

الجمهورية الجزائرية الديمقراطية والشعبية



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي



MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE

Université d'Ibn-Khaldoun de Tiaret

FACULTE DES MATHEMATIQUES ET DE L'INFORMATIQUE

Département : INFORMATIQUE

MEMOIRE

Pour l'obtention du diplôme de MASTER

En : Informatique

Spécialité : Génie informatique

Par : Daniele Bernard Emmanuelle LELOU

Thème :

INTERPRETATION DES  
IMAGES DE  
MAMMOGRAPHIE

Soutenu publiquement, le 16 / 11 / 2020 , devant le jury composé de :

Mr. CHIKHAOUI Ahmed	MCB	Université de Tiaret	Président
Mr. BAGHDADI Mohamed	MCB	Université de Tiaret	Encadreur
Mr. MEZZOUG Karim	MAA	Université de Tiaret	Examineur

Année Universitaire : 2019-2020

# Remerciements

« Celui qui ne remercie pas les gens ne remerciera pas Dieu »

*Je tiens à remercier tout d'abord, DIEU qui m'a donnée la force et le courage afin de parvenir à élaborer ce modeste travail malgré les difficultés et la longue attente dû à la COVID-19*

*Je tiens à exprimer toute ma gratitude auprès de mon directeur de mémoire Monsieur BAGHDADI MOHAMED pour la confiance qu'il m'a accordée et à le féliciter pour sa patience envers moi, qui m'a ouvert la porte pour le domaine de recherche et m'a inspiré avec sa méthode de travail et son humanité.*

*J'adresse mes remerciements aux membres du jury Mr. CHIKHAOUI AHMED et Mr. MEZZOUG KARIM qui m'ont fait l'honneur d'accepter d'évaluer ce travail.*

*Je remercie également :*

*Tous les professeurs à l'Université de Tiaret au département d'Informatique qui m'ont assisté tout au long de mon cursus Universitaire, pour le temps qu'ils m'ont consacré, pour leurs conseils et encouragements ainsi que l'orientation qu'ils m'ont accordé qui répond parfaitement à mon projet d'étude.*

*Enfin, adressons nos remerciements à toutes les personnes qui de près ou de loin m'ont aidée et soutenus durant cette période.*

## *Dédicaces*

*Je dédie ce modeste travail à :*

*A mes parents, pour leur soutien et tous les efforts qu'ils ont fournis le long de mon parcours et je leur souhaite bonne santé et longue vie.*

*A mes frères et ma sœur.*

*A tous ceux qui m'ont aidée de loin ou de près durant mes études.*

## Résumé

Le cancer du sein est l'une des causes les plus courantes de décès chez la femme. Une femme sur huit risque d'en être atteinte durant sa vie. Son dépistage à un stade précoce avant que la maladie n'ait eu le temps de s'étendre, revêt une importance capitale. Du fait de son diagnostic tardif, il en résulte souvent un traitement douloureux. Grâce à la mammographie numérique et aux systèmes d'aide à la décision, les chances de survie ont été considérablement augmentées. Le défi est de détecter des cancers qui risquent d'être manqués par le radiologue. L'objectif de ce thème est le développement d'un système de détection et de classification des anomalies suspectes sur des clichés mammographiques. Le système qu'on propose se voit divisé en plusieurs étapes : une phase de prétraitement qui vise à améliorer la qualité des images de mammographie pour aider à la détection du cancer du sein à l'aide de 5 (cinq) algorithmes d'approche floue. Ensuite, la segmentation à l'aide de la méthode de seuillage de Otsu permet de trouver des sections plus petites à couper et à analyser à l'étape suivante afin de segmenter les structures suspectes. La troisième étape est l'extraction des caractéristiques, après avoir essayé certaines fonctionnalités, nous avons retenus 6 caractéristiques qui représentent au mieux les images mammographiques. Nous avons utilisé la méthode MCNG (Matrice de co-occurrence de niveau gris) pour le calcul de ces caractéristiques. Enfin l'étape de classification dans laquelle nous utilisons un système basé sur la classification à l'aide des réseaux de neurones qui divise 6 entrées de l'extraction de fonctionnalités en deux sorties (normale et cancer) avec une taille de couche cachée égale à 5. Nous avons appliqué notre système sur plusieurs images mammographiques, les résultats obtenus sont encourageants.

### Mot clés :

Cancer du sein, Segmentation des masses, Systèmes d'aide à la détection (CADe), logique floue, Algorithme de seuillage, les réseaux de neurones artificiels, Matrice de co-occurrence de niveau gris (MCNG).

## **Abstract Summary**

Breast cancer is one of the most common causes of death in women. One in eight women is at risk of developing it in her lifetime. Screening at an early stage before the disease has a chance to spread is of paramount importance. Because of his late diagnosis, it often results in painful treatment. Thanks to digital mammography and decision support systems, the chances of survival have been greatly increased. The challenge is to detect cancers that may be missed by the radiologist. The objective of this theme is the development of a system for detecting and classifying suspicious abnormalities on mammographic images. The proposed system is divided into several stages: a pre-treatment phase that aims to improve the quality of mammography images to help detect breast cancer using 5 (five) fuzzy approach algorithms. Then, segmentation using The Otsu threshold method allows smaller sections to be cut and analyzed in the next step to segment suspicious structures. The third step is the extraction of features, after trying some features, we have selected 6 features that best represent mammographic images. We used the MCNG method (Grey Level Co-occurrence Matrix) to calculate these characteristics. Finally the classification stage in which we use a system based on classification using neural networks that divides 6 inputs of feature extraction into two outputs (normal and cancer) with a hidden layer size equal to 5. We applied our system to several mammographic images; the results are encouraging.

### **Key words:**

Breast Cancer, Mass Segmentation, Detection Assistance Systems (CADe), Blurred Logic, Threshold Algorithm, Artificial Neural Networks, Grey Level Co-occurrence Matrix (MCNG).

سرطان الثدي هو واحد من الأسباب الأكثر شيوعا للوفاة في النساء. واحدة من كل ثماني نساء معرضة لخطر تطورها في حياتها. الفحص في مرحلة مبكرة قبل أن يكون للمرض فرصة للانتشار هو من الأهمية بمكان. بسبب تشخيصه المتأخر ، فإنه غالبا ما يؤدي إلى علاج مؤلم. وبفضل التصوير الرقمي للثدي ونظم دعم القرار ، زادت فرص البقاء على قيد الحياة إلى حد كبير. ويتمثل التحدي في الكشف عن السرطانات التي قد يغيب عنها أخصائي الأشعة. والهدف من هذا الموضوع هو وضع نظام للكشف عن التشوهات المشبوهة وتصنيفها على الصور الشعاعية للثدي. ينقسم النظام لمقترح إلى عدة مراحل: مرحلة ما قبل العلاج التي تهدف إلى تحسين جودة صور التصوير الشعاعي للثدي للمساعدة في 5 (خمسة) خوارزميات النهج الضبابي. ثم، تجزئة باستخدام الأسلوب عتبة أوتسو يسمح المقاطع أصغر أن يتم قطع وتحليلها في الخطوة التالية لتجزئة الهياكل المشبوهة. الخطوة الثالثة هي استخراج الميزات، بعد محاولة بعض الميزات، اخترنا 6 ميزات التي تمثل الصور الماموغرافية أفضل. استخدمنا طريقة MCNG ( التواجد المشترك للمستوى الرمادي) لحساب هذه الخصائص. أخيرا مرحلة التصنيف التي نستخدم فيها التصنيف باستخدام الشبكات العصبية التي تقسم 6 مدخلات لاستخراج المعالم إلى مخرجين (عادي وسرطان) بحجم طبقة مخفية يساوي 5 .

طبقتنا نظامنا على العديد من الصور الماموغرافية، والنتائج مشجعة.

#### الكلمات الرئيسية:

(CADE) ، وعدم وضوح المنطق، خوارزمية العتبة، والشبكات

العصبية الاصطناعية، ومصفوفة التواجد المشترك المستوى الرمادي.(MCNG)

# Table des matières

<b>Introduction générale</b> .....	1
------------------------------------	---

## **CHAPITRE 1 : Généralités sur le cancer du sein**

I.1 Introduction.....	5
I.2 L'anatomie macroscopique du sein .....	5
I.3 Définition du cancer du sein .....	6
I.4 Les types d'anomalies dans le sein.....	7
I.4.1 Macro-calcification .....	8
I.4.2 Microcalcification .....	8
I.4.3 Les masses .....	8
I.4.3.1 La forme.....	8
I.4.3.2 Le contour .....	9
I.4.3.3 Densité mammaire .....	10
I.5 La mammographie [4].....	12
I.5.1 Principe .....	12
I.5.2 Types de mammographie .....	12
I.5.3 Mammographie numérique.....	13
I.5.3.1 Avantages et inconvénients de la mammographie numérique [26] .....	13
I.6 Relation entre les images de mammographie et l'anatomie du sein.....	15
I.7 Le mammographe .....	15
I.8 La classification des anomalies mammographiques .....	16
I.9 Conclusion .....	17

## **CHAPITRE 2 : Interprétation des images de mammographies**

II.1 Introduction .....	19
II.2 Définition de traitement d'image .....	19
II.2.1 Exemples de traitement d'images [22] .....	20
II.3 Cadre général d'un système d'interprétation des images médicales .....	21
II.3.1 Acquisition.....	21
II.3.2 Prétraitement d'images.....	22
II.3.2.1 Modification d'histogramme des mammographies .....	22
II.3.2.2 Filtrage des mammographies .....	23
II.3.3 Segmentation d'images.....	24
II.3.4 Interprétation des images.....	24

II.4 Conclusion.....	25
----------------------	----

**CHAPITRE 3 : Etat de l'art sur les méthodes de segmentation des images mammographiques**

III.1 Introduction .....	27
III.2 Principe de la segmentation.....	27
III.3 Objectif de la segmentation.....	28
III.4 Les différentes approches de segmentation.....	30
III.4.1 Segmentation par approche contour .....	31
III.4.1.1 Filtre Sobel .....	32
III.4.1.2 Canny-Dérivée .....	33
III.4.1.3 Les Modèles déformables.....	34
III.4.2 Segmentation par approche régions .....	37
III.4.2.1 Méthodes de croissance de régions .....	40
III.4.2.2 Approche fonctionnant par fusion et division de région .....	41
a) <i>Segmentation par division de région (Split)</i> .....	41
b) <i>Segmentation par fusion de régions (Merge)</i> .....	42
c) <i>Segmentation par division/fusion (Split and Merge)</i> .....	43
III.4.3 Approches Contours - Régions .....	43
III.4.3.1 Coopération séquentielle .....	44
III.4.3.2 Coopération des résultats .....	44
III.4.3.3 Coopération mutuelle .....	45
III.4.5 Les méthodes basées sur l'estimation itérative de paramètres .....	46
III.4.6 la Segmentation basée sur la classification .....	46
III.4.6.1 Méthodes de segmentation basées sur le seuillage d'histogramme .....	47
c) <i>Mise en œuvre du seuillage</i> .....	48
III.4.6.2 la méthode de K-means .....	49
III.4.7 Segmentation par Théorie des graphes.....	50
III.4.8 Segmentation par approche structurelle .....	50
III.4.8.1 La méthode de la ligne de partage des eaux.....	50
III.5 Aperçu de quelques travaux pour la détection de pathologies mammographiques .....	54
III.5.1 Les travaux de A. RICK.....	54
III.5.2 Les travaux de C.VACHIER .....	55
III.5.3 Les travaux de Rangayyanne.....	57
III.5.4 Les travaux de Boujelben.....	58
III.6 Conclusion.....	60

## CHAPITRE 4 : Approche proposée

IV.1 Introduction .....	62
IV.2 Objectifs de notre étude .....	62
IV.3 Présentation du système général .....	63
IV.3.1 Pré-traitement.....	64
VI.3.1.1 Algorithme de distribution des possibilités.....	64
IV.3.2 Segmentation.....	65
IV.3.2.1 La méthode de OTSU .....	65
IV.3.2.2 Détermination des régions d'intérêt.....	67
IV.3.3 Extraction des caractéristiques.....	67
IV.3.3.1 Algorithme MCNG (Matrice de co-occurrence de niveau gris) .....	67
IV.3.4 Classification avec les réseaux de neurones .....	68
IV.4 Conclusion.....	71

## CHAPITRE 5 : Expérimentations et résultats obtenus

V.1 Introduction .....	73
V.2 Les Bases de données des images de mammographie .....	74
V.3 Implémentation et présentation de l'interface .....	74
V.3.1 Langage de programmation .....	74
V.4 Expérimentation et résultats .....	75
V.4.1 résultats obtenus dans l'étape de prétraitement .....	75
V.4.2 Résultats de l'expérience d'amélioration de la qualité de l'image.....	75
V.4.3 Résultats de l'étape de segmentation.....	76
V.4.4 Résultat de l'étape d'extraction des caractéristiques .....	77
V.4.5 Résultats de l'étape de classification.....	78
V.5 Discussions.....	81
V.6 Conclusion.....	81
<b>Conclusion générale et perspectives.....</b>	<b>83</b>
<b>Références.....</b>	<b>87</b>

# Liste des figures

## Chapitre 1

<b>Figure 1.1</b> - La structure du sein.....	8
<b>Figure 1.2</b> - Cellules cancéreuses.....	9
<b>Figure 1.3</b> - Les différentes formes possibles d'une masse : (a) ronde, (b) ovale, (c) lobulée, (d) irrégulière.....	9
<b>Figure 1.4</b> - Les différents contours possibles d'une masse : (a) circonscrit, (b) micro-lobulé, (c) masqué, (d) indistinct, (e) spéculé.....	11
<b>Figure 1.5</b> - Différents types de densité mammaire [5].....	12
<b>Figure 1.6</b> - Les composants d'un mammographe.....	17

## Chapitre 2

<b>Figure 2.1</b> - Chaîne de Vision par ordinateur.....	21
<b>Figure 2.2</b> - La chaîne de traitement et d'analyse des images.....	28

## Chapitre 3

<b>Figure 3.1</b> - Segmentation d'image en 5 régions.....	28
<b>Figure 3.2</b> - Classification des différentes méthodes de segmentation.....	30
<b>Figure 3.3</b> - Illustration de détection de contours par les différents filtres [42].....	31
<b>Figure 3.4</b> - Exemples de contours, (a) Contour idéal, (b) rampe, (c) contour réel.....	32
<b>Figure 3.5</b> - Détection de contour par le filtre Sobel.....	33
<b>Figure 3.6</b> - Détection de contour par Canny-Dérivée.....	34
<b>Figure 3.7</b> - Principe des contours actifs.....	35
<b>Figure 3.8</b> - Exemple de segmentation de masse par contours actifs (snakes) [Mustafa 2019].....	37
<b>Figure 3.9</b> - Processus de division de l'image I utilisant le tétra-arbre.....	38
<b>Figure 3.10</b> - Exemple de diagramme de Voronoï.....	39
<b>Figure 3.11</b> - Le graphe d'adjacence des régions, (a) image d'étiquettes et (b) graphe d'adjacence de régions.....	39
<b>Figure 3.12</b> - Les différentes méthodes de segmentation par approche région.....	40
<b>Figure 3.13</b> - Segmentation par le principe de la croissance de régions [28].....	41
<b>Figure 3.14</b> - Graphe d'adjacence de régions.....	43
<b>Figure 3.15</b> - Schéma de la coopération séquentielle.....	44

<b>Figure 3.16</b> - Synoptique de la coopération des résultats.....	45
<b>Figure 3.17</b> - Synoptique de la coopération mutuelle.....	45
<b>Figure 3.18</b> - Les méthodes de segmentation par estimation itérative de paramètres.....	46
<b>Figure 3.19</b> - Seuillage par utilisation d'histogramme.....	47
<b>Figure 3.20</b> - L'histogramme d'une image mammographie.....	47
<b>Figure 3.21</b> - Elément de relief.....	51
<b>Figure 3.22</b> - Principe de la procédure d'immersion.....	51
<b>Figure 3.23</b> – Principe de la ligne de partage des eaux.....	52
<b>Figure 3.24</b> - Délimitation des composantes connexes.....	52
<b>Figure 3.25</b> - Processus de LPE par inondation, (a) avec contrainte et (b) sans contrainte....	53
<b>Figure 3.26</b> - Sous-ensembles flous adaptatifs [47].....	55
<b>Figure 3.27</b> - Les principales étapes de l'algorithme de détection des opacités du sein [46]..	56
<b>Figure 3.28</b> - Segmentation des sur-densités en utilisant la dynamique [46].....	57
<b>Figure 3.29</b> - Segmentation par croissance de régions permettant de décrire des contours flous [45].....	58
<b>Figure 3.30</b> - Résultat de la transformation logarithmique : (a) image originale, b) image améliorée [48].....	59
<b>Figure 3.31</b> - Résultats de la binarisation : (a) image améliorée, (b) méthode d'Otsu, (c) Méthode Maximum-entropie, (d) Méthode de la « Maximum corrélation » [48].....	59
<b>Figure 3.32</b> - Résultats de l'orientation [48].....	59

## Chapitre 4

<b>Figure 4.1</b> - Schéma général de l'approche proposée .....	63
<b>Figure 4.2</b> - Fonction de distribution de possibilité.....	64
<b>Figure 4.3</b> – Architectures des réseaux de neurones .....	69
<b>Figure 4.4</b> - Nœuds uniques sur le réseau MLP .....	69
<b>Figure 4.5</b> - MLP avec une couche cachée .....	70

## Chapitre 5

<b>Figure 5.1</b> - Exemple d'amélioration de la qualité de l'image.....	75
<b>Figure 5.2</b> - Résultats d'expériences à l'aide de données de mammographie normales.....	76
<b>Figure 5.3</b> - Résultats des essais expérience à l'aide de données Cancer de mammographie du cancer.....	76

<b>Figure 5.5</b> – (Etape 1 de notre système d'interprétation (début)).....	78
<b>Figure 5.6</b> – (a) l'image d'origine, (b) l'image après amélioration, c) l'image après seuillage.....	79
<b>Figure 5.7</b> – (a) l'image d'origine, (b) l'image après amélioration, c) l'image après seuillage.....	79
<b>Figure 5.8</b> – (a) l'image d'origine, (b) l'image après amélioration, c) l'image après seuillage.....	80
<b>Figure 5.9</b> – (a) l'image d'origine, (b) l'image après amélioration, c) l'image après seuillage.....	80

## Liste des Tableaux

<b>Tableau 1.1</b> - La classification BI-RADS de la densité mammaire [5].....	11
<b>Tableau 5.1</b> – Exemple des différentes valeurs des caractéristiques obtenues pour des images cancéreuses et normales.....	78

# *INTRODUCTION GENERALE*

**L**e domaine du traitement de l'image est un domaine qui est encore en expansion et qui pénètre tous les secteurs d'activités, des environnements industriels et professionnels au monde des applications grand public. Malgré les progrès significatifs réalisés ces dernières années, l'image reste aujourd'hui encore un sujet difficile et les solutions actuelles ne permettent pas toujours de résoudre de manière satisfaisante de nombreux problèmes importants. Le traitement et l'analyse d'images par ordinateur occupe par ailleurs une place importante dans le domaine médical.

**L**e cancer du sein représente l'un des enjeux prépondérants dans le domaine de la santé publique. En effet, il s'agit d'un fameux cancer menaçant la vie de la plupart des femmes. Environ une femme sur dix est touchée par cette maladie durant son existence. Cependant, la réduction du taux de mortalité causée par ce type de cancer ainsi que la favorisation des chances de guérison ne sont possibles que si la tumeur a été prise en charge dès les premiers stades de son apparition. Dans l'objectif d'assurer le dépistage précoce d'une telle tumeur, les radiologues ont été amenés à augmenter la fréquence des mammographies notamment pour la tranche d'âge 40-50 ans. Car la mammographie est la technique de référence incontournable pour l'exploration du sein, la plus performante en matière de surveillance, elle permet de mettre les anomalies en évidence telle que les opacités et les microcalcifications cependant tous les radiologues reconnaissent la difficulté de l'interprétation des mammographies parce qu'elle représente une modalité d'image complexe à interpréter à cause de la variété de densité des tissus, des structures compliquées du sein, et de la diversité des tumeurs en termes de type, de forme, de contours, etc.

**L**e but de notre travail est d'apporter aux médecins une aide essentielle pour la détection

précoce du cancer du sein mais aussi lors de la phase préopératoire, en améliorant l'analyse et le traitement d'images de manière à faciliter la planification de l'intervention chirurgicale. Ceci permettrait de gagner en précision du geste chirurgicale avec toutes les conséquences qui s'en suivent (minimiser les complications, réduction du temps d'hospitalisation).

La conception d'un système de détection assistée par ordinateur (CAD) représente un système pour aider les radiologistes dans l'interprétation des mammographies pour le dépistage des masses et des calcifications.

En dépit des efforts faits par les chercheurs, l'automatisation de détection des pathologies mammaires demeure toujours difficile. Au cours des dernières années, il y a eu des efforts importants dans le développement d'algorithmes pour la détection des calcifications dans les images de mammographie.

**Objectif :**

**D**ans ce mémoire, l'objectif principal est l'interprétation des images médicales pour

l'amélioration automatique et la segmentation des masses dans les images de mammographies, c.-à-d. il s'agit de concevoir un système de reconnaissance des images mammographiques. La reconnaissance est basée sur une segmentation de ces images par une approche de segmentation pour l'extraction des informations pertinentes nécessaires, par la suite, dans la procédure de décision et de classification d'anomalies ; tout en nous concentrant sur la qualité de la segmentation car un traitement ultérieur dépend étroitement du résultat de cette dernière.

## Organisation du manuscrit :

Ce document est organisé en cinq chapitres :

- Dans le premier chapitre, nous allons étaler les généralités sur le sein, le cancer du sein ainsi que la mammographie pour situer d'une part le contexte dans lequel la contribution que nous cherchons à développer va être utilisée et d'autre part pour faciliter la compréhension de la problématique.
- Dans le deuxième chapitre, nous survolons brièvement les systèmes d'interprétation assistés par ordinateur et surtout en mammographies. Ils présentent leurs objectifs ainsi que leurs fonctionnements. Il met en relief les concepts de base, les différentes théories ainsi que les applications de reconnaissances.
- Le chapitre trois présente un état de l'art sur les méthodes de segmentation des images mammographiques car c'est l'étape la plus importante du système CAD.
- Le chapitre 4 est consacré à la description de notre contribution, étale le modèle proposé pour la seconde lecture des mammographies. Nous présentons d'une manière détaillée les étapes dédiées à la conception ainsi que les approches proposées pour localiser les régions suspectes par le système d'aide à la détection que nous voulons développer.
- Le dernier chapitre dévoile les différentes étapes de la partie expérimentale pour la conception de notre système d'aide au diagnostic. Il présente les différents résultats obtenus.

**N**ous terminons notre mémoire par une conclusion générale qui présente une synthèse

des contributions apportées ainsi que quelques perspectives ouvertes par cette étude.

# **CHAPITRE I :**

## ***Généralités sur le cancer du sein***

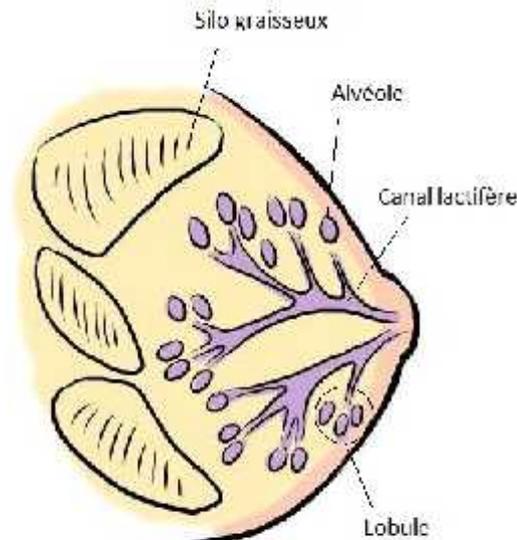
## **I.1 Introduction**

Le cancer du sein est la maladie la plus courante chez la femme dans le monde. Il est l'une des principales causes de mortalité féminine. Le taux de mortalité du cancer du sein a globalement diminué de 36 % entre 1989 et 2012, après avoir lentement augmenté (0,4% par an) depuis 1975 [38]. Le dépistage précoce constitue l'une des armes les plus efficaces contre cette maladie. Un cancer du sein détecté à un stade précoce est plus facile à traiter, et comporte moins de risque, et de séquelles ; il permet de diminuer la mortalité de 15% à 20% [15].

La mammographie est la technique d'imagerie la plus sensible pour détecter des tumeurs à un stade précoce, elle est actuellement la principale investigation dans le dépistage du cancer du sein ; c'est un examen radiologique qui utilise des rayons X pour produire des images de haute résolution, de la structure intérieure du sein. Les images permettent de détecter des zones de surdensité pouvant être évocatrice de cancer. Dans ce chapitre, nous aborderons les aspects médicaux du cancer du sein, ses différentes formes, et la place de la mammographie numérique dans le diagnostic précoce du cancer.

## **I.2 L'anatomie macroscopique du sein**

Le sein est un organe pair et symétrique de forme hémisphérique, situé en avant du thorax, entre la troisième et la cinquième côte, au-dessus du muscle grand pectoral. Une coupe longitudinale passant par le mamelon, laisse entrevoir l'architecture de la glande mammaire : elle se divise en 15 à 20 unités glandulaires indépendantes, lobes mammaires, divisés en lobules par des cloisons conjonctives, qui s'achèvent en douzaines de minuscules bulbes, acini, base de tout raisonnement physiopathologique du sein. Les lobes, lobules et bulbes sont reliés entre eux par les canaux galactophores, qui se terminent par un canal unique s'abouchant au niveau du mamelon, situé au centre d'une zone pigmentée, l'aréole. La plus grande partie du sein est constituée de tissu adipeux qui comble l'espace situé entre les différentes structures du sein (voir figure 1.1). En période d'allaitement, les lobules fabriquent du lait et les canaux le transportent vers le mamelon où il est tété par le bébé. Les tissus mammaires baignent dans des liquides lymphatiques, qui permettent d'éliminer les déchets, les cellules mortes et autres débris. La circulation lymphatique du sein est importante dans la mesure où elle va déterminer les lieux où l'on retrouvera des métastases lors de cancer.



**Figure 1.1 - La structure du sein.**

### **I.3 Définition du cancer du sein**

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

La figure 3.2 présente les étapes de développement des cellules cancéreuses dans le sein.

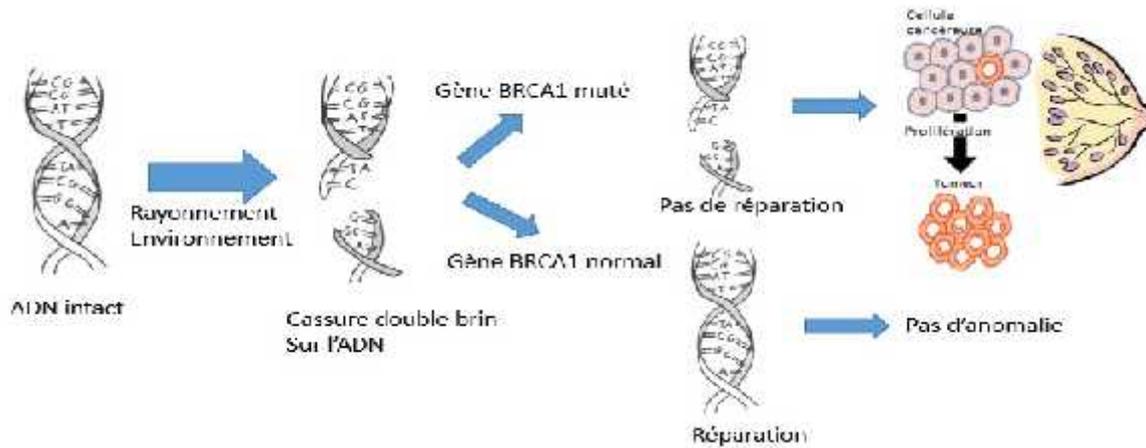


Figure 1.2 - Cellules cancéreuses.

## I.4 Les types d'anomalies dans le sein

À la mammographie, on note trois signes nettement évocateurs d'un cancer :

- ❖ Une opacité irrégulière (blanche sur la radiographie).
- ❖ La présence des microcalcifications.
- ❖ Une Modification de l'architecture normale de la glande mammaire.

L'échographie, quant à elle, permet de détecter un cancer qui apparaît comme une masse noire/grise irrégulière déjà détectable dès 3-4 mm. Elle permet parfois de détecter les microcalcifications (lorsqu'elles sont abondantes) avec des appareils performants de dernière génération [27]. La généralisation du dépistage du cancer du sein, qu'il soit du dépistage individuel ou du dépistage de masse organisé, amène à découvrir beaucoup d'anomalies purement radiologiques. Parmi ces anomalies, on trouve :

- ❖ Macro-calcification.
- ❖ Microcalcification.
- ❖ Masse.

### **I.4.1 Macro-calcification**

Ce sont des gros dépôts calciums dans le sein qui sont associés à des affections bénignes comme par exemple une ancienne plaie, infection ou un kyste calcifié [38].

### **I.4.2 Microcalcification**

Ce sont des dépôts calciums minuscules ( $< 0,5$  mm) qui ressemblent à des grains de sable dans le sein. Ce sont des anomalies fréquentes et souvent tout à fait anodines. Elles sont facilement vues en mammographie du fait de leur forte densité : elles apparaissent comme des ponctuations blanches sur le gris des tissus. Elles représentent la majorité des problèmes rencontrés. Sur les mammographies de dépistage. Quand elles sont nombreuses, de forme irrégulière et groupées, cela peut être un signe de cancer du sein. Dans ce cas une biopsie est indiquée. Cependant la majorité des microcalcifications sont bénignes.

Le seul signe précoce de cancer que l'on peut observer sur une mammographie est la présence de microcalcifications. Dans 40% à 70% des cas, selon les études, ces microcalcifications sont associées à un processus malin [38].

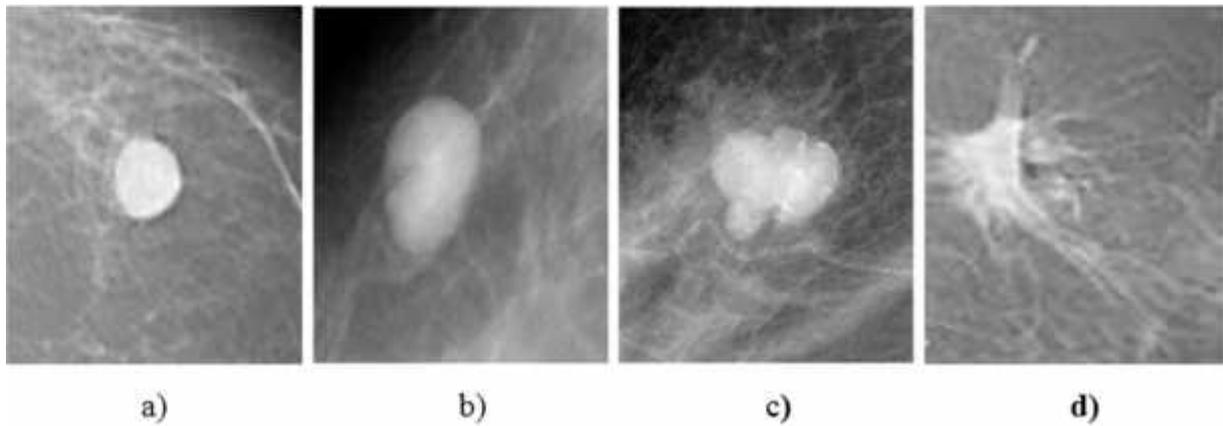
### **I.4.3 Les masses**

Une masse est une opacité importante occupant un espace dans le sein et vue comme une tâche blanche sur l'image mammographie. Il peut s'agir d'un kyste (collection liquidienne non cancéreuse) ou d'une lésion solide, qui peut correspondre à un cancer de sein.

Différents attributs permettent aux médecins de la décrire en vue de déterminer leur nature : sa forme, son contour et sa densité.

#### **I.4.3.1 La forme**

Selon la description du BIRADS (Breast Imaging Reporting and Data System), les masses mammaires peuvent avoir la forme ronde, ovale, lobulée ou irrégulière [15]. Un exemple de ces formes est montré sur la figure 1.3.



**Figure 1.3 - Les différentes formes possibles d'une masse : (a) ronde, (b) ovale, (c) lobulée, (d) irrégulière.**

**Une masse Ronde :** Il s'agit de masse sphérique, circulaire ou globuleuse.

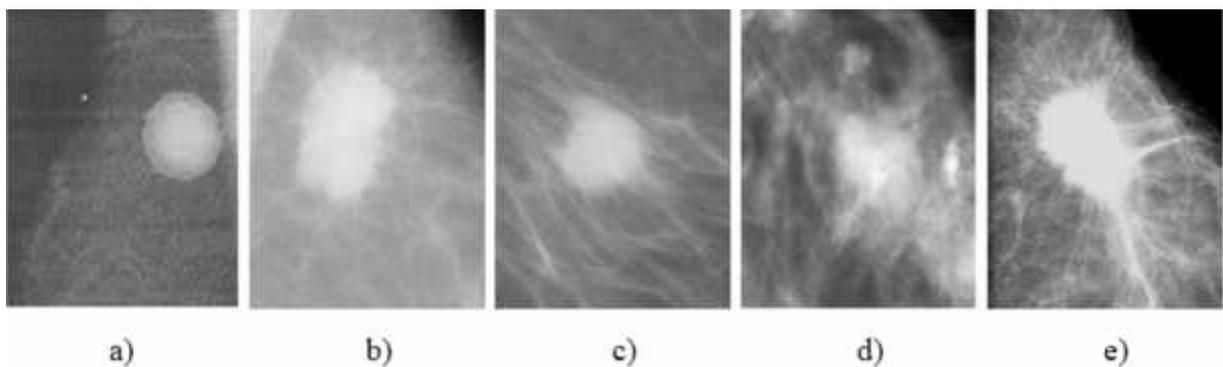
**Une masse Ovale :** Elle présente une forme elliptique (ou en forme d'œuf).

**Une masse Lobulée :** La forme de la masse présente une légère ondulation.

**Une masse Irrégulière :** Cette appellation est réservée aux masses dont la forme est irrégulière.

### I.4.3.2 Le contour

Le contour des masses mammaires est soit : circonscrit, micro lobulé, masqué, indistinct ou spéculé [15] (voir la figure 1.4). On détaille dans ce que suivent ces différentes notions :



**Figure 1.4 - Les différents contours possibles d'une masse : (a) circonscrit, (b) micro-lobulé, (c) masqué, (d) indistinct, (e) spéculé.**

a) **Circonscrit** : Il s'agit d'une transition brusque entre la lésion et le tissu environnant. Le contour est alors net et bien défini. Pour qu'une masse soit qualifiée de circonscrite, il faut qu'au moins 75% de son contour soit nettement délimité.

b) **Micro-lobulé** : Dans ce cas, de courtes dentelures du contour créent de petites ondulations

c) **Masqué** : Un contour masqué est un contour qui est caché par le tissu normal adjacent. Ce terme est employé pour caractériser une masse circonscrite dont une partie du contour est cachée.

d) **Indistinct** : Dans ce cas, le contour est mal défini. Ce caractère indistinct (le contraire de circonscrit) peut correspondre à une infiltration.

e) **Spéculé** : La masse est caractérisée par des lignes radiaires prenant naissance sur le contour de la masse. Ces lignes radiaires sont appelées les spécules.

### **I.4.3.3 Densité mammaire**

Le sein est constitué de différents types de tissus : le tissu glandulaire et conjonctif et la graisse. La densité des seins décrit la quantité relative des différents tissus dans le sein. Les seins denses présentent moins de tissu graisseux et plus de tissu conjonctif et glandulaire. La recherche montre que le risque de présenter un cancer du sein est 4 à 6 fois plus élevé chez les femmes qui ont des seins denses que chez celles qui n'ont pas des seins denses.

Le tissu mammaire change au fur et à mesure que nous vieillissons ; il devient généralement moins dense avec l'âge, bien que certaines femmes aient toujours des seins denses sans égard à leur âge [28]. La classification BI-RADS (Breast Imaging Reporting and Data System) de l'American College of Radiology permet de classer les seins en quatre groupes pour lesquelles le risque de rater un cancer en raison du caractère hétérogène du tissu fibro-glandulaire augmente lorsque l'on passe de la catégorie 1 à la catégorie 4 (figure 1.5). Une description des différents types de densité mammaire est donnée dans le tableau 1.1.

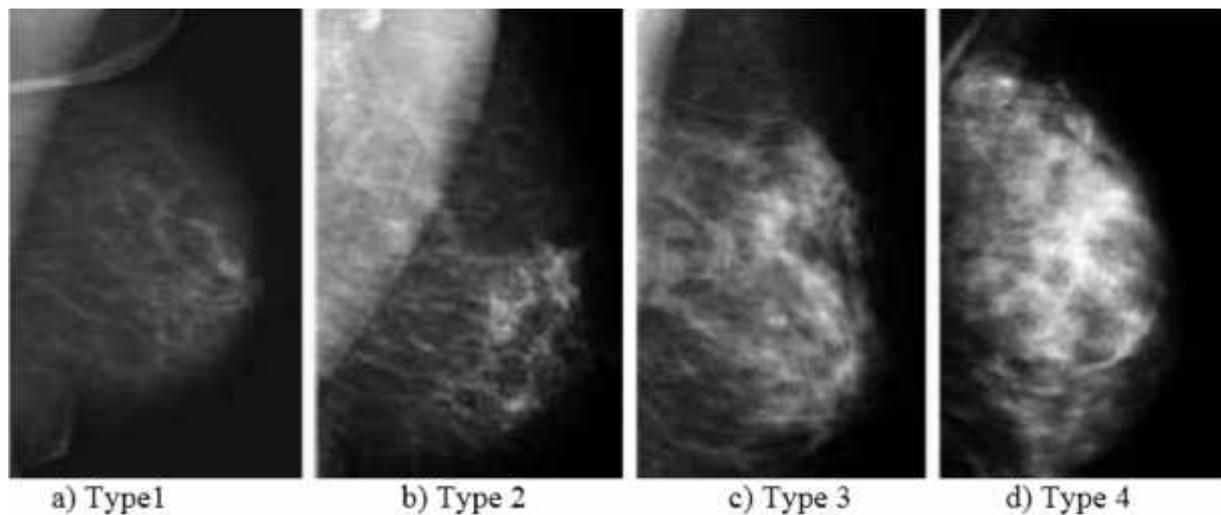


Figure 1.5 - Différents types de densité mammaire [29].

Catégorie BI-RADS*	DM exprimée en % de tissu glandulaire	
<b>A</b>	< 25%	Seins presque entièrement graisseux
<b>B</b> Majorité des femmes	25à49%	Seins composés de zones de densités Fibro-glandulaires éparses
<b>C</b>	50à74%	Seins denses de façon hétérogène, pouvant masquer des petites masses
<b>D</b> Risque augmenté jusqu'à 2 fois par rapport à B	75%	Seins extrêmement denses, Diminuant la sensibilité de la Mammographie

Tableau 1.1 - La classification BI-RADS de la densité mammaire [29].

## I.5 La mammographie [4]

La mammographie réalise une projection du volume du sein sur le plan de l'image. Elle permet d'analyser la glande mammaire grâce aux différences de l'atténuation des différents types de tissu. L'avantage de la visualisation de l'ensemble des tissus mammaires sur une seule image est directement lié à l'un de ses plus grands défauts, la superposition des différents tissus traversés par le même rayon et projetés sur un seul endroit du détecteur. Cette superposition est une source d'incertitude, puisqu'il n'est plus possible de distinguer sur l'image, des surdensités réelles qui correspondent à une région radio-opaque dans le tissu en trois dimensions, et des superpositions de plusieurs tissus avec des densités relativement faibles.

### I.5.1 Principe

La mammographie est un examen radiologique des seins qui permet de visualiser avec précision leur structure interne, elle a pour objectif le dépistage du cancer du sein à un stade précoce pour améliorer les chances de guérison. Cette technique permet de recueillir des images grâce à l'utilisation de rayons X. Dans la plupart des cas, 2 clichés sont réalisés sur chaque sein (face et oblique). On distingue généralement deux types de mammographies, la mammographie analogique et la mammographie numérique.

### I.5.2 Types de mammographie

On peut diagnostiquer trois types de mammographies [26]:

- ❖ **La mammographie diagnostique** : est effectuée chez des personnes symptomatiques une masse a été palpée dans le sein; Il existe un écoulement sanguinolent à travers le mamelon; etc. Dans ces situations, la mammographie est réalisée pour caractériser la masse palpée ou pour rechercher une tumeur. Ce type de mammographie est toujours pris en charge par les assurances maladie.
- ❖ **Mammographie préventive** : peut être demandée par le médecin, sans qu'il y ait de symptômes particuliers, en vue de s'assurer qu'il n'y a aucun point d'appel de cancer du sein ; ceci surtout chez les *femmes " à risques "*.
- ❖ **La mammographie de dépistage** : recherche la présence d'une tumeur chez des personnes non symptomatiques. Elle est réalisée chez les femmes entre 50 et 69 ans par recommandation du médecin en parallèle avec les examens médicaux généraux réguliers.

### **I.5.3 Mammographie numérique**

La mammographie numérique est une alternative à la mammographie conventionnelle. Dans le cas de cette dernière, l'image est placée sur pellicule tandis que pour la mammographie numérique, l'image est transmise directement à l'ordinateur. Ainsi, les images sauvegardées peuvent être manipulées facilement, offrant à la radiologue différents moyens d'interprétation. La mammographie numérique est une technologie récente et n'est donc pas aussi largement disponible que la mammographie conventionnelle. Il a été démontré que pour certaines femmes, la mammographie numérique donne de meilleurs résultats que la mammographie film-écran. Les auteurs d'une étude phare intitulée Digital Mammographic Imaging Screening Trial (DMIST) [26] ont découvert qu'en général, la mammographie numérique n'est pas plus efficace que la mammographie conventionnelle mais qu'elle s'avère beaucoup plus efficace pour les femmes âgées de moins de 50 ans, celles ayant des seins denses et les femmes pré-ménopausées ou ménopausées de tous âges.

#### **I.5.3.1 Avantages et inconvénients de la mammographie numérique [26]**

##### **a) Les avantages**

La mammographie numérique possède quelques avantages par rapport à la mammographie conventionnelle.

- ❖ **Surface dynamique plus grande** : la mammographie numérisée comparée à la mammographie analogique a en général une zone d'activité dynamique plus importante avec une relation linéaire. La mammographie numérique permet également une modification du contraste afin de mieux explorer certaines zones denses, ce qui n'est pas possible avec la mammographie analogique.
- ❖ **Complexité du travail sur l'image** : le travail sur l'image est nécessaire pour avoir une bonne représentation de l'ensemble de l'acquisition. Il consiste en un réglage optimal de la fenêtre, une soustraction de l'énergie pour avoir une bonne visualisation des microcalcifications, la tomosynthèse qui permet de résoudre les problèmes de superposition *etc.*
- ❖ **Archivage copie et envoi sans pertes** : les données numérisées peuvent être stockées de façon électronique, rappelées à tout moment et dupliquées sans perte d'informations. Elles peuvent également être transférées électroniquement, par télémammographique. Ainsi, l'examen des images et leur interprétation peuvent être faits dans un lieu différent de celui de la production de l'image. Les distances peuvent

être considérables. Par ailleurs, les examens peuvent être envoyés immédiatement, pour une deuxième lecture.

- ❖ **Fin des procédures de développement** : si les clichés ne sont pas stockés sur une copie, il n'est plus besoin ni de machine à développer, ni de chimie, ni de chambre noire. Les cassettes ne sont plus nécessaires. Jusqu'à présent, pour transférer les données, ils ont continué à faire des copies laser. Les coûts pour le stockage sont élevés. Comme toutes les autres données numérisées, la mammographie numérisée peut être générée par un système comme le PACS (Picture Archiving and Communication System).

***b) Les inconvénients***

Cette technique a quelques inconvénients tels que :

- ❖ **Résolution limitée du récepteur d'images** : le facteur le plus important qui limite ces systèmes plein champ est la résolution de l'image. Pour mettre en évidence des microcalcifications, la mammographie nécessite un contraste élevé et une résolution locale très importante. Malgré certaines faiblesses, le système couple écran/film arrive à répondre aux attentes des médecins. Dans ce système, la résolution n'est limitée que par le couple écran/film. On obtient une résolution locale qui peut aller jusqu'à 16 pixel/mm. Dans les systèmes numériques, la résolution est déterminée par la taille du pixel.
- ❖ **Résolution limitée des moniteurs** : les capteurs actuellement utilisés pour la mammographie plein champ ne remplissent pas toutes les conditions de résolution. Les moniteurs permettent une résolution satisfaisante pour la mammographie mais restent trop coûteux pour un usage généralisé.
- ❖ **Manque de standard pour l'assurance de qualité** : Les normes d'assurance de qualité et les exigences du matériel pour l'examen, le travail du film, la documentation, doivent encore être préparée et validées
- ❖ **Importance numérique des données** : Le problème principal de la mammographie numérique est le nombre élevé de données disponibles (contraste élevé et résolution très élevée). Comparé à une image IRM (0,5 à 1,0 Mbyte), un seul cliché de mammographie d'un format 18 x 14 cm, nécessite jusqu'à 30 Mbyte, cette grande quantité de données entraîne des coûts importants liés au temps de transfert d'une mammographie entière.

- ❖ **Prix élevé à l'achat et de la maintenance** : Les coûts pour l'achat de ces systèmes et pour la maintenance sont actuellement de trois à cinq fois plus importants que ceux de la mammographie conventionnelle et ne peuvent pas être compensés par le remboursement des actes radiographiques réalisés.

## **I.6 Relation entre les images de mammographie et l'anatomie du sein**

L'atténuation du faisceau de rayons X dépend du type des différents tissus mammaires traversés. En effet, la graisse est considérée comme une zone radio transparente, elle apparaît très sombre sur un cliché de mammographie. En revanche, le tissu fibro glandulaire et le calcium (qui est le composant essentiel des lésions mammaires) sont des éléments radio opaques qui apparaissent claires sur un cliché de mammographie. D'après les informations relatives à l'anatomie et à la radio transparence, on peut affirmer que l'aspect général d'une mammographie est sombre tandis que les zones contenant des microcalcifications ou des masses (composées de calcium) sont plus claires [4].

## **I.7 Le mammographe**

C'est l'appareil qui réalise les mammographies, il se compose d'un tube radiogène générateur de rayons X de faible énergie et d'un système de compression du sein (Figure 1.6). Le sein est placé entre le plateau porte-film et une plaque amovible, afin d'y être comprimé. La radiographie est réalisée sur des films argentiques ou sur des systèmes de radiologie digitale de haute qualité. L'analyse de la glande mammaire est réalisée grâce aux différences de l'atténuation des différents types de tissu. L'anatomie du sein permet d'établir la relation entre la nature du tissu mammaire et l'infiltration des rayons X.

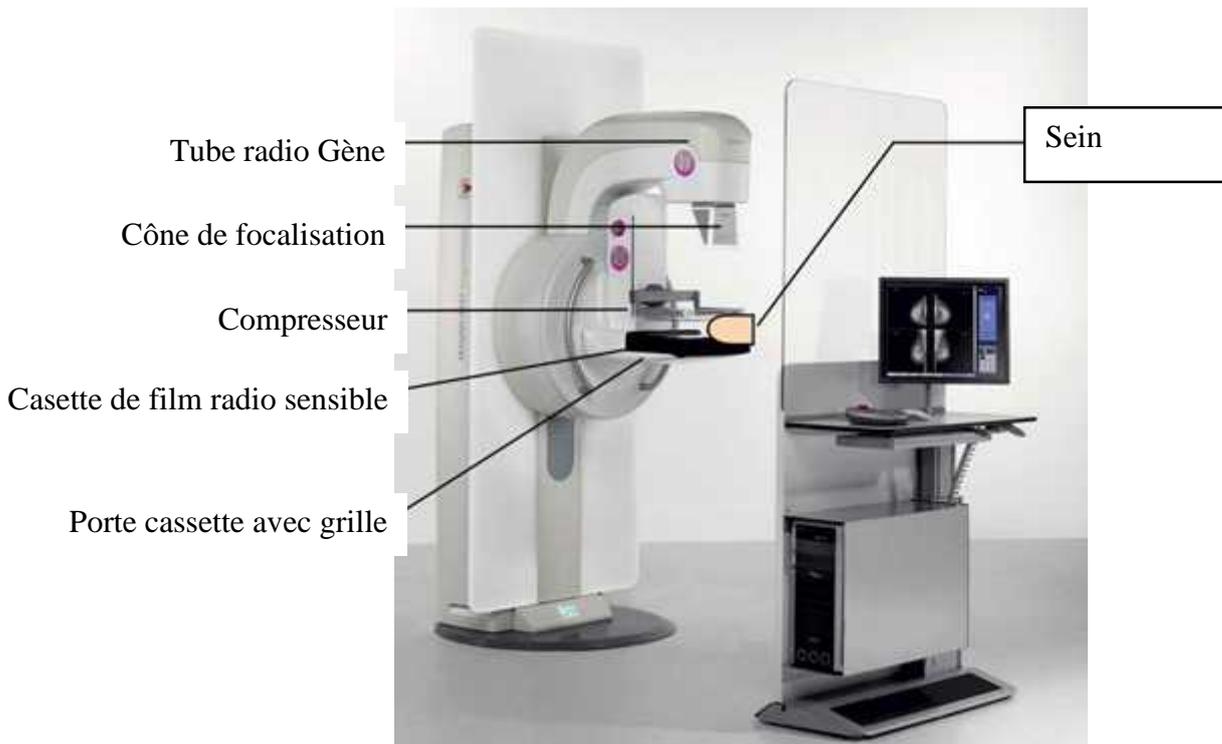


Figure 1.6 - Les composants d'un mammographe.

## I.8 La classification des anomalies mammographiques

La classification des résultats d'une mammographie comprend 6 catégories. Ces dernières ont été définies par l'American College of Radiology (ACR) et sont utilisées par la majorité des professionnels de santé à travers le monde. Les données obtenues à partir d'un examen mammographique peuvent ainsi être classées en :

- ❖ **ACR1** : cela signifie que la mammographie n'a révélé aucun problème de santé, la poitrine de la patiente est parfaitement normale.
- ❖ **ACR2** : l'examen mammographique a révélé des anomalies bénignes sans danger. Elles ne nécessitent pas d'examens complémentaires ni de surveillance particulière.
- ❖ **ACR3** : cette classification indique la présence d'une anomalie d'apparence bénigne, mais nécessitant tout de même une surveillance à court terme, d'une durée de 3 à 6 mois.

La surveillance pourra prendre la forme de nouveaux examens radiologiques, selon une échéance définie par le radiologue et en association avec le médecin traitant.

- ❖ **ACR4** : cette catégorie signifie qu'une lésion douteuse ou indéterminée a été détectée et qu'une biopsie et/ou une cytoponction sont nécessaires pour apporter un diagnostic précis. La lésion identifiée peut être bénigne, précancéreuse ou indiquer que la patiente souffre d'un cancer.
- ❖ **ACR5** : l'examen mammographique a relevé la présence d'une lésion évocatrice d'un cancer. Les patientes se voient systématiquement prescrire une biopsie ou une exérèse, une intervention chirurgicale servant à retirer un élément nuisible ou inutile à l'organisme, comme une tumeur ou un organe.
- ❖ **ACR0** : un tel classement indique que les résultats nécessitent un examen complémentaire, par exemple une lecture par un autre médecin. De nouveaux examens peuvent également être prescrits.

La classification **ACR** des résultats de mammographies est un outil particulièrement utile permettant d'évaluer l'état de santé d'une patiente et contribue à améliorer la prévention du cancer du sein.

## **I.9 Conclusion**

Dans ce chapitre, nous avons introduit quelques notions générales concernant l'anatomie et le cancer de sein ainsi que les types d'anomalies dans le sein. Une grande attention a été consacrée à l'étude des différents types des pathologies mammaires. L'étude des différentes lésions mammaires et leur classification montre, d'une part, la complexité de l'étape de détection des lésions et, d'autre part, la complexité de l'étape de décision (mammographie bénigne ou maligne). Alors, dans ces deux cas, la prise de décision par le radiologue est une tâche suffisamment délicate et primordiale pour la survie ou non d'une femme. Pour cela, plusieurs études de recherche ont été orientées vers l'automatisation de la lecture des mammographies et de la prise de décision.

## **CHAPITRE II :**

### *Le processus d'interprétation des images de mammographies*

## **II.1 Introduction**

L'imagerie médicale est l'un des domaines de la médecine qui a le plus progressé ces vingt dernières années. Elle est le procédé par lequel un médecin peut examiner l'intérieur du corps d'un patient sans l'opérer. Cela, en créant une image visuelle compréhensible d'une information à caractère médical dans le but d'établir un diagnostic et de faire un suivi approprié du traitement.

Une image numérique est une matrice de pixels repérés par leur coordonnées  $(x,y)$ . S'il s'agit d'une image couleur, un pixel est codé par 3 composantes  $(r,g,b)$  (chacune comprise au sens large entre 0 et 255), représentant respectivement les "doses" de rouge, vert et bleu qui caractérisent la couleur du pixel. S'il s'agit d'une image en niveau de gris, il est codé par la composante comprise au sens large entre 0 et 255, représentant la luminosité du pixel. Avec la parole, l'image constitue l'un des moyens les plus importants qu'utilise l'homme pour communiquer avec autrui. C'est un moyen de communication universel dont la richesse du contenu permet aux êtres humains de tout âge et de toute culture de se comprendre. C'est aussi le moyen le plus efficace pour communiquer, chacun peut analyser l'image à sa manière, pour en dégager une impression et d'en extraire des informations précises.

Le traitement des images médicales a pour but d'extraire à partir des images acquises, les informations utiles au diagnostic, de révéler des détails difficiles à percevoir à l'œil nu, tout en évitant la création d'artefacts, faussement informatifs. Pour cela le traitement fait appel à des outils, des algorithmes, qui permettent d'agir sur l'image numérisée. La reconstruction de forme, les segmentations, les quantifications, l'analyse fonctionnelle, jusqu'aux simulations (organes virtuels, maladies virtuels), tous ces outils de traitement ont contribué à l'amélioration de la qualité des images acquises, à leur interprétation et surtout à une meilleure approche au diagnostic.

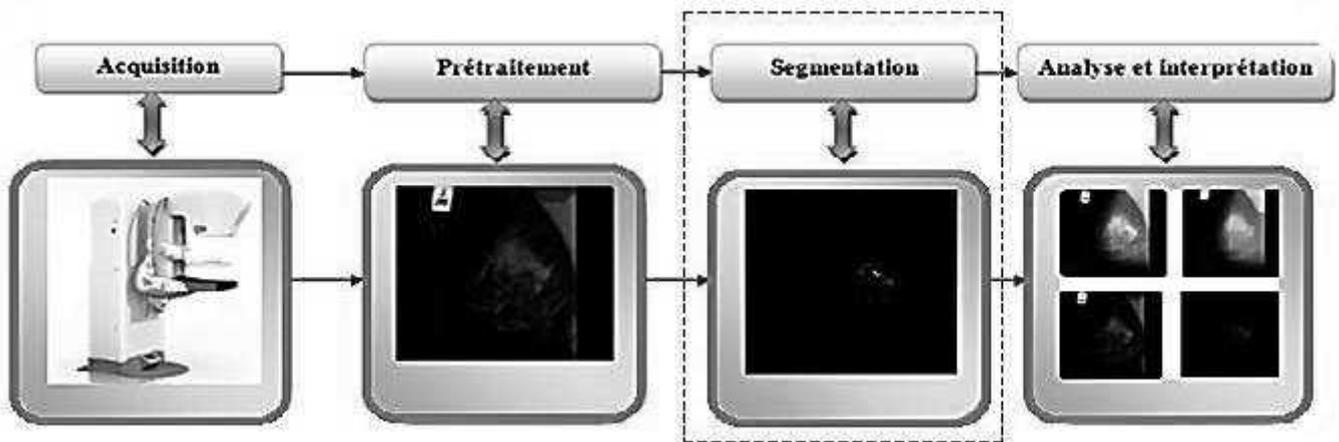
Dans ce chapitre nous donnons une description du système d'interprétation des images médicales et nous détaillons leurs différentes étapes.

## **II.2 Définition de traitement d'image**

L'analyse d'une image cherche à passer d'une grande quantité d'informations de bas niveau (i.e. les pixels) à une description symbolique de haut niveau pouvant aider à obtenir une interprétation structurelle de cette image. La distinction entre traitement de bas ou de haut

niveau est liée au contenu sémantique des entités traitées et extraites de l'image. **Les traitements de bas niveau** opèrent en général sur les grandeurs calculées à partir des valeurs attachées à chaque point de l'image sans faire nécessairement la liaison avec la réalité qu'elles représentent. Par exemple, la détection des contours est un traitement de bas niveau qui est effectué « sans comprendre » l'image. Le contour détecté peut très bien ne pas correspondre à un bord d'objet dans la réalité et ne présente aucun intérêt. Ainsi les traitements de bas niveau opèrent plutôt sur des données de nature numérique [39].

A l'opposé, **les traitements de haut niveau** s'appliquent à des entités de nature symbolique associées à une représentation de la réalité extraite de l'image. Ils sont relatifs à l'interprétation et à la compréhension de l'image et sont exprimés avec des mots du vocabulaire de l'application. Par exemple, des zones d'une image aérienne peuvent être caractérisées par leur forme (rectangulaire, linéique,...), être étiquetées avec les termes : bâtiment, route, bosquet, ombre, etc [2]. Les étapes que comprend le traitement d'images sont représentées dans la figure 3.1.



**Figure 2.1 - Chaîne de Vision par ordinateur.**

### **II.2.1 Exemples de traitement d'images [22]**

- ❖ Amélioration : augmenter la qualité de la perception visuelle qu'on a d'une image
- ❖ Restauration : compenser les dégradations (bruit, flou, ...)
- ❖ Compression : stocker et transférer efficacement les données
- ❖ Segmentation : limiter les "objets"
- ❖ Reconstruction 3D : obtenir un volume à partir de plans (images 2D)
- ❖ Représentation : modéliser
  - Bas niveau : texture, couleur, forme, etc.

- Haut niveau : caractéristiques (*features*), apprentissages statistiques, graphes.
- ❖ Analyse : convertir l'information
- ❖ Reconnaissance / Compréhension : identifier le contenu

## II.3 Cadre général d'un système d'interprétation des images médicales

Un système d'interprétation d'images numériques peut être divisé en plusieurs étapes. Une phase de prétraitements qui suit l'acquisition et la numérisation de l'image. Elle permet essentiellement de réduire une quantité importante de bruits. La phase de segmentation consiste à isoler les uns des autres les objets présents dans l'image. Suite à cette étape, vient l'étape de l'interprétation qui vise à reconnaître les pathologies recherchées dans le but d'aider le médecin dans son diagnostic [31].

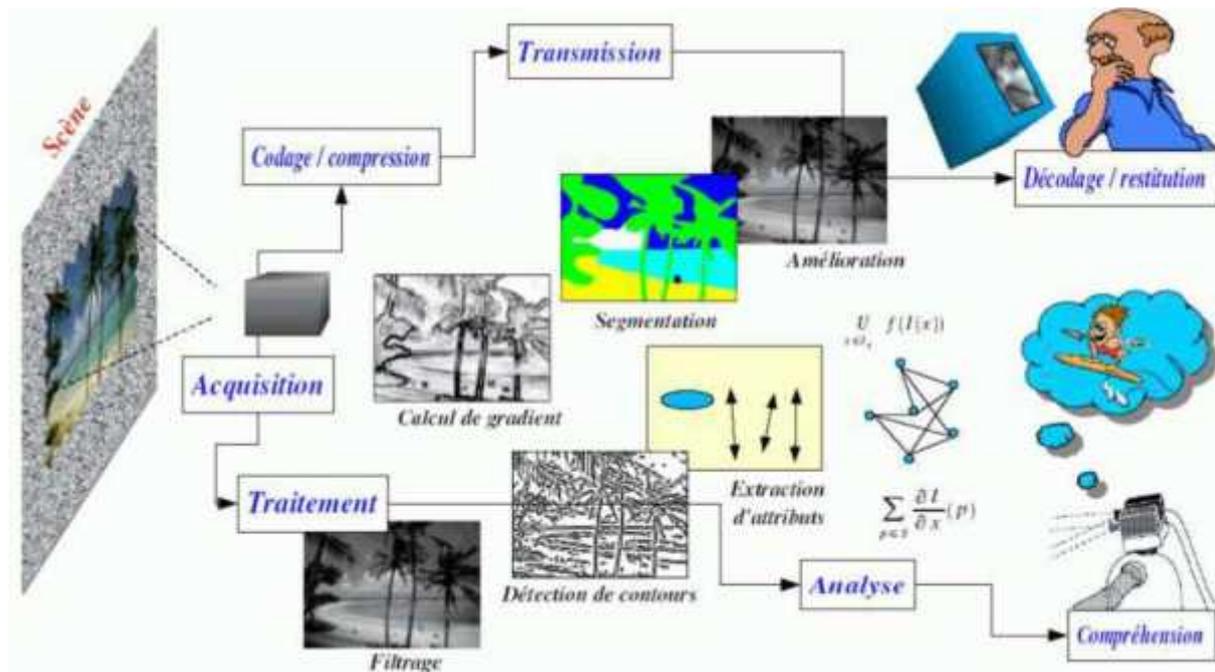


Figure 2.2 - La chaîne de traitement et d'analyse des images.

### II.3.1 Acquisition

L'acquisition d'images constitue un des outils essentiels de toute chaîne de conception et de production d'images. Pour pouvoir manipuler une image sur un système informatique, il est avant tout nécessaire de lui faire subir une transformation qui la rendra lisible et manipulable par ce système. Le passage de cet objet externe (l'image d'origine) à sa

représentation interne (dans l'unité de traitement) se fait grâce à une procédure de numérisation. Ces systèmes de saisie peuvent être classés en deux catégories principales : les caméras numériques et les scanners. A ce niveau, notons que le principe utilisé par le scanner est de plus en plus adapté aux domaines professionnels utilisant le traitement de l'image comme la télédétection, la médecine, etc.

### **II.3.2 Prétraitement d'images**

Une première étape, souvent utilisée, est donc la préparation des images avant la détection. En effet, les structures que l'on recherche n'étant pas toujours facilement discernables, une étape de prétraitement destinée à les mettre en évidence peut faciliter leur détection [44].

Le cancer du sein (comme tous les cancers d'une manière générale) doit être détecté dans sa première phase pour maximiser les chances de survie. Sauf que dans cette phase, il s'avère très difficile de repérer, à l'œil nu, la pathologie dans le tissu mammaire environnant sans avoir recours à un prétraitement spécifique de l'image acquise. D'où, l'objectif principal de cette étape est d'augmenter le contraste entre la lésion mammaire (que ce soit masse ou microcalcifications) et le reste de l'image pour faciliter les traitements ultérieurs. Sachant que dans le cas où une région d'intérêt diffère en luminance de moins de 2% du reste de l'image, elle demeure indiscernable à l'œil nu. Le prétraitement des images mammographiques est connu sous le nom de rehaussement ou d'amélioration du contraste.

#### **II.3.2.1 Modification d'histogramme des mammographies**

##### *a) Expansion de la dynamique*

La modification de l'histogramme consiste à répartir les fréquences d'apparition des pixels sur la largeur de l'histogramme pour faire varier les contrastes de façon définies et différentes selon l'intervalle de niveaux de gris considéré. Cette transformation ne fait qu'améliorer la qualité visuelle de l'image, l'information présente ne change pas.

##### *b) Egalisation de l'histogramme*

L'égalisation de l'histogramme est utilisée en cas des images de mauvaise qualité (mauvais contraste, images trop sombres ou trop claires, mauvaise répartition des niveaux d'intensité, etc.) [14]

Cette transformation consiste à rendre le plus plat possible l'histogramme des niveaux de gris de l'image. On désire que chaque niveau de gris soit équitablement représenté dans l'image.

**c) Inversion d'histogramme**

Consiste à inverser les valeurs des pixels par rapport à la moyenne des valeurs possibles. Dans une image de mammographie pathologique cette opération permettra de mieux visualiser les opacités (en noir sur fond clair sera mieux perçue qu'en blanc sur fond noir).

**II.3.2.2 Filtrage des mammographies**

Pour améliorer la qualité visuelle de l'image, on doit éliminer les effets des bruits en lui faisant subir un traitement appelé filtrage. Le filtrage est une opération qui consiste à appliquer une transformation à tout ou partie d'une image numérique. Le principe du filtrage est de modifier la valeur des pixels d'une image, généralement dans le but d'améliorer son aspect [13].

**a) Filtre passe bas (lissage)**

Ce filtre n'affecte pas les composantes de basse fréquence dans les données d'une image, mais doit atténuer les composantes de haute fréquence.

**b) Filtre passe haut (accentuation)**

Le renforcement des contours et leur extraction s'obtiennent dans le domaine fréquentiel par l'application d'un filtre passe haut. Le filtre digital passe haut a les caractéristiques inverses du filtre passe bas : Il n'affecte pas les composantes de haute fréquence d'un signal, mais doit atténuer les composantes de basses fréquences.

**c) Filtre de Gauss**

C'est un filtre linéaire passe bas. Les valeurs des coefficients sont déterminées selon une fonction Gaussienne. L'intérêt du filtre gaussien est que le degré de filtrage peut être facilement réglé à travers le paramètre qui est l'écart type.

Soit  $A[x, y]$  une image originale et  $B[x, y]$  l'image filtrée tel que :

$$B(x, y) = G(x, y) * A(x, y) \tag{2.1}$$

**d) Filtre médian (non linéaire)**

Souvent, les filtres moyenneurs ont tendance à rendre l'image floue et donc à perdre de l'information sur les contours caractérisés par des fortes variations d'intensité. Pour diminuer cet effet, on ne moyenne plus sur le voisinage mais on prend la valeur médiane sur ce voisinage : on parle d'un filtre médian.

### **II.3.3 Segmentation d'images**

La segmentation d'image est une opération de traitement d'images qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis. Elle consiste à diviser l'ensemble de pixels en régions connexes, homogènes et bien séparées. Ces régions possèdent une certaine uniformité pour une ou plusieurs caractéristiques (intensité, couleur, texture, ...) et sont différentes pour au moins une de ses caractéristiques des régions voisines. En analyse d'images, on distingue les traitements de bas niveau et les traitements de haut Niveau. Cette distinction est liée au contenu sémantique des entités traitées et extraites de l'image. Les traitements de bas niveau opèrent en général, sur les grandeurs calculées à partir des valeurs attachées à chaque point de l'image sans faire nécessairement la liaison avec la réalité qu'elles représentent. Par exemple, la détection des contours est un traitement de bas niveau qui est « effectué sans comprendre » l'image.

Le contour détecté peut très bien ne pas correspondre à un bord d'objet dans la réalité et ne présenter aucun intérêt, ainsi les traitements de bas niveau opèrent plutôt sur des données de nature numérique.

A l'opposé, les traitements de haut niveau s'appliquent à des entités de nature symbolique associées à une représentation de la réalité extraite de l'image. Ils sont relatifs à l'interprétation et à la compréhension de l'image et sont exprimés avec des mots du vocabulaire de l'application. Par exemple, des zones d'une image aérienne peuvent être caractérisées par leur forme (rectangulaire, linéique,), être étiquetées avec les termes : bâtiment, route, bosquet, ombre, etc. [52].

Plusieurs techniques sont proposées dans la littérature Ou chacune a ses avantages et ses inconvénients.

La segmentation est l'étape la plus importante car l'extraction de primitives, la détection d'une position d'un objet, ou la reconnaissance d'un objet dépendent fortement de la qualité de la segmentation. Son but est de faciliter l'extraction des éléments qui la composent. Nous allons détailler les différentes méthodes de segmentation dans le chapitre suivant.

### **II.3.4 Interprétation des images**

Cette étape est la plus décisive d'un système d'aide au diagnostic des tumeurs. Le rôle d'un système d'aide à l'interprétation d'images médicales est de fournir au médecin des informations symboliques de haut niveau sur le contenu de l'image. Dans un but d'aide au

diagnostic, ces informations doivent expliciter les différentes régions d'intérêt ainsi que les aspects pathologiques des structures présentes dans l'image. Du point de vue des traitements impliqués, le domaine de l'interprétation d'images médicales s'intègre dans le thème de recherche plus général du traitement de l'information visuelle [49]. En vision par ordinateur, il est généralement admis que l'analyse d'une image s'effectue par une série de procédures qui forment la chaîne de traitement de l'image.

## **II.4 Conclusion**

Le diagnostic de cancer du sein est souvent facile lorsqu'il se manifeste par une tumeur. Il est plus difficile dans les autres cas. Cependant ce diagnostic ne doit pas être retardé car la vie de la patiente en dépend.

Le système d'interprétation des images médicales et de mammographie en particulier est constitué de plusieurs étapes : (1) amélioration et filtrage des images, (2) segmentation et classification des images et (3) l'interprétation qui a pour rôle de fournir au médecin des informations symboliques de haut niveau sur le contenu de l'image afin de l'aider à sa prise de décision finale.

Dans ce chapitre, nous avons brièvement parlé des différentes étapes du système d'interprétation des images médicales. Le chapitre suivant sera consacré à l'étape de segmentation des images mammographiques qui est considérée comme l'une des étapes la plus importante du système d'interprétation et dont les résultats affectent également toutes les autres étapes.

## **CHAPITRE III :**

***Etat de l'art sur les méthodes  
de segmentation des images  
mammographiques***

## **III.1 Introduction**

La chaîne de traitement du système CAD implique une phase de prétraitement et une phase d'extraction de caractéristiques de chaque région comme la texture, la compacité, la surface, le périmètre, etc., une phase de segmentation et une dernière phase de classification qui fait appel à des techniques de l'intelligence artificielle telles que : les réseaux de neurones, la logique floue, les algorithmes génétiques...

La segmentation est l'étape la plus importante du système CAD, elle vise à décomposer complètement ou partiellement une image mammographique en régions qui possèdent une corrélation forte avec les objets d'intérêt (à détecter) dans l'image. Elle permet ensuite de visualiser et de quantifier ces objets. Un bon résultat de segmentation ne permet pas forcément une bonne interprétation, mais nous ne pouvons pas obtenir une bonne interprétation à partir d'un mauvais résultat de segmentation [39]. Dans ce chapitre, nous allons esquisser un état de l'art des méthodes de segmentation d'images qui sont largement utilisées pour la détection des pathologies mammaires. Cette Revue des différentes techniques et approches n'a pas pour but d'être exhaustive mais de montrer la grande diversité qui existe au sein de ce vaste domaine.

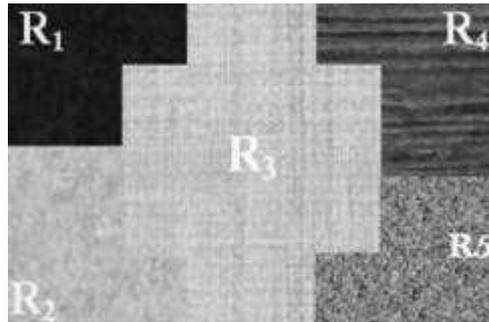
## **III.2 Principe de la segmentation**

La segmentation selon le dictionnaire Larousse, est la division d'une unité anatomique en plusieurs éléments. En traitement d'image, c'est le processus de partitionnement d'une image numérique en plusieurs régions ou en un ensemble de pixels.

Pavlidis [19] l'a définie de la manière suivante : « Etant donné un critère d'uniformité, la segmentation vis-à-vis de ce critère est une partition de l'image en régions connexes, homogènes et uniformes, tel que la réunion de deux sous-ensembles quelconques adjacents, ne soit jamais uniforme ». Mathématiquement, elle peut être modélisée de la façon suivante :

Soit une image  $I$ , la segmentation consiste à la diviser en  $K$  régions  $R_i$ , avec  $1 \leq i \leq K$ . Nous avons les relations suivantes (voir figure 3.1) :

1.  $\forall i, R_i \neq \emptyset$
2.  $\forall i, j; i \neq j, R_i \cap R_j = \emptyset$
3.  $I = \bigcup_i R_i$



**Figure 3.1 - Segmentation d'image en 5 régions.**

En d'autres termes, les conditions suivantes doivent être vérifiées :

- ❖ Chaque pixel doit appartenir à une classe.
- ❖ Les régions doivent être homogènes.
- ❖ Les régions doivent être deux à deux disjointes.

La segmentation permet une extraction de données en regroupant les pixels formant l'image à classifier de telle sorte que les pixels appartenant à une classe soient plus similaires entre eux que ceux des classes différentes. On affecte à chaque entité de la scène traitée une étiquette indiquant son appartenance à une classe particulière. L'entité utilisée caractérise généralement un pixel ou un ensemble de pixel, tandis que l'étiquette constitue un thème choisi par l'utilisateur.

On regroupe généralement les algorithmes de segmentation en trois grandes classes [30] :

- ❖ la segmentation en régions.
- ❖ la segmentation en contours.
- ❖ la segmentation hybride.

L'objectif de la détection de contours est la détermination des points de l'image séparant des zones de caractéristiques différentes. Chaque méthode de détection de contours est justifiée de manière plus ou moins explicite par la définition du type de transitions recherchées. De manière duale, le but de la segmentation en régions est de partitionner l'image en ensembles de points connexes possédant des propriétés d'homogénéité.

### **III.3 Objectif de la segmentation**

L'objectif de la segmentation d'image est de partitionner l'image en zones d'intérêt correspondant à des objectifs de la scène d'où elle est issue. Elle permet d'obtenir une représentation de l'information contenue dans l'image et constitue un premier pas vers son

interprétation. Dans le cas de la segmentation d'images médicales, l'objectif est [11] d'étudier les structures anatomiques.

- ❖ Identifier les régions d'intérêts, localisation de la tumeur, lésions et d'autres anomalies.
- ❖ Mesure du volume des tissus pour mesurer la croissance de la tumeur.
- ❖ Aide à la planification de traitement préalable à la radiologie, dans le calcul de la dose du rayonnement.

La segmentation est basée sur [45] :

- ❖ Un ensemble d'entités,
- ❖ Un ensemble d'attributs caractérisant ces entités,
- ❖ Des relations topologiques entre ces entités,
- ❖ Des attributs relationnels

On cherche des partitions de données ayant des propriétés intrinsèques par rapport aux attributs et aux relations topologiques (la 4 connexité et la 8 connexité).

On peut donc caractériser un problème de segmentation d'images par un ensemble de critères d'homogénéité déterminant les propriétés des partitions de l'image que l'on cherche. Le critère qui définit l'homogénéité est donc un point déterminant des performances de la segmentation. Les principaux critères utilisés sont le niveau de gris, la couleur pour les images couleur et la texture [54]. Dans la littérature [54], de nombreuses techniques de segmentation sont proposées, mais la plupart d'entre elles ont besoin de plusieurs paramètres, dont le réglage exige souvent l'expertise humaine.

L'objectif de la segmentation d'image est de partitionner l'image en zones d'intérêt correspondant à des objectifs de la scène d'où elle est issue. Elle permet d'obtenir une représentation de l'information contenue dans l'image et constitue un premier pas vers son interprétation. Dans le cas de la segmentation d'images médicales, l'objectif est [51]:

- ❖ D'étudier les structures anatomiques.
- ❖ Identifier les régions d'intérêts, localisation de la tumeur, lésions et d'autres anomalies.
- ❖ Mesure du volume des tissus pour mesurer la croissance de la tumeur.
- ❖ Aide à la planification de traitement préalable à la radiologie, dans le calcul de la dose du rayonnement.

La segmentation est basée sur [30]:

- ❖ Un ensemble d'entités.
- ❖ un ensemble d'attributs caractérisant ces entités.
- ❖ des relations topologiques entre ces entités.
- ❖ des attributs relationnels.

On cherche des partitions de données ayant des propriétés intrinsèques par rapport aux attributs et aux relations topologiques (la 4 connexité et la 8 connexité).

On peut donc caractériser un problème de segmentation d'images par un ensemble de critères d'homogénéité déterminant les propriétés des partitions de l'image que l'on cherche. Le critère qui définit l'homogénéité est donc un point déterminant des performances de la segmentation. Les principaux critères utilisés sont le niveau de gris, la couleur pour les images couleur et la texture [30].

### III.4 Les différentes approches de segmentation

Il existe de nombreuses méthodes de segmentation [5]. Dans cette section, nous présenterons les diverses techniques connues en les organisant selon l'approche qui les régit. Ainsi, nous avons retenu six approches. Cette classification est représentée dans la figure 3.2 [33].

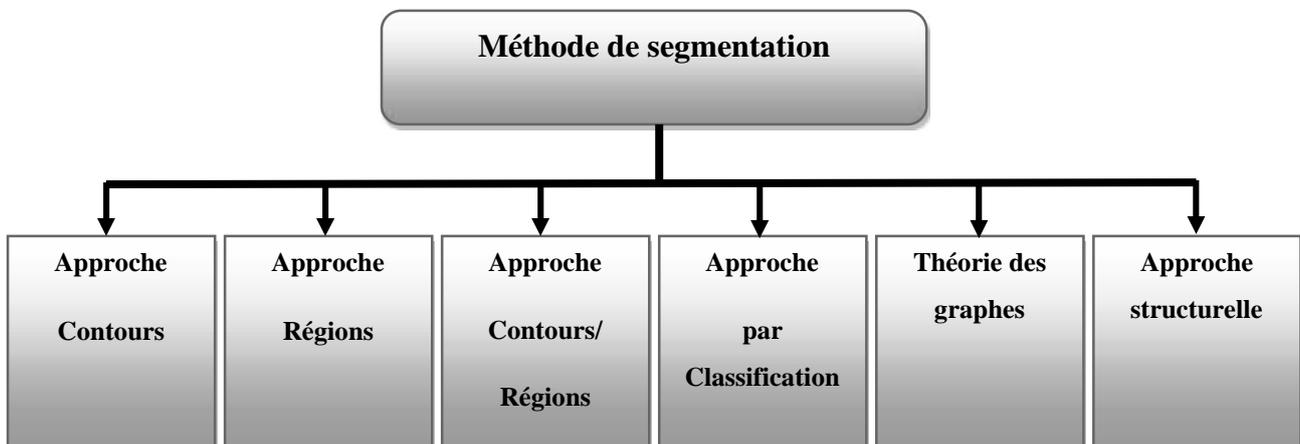


Figure 3.2 - Classification des différentes méthodes de segmentation.

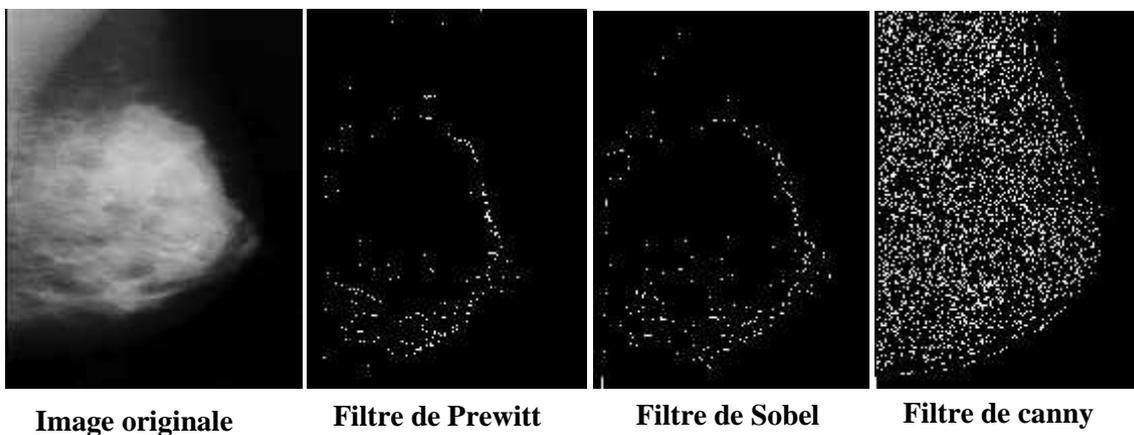
#### III.4.1 Segmentation par approche contour

Un contour est un ensemble de pixels formant une frontière entre deux ou plusieurs régions voisines. L'épaisseur d'un contour est d'un ou plusieurs pixels et il est défini par une

variation « brusque ». La détection des contours dans une image est un problème très étudié dans le domaine du traitement et de l'analyse de l'image. De façon générale, un contour est défini comme le lieu de variation significative de l'information "niveau de gris". Par conséquent, trouver les contours dans une image revient à évaluer la variation du niveau de gris en chaque pixel de l'image [40]. Les méthodes d'extraction de contours s'appuient sur la détection des discontinuités dans l'image et peuvent être divisées en trois classes :

- ❖ Les méthodes dérivatives.
- ❖ Les méthodes analytiques.
- ❖ Les méthodes basées sur les contours actifs.

Nous pouvons citer les approches se basant sur la méthode dérivative comme l'opération de gradient, l'opération Laplacien et les différents filtres à savoir : filtre de sobel, prewitt et roberts ou bien des approches analytiques comme le filtre de canny. Ce genre de techniques est peu exploitables car elles peuvent donner leurs contours non fermés et restent sensible au bruit. La troisième approche pour la détection de contours est basée sur les contours actifs (snakes). Cette méthode sera bien détaillée par la suite. La figure 3.3 illustre les résultats obtenus avec quelques filtres avec l'image de mammographie [40].



**Figure 3.3 - Illustration de détection de contours par les différents filtres [40].**

On distingue trois types de contours simples (Figure 3.4):

- ❖ **Marche d'escalier** : le contour est net (contour idéal);
- ❖ **Rampe** : le contour est plus flou ;
- ❖ **Toit** : il s'agit d'une ligne sur un fond uniforme ;

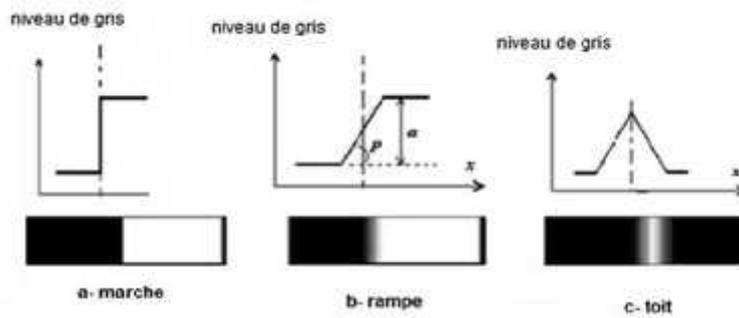


Figure 3.4 - Exemples de contours, (a) Contour idéal, (b) rampe, (c) contour réel.

Les approches contours se caractérisent par le fait qu'elles ne prennent en compte que des informations sur le contour des objets. Les premiers modèles de segmentation s'appuient sur des détections de changement rapide d'intensité lumineuse ou de couleur, c'est-à-dire les contours de l'image. L'application de détecteurs de contours sous la forme de filtres dérivatifs permet d'obtenir les contours des objets présents dans la scène. Nous pouvons citer les approches se basant sur les différences finies comme l'opérateur de gradient, l'opérateur laplacien, les filtres de Sobel, Prewitt, Roberts ou bien des approches reposant sur des critères d'optimalité comme les filtres de Canny-Deriche. Mais ce genre de techniques est peu exploitable car elles donnent des contours non fermés, bruités ou des contours non détectés [24].

### III.4.1.1 Filtre Sobel

Le filtre de Sobel non-linéaire est un passe-haut destiné à extraire le contour extérieur des objets. Il met également en évidence les variations d'intensité le long de l'axe vertical et horizontal. Chaque pixel se voit affecter une valeur égale à la valeur maximum des deux gradients verticaux et horizontaux de la convolution de Prewitt classique contrairement au Prewitt, le filtre de Sobel assigne une valeur plus importante aux voisins horizontaux et verticaux du pixel central.

Les deux filtres extraient les contours des objets. En raison des différences sur les noyaux utilisés, le Prewitt NL (non linéaire) a tendance à extraire des contours incurvés alors que le Sobel NL extrait des contours carrés. Ceci est notable sur les petits objets isolés. Le calcul de gradient est mené par l'intermédiaire des deux masques  $h_x$  et  $h_y$ . Le masque  $h_x$  se déduit du premier masque ( $h_x$ ) par une rotation de  $\pi/2$ .

$$h_x = \begin{pmatrix} -1 & 0 & 1 \\ -c & 0 & c \\ -1 & 0 & 1 \end{pmatrix} \quad h_y = \begin{pmatrix} -1 & -c & -1 \\ 0 & 0 & 0 \\ 1 & c & 1 \end{pmatrix}. \quad (3.1)$$

Un exemple d'application de ce filtre sur une image mammographique est montré sur la figure 3.5.

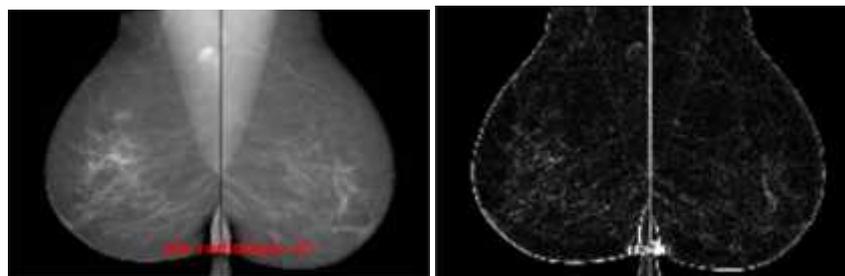


Image originale

Filtre de Sobel

**Figure 3.5 - Détection de contour par le filtre Sobel.**

### III.4.1.2 Canny-Dérivée

Canny a proposé une étude théorique de la détection de contour. Son étude s'est limitée au cas de la dimension 1, c'est à dire la détection des variations dans un signal bruité. Il est le premier à avoir formalisé trois critères que doit valider un détecteur de contour :

- ❖ Détection : robustesse au bruit.
- ❖ Localisation : précision de la localisation du point contour.
- ❖ Unicité : une seule réponse par contour.

A chaque critère est associée une formule mathématique. La maximisation de ces critères conduit à la résolution d'une équation différentielle dont la solution est le filtre  $f$ , qui permet la détection du contour, i.e. la position du contour correspond à :

$$\max (I * f) (x). \quad (3.2)$$

En fixant des conditions initiales, Canny a montré que cette solution générale pouvait être approximée par :

$$f(x) = -(x / \text{tau}^2) e^{-\frac{x^2}{2\text{tau}^2}} \quad (3.3)$$

C'est à dire la dérivée du filtre gaussien. Le passage à un espace à 2 dimensions (une image) se fait alors simplement car le filtre gaussien étant séparable, on a :

$$f(x,y) = f(x) \cdot f(y). \quad (3.4)$$

Le calcul du gradient sur une image se ramène donc à deux balayages (un en ligne et un en colonne) [7]. Partant d'autres conditions initiales, Dérêche a proposé un filtre différent dont la forme simplifiée est :

$$f(x) = (sx)e^{-\alpha |x|} \quad (3.5)$$

Un exemple d'application de ce filtre sur une image mammographique est montré sur la figure 3.6 [50].

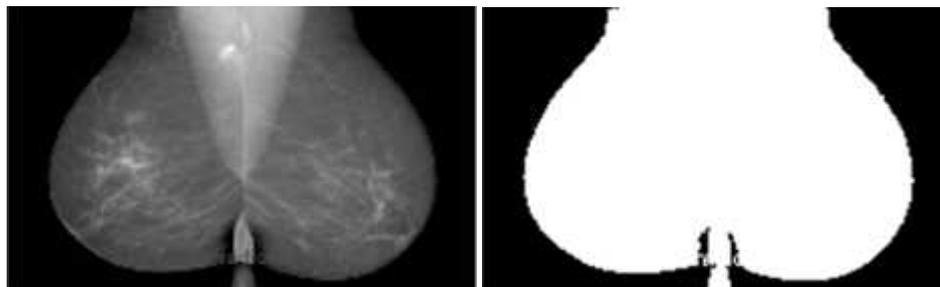


Image originale

filtre de canny dériche

**Figure 3.6 - Détection de contour par Canny-Dérêche.**

### III.4.1.3 Les Modèles déformables

Pour surmonter certaines limitations des méthodes de segmentation par approches région, des méthodes par approches contour sont utilisées pour chercher des frontières explicites ou implicites entre les régions correspondant aux différents types de tissus. Dans les dernières années, les modèles déformables, une des méthodes les plus populaires des méthodes de segmentation contour, ont été largement utilisés dans la segmentation d'images. L'idée derrière les modèles déformables est assez simple. L'utilisateur détermine une estimation initiale pour le contour qui est ensuite déformé par les forces dérivées de l'image jusqu'à la délimitation des objets désirés.

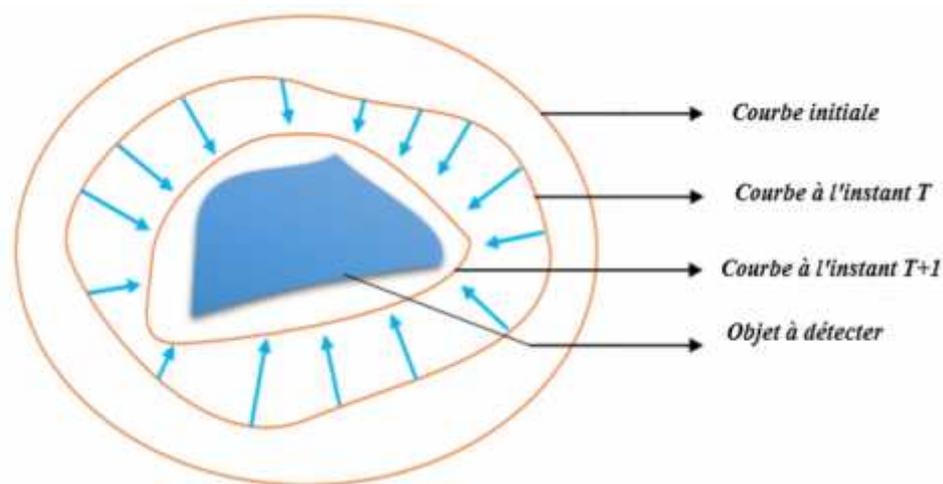
#### a) Contour actif (snake)

La segmentation par contours actifs, comme l'ensemble des méthodes de segmentation a pour but de partitionner une image en différentes régions d'intérêt présentant une similarité suffisante au sens d'un critère donné (statistique par exemple) et ce au moyen d'une ou plusieurs courbes dont la ou les position(s) évolue(nt) dans le temps en accord avec le critère

de segmentation choisi. Ces courbes évoluent d'une forme initiale vers les frontières d'un objet d'intérêt, sous l'action d'une force.

Cette force est basée sur la notion d'énergie interne et externe, le but étant de minimiser l'énergie totale présente le long de la courbe. Des contraintes permettent de conserver une courbe lisse avec des points équidistants tout en laissant un certain champ libre pour les déformations.

Le principe de cette méthode est de déplacer les points pour les rapprocher des zones de fort gradient tout en conservant les caractéristiques initiales comme la courbure, changement de forme du contour ou d'autre contraintes liées au positionnement des points [16] (voir figure 3.7).



**Figure 3.7 - Principe des contours actifs.**

Un contour actif (snake) est un ensemble de points qu'on va tenter de déplacer pour leur faire épouser une forme. Il s'agit d'une technique d'extraction de données utilisée en traitement d'images. La courbe représentant le contour est définie sous la forme paramétrique de l'équation suivante :

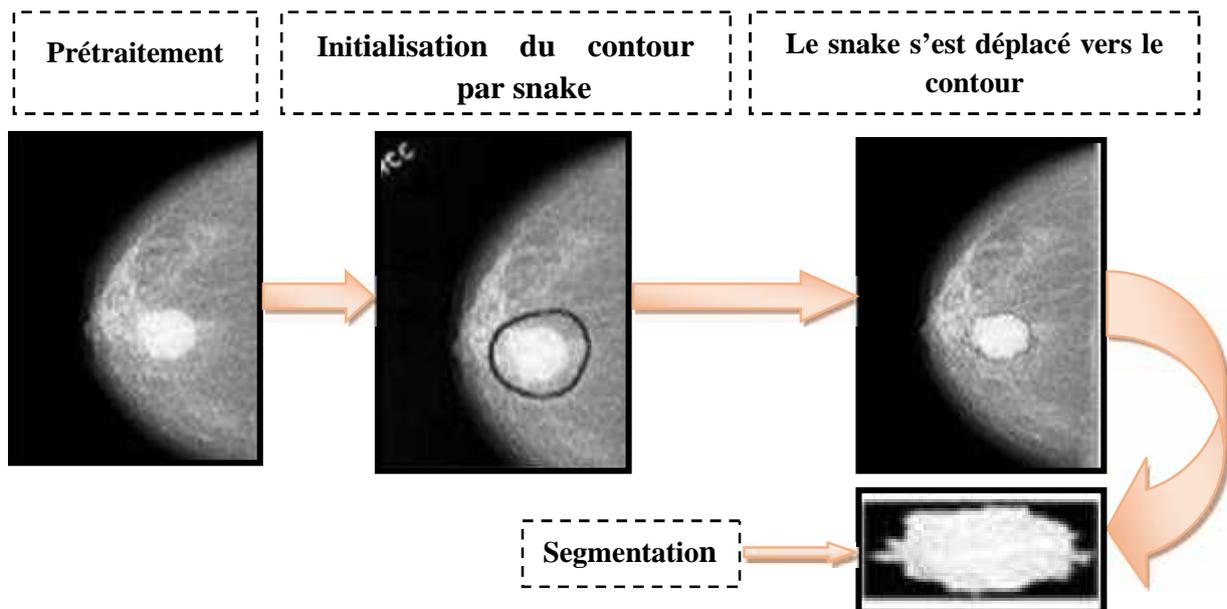
$$C = (x(s,t), y(s,t)) \quad (3.6)$$

L'idée de cette méthode est de déplacer les points pour les rapprocher des zones de fort gradient tout en conservant des caractéristiques comme la courbure du contour ou la répartition des points sur le contour ou d'autres contraintes liées à la disposition des points. Au démarrage de l'algorithme, le contour est disposé uniformément autour de l'objet à détecter puis il va se rétracter pour en épouser au mieux ses formes. De la même manière, un contour actif peut aussi se dilater et tenter de remplir une forme. Il sera alors situé à l'intérieur de

celle-ci au démarrage de l'algorithme. A chaque itération, l'algorithme va tenter de trouver un meilleur positionnement pour le contour pour minimiser ses dérivés par rapport aux contraintes utilisées. L'algorithme s'arrêtera lorsqu'il ne sera plus possible d'améliorer le positionnement ou simplement quand le nombre maximum d'itérations aura été atteint. On utilise les notions d'énergies interne et externe pour caractériser respectivement la forme du contour et tous les éléments qui lui sont propres et, le positionnement du contour sur l'image en tenant compte des lignes de gradient.

**b) Mise en œuvre du contour actif**

Le contour actif évolue pour minimiser une énergie globale. Pour cela, une méthode simple et intuitive est d'observer les pixels voisins immédiats de chaque point du contour pour déterminer pour chacun d'eux l'énergie globale du snake. Chaque meilleur voisin devenant un point du contour. Il est nécessaire que le contour possède toujours suffisamment de points pour être sûr de bien calculer son énergie globale et en particulier son énergie externe qui sera plus précise en tenant compte davantage de points. C'est la raison pour laquelle il peut s'avérer pertinent de rajouter ou de supprimer des points à chaque itération si des contraintes ne sont pas suffisamment respectées. Par exemple, on pourra rajouter un point au snake si ses voisins sont trop éloignés. A l'inverse, on pourra supprimer un point s'il est trop près de ses voisins. La figure 3.8 présente le résultat de l'application du contour actif (snake) sur une image de mammographie.



**Figure 3.8 - Exemple de segmentation de masse par contours actifs (snakes) [41].**

### III.4.2 Segmentation par approche régions

La segmentation par région est une approche dans laquelle on cherche à construire des surfaces en regroupant des pixels voisins suivant un critère d'homogénéités [32]. Elle peut être définie comme une partition d'une image  $I$  en une ou plusieurs régions  $R_1, \dots, R_n$  telles que [16] :

$$I = \bigcup_{i=1}^n R_i \quad \text{et} \quad R_i \cap R_j = \emptyset \quad \text{pour } i \neq j \quad (3.7)$$

La segmentation par région crée un ensemble de régions qui ont les propriétés suivantes :

- ❖ La réunion de toutes les régions donne l'image entière.
- ❖ Les régions sont connexes, tous les pixels d'une même région sont jointifs.
- ❖ Tous les pixels d'une même région sont homogènes entre eux.

Les algorithmes de segmentation basés sur les régions nécessitent l'utilisation de structure de données particulières tel que :

**Le tétra-arbre (Quadtree) :** Le tétra-arbre est une structure de données très commune de par sa simplicité et d'autre part son faible temps de calcul. Il est une arborescence dont la racine est l'image toute entière et dont chaque nœud parent (sauf les nœuds terminaux) possède exactement 4 fils (voir figure 3.9). Il est défini de manière récursive : l'image est partagée d'abord en quatre blocs. À chacun de ces blocs est ensuite associé un nœud fils de la racine. Puis le processus de découpage en quatre quarts est itéré pour chacun des fils sans chevauchement des blocs. L'analyse récursive s'arrête lorsque chaque sous-bloc respecte un prédicat d'homogénéité. Après cette phase de division des petites régions, certains blocs adjacents présentent des caractéristiques calorimétriques identiques d'où la nécessité de les fusionner. Cette fusion s'arrête lorsqu'il n'existe plus de couple qui respecte le prédicat de fusion [38].

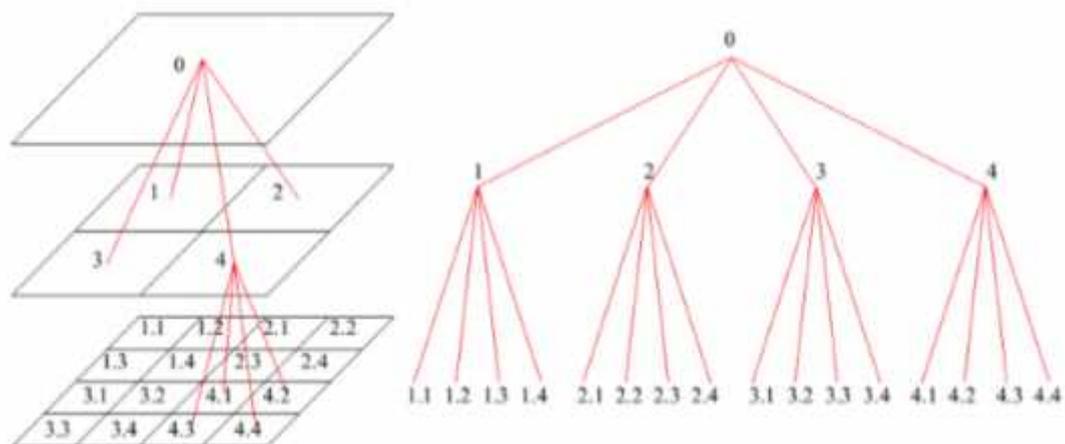


Figure 3.9 - Processus de division de l'image I utilisant le tétra-arbre.

**Le diagramme de Voronoï :** L'approche par diagramme de Voronoï s'inscrit dans un processus itératif de divisions et de fusions et peut être considéré comme une amélioration de la segmentation par analyse d'un quadtree. La phase de division n'est plus réalisée par un découpage en régions de forme carrée, mais par un découpage en polygones de Voronoï qui s'adaptent aux formes des régions présentes dans l'image.

Le diagramme de Voronoï génère une partition de l'image à partir de germes. À chaque germe est associée une région de Voronoï constituée par l'ensemble des pixels les plus proches de ce germe. La figure 3.10 montre un exemple de diagramme de Voronoï. Dans cette figure, les germes correspondent aux points rouges. Les pixels les plus proches d'un germe sont ceux qui sont inscrits dans le polygone noir centré sur le germe [38].

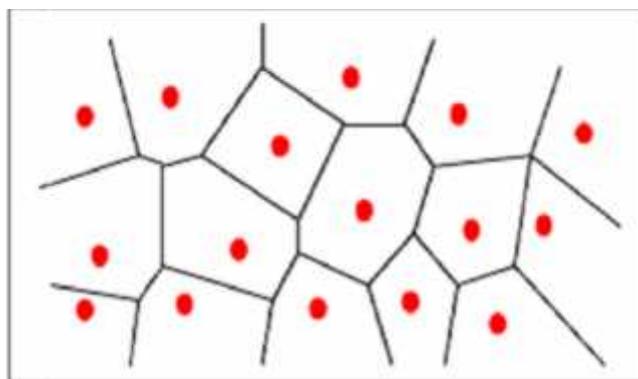
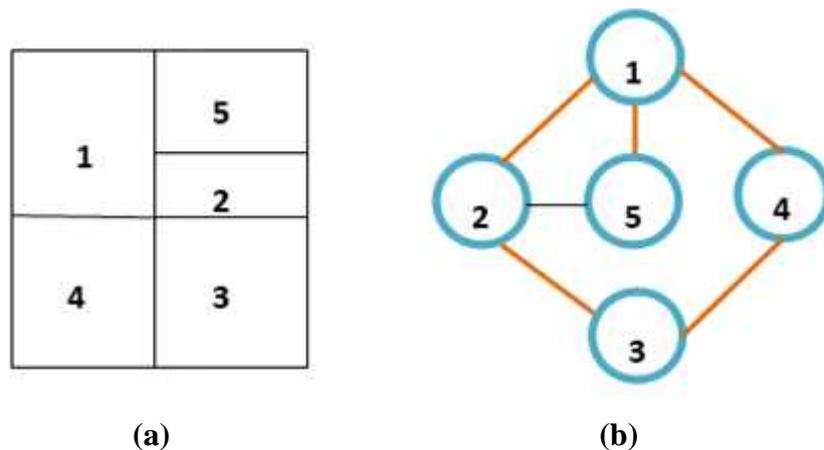


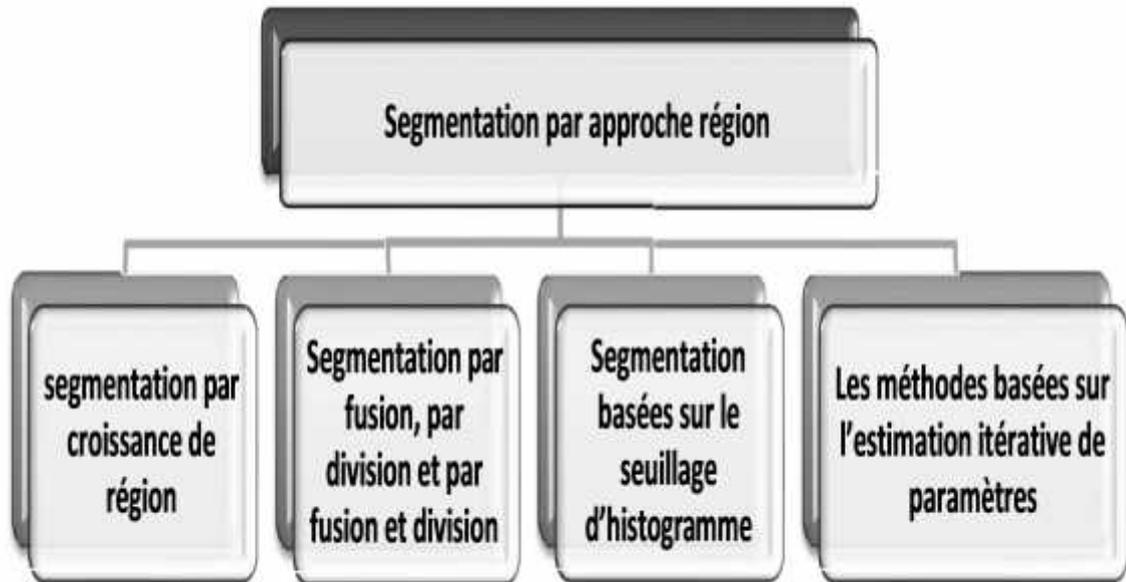
Figure 3.10 - Exemple de diagramme de Voronoï.

**Le graphe d'adjacence des régions :** La fusion des régions ne s'opère pas nécessairement après un algorithme de division de l'image, mais peut être accomplie après un algorithme de segmentation ayant provoqué une sur-segmentation. Ces approches de fusion se basent généralement sur l'analyse du graphe d'adjacence des régions ou RAG (Region Adjacency Graph). Les régions  $y$  sont représentées par les nœuds du graphe et l'information d'adjacence entre régions est symbolisée par les arêtes. Ce processus est représenté dans la figure 3.11. Sur ce graphe, des régions adjacentes vérifiant certains critères de similarité spatiale ou spectrale peuvent fusionner afin de résoudre le problème de la sur-segmentation [38].



**Figure 3.11 - Le graphe d'adjacence des régions, (a) image d'étiquettes et (b) graphe d'adjacence de régions.**

Cette approche se distingue par exemple, des segmentations par contours ou par seuillage dans lesquelles les régions créées ne possèdent pas toutes ces propriétés. Les méthodes les plus importantes pour cette approche sont citées dans la figure 3.12.



**Figure 3.12 - Les différentes méthodes de segmentation par approche région.**

### **III.4.2.1 Méthodes de croissance de régions**

Les méthodes de segmentation par croissance de régions sont basées sur l'utilisation de points amorces appelés « germes » et qui sont choisis soit, d'une façon manuelle soit, d'une façon automatique. Ces germes désignent les points ou régions de départ au sein de l'image à segmenter. En s'appuyant sur des mesures de similarité, chaque germe est comparé à son voisinage spatial immédiat (voir figure 3.13). En fonction de ces mesures, les régions similaires adjacentes sont fusionnées aux germes de départ donnant lieu à des régions de plus grandes tailles. En considérant les régions ainsi obtenues, le processus est ensuite itéré jusqu'à l'épuisement des régions susceptibles d'être fusionnées [10].

Ce type de segmentation consiste à faire croître des régions en y ajoutant successivement les pixels adjacents qui satisfont un critère d'homogénéité. La croissance s'arrête lorsque tous les pixels ont été traités. L'étape initiale consiste à sélectionner les germes des régions qui correspondent généralement à un pixel. Puis, les régions sont construites en y ajoutant successivement les pixels qui leur sont connexes et qui vérifient un critère de similarité calorimétrique. La plupart des algorithmes de croissance de régions sont régis par un parcours de l'image selon le balayage de gauche à droite et de haut en bas.

L'avantage de la croissance de régions est de préserver la forme de chaque région de l'image. Cependant, une mauvaise sélection des germes ou un choix du critère de similarité mal adapté peut entraîner des phénomènes de sous-segmentation ou de sur-segmentation

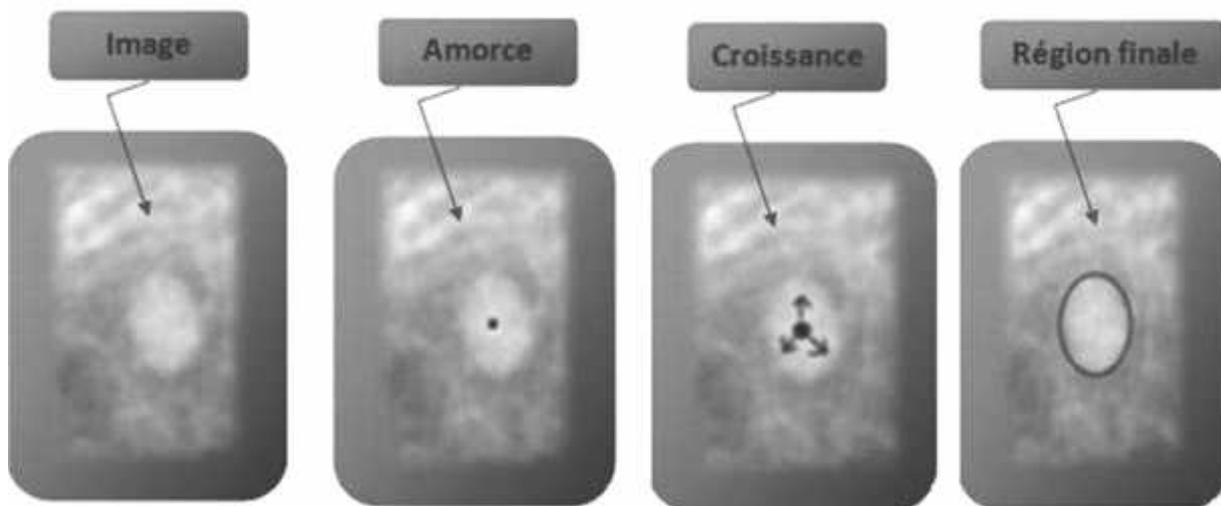
Pour pallier à ces erreurs, des paramètres ont été rajoutés :

- ❖ Prise en compte de la similarité de deux régions.
- ❖ Faire croître plusieurs régions en même temps [16].

D'une façon générale, les méthodes de croissance de régions offrent l'avantage de fournir des contours bien délimités pour les régions obtenues. Néanmoins, la localisation de ces contours souffre d'imprécisions importantes car il s'agit d'une croissance basée régions et non pixélique [18].

Les inconvénients de cette méthode sont :

- ❖ Méthode récursive avec risques de débordements (piles).
- ❖ Influence de la position initiale des germes.
- ❖ Choix des conditions d'arrêt de croissance des régions.
- ❖ Temps de calcul important.
- ❖ Tendance des algorithmes à trouver un nombre trop important de régions par rapport au nombre d'objets présents dans l'image.
- ❖ Méthode sensible au bruit.



**Figure 3.13 - Segmentation par le principe de la croissance de régions [2].**

### **III.4.2.2 Approche fonctionnant par fusion et division de région**

#### **a) Segmentation par division de région (Split)**

Ce sont des approches qui opèrent en découpant l'image originale en régions plus petites suivant un critère d'hétérogénéité. Le découpage s'arrête quand il n'y a plus de régions inhomogènes [8]. Le principe de cette technique est de considérer l'image elle-même comme région initiale, qui par la suite est divisée en régions. Le processus de division est réitéré sur

chaque nouvelle région (issue de la division) jusqu'à l'obtention des classes homogènes [6].

L'approche de segmentation par division de régions consiste à diviser l'image originale en régions homogènes au sens d'un critère donné. Ce processus est récursif et considère que la région initiale correspond à l'image à analyser. Si une région ne respecte pas un prédicat d'homogénéité, elle est divisée en quatre sous régions de taille égale. Chaque sous-région est ensuite analysée. L'algorithme récursif s'arrête lorsque toutes les régions respectent le prédicat D'homogénéité. A cause des divisions en quatre régions, cette méthode est plutôt adaptée à des images carrées ayant un nombre de lignes et de colonnes égal à une puissance de deux et dans lesquelles les régions sont de forme rectangulaire. D'autre part, cette méthode a tendance à faire apparaître des effets de bloc dû aux fautes qu'elle peut donner souvent naissance à une segmentation trop fine [16].

**b) Segmentation par fusion de régions (Merge)**

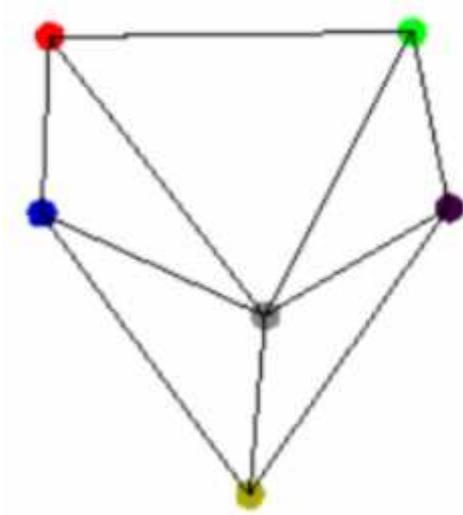
Les techniques de réunion (région merging) sont des méthodes ascendantes où tous les pixels sont visités. Ces méthodes opèrent en regroupant des pixels ou des groupes de pixels selon un critère d'homogénéité pour obtenir un ensemble de régions homogènes [3]. Elles suivent une hiérarchie arborescente. Plusieurs règles de regroupement ont été proposées dans la littérature. Certaines de ces règles mettent en jeu :

- ❖ Des propriétés statistiques telles que la moyenne ou la variance des niveaux de gris des régions, le gradient moyen des frontières des régions, le contraste maximum des régions, ou d'autres statistiques locales qui expriment l'état de la surface des régions ;
- ❖ Des propriétés géométriques ou morphologiques telle que l'élongation ou la compacité des régions. Deux régions sont regroupées si par exemple un facteur de forme est conservé ou amélioré après leur fusion.

Les inconvénients de cette méthode se situent à trois niveaux :

- ❖ Cette méthode dépend du critère de fusion qui peut influencer sur le résultat final de la segmentation.
- ❖ Elle peut introduire l'effet de sous-segmentation.
- ❖ Le résultat final (image segmentée) dépend fortement du choix des régions de départ.
- ❖ L'approche de fusion de régions est principalement fondée sur l'analyse d'un graphe d'adjacence de régions (figure 3.14) qui analyse une image pré-segmentée, constituée d'un ensemble de régions. C'est une structure de données constituée d'un graphe non-orienté dont chaque nœud représente une région et chaque arête représente une adjacence

entre deux régions. L'analyse du graphe d'adjacence de régions permet de fusionner des régions d'une image sur-segmentée. Le procédé consiste à fusionner deux nœuds reliés par une arête à condition qu'ils respectent un critère de fusion. Les méthodes d'analyse des graphes d'adjacence de régions se distinguent selon l'ordre de parcours des différents nœuds du graphe et selon les critères de fusion [16].



**Figure 3.14 - Graphe d'adjacence de régions.**

**c) Segmentation par division/fusion (*Split and Merge*)**

Proposées par Horowitz [25], elles réunissent l'ensemble des algorithmes utilisés dans les techniques précédemment citées (méthodes par fusion et méthodes par division).

La segmentation par division-fusion est caractérisée par la division au préalable de l'image en régions homogènes disjointes qui respectent des critères globaux. Puis elle consiste à fusionner des régions adjacentes qui vérifient des critères locaux (exemple le critère d'homogénéité calorimétrique). Parmi les structures de données permettant d'effectuer cette approche :

- ❖ Le tétra-arbre.
- ❖ Le diagramme de Voronoï.
- ❖ Le graphe d'adjacence des régions [38].

### **III.4.3 Approches Contours - Régions**

Plusieurs chercheurs s'accordent pour affirmer que la coopération entre les segmentations par régions et par contours contribue à une meilleure prise en compte des caractéristiques des

entités de l'image et, par conséquent, à une meilleure segmentation. En effet, les algorithmes combinant les techniques de segmentation basées sur les régions et celles basées sur les contours prennent avantage de la nature complémentaire de l'information sur la région et sur le contour. Ainsi, une segmentation par coopération régions-contours peut être exprimée comme une entraide entre ces deux concepts afin d'améliorer le résultat final. Elle exploite les avantages de ces deux types de segmentation pour aboutir à un résultat de segmentation plus précis et plus fidèle que celui obtenu à l'aide d'une seule technique. La segmentation d'images par coopération régions-contours peut se faire en fonction de trois approches

### III.4.3.1 Coopération séquentielle

Il s'agit d'une coopération (figure 3.15) dans laquelle l'une des techniques de segmentation (région ou contour) est réalisée en premier lieu ; son résultat va être exploité par l'autre technique pour renforcer la définition des critères ou des paramètres de la segmentation.

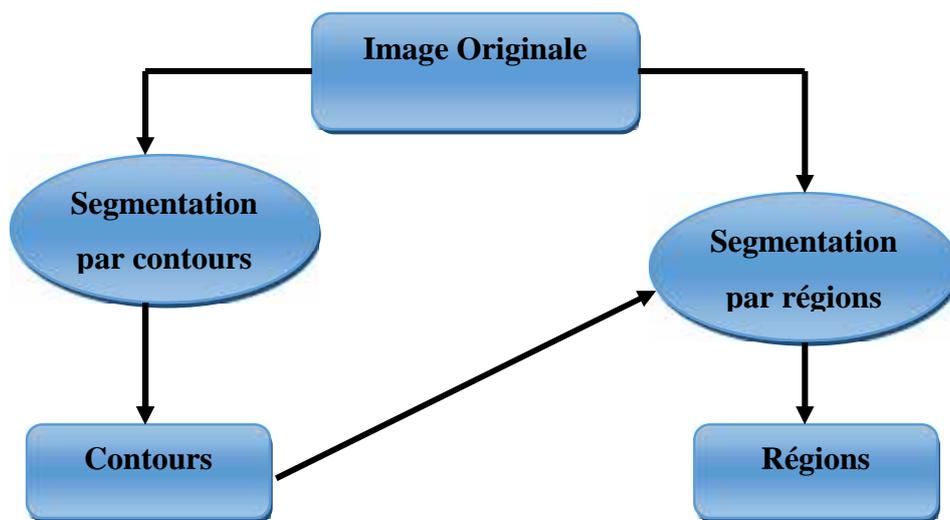


Figure 3.15 - Schéma de la coopération séquentielle.

### III.4.3.2 Coopération des résultats

Les deux types de la segmentation seront réalisés indépendamment ; la coopération (figure 3.16) concernera leurs résultats qui seront intégrés afin d'atteindre une meilleure segmentation

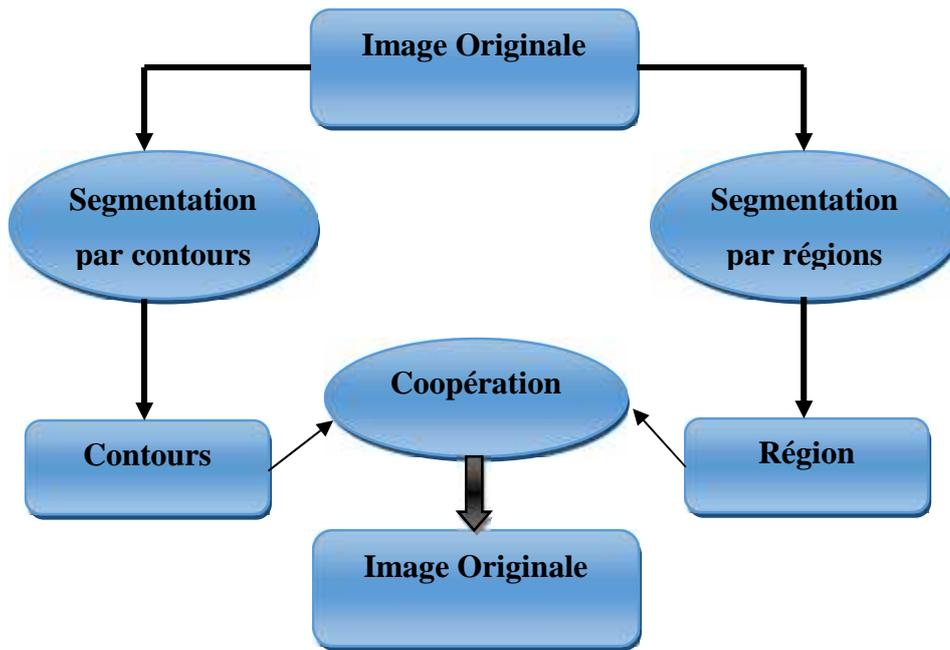


Figure 3.16 - Synoptique de la coopération des résultats.

### III.4.3.3 Coopération mutuelle

Les deux types de segmentations coopéreront mutuellement (figure 3.17) au cours de leur processus d'exécution. L'information échangée sert à aider la technique de segmentation dans la prise de décision dans le cas de lacunes ou d'informations insuffisantes. La coopération permet de prendre des décisions plus sûres et plus fiables.

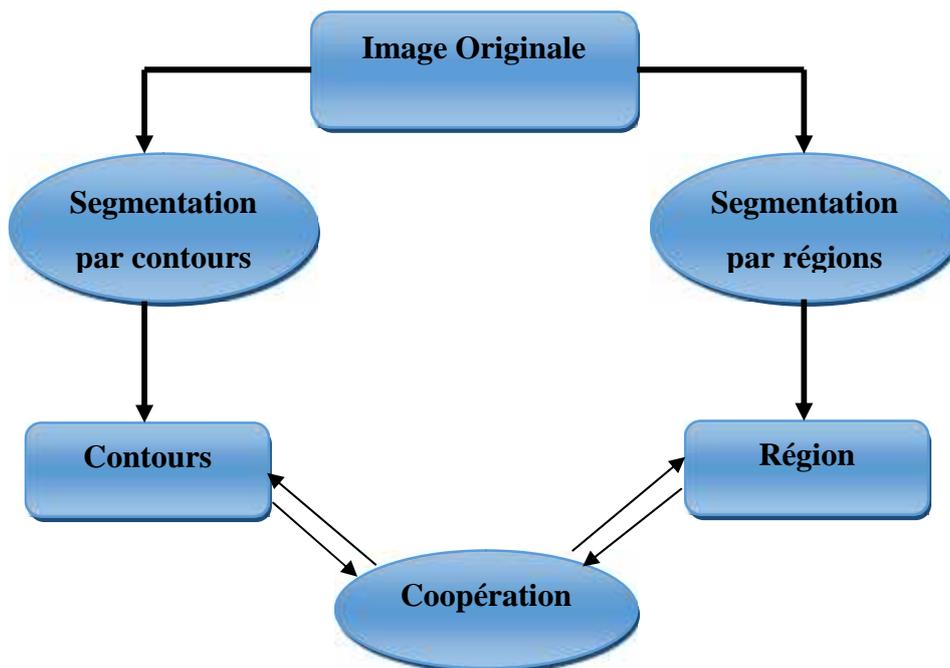


Figure 3.17 - Synoptique de la coopération mutuelle.

### III.4.5 Les méthodes basées sur l'estimation itérative de paramètres

Les méthodes de segmentation d'images comportent souvent des paramètres qui peuvent être fixés d'une façon empirique ou grâce à une méthode automatique d'estimation. Les méthodes basées sur l'estimation itérative des paramètres de segmentation consistent dans un premier temps, à appliquer une initialisation de ces paramètres et à réaliser la segmentation en se basant sur ces valeurs. Dans un second temps, le résultat de la segmentation est utilisé afin d'apprécier l'adéquation des paramètres et de réaliser par la suite un ajustement de leurs valeurs, puis de recommencer le processus de segmentation. Cette opération d'ajustement est ensuite répétée d'une façon itérative jusqu'à la stabilisation des valeurs obtenues pour les paramètres de segmentation. Les itérations de segmentation prennent fin en fonction d'un critère lié à la stabilisation de l'évolution des paramètres ou à la qualité de la segmentation obtenue. La figure 3.18 illustre le principe du processus lié à cette technique.

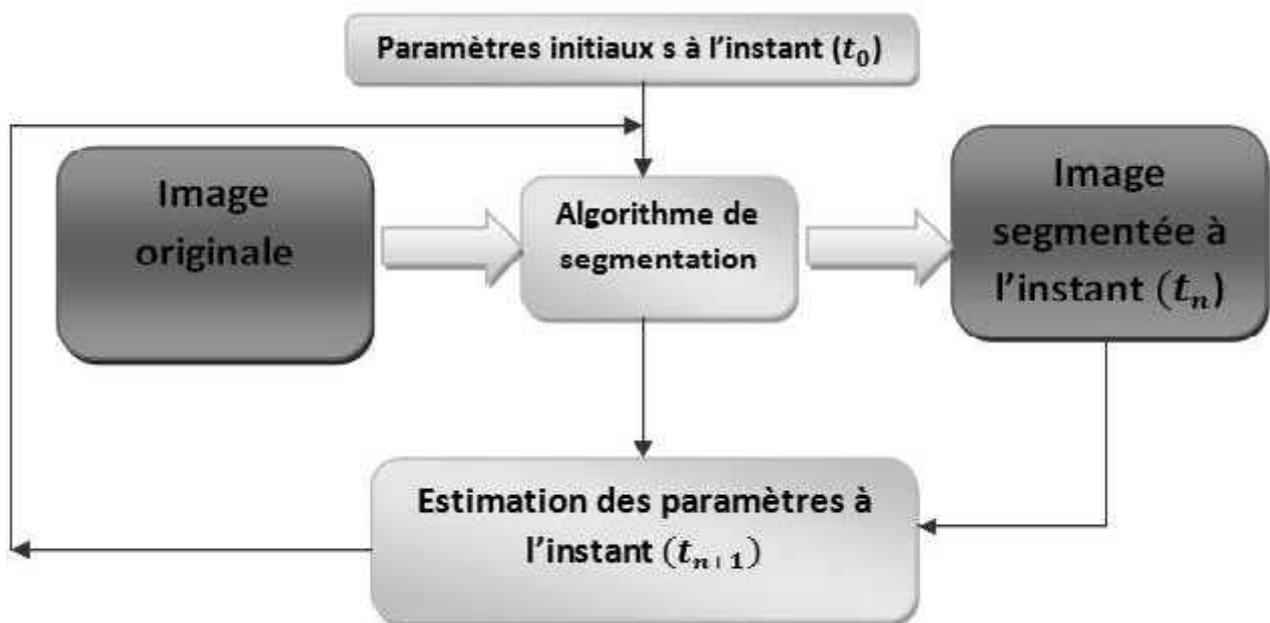


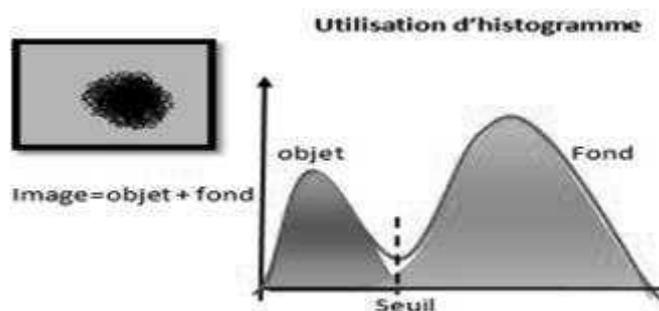
Figure 3.18 - Les méthodes de segmentation par estimation itérative de paramètres.

### III.4.6 la Segmentation basée sur la classification

La segmentation basée sur la classification travaille sur des histogrammes de l'image, on citera essentiellement la méthode du seuillage et k-means.

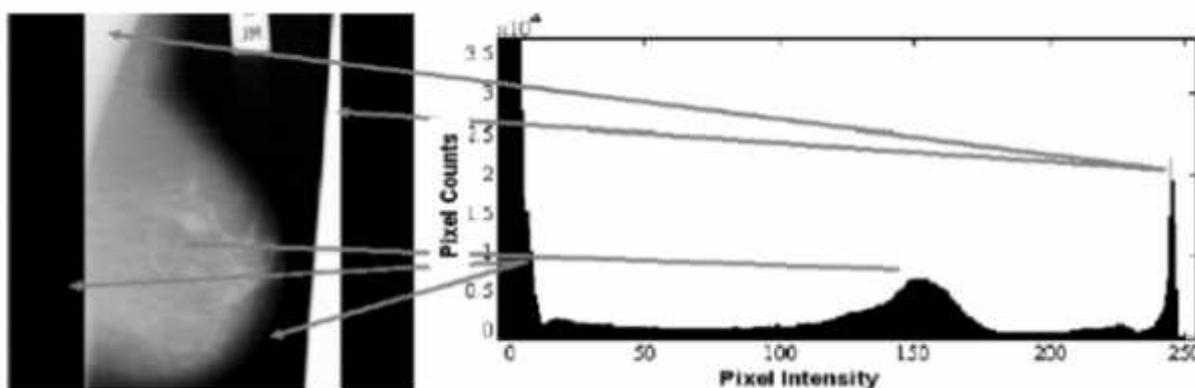
### III.4.6.1 Méthodes de segmentation basées sur le seuillage d'histogramme

Il s'agit de méthodes de base en segmentation d'image [21]. Le principe général du seuillage consiste à chercher une valeur appropriée de seuil puis de classifier tous les pixels de l'image selon la valeur de leurs niveaux de gris par rapport à ce seuil afin de séparer les régions d'intérêt et l'arrière-plan de l'image [20] (voir figure 3.19).



**Figure 3.19 - Seuillage par utilisation d'histogramme.**

Le seuil est une valeur numérique correspondant à un paramètre de l'image (exemple : le niveau de gris), et l'histogramme des niveaux de gris est utilisé afin d'en déduire le seuil [20]. L'image 3.20 présente un exemple de l'histogramme d'une image mammographique.



**Figure 3.20 - L'histogramme d'une image mammographie.**

Plusieurs méthodes de détermination de cette valeur de seuil ont été appliquées pour la segmentation d'images mammographiques. Certaines méthodes de détermination du seuil sont basées sur l'utilisation de paramètres autres que le niveau de gris, tels que l'entropie [21] ou l'entropie de Tsallis [21]. A titre d'exemple, Otsu [21] a formulé le problème de seuillage d'une image comme étant un problème d'analyse discriminante itérative qui permet de choisir

une valeur optimale du seuil. Le critère utilisé pour le choix du seuil est basé sur la maximisation d'une mesure statistique de séparation entre les classes. Dans tous les cas, le seuil obtenu via les méthodes citées ci-dessus sert finalement à la classification des pixels de l'image selon leurs niveaux de gris. D'une façon générale, les méthodes de seuillage peuvent être classées en deux catégories :

**a) La méthode de seuillage global [21] (un seuil pour toute l'image)**

Elle est considérée comme la méthode de référence dans le domaine du seuillage d'histogramme. Ces méthodes sont largement utilisées en segmentation d'images de mammographies afin de détecter les zones de tumeur ou des calcifications.

Le principe de cette méthode consiste à faire une séparation des pixels d'une image en deux classes  $c_1$ (fond),  $c_2$ (objet) à partir d'un seuil  $S$ . La classe « fond » regroupe tous les pixels ayant un niveau de gris inférieur au seuil  $S$  alors que la classe « objet » contient tous les pixels de niveau de gris supérieur à  $S$ . Dans la mammographie, la classe  $c_2$  représente les calcifications et la classe  $c_1$  représente des éléments de hautes fréquences causées par les tissus denses.

**b) Les méthodes de seuillage local (un seuil pour une portion de l'image)**

Ces méthodes visent à affiner localement la valeur du seuil pour mieux identifier les régions d'intérêts. La valeur du seuil est déterminée en se limitant aux informations contenues dans le voisinage local de chaque pixel. Ces méthodes ont souvent montré une meilleure efficacité de détection par rapport aux méthodes de seuillage global. Notons que les méthodes de seuillage local n'ont pas été seulement utilisées pour la segmentation d'image, mais ont aussi été exploitées comme une étape de prétraitement dédiée à d'autres algorithmes tels que ceux fondés sur les champs markoviens.

**c) Mise en œuvre du seuillage**

Le seuillage le plus simple est le seuillage global, il a pour principe : Soit  $p$  un pixel et  $V$  une fonction définissant le niveau de gris et  $S$  un seuil, alors le seuillage global est décrit par la formule suivante :

$$V(p) \begin{cases} 0 \text{ si } V(p) < S \\ 1 \text{ si } V(p) \geq S \end{cases} \quad (3.8)$$

Le résultat sera une image binaire. Le seuillage global est certes rapide, mais il ne donne pas de pixels contigus, et le bruit peut donner de faux éléments. Le seuillage local détermine pour chaque pixel un seuil en fonction de la luminosité de son voisinage. Le seuillage adaptatif se résume en la séparation de l'image en sous images et de traiter chacune avec son propre seuil (Le choix de la dimension des sous-images est critique).

Dans tous les cas, le problème majeur reste la détermination du seuil, il existe quelques méthodes de détermination telles que :

- ❖ La recherche des vallées de l'histogramme, en considérant que les vallées significatives correspondent aux pixels de la frontière des régions.
- ❖ La segmentation de l'histogramme (développée par Otsu) qui ne s'applique que dans le cas de segmentation de l'image en deux régions.

#### **III.4.6.2 la méthode de K-means**

Cette méthode consiste à rassembler les pixels en  $K$  groupes (clusters).  $K$  étant un paramètre préfixé qui détermine le nombre de régions. Et après avoir déterminé les paramètres (couleur par exemple) de toute région, chaque point est affecté au centre le plus proche, pour recalculer ensuite le paramètre de chaque région, jusqu'à ce qu'elles soient stables [20].

##### **a) Algorithme K-means**

L'algorithme des K-Means est l'une des techniques de clustering non supervisée les plus utilisées. La méthode consiste à placer aléatoirement dans l'espace  $K$  "centroïdes" afin de déterminer  $K$  clusters. On affecte à chacun de ces "centroïdes" les objets les plus proches, puis on calcule la position moyenne des objets associés aux "centroïdes" que l'on déplace en ce point. Les opérations d'affectation d'objet et de déplacement du centroïde sur la moyenne repérée jusqu'à ce que chaque centroïde ait atteint une position stable. Malgré sa simplicité, cet algorithme se révèle efficace. Toutefois il est nécessaire de prédéterminer le nombre de catégories et la position de départ des prototypes qui a un impact sur le découpage en classes. Les principales étapes de cet algorithme sont :

- (1) Choix aléatoire de la position initiale des  $K$  clusters.
- (2) Réaffecter les objets à un cluster suivant un critère de minimisation des distances (généralement selon une mesure de distance euclidienne).
- (3) Une fois tous les objets placés, recalculer les  $K$  cent-roïdes.

(4) Répéter les étapes 2 et 3 jusqu'à ce que plus aucune réaffectation ne soit faite.

Le principal avantage de cette méthode est que le nombre de régions est connu au préalable (il n'y aura pas de problème de sur-segmentation ou sous-segmentation), mais l'inconvénient réside en la difficulté de déterminer le nombre de clusters et l'incohérence des régions (deux objets éloignés peuvent appartenir à la même région).

### **III.4.7 Segmentation par Théorie des graphes**

L'histoire de la théorie des graphes débute peut-être avec les travaux d'Euler au XVIII<sup>e</sup> siècle et trouve son origine dans l'étude de certains problèmes, tels que celui des ponts de Königsberg. Dans les approches de segmentation utilisant la théorie des graphes, l'idée directrice est de créer un graphe à partir de l'image selon des procédés assez simples et de travailler sur ces graphes pouvant ainsi utiliser toute la panoplie d'outils développés dans le cadre de cette théorie. Les méthodes de segmentation basées sur la théorie des graphes évaluent les arêtes puisque ce sont-elles qui permettent de donner les caractéristiques de l'image au graphe. Un état de l'art important a été fait par Faliu Yi et Inkyu Moon [54] où ils ont décrit le concept des théories de graphes basé sur la segmentation d'images et ils l'ont classifié en trois différentes catégories : coupe de graphe accéléré (speed up-based graph cut), Coupe de graphe interactif (interactive-based graph cut) et Coupe de graphe en fonction de forme a priori (shape prior-based graph cut).

### **III.4.8 Segmentation par approche structurelle**

#### **III.4.8.1 La méthode de la ligne de partage des eaux**

Cet algorithme permet de décomposer l'image en régions homogènes. Une image en niveau de gris peut être perçue comme un relief topographique. Le niveau de gris d'un pixel de l'image est interprété comme son altitude dans le relief topographique. Un point est d'autant plus élevé dans le relief qu'il est clair dans l'image, les pixels sombres correspondent donc aux vallées et bassins du relief alors que les pixels clairs correspondent aux collines et lignes de crêtes (voir figure 3.21).

Une des difficultés à la mise en œuvre de cette analogie intuitive est qu'elle laisse beaucoup de liberté quant à sa formalisation. Il existe deux classes principales d'implémentation : l'une est basée sur un algorithme récursif d'immersion et une autre basée sur l'utilisation de fonctions de distances géodésiques.

Pour l'obtenir, il faut imaginer l'immersion d'un relief dans de l'eau (voir figure 3.22), en précisant que l'eau ne peut pénétrer dans les vallées que par ses minima. La ligne de partage des eaux est représentée par les points où deux lacs disjoints se rejoignent au cours de l'immersion. L'application de la LPE sur une image permet de détecter ses contours.

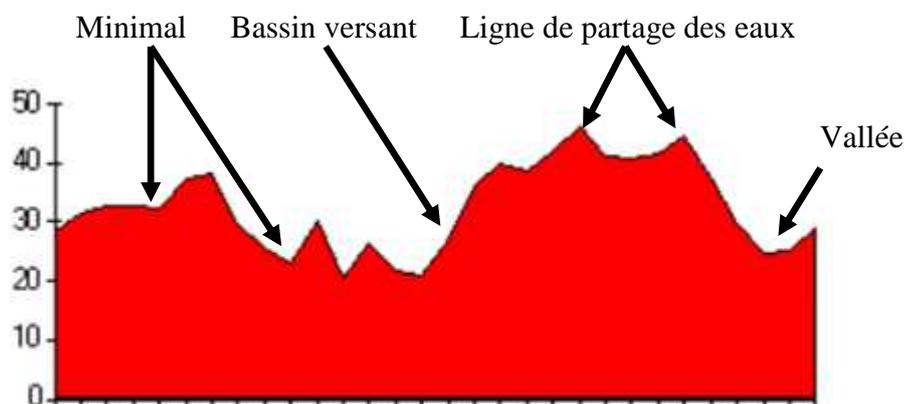


Figure 3.21 - Elément de relief.

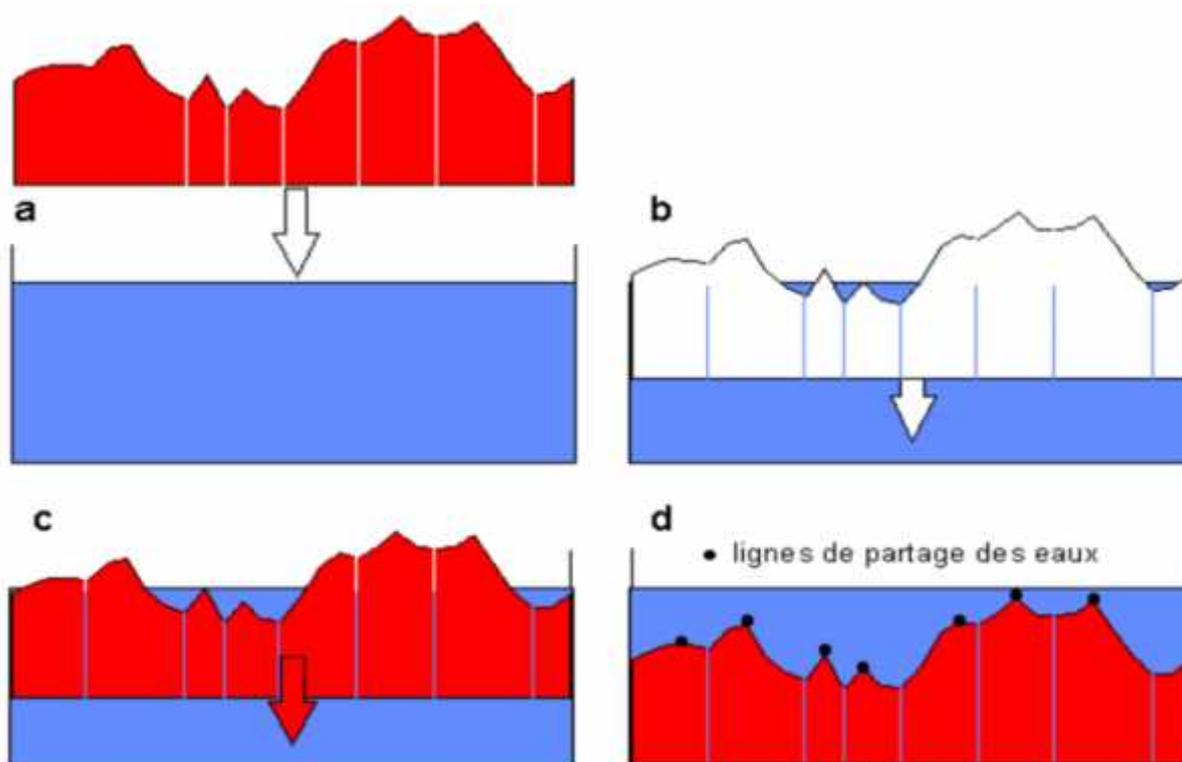


Figure 3.22 - Principe de la procédure d'immersion.

Cependant, lorsqu'il est appliqué à une image gradient construite à partir de l'image original, les minima correspondent alors aux zones homogènes et les contours de la LPE à la plus forte

transition trouver entre deux régions. On peut donc définir les contours des régions d'une image comme la LPE de l'image gradient [36].

La Ligne de Partage des Eaux (LPE) utilise une description des images en termes géographiques. Elle ne s'applique pas à l'image originale mais, à l'image de son gradient morphologique où, le niveau de gris de chaque point correspond à une altitude. Il est alors possible de définir la ligne de partage des eaux (figure 3.23) comme étant la crête formant la limite entre deux bassins versants.

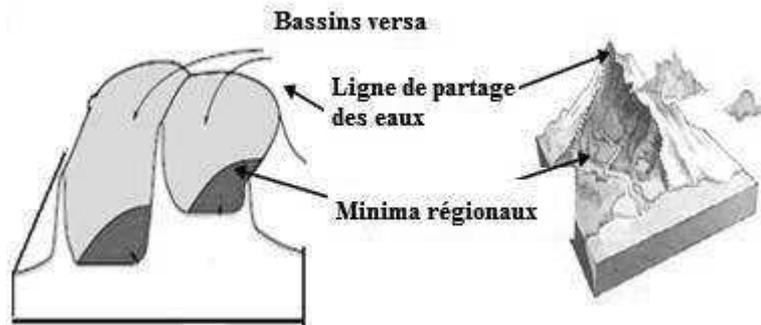


Figure 3.23 – Principe de la ligne de partage des eaux.

En d'autres termes, cette méthode consiste à « immerger » l'espace et à compter le nombre « d'îles ». De manière plus formalisée, on compte le nombre de composantes connexes (figure 3.24) aux ensembles  $E_h = \{x / f(x) \leq h\}$  où  $f$  représente la densité.

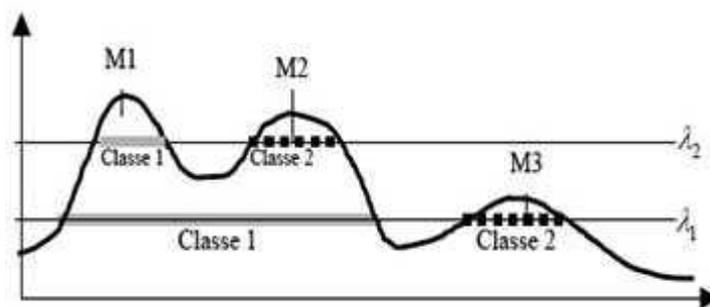


Figure 3.24 - Délimitation des composantes connexes.

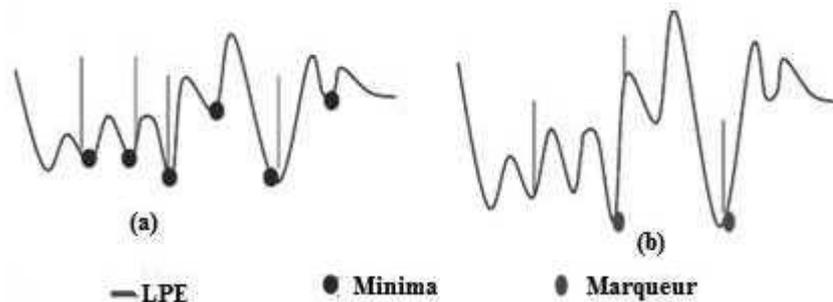
#### a) La LPE par l'inondation de relief

La LPE produite par inondation peut être formulée à l'aide d'une distance topographique, on prolonge la surface dans un lac (bassin versant), l'eau va passer par les trous (donc les minima locaux) de manière à ce que le niveau d'eau monte à vitesse constante et soit uniforme dans tout le bassin versant. Quand les eaux issues de deux minima différents se

rencontrent, on monte une digue pour qu'elles ne se mélangent pas. Lorsque la totalité de la surface topographique aura été engloutie, seuls les barrages émergeront, délimitant les bassins versants au nombre des minima locaux. A la fin de l'immersion, l'ensemble des tiges constituent un ensemble de contours fermés, qu'on appelle la ligne de partage des eaux.

**b) La LPE sous la contrainte des marqueurs**

Le principe de la modification d'homotopie du gradient est d'imposer les marqueurs des régions à segmenter comme minima du gradient en supprimant tous les autres minima indésirables qui sont à l'origine de toute sur segmentation. On inonde ensuite ce gradient à partir de tous les marqueurs. Une et une seule ligne de partage des eaux est alors présente entre chaque marqueur (voir figure 3.25). Elle a tendance à se situer sur le contour des objets à segmenter qui ont déjà été pré-détectés par le gradient. Par définition, on entend par marqueur une ou plusieurs composante(s) connexe(s) permettant de localiser les régions devant être segmentées dans l'image. Notons  $f_M$  l'image (binaire) des marqueurs définis comme



**Figure 3.25 - Processus de LPE par inondation, (a) avec contrainte et (b) sans contrainte.**

**c) Les points clefs de la segmentation par LPE**

Les algorithmes de segmentation par ligne de partage des eaux consistent en trois étapes [9] :

- ✓ Extraire des marqueurs des régions à segmenter.
- ✓ Déterminer l'image sur laquelle on calcule la LPE (image gradient le plus souvent).
- ✓ Calculer la LPE associée aux marqueurs.

La LPE s'avère être une technique puissante de segmentation à partir du moment où les étapes préparatoires (1 et 2) qui lui sont associées sont correctement effectuées.

## **III.5 Aperçu de quelques travaux pour la détection de pathologies mammographiques**

Les attributs pour la détection des lésions dans les images mammographiques proposés dans la littérature sont très variés et on retrouve des adaptations de toutes les méthodes du traitement d'images mettant en œuvre l'analyse multi résolution selon les travaux de Liu [35], la segmentation par croissance de régions adaptative dans les travaux de Rangayyan [47-46], la détection par la morphologie mathématique mise en exergue par la ligne de partage des eaux dans les travaux de Vachier [53], pour la détection des masses. L'extraction des calcifications par des transformations morphologiques se basant sur le chapeau haut de forme dans les travaux de Rick [48], ainsi que la détection de masses en utilisant les contours actifs par les travaux de Boujelben [12] etc.

### **III.5.1 Les travaux de A. RICK**

Rick [48] a proposé un prototype de classification d'images ou partie d'image par rapport à une base de connaissance se basant sur les systèmes d'apprentissage adaptatif combinant le traitement de l'incertitude (logique floue) et la variabilité de l'image. Cette variabilité se traduit en un ensemble de paramètres à prendre en considération lors de l'étape d'apprentissage et lors de la classification d'une nouvelle image (étape d'utilisation du prototype) [48]. Ce prototype se repose sur la modélisation des densités des histogrammes des attributs par une somme de fonctions paramétriques. Elle utilise des sommes de composantes gaussiennes pour lesquelles il existe des méthodes d'estimation (algorithme EM). Les liens entre les paramètres des composantes et leurs variations sont estimés à partir de la base d'apprentissage avec une démarche en deux étapes : le regroupement des composantes dans des groupes, pour établir un ordre unique de ces composantes pour toutes les images, et une régression linéaire des interactions entre paramètres. L'utilisation du système de classification sur de nouveaux exemples passe par l'introduction de cette connaissance dans l'algorithme d'estimation des paramètres PEM (Prototype-EM). Le schéma général de cette méthode est présenté sur la figure 3.26.

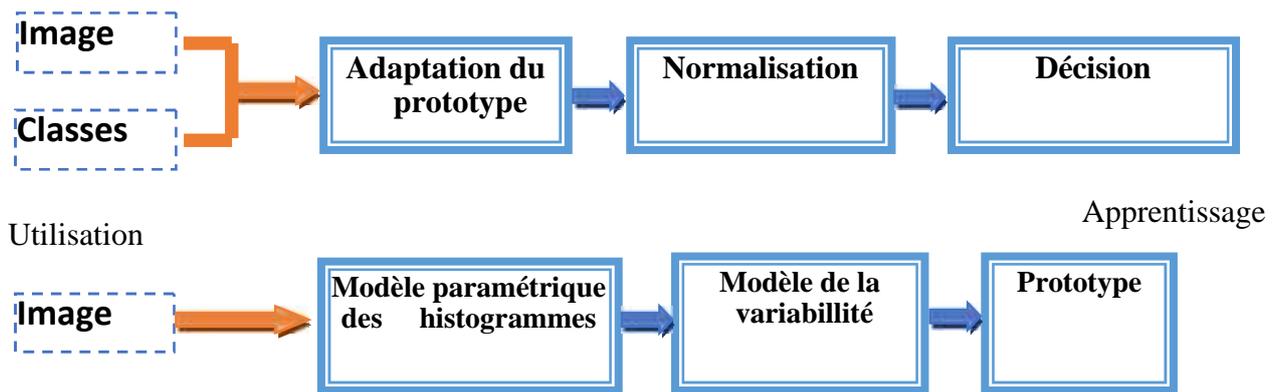


Figure 3.26 - Sous-ensembles flous adaptatifs [48].

Les résultats obtenus montrent l'utilité des fonctions d'appartenance adaptatives dans le traitement d'images pour la classification, et aussi la possibilité de calculer l'estimation des variabilités entre les images. Néanmoins, le prototype se base sur le choix des paramètres de modélisation des histogrammes. Il faut donc trouver les bonnes fonctions de paramétrage ainsi que le nombre de composantes idéal pour décrire la densité dans les histogrammes. Il reste aussi que les résultats obtenus ne sont pas quantitatifs et sont juste des appréciations des experts et des comparaisons avec des travaux antérieurs se basant sur des classifications classiques non adaptatives.

### III.5.2 Les travaux de C. VACHIER

Les travaux de C. VACHIER [53] se basent sur une technique de segmentation d'images puissante et souple issu de la morphologie mathématique basée sur le concept de la Ligne de Partage des Eaux avec la contrainte de marquage ce qui permet d'obtenir une segmentation de bonne qualité. Le système proposé permet une détection automatique des opacités du sein qui repose sur les étapes illustrées dans la figure 3.27.

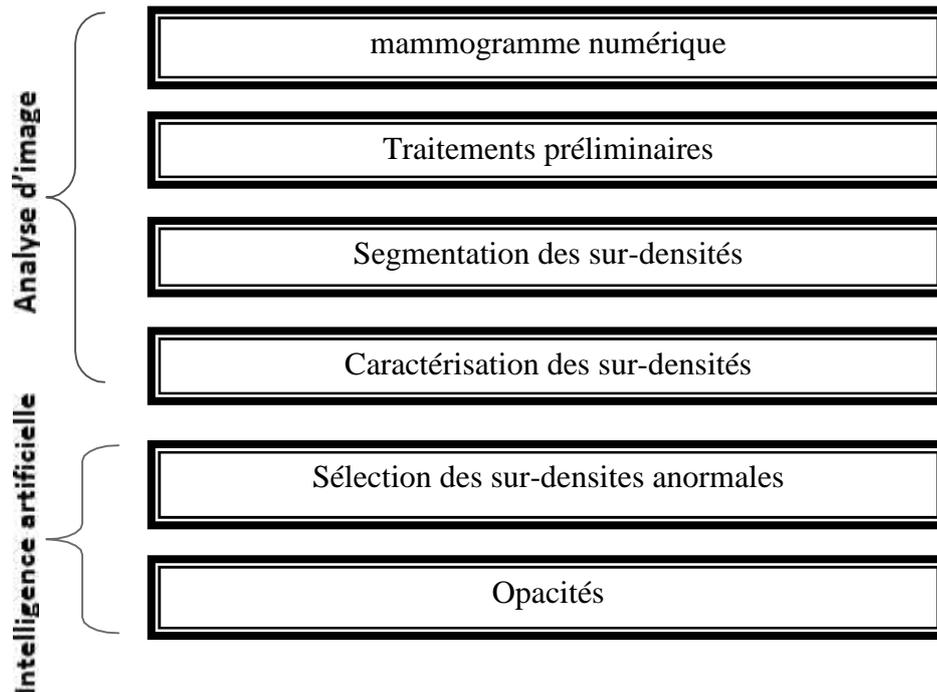


Figure 3.27 - Les principales étapes de l'algorithme de détection des opacités du sein [53].

❖ **Traitement préliminaire** : elle permet de réduire la fenêtre de travail (la glande mammaire peut ne pas occuper qu'une petite partie du cliché radiographique), ce qui permet de réduire les temps des traitements qui suivront l'algorithme utilisé pour la segmentation de la glande mammaire qui est basé sur l'utilisation de la fonction d'extinction surfacique et de la LPE [53].

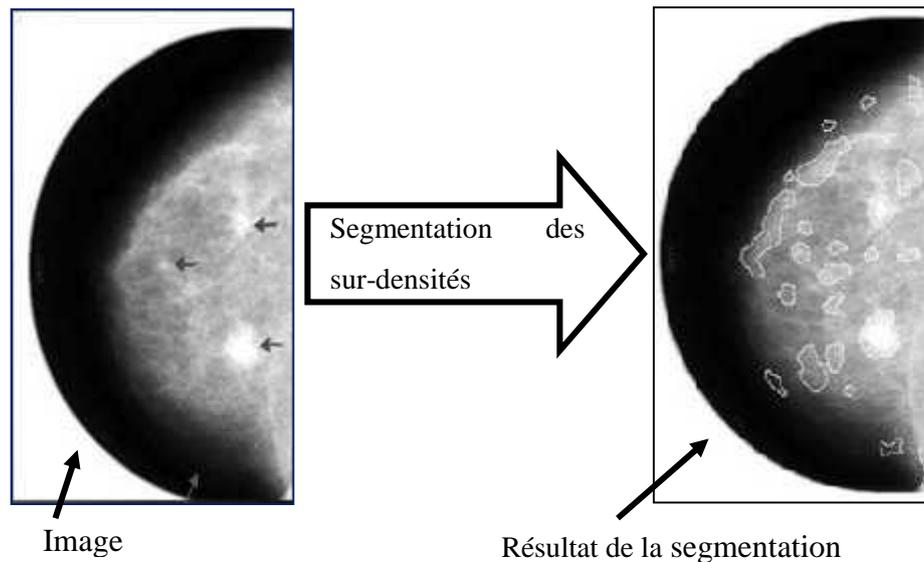
❖ **Segmentation des sur-densités** : elle consiste à extraire les régions d'intérêt dans la glande mammaire (les sur-densités), l'algorithme utilisé ici est toujours basé sur la LPE ; les marqueurs des sur-densités sont obtenus en considérant les maxima de plus forte dynamique [53].

❖ **Caractérisation des sur-densités** : elle permet de créer pour chaque sur-densité (un candidat) une carte d'identification contenant toutes les caractéristiques et indices nécessaires à la détection d'une anomalie (pathologie) (utilisés par les radiologues) [53].

❖ **Sélection des sur-densités anormales** : elle consiste à comparer chaque carte d'identification d'un candidat à la carte d'identité de la pathologie à rechercher et ainsi, la confrontation de ces deux modèles permet de conclure pour chaque candidat, si oui ou non il correspond à une lésion.

Les résultats obtenus par cette méthode sont très prometteurs avec un taux de performance

d'environ 73.5%. Néanmoins, le système a été testé sur un petit échantillon d'images et pour une seule pathologie. De plus, la méthode de sélection des anomalies est très primitive et lente. Les indices de détections tels qu'ils sont pris sont insuffisants car chaque indice dans une carte d'identification peut avoir un poids variant d'une pathologie à une autre, sans parler de la méthode LPE qui est une méthode paramétrable, donc une erreur de paramétrage peut aboutir à des résultats très mauvais [53]. Un exemple des résultats de cette méthode est présenté sur la figure 3.28.



**Figure 3.28 - Segmentation des sur-densités en utilisant la dynamique [53].**

### **III.5.3 Les travaux de Rangayyan**

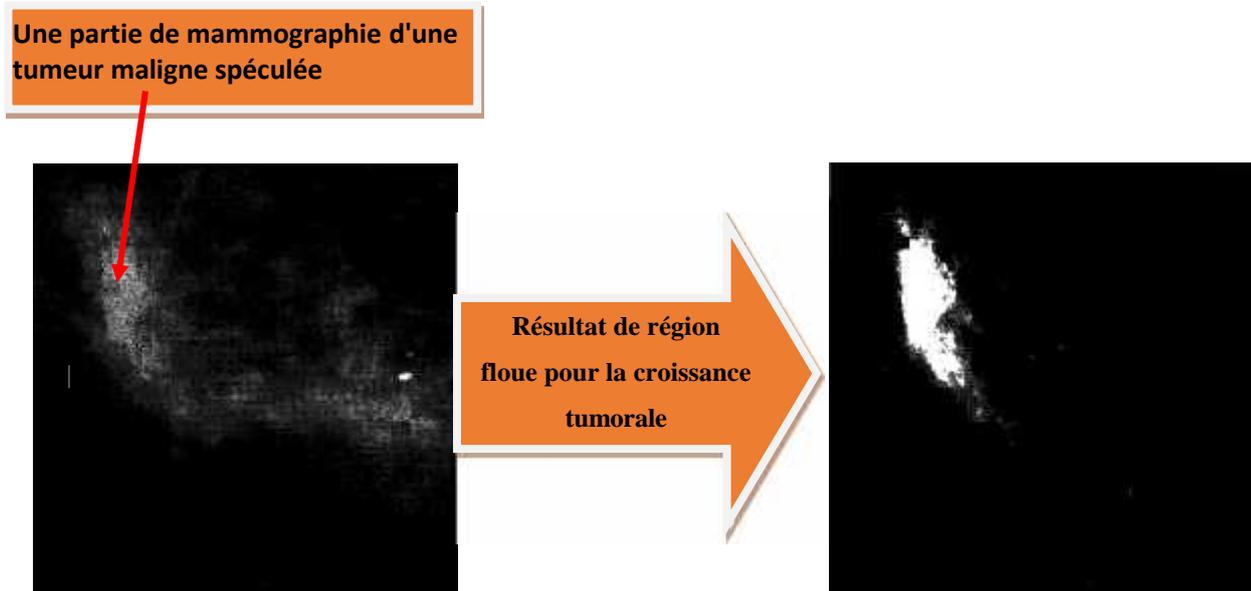
Rangayyan et al. [46] ont proposé une méthode de croissance de régions permettant de décrire des contours flous et ceci afin de réaliser la détection de zones tumorales. La méthode proposée consiste tout d'abord à choisir un point amorce qui constituera le point de départ pour la segmentation d'une telle zone. Une valeur d'appartenance unitaire est attribuée à ce pixel. Ensuite, les pixels voisins du point amorce seront agrégés d'une manière itérative suivant la règle [46] :

*Si* le niveau de gris d'un pixel est proche de la valeur moyenne de la zone et le degré de l'homogénéité de la zone ne change pas beaucoup entre deux itérations successives  
**Alors l'appartenance** du pixel est unitaire  
*Sinon*  
**Si** le niveau de gris est très loin de la moyenne

**Alors** l'appartenance du pixel est nulle

**Sinon** l'appartenance du pixel est entre zéro et l'unité.

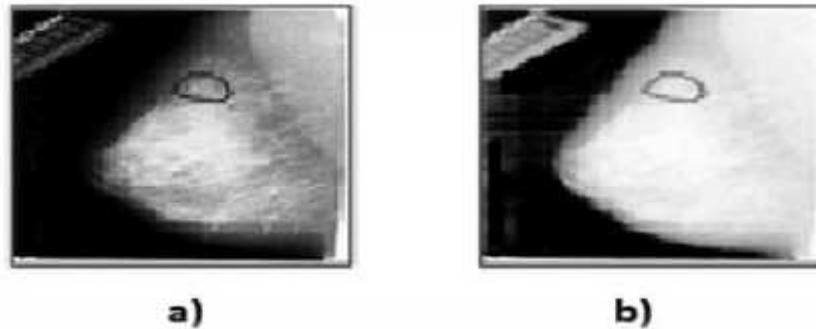
Dans ce dernier cas, le degré d'appartenance est calculé en utilisant le niveau de gris du pixel et la valeur moyenne de la zone. Lorsque cette agrégation itérative est finalisée, l'ensemble des pixels appartenant à la zone tumorale (resp. fond d'image) auront une valeur d'appartenance unitaire (resp. zéro). Les pixels ayant des valeurs d'appartenance partielles (i.e.  $\in ] 0, 1 [$ ) sont considérés comme étant des pixels appartenant aux zones frontières qui séparent la zone tumorale du fond. Cette méthode a été appliquée sur 18 mammographies. [46]. Un exemple des résultats de la segmentation d'une mammographie tumorale est présenté sur la figure 3.29.



**Figure 3.29 - Segmentation par croissance de régions permettant de décrire des contours flous [46].**

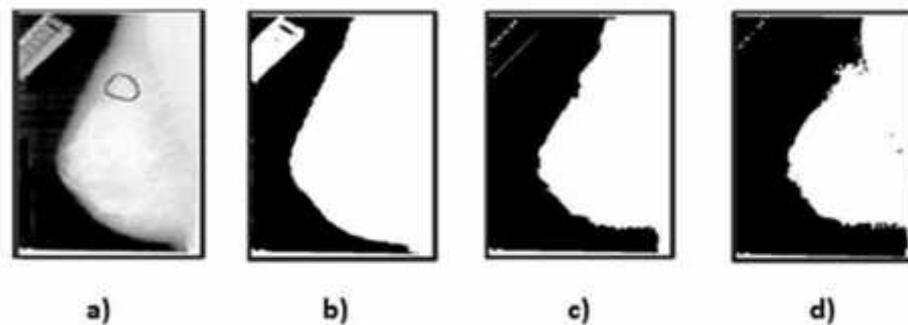
### **III.5.4 Les travaux de Boujelben**

Le travail de recherche appréhendé par Boujelben se rapporte à la détection des calcifications dans une image mammographique suivant les ensembles de niveaux [12]. La première phase consiste à appliquer sur l'image une transformation logarithmique pour améliorer le contraste des régions proches du contour de la tumeur (figure 3.30).



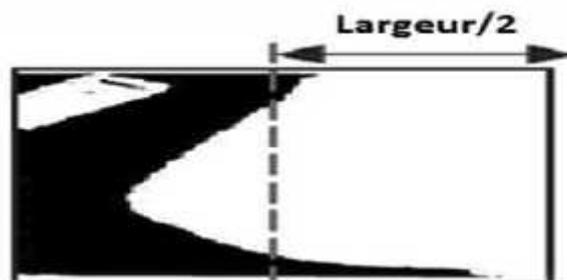
**Figure 3.30 - Résultat de la transformation logarithmique : (a) image originale, b) image améliorée [12].**

La seconde étape consiste en la binarisation de l'image pour recueillir l'enveloppe de la glande mammaire (figure 3.31).



**Figure 3.31 - Résultats de la binarisation : (a) image améliorée, (b) méthode d'Otsu, (c) Méthode Maximum-entropie, (d) Méthode de la « Maximum corrélation » [12].**

La troisième étape consiste à déterminer l'orientation de la tumeur (Figure 3.32).



**Figure 3.32 - Résultats de l'orientation [12].**

## **III.6 Conclusion**

Dans ce chapitre, nous avons proposé une classification des méthodes de segmentation des images mammographiques en explicitant les techniques les plus connues et les plus employées. Cependant chacune d'entre elles a ses qualités et ses défauts, et c'est sur cette base que doit se faire le choix de l'emploi de l'une ou l'autre de ces techniques. ? L'analyse des principes et des performances des différentes méthodes de segmentation permet de déduire les remarques suivantes :

- ❖ Etant donné le faible contraste et la nature fortement texturée des images mammographiques, toutes les méthodes de segmentation proposées dépendent très fortement des valeurs choisies des paramètres (seuils, valeurs moyennes, variances, etc.) et des modèles exploités par ces méthodes (densité de probabilités, fonction d'appartenance ...). Par conséquent, une faible erreur d'estimation de ces paramètres ou de ces modèles peut aboutir à des résultats de segmentation de qualités médiocres en termes de taux d'erreur au niveau pixélique. Cette mauvaise estimation risque également de compromettre la détection des petites régions contenues dans les images mammographiques.
- ❖ Les méthodes de segmentation par croissance de régions se basant sur les mesures liées aux niveaux de gris (y compris des mesures texturales), ou des mesures probabilistes apportent une bonne identification initiale des régions d'intérêt mais souffrent de l'inconvénient majeur d'une localisation imprécise des contours de ces régions.
- ❖ L'approche de segmentation par les contours actifs présente de bons résultats au niveau de la localisation des contours des régions d'intérêt, à condition que l'initialisation de ces contours ne soit pas trop éloignée des contours finaux. Néanmoins, la nature texturée des images mammographiques fait souvent aboutir à de multiples faux contours à l'intérieur des régions détectées.
- ❖ Les méthodes de segmentation basées sur le concept de ligne de partage des eaux est puissante et souple, mais le choix des paramètres reste le maillon faible de cette méthode.

# **CHAPITRE IV :**

*Approche proposée*

## **IV.1 Introduction**

Après une étude de l'état de l'art sur les systèmes assistés par ordinateur en imagerie mammaire, et exploration des différentes approches de détection et de segmentation, dans ce chapitre nous exposons l'architecture de notre système d'aide à la diagnostic automatique du cancer de sein à partir d'images mammographiques, nous présentons notre approche qui se base sur 4 étapes principales: le prétraitement, la segmentation, l'extraction et la sélection des caractéristiques, et la classification.

L'extraction des descripteurs est une étape fondamentale dans tout système de traitement d'images mammographiques, car ce sont eux qui vont traduire l'aspect pathologique et la spécificité des lésions par des quantités facilement utilisables par les algorithmes de classification. Toutefois, les systèmes conçus sur de telles approches souffrent de plusieurs limitations qui justifient leur absence en milieu médical. Parmi ces limitations, nous pouvons évoquer le fait que les performances obtenues varient en fonction des descripteurs utilisés et du jeu de données ayant servi pour la validation, de plus, la combinaison de certains descripteurs peut dégrader les performances de ces systèmes, ce qui oblige à sélectionner de façon minutieuse les descripteurs les plus pertinents. En termes de santé publique, l'enjeu primordial est de proposer, au médecin, un outil performant, intelligent et évolutif permettant d'augmenter sa sensibilité aux lectures des images tout en attirant son attention sur des lésions potentiellement significatives.

## **IV.2 Objectifs de notre étude**

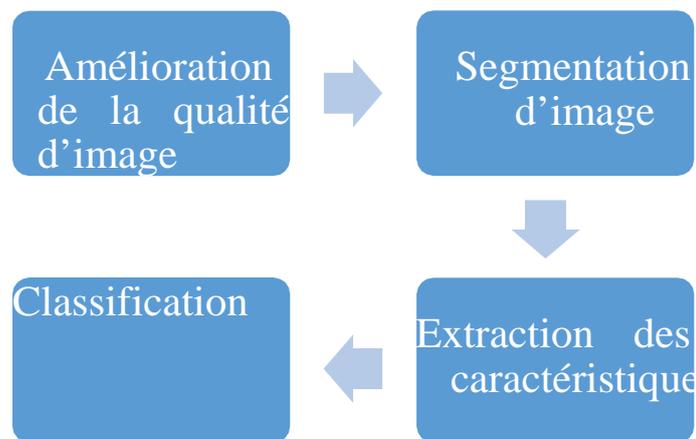
Notre objectif principal est d'élaborer un système d'aide au diagnostic en mammographie, quasi automatique pouvant être associé à un appareil de mammographie et destiné aux radiologues comme deuxième lecture pour une détection précoce du cancer du sein par l'un des premiers signes pathologiques : les masses mammaires. Donc, ce système ne doit pas être vu comme une tentative de remplacer le médecin mais de lui proposer des outils performants qui l'aident dans la lourde tâche d'analyse des clichés de mammographie lors des campagnes de dépistage.

La détection des opacités du sein sur les mammographies n'est pas une opération mathématique qui permet d'aboutir à un résultat absolument certain. La preuve en est que, dans les cas litigieux, le radiologue a généralement recours à d'autres techniques plus précises (l'histologie notamment) pour valider son diagnostic. Les mammographies sont utilisées pour déceler d'éventuels signes pathologiques et décider de la nécessité d'exams

complémentaires plus approfondis. Le but de notre étude se définit, dans ce cadre, comme une aide au diagnostic médical ; c'est-à-dire que l'on cherche à effectuer automatiquement une lecture de la mammographie, qui sera ensuite comparée à celle du médecin.

### IV.3 Présentation du système général

Le système général de notre approche est constitué de quatre étapes comme montré dans la figure 4.1.



**Figure 4.1 - Schéma général de l'approche proposée**

- ❖ Amélioration de la qualité de l'image : cette première étape est réalisée pour améliorer l'image de mammographie afin de faciliter son traitement ultérieur.
- ❖ Segmentation de l'image : l'étape de segmentation vise à diviser l'image en segments afin d'obtenir des formes appropriées pour le processus de traitement.
- ❖ Extraction de fonctionnalités (caractéristiques) : cela est fait pour obtenir les caractéristiques des formes contenues dans l'image.
- ❖ Classification : après avoir obtenu des informations de l'image, l'image mammographique est classée en deux classes : cancéreuse ou normale.

### IV.3.1 Pré-traitement

#### IV.3.1.1 Algorithme de distribution des possibilités

La méthode de distribution des possibilités [1] améliore l'image en utilisant le principe de l'approche de la logique floue. Cette méthode utilise 5 paramètres :  $\alpha$ ,  $\beta_1$ ,  $\gamma$ ,  $\beta_2$  et  $\max$ , comme le montre la figure 4.2.

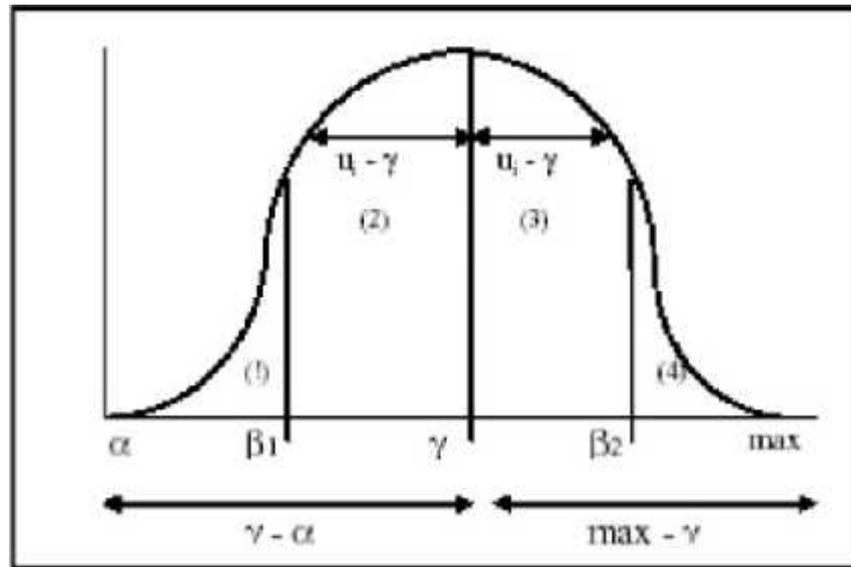


Figure 4.2 - Fonction de distribution de possibilité.

Parmi les paramètres requis,  $\alpha$  représente la valeur de distribution minimale,  $\gamma$  représente la valeur de distribution moyenne et le  $\max$  représente la valeur de distribution maximale. La fonction de transformation floue pour obtenir la valeur entière est définie comme suit :

- ❖  $\alpha$  = min de la distribution
- ❖  $\beta_1 = (\alpha + \gamma) / 2$
- ❖  $\gamma$  = moyenne
- ❖  $\beta_2 = (\max + \gamma) / 2$
- ❖  $\max$  = max de la distribution

Le principe de la méthode est de fournir une nouvelle valeur d'intensité en pixels entre  $\beta_1$  et  $\beta_2$ , et entre  $\alpha$  et  $\max$  avec la valeur de direction opposée à la valeur moyenne  $\gamma$ . Les règles floues ci-dessous sont utilisées pour améliorer le contraste de l'image.

- ❖ Règle -1: Si  $u_i < \beta_1$  Alors  $P = 2((u_i - \alpha) / (\beta_1 - \alpha))^2$
- ❖ Règle -2: Si  $\beta_1 < u_i < \beta_2$  Alors  $P = 1 - 2((u_i - \gamma) / (\beta_2 - \gamma))^2$
- ❖ Règle -3: Si  $u_i > \beta_2$  Alors  $P = 1 - 2((u_i - \max) / (\max - \beta_2))^2$

❖ Règle -4: Si  $2 u_i \max$  Alors  $P=2((u_i-)/(max-))^2$

Où  $u_i = f(x,y)$  est l'intensité du pixel.

Les règles qui améliorent le niveau de gris des pixels dont la valeur est entre 1 et 2 sont représentées par les règles 2 et 3. Les étapes de l'algorithme de distribution des possibilités peuvent être expliquées comme suit :

#### Étape -1 : initialisation des paramètres

Déterminer le 1 obtenu à partir du calcul  $(\min+\text{moyen})/2$  et 2 obtenus à partir du calcul  $(\max+\text{moyen})/2$  où :

- ✓ Min : valeur minimale d'intensité de pixel dans l'image.
- ✓ Max : valeur maximale d'intensité de pixel dans l'image.
- ✓ Moyen : valeur moyenne d'intensité de pixel sur l'image

**Étape -2: fuzziication :** Pour tous les pixels de la matrice de données, appliquer des règles floues (Règle-1, Règle-2, Règle-3 et Règle-4) pour obtenir de nouvelles valeurs de niveau de gris.

**Étape -3 : Modification :** Pour tous les pixels, on calcule la valeur floue des données obtenues à partir de la valeur de niveau de gris nouvellement classée en deux.

**Étape -4 : Défuzziication :** Pour tous les pixels de l'image, le calcul des données améliorées se fait en multipliant les données floues par les données d'intensité de chaque pixel.

## IV.3.2 Segmentation

### IV.3.2.1 La méthode de OTSU [43]

La méthode Otsu est une méthode de segmentation itérative qui vise à minimiser la variance inter classe de l'histogramme d'image jusqu'à convergence et sépare parallèlement l'histogramme en deux classes.

Le but de l'algorithme de seuillage d'Otsu est de segmenter les images d'une manière différenciée en 2 classes, à savoir l'arrière-plan (valeur définie sur 0) et l'objet (valeur définie sur 1) en utilisant un certain niveau comme délimiteur. Les différentes étapes pour trouver le seuil d'Otsu sont les suivantes :

❖ Etape 1 : A partir de l'image améliorée on calcule :

- Le nombre de pixels avec le niveau de gris  $i$ , représenté par  $n_i$ ,
- Le nombre de pixels de l'image toute entière  $N$  définis par  $N=n_1+n_2+\dots+n_n$ .
- La valeur de probabilité  $p$  du nombre de pixels divisé par le total du nombre de pixels.

$$\text{Caractéristique de forme } p_i = n_i/N, p_i \geq 0, \sum_{i=1}^L p_i = 1,$$

❖ Etape 2 : On divise les pixels en 2 classes,  $C_0$  comme arrière-plan, et  $C_1$  comme objets en utilisant les formules suivantes :

$$w_0 = \Pr(C_0) = \sum_{i=1}^k p_i = w(k) \quad (4.1)$$

$$w_1 = \Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - w(k)$$

$$u(k) = \sum_{i=1}^k i * p_i \quad (4.2)$$

❖ Etape 3 : On définit La classe de variance avec la formule suivante :

$$\begin{aligned} \dagger_w^2 &= w_0 \dagger_0^2 + w_1 \dagger_1^2 \\ \dagger_B^2 &= w_0 (\sim_0 - \sim_r)^2 + w_1 (\sim_1 - \sim_r)^2 \\ &= w_0 w_1 (\sim_1 - \sim_0)^2 \end{aligned} \quad (4.3)$$

$$\dagger_r^2 = \sum_{i=1}^L w_0 (i - \sim_r)^2 p_i \quad (4.4)$$

❖ Etape 4 : une recherche de valeur maximale de toutes les valeurs de variance est effectuée et une valeur de seuil est déterminée à l'aide de la moyenne de la valeur de variance trouvée. S'il n'est pas trouvé, la valeur de seuil prend la valeur 0.

$$\dagger_B^2(k^*) = \max_{1 \leq k \leq L} \dagger_B^2(k). \quad (4.5)$$

$$S^* = \{k; w_0 w_1 = w(k)[1 - w(k)] > 0, \text{ ou } 0 < w(k) < 1\}; \quad (4.6)$$

### IV.3.2.2 Détermination des régions d'intérêt

L'image segmentée est rognée de sorte que l'objet traité devient plus concentré. Pour rogner une image, il est nécessaire de spécifier la bordure supérieure, la bordure inférieure, la bordure droite et la bordure gauche des contours des objets. Pour obtenir ces limites on parcourt tous les pixels de l'image horizontalement et verticalement (à partir du haut, du bas, de la droite ou de la gauche), si dans ce parcours on trouve un pixel avec l'intensité 1, alors ce pixel devient une limite de l'image. Une fois cette limite est trouvée, la prochaine étape consiste à rogner l'image à partir de la limite précédemment définie.

### IV.3.3 Extraction des caractéristiques

#### IV.3.3.1 Algorithme MCNG (Matrice de co-occurrence de niveau gris)

La matrice de co-occurrence des niveaux de gris (MCNG) [42] est une matrice bidimensionnelle qui représente les variations des niveaux de gris dans l'image. Les mesures telles que la corrélation, l'énergie, le contraste, l'entropie, l'homogénéité sont des calculs de MCNG. L'ensemble des caractéristiques est calculé de la manière suivante :

##### La moyenne :

$$\sim = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N p(i, j) \quad (4.7)$$

M et N sont respectivement la largeur et la hauteur de l'image.

P(i,j) est le niveau de gris du pixel avec les coordonnées (i,j).

##### La variance :

$$\text{var} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (p(i, j) - \sim)^2 \quad (4.8)$$

##### L'écart type :

$$\dagger = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (p(i, j) - \sim)^2} \quad (4.9)$$

##### Le constraste :

$$\text{contraste} = \sum_{i=1}^M \sum_{j=1}^N (i - j)^2 P(i, j) \quad (4.10)$$

Pour calculer le contraste on construit une matrice de contraste dont la taille (ligne)  $i$  x (colonne)  $j$ . Lorsque la ligne ( $i$ ) et la colonne ( $j$ ) sont les mêmes, la valeur du pixel est en diagonale ( $ij$ ) = 0. Ces valeurs représentent des pixels complètement similaires au voisin de ces pixels, de sorte qu'ils sont pondérés 0.

Si  $i$  et  $j$  diffèrent de 1 (léger changement de contraste), il y a une petite différence, et le poids est de  $1^2 = 1$ .

Si  $i$  et  $j$  diffèrent de 2 (changement de contraste considérable), contraste augmente et le poids est de  $2^2 = 4$ . Le poids continue d'augmenter de façon exponentielle lorsque ( $ij$ ) augmente.

**Le Kurtosis :**

$$K = \left\{ \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \left[ \frac{p(i, j) - \sim}{\dagger} \right]^4 \right\} - 3 \quad (4.11)$$

**Equation de lissage :**

$$R = 1 - \frac{1}{1 + \dagger^2} \quad (4.12)$$

### **IV.3.4 Classification avec les réseaux de neurones**

Daté de millions d'années d'existence, le cerveau humain est la machine de traitement de l'information la plus performante. En effet, la lecture d'un texte, la reconnaissance d'un visage ou la résolution d'un problème (prise de décision) sont des opérations propres au cerveau.

L'élément de base constituant le cerveau est le neurone (nom attribué par les biologistes à cet élément) figure 4.3

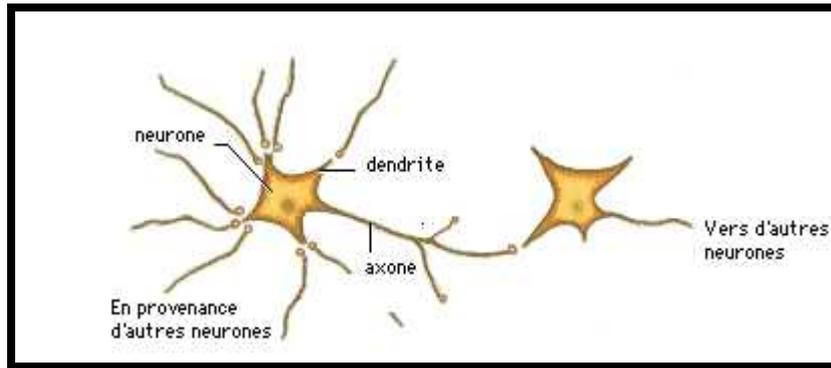


Figure 4.3 - Connexions inter neuronales chez l'humain.

L'activité collective et simultanée des milliards de ces neurones ou cellules nerveuses, confère au cerveau des capacités inaccessibles aux ordinateurs (même les plus puissants). Et pourtant chaque neurone effectue une opération simple sur l'ensemble des signaux qu'il reçoit des autres neurones par le biais des dendrites. En effet, suivant l'intensité de ses signaux le neurone va décider d'envoyer ou non un signal à ses congénères le long de son axone. Cependant, la communication entre axones et dendrites s'effectue en des points appelés synapses.

Les réseaux de neurones artificiels est une architecture artificielle inspirée à partir d'un tel fonctionnement (du cerveau). La figure 4.4 présente les deux réseaux neuronaux biologiques et artificiels.

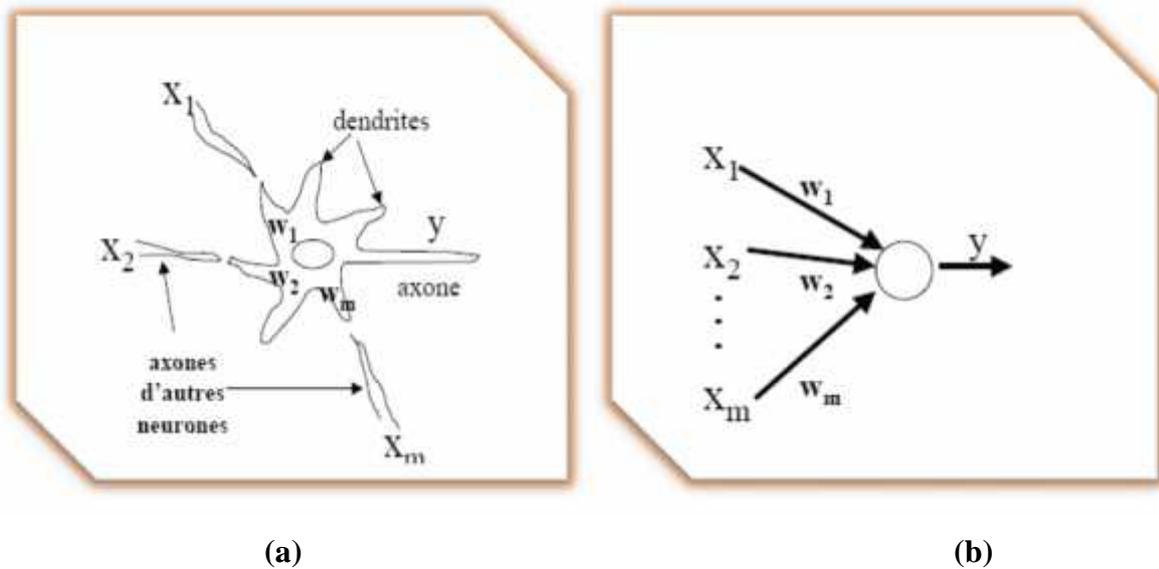


Figure 4.4 - (a) Neurone biologique et (b) Neurone artificiel.

Dans cette section nous dresserons un aperçu général sur ce type d'architecture. Nous présenterons en particulier la modélisation d'un neurone biologique puis nous donnerons différentes architectures basées sur ce modèle et le modèle multicouche utilisée dans notre approche.

#### **IV.3.4 .1 Définition d'un réseau neuronal**

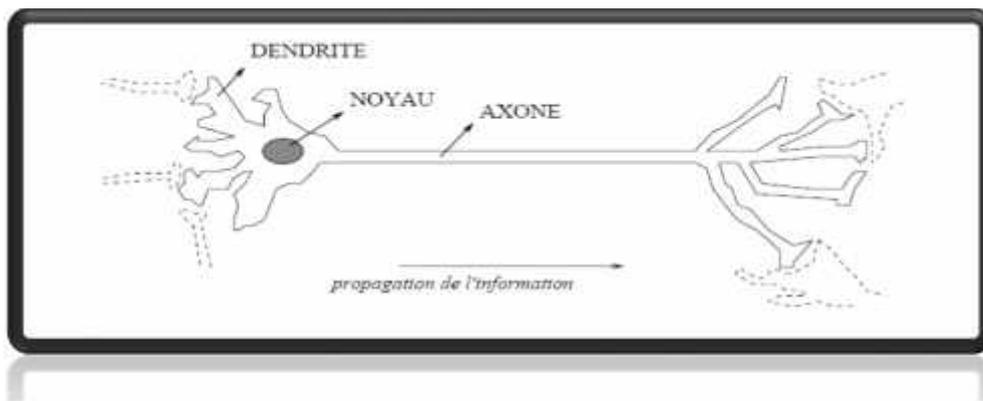
Un réseau neuronal peut être vu comme un modèle du fonctionnement d'une partie du cerveau, lequel se compose d'environ  $10^{11}$  cellules appelées neurones. Ces neurones sont connectés entre eux par des liens entre l'axone d'un neurone et la dendrite d'un autre neurone. On estime qu'un cerveau humain comporte environ  $10^{15}$  liens inter-neuronaux. L'axone est un prolongement du neurone émetteur qui transmet un signal à la dendrite du neurone récepteur. Il y a des liens entre les neurones qui sont forts et des liens plus faibles. Mais lorsque la quantité combinée de matériel chimique provenant des différents axones atteint un certain seuil, le neurone récepteur émettra à son tour un signal de sortie via son axone, à l'intention des autres neurones auxquels il est connecté.

#### **IV.3.4.2 Le neurone biologique**

Les humains peuvent traiter beaucoup de types d'information plus rapidement que les machines. Par exemple, comprendre la langue parlée ou reconnaître des objets dans une scène visuelle, ces opérations prennent quelques cent millisecondes pour les cerveaux mais beaucoup plus pour un ordinateur si même l'opération est possible!

Les unités de l'informatique de base du cerveau humain (neurones) ayez une vitesse de l'informatique de base de quelques millisecondes, beaucoup plus lente que les transistors dans un ordinateur moderne.

**a) La cellule nerveuse**



**Figure 4.5 - Une cellule nerveuse.**

Le type le plus commun de neurone biologique se présente comme une cellule dont le corps est doté de deux types de prolongements (voir figure 4.5) :

- Des ramifications courtes et buissonnantes : les dendrites.
- Une longue fibre qui se termine par une arborisation : l'axone.

Ces prolongements permettent au neurone d'établir des connexions avec d'autres cellules nerveuses (ou avec des cellules musculaires dans le cas des neurones moteurs).

**b) Le corps cellulaire**

Il contient le noyau du neurone ainsi que la machine biochimique nécessaire à la synthèse des enzymes. Ce corps cellulaire de forme sphérique ou pyramidale contient aussi les autres molécules essentielles à la vie de la cellule. Sa taille est de quelques microns de diamètre.

**c) Les dendrites**

Ce sont de fines extensions tubulaires qui se ramifient autour du neurone et forment une sorte de vaste arborescence. Les signaux envoyés au neurone sont captés par les dendrites. Leur taille est de quelques dizaines de microns de longueur.

**d) L'axone**

C'est le long de l'axone que les signaux partent du neurone. Contrairement aux dendrites qui se ramifient autour du neurone, l'axone est plus long et se ramifie à son extrémité ou il se

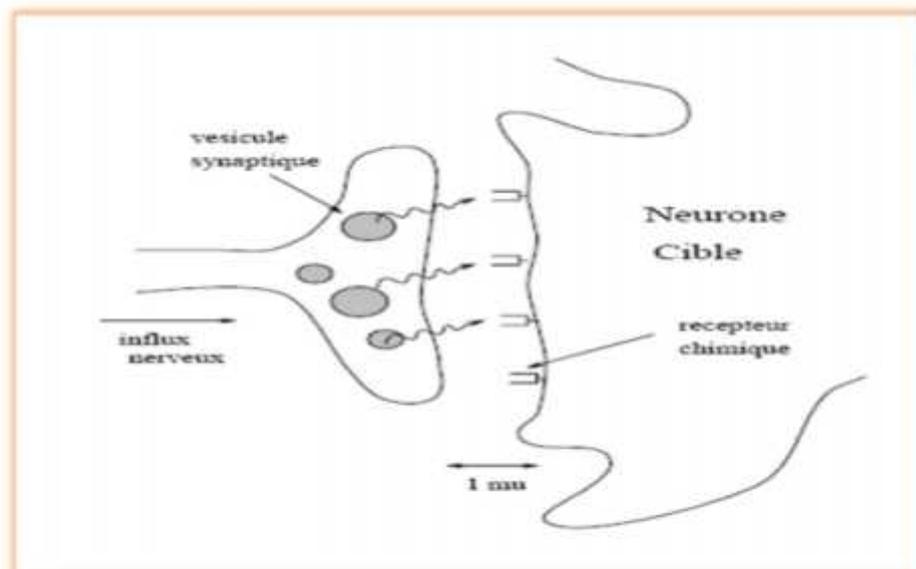
connecte aux dendrites des autres neurones. Sa taille peut varier entre quelques millimètres à plusieurs mètres.

*e) La transmission synaptique*

Comment se fait la transmission de l'influx nerveux lorsqu'il arrive à une extrémité de l'axone (synapse) ? Les expériences montrent que les cas de transmission électrique sont rares. Généralement, la transmission met en œuvre des processus chimiques par l'intermédiaire de molécules organiques : les neuromédiateurs.

La terminaison synaptique contient des vésicules de neuromédiateurs (voir figure 4.6). L'arrivée de l'onde de dépolarisation provoque une augmentation de la perméabilité de la membrane aux neuromédiateurs. Ceux-ci se trouvent donc libérés dans l'espace inter synaptique, et viennent se fixer sur des récepteurs chimiques situés sur le neurone cible. Lorsque ces récepteurs captent des neuromédiateurs, ils provoquent une augmentation de la perméabilité de la membrane du neurone cible aux ions sodium.

L'entrée d'ions sodium dans la cellule peut ainsi déclencher un nouveau potentiel d'action suivant les principes précédemment énoncés (en pratique il faut souvent plusieurs dépolarisations successives pour libérer une quantité suffisante de neuromédiateur).



**Figure 4.6 - Une transmission synaptique.**

### IV.3.4.3 Neurone artificiel

Un Réseau de Neurones Artificiels (RNA) est un ensemble de neurones formels (d'unités de calcul simples, de nœuds processeurs) associés en couches (ou sous-groupes) et fonctionnant en parallèle (figure 4.7).

Dans un réseau, chaque sous-groupe fait un traitement indépendant des autres et transmet le résultat de son analyse au sous-groupe suivant. L'information donnée au réseau va donc se propager couche par couche, de la couche d'entrée à la couche de sortie, en passant soit par une, une ou plusieurs couches intermédiaires (dites couches cachées). Il est à noter qu'en fonction de l'algorithme d'apprentissage, il est aussi possible d'avoir une propagation de l'information à reculons ("*back propagation*"). Habituellement (excepté pour les couches d'entrée et de sortie), chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante.

Les RNA ont la capacité de stocker de la connaissance empirique et de la rendre disponible à l'usage. Les habiletés de traitement (et donc la connaissance) du réseau vont être stockées dans les poids synaptiques, obtenus par des processus d'adaptation ou d'apprentissage.

En ce sens, les RNA ressemblent donc au cerveau car non seulement, la connaissance est acquise au travers d'un apprentissage mais de plus, cette connaissance est stockée dans les connexions entre les entités, soit dans les poids synaptiques.

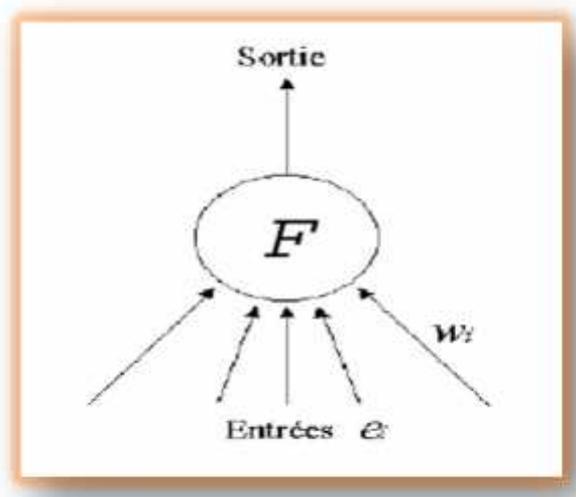


Figure 4.7 - Neurone artificiel.

On peut classer les RNA en deux grandes catégories : Les réseaux "FEED-FORWARD" et Les réseaux "FEED-BACK".

*a) Les réseaux "FEED-FORWARD"*

Appelés aussi "réseaux de type Perceptron", ce sont des réseaux dans lesquels l'information se propage de couche en couche sans retour en arrière possible. Ils sont divisés en trois classes :

*Le perceptron monocouche :* C'est historiquement le premier RNA, c'est le Perceptron de *Rosenblatt*. C'est un réseau simple, puisqu'il ne se compose que d'une couche d'entrée et d'une couche de sortie. Il est calqué, à la base, sur le système visuel et de ce fait a été conçu dans un but premier de reconnaissance des formes. Cependant, il peut aussi être utilisé pour faire de la classification et pour résoudre des opérations logiques simples (telle "ET" ou "OU"). Sa principale limite est qu'il ne peut résoudre que des problèmes linéairement séparables. Il suit généralement un apprentissage supervisé selon la règle de correction de l'erreur.

*Le perceptron multicouche (PMC) :* C'est une extension du précédent, avec une ou plusieurs couches cachées entre l'entrée et la sortie. Chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante (excepté pour les couches d'entrée et de sortie) et il n'y a pas de connexions entre les cellules d'une même couche. Les fonctions d'activation utilisées dans ce type de réseaux sont principalement les fonctions à seuil ou sigmoïdes. Il peut résoudre des problèmes non-linéairement séparables et des problèmes logiques plus compliqués, et notamment le fameux problème du XOR. Il suit aussi un apprentissage supervisé selon la règle de correction de l'erreur.

*Les réseaux à fonction radiale :* Ce sont les réseaux que l'on nomme aussi RBF. L'architecture est la même que pour les PMC cependant, les fonctions de base utilisées ici sont des fonctions Gaussiennes. Les RBF seront donc employés dans les mêmes types de problèmes que les PMC à savoir, en classification et en approximation de fonctions, particulièrement. L'apprentissage le plus utilisé pour les RBF est le mode hybride et les règles sont soit, la règle de correction de l'erreur soit, la règle d'apprentissage par compétition.

***b) Les réseaux "FEED-BACK"***

Appelés aussi "réseaux récurrents", ce sont des réseaux dans lesquels il y a retour en arrière de l'information.

*Les cartes auto-organisatrices de Kohonen :* Ce sont des réseaux à apprentissage non-supervisé qui établissent une carte discrète, ordonnée topologiquement, en fonction de patterns d'entrée. Le réseau forme ainsi une sorte de treillis dont chaque nœud est un neurone associé à un vecteur de poids. La correspondance entre chaque vecteur de poids est calculée pour chaque entrée. Par la suite, le vecteur de poids ayant la meilleure corrélation, ainsi que certains de ses voisins, vont être modifiés afin d'augmenter encore cette corrélation.

*Les réseaux de Hopfield :* Les réseaux de Hopfield sont des réseaux récurrents et entièrement connectés. Dans ce type de réseau, chaque neurone est connecté à chaque autre neurone et il n'y a aucune différenciation entre les neurones d'entrée et de sortie. Ils fonctionnent comme une mémoire associative non-linéaire et sont capables de trouver un objet stocké en fonction de représentations partielles ou bruitées. L'application principale des réseaux de Hopfield est l'entrepôt de connaissances mais aussi la résolution de problèmes d'optimisation. Le mode d'apprentissage utilisé ici est le mode non-supervisé.

**IV.3.4.4 Apprentissage**

Pour un RNA, l'apprentissage peut être considéré comme le problème de la mise à jour des poids des connexions au sein du réseau, afin de réussir la tâche qui lui est demandée. L'apprentissage est la caractéristique principale des RNA et il peut se faire de différentes manières et selon différentes règles. Il existe quatre types d'apprentissages :

***a) Le mode supervisé***

Dans ce type d'apprentissage, le réseau s'adapte par comparaison entre le résultat qu'il a calculé, en fonction des entrées fournies, et la réponse attendue en sortie. Ainsi, le réseau va se modifier jusqu'à ce qu'il trouve la bonne sortie, c'est-à-dire celle attendue, correspondant à une entrée donnée.

***b) Le renforcement***

Le renforcement est en fait une sorte d'apprentissage supervisé et certains auteurs le classe d'ailleurs, dans la catégorie des modes supervisés. Dans cette approche le réseau doit

apprendre la corrélation entrée/sortie via une estimation de son erreur, c'est-à-dire du rapport échec/succès. Le réseau va donc tendre à maximiser un index de performance qui lui est fourni, appelé signal de renforcement. Le système étant capable ici, de savoir si la réponse qu'il fournit est correcte ou non, mais il ne connaît pas la bonne réponse.

**c) Le mode non-supervisé (ou auto-organisationnel)**

Dans ce cas, l'apprentissage est basé sur des probabilités. Le réseau va se modifier en fonction des régularités statistiques de l'entrée et établir des catégories, en attribuant et en optimisant une valeur de qualité, aux catégories reconnues.

**d) Le mode hybride**

Le mode hybride reprend en fait les deux autres approches, puisqu'une partie des poids va être déterminée par apprentissage supervisé et l'autre partie par apprentissage non-supervisé.

### IV.3.4.5 Neurone formel

Le premier neurone formel est apparu en 1943. On le doit à *Mac Culloch* et *Pitts*. Voici un schéma de leur modèle de neurone formel :

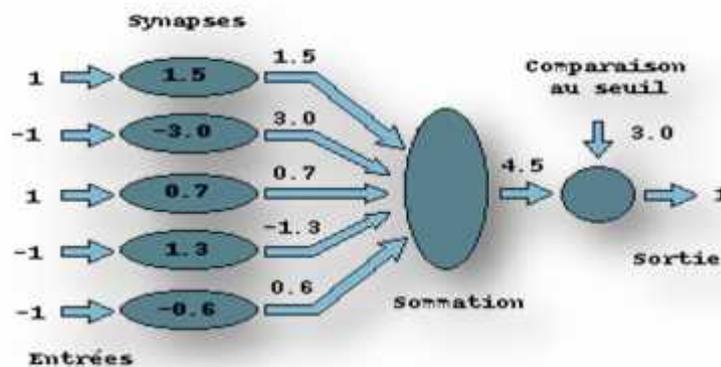


Figure 4.8 - Le neurone formel.

Le neurone formel est donc une modélisation mathématique qui reprend les principes du fonctionnement du neurone biologique, en particulier la sommation des entrées. Sachant qu'au niveau biologique, les synapses n'ont pas toutes la même « valeur » (les connexions entre les neurones étant plus ou moins fortes), les auteurs ont donc créé un algorithme qui pondère la

somme de ses entrées par des poids synaptiques (coefficients de pondération). De plus, les 1 et les -1 en entrée sont là pour figurer une synapse excitatrice ou inhibitrice.

a) *Interprétation mathématique*

D'un point de vue mathématique, le neurone formel peut être représenté de la manière suivante :

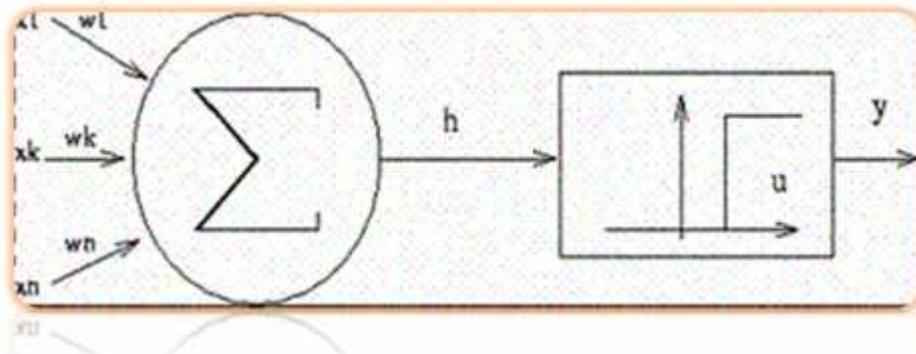


Figure 4.9 - Représentation mathématique du neurone formel.

Pour un nombre compris entre  $j$  ( $=1$ ) et un nombre quelconque  $n$ , le neurone formel va calculer la somme de ses entrées ( $x_1, \dots, x_n$ ), pondérées par les poids synaptiques ( $w_1, \dots, w_n$ ), et la comparer à son seuil. Si le résultat est supérieur au seuil, alors la valeur renvoyée est 1, sinon la valeur renvoyée est 0. D'où la formule (avec  $f$  = fonction seuil) :

$$y = f \sum_{j=1}^n w_j x_j - u \tag{4.13}$$

b) *Fonctions d'activation*

Dans sa première version, le neurone formel était donc implémenté avec une fonction à seuil (figure 4.10 a), mais de nombreuses versions existent. Ainsi le neurone de *McCulloch* et *Pitts* a été généralisé de différentes manières, en choisissant d'autres fonctions d'activations, comme les fonctions linéaires par morceaux (figure 4.10 b), des sigmoïdes (figure 4.10 c) ou des gaussiennes (figure 4.10 d) par exemples.

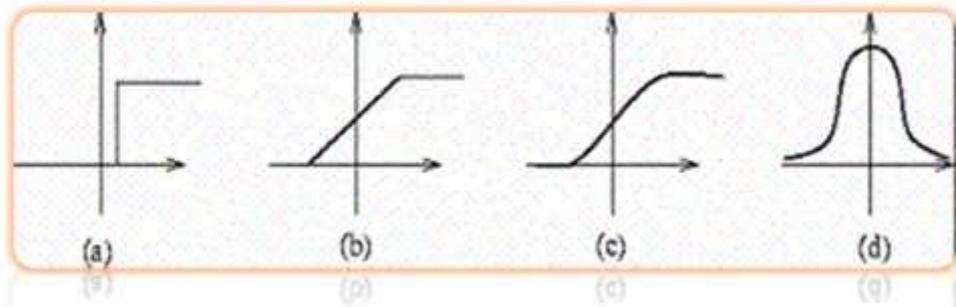


Figure 4.10 - Les fonctions d'activation : à seuil(a) par morceaux (b), des sigmoïdes (c) et des gaussiennes (d).

#### IV.3.4.6 Le réseau de neurone multi couche (MLP)

Le réseau de neurones multi couche est un modèle créé pour imiter les fonctions d'apprentissage du cerveau humain. Dans les réseaux neuronaux, les neurones sont regroupés en couches, appelées couches neuronales. Habituellement, chaque neurone d'une couche est connecté à tous les neurones de ses couches arrière et avant (à l'exception de l'entrée et de la sortie). L'information envoyée dans un réseau neuronal, la propagande de couches – par – des couches allant de l'entrée à la sortie sans ou à travers une ou plusieurs couches cachées.

Selon l'algorithme utilisé, l'information peut également être propagée vers l'arrière (rétropropagation). La figure 4.11 montre le réseau neuronal avec trois couches de neurones.

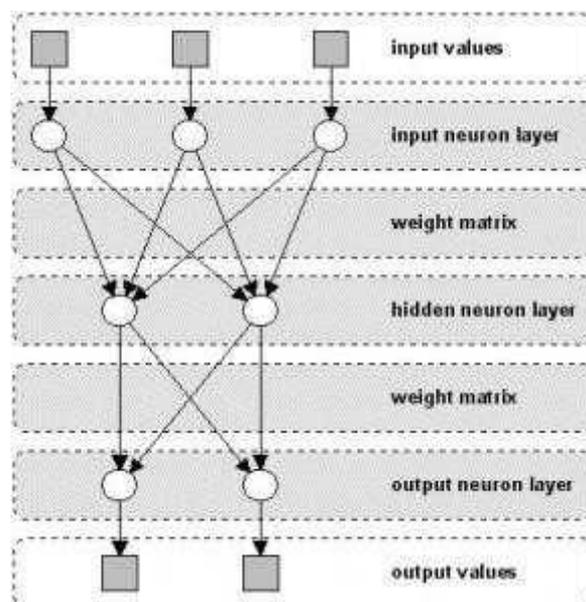


Figure 4.11 – Architectures des réseaux de neurones multi couches.

Le Multi Layer Perceptron (MLP) est un type de réseau de neurones très populaire et qui donne généralement de bons résultats de classification.

Le MLP dispose d'une couche d'entrée, de plusieurs couches cachées et d'une couche de sortie. La figure 4.12 donne un aperçu de l'architecture d'un seul nœud du réseau MLP, Le point  $i$  signifie le numéro de neurone dans le réseau et  $g$  est sa fonction d'activation.

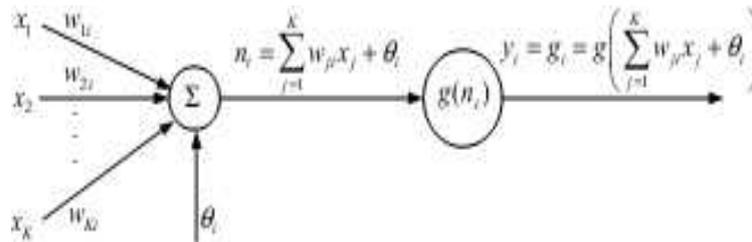


Figure 4.12 - Nœuds uniques sur le réseau MLP.

Avec  $x_k$ ,  $k=1, \dots, K$  est l'entrée du modèle,  $W$  la matrice des poids et  $\theta$  le biais. Le résultat  $n_i$  est une entrée pour activer la fonction d'activation  $g$ . il existe plusieurs fonctions d'activation mais les deux couramment utilisées en MLP sont : la fonction *tangente* hyperbolique (*tanh*) et la fonction sigmoïde. Les tangentes hyperboliques sont définies comme :

$$\tanh(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (4.14)$$

Ainsi, le point de sortie devient comme ceci :

$$y_i = g_i = g \left( \sum_{j=1}^k w_{ji} x_j + n_i \right) \quad (4.15)$$

En connectant plusieurs points en série et en parallèle, le réseau MLP sera formé comme à la figure 4.13.

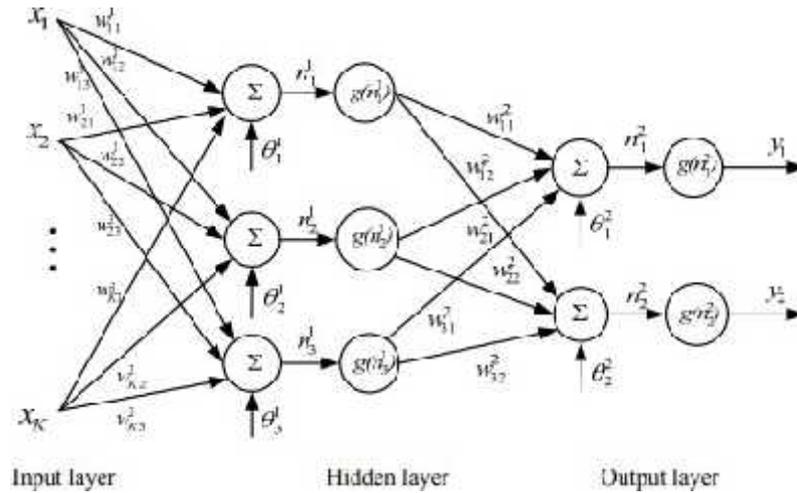


Figure 4.13 - MLP avec une couche cachée.

La sortie  $y_i, i=1,2$ , du MLP est donnée comme suit :

$$y_i = g \left( \sum_{j=1}^3 w_{ji}^2 g(n_j^1) + n_j^2 \right) = g \left( \sum_{j=1}^3 w_{ji}^2 g \left( \sum_{k=1}^K w_{kj}^1 x_k + n_j^1 \right) + n_j^2 \right) \quad (4.16)$$

À partir de l'équation ci-dessus, on peut conclure que le réseau MLP est un réseau avec des paramètres non linéaires de l'entrée  $x - R_k$  à la sortie  $y - R_m$  (ici  $m = 3$ ). Les paramètres sont les poids  $w_k$  et  $\theta_k$  les biais. La fonction d'activation  $g$  est généralement supposée être la même sur chaque couche.

Ce type de réseau de neurone est utilisé dans notre approche pour classifier les images mammographiques en image normales et cancéreuses.

Ce processus de classification est effectué en deux étapes :

Dans la première étape on fait apprendre le réseau MLP les caractéristiques des images cancéreuses. Et dans l'étape de test on présente au réseau des images quelconques avec leurs caractéristiques et le réseau peut automatiquement les classifier soit en images normales ou cancéreuses.

## **IV.4 Conclusion**

La recherche en imagerie médicale est une des disciplines les plus actives en traitement d'images. Ses récentes découvertes permettent non seulement un meilleur diagnostic mais offrent aussi de nouveaux espoirs de traitement pour de nombreuses pathologies.

Dans notre démarche, nous nous sommes lancé le défi de développer un système simple et performant d'interprétation des images mammographiques dans le but de pouvoir faire une détection précoce du cancer du sein. En effet, le cancer du sein constitue la cause de décès la plus fréquente chez la femme [38]. L'étape du diagnostic, est une étape cruciale dans la lutte contre le cancer du sein. Par conséquent, la détection du cancer, l'analyse et le traitement du cancer sont devenus un grand axe de la recherche.

L'amélioration de la qualité visuelle des clichés mammographiques et la segmentation de son contenu en régions reste un problème important pour l'aide à la détection précoce du cancer du sein par les spécialistes du domaine. En effet, l'interprétation de ce type d'images n'est pas aisée à cause de la forte densité mammaire d'une part et du très faible contraste entre les tissus sains et les zones cancéreuses d'autre part.

Par conséquent, la conception d'un système de détection assistée par ordinateur (CAD) représente un système d'aide pour les radiologistes dans l'interprétation des mammographies pour le dépistage des masses et des calcifications [31].

Nous avons proposé dans ce chapitre notre approche d'interprétation des images mammographiques, qui s'appuie sur des connaissances du domaine. Notre approche permet la détection des images cancéreuses en classifiant les images mammographiques en images normales et cancéreuses.

# **CHAPITRE V :**

*Expérimentations et résultats  
obtenus*

## **V.1 Introduction**

Le rôle d'interprétation des images mammographiques est de fournir au médecin des informations symboliques de haut niveau sur le contenu de l'image. Dans un but d'aide au diagnostic, ces informations doivent expliciter les différentes régions d'intérêt ainsi que les aspects pathologiques des structures présentes dans l'image. L'interprétation des images médicales représente un outil efficace pour le traitement de l'information visuelle, elle permet inévitablement non seulement de détecter et de localiser la tumeur mais aussi de spécifier ou plutôt de prévoir la gravité de la tumeur, en termes de malignité/bénignité, en exploitant les caractéristiques extraites de l'image. L'interprétation d'images est un problème complexe qui peut être défini comme l'extraction automatique de la sémantique d'une image. Cependant, cette sémantique n'est pas toujours explicite dans l'image elle-même. Elle dépend d'une part des connaissances a priori sur le domaine et d'autre part du contexte de l'interprétation.

### **Pourquoi nous détectons les opacités mammaires ?**

- ❖ De paramètres quantitatifs pour préciser la nature des lésions : Selon le standard
- ❖ Opacités mammaires doit être détecté dans sa première phase pour maximiser les chances de survie.
- ❖ Réduire l'erreur commise par les radiologues : La classification des lésions mammaires par un radiologue est une classification humaine subjective qui peut facilement classifier la même lésion de deux manières différentes.
- ❖ Aider les radiologues dans leurs interprétations : Plusieurs études ont montré que les radiologues manquent entre 4% et 38% de détection de cancers et que ce taux s'améliore de 15% en utilisant une deuxième lecture vue que l'interprétation est souvent difficile et dépend de l'expertise du radiologue.

Dans le présent chapitre nous détaillons la partie expérimentale de notre travail. Nous décrivons les expérimentations réalisées, les résultats expérimentaux obtenus, les bases d'images mammographiques utilisées et le logiciel développé.

## **V.2 Les Bases de données des images de mammographie**

Afin de tester notre système CAD proposé nous avons utilisé la base d'images médicales DDSM (Digital Database for Screening Mammography). Cette base est la plus grande base publique de mammographie contenant 2620 dossiers de patients classés en images normales et cancéreuses [17].

La deuxième base utilisée est la base mammographique mini-MIAS [55].

Aussi nous avons utilisé des images mammographiques obtenus du centre hospitalier régional d'Oran (EHU Oran).

## **V.3 Implémentation et présentation de l'interface**

Il est nécessaire de fournir à l'utilisateur un outil permettant de visualiser les images obtenues par les segmentations décrites aux chapitres précédents. Pour cela, une interface utilisateur a été développée. Nous présenterons brièvement l'environnement de ce logiciel puis détaillerons les étapes de sa réalisation et les différentes structures qui la constituent, ainsi que les résultats obtenus par les différents algorithmes élaborés.

### **V.3.1 Langage de programmation**

Notre application a été programmée par le langage de programmation Matlab de version (R2011). Notre choix est justifié par la souplesse et la facilité du travail. Matlab (matrix laboratory) est une bibliothèque très riche et simple à utiliser. Elle nous a permis d'utiliser ses propres fonctions [37].

Matlab est un langage très performant utilisé pour :

- ❖ Le calcul scientifique
- ❖ L'acquisition des données
- ❖ Le développement des algorithmes
- ❖ La modélisation et simulation
- ❖ L'analyse, l'exploration et visualisation des données et des résultats

L'outil Guide (GUI) nous a bien facilité la tâche de la réalisation de notre interface graphique.

Elle est assurée en général par la création des :

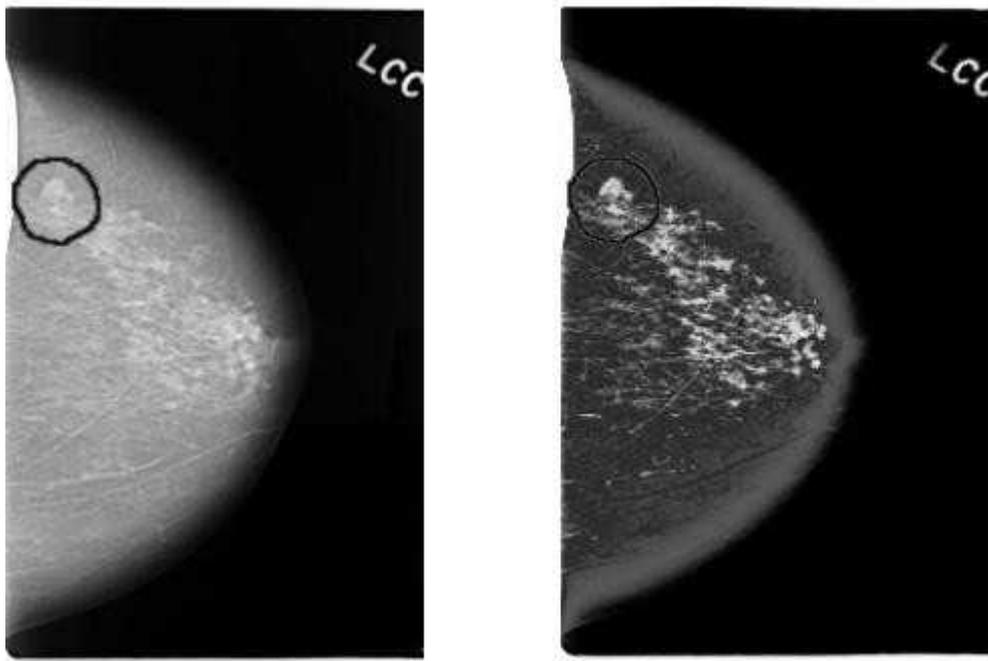
- ❖ Fenêtres qui permettent l'affichage des images.
- ❖ L'utilisation d'un menu principale avec des sous menus qui renferment toutes les fonctionnalités de ce logiciel.
- ❖ Une barre d'outils qui facilite l'accès directe aux fonctionnalités d'application

## **V.4 Expérimentation et résultats**

Dans cette section nous allons évaluer notre système proposé pour l'interprétation des images de mammographie pour la détection du cancer.

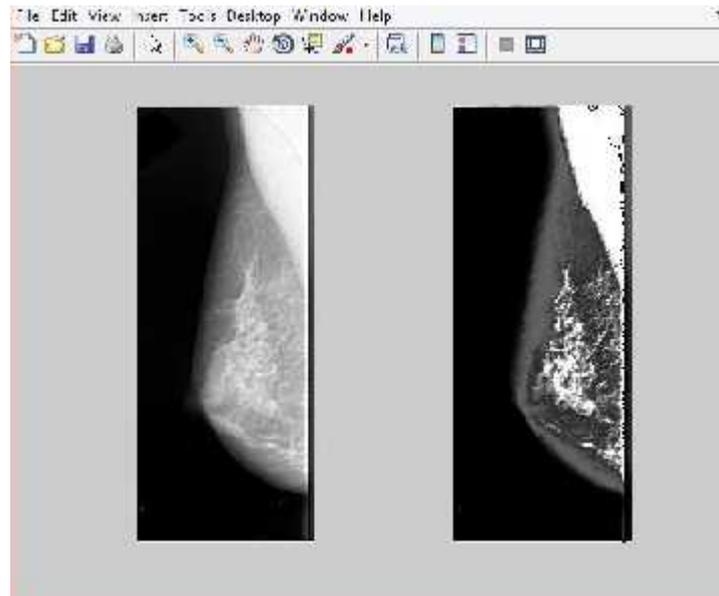
### **V.4.1 résultats obtenus dans l'étape de prétraitement (amélioration de la qualité de l'image)**

Plusieurs expériences ont été effectuées sur les différentes bases d'images. La figure 5.1 est un exemple de mammographie de cancer. L'image de gauche est l'image originale tandis que l'image de droite est l'image après une amélioration de sa qualité en utilisant l'algorithme de distribution de possibilité.

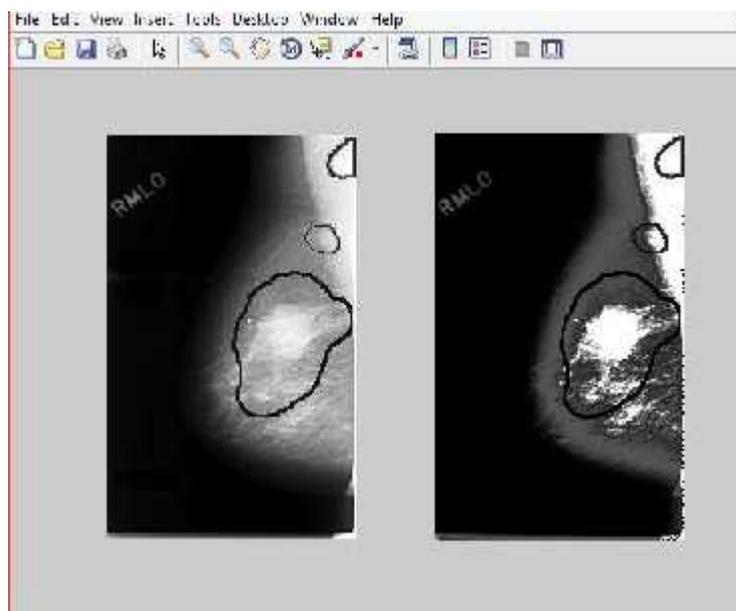


**Figure 5.1 - Exemple d'amélioration de la qualité de l'image.**

Les figures 5.2 la figure 5.3 présentent les résultats de l'étape de prétraitement sur des images normales et cancéreuses. Les figures 5.2 est le résultat d'une expérience avec un ensemble de données de mammographie normale, tandis que la figure 5.3 est le résultat d'une expérience avec un ensemble de données de mammographie du cancer.



**Figure 5.2 - Résultat de l'amélioration d'une image mammographique normale.**



**Figure 5.3 - Résultat de l'amélioration d'une image mammographique du cancer.**

## **V.4.2 Résultats de l'étape de segmentation**

La figure 5.4 montre les résultats de la segmentation d'une image mammographique en utilisant la méthode de OTSU.

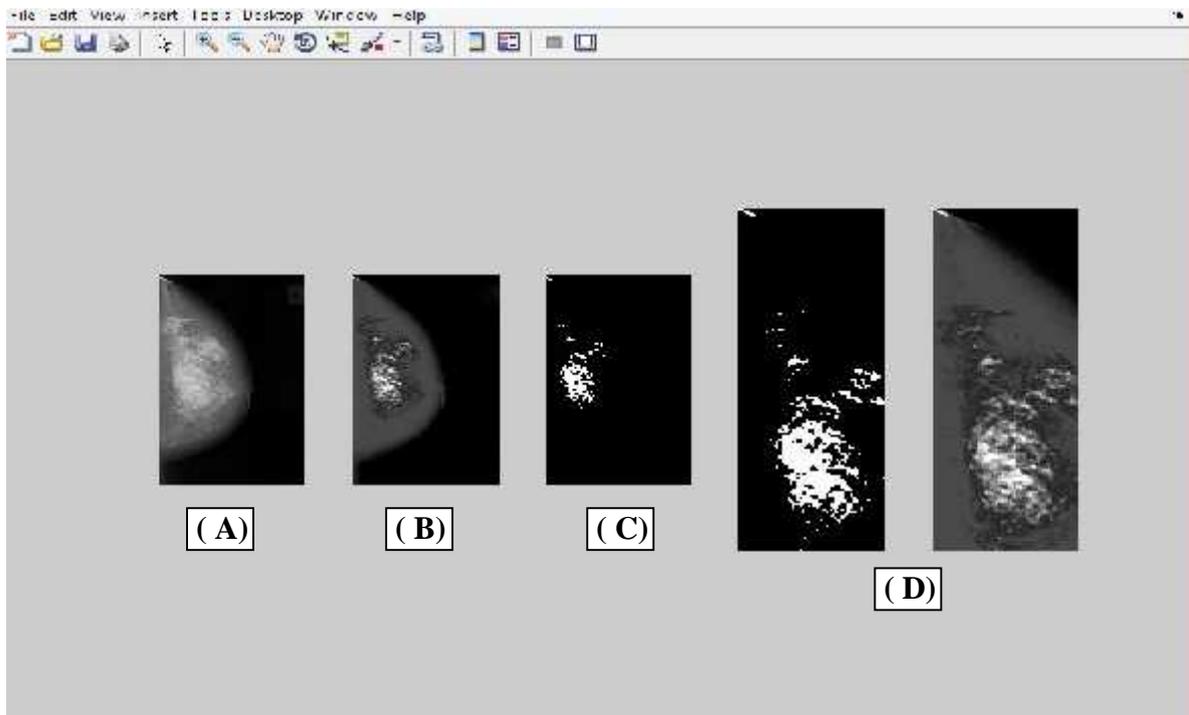


Figure 5.4 – Résultat de l'étape de segmentation, (A) l'image originale, (B) l'image après amélioration, (C) l'image après segmentation et (D) un zoom sur l'objet d'intérêt.

#### V.4.3 Résultat de l'étape d'extraction des caractéristiques

Cette étape permet l'extraction des caractéristiques, nous avons utilisé 6 caractéristiques. Ces derniers sont utilisés dans l'étape de classification finale. Le tableau 5.1 donne les valeurs de caractéristiques obtenues à partir d'un ensemble d'images normales et cancéreuses.

Images	contrast	Ecart-Type	curtosis	Moyenne	variance	lissage
cancer001	1.027975	0.628802	9.451290	0.539653	0.395393	0.283356
cancer002	1.398750	0.685388	9.736556	0.578350	0.469757	0.319615
cancer004	0.737317	0.605021	7.130871	0.527126	0.366050	0.267962
cancer005	0.649256	0.952167	3.081397	0.956720	0.906621	0.475512
cancer008	0.791496	0.592010	14.389570	0.418011	0.350476	0.259520
normal001	0.192300	0.148253	34.168940	0.188447	0.021979	0.021506
normal002	0.369617	0.225630	27.856770	0.224310	0.050909	0.048443
normal003	0.313964	0.173380	6.763178	0.195019	0.0300600	0.029183
normal005	0.750176	0.444575	15.891920	0.410340	0.197647	0.165030

**Table 5.1 – Les différentes valeurs des caractéristiques obtenues pour des images cancéreuses et normales.**

#### **V.4.4 Résultats de l'étape de classification**

L'étape de classification est la dernière étape du système d'interprétation permettant la classification des images en images normales ou cancéreuses.

Le réseau MLP avec 5 couches est utilisé pour la classification. L'apprentissage du réseau est effectué avec la base mini-MIAS.

Après l'obtention des poids finaux de l'étape d'apprentissage nous pouvons tester le réseau avec n'importe quelle image et ce dernier peut détecter les images qui présentent un cancer.

Les figures 5.5, 5-6, 5-7 et 5.8 présentent les différentes étapes de notre système d'interprétation, la figure 5.9 montre le résultat final de classification. La décision finale de notre système pour l'image de cet exemple est qu'elle est cancéreuse. La figure 5.9 présente une portion de code montrant les différentes valeurs des caractéristiques obtenues de cette image cancéreuse.

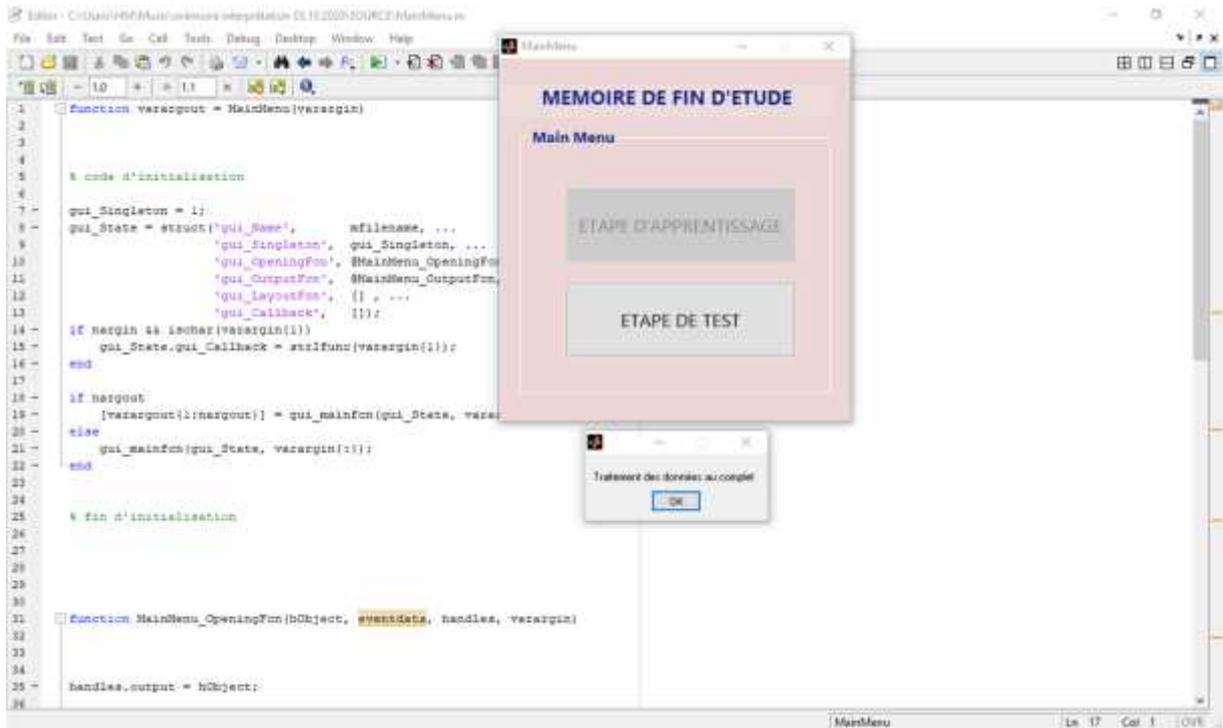


Figure 5.5 – Etape 1 de notre système d'interprétation (début).

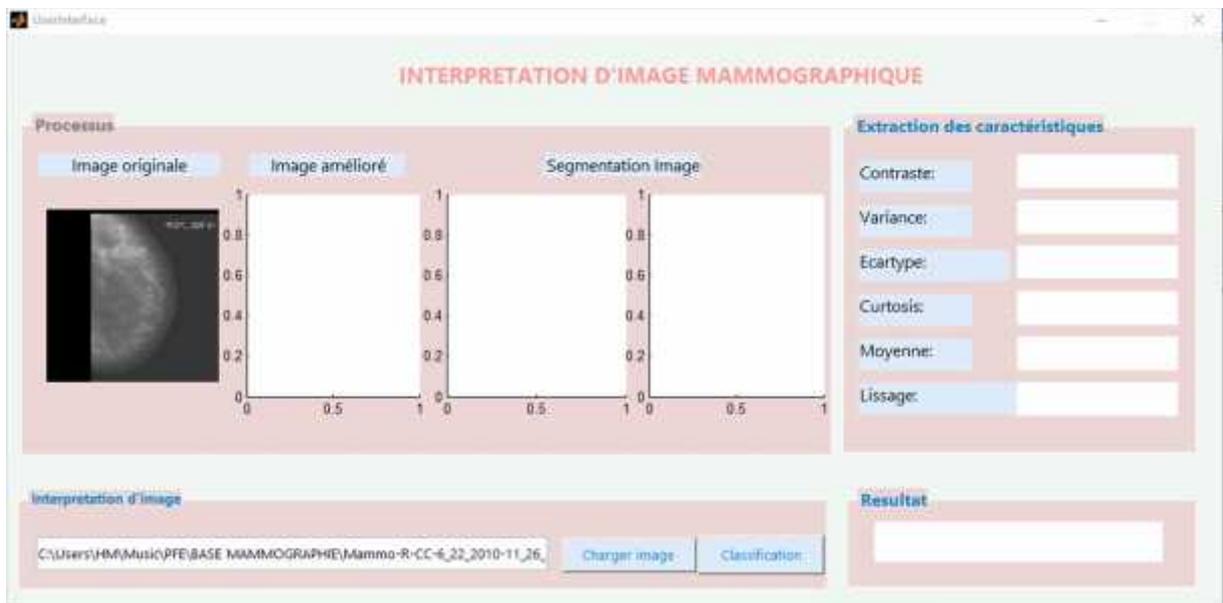


Figure 5.6 – Etape 2 de notre système d'interprétation (chargement de l'image test).

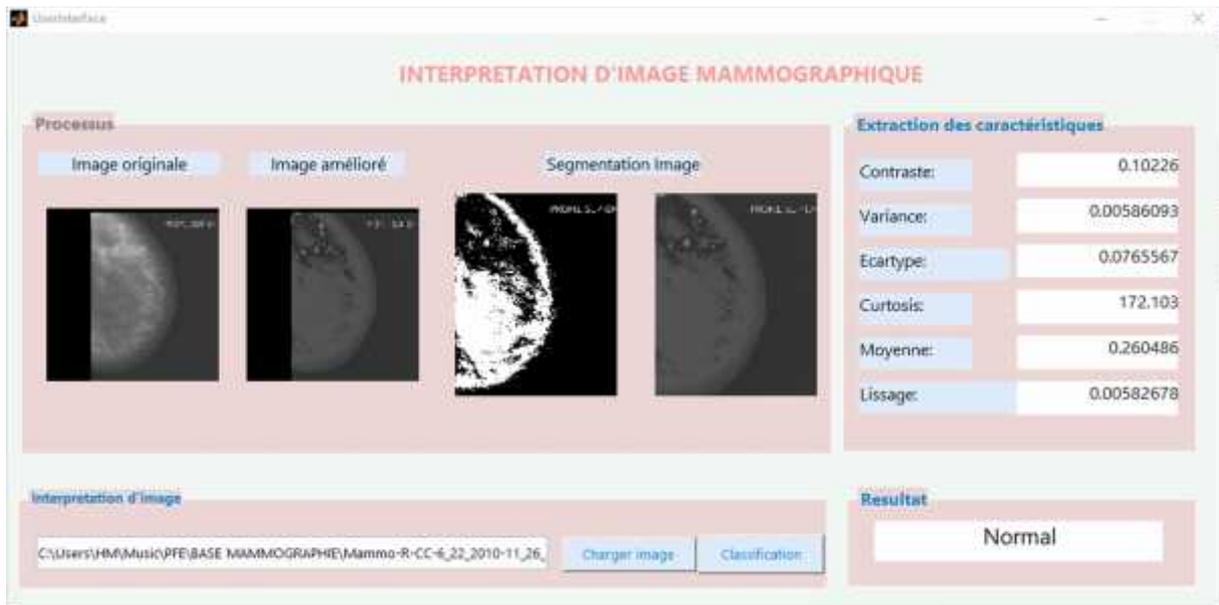


Figure 5.7 – Etape 3 de notre système d’interprétation (amélioration, segmentation et extraction des caractéristiques).

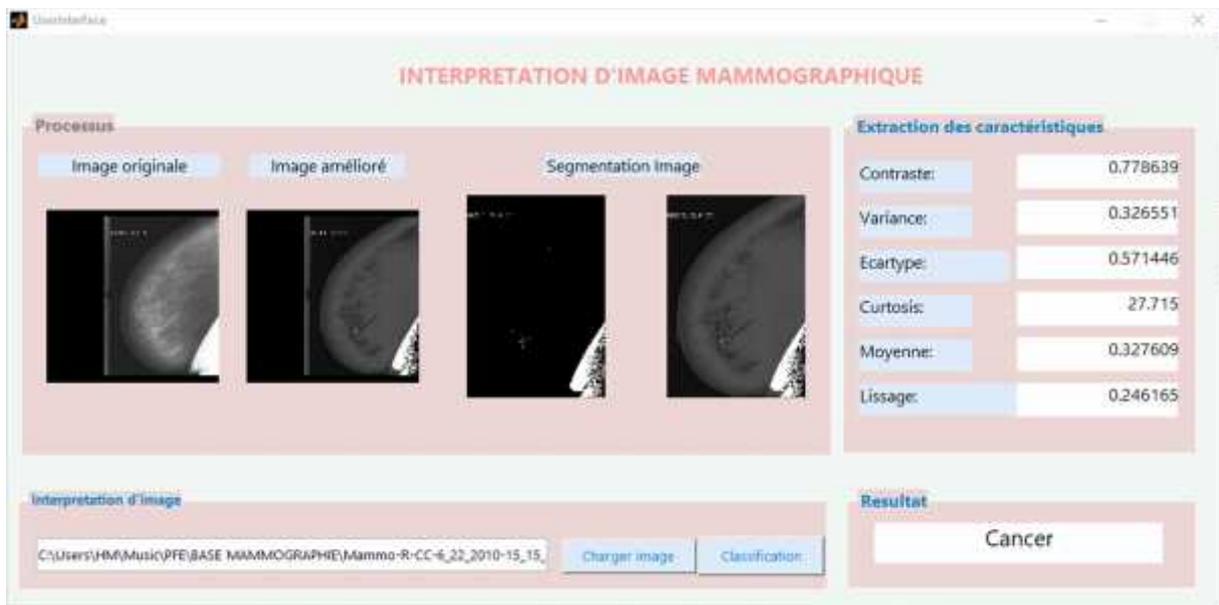
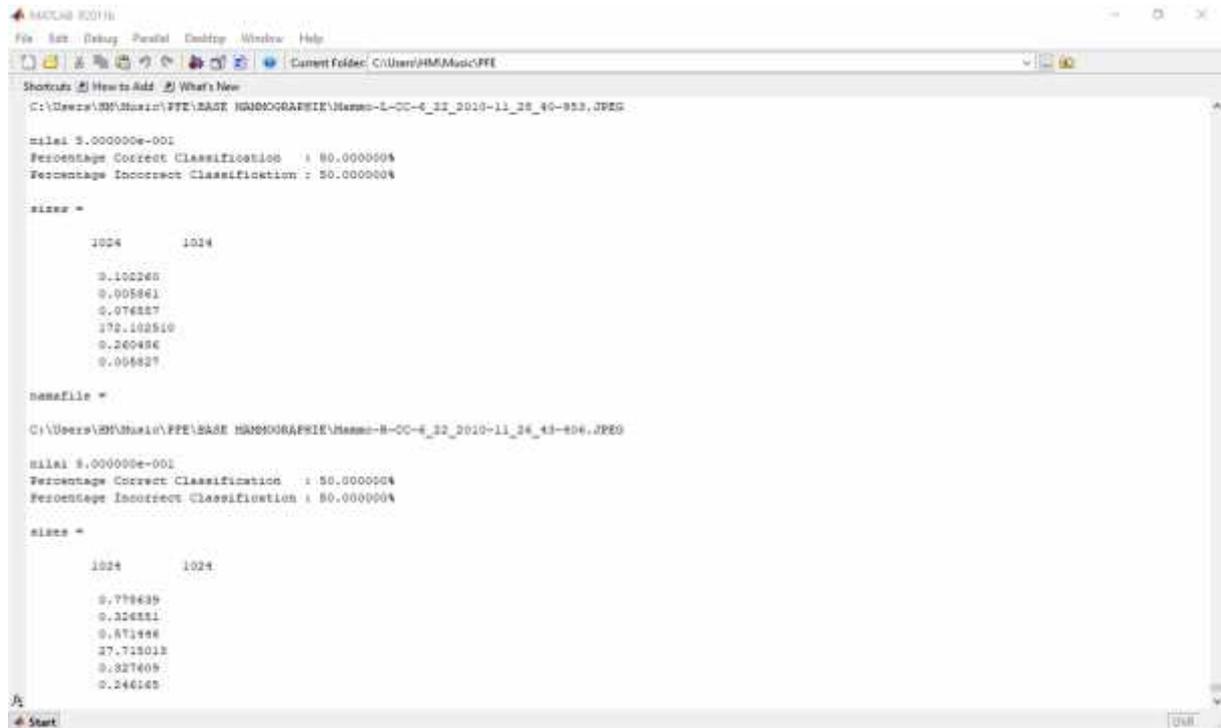


Figure 5.8 – Etape 4 de notre système d’interprétation (classification finale).



```

C:\Users\BHV\Music\FFE\BASE MAMMOGRAPHIE\Hamm-B-CC-8_22_2010-11_26_43-833.JPG

trial 3.000000e-001
Percentage Correct Classification : 50.000000%
Percentage Incorrect Classification : 50.000000%

size =

    1024    1024

    0.100268
    0.005861
    0.076887
    172.182510
    0.200484
    0.006827

namefile =

C:\Users\BHV\Music\FFE\BASE MAMMOGRAPHIE\Hamm-B-CC-8_22_2010-11_26_43-836.JPG

trial 3.000000e-001
Percentage Correct Classification : 50.000000%
Percentage Incorrect Classification : 50.000000%

size =

    1024    1024

    0.778439
    0.124811
    0.871988
    17.715013
    0.827409
    0.246145

```

Figure 5.9 – Une portion du code montrant les valeurs des caractéristiques obtenues.

## V.5 Conclusion

Dans la vision par ordinateur, ce vaste domaine de recherche à la croisée des chemins entre les mathématiques, le traitement du signal et l'intelligence artificielle, la segmentation des images est une tâche très délicate et nullement aisée. Elle demande une connaissance précise des images, leur nature et le domaine d'application. C'est ainsi que nous avons pris en considération l'aspect médical durant l'étude de notre sujet, pour mieux comprendre notre sujet. Dans cette optique, nous avons proposé une nouvelle approche pour la détection des calcifications sur les images mammographiques.

Nous avons présenté dans ce chapitre l'implémentation de notre système d'aide au diagnostic des images mammographiques ainsi que les résultats obtenus sur plusieurs bases d'images de mammographie.

Les résultats obtenus ont été encourageants pour toutes les expériences.

# Conclusion Générale et Perspectives

## **Conclusion générale**

Dans ce travail, nous sommes intéressés à l'étude sur les imageries médicales pour l'aide au diagnostic du cancer du sein, notamment à la description des masses mammaires. En effet, les statistiques confirment que le cancer du sein représente une menace prépondérante pour la vie de la femme. Cependant, une telle menace n'est gérée que par la prise en charge rapide de la maladie pour maximiser les chances de survie.

Nous avons proposé une méthodologie globale d'interprétation d'images, qui s'appuie sur des connaissances du domaine. Nous avons également montré comment ces connaissances du domaine devaient être traduites en des connaissances de scène.

La recherche en imagerie médicale est une des disciplines les plus actives en traitement d'images. Ses récentes découvertes permettent non seulement un meilleur diagnostic mais offrent aussi de nouveaux espoirs de traitement pour de nombreuses pathologies.

Dans notre démarche, nous nous sommes lancés le défi d'améliorer la qualité visuelle des clichés d'images mammographiques dans le but de pouvoir faire une détection précoce du cancer du sein. En effet, le cancer du sein constitue la cause de décès la plus fréquente chez la femme. L'étape du diagnostic, est une étape cruciale dans la lutte contre le cancer du sein. Par conséquent, la détection du cancer, l'analyse et le traitement du cancer sont devenus un grand axe de la recherche.

- ❖ L'amélioration de la qualité d'image des données de mammographie pour détecter le cancer du sein à l'aide de 5 (cinq) algorithmes d'approche floue. L'un des algorithmes utilisés dans l'achèvement de cette tâche est l'algorithme de distribution des possibilités. Dans l'ensemble, cet algorithme abaisse la grisaille des pixels qui ont une valeur grisâtre entre 1 et 2. D'après les résultats de l'exécution du programme, on peut conclure que cet algorithme conserve une valeur très sombre et très brillante avec une approche floue, tandis que la valeur qui est autour de la moyenne (moyenne) est évaluée dans la direction opposée à la valeur moyenne. Ainsi, le résultat est une image qui accentue le côté de l'image d'un ensemble de données très lumineuses et très sombres, qui représente le tissu cutané et peut être traité pour analyser le cancer du sein.
- ❖ La segmentation à l'aide de seuils trouve des sections plus petites à couper et à analyser à l'étape suivante.
- ❖ Au stade d'extraction de la fonctionnalité, après avoir essayé certaines

fonctionnalités, nous avons trouvé 6 caractéristiques qui représentent les mammographies. Nous prenons de GLCM, les caractéristiques d'intensité, et l'intensité des caractéristiques histogrammes. De la fonction GLCM nous prenons la valeur de *contraste*, de la fonction d'intensité, nous prenons la valeur *moyenne*, la *variance*, *l'écart type*, de la fonction d'intensité histogramme, nous prenons la *curtose* et la *douceur*.

- ❖ La classification à l'aide d'un algorithme de réseau neuronal divise 6 entrées de l'extraction de fonctionnalités en deux sorties (normale et *cancer*) avec une taille de *couche cachée* de 5. Formation des données à l'aide de 60 données de mammographie (30 données normales et 30 données sur le *cancer*) et avec des données) test 18 mammographies de données (8 données normales et 10 données sur le cancer). D'après les données de test, 3 données ne sont pas précises dans la classification.

## **Perspectives et suites possibles de ce travail...**

Ce mémoire a permis de dégager de nombreuses perspectives que nous les résumons dans ces quelques lignes :

1. Une forte liaison relie le concept de segmentation et classification : après l'étape de l'extraction des paramètres pertinents une classification peut être envisagée pour identifier une anomalie en type bénigne ou maligne.
2. Une extension logique de cette approche est l'automatisation de système. Bien entendu cette liste est non exhaustive et un bon nombre d'extensions supplémentaires peuvent certainement être imaginées.
3. Lorsque les masses mammaires sont importantes, nous proposons une étape de description de ces formes obtenues à partir de l'étape de segmentation car une telle caractérisation des masses est délicate. En effet, la considération d'une lésion maligne en tant que bénigne pose un problème puisque la patiente réellement atteinte d'un cancer n'aura pas le traitement adéquat pour maximiser ses chances de survie. En revanche, considérer une lésion bénigne en tant que maligne fait croître le taux des biopsies inutiles sans oublier l'impact psychologique d'une telle information sur les patientes.
4. L'étude locale de la texture de ces différentes masses mammaires. En effet, en plus de la classification des masses en malignes/bénignes, l'étude de la nature de la texture

permet de différencier celle-ci entre les différentes classes et élaborer ainsi les niveaux d'atteinte comme celles utilisées par l'ACR (American College of Radiology).

5. L'utilisation des modèles déformables pour l'identification des foyers de microcalcification semble très intéressante car ces modèles se basent sur l'évolution de courbes ou les méthodes Levels set (ensembles de niveaux). Les courbes et les surfaces évoluent par le biais de mesures géométriques résultant d'une évolution, et donc indépendantes de paramètres, les courbes sont représentées comme un ensemble de niveaux d'une fonction de dimension supérieure. Les changements de topologie sont pris en compte automatiquement, le modèle de Chunming Li serait intéressant dans ce cas-là.
6. Une autre perspective est de considérer dans le cas de la gaussienne avec une grande valeur de  $\sigma$  pour obtenir une énergie potentielle créant une large vallée au niveau des contours des masses ou calcifications. La valeur de  $\sigma$  est ensuite réduite pour affiner le plaquage du modèle contre le contour.

# **Références bibliographiques**

## REFERENCES

- [1] A. B. Aboul Ella Hassanien. A Comparative Study on Digital Mamography Enhancement Algorithms Based on Fuzzy Theory, *Studies in Informatics and Control*, Vols. 12, No.1, pp. 21-31, 2003.
- [2] A. Ammadi. Détection automatique de phénomènes présents dans une image satellitaire en composition colorée, Mémoire de master. L'université mohammed 5, Maroc, 2008.
- [3] O. Assas. Classification floue des images, Doctorat en sciences Université de Batna 2013.
- [4] Association de dépistage : Haute autorité de la santé : « Place de la mammographienumérique dans le dépistage organisé du cancer du sein » Synthèse et perspectives.  
[http://www.hassante.fr/portail/upload/docs/application/pdf/synthese\\_mammo\\_numerique.pdf](http://www.hassante.fr/portail/upload/docs/application/pdf/synthese_mammo_numerique.pdf)  
(Consulté en Octobre 2020).
- [4] H. Ayadi, B. Bouanoun. Système d'aide à la détection de maladies sur image médicale basé sur les réseaux de neurones, mémoire d'ingénieur d'état en informatique, 2010.
- [5] B. Banrezes. La modélisation 3D : contribution et nouvelles perspectives en Neuroanatomie fonctionnelle, Thèse de doctorat. L'Université de Paris 6, France, 2002.
- [6] J.C. Baillie. Cours de Segmentation, Module D9 : Traitement d'images et vision artificielle. ENSA, 2003.
- [7] M. H. Bendaoud. Développement de méthodes d'extraction de contours sur des images à niveaux de gris, thèse pour l'obtention du diplôme de doctorat, Université des sciences et de la technologie d'Oran USTO-MB, 2017.
- [8] U. C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, M. Heynen. Multi resolution object-oriented fuzzy analysis of remote sensing data for GIS-ready information, *ISPRS Journal of Photo grammetry & Rmote Sensing*, 58(3-4), pp.239-258, 2004.
- [9] T. R Berry, K. D. Curtin. Heart disease and breast cancer perceptions: Ethnic differences and relationship to attentional bias, *Journal of Health Psychology Open*, Volume: 3, issue: 2, juillet 2016.
- [10] L.S.A. Bins, L. M. G. Foncseca, G.J. Erthal, F. M. Ii. Satellite imagery segmentation : A region growing approach, Dans 8 Brazilian Symposium Remote Sensing, pp.677-680, 1996.
- [11] P. Bolon, J. Chassery, J. Cocquerez, D. Demigny, C. Graffigne, et al.. Analyse d'images : Filtrage et segmentation. Jean-Pierre Cocquerez; Sylvie Philipp. MASSON, 1995, Enseignement de la physique, 2-225-84923-4. hal-00706168.

- [12] A. Boujelben, H. Tmar, J. Mnif, M. Abid, Automatic application level set approach in detection calcifications in mammographic image, *International Journal of Computer Science & Information Technology (IJCSIT)*, Vol 3, No 4, pp 1-14, 2011.
- [13] K. Chakib. Compression des images fixes par les approximations fractales basée, mémoire de fin d'études, 1999.
- [14] M. T. Cheikh. Amélioration des images par un modèle de réseau de neurones (Comparaison avec les filtres de base), Mémoire de fin d'études pour l'obtention du diplôme de Master en Informatique, Université Abou-Bakr Belkaid –Tlemcen, 2011.
- [15] I. CheikhRouho. Description et classification des masses mammaires pour le diagnostic du cancer du sein, thèse pour l'obtention du titre de docteur en traitement des images et du signal, Université d'Evry-Val d'Essonne, France, 27 juin 2012.
- [16] C. Darolti, E. Barth . Active Contours that Grow and Compete Driven by Local Region Descriptors, *Proceedings of SPIE - The International Society for Optical Engineering* 7245:72450, Février 2009.
- [17] “DDSM: Digital Database for Screening Mammography”. En ligne sur : <http://marathon.csee.usf.edu/Mammography/Database.html>. (Consulté en mars 2020).
- [18] C. D. Djalal, Kermad, C. Kacem. Automatic Image Segmentation System through Iterative Edge-Region Co-Operation, *Image and Vision Computing*, vol. 20, no.8, 2002.
- [19] F. Eddaoudi. Aide à la détection des pathologies du sein par les techniques de segmentation et de codage de textures des images mammographiques, thèse de doctorat, université mohammed v-Agdal, Rabat, Maroc , 2012.
- [20] M. A. El-Sayed, A. A. Ali, M. E. Hussien , H. A. Sennary. A multi-level threshold method for edge detection and segmentation based on entropy, *Computers, Materials & Continua*, vol. 63, no.1, pp. 1–16, 2020.
- [21] M. A. Guerroudji. Segmentation des clichés mammographiques en vue de la détection des foyers de microcalcifications mammaires : application à la base de données CHU de Tizi-Ouzou, Thèse de doctorat LMD, Université Mouloud Mammeri, 2017.
- [22] A. Gupta. Current research opportunities of image processing and computer vision, *Computer Science* 20(4), décembre 2019.
- [23] I. Hadjidj. Approche morphologique pour la segmentation d'images médicales, mémoire de magister en électronique biomédicale, Université Abou-Bakr Belkaid de Tlemcen, 2011.
- [24] A. Herbulot. Mesures statistiques non-paramétriques pour la segmentation d'images et de vidéos et minimisation par contours actifs, thèse pour obtenir le titre de Docteur en Sciences Automatique, Traitement du Signal et des Images, Universités Nice Sophia-Antipolis, France, 10 octobre 2007.

- [25] S. L. Horowitz, T. Pavlidis. Picture segmentation by tree transversal algorithm, J.ACM., Vol. 32, 2, pp. 368 – 388, 1976.
- [26] <http://www.info-radiologie.ch/mammographie.php> (Consulté en octobre 2020).
- [27] <https://www.bordet.be/fr/infosmed/brochure/sein/techniq.htm> (Consulté en octobre 2020).
- [28] <https://www.cancer.ca/en/about-us/a-future-without-breast-cancer/>, (Consulté en octobre 2020).
- [29] <http://tpebaudechahdeldicque.e-monsite.com/pages/ii-les-differents-principes-d-imagerie-medicale-2/ii-les-differents-principes-d-imagerie-medicale.html> (Consulté en octobre 2020).
- [30] Image Segmentation in 2020: Architectures, Losses, Datasets, and Frameworks, <https://neptune.ai/blog/image-segmentation-in-2020>, publié le 9 mars, 2020.
- [31] M.C. Jaulent, P. Degoulet. Systèmes d'aide à l'interprétation des images médicales: application à l'interprétation des angiographies numérisées, service d'informatique médicale, hôpital broussais paris, 1992.
- [32] I. Kessour, I. TALI. Simulation des contours actifs par les colonies de fourmis, Pour l'obtention du diplôme d'Ingénieur d'Etat en Informatique, 2011.
- [33] J. Lecoer, C. Barillot. Segmentation d'images cérébrales : Etat de l'art, Rapport de recherche, Institut INI, version révisée en 2008.
- [34] D. Lingrand. Introduction au Traitement d'images, Paris, livre : Éditions Vuibert, 2<sup>ème</sup> édition, 256 pages, ISBN : 978-2-7117-4866-2, 2008.
- [35] S. Liu , E.J. Delp. Multiresolution detection of stellate lesions in mammograms, IEEE transactions on medical imaging, pp 10-112, 1997.
- [36] Y. Lu, Z. Jiang, T. Zhou, S. Fu. An Improved Watershed Segmentation Algorithm of Medical Tumor Image, IOP Conf. Series: Materials Science and Engineering 677 (2019) 042028, 2019.
- [37] "MathWorks," En ligne sur : <http://www.mathworks.com/>. (Consulté en mars 2020)
- [38] C. Mechta. Interprétation des images de mammographie par l'ontologie, mémoire pour l'obtention du diplôme de magister en informatique, Université des sciences et de la technologie d'Oran USTOMB, 2013.
- [39] M. Meliani. Segmentation d'image par coopération régions-contours, mémoire en vue de l'obtention du diplôme de magister en informatique, Ecole nationale supérieure en informatique Oued-Smar, Alger, 2012.
- [40] The mini-MIAS database of mammograms, disponible en ligne en 2020 sur:

<https://www.repository.cam.ac.uk/handle/1810/250394>.

- [41] M. Mustafa, N.O. Rashid, R. Samad. Breast Cancer Segmentation Based On GVF snake, Conference: 2019 IEEE Conference on Biomedical Engineering and Sciences (IECBES), December 2019.
- [42] R. Nithy, B. Santhi. Classification of Normal and Abnormal Patterns in Digital Mammograms for Diagnosis of Breast Cancer, International Journal of Computer Applications (0975 – 8887), Vols. 28-No.6, pp. 21-25, 2011.
- [43] N. Otsu. A Threshold Selection Method from Gray-Level Histogram, Transactions on Systems, Man, and Cybernetics, Vols. SMC-9 No.1, pp. 62-66, 1979.
- [44] G. Palma. Détection automatique des opacités en tomosynthèse numérique du sein, Thèse de doctorat en Signal et images, Télécom Paris Tech, France . février 23, 2019.
- [45] Priyanka, B. Singh. A review on brain tumor detection using segmentation, IJCSMC, Vol. 2, issue 7. pages (48-54), Juillet 2013.
- [46] R.M. Rangayyan, D. Guliato, W.A. Carnielli, J.A. Zuffo et J.E.L. Desautels, Segmentation of breast tumors in mammograms by fuzzy region growing, IEEE. Proceedings of the 20th Annual International Conference of the, vol. 2, pp. 1002-1005, 1998.
- [47] R.M. Rangayyan. Recent Advances in Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer”, Editor(s) : Jasjit S. Suri; SPIE Press Book, April 2006.
- [48] A. Rick. Représentation de la variabilité dans le traitement d’images flou, application à la mammographie numérique, Thèse de doctorat, Université Paris VI, 1999.
- [49] F. Sadoughi, Z. Kazemy. Artificial intelligence methods for the diagnosis of breast cancer by image processing: a review, Breast Cancer (Dove Med Press), 10: 219–230, 2018.
- [50] N. Safdarian, Mohammad Reza Hediyezhadeh. Detection and Classification of Breast Cancer in Mammography Images Using Pattern Recognition Methods, journal of Multidisciplinary Cancer Investigation, Volume 3, Issue 4, Octobre 2019.
- [51] N. Sharma, L. M. Aggarwal. Automated medical image segmentation techniques, Journal of Medical Physics, 35(1): 3–14, 2010.
- [52] T. Rachida. Segmentation des images médicales par ondelettes, mémoire de magister en imagerie vision artificielle et robotique médicale, Université des sciences et de la technologie d’Oran, 2011.
- [53] C. Vachier. Extraction de caractéristiques, segmentation d’image et morphologie mathématique, thèse de doctorat, école des mines de Paris, 1995.
- [54] F. Yi, I. Moon. Extraction of target specimens from bioholographic images using interactive graph cuts. Journal of Biomedical Optics 18(12):126015, décembre 2013.