en un nombre k de clusters disjoints. L'algorithme K-means se compose de deux phases distinctes [73].

Dans la première phase, il calcule le centroïde k et dans la deuxième phase, il affecte chaque point au cluster qui a le centroïde le plus proche du point de données respectif. Il existe différentes méthodes pour définir la distance du centroïde le plus proche. L'une des méthodes les plus utilisées est la distance euclidienne. Une fois le regroupement effectué, il recalcule le nouveau barycentre de chaque cluster et sur la base de ce centroïde, une nouvelle distance euclidienne est calculée entre chaque centre et chaque point de données est attribué au cluster qui a une distance euclidienne minimale. Chaque cluster de la partition est défini par ses objets membres, et par son centroïde. Le centre de gravité de chaque cluster est le point auquel la somme des distances de tous les objets de ce cluster est minimisée (Figure 3.15). Ainsi, K-means est un algorithme itératif qui minimise la somme des distances de chaque objet à son centroïde de cluster, sur tous les clusters [73].

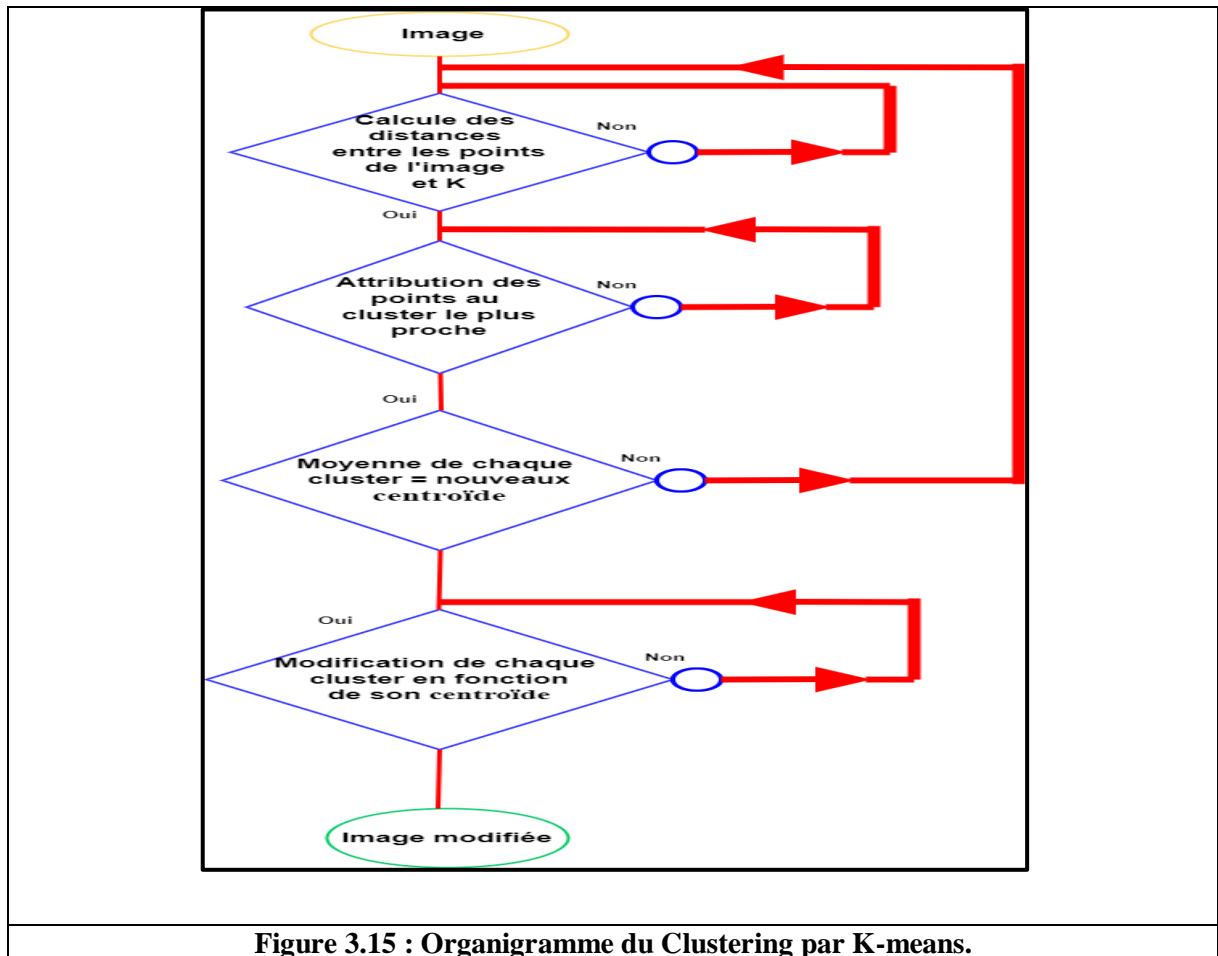


Figure 3.15 : Organigramme du Clustering par K-means.

Algorithme de K-means :

Algorithme
 Entrée : Image originale I
 Résultat : centres de cluster $C = \{c_1, \dots, c_k\}$
 $K=N$
 $S1 = S2 = \dots = Sk = \{\}$; partition
 Pour $i \in \{1, \dots, N\}$ faire
 $I = \arg \min_{j \in \{1, \dots, k\}} \|X_i - c_j\|^2$
 $S1 = S1 \cup \{xi\}$; xi est l'ensemble de données d'apprentissage
 fin
 pour $j \in \{1, \dots, k\}$ faire
 $c_j = \frac{\sum_{ni \in \{1, \dots, N\}, xi \in S_j} xi}{|S_j|}$
 fin

Nous avons utilisé l'algorithme de k-means pour ses nombreux avantages tels :

- Facile à interpréter.
- Faible coût de calcul.
- Précision.
- Convient aux gros data sets.
- Simple

3.4.6. Fusion de plusieurs segmentations alternatives

Dans cette section, nous décrivons les règles de fusion, et les approches d'ensemble correspondantes, qui sont utilisées pour générer une image de segmentation fusionnée qui combine des informations au niveau des pixels à partir des images segmentées par K-means et par CHF.

On s'est inspirée des travaux de **Iván Cabriaa** et **Iker Gondra** [74], pour l'utilisation de deux règles de fusion différentes qui combinent l'information au niveau pixel à partir des images segmentées et générées par l'algorithme du k-means (noté $S_{k-means}$), et des images segmentées et générées par l'algorithme du CHF (noté $S_{Top-hat}$),

- La première règle de fusion est l'intersection, qui sera notée INT, qui se traduit par une segmentation d'intersection fusionnée S_{INT} . Cette segmentation est générée en étiquetant un pixel pi comme pixel Mc si et seulement s'il est étiqueté comme pixel Mc dans chacune des segmentations $S_{k-means}$ et $S_{Top-hat}$, c'est à dire $pi \in S_{INT}$ si et seulement si :

$$pi \in S_{k-means} \cap S_{Top-hat}$$
 [74].

L'algorithme 1 résume cette approche de fusion par intersection.

Algorithme1 :

Entree: Image originale I

Segmentation $S_{k-means}$ générée par K-means

Segmentation $S_{Top-hat}$ générée par Top-hat

$S_{INT} = \emptyset$

Pour chaque $pi \in I$ faire

Si $pi \in S_{k-means} \cap S_{Top-hat}$ alors

$S_{INT} = S_{INT} \cup \{pi\}$

Fin Si

Fin pour

- La deuxième règle de fusion est l'union, qui sera notée UNI , et aboutit à une segmentation syndicale fusionnée S_{UNI} [74]. Cette segmentation est générée en étiquetant un pixel pi comme un pixel Mc si et seulement s'il est étiqueté comme pixel Mc par l'une (au moins une) des segmentations $S_{k-means}$ et $S_{Top-hat}$, c'est-à-dire : $pi \in S_{UNI}$ si et seulement si $pi \in S_{k-means} \cup S_{Top-hat}$.

L'algorithme 2 résume cette approche de segmentation.

Algorithme2 :

Entree: Image originale I

Segmentation $S_{k-means}$ générée par K-means

Segmentation $S_{Top-hat}$ générée par Top-hat

$S_{UNI} = \emptyset$

Pour chaque $pi \in I$ faire

Si $pi \in S_{k-means} \cup S_{Top-hat}$ alors

$S_{UNI} = S_{UNI} \cup \{pi\}$

Fin Si

Fin pour

3.5. Discussion

Dans ce chapitre, nous avons évoqué un ensemble de méthodes pour détecter les microcalcifications (Mcs) dans les images mammographiques, et nous avons cité les différentes techniques utilisées dans le traitement ainsi que leurs avantages. La bonne connaissance de ces techniques, nous sert comme bagage pour la détection des Mcs. Nous avons présenté les différentes étapes de notre approche : la détection de la région d'intérêt et la segmentation des Mcs. Nous allons présenter l'implémentation et les résultats de ces méthodes que nous avons obtenues, dans le prochain chapitre.

Chapitre 4

Résultats et Interprétations

Dans ce dernier chapitre, nous allons tous d'abord présenter la base de données utilisée dans la validation de notre système de détection et de segmentation des Mcs, puis l'environnement et le langage d'implémentation dans lequel nous avons développé notre approche. Par la suite, nous présenterons les résultats obtenus par l'approche proposée ainsi qu'une présentation de l'interface de l'application réalisée.

4.1. Base de données

Une base de données est un ensemble d'informations qui est organisé de manière à être facilement accessible, géré et mis à jour. Elle est utilisée par les organisations comme méthode de stockage, de gestion et de récupération de l'information [75].

Dans notre travail, nous avons utilisé une base des données téléchargées BMD 2021 (BREAST MICRO-CALCIFICATION DATASET) [2]. Elle contient 100 fichiers de mammographies numériques, provenant de divers hôpitaux. Cet ensemble de données se compose de deux dossiers : cas normaux et cas suspects. Les deux dossiers contiennent 50 sous-dossiers qui représentent 50 patients (50 dossiers, 50 patients dans chaque cas). À l'intérieur de chaque dossier, il y a 8 fichiers : les quatre mammographies au format DICOM de vue Cranio-Caudal (CC) et Medio-Lateral oblique (MLO) pour le sein gauche et droit (CC_prior, CC_recent, MLO_prior, MLO_recent) comme le montre la figure 4.1, ainsi que les images leur correspondant au format PNG segmentées par les experts (vérité terrain) [76].

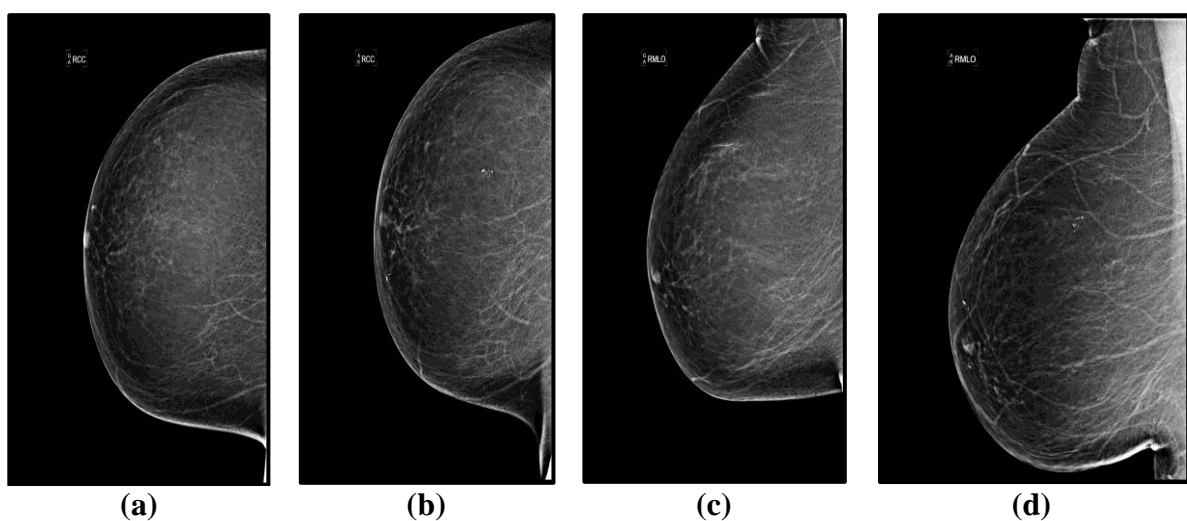


Figure 4.1: Exemples d'images de la base des données [2]. (a) Image CC_prior, (b) image CC_recent, (c) image MLO_prior, (d) imageMLO_recent.

4.2. Environnement et langage d'implémentation

L'implémentation de notre algorithme a été réalisée à l'aide du langage de programmation Python.

4.2.1. Logiciel et matériel

Ce travail a été fait sur un ordinateur portable **DELL** caractérisé par :

- un processeur Intel® Core™ I5-4300M CPU 2.6 GHz.
- une mémoire RAM de 8.0 Go.
- Un système d'exploitation Windows 7 avec 64 bits.

Et un autre ordinateur portable DELL caractérisé par :

- Un processeur INtel® Celeron® 2957U CPU 1.4 GHz.
- Une mémoire RAM de 2.0 Go.
- Un système d'exploitation Windows 7 avec 32 bits.

4.2.2. Langage de programmation utilisé

L'implémentation de notre programme a été faite avec le langage Python 3 (version 2018). C'est un langage de programmation très utilisé dans le domaine du machine learning, du big data et de la data science [77]. Le langage Python est conçu pour optimiser la productivité des programmeurs en offrant des outils de hauts niveaux et une syntaxe simple à utiliser. Il est également appréciée par les pédagogues qui y trouvent un langage ou la syntaxe, clairement séparée des mécanismes de bas niveau, permet une initiation plus aisée aux concepts de base de la programmation [78].

La Python Software Foundation est l'association qui organise le développement de Python et anime la communauté de développeurs et d'utilisateurs. Ce langage de programmation présente de nombreuses caractéristiques intéressantes [79] :

- Il est multiplateforme, c'est-à-dire qu'il fonctionne sur de nombreux systèmes d'exploitation : Windows, Mac OS X, Linux, Android, iOS, depuis les mini-ordinateurs Raspberry Pi jusqu'aux supercalculateurs.
- Il est gratuit, il peut être installé sur autant d'ordinateurs que l'on veut.
- C'est un langage de haut niveau. Il demande relativement peu de connaissance sur le fonctionnement d'un ordinateur pour être utilisé.

- C'est un langage interprété, un script Python n'a pas besoin d'être compilé pour être exécuté, contrairement à des langages comme le C ou le C++.
- Il est orienté objet, c'est-à-dire qu'il est possible de concevoir en Python des entités qui miment celles du monde réel (une cellule, une protéine, un atome, etc.) avec un certain nombre de règles de fonctionnement et d'interactions.
- Il est relativement simple à prendre en main.
- Enfin, il est très utilisé en bio-informatique et plus généralement en analyse de données.

Par conséquent, nous avons choisi de développer notre méthode à l'aide du langage Python, du fait de ses nombreux avantages

Anaconda : est une plate-forme informatique gratuite. Il est possible de l'installer en fonction du système d'exploitation Windows, Linux, Mac OS, il se compose des distributions Python et R et du gestionnaire de paquets appelé conda. Anaconda fournit un ensemble de bibliothèques et packages préinstallés. Certains d'entre eux sont NumPy, SciPy, Pandas, Scikit Learn, Nltk et Jupiter [80] (**Figure 4.2**).

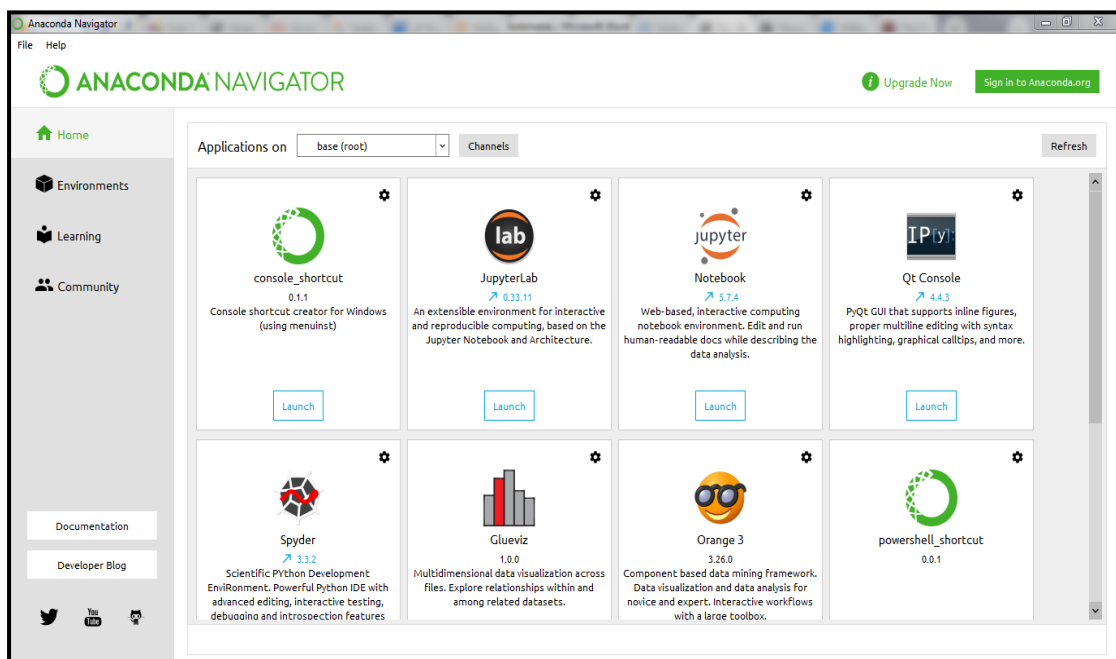


Figure 4.2: Interface d'Anaconda.

Bibliothèques nécessaires pour notre travail.

Tensorflow : c'est un framework de Google, pour faire du Deep Learning, et est un outil, open source, d'apprentissage automatique, développé par Google. Le code source a été ouvert le 9 novembre 2015 par Google. Tensorflow est l'un des outils les plus utilisés, dans le domaine de

l'apprentissage automatique. On va utiliser Keras, qui est un framework, s'appuyant sur Tensorflow (il simplifie les commandes, et ajoute certaines fonctionnalités) [81].

Keras : est une bibliothèque open source, écrite en python, qui peut fonctionner, sur Theano ou Tensorflow, et permet d'interagir avec les algorithmes de réseaux de neurones profonds, et de machine learning. Elle a été initialement écrite par François Chollet. Elle a été développée, pour l'implémentation des modèles d'apprentissage en profondeur, aussi rapidement et facilement que possible [81].

NumPy : est une extension du langage de programmation Python, destinée à manipuler des matrices, ou tableaux multidimensionnels, ainsi que des fonctions mathématiques, opérant sur ces tableaux. Plus précisément, cette bibliothèque logicielle, libre et open source, fournit de multiples fonctions, permettant notamment de créer directement un tableau, depuis un fichier, ou au contraire, de sauvegarder un tableau dans un fichier, et de manipuler des vecteurs, matrices et polynômes [81].

OpenCV : pour Open Computer Vision, est une bibliothèque graphique libre, initialement développée par Intel, spécialisée dans le traitement d'images, en temps réel. La société de robotique Willow Garage et la société ItSeez, se sont succédées, au support de cette bibliothèque [81].

Pillow : Python Imaging Library (ou PIL) est une bibliothèque de traitement d'images, pour le langage de programmation Python. Elle permet d'ouvrir, de manipuler, et de sauvegarder différents formats de fichiers graphiques [81].

Matplotlib : est une bibliothèque du langage de programmation Python, destinée à tracer, et à visualiser des données, sous forme graphique. Elle peut être combinée avec les bibliothèques Python, de calcul scientifique NumPy et SciPy [81].

Description de l'interface Spyder

Dans notre travail, nous avons utilisé l'éditeur Spyder qui permet le développement des différents programmes. **La figure 4.3** représente l'interface de Spyder.

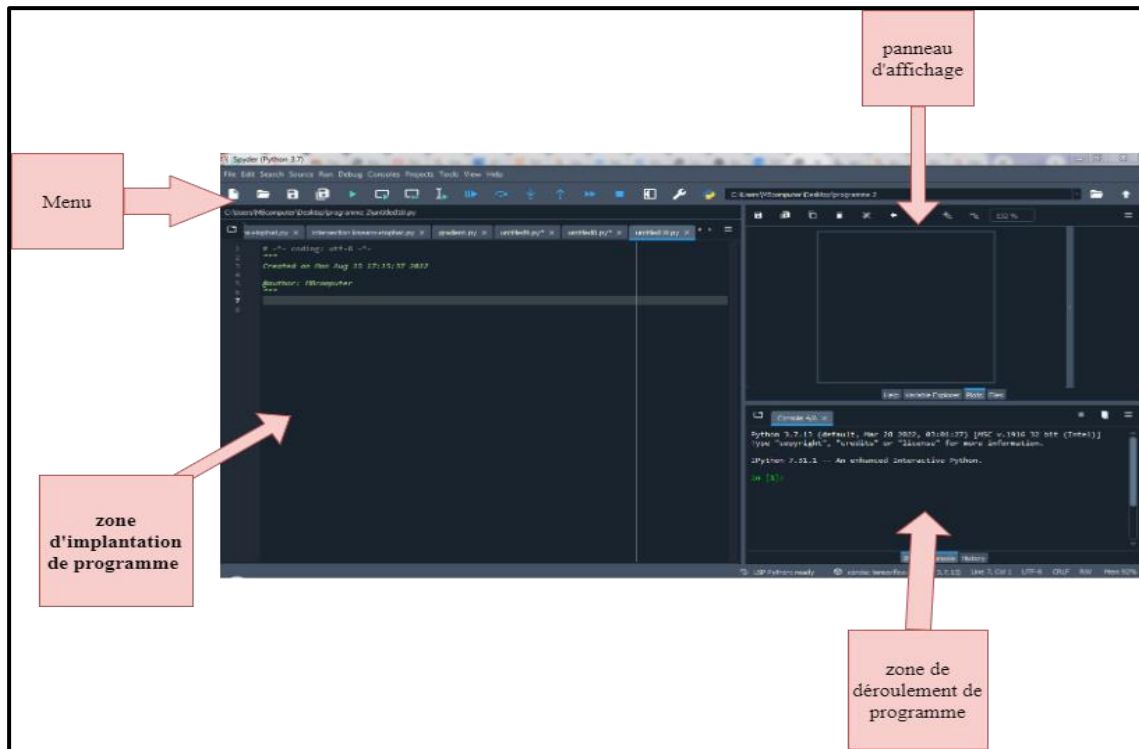


Figure 4.3: Description de l'interface Spyder.

4.3. Implémentation de l'algorithme

Dans cette partie, nous allons présenter toutes les étapes de notre approche en détail, dont l'objectif est de détecter et de segmenter les microcalcifications mammographiques. **La figure 4.4** présente la procédure générale adoptée pour le test et l'évaluation de l'approche proposée. Une détection de la région d'intérêt est nécessaire pour faciliter notre travail, qui sera suivie de l'étape de segmentation.

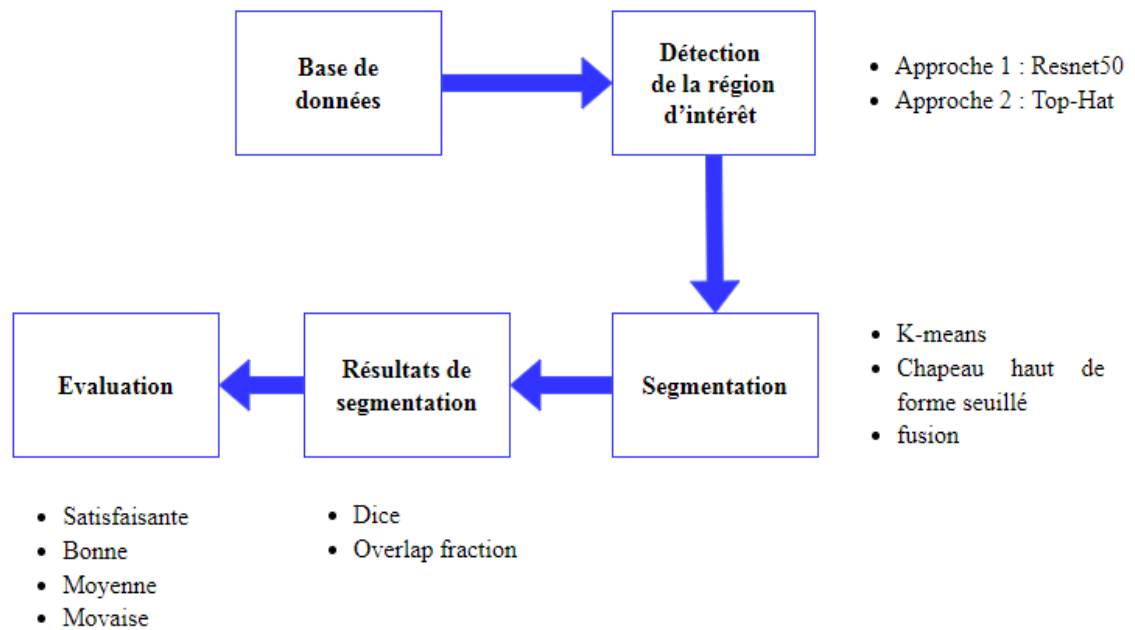


Figure 4.4: Schéma des différentes étapes de l'approche proposée.

4.4. Résultats de la détection automatique de la région d'intérêt

Le but de cette approche est la détection de la région d'intérêt qui présente le foyer des Mcs) et l'élimination des régions non suspectes, pour rendre notre traitement plus performant, plus ciblé et supprimer parallèlement tous types de bruit. Nous avons procédé à deux approches pour extraire la région d'intérêt (ROI).

4.4.1. Résultats de détection automatique de la ROI par apprentissage en profondeur : Approche 1

Comme nous avons expliqué dans le chapitre précédent, nous avons tous d'abord effectué une étape de préparation des données d'entrées au classifieur. Nous avons effectué une division des images en patches de taille de 100x100, puis nous avons trié les patches résultants de la division en se basant sur les images de la vérité terrain. Le patch qui contient un foyer de Mcs est répertorié dans un fichier Mcs et le patch qui ne contient pas un foyer de Mcs est répertorié dans un fichier non Mcs (**figure 4.5**). Avec cette méthode, 148 images pour chaque classe ont été sélectionnées.

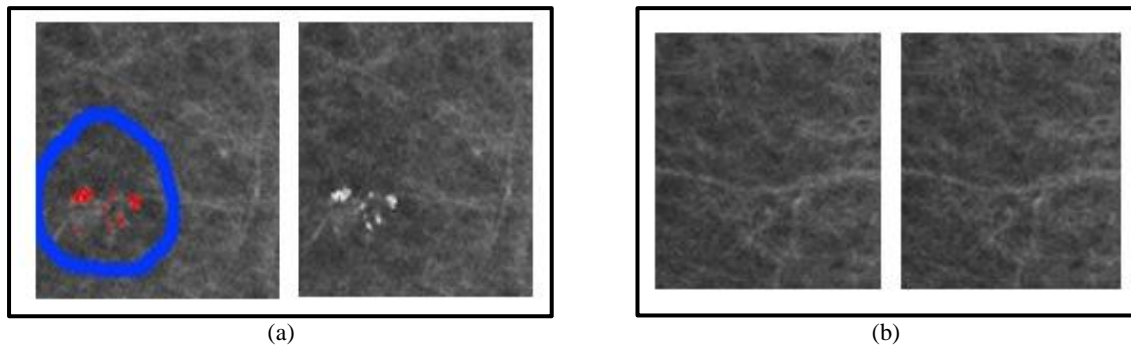


Figure 4.5: Exemples de classification des patches en classe Mcs et non Mcs. (a) Patch répertorié en classe Mcs, (b) patch répertorié en classe non Mcs.

Les deux classes sont considérées comme entrées de notre classificateur, pour un apprentissage en profondeur avec l'architecture du Resnet50. Nous avons divisé notre base de données en deux parties, 70% des images sont utilisées pour la phase d'apprentissage et 30% sont utilisées pour le test.

Evaluation de l'approche 1 proposée

Afin d'évaluer les performances de notre modèle de classification, on considère deux résultats pour montrer l'efficacité de notre classificateur.

- **Matrice de confusion**

La matrice de confusion est en quelque sorte un résumé des résultats et un outil pour évaluer les performances de type de classification. C'est un tableau à 4 valeurs de comparaison, entre les résultats prédits et les résultats réels.

Résultats et interprétation

La Figure 4.6 représente la matrice de confusion de chaque classe (classe Mcs et classe non Mcs) :

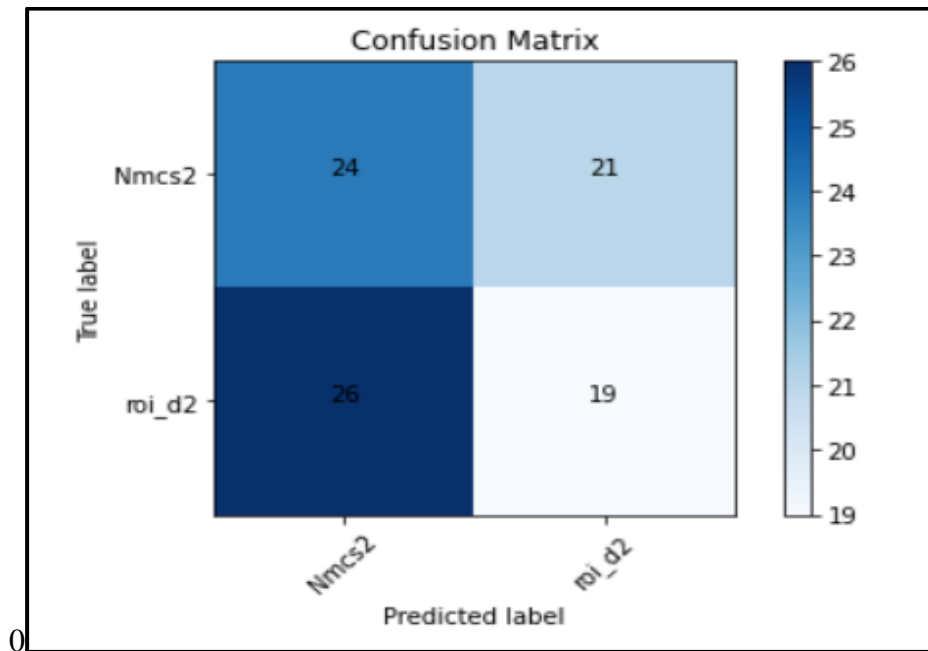


Figure 4.6 : Matrice de confusion.

- **Vrais positifs (VP)** : c'est le nombre de cas réellement positifs (Mcs), prédits par le modèle comme positifs. D'après l'exemple, sur les 90 images, 24 images sont correctement prédites (VP).
- **Vrais négatifs (VN)** : ce sont des cas non Mcs et classés comme « non-Mcs », donc le nombre de vrais négatifs est égal à 19.
- **Faux positifs (FP)** : ce sont les cas négatifs que le modèle a "faussement" prédit comme positifs, donc ce ne sont pas des Mcs que le modèle prédit comme Mcs, donc le nombre de faux positifs est égal à 21.
- **Faux négatifs (FN)** : ce sont les cas positifs que le modèle a "faussement" prédit comme négatifs, donc ce sont des Mcs que le modèle prédit comme Non Mcs, donc le nombre de faux négatifs est égal à 26.

D'après les résultats obtenus de la matrice de confusion, le modèle du ResNet50 s'est trompé sur la moitié des cas, pour cela le modèle obtenu n'est pas assez puissant pour pouvoir être utilisé dans la détection de la région d'intérêt. La principale cause est le manque de données pour l'apprentissage du modèle, car comme on le sait, les CNN ont besoin d'un grand nombre

de données, et notre base de données ne contient que 148 images de Mcs. La deuxième cause peut être due au fait que les Mcs sont très petites et peuvent être considérées comme du bruit.

- **Précision**

Elle représente la proportion de prédictions correctes parmi les points que l'on a prédits positifs [82]. Elle est donnée par l'équation (4.1).

$$P = \frac{Vp+Vn}{Vp+Vn+Fp+Fn} \quad (4.1)$$

Nous avons testé la précision de notre modèle selon le nombre d'époques utilisées dans l'apprentissage.

Epoques	Précision (Accuracy) %
7	68,22
17	75,55
21	82,44

Tableau 4-1 : Résultats de la précision en fonction du nombre d'époques.

D'après les résultats obtenus, nous remarquons que la précision s'augmente avec le nombre d'époques.

4.4.2. Résultats de détection automatique de la ROI par les outils morphologiques : Approche 2

- **Prétraitement de l'image**

Dans cette partie, notre approche a été mise en œuvre pour détecter la région d'intérêt en se basant sur l'utilisation des outils morphologiques. Tout d'abord, nous avons redimensionné la taille des images qui sont assez grandes (3328x4096), ensuite, on a constaté que le mamelon, l'aréole et la peau apparaissaient d'une manière très lumineuse (blanche) et leurs intensités se rapprochaient de celles des Mcs (**Figure 4.7**). Pour cette raison, nous avons réalisé un prétraitement pour tenter d'éliminer ces parties, qui représentent des régions non significatives, et leurs suppressions cibleront mieux la détection.

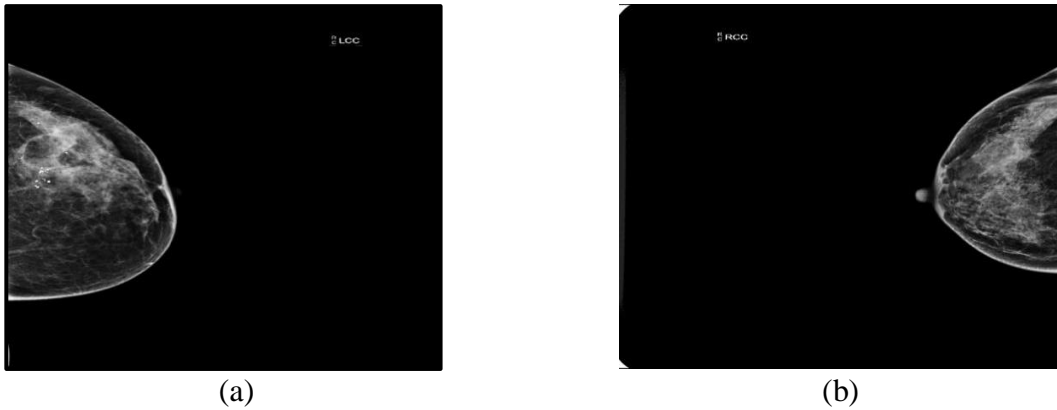


Figure 4.7 : Deux exemples d'images mammographiques.

Pour éliminer les parties indésirables sur notre image (mamelon, auréole et peau), nous avons utilisé l'érosion avec deux itérations et un élément structurant (ES) carré de taille 19x19 choisi par test.

Résultats et interprétation

La **figure 4.8** donne les résultats de la suppression des régions indésirables, obtenus par érosion morphologique.

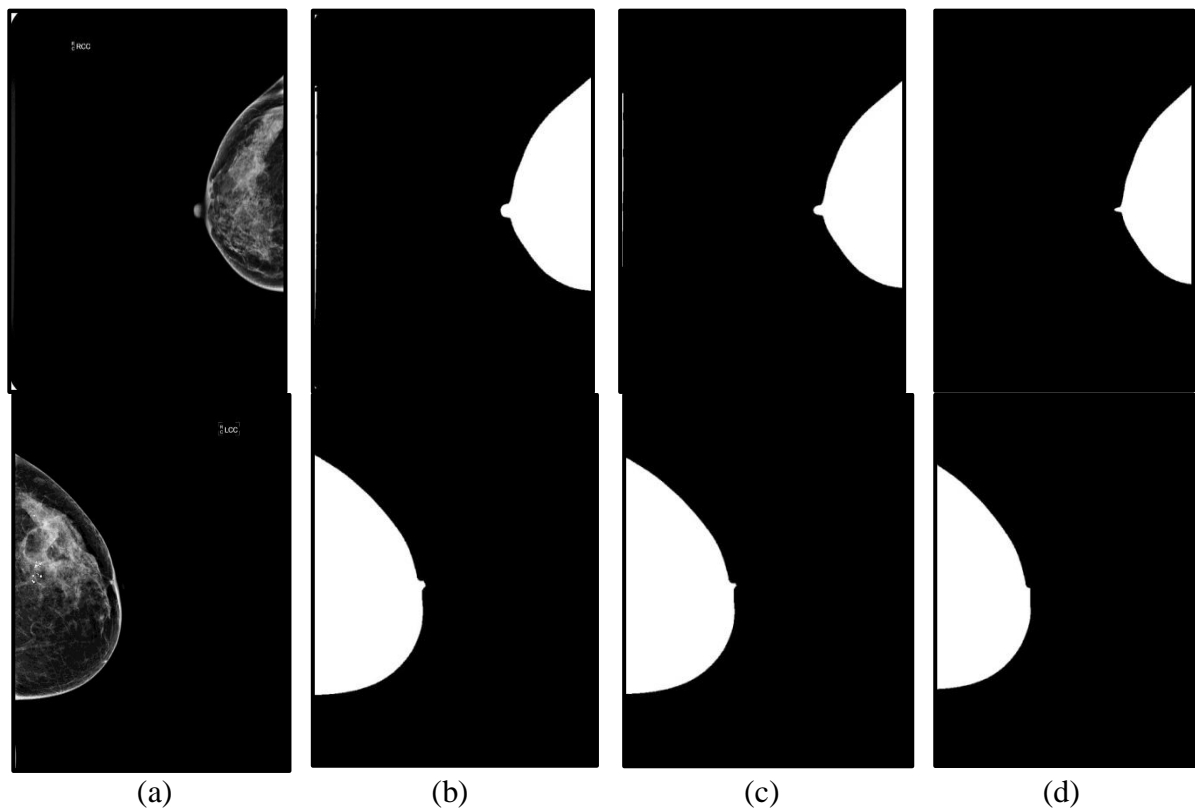


Figure 4.8: Masque du sein sans le mamelon, l'aréole et la peau. Colonne (a) image originale, colonne (b), (c), (d) image (a) érodée avec un élément structurant de taille (5x5), (9x9), et (19x19) respectivement.

Les résultats du masque du sein obtenu par érosion avec les éléments structurants carrés (5x5) et (9x9) ne sont pas suffisants, pour éliminer les parties indésirables sur notre image (mamelon, auréole et peau) ; pour cela on a augmenté l'élément structurant carré jusqu'à (19x19).

Après obtention du masque du sein sans la région du mamelon, de l'aréole et de la peau, nous effectuons une multiplication avec l'image originale pour obtenir la région à traiter (**Figure 4.9**). La multiplication consiste à faire coïncider géométriquement, l'image originale avec le résultat du masque obtenu.

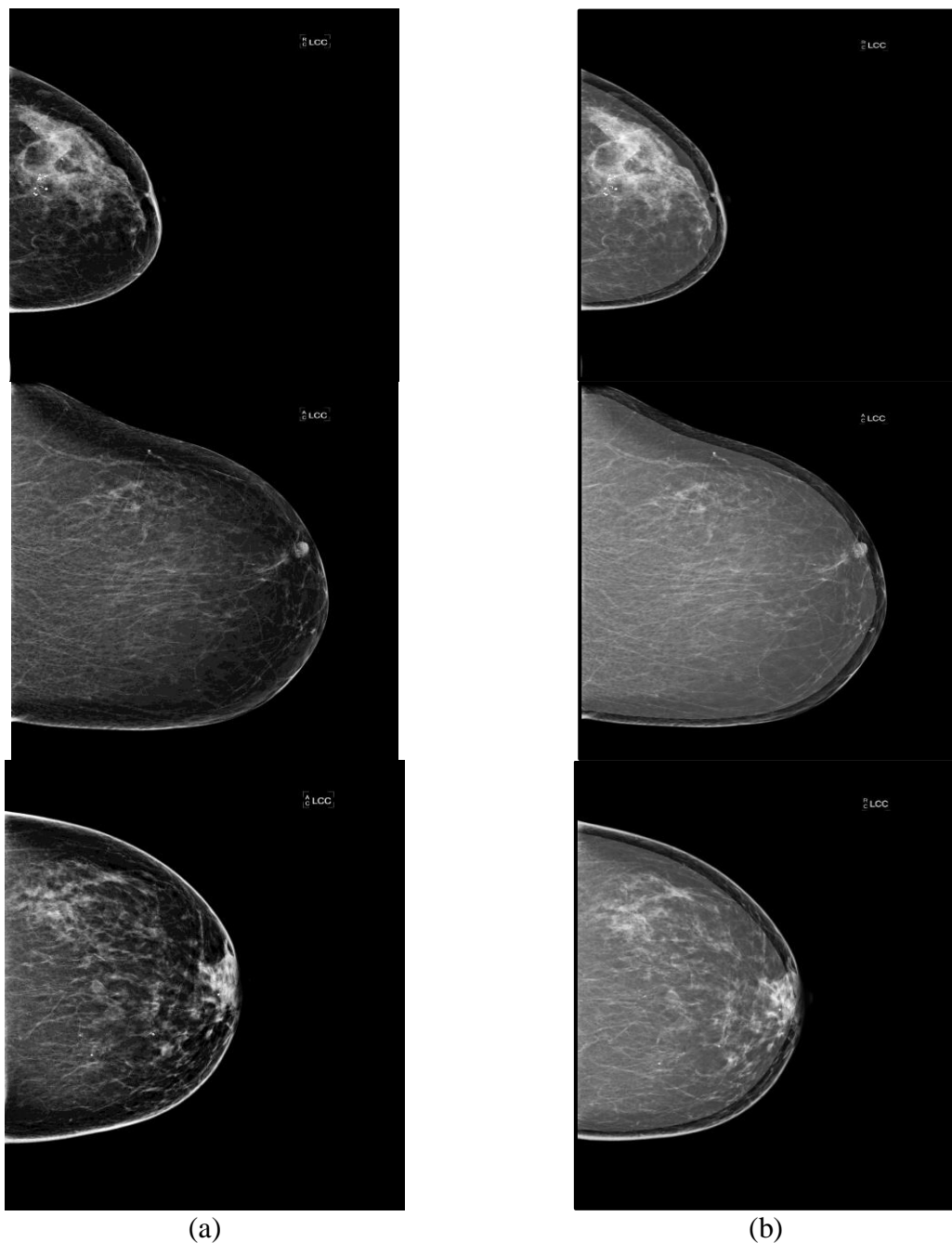


Figure 4.9: Résultats du prétraitement du sein. Colonne (a) image originale, colonne (b) image prétraitée.

- **Segmentation par le chapeau haut de forme (CHF)**

Après l'étape de prétraitement, on passe par la suite à la détection des petites structures qui représentent les Mcs. À l'aide de chapeau haut de forme avec un élément structurant (5x5) (choisi par le test), nous avons réussi à filtrer les petites structures de l'image (figures ci-dessous).

Résultats et interprétation

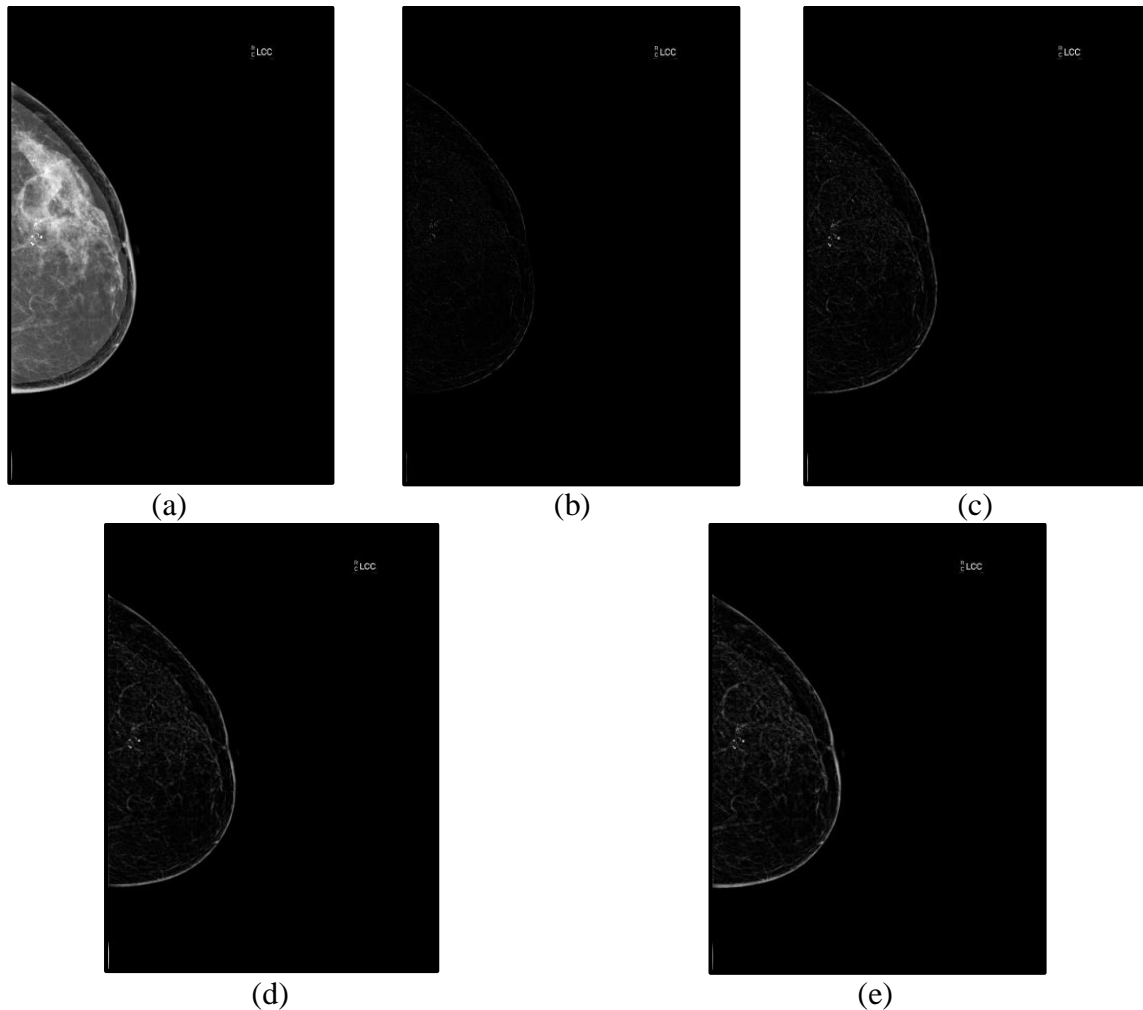


Figure 4.10: Résultats de l'application du chapeau haut de forme. (a) image prétraitée, (b), (c), (d) et (e) application du CHF avec élément structurant de taille (3x3), (5x5), (7x7) et (9x9) respectivement.

La **figure 4.10** représente les différents tests effectués pour trouver la taille adéquate de l'élément structurant. On remarque bien qu'avec la taille 3x3, les Mcs n'apparaissent pas bien, tandis qu'avec la taille 5x5, remarque qu'ils apparaissent mieux, mais avec l'augmentation à la taille 9x9, d'autres structures apparaissent aussi, comme les tissu conjonctifs et les fibres, pour cela, nous avons choisi la taille 5x5.

Une étape de seuillage a été effectuée sur les résultats du CHF, pour ne prendre que les Mcs qui représentent les particules les plus lumineuses du sein prétraité (**Figure 4.11**).

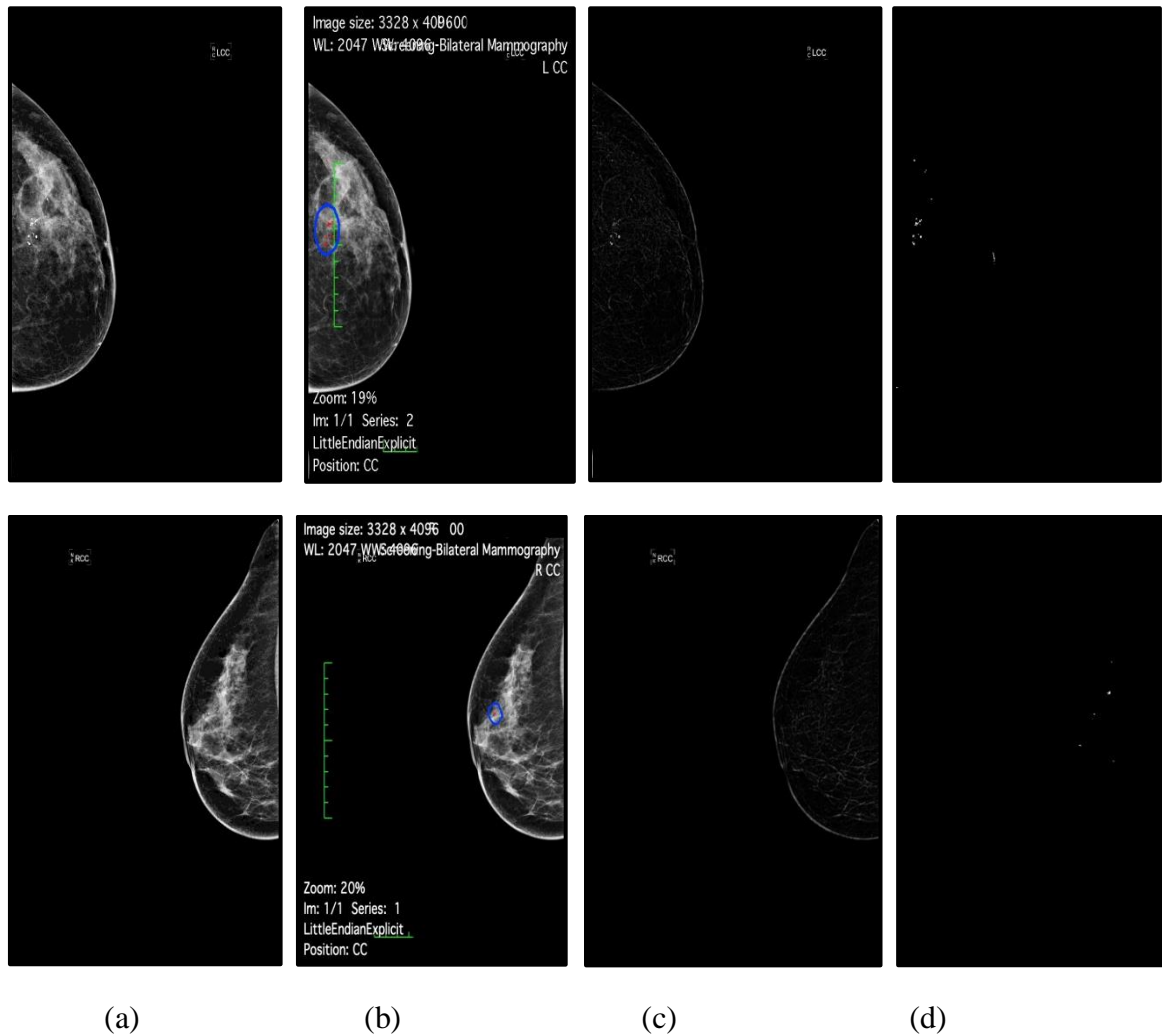


Figure 4.11: Résultats de l'étape de segmentation par CHF seuillée. Colonne (a) image originale, colonne (b) image vérité de terrain, colonne (b) image après application du CHF et colonne (c) image CHF seuillée.

On remarque la **figure 4.11**, qu'après le seuillage il reste encore d'autres faux positifs qui ne représente pas la région recherchée c'est pour cela qu'une autre étape est plus que nécessaire pour améliorer notre résultat de segmentation des Mcs. Nous avons donc ajouté une autre étape pour mieux cibler le foyer des Mc et aussi réduire la zone de traitement pour la segmentation.

- **Détection de la région d'intérêt**

Après l'étape de prétraitement et de segmentation des Mcs, vient l'étape de localisation du foyer des Mcs, et la suppression de tous les autres faux positifs qui résident de l'étape de segmentation précédente, on passe alors à une étape de sélection de la région qui nous intéresse et qui ne contient que le foyer des Mcs.

Cette approche consiste en une division de l'image en grille de taille (100x100), ensuite on ne sélectionne que le patch avec la plus grande surface (la surface qui représente le foyer des Mcs) (**Figure 4.12**).

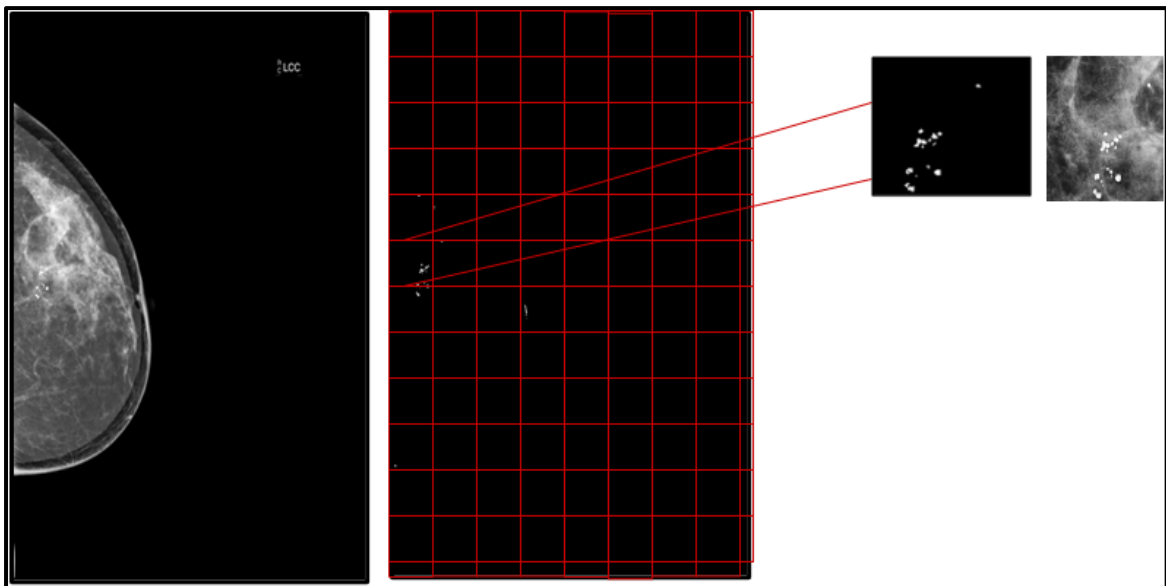


Figure 4.12: Exemple des étapes de détection de la ROI.

Après la sélection de patch avec la grande surface, il est nécessaire de reconstruire notre ROI par la multiplication de la région binaire avec son équivalent dans l'image originale. Cette étape a été effectuée par l'indexation de coordonnée de patch.

Résultats et interprétation

La **figure 4.13** donne le résultat de la détection du foyer des Mcs (ROI) pour deux images.

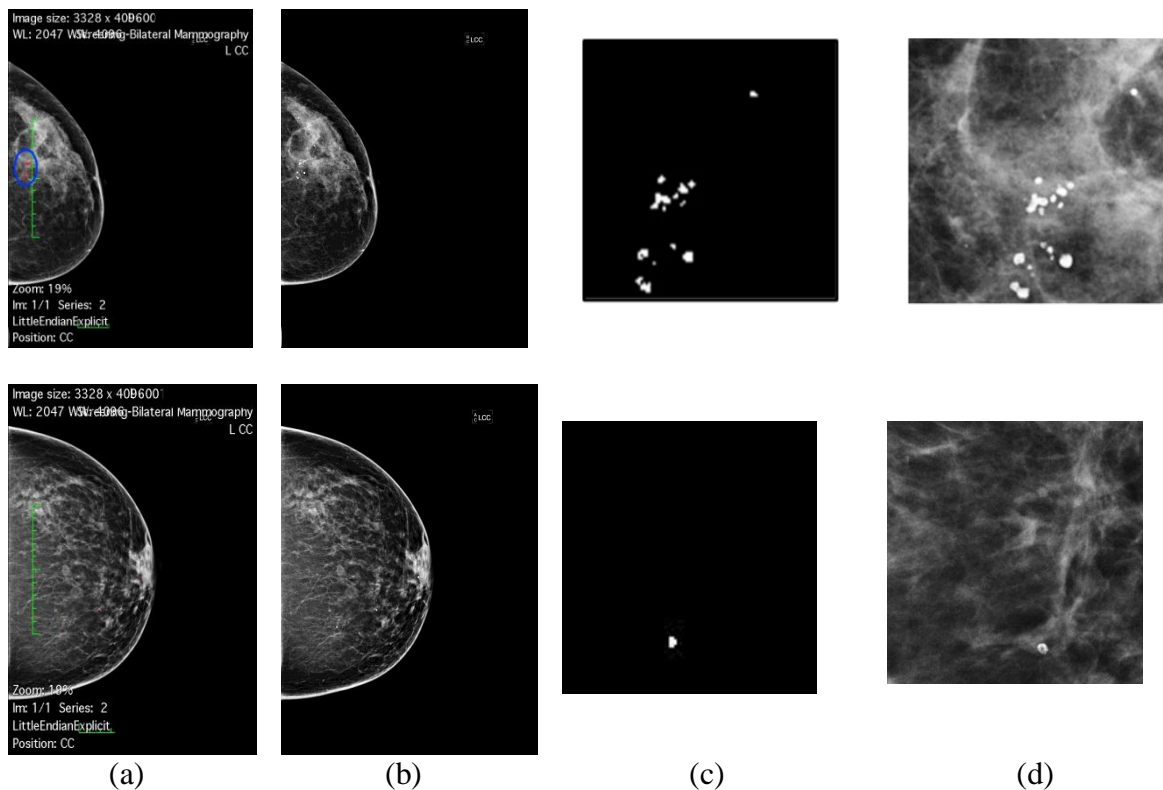


Figure 4.13: Résultats de l'étape de détection des Mcs. Colonne (a) image vérité de terrain, Colonne (b) image originale, colonne (c) masque du foyer des Mcs détectés et colonne (d) foyer des Mcs détectés en NG.

Les résultats obtenus sur la figure 3.14 présente les résultats de l'étape de détection du foyer des Mcs, on remarque bien que la forme des Mcs est déformée, et cela à cause de l'ouverture de CHF, néanmoins l'approche est efficace, pour la détection de l'emplacement du foyer. Cet inconvénient peut être corrigé par la deuxième étape de notre approche, qui consiste à la segmentation des Mcs pour bien affiner les contours des Mcs.

4.4.3. Résultats de la segmentation des microcalcifications

Après la détection de la région d'intérêt, vient l'étape de la segmentation des ROIs des microcalcifications pour affiner les contours des Mcs. Nous avons tout d'abord labélisé les différentes structures de la ROI par une labélisation avec l'algorithme du K-means, et ensuite une segmentation basée sur le CHF seuillé, et à la fin, nous avons fusionné entre les deux méthodes.

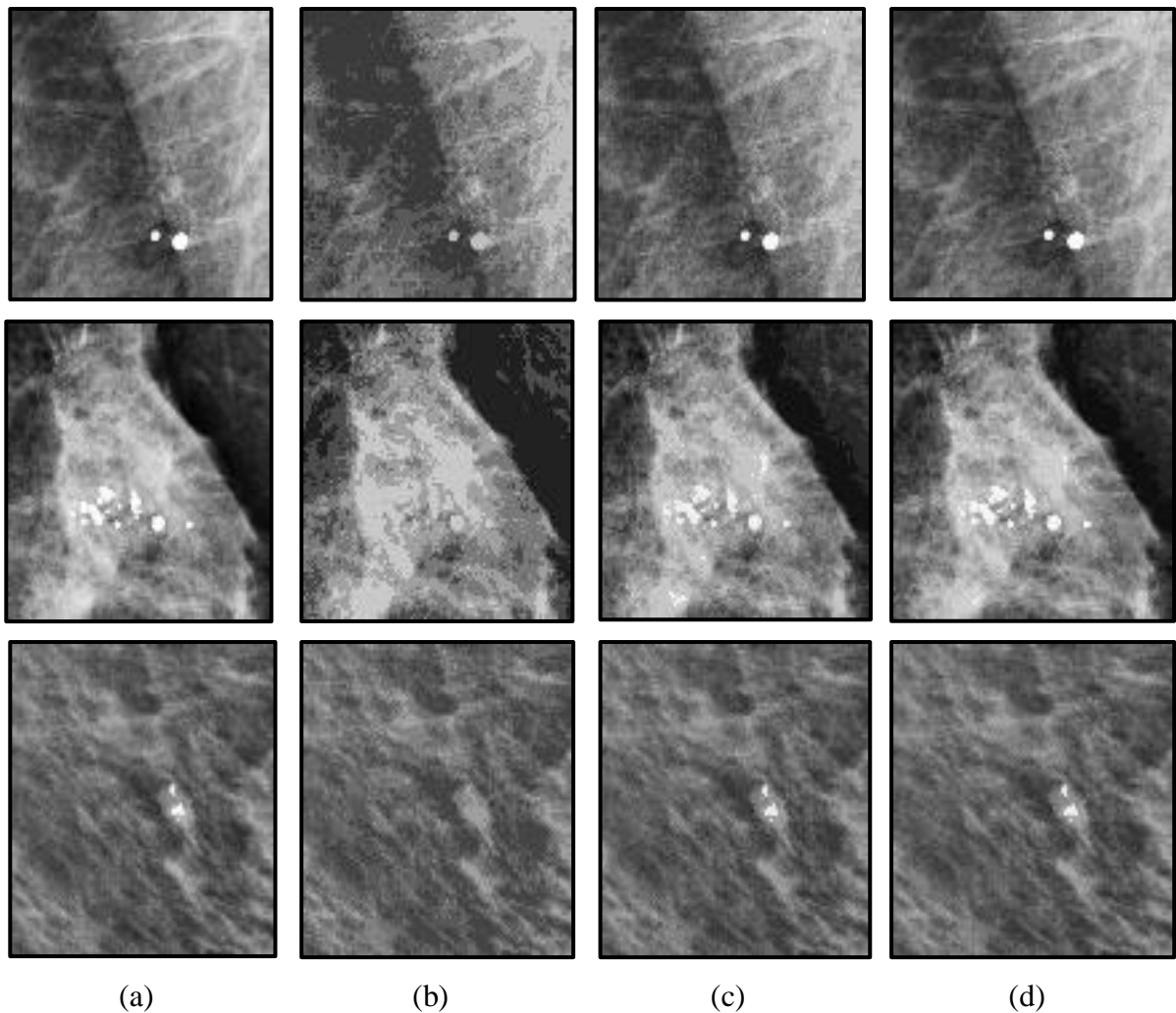
4.4.4. Résultats de la labélisation par K-means

La labélisation avec l'algorithme K-means nous a permis de classer les différentes structures que contient la région d'intérêt pour tenter de classer les Mcs. Nous avons testé plusieurs nombre k de classes (clustres) disjointes pour avoir un meilleur résultat, et pour différencier entre les tissus mammaires et les microcalcifications.

La détection des Mcs est très complexe à cause de leurs frontières mal définies, avec les cellules saines. Nous avons fait varier k entre 5 et 15.

Résultats et interprétation

La figure 4.14 donne les résultats de la détection des Mcs pour les valeurs $k = 5, 10$ et 15 respectivement.



**Figure 4.14: Résultats de la classification par K-means avec différents nombres de cluster k .
Colonne (a) Image originale, colonnes (b), (c) et (d) $k=5, k=10$ et $k=15$ respectivement.**

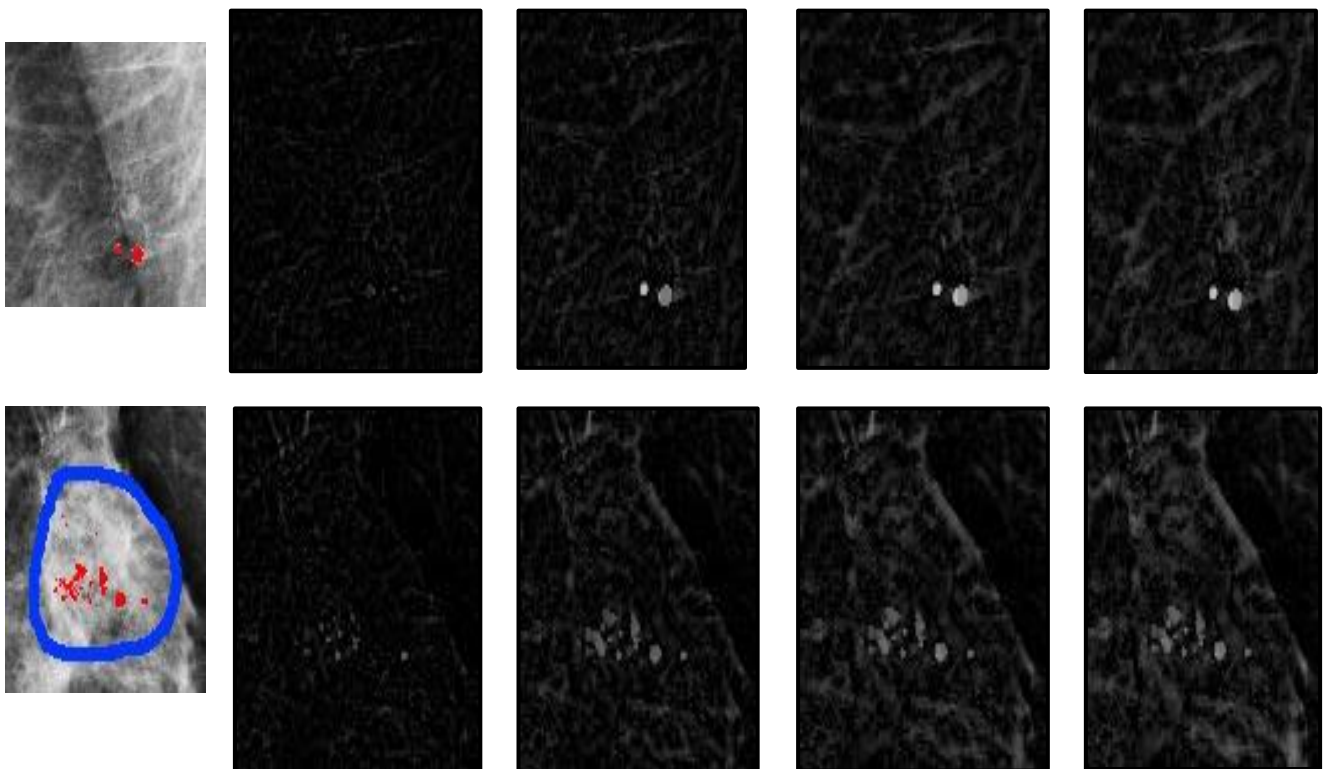
D'après les résultats obtenus dans la **figure 4.14** on a choisi de prendre le nombre de cluster k égal à 10, parce que la structure des tissus mammaires y apparait clairement et les microcalcifications ont une intensité élevée, cela est mieux qu'avec les clusters 5 et 15.

4.4.5. Résultats de la segmentation des microcalcifications par le CHF seuillé

La deuxième étape de notre approche consiste à appliquer un chapeau haut de forme seuillé (Top-Hat seuillé) sur la ROI. Du fait de la petite taille des Mcs, l'application de ce genre d'outils morphologiques permet de les détecter.

Résultats et interprétation

Nous avons appliqué le CHF avec des éléments structurants de différentes tailles pour choisir celui ou ceux qui donnent les meilleurs résultats, comme le montre la **figure 4.15**.



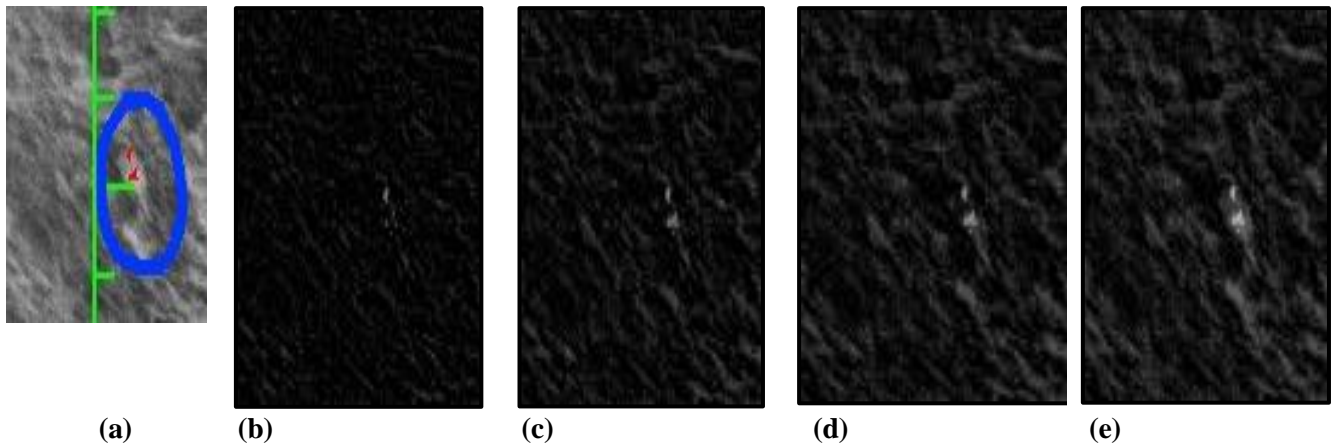


Figure 4.15:Résultats de la segmentation par CHF avec différents éléments structurants. Colonnes (a)image de vérité de terrain, (b), (c), (d) et (e) application du CHF avec élément structurant (3x3), (5x5), (7x7) et (9x9) respectivement.

Le choix de l'élément structurant est basé sur l'apparence des microcalcifications sur la ROI.

- Un élément structurant de taille (3x3) ne fait pas apparaître les microcalcifications
- Un élément structurant de taille (5x5) fait mieux apparaître les microcalcifications que l'élément structurant de taille (3x3), de plus les limites sont claires et proches de celles de la vérité terrain (GT)
- Les résultats du CHF avec un élément structurant (7x7) fait apparaître les microcalcifications dans la ROI, mais le fond (les autres tissus mammaires) est plus clair.
- Les résultats du CHF avec un élément structurant (9x9) fait apparaître les microcalcifications dans la ROI, mais leurs limites sont déformées par rapport à la vérité terrain (GT).

Après plusieurs tests, nous avons choisi le CHF avec un élément structurant (5x5), parce que les Mcs y apparaissent mieux, et parce qu'ils sont plus lumineux et le fond y apparaît plus sombre.

Cette méthode a été suivie par un seuillage par rapport à la valeur maximale de l'intensité des pixels, nous avons effectué un seuillage avec un seuil égale à la valeur maximale de l'intensité et qui est égale à 90 (choisit par test), qui permet de ne garder que les petites structures (résultats illustrés sur **la figure 4.16**).

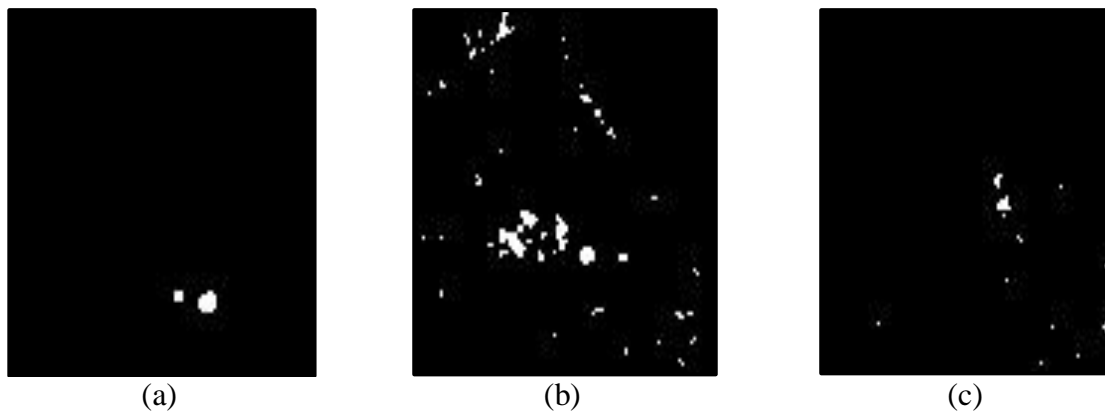


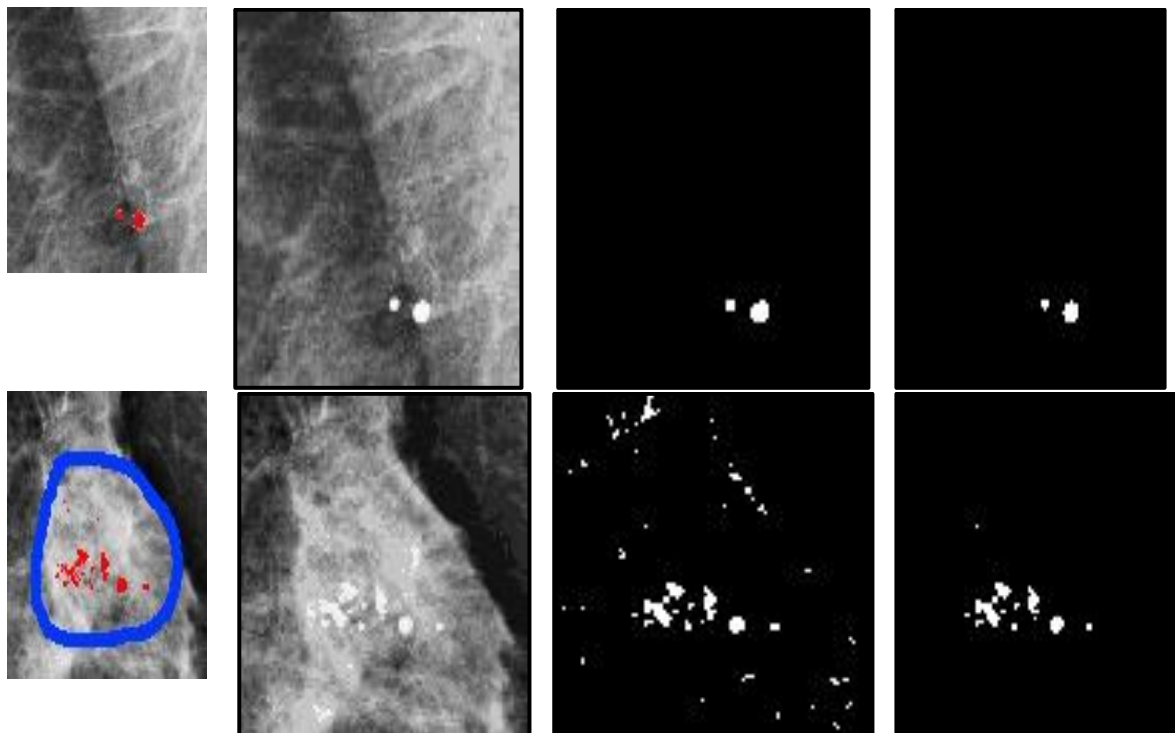
Figure 4.16: Exemple des différentes images seuillées.

4.4.6. Résultats de la segmentation des microcalcifications par la fusion

Nous avons fusionné les résultats obtenus par les deux approches appliquées sur la ROI : la labélisation K-means et le CHF seuillé. Cette étape consiste à fusionner pour chaque pixel l'information obtenue par chacune des modalités, pour améliorer les résultats de la segmentation effectuée par les deux approches utilisées.

Résultats et interprétation

La figure 4.17 représente les résultats de la fusion des images par l'intersection.



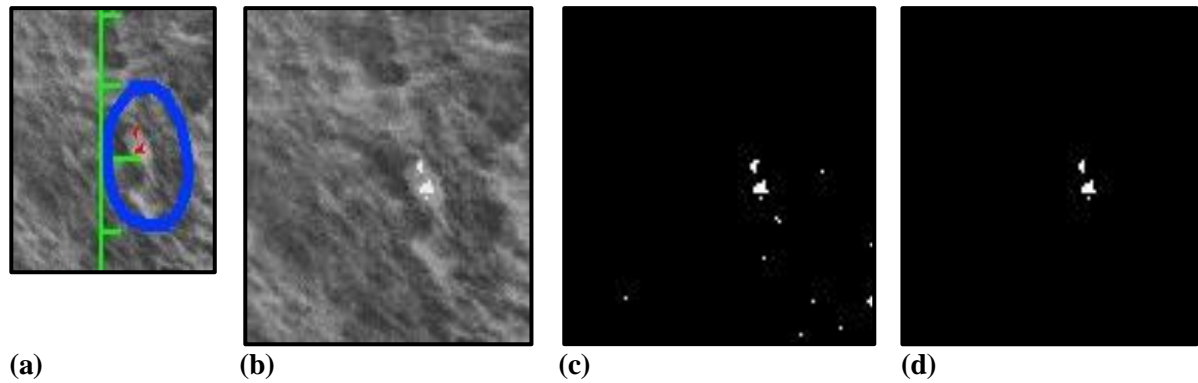


Figure 4.17: Résultats de l'étape de la fusion par l'union. La colonne (a) représente la vérité de terrain, La colonne (b) représente les résultats de la labélisation K-means, la colonne (c) représente les résultats du CHF seuillé, la colonne (d) représente les résultats de la fusion.

D'après les résultats de la **figure 4.17**, nous pouvons remarquer qu'avec la méthode K-means, la forme et la taille des microcalcifications sont bien restituées. Contrairement à la méthode du CHF, qui élimine les structures les plus petits que l'élément structurant et garde que les plus grands Mcs, à cause de l'ouverture effectuée dans la méthode, mais la fusion par l'intersection réussie à combler les inconvénients des deux méthodes et aboutie à des résultats proches de ceux des experts.

4.4.7. Résultats expérimentaux de la segmentation

Les mesures de la précision de la segmentation concernant la similitude entre le masque des microcalcifications détectées et le masque des microcalcifications noté GT (détectées par les experts). Les indices principaux sont estimés sur la base de l'intersection et de l'union de deux régions P et G, où P désigne une région de microcalcifications potentielle détectée (avec $|P|$ cardinal de P) et G désigne la vraie région de microcalcification dans le masque GT (avec $|G|$ cardinal de G) [83].

- **L'indice de similarité Dice (Dice similarity)**

L'indice de similarité (SI) connu aussi par Dice similarity indice quantifie le quotient de similarité entre le P est la région G, Il est défini par l'équation (4.2)

$$SI = \frac{2 * |P \cap G|}{|P| + |G|} \quad (4.2)$$

• Fraction de chevauchement (the overlap fraction)

L'OF (Overlap Fraction) quantifie la sous-segmentation de la région P segmentée de la région P segmentée candidate par rapport à la région G, elle se calcule comme suit :

$$OF = \frac{|P \cap G|}{|G|} \tag{4.3}$$

Le tableau ci-dessous représente quelques résultats des différentes étapes de notre approche de segmentation en termes de Dice et d'OF.

Images	Dice de K means	Dice de CHF	Dice de l'intersection	Dice de l'union	OF
Image 1	26,28%	71,64%	70,96%	68,57%	54,54%
Image 2	87,17%	66,66%	89,47%	68,57%	38,28%
Image 3	5,22%	41,02%	76,92%	27,34%	35,75%
Image 4	64,70%	66,66%	64,70%	66,66%	55,00%
Image 5	70,05%	80,00%	84,21%	80,00%	20,00%
Image 6	75,26%	63,89%	81,19%	63,44%	44,78%
Image 7	76,19%	68,42%	70,50%	70,83%	13,04%

Tableau 4-2: Résultats des différentes étapes de la méthode proposée, en termes de dice et d'OF.

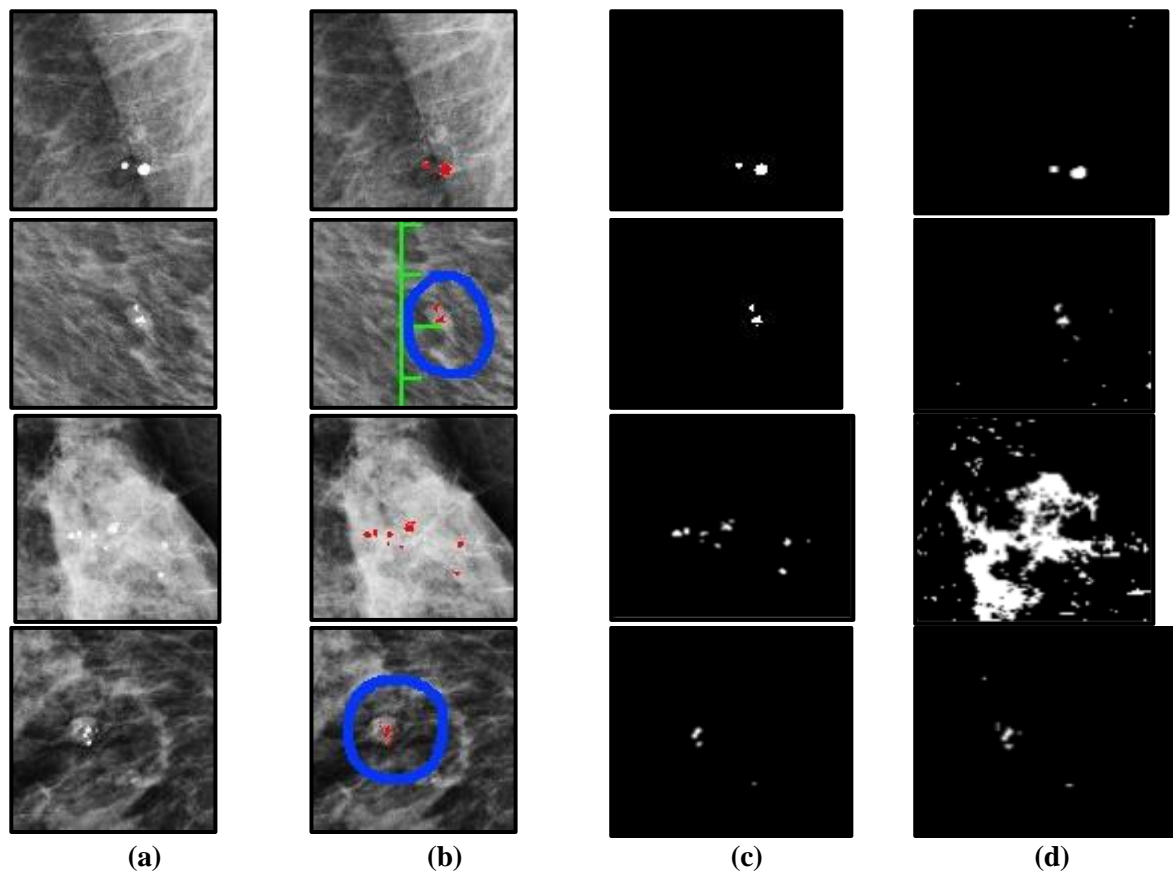


Figure 4.18: Résultats de la fusion. La colonne (a) représente l'image originale, colonne (b) image de vérité de terrain, colonne (c) résultats de la fusion par intersection, colonne (d) résultats de la fusion par union.

On remarque bien d'après les résultats présentés dans le tableau 4-2 et la figure 4.18, que les résultats de la méthode de fusion par intersection affichent de meilleurs résultats, et l'approche de fusion par intersection améliore nettement les résultats du K-means et du CHF utilisés séparément.

La figure 4.19 présente graphiquement quelques exemples du résultat final de notre approche de segmentation par la fusion. Les microcalcifications segmentées par notre approche de fusion par intersection y sont représentées en bleu, alors que les microcalcifications segmentées par les experts (GT) de la base de données y sont représentées en rouge.

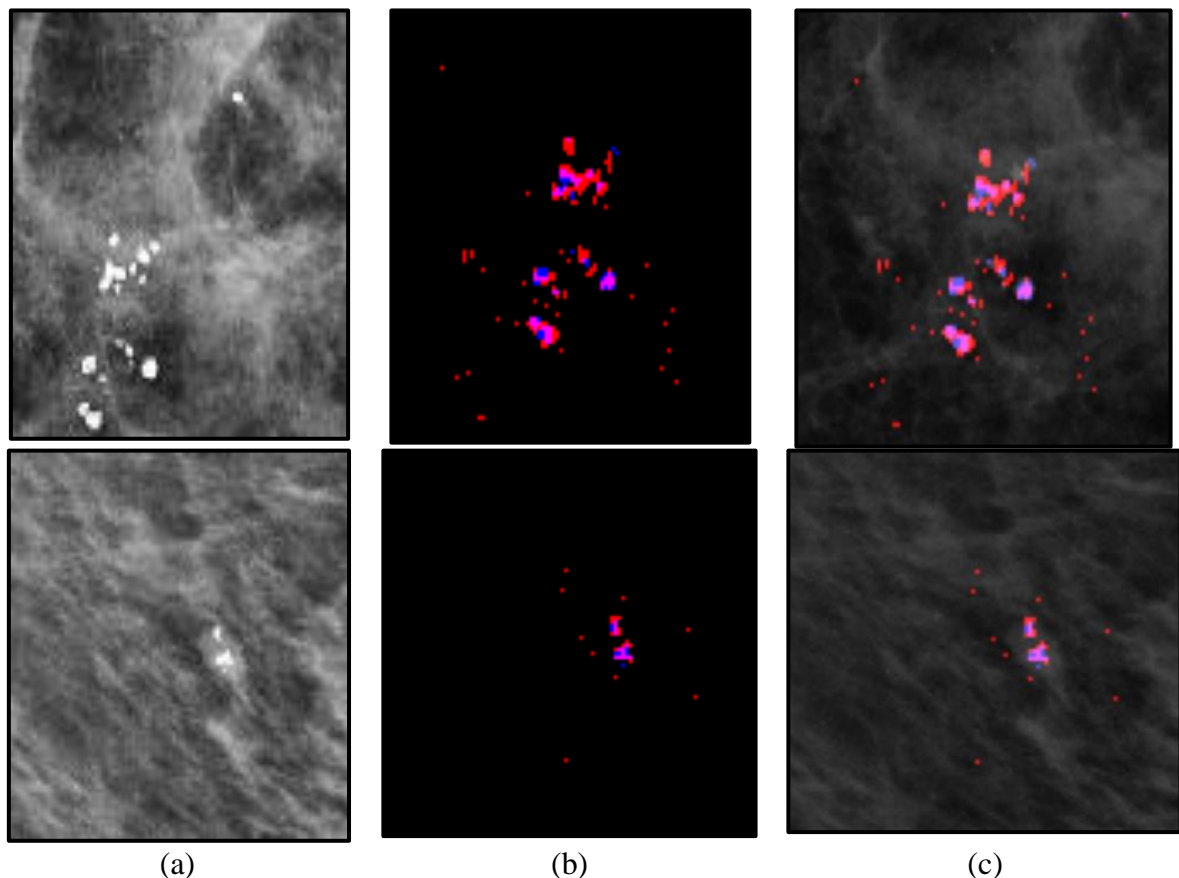


Figure 4.19: Résultats de l'approche proposée. La colonne (a) représente la ROI détectée automatiquement par CHF-, la colonne (b) représente la superposition des résultats de fusion avec GT, la colonne (c) représente la superposition des résultats de fusion avec la vérité terrain (GT) sur l'image en NG.

4.5. Présentation de l'interface graphique

Dans cette partie, nous présenterons notre interface graphique sous Python.

Une interface graphique ou GUI (Graphical User Interface) est un environnement d'affichage, basée sur l'emploi d'éléments graphiques tels que les fenêtres, les barres de menus, les boutons ...

Dans notre cas, nous créons une fenêtre graphique contenant l'application de l'approche, afin de faciliter à l'utilisateur la détection et la segmentation des Mcs.

Notre interface est mise en œuvre sous un système d'authentification (**Figure 4.20**).

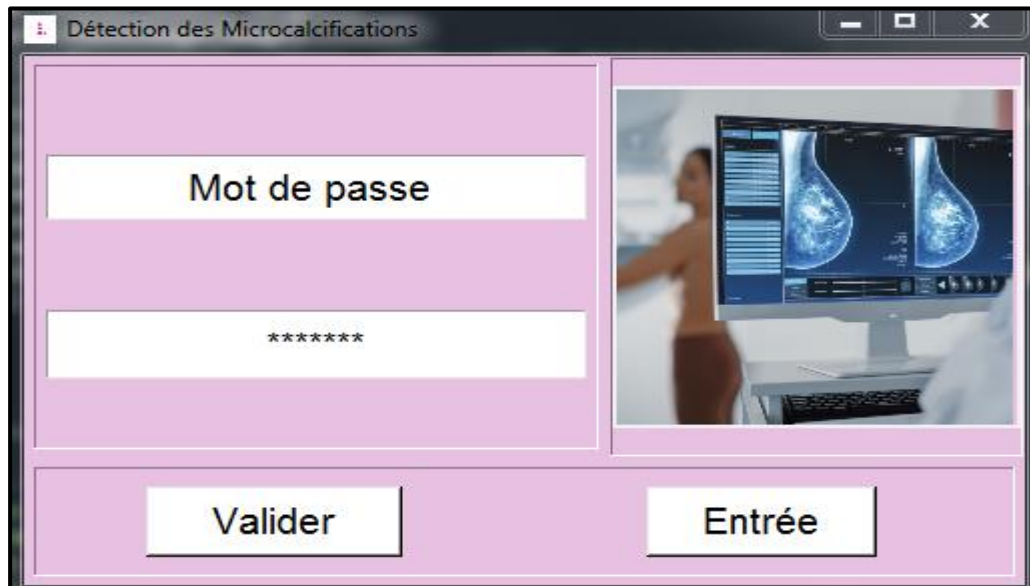


Figure 4.20: Fenêtre de système d'authentification.

Cette fenêtre se compose de :

- Un champ d'introduire du mot de passe.
- Deux boutons :
 - Un pour valider notre mot de passe.
 - L'autre pour accéder à l'interface principale.

Après la validation de notre mot de passe, une fenêtre qui représente notre interface graphique s'affiche sur l'écran (**Figure 4.21**).

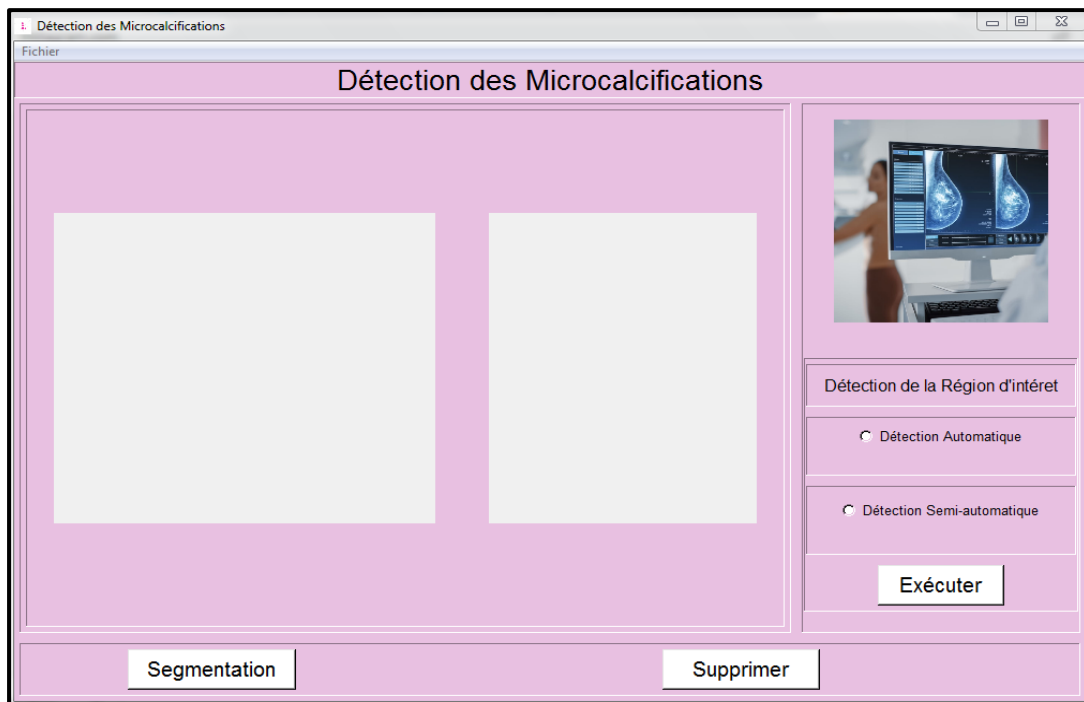


Figure 4.21: Fenêtre de l'interface.

Notre interface se compose de :

- Un Menu sur la barre d'outils nommé Fichier, qui contient un sous menu :
 - Ouvrir : pour lire l'image à segmenter
 - Quitter : pour fermer la fenêtre d'interface.
- Trois boutons :

Un bouton d'exécution pour la détection de la région d'intérêt avec deux approches :

- Approche automatique (développée par nous-même)
 - Approche semi-automatique, ou c'est à l'utilisateur de sélectionner la région suspectée.
- Un bouton de segmentation de l'image.
 - Un bouton pour supprimer les résultats précédents.
- Deux fenêtres d'affichages :
 - Une fenêtre pour afficher l'image originale.
 - Une autre pour afficher le résultat de segmentation de l'image.

4.6. Application de l'interface

- **Lecture de l'image**

La lecture de l'image se fait à l'aide du menu << Fichier >> et du bouton << Ouvrir >> (figure 4.22)

1. Cliquer sur le bouton Fichier.
2. Cliquer sur le bouton Ouvrir.

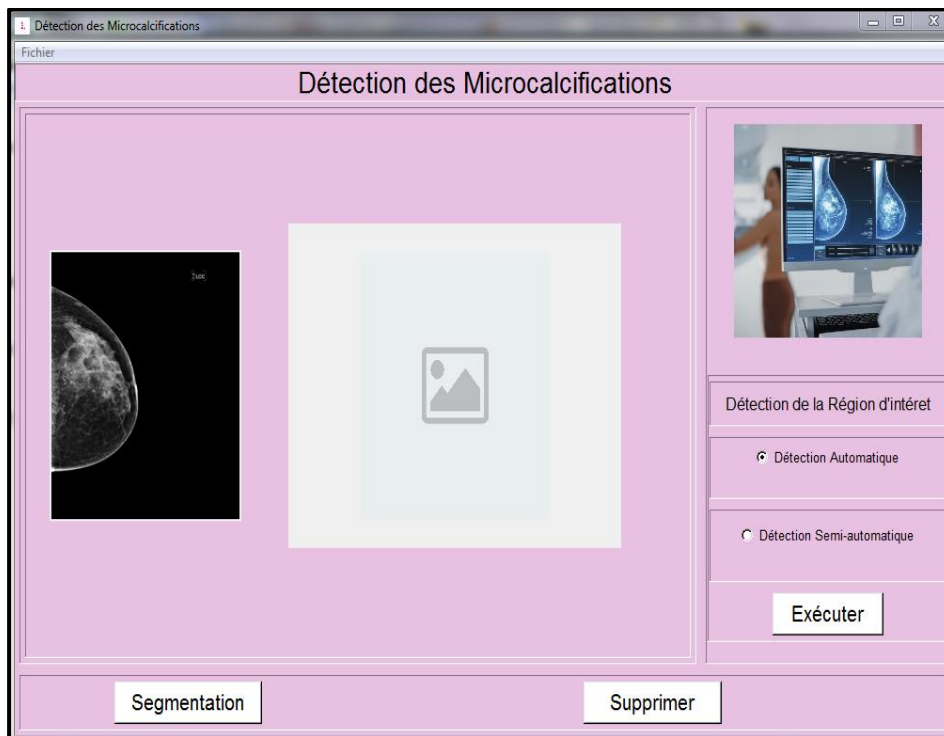


Figure 4.22: Lecture d'image.

- **Détection de la région d'intérêt**

Détection automatique

Les étapes à suivre pour la détection automatique de la région d'intérêt :

1. Sélectionner le choix << Détection Automatique >>
2. Cliquer sur le bouton << Exécuter >> le résultat obtenu sera affiché dans la fenêtre correspondante (Figure 4.23).

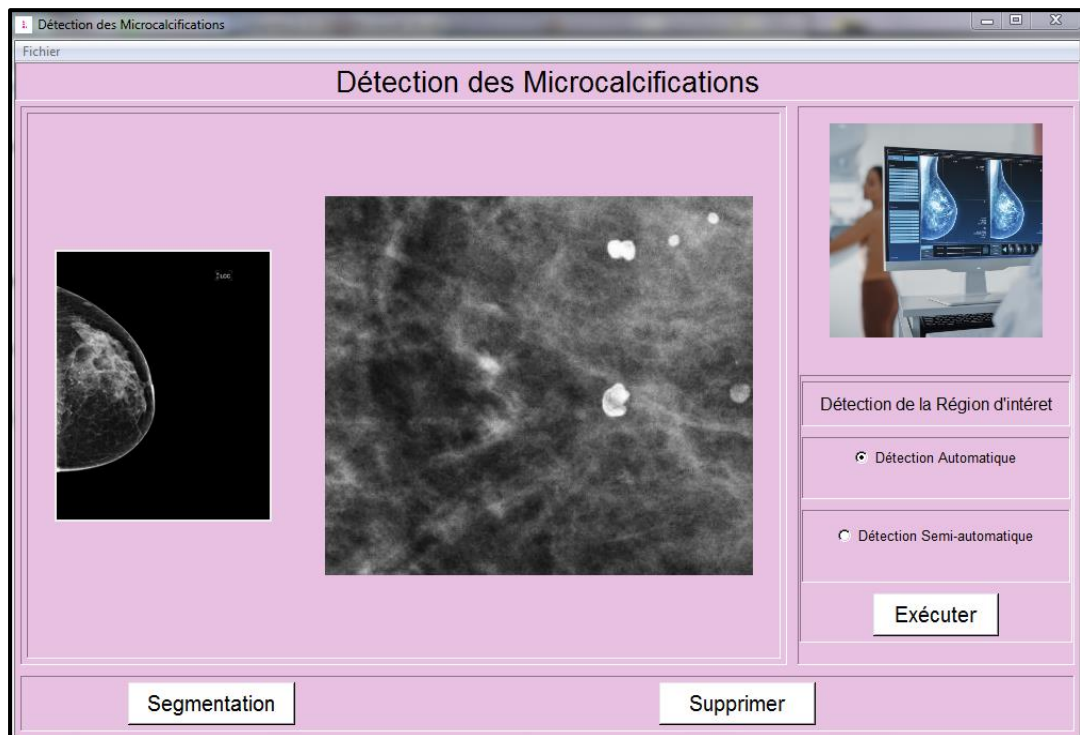


Figure 4.23 : Détection automatique de la région d'intérêt.

Détection semi-automatique

Les étapes à suivre pour la détection semi-automatique de la région d'intérêt sont :

1. Sélectionner le choix << Détection Semi-automatique >>.
2. Cliquer sur le bouton << Exécuter >>.
3. Une fenêtre s'affiche (**Figure 4.24**).
4. Cropper la région souhaitée avec la souris (Choisir la région suspecte sur l'image originale avec la souris)
5. le résultat obtenu sera affiché dans la fenêtre correspondante (**Figure 4.25**)



Figure 4.24: Image à cropper.

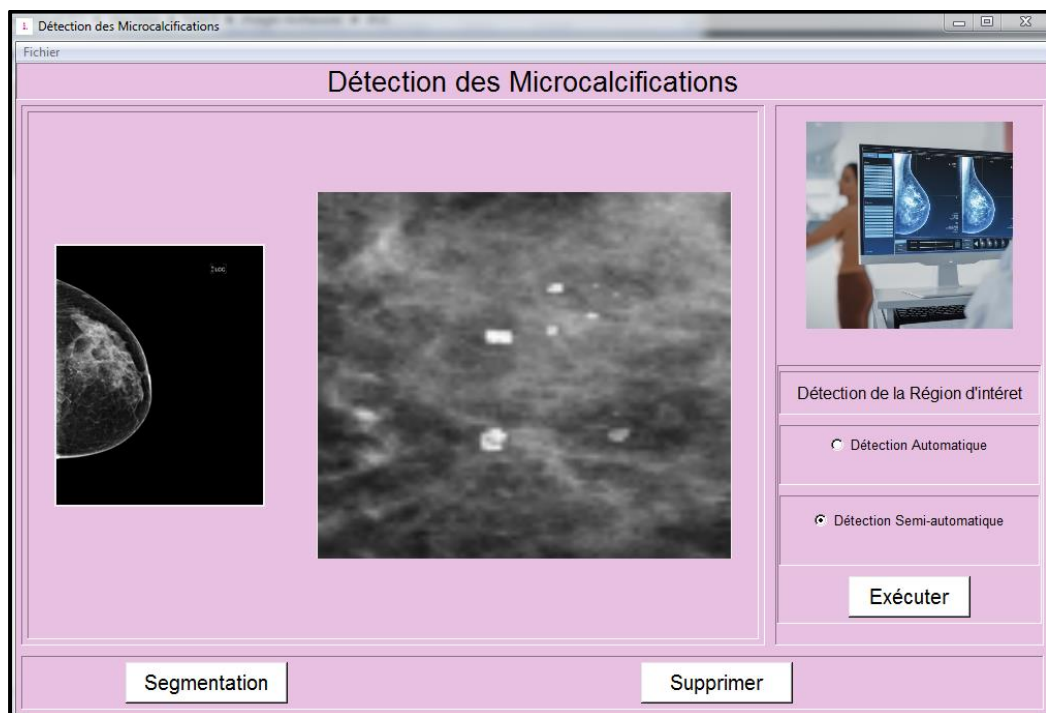


Figure 4.25 : Détection semi-automatique de la région d'intérêt.

• Segmentation des Mcs

Pour segmenter les microcalcifications, il faut :

1. Cliquer sur Segmentation.
2. Le résultat s'affiche sur la fenêtre correspondante (**Figure 4.26**).

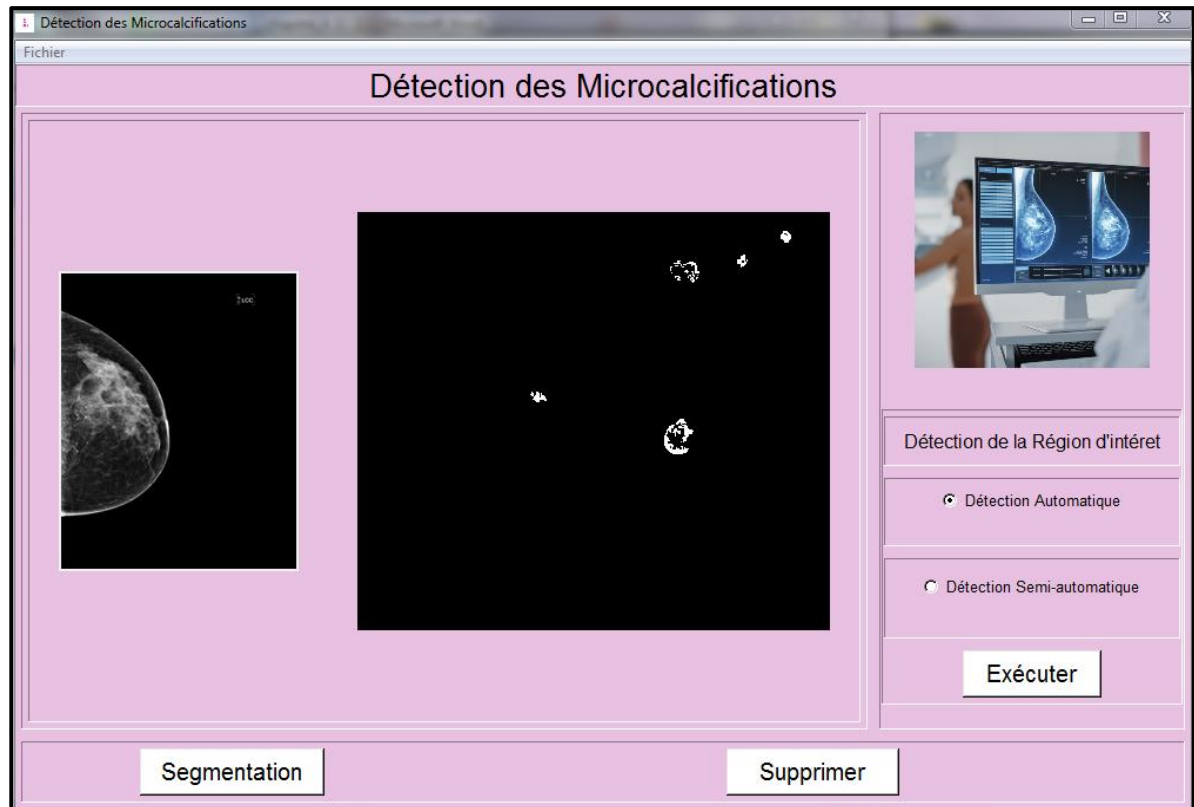


Figure 4.26 : Image segmentée.

A la fin de toute l'opération on supprime les résultats précédents pour refaire une autre segmentation.

1. Cliquer sur le bouton supprimer
2. Les résultats précédents seront supprimés

4.7. Discussion

Dans ce chapitre, nous avons présenté les différentes étapes ainsi que les résultats des deux approches de détection de la région d'intérêt et de l'étape de segmentation. Selon les résultats obtenus, nous avons choisi la détection de la ROI avec les outils morphologiques pour localiser la segmentation des Mcs.

Finalement, nous avons réalisé une interface graphique qui sert à faciliter l'utilisation de l'algorithme.

Conclusion Générale

Conclusion Générale

Dans ce travail de projet de fin d'études, nous avons tenté d'apporter un nouvel éclairage à l'analyse des images mammographiques pour une aide au diagnostic précoce. Toutefois, et au regard de la technologie avancée dans le secteur médical, la détection du cancer du sein reste un réel problème de santé publique, et un sujet de recherche d'actualité mais difficile.

Nous avons tenté de concevoir un système performant de segmentation d'images mammographiques, ayant pour objectif d'aider les radiologues à dépister le cancer du sein à un stade précoce. Ceci s'effectue à l'aide d'un processus de détection automatique de foyer de microcalcifications, qui présentent le plus souvent les signes primaires d'un cancer du sein. Nous avons essayé de concevoir un système d'analyse des clichés mammographiques, en vue de la détection des microcalcifications, et ainsi de mettre en évidence trois paramètres dont les experts ont besoin : la position, le nombre et la forme.

Dans ce contexte, nous avons exposé une bibliographie sur les travaux de recherche, en termes de la détection des Mcs. Par une analyse profonde de ces articles, nous avons déduit des outils morphologiques et l'apprentissage profond (ResNet 50) qui sont des outils mathématiques par excellence qui répondent à nos besoins, à mettre à la fin, et qui permettent de prendre en compte, efficacement, des notions de taille, de forme, de contraste et conserve l'information sur le contour, notion importante dans le diagnostic des lésions mammaires ce qui justifie notre choix.

Notre approche est constituée de deux étapes complémentaires, la première consiste à détecter la région d'intégrer (foyer des Mcs), pour cela, nous avons utilisé l'apprentissage profond avec l'architecture ResNet 50, mais malheureusement vue la petite taille des Mcs et leurs intensités proches de celles des tissus conjonctifs les entourant, le model du ResNet 50 a donné de mauvaise prédiction, et n'a pas su différencier les régions des Mcs des régions non Mcs. De ce fait, on s'est tournées sur d'autres outils mathématiques, et les méthodes dérivées de la morphologie mathématique, telles que le chapeau haut de forme.

L'application du chapeau haut de forme a donné de bons résultats. Seulement la forme des microcalcifications est légèrement différente de celle de l'image originale, cela revient aux étapes qu'il faut suivre pour obtenir le chapeau haut de forme (ouverture). En outre, cette approche représente un filtre spécialisé de détection des microcalcifications. La détection des régions d'intérêt présente des résultats très satisfaisants et encourageants, du fait qu'on a pu détecter les foyers de Mcs sur un grand nombre d'images de la base de données utilisées.

Conclusion Générale

La deuxième étape, qui est la segmentation des foyers de Mcs, permet de voir mieux et d'une manière plus précise le nombre et la forme des Mcs, car ces deux critères sont importants pour l'établissement d'un bon diagnostic. Nous utilisons une règle de fusion qui combine l'information au niveau pixel à partir des images segmentées et générées par l'algorithme du K-means et le top-hat, cette approche nous permet de mieux affiner la précision de la segmentation et de diminuer les faux positifs obtenus.

Finalelement, nous avons mis en œuvre un protocole de validation de nos résultats pour juger quantitativement leur qualité, nous avons pris en considération : l'indice de Dice, la sensibilité et la fraction de chevauchement. L'évaluation a été réalisée sur une base de données d'images mammographiques connues (BREAST MICRO-CALCIFICATION DATASET). Les résultats de la segmentation peuvent être améliorés en utilisant d'autres méthodes de fusion.

Perspectives

Notre travail de projet de fin d'études a été consacré à la détection des microcalcifications mammaires. Il serait, donc, intéressant de tester son efficacité sur la caractérisation afin de déterminer leur nature en termes de malignité ou bénignité. Nous proposons aussi d'effectuer une classification du Gal, par des méthodes de reconnaissance de formes ou d'autres outils d'intelligence Artificielle.

Références bibliographiques

Références bibliographiques

- [1] J. Brettes, C. Mathelin, B. Gairard, J. Bellocq. Cancer du sein. Paris : Elsevier Masson, 2007. 358 p. ISBN : 978-2-294-01813-8
- [2] J. Métivier. Radiothérapie externe du sein : amélioration de l'uniformité de la dose au volume cible. Mémoire présenté à l'Université Laval pour l'obtention du grade de Maître des Sciences (M.Sc.), Sep 2003.
- [3] l'article de comprendre le cancer du sein : l'INSTITUT NATIONAL DU CANCER, Disponible sur le net : www.e-cancer.fr, ISBN: 978-2-913 495-30-2, Validation : janvier 2007.
- [4] I. Hadjidj, Approche Morphologique pour la Segmentation d'Images Médicales, Mémoire présenté à l'Université de Tlemcen pour l'obtention de Diplôme de Magister en Électronique Biomédicale, 2011.
- [5] J. Morère, F. P. Plorca, M. Aapro, R. Salmon. Le cancer du sein. Paris, Springer, Oct 2007, 313 p, Oncologie pratique, ISBN : 978-2-287-36071-8.
- [6] « MX-300 - Mammographe analogique by Genoray MedicalExpo ». <https://www.medicaexpo.fr/prod/genoray/product-72338-674152.html> (consulté le 20 septembre 2022).
- [7] P. Haehnel. Mammographie- 83 exercices de radiodiagnostic pour étudiants et praticiens. Paris : Vigot, 1996. 137p. Exercices de radiodiagnostic. ISBN : 2-7114- 1049.
- [8] K. Chakib, Compression des images fixes par les approximations fractales basée, Mémoire de fin d'études, 1999.
- [9] « Mammographie 2D | Imagerie Duroc ». <https://www.imagerieduroc.com/examens/centre-specialise-en-senologie> (consulté le 20 septembre 2022).
- [10] J. Stines, « Les cancers du sein : les images visibles sur une seule incidence », *Journal de Radiologie*, vol. 85, n° 12, Part 2, p. 2083-2094, déc. 2004, doi: 10.1016/S0221-0363(04)97786-4.
- [11] A. LE TREUT, M. H. DILHUYDY "Mammographie, guide d'interprétation". Arnette (1989).
- [12] H. Chekkaf, I. Touil, Segmentation des masses dans les images mammographiques, Mémoire de fin d'études pour l'obtention du diplôme de Master en Informatique, 2012.

Références bibliographiques

- [13] I. Cheikhrouhou, Description et classification des masses mammaires pour le diagnostic du cancer du sein, Sciences pour l'ingénieur Traitement des images et du signal, Thèse, Juin 2012.
- [13] A. Le Treut, M. H. Dilhuydy, "Mammography : a guide to interpretation", Ed Mosby, 1991.
- [15] Comprendre le cancer du sein, Centre de Sénologie des Coteaux, <http://www.centre-des-coteaux.be/site.php?r=comprendre>, 2012.
- [16] « 1: Mammographie d'un sein normal en vue de profil. En comparant avec le... | Download Scientific Diagram ». https://www.researchgate.net/figure/Mammographie-dun-sein-normal-en-vue-de-profil-En-comparant-avec-le-schema-de-la_fig4_281600112 (consulté le 20 septembre 2022).
- [17] M. LE GAL, JC. DURANT, M. LAURENT, D. PELLIER "Valeur diagnostique des microcalcifications groupées découvertes par mammographie". Bull cancer, Paris 1984, 77 : 57-64.
- [18] A.LE TREUT, M.H.DILHUYDY "Mammographie, guide d'interprétation". Arnette (1989).
- [19] M. Aïcha, « Détection des microcalcifications par les modèles Markoviens pour des tissus mammaires », p. 139.
- [20] « Fibrose kystique du sein: symptômes et traitements | Creapharma ». <https://www.creapharma.ch/fibrose-kystique-sein.htm> (consulté le 15 mars 2022).
- [21] « Dépistage : une IA qui anticipe le cancer du sein ? » https://hitek.fr/actualite/depistage-ia-anticipe-cancer-sein_19578 (consulté le 10 juin 2022).
- [22] « Adénofibrome du sein - Docteur Benchimol: Gynécologue-obstétricien à Paris, France ». <https://www.docteur-benchimol.com/adenofibrome-du-sein.html> (consulté le 2 mars 2022).
- [23] M. LE GAL, JC. DURANT, M. LAURENT, D. PELLIER "Valeur diagnostique des microcalcifications groupées découvertes par mammographie". Bull cancer, Paris 1984, 77 : 57-64.
- [22] « Adénofibrome du sein - Docteur Benchimol: Gynécologue-obstétricien à Paris, France ». <https://www.docteur-benchimol.com/adenofibrome-du-sein.html> (consulté le 2 mars 2022).
- [24] A. Boucher et N. Vincent, « Segmentation du muscle pectoral sur une mammographie », Article de revue p. 9.

Références bibliographiques

- [25] T. Balakumaran et I. Vennila, « Detection of microcalcification clusters in digital mammograms using Multiresolution based foveal algorithm », Article de colloque in *2011 World Congress on Information and Communication Technologies*, Mumbai, India, déc. 2011, p. 657-660. doi: 10.1109/WICT.2011.6141323.
- [26] Y. Kabbadj, F. Regragui, et M. M. Himmi, « Microcalcification detection using a fuzzy inference system and support vector machines », Article de colloque in *2012 International Conference on Multimedia Computing and Systems*, Tangiers, Morocco, mai 2012, p. 312-315. doi: 10.1109/ICMCS.2012.6320216.
- [27] S. H. Lewis et Aijuan Dong, « Detection of breast tumor candidates using marker-controlled watershed segmentation and morphological analysis », Article de colloque in *2012 IEEE Southwest Symposium on Image Analysis and Interpretation*, Santa Fe, NM, avr. 2012, p. 1-4. doi: 10.1109/SSIAI.2012.6202438.
- [28] Y. Zhang, Y. Lan, et H. Ren, « Image Enhancement and Its Effects on Segmentation for Mammographic Masses », Article de colloque in *2012 Fifth International Symposium on Computational Intelligence and Design*, Hangzhou, China, oct. 2012, p. 423-426. doi: 10.1109/ISCID.2012.112.
- [29] P. Suapang, C. Naruephai, M. Thongyoun, et S. Chivaprecha, « Mammographic masses segmentation based on morphology », Article de colloque in *The 5th 2012 Biomedical Engineering International Conference*, Muang, Ubon Ratchathani, Thailand, déc. 2012, p. 1-4. doi: 10.1109/BMEiCon.2012.6465503.
- [30] E. Malar, A. Kandaswamy, D. Chakravarthy, et A. Giri Dharan, « A novel approach for detection and classification of mammographic microcalcifications using wavelet analysis and extreme learning machine », *Computers in Biology and Medicine*, Article de revue vol. 42, n° 9, p. 898-905, sept. 2012, doi: 10.1016/j.compbiomed.2012.07.001.
- [31] M. Mohamed Fathima, D. Manimegalai, et S. Thaiyalnayaki, « Automatic detection of tumor subtype in mammograms based On GLCM and DWT features using SVM », Article de colloque in *2013 International Conference on Information Communication and Embedded Systems (ICICES)*, Chennai, févr. 2013, p. 809-813. doi: 10.1109/ICICES.2013.6508213.

Références bibliographiques

- [32] S. K. Wajid, A. Hussain, et B. Luo, « An efficient Computer Aided Decision Support System for breast cancer diagnosis using Echo State Network classifier », Article de colloque in *2014 IEEE Symposium on Computational Intelligence in Healthcare and e-health (CICARE)*, Orlando, FL, USA, déc. 2014, p. 17-24. doi: 10.1109/CICARE.2014.7007829.
- [33] M. Zouari, A. Dammak Masmoudi, et D. S. Masmoudi, « A non linear stretching image enhancement technique for microcalcification detection », Article de colloque in *2014 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, mars 2014, p. 193-197. doi: 10.1109/ATSIP.2014.6834605.
- [34] M. Mustra et M. Grgic, « Detection of areas containing microcalcifications in digital mammograms », Article de colloque in *IWSSIP 2014 Proceedings*, mai 2014, p. 51-54.
- [35] O. Intharasombat et P. Piamsa-nga, « Adaptive window and adaptive threshold method for microcalcification detection », Article de colloque in *2014 International Computer Science and Engineering Conference (ICSEC)*, Khon Kaen, Thailand, juill. 2014, p. 446-451. doi: 10.1109/ICSEC.2014.6978238.
- [36] M. I. Chacon-Murguia et A. J. Villalobos-Montiel, « A fuzzy computer aided diagnosis system using breast thermography », Article de colloque in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, Brooklyn, NY, USA, avr. 2015, p. 105-108. doi: 10.1109/ISBI.2015.7163827.
- [37] I. I. Esener, S. Ergin, et T. Yuksel, « A new ensemble of features for breast cancer diagnosis », Article de colloque in *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, mai 2015, p. 1168-1173. doi: 10.1109/MIPRO.2015.7160452.
- [38] Hai Su *et al.*, « Robust automatic breast cancer staging using a combination of functional genomics and image-omics », Article de colloque in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, août 2015, p. 7226-7229. doi: 10.1109/EMBC.2015.7320059.
- [39] M. Tan, W. Qian, J. Pu, H. Liu, et B. Zheng, « A new approach to develop computer-aided detection schemes of digital mammograms », Article de revue *Phys Med Biol*, vol. 60, n° 11, p. 4413-4427, juin 2015, doi: 10.1088/0031-9155/60/11/4413.

Références bibliographiques

- [40] D.-R. Chen, C.-L. Chien, et Y.-F. Kuo, « Computer-Aided Assessment of Tumor Grade for Breast Cancer in Ultrasound Images », Article de revue *Computational and Mathematical Methods in Medicine*, vol. 2015, p. 1-6, 2015, doi: 10.1155/2015/914091.
- [41] A. F. Khalaf et I. A. Yassine, « Novel features for microcalcification detection in digital mammogram images based on wavelet and statistical analysis », Article de colloque in *2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, QC, Canada, sept. 2015, p. 1825-1829. doi: 10.1109/ICIP.2015.7351116.
- [42] M. V. Sainz de Cea et Y. Yang, « Case-adaptive decision rule for detection of clustered microcalcifications in mammograms », Article de colloque in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, New York, NY, avr. 2015, p. 1147-1150. doi: 10.1109/ISBI.2015.7164075.
- [43] Q. Zhang *et al.*, « Deep learning based classification of breast tumors with shear-wave elastography », Article de revue *Ultrasonics*, vol. 72, p. 150-157, déc. 2016, doi: 10.1016/j.ultras.2016.08.004.
- [44] D. Cascio, V. Taormina, L. Abbene, et G. Raso, « A Microcalcification Detection System in Mammograms based on ANN Clustering », Article de colloque in *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, Sydney, Australia, nov. 2018, p. 1-4. doi: 10.1109/NSSMIC.2018.8824729.
- [45] V. A. Karale *et al.*, « A Screening CAD Tool for the Detection of Microcalcification Clusters in Mammograms », Article de revue *J Digit Imaging*, vol. 32, n° 5, p. 728-745, oct. 2019, doi: 10.1007/s10278-019-00249-5.
- [46] T. M. A. Basile *et al.*, « Microcalcification detection in full-field digital mammograms: A fully automated computer-aided system », Article de revue *Physica Medica*, vol. 64, p. 1-9, août 2019, doi: 10.1016/j.ejmp.2019.05.022.
- [47] A. S. B. Azam, A. A. Malek, A. S. Ramlee, N. D. S. M. Suhaimi, et N. Mohamed, « Segmentation of Breast Microcalcification Using Hybrid Method of Canny Algorithm with Otsu Thresholding and 2D Wavelet Transform », in *2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Penang, Malaysia, août 2020, p. 91-96. doi: 10.1109/ICCSCE50387.2020.9204950.

Références bibliographiques

- [48] A. Touil, K. Kalti, P.-H. Conze, B. Solaiman, et M. A. Mahjoub, « Automatic detection of microcalcification based on morphological operations and structural similarity indices », *Biocybernetics and Biomedical Engineering*, vol. 40, n° 3, p. 1155-1173, juill. 2020, doi: 10.1016/j.bbe.2020.05.002.
- [49] B. Mazumder, S. M. T. Islam, M. M. Rahman, et M. Nurullah, « Stationary Wavelet Based Energy Feature Extraction for Detection and Classification of Mammographic Microcalcifications », in *2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, Dhaka, Bangladesh, déc. 2020, p. 1-5. doi: 10.1109/STI50764.2020.9350417.
- [50] A. A.-M. Bulbul, H. Rahaman, S. Biswas, M. B. Hossain, et A.-A. Nahid, « Design and numerical analysis of a PCF-based bio-sensor for breast cancer cell detection in the THz regime », *Sensing and Bio-Sensing Research*, vol. 30, p. 100388, déc. 2020, doi: 10.1016/j.sbsr.2020.100388.
- [51] M. Desai et M. Shah, « An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN) », *Clinical eHealth*, vol. 4, p. 1-11, 2021, doi: 10.1016/j.ceh.2020.11.002.
- [52] R. Murakami, N. Uchiyama, H. Tani, T. Yoshida, et S. Kumita, « Comparative analysis between synthetic mammography reconstructed from digital breast tomosynthesis and full-field digital mammography for breast cancer detection and visibility », *European Journal of Radiology Open*, vol. 7, p. 100207, 2020, doi: 10.1016/j.ejro.2019.12.001.
- [53] J. Bai, R. Posner, T. Wang, C. Yang, et S. Nabavi, « Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review », *Medical Image Analysis*, vol. 71, p. 102049, juill. 2021, doi: 10.1016/j.media.2021.102049.
- [54] V. Gupta, M. Vasudev, A. Doegar, et N. Sambyal, « Breast cancer detection from histopathology images using modified residual neural networks », *Biocybernetics and Biomedical Engineering*, vol. 41, n° 4, p. 1272-1287, oct. 2021, doi: 10.1016/j.bbe.2021.08.011.
- [55] H. Cao, S. Pu, W. Tan, et J. Tong, « Breast mass detection in digital mammography based on anchor-free architecture », *Computer Methods and Programs in Biomedicine*, vol. 205, p. 106033, juin 2021, doi: 10.1016/j.cmpb.2021.106033.

Références bibliographiques

- [56] J. An, H. Yu, R. Bai, J. Li, Y. Wang, et R. Cao, « Detection and Segmentation of Breast Masses Based on Multi-Layer Feature Fusion », *Methods*, vol. 202, p. 54-61, juin 2022, doi: 10.1016/j.ymeth.2021.04.022.
- [57] R. Sánchez-Cauce, J. Pérez-Martín, et M. Luque, « Multi-input convolutional neural network for breast cancer detection using thermal images and clinical data », *Computer Methods and Programs in Biomedicine*, vol. 204, p. 106045, juin 2021, doi: 10.1016/j.cmpb.2021.106045.
- [58] Y. Zhang *et al.*, « Automatic Detection and Segmentation of Breast Cancer on MRI Using Mask R-CNN Trained on Non-Fat-Sat Images and Tested on Fat-Sat Images », *Academic Radiology*, vol. 29, p. S135-S144, janv. 2022, doi: 10.1016/j.acra.2020.12.001.
- [59] M. Kasdi, « L'apprentissage profond (Deep Learning) pour la classification et la recherche d'imagerie par le contenu », Mémoire de Master en Informatique, Université de Ouargla, Algérie, 2017
- [60] « Deep Learning : définition, applications, avantages et inconvénients », *Retengr*, 22 janvier 2021. <https://www.retengr.com/2021/01/22/deep-learning-definitions-applications-avantages-inconvenients/> (consulté le 20 septembre 2022).
- [61] « Appréhendez le Deep Learning ou l'apprentissage profond », *OpenClassrooms*. <https://openclassrooms.com/fr/courses/6417031-objectif-ia-initiez-vous-a-lintelligence-artificielle/6823506-apprenez-le-deep-learning-ou-lapprentissage-profond> (consulté le 20 septembre 2022).
- [62] « CS 230 - Pense-bête de réseaux de neurones convolutionnels ». <https://stanford.edu/~shervine/l/fr/teaching/cs-230/pense-bete-reseaux-neurones-convolutionnels> (consulté le 20 septembre 2022).
- [63] Y. S. Obam, « Comprendre les Réseaux de Neurones Convolutifs (CNN) », *Medium*, 2 septembre 2020. <https://yannicksergeobam.medium.com/comprendre-les-r%C3%A9seaux-de-neurones-convolutifs-cnn-d5f14d963714> (consulté le 20 septembre 2022).
- [64] T. rédac, « Convolutional neural network-Deep Learning - DataScientest », *Formation Data Science | DataScientest.com*, 25 juin 2020. <https://datascientest.com/convolutional-neural-network> (consulté le 20 septembre 2022).
- [65] « Découvrez les différentes couches d'un CNN », *OpenClassrooms*. <https://openclassrooms.com/fr/courses/4470531-classez-et-segmentez-des-donnees-visuelles/5083336-decouvrez-les-differentes-couches-dun-cnn> (consulté le 20 septembre 2022).

Références bibliographiques

- [66] A. S. Pooling David H. v Hubel, Torsten Wiesel, Max, « Deep Learning. 2 Plan réseaux de neurones artificiels Définition Fonction d'activation Fonction de cout Propagation et rétropropagation Algorithme d'optimisation. - ppt télécharger ». <https://slideplayer.fr/slide/13975436/> (consulté le 20 septembre 2022).
- [67] Juri'Predis, « Machine learning : comprendre les réseaux de neurones », *Juri'Predis*, 29 avril 2019. <https://juripredis.com/blog/démystifier-le-machine-learning-partie-2-les-reseaux-de-neurones-artificiels/> (consulté le 20 septembre 2022).
- [68] « Les réseaux de neurones ». https://ml4a.github.io/ml4a/fr/neural_networks/ (consulté le 20 septembre 2022).
- [69] « DataScienceToday - Les réseaux de neurones convolutifs ». <https://datasciencetoday.net/index.php/en-us/deep-learning/173-les-reseaux-de-neurones-convolutifs> (consulté le 20 septembre 2022).
- [70] « Morphologie Mathématique — Documentation Traitement et analyse d'images 1 ». <https://perso.esiee.fr/~perretb/I5FM/TAI/morpho/index.html> (consulté le 21 septembre 2022)
- [71] « MRI segmentation fusion for brain tumor detection », *Information Fusion*, vol. 36, p. 1-9, juill. 2017, doi: 10.1016/j.inffus.2016.10.003.
- [72] A. Touil, K. Kalti, P.-H. Conze, B. Solaiman, et M. A. Mahjoub, « Automatic detection of microcalcification based on morphological operations and structural similarity indices », *Biocybernetics and Biomedical Engineering*, vol. 40, n° 3, p. 1155-1173, juill. 2020, doi: 10.1016/j.bbe.2020.05.002.
- [73] « Y. Ben et Salem, « Fusion d'images multimodales pour l'aide au diagnostic du cancer du sein»
- [74] I. Cabria et I. Gondra, « MRI segmentation fusion for brain tumor detection », *Information Fusion*, vol. 36, p. 1-9, juill. 2017, doi : 10.1016/j.inffus.2016.10.003.
- [75] « Qu'est-ce qu'une base de données | Oracle France ». <https://www.oracle.com/fr/database/definition-base-de-donnees.html> (consulté le 14 août 2022).
- [76] K. Loizidou, G. Skouroumouni, C. Nikolaou, et C. Pitris, « Breast Micro-Calcification Dataset ». Zenodo, 30 juin 2021. doi: 10.5281/zenodo.5036062.
- [77] B. L., « Python : tout savoir sur le principal langage Big Data et Machine Learning », *LeBigData.fr*, 31 juillet 2022. <https://www.lebigdata.fr/python-langage-definition> (consulté le 15 août 2022).

Références bibliographiques

- [78] « Python (langage) - Définition et Explications », *Techno-Science.net*.
<https://www.techno-science.net/glossaire-definition/Python-langage.html> (consulté le 15 août 2022).
- [79] « Différence entre la programmation Anaconda et Python / La programmation », *La différence entre des objets et des termes similaires*.
<https://fr.sawakinome.com/articles/programming/difference-between-anaconda-and-python-programming.html> (consulté le 15 août 2022).
- [80] « cours-python.pdf ». Consulté le: 15 août 2022. [En ligne]. Disponible sur:
<https://python.sdv.univ-paris-diderot.fr/cours-python.pdf>
- [81] D. Hellmann, *The Python Standard Library by Example*, 1er édition. Upper Saddle River, NJ: Addison Wesley, 2011.
- [82] « Évaluez un algorithme de classification qui retourne des valeurs binaires », *OpenClassrooms*.<https://openclassrooms.com/fr/courses/4297211-evaluez-les-performances-dun-modele-de-machine-learning/4308256-evaluez-un-algorithme-de-classification-qui-retourne-des-valeurs-binaires> (consulté le 27 août 2022).
- [83] A. Touil, K. Kalti, P.-H. Conze, B. Solaiman, et M. A. Mahjoub, « Automatic detection of microcalcification based on morphological operations and structural similarity indices », *Biocybernetics and Biomedical Engineering*, vol. 40, n° 3, p. 1155-1173, juill. 2020, doi: 10.1016/j.bbe.2020.05.002.