

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

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

جامعة سعد دحلب البليدة
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Projet de Fin d'Études

présenté par

Djeboua Ibrahim

pour l'obtention du diplôme de Master en Électronique option 'Traitement de
l'Information et Systèmes Electroniques'

Thème

Détection connexionniste de tumeurs mammaires pour l'aide à la décision

Proposé par: Reguieg F. Zohra & Benblidia Nadjia

Année Universitaire 2014-2015

Dédicace.....

....Je dédie ce mémoire à :

- Mes parents :*

Ma mère, qui a œuvré pour ma réussite, de par son amour, son soutien, tous les sacrifices consentis et ses précieux conseils, pour toute son assistance et sa présence dans ma vie, reçois à travers ce travail aussi modeste soit-il, l'expression de mes sentiments et de mon éternelle gratitude.

Mon père, qui peut être fier et trouver ici le résultat de longues années de sacrifices et de privations pour m'aider à avancer dans la vie. Puisse Dieu faire en sorte que ce travail porte son fruit : Merci pour les valeurs nobles, l'éducation et le soutien permanent venu de toi.

- Mes frères: Halim, Youcef, Kheireddine ;*
- Mes Sœur: Yousra et Hadjer;*
- Toute ma famille;*
- Mes amis: Hicham, Ahcene, Hassen, Abubakar, Mounir, Billel, Abdelmoumen;*
- Mes collègues (e) de classe;*
- Mes professeurs du département d'électronique(TISE).....*



Remerciements

En préambule à ce mémoire, je souhaiterais adresser mes remerciements les plus sincères à DIEU, pour m'avoir donné la force dans les moments difficiles de réaliser ce mémoire, ainsi qu'aux personnes qui m'ont apporté leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite, de cette formidable année universitaire.

Je tiens à remercier sincèrement *Mlle REGUIEG.Fz*, qui, en tant que Directeur de mémoire, s'est toujours montrée à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'elle a bien voulu me consacrer et sans qui ce mémoire, n'aurait jamais vu le jour.

Je tiens aussi à remercier sincèrement ma co-promotrice *Mlle Benblidia Nadja*, pour ses encouragements, son aide et ses conseils judicieux.

Mes remerciements s'adressent également à *Mr DJERBOUA Amer*, pour sa générosité et la grande patience dont il a su faire preuve, malgré sa charge professionnelle.

J'exprime ma gratitude aux collègues rencontrés lors des recherches effectuées et, qui ont accepté de répondre à mes questions avec gentillesse.

Je n'oublie pas mes parents pour leur contribution, leur soutien et leur patience.

Enfin, j'adresse mes plus sincères remerciements à tous mes proches et amis, qui m'ont toujours soutenu et encouragé au cours de la réalisation de ce mémoire.

Merci à tous et à toutes.

ملخص :

ان تشخيص أمراض الثدي وتصنيفها يشكل حاليا اهتماما كبيرا. في هذا العمل، قمنا بدراسة لمختلف تقنيات التجزئة، من أجل تحليل أشكال الجماهير وتكلسات. وقد سمحت لنا هذه الدراسة بالمضي قدما نحو التعاون بين التجزئة حسب المنطقة المتنامية ومجموعة مستوى النموذج. بعد تجزئة الصور نقوم بالتمييز بينها، من خلال المعلومات المورفولوجية والتكوينية. وأخيرا، فإن تصنيف يكون عن طريق شبكة عصبية تعتمد على أساس المستقبلات متعدد الطبقات

كلمات البحث: أمراض الثدي، التعاون، تجزئة حسب المنطقة المتنامية، مجموعة المستوى، المعلومات التكوينية و المورفولوجية و الشبكة العصبية، متعدد الطبقات المستقبلات

Résumé: le diagnostic des pathologies mammaires et leur classification, suscitent actuellement un grand intérêt. Dans ce travail, un système de détection de pathologies mammographiques est réalisé, pour une aide à la décision. Ce système adopté, opère en quatre étapes: la première permet le prétraitement de l'image suivant les opérateurs de la morphologie mathématique. La seconde, effectue une extraction des structures mammaires suivant une coopération entre une croissance de régions et les modèles déformables. La troisième étape, consiste, à caractériser ces régions d'intérêt, grâce à des paramètres morphologiques et texturaux. La dernière étape, permet une classification des masses et des calcifications, basée sur le perceptron multicouche.

Mots clés : image mammographique, segmentation par croissance de régions, levels set, paramètres de forme et de texture, perceptron multicouche, aide à la décision.

Abstract: the diagnosis of breast diseases and classification currently of great interest. In this work, a study on the various segmentation techniques is realized. In order to analyze the forms of masses and calcifications. This study has allowed us to move towards cooperation between segmentation by region growing and model level set. Once the segmented images were characterized by morphological and textural parameters. Finally, the classification is by a neural network-based multilayer perceptron.

Keywords: mammographic image, segmentation by region growing, level set, textural and morphological parameters, multilayer perceptron, decision assistance.

Table des matières

INTRODUCTION GENERALE.....	1
CHAPITRE 1 CONTEXTE MAMMOGRAPHIQUE.....	3
1.1 Introduction	3
1.2 Anatomie du sein	3
1.3 Cancer du sein.....	4
1.3.1 Types du cancer	5
1.3.2 Signes et symptômes du cancer du sein	6
1.4 La mammographie	7
1.4.1 Le déroulement d'une mammographie	8
1.5 Anomalies mammaires	9
1.5.1 Les masses.....	9
1.5.2 Les calcifications	12
1.5.3 Distorsions architecturales	15
1.6 Les résultats d'une mammographie : la classification ACR	16
1.7 Conclusion	18
CHAPITRE 2 PRETRAITEMENT ET SEGMENTATION EN IMAGERIE MAMMOGRAPHIQUE.....	20
2.1 Introduction.....	20
2.2 Système de détection automatique	21
2.3 Prétraitement	22
2.4 Introduction à la morphologie mathématique	23
2.4.1 Opérateurs de base de la morphologie mathématique	23
2.4.2 Opérateurs morphologiques transformés	25
2.5 Segmentation	28
2.5.1 Approche région	30
2.5.2 Approche contour	34
2.6 Les modèles déformables	34
2.6.1 Les contours actifs (Snakes)	35
2.6.2 Les ensembles de niveaux (Level sets)	36
2.6.3 Le modèle de Cahn et Vese	38
2.6.4 Modèle de Lankton	41
2.7 Approche Coopérative	42
2.7.1 Coopération séquentielle	42
2.7.2 Coopération des résultats	43
2.7.3 Coopération mutuelle	43

2.8 Parcours préliminaire aux travaux du laboratoire LATSI	44
2.9 Méthodologie adoptée pour les masses et les calcifications	44
2.9.1 Méthodologie adoptée pour les masses	44
2.9.2 Méthodologie adoptée pour les calcifications	45
2.10 Conclusion.....	46
CHAPITRE 3 CARACTERISATION ET CLASSIFICATION CONNEXIONNISTE	47
3.1 Introduction	47
3.2 Descripteurs de texture	47
3.2.1 Notion de texture	47
3.2.2 Types de texture	48
3.2.3 Méthode d'analyse de texture	49
3.2.4 Matrice de cooccurrence	50
3.3 Descripteurs de forme	52
3.4 Classification connexionniste	55
3.4.1 Les réseaux de neurones.....	56
3.4.2 Architecture	58
3.4.3 L'apprentissage des réseaux de neurones	63
3.4.4 Avantages et inconvénients des réseaux de neurones	66
3.5 Conclusion	67
CHAPITRE 4 APPLICATION "MAMMOCONNEX" SUR LES IMAGES MAMMOGRAPHIQUES..68	
4.1 Introduction.....	68
4.2 Environnement de travail.....	68
4.2.1 Matériel utilisé.....	68
4.2.2 Présentation de l'outil Matlab.....	68
4.3 Présentation du système développé.....	71
4.3.1 Images utilisées.....	73
4.4 Résultats et interprétation.....	74
4.4.1 Prétraitement.....	74
4.4.2 Extraction des régions d'intérêt.....	76
4.4.3 Extraction des caractéristiques.....	78
4.4.4 La classification.....	84
4.4.5 Evaluation des résultats.....	85
4.5 Conclusion.....	86
CONCLUSION GENERALE.....	87
BIBLIOGRAPHIE	89

Liste des figures :

Figure 1.1 : Structure du sein [3].....	4
Figure 1.2 : La mammographie numérique et le cliché associé [8][9].....	8
Figure 1.3 : Mammographie de dépistage avec quatre incidences [9].....	10
Figure 1.4 : Les différentes densités mammaires [8].....	11
Figure 1.5 : Les types des masses [8].....	11
Figure 1.6 : Les contours des masses [8].....	12
Figure 1.7 : Les différents types de micro-calcifications [8].....	14
Figure 1.8 : Incidences cranio-caudales des deux seins [8].....	15
Figure 1.9 : Classification des images mammographiques [10].....	18
Figure 2.1 : Schéma général d'une chaîne de détection des micro-calcifications ou d'opacités en mammographie 2D [12].....	21
Figure 2.2 : érosion d'une image mammographique.....	24
Figure 2.3 : Dilatation d'une image mammographique.....	25
Figure 2.4 : Ouverture d'une image mammographique par un élément structurant circulaire.....	25
Figure 2.5 : Fermeture d'une image mammographique par un élément structurant circulaire.....	26
Figure 2.6 : Organigramme du prétraitement des images mammographique.....	28
Figure 2.7 : Application de l'algorithme de croissance de région sur les images mammographiques [21].....	33
Figure 2.8 : Principe de contour actif.....	35
Figure 2.9 : Segmentation (contour actif) d'une masse circonscrite [24].....	36

Figure 2.10 : Fonction d'ensembles des niveaux [22].....	38
Figure 2.11 : Principe de la coopération séquentielle [28].....	42
Figure 2.12 : Principe de la coopération des résultat [28].....	43
Figure 2.13 : Principe de la coopération mutuelle [28].....	43
Figure 2.14 : Processus de détection des masses.....	45
Figure 2.15 : Processus de détection des calcifications.....	46
Figure 3.1 : Exemple de texture aléatoire [36].....	48
Figure 3.2 : Exemple de texture périodique [36].....	48
Figure 3.3 : Exemple de texture directionnelle [36].....	49
Figure 3.4 : Plus proches voisins des x selon 4 directions [36].....	51
Figure 3.5 : Schéma d'un neurone biologique [42].....	56
Figure 3.6 : Structure d'un neurone artificiel [42].....	57
Figure 3.7 : Différents types de fonctions de transfert pour le neurone artificiel [42].....	58
Figure 3.8 : Schéma d'un PMC [42][43].....	60
Figure 3.9 : Représentation de la fonction sigmoïde [42].....	65
Figure 4.1 : Fenêtre principale du GUI.....	71
Figure 4.2 : Synoptique du système d'aide à la décision mammographique "Mammoconnex".....	72
Figure 4.3 : Images de la base MIAS (a)Masse bénigne, (b)Masse maligne, (c)Calcifications malignes [44].....	73
Figure 4.4 : Prétraitement d'une masse bénigne [mdb021], (a)Chapeau haut de forme, (b)Superposition.....	74

Figure 4.5 : Prétraitement d'une masse maligne [mdb028], (a)Chapeau haut de forme, (disque de 100), (b)Superposition.....	75
Figure 4.6 : Prétraitement des calcifications malignes [mdb209], (a)Découpage de la zone d'intérêt, (b)Superposition avec Chapeau haut de forme (disque=5).....	76
Figure 4.7 : Détection d'une masse maligne « mdb028 »	77
Figure 4.8 : Détection d'une masse bénigne « mdb021 »	77
Figure 4.9 : Détection des calcifications malignes « mdb209 ».....	78
Figure 4.10 : l'influence du contraste sur les masses.....	86
Figure 4.11 : l'influence de l'entropie sur les masses.....	82
Figure 4.12: L'influence du contraste sur les calcifications.....	83
Figures 4.13 : l'influence de l'entropie sur les calcifications.....	83
Figure 4.14 : Principales étapes de l'apprentissage	84
Figure 4.15 : Performances du réseau	85

Liste des tableaux :

<i>Tableau 4.1 : Paramètres des masses bénignes.....</i>	<i>79</i>
<i>Tableau 4.2 : Paramètres des masses malignes.....</i>	<i>80</i>
<i>Tableau 4.3 : Paramètres des calcifications malignes.....</i>	<i>81</i>
<i>Tableau 4.4 : Paramètres des calcifications bénignes.....</i>	<i>81</i>
<i>Tableau 4.5 : Matrice de confusion pour la catégorie des tests.....</i>	<i>86</i>

Introduction générale

Le 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 huit est touchée par cette maladie durant son existence [1]. 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 est 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.

Face à l'augmentation du nombre de mammographies pendant ces dernières décennies, différents travaux de recherche font l'effort, soit pour détecter automatiquement les lésions mammaires à travers des systèmes de détection assisté par ordinateur, soit pour interpréter automatiquement les mammographies à travers des systèmes de diagnostic assisté par ordinateur.

La mammographie est la modalité qui permet l'acquisition d'images du sein en plusieurs vues, selon un angle de projection et des rayons X, traversant le tissu. Comme tout processus d'acquisition, l'image résultat présente certains artefacts qui, peuvent affecter sa qualité. De même, les différentes tumeurs cancérologiques possibles se concentrent à l'intérieur du sein, rendant leur détection difficile. C'est dans ce contexte, que s'inscrit le travail que nous allons présenter dans notre mémoire, intitulé « détection connexionniste de tumeurs mammaires pour l'aide à la décision ».

Le travail appréhendé dans ce projet, suscite l'intérêt de l'équipe de l'imagerie du laboratoire « LATSI (laboratoire du traitement du signal et de l'image) de l'université de Saad Dahlab de Blida 1 », depuis près de deux décennies, pour la réalisation d'un système d'aide au diagnostic, des clichés mammographiques.

Le but de notre étude, est alors la détection des pathologies mammaires. Dans le cas des masses, elle se fait par une segmentation coopérative entre une croissance de régions et les modèles déformables; dans le cas calcifications, elle s'effectue par les modèles déformables. L'extraction de ces régions, est réalisée après une étape de prétraitement par les opérateurs de la morphologie mathématique. Les régions d'intérêt obtenues sont caractérisées par une analyse morphologique et texturale, pour leur classification, par le biais du perceptron multicouche, pour l'aide à la décision.

Le manuscrit du mémoire est organisé en quatre chapitres.

Dans le premier chapitre, nous introduisons le contexte mammographique, avec des notions sur le cancer du sein, afin de justifier la problématique de notre projet.

Dans le deuxième chapitre, nous commençons tout d'abord, par l'étude des méthodes de prétraitement adoptées dans le cadre de notre travail. Ensuite, nous introduisons l'étude de différentes méthodes de segmentation en régions et les modèles déformables, pour fixer le choix de la méthode de segmentation la plus appropriée, aux types d'images considérées.

Dans le troisième chapitre, nous commençons, par aborder les descripteurs de texture et de forme choisis, pour caractériser les pathologies étudiées dans ce mémoire. La dernière partie est consacrée à la classification connexionniste à l'aide du perceptron multicouche pour une prise de décision.

Dans le quatrième chapitre, nous présentons l'application développée "Mammoconnex" sur les images mammographiques réelles.

1.1 Introduction

Le cancer du sein est le cancer le plus fréquent chez la femme et celui, qui cause le plus de décès dans le monde. Une femme sur huit sera confrontée, au cancer du sein au cours de sa vie [1]. Cependant, il est l'un des cancers que l'on dépiste et que l'on soigne le mieux, s'il est détecté à un stade précoce. Avec 9000 nouveaux cas enregistrés chaque année, le cancer du sein en Algérie connaît une « recrudescence épidémique » et en constitue le type le plus répandu chez la femme. A travers le monde, on dénombre 1 million de sujets atteints par cette pathologie [2]. Son incidence est la plus importante entre 50 et 74 ans; ce qui justifie le choix de cette tranche d'âge pour son dépistage par la mammographie, principale technique d'investigation dans la détection et le diagnostic du cancer du sein.

Dans ce chapitre, on va s'intéresser au contexte médical de notre étude, suivant des notions sur l'anatomie du sein, le cancer du sein, ses types et ses symptômes, pour appréhender la problématique de cette initiation à la recherche.

1.2 Anatomie du sein

Le sein est composé d'une glande mammaire (Figure 1.1), de fibres de soutien (ligaments de Cooper) et de graisse (tissu adipeux); le tout est recouvert par la peau. La quantité de chacune de ses composantes, peut varier d'une femme à l'autre. Le sein est situé par-dessus le muscle pectoral. Il se compose également de nerfs, de vaisseaux sanguins et lymphatiques. La glande mammaire est divisée en 15 à 20 lobes, composés de lobules. Ceux-ci sont reliés à des canaux qui, se rendent sous le mamelon (situé au centre du sein).

On peut également, observer des chaînes de ganglions lymphatiques, qui filtrent les microbes et protègent le corps contre l'infection et la maladie. Le cancer du sein peut se développer, tant au niveau d'un canal galactophore que d'un lobule et il peut également, s'infiltrer au niveau des ganglions lymphatiques [3] [4].

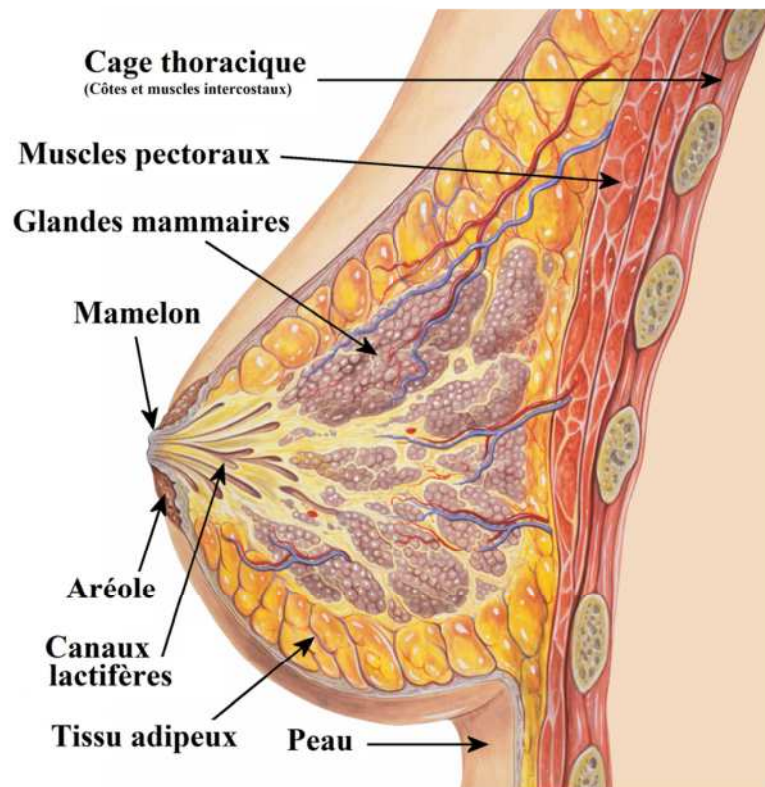


Figure 1.1 : Structure du sein [3]

1.3 Cancer du sein [5][6]

Le cancer du sein correspond à la présence de cellules anormales dans un sein; cellules qui se multiplient de façon anarchique, pour former une tumeur maligne. Selon le type du sein, ces cellules peuvent rester confinées dans le sein ou migrer, vers les ganglions avoisinants, voire le reste du corps. Selon, le type de cellules à l'origine de la pathologie et, selon l'aspect de la tumeur, on distingue différents cancers du sein. Par exemple, un cancer du sein, qui touche les cellules qui bordent les canaux sera dit « **canalaire** » et, un cancer qui affecte les cellules des lobules sera dit : « **Lobulaire** ».

Grâce aux examens cliniques, les médecins déterminent le stade du cancer suivant plusieurs critères dont: la taille et l'infiltration de la tumeur, l'atteinte ou non des ganglions lymphatiques et la présence ou non de métastases

Il existe différents types de cancer du sein. On peut diviser les carcinomes mammaires en deux grandes catégories: cancer non invasif (in situ), cancer invasif (infiltrant).

1.3.1 Types du cancer du sein

a. Cancer du sein non-invasif

Le carcinome canalaire In situ/intra-canalaire (ccis), est le type le plus fréquent de cancer du sein non-invasif chez la femme. C'est une forme de cancer très précoce. Comme son nom l'indique, cette forme de cancer se développe à l'intérieur des canaux de lactation du sein. On parle de cancer du sein « **non invasif** » ou « **in situ** », lorsque la tumeur reste dans le tissu d'origine et n'envahit pas les tissus voisins. Les cellules anormales ne se dispersent pas à l'extérieur des canaux de lactation. On diagnostique beaucoup plus fréquemment ce type de cancer depuis l'utilisation plus répandue de la mammographie [7].

Il existe 2 types principaux d'adénocarcinome, soit :

- 1)** le carcinome canalaire prend naissance dans le revêtement des canaux galactophores du sein.
- 2)** le carcinome lobulaire prend naissance dans les glandes productrices de lait (lobules) du sein.

b. Cancer du sein invasif

Cette catégorie de cancer est aussi appelée cancer infiltrant. Ce terme signifie que les cellules cancéreuses traversent la paroi du canal de lactation. Le cancer est souvent palpable et a la possibilité de se propager aux ganglions de l'aisselle. Dans la plupart des cas, le cancer invasif est encore bien guérissable; cependant, si on néglige de le traiter il pourrait se propager dans diverses parties du corps pour générer des métastases.

Il existe plusieurs types de cancer du sein invasif, à savoir le carcinome lobulaire infiltrant et le carcinome canalaire infiltrant.

Une forme distincte de cancer du sein invasif est le cancer du sein inflammatoire qui se caractérise principalement par un sein qui peut devenir rouge, enflé et chaud. La peau du sein peut aussi prendre l'aspect d'une orange. Il passe souvent pour une banale mastite ou une infection.

Une autre forme particulière de cancer du sein, appelée la maladie de Paget se caractérise par une petite plaie au niveau du mamelon que ne guérit pas. Encore une fois, si une petite ulcération survient au niveau du mamelon, ce n'est pas nécessairement un cancer.

1.3.2 Signes et symptômes du cancer du sein

Les symptômes listés ci-dessous, ne signifient pas nécessairement qu'il s'agit d'un cancer du sein. Cependant, il est important de le prendre en charge le plus tôt possible.

1. Une boule ou une masse dans un sein est, le signe d'un cancer du sein le plus couramment observé.
2. Une ou plusieurs masse(s) dures à l'aisselle signifient, parfois qu'un cancer du sein s'est propagé aux ganglions axillaires. Les ganglions restent toutefois indolores.
3. Des modifications de la peau du sein et du mamelon, pourraient générer une tumeur bénigne.
4. Un changement de la taille et de la forme du sein, n'indiquent pas forcément une tumeur maligne.

Il est donc recommandé, de demander un avis médical dès que l'on repère une anomalie. Il ne faut pas attendre et ne négliger aucun signe inhabituel car Le cancer du sein, ne donne pas de signe au début de la maladie et quand la tumeur est suffisamment grosse, elle est palpable au niveau du sein ou de l'aisselle.

1.4 La mammographie [6] [7] [8]

Une mammographie est une radiographie des seins, réalisée par un appareil spécifique appelé mammographe (figure 1.2 (a)). Elle permet d'obtenir des images de l'intérieur du sein, à l'aide de rayons X et de détecter ainsi d'éventuelles anomalies. Autrement dit, c'est un examen qui permet d'orienter le médecin dans son diagnostic (pour repérer des nodules, des opacités, des microcalcifications, qui peuvent correspondre à des lésions suspectes).

Une mammographie peut être réalisée, soit dans le cadre d'un dépistage du cancer du sein (mammographie de dépistage) soit en présence de symptômes (mammographie de diagnostic).

Il existe actuellement deux types de mammographies: les mammographes conventionnels qui, permettent l'obtention des films radiologiques habituels et les mammographes numériques, qui utilisent un détecteur numérique permettant l'amélioration de la qualité des clichés mammographiques.

a. Mammographie Conventionnelle

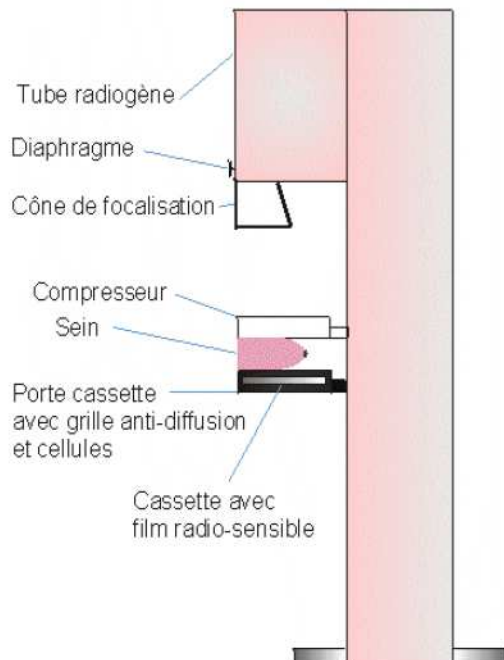
Le sein est interposé entre un tube émetteur de rayons X et un couple « écran– film photosensible ». Les rayons X traversent le sein où ils sont absorbés, selon la densité des tissus rencontrés et se reflètent sur l'écran, ce qui entraîne la production d'une lumière qui impressionne le film argentique. L'image obtenue traduit la densité des tissus traversés (le blanc représente l'absorption importante des tissus denses glandulaires et fibreux et, le noir exprime la faible absorption du grasseux).

b. Mammographie numérique

Il existe deux techniques pour obtenir une mammographie numérique. La première, la numérisation secondaire, consiste à transformer l'image du film argentique en image numérique par l'intermédiaire d'un scanner ou d'une caméra CCD qui affecte une valeur numérique déterminée à chaque niveau de gris. La seconde, la numérisation directe, utilise des mammographes numériques plein champ où la plaque et son film photographique; sont remplacés par des détecteurs qui captent le rayonnement X en mode numérique et

construisent l'image, à partir d'un nombre fini de points dont, chacun a une valeur appropriée.

Mammographe



(a)



(b)

Figure 1.2 : la mammographie numérique et le cliché associé [8] [9]

(a) Mammographe [8]

(b) Cliché associé [9]

1.4.1 Le déroulement d'une mammographie

Le radiologue ou le manipulateur de radiologie, positionne la patiente correctement sur l'appareil. La mammographie est réalisée en position debout, torse nu.

Pour obtenir une analyse de bonne qualité, le sein est comprimé entre deux plaques pour bien l'étaler et éviter les superpositions des images (une compression insuffisante peut empêcher de détecter une anomalie). La sensation de compression du sein, est désagréable pour certaines femmes, mais cela ne dure que quelques secondes. Il existe une sécurité qui limite la pression maximale du sein. Afin de mieux examiner le sein, plusieurs clichés ou incidences (habituellement 2 par seins) sont réalisés sous différents angles. La mammographie de base dure en moyenne de 10 à 15 minutes.

Dans certaines conditions, notamment quand les seins sont trop denses (cas d'une adolescente ou d'une jeune femme), la mammographie est moins performante. On a alors recours, à une échographie mammaire.

1.5 Anomalies mammaires [8][9]

La très grande majorité des anomalies détectées, sur la mammographie de dépistage ne correspondront donc finalement pas à un cancer. Elles doivent néanmoins toutes être prises au sérieux et seules les investigations complémentaires, permettront de déterminer s'il s'agit ou non d'une lésion maligne.

Les anomalies mammographiques détectées le plus fréquemment sont les masses et, les calcifications.

- ✓ Les masses (dont on précise la forme, la taille, les caractéristiques des bords) et les distorsions architecturales;
- ✓ Les calcifications (dont on peut préciser la taille, le nombre, la morphologie, la distribution et l'hétérogénéité).

1.5.1 Les masses

Les anomalies mammographiques détectées le plus fréquemment sont les masses (figure 1.3). La masse se définit comme une lésion occupant un espace, vue sur deux incidences différentes.

Les masses se caractérisent par leur forme ronde, ovale, lobulée ou irrégulière, leur contour circonscrit, micro-lobulé, indistinct, masqué ou spiculé et leur densité. L'analyse de ces différentes caractéristiques, permet de déterminer s'il s'agit d'une anomalie bénigne ou suspecte. Si l'anomalie est bénigne, la mammographie de dépistage reste négative. Dans les autres cas, le dépistage est positif et des investigations complémentaires doivent être entreprises.

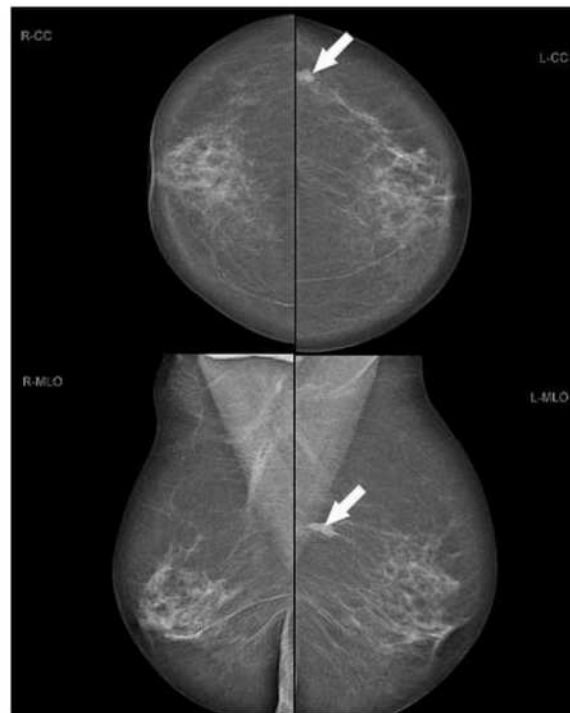


Figure 1.3 : Mammographie de dépistage avec quatre incidences (RCC = Cranio-caudal droit, LCC = Cranio-caudal gauche, RMLO = médio-latéral oblique droit, LMLO = médio-latéral oblique gauche). Une opacité bosselée centimétrique est visible dans le quadrant supéro-externe à gauche (flèche) [9].

a. L'analyse des opacités mammaires

Les masses sont caractérisées par: leur taille, leur forme, leurs contours et leur densité.

- ✓ **La taille.** Elle est variable de quelques millimètres, à plusieurs centimètres. Cependant la taille ne prédit pas le caractère malin, sauf sur des clichés successifs lorsque l'on voit la taille augmenter régulièrement.
- ✓ **La densité.** Ces masses sont en général denses (plus blanches) que le tissu mammaire adjacent. La densité (figure 1.4) peut être appréciée en utilisant la terminologie suivante:
 - tissu radio-transparent, graisseux, adipeux: ce terme est utilisé pour toute lésion contenant de la graisse, tel qu'un kyste huileux, un lipome, une galactocèle, ainsi que des lésions mixtes telles qu'un hamartome ou adéno-lipome.
 - Faible densité: il s'agit d'un tissu hypodense par rapport au tissu conjonctivo-glandulaire mais, sans inclusion de tissu adipeux.

- Densité moyenne: c'est un tissu isodense c'est à dire comparable au tissu conjonctivo-glandulaire.
- Forte densité: c'est un tissu hyperdense par rapport au tissu conjonctivo-glandulaire.

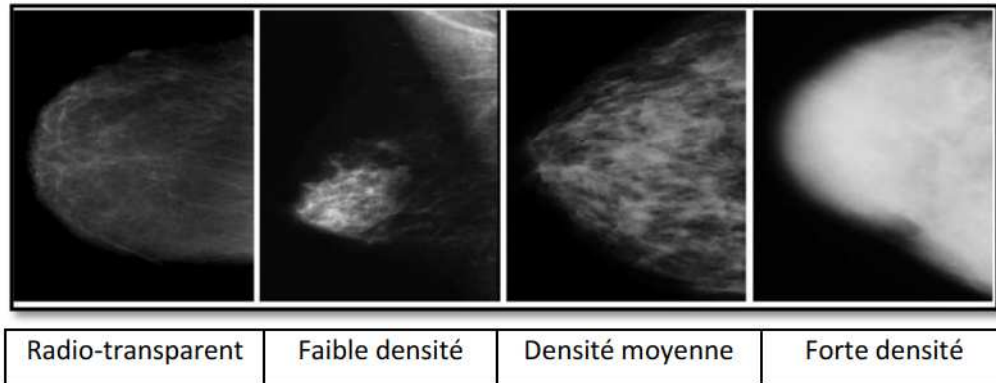


Figure 1.4 : les différentes densités mammaires [8]

- ✓ **La forme** : on peut envisager différentes formes de masse (figure 1.5) :
- *Ronde*: il s'agit d'une masse sphérique, circulaire ou globuleuse (a).
 - *Ovale*: elle présente une forme elliptique (ou en forme d'œuf) (b).
 - *Lobulée*: la forme de la masse présente une légère ondulation (c).
 - *Irrégulière*: cette appellation est réservée aux masses, dont la forme est aléatoire (d).

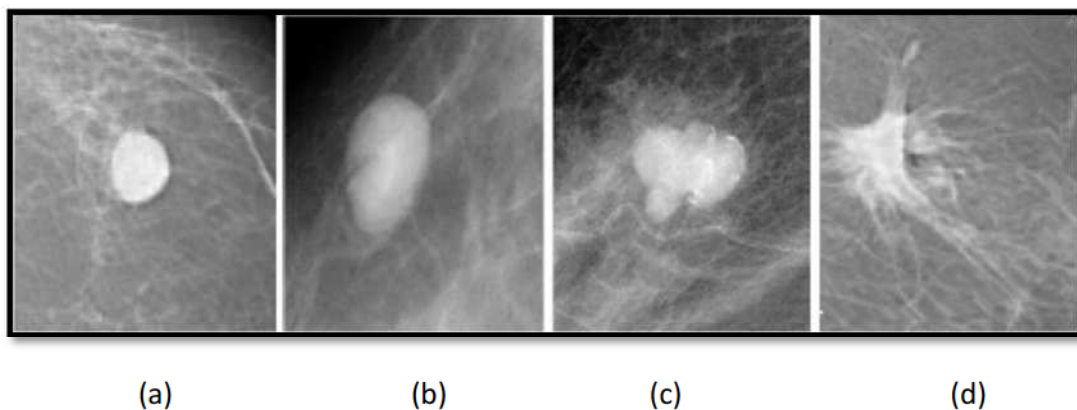


Figure 1.5 : les types de masses [8]

(a) : forme ronde (b) : forme ovale (c) : forme lobulée (d) : forme irrégulière

- ✓ **Les contours** : c'est le meilleur critère d'études des opacités. La figure 1.6 montre cinq types de contours des masses.

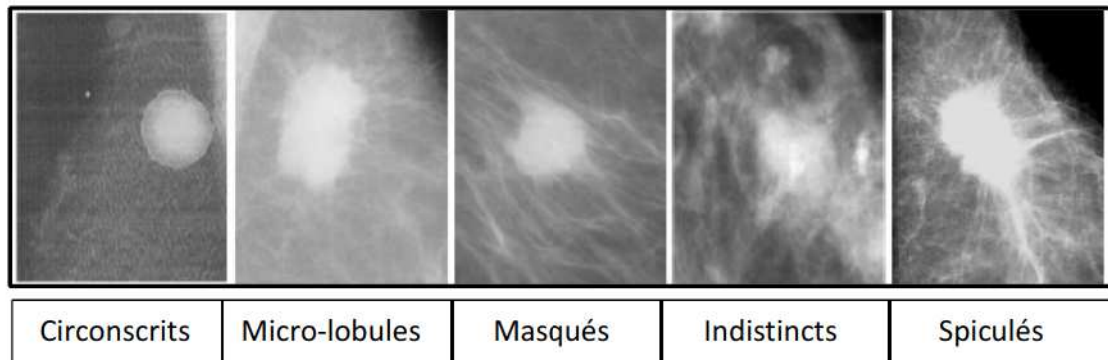


Figure 1.6 : Les contours des masses [8]

- **Circonscrit**: il s'agit d'une transition brusque entre la lésion et le tissu environnant. Le contour est net et bien défini. Pour qu'une masse soit qualifiée de circonscrite, il faut qu'au moins 75% de son contour, soit nettement délimité.
- **Micro-lobulé**: dans ce cas, de courtes dentelures du contour créent de petites ondulations.
- **Masqué**: un contour masqué est un contour qui est caché par le tissu normal adjacent.
- **Indistinct « Mal défini »**: dans ce cas, le contour mal défini, peut correspondre à une infiltration.
- **Spiculé**: la masse est caractérisée par des lignes radiaires, prenant naissance sur le contour de la masse. Ces lignes radiaires sont appelées, les spicules.

1.5.2 Les Calcifications

Les calcifications mammaires sont des dépôts de calcium qui se forment dans le tissu du sein. Ces calcifications ne sont pas liées à la quantité de calcium qu'une personne, absorbe par son alimentation ou les suppléments qu'elle prend.

Les calcifications mammaires sont assez courantes et on les détecte, souvent lors d'une mammographie de dépistage. Elles apparaissent sous forme de points blancs, à la mammographie. Le radiologiste observe la taille, la forme et la disposition des calcifications et note le tout dans les résultats de la mammographie.

La plupart des calcifications mammaires, ne sont pas associées au cancer. Cependant, certaines caractéristiques des calcifications, comme une forme irrégulière ou certains regroupements, peuvent être inquiétantes.

Il existe 2 types de calcifications:

a. Macrocalcifications

Les macrocalcifications sont des dépôts grossiers de calcium, dans le sein. Elles sont plutôt fréquentes, chez les femmes âgées de plus de 50 ans. On les associe souvent à des changements bénins qui, se produisent dans le sein et qui sont causés par: le vieillissement des artères du sein, l'inflammation des tissus du sein (mastite), les masses bénignes.

Les macrocalcifications ont un aspect très distinct, à la mammographie et on les associe habituellement, à des affections bénignes. La plupart des macrocalcifications, ne requièrent pas de biopsie.

b. Microcalcifications

Les microcalcifications sont de minuscules dépôts de calcium, dans le sein. Elles peuvent laisser croire que l'activité est accrue dans certaines cellules du sein. Lorsque ces cellules sont plus actives, elles absorbent plus de calcium du corps.

Les microcalcifications peuvent être isolées ou regroupées en grappes, dans une région du sein. Il est possible qu'elles indiquent la présence d'un petit cancer comme le carcinome canalaire in situ (CCIS).

c. La classification de Le Gal

Cette classification cherche à évaluer le caractère dangereux, ou non des calcifications observées. On distingue 5 types de microcalcifications (figure 1.7).

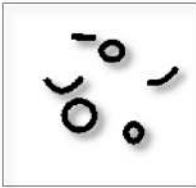
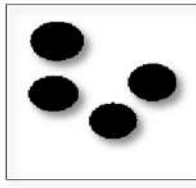
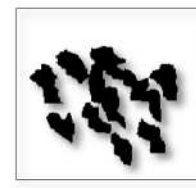
Type I	Type II	Type III	Type IV	Type V
Images rondes planes	Images régulières	Grains de sel	Points irréguliers	Vermiculaires
				
0% de cancers	39% de cancers	39% de cancers	59% de cancers	96% de cancers

Figure 1.7 : Les différents types de microcalcifications [8]

Elles sont associées à des images de masse, ou de modifications architecturales et peuvent se retrouver, dans des pathologies malignes ou bénignes. On analyse les calcifications, selon leur taille, leur forme, leur nombre et leur distribution.

En règle générale, des calcifications de grande taille régulière, rondes ou ovales sont synonymes de tumeur bénigne, les calcifications petites, irrégulières, de formes variées sont souvent associées avec une pathologie maligne.

✓ Analyse de la taille

En règle générale, les microcalcifications, sont associées à un processus malin et les macrocalcifications, à un processus bénin. Les calcifications sont d'abord, peu visibles puis augmentent de taille.

On peut observer des calcifications dont la taille est d'environ 0.2 à 0.3 mm. Les calcifications de moins de 0.5 mm, sont souvent associées à des cancers; celles de plus de 2 mm sont assez typiques de lésions bénignes.

✓ Analyse du nombre

Les petites calcifications nombreuses, sont suspectes. La plupart des radiologues considèrent qu'au-dessus de 4 à 6 (micro-) calcifications, il convient d'être prudent dans l'interprétation.

✓ Analyse de la morphologie

L'analyse de la morphologie, apporte beaucoup de renseignements et permet le plus souvent de séparer les microcalcifications bénignes des malignes. Les calcifications arrondies ou ovales, uniformes dans leur taille et leur forme, sont probablement bénignes. A l'inverse, celles qui sont irrégulières dans leur taille, ressemblant par leur hétérogénéité à des 'débris de verre', sont probablement malignes.

1.5.3 Les distorsions architecturales

Le troisième type d'anomalie détecté sur la mammographie de dépistage est la distorsion architecturale. C'est une anomalie subtile qui se caractérise par l'interruption de l'architecture normale du sein sans qu'une modification de densité ne soit visible (figure 1.9). Une cicatrice chirurgicale provoque typiquement une distorsion architecturale. En l'absence d'antécédent chirurgical, des investigations complémentaires doivent impérativement être entreprises.

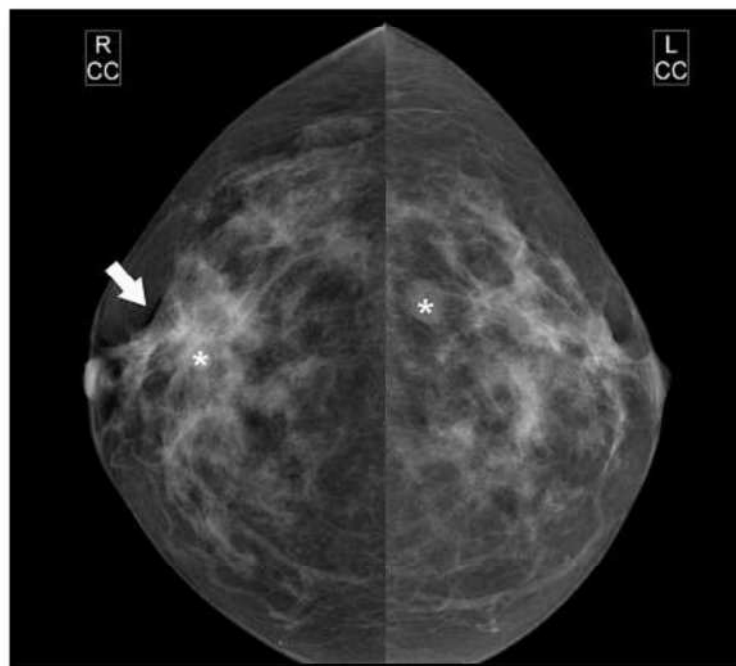


Figure 1.8 : Incidences Cranio-caudales des deux seins. Des opacités infra-centimétriques rondes d'allure bénigne sont visibles des deux côtés (*). Dans la région rétro-aréolaire droite, une distorsion architecturale est reconnaissable (flèche) [8].

1.6 Les résultats d'une mammographie: la classification ACR [10]

La mammographie est un examen qui permet d'orienter le médecin dans son diagnostic. On utilise le système BIRADS de l'*American College of Radiology* (ACR) pour classer les images mammographiques, en 6 catégories.

- **ACR 0** : C'est une classification « d'attente », qui s'utilise en situation de dépistage ou dans l'attente d'un second avis, avant que le second avis soit obtenu ou que le bilan d'imagerie soit complété et qu'ils permettent une classification définitive. Cette classification nécessite des investigations complémentaires en comparaison avec les documents antérieurs, les incidences complémentaires, les clichés centrés comprimés, ainsi que l'échographie.
- **ACR 1** : mammographie normale.
- **ACR 2** : Il existe des anomalies bénignes ne nécessitant ni surveillance ni examen complémentaire. Cette classification est obtenue dans le cas des lésions suivantes :
 - Opacité ronde avec macrocalcifications (adénofibrome ou kyste).
 - Ganglion intramammaire.
 - Opacité(s) ronde(s) correspondant à un/des kyste(s) typique(s) en échographie.
 - Image(s) de densité grasseuse ou mixte (lipome, hamartome, galactocèle, kyste huileux).
 - Cicatrice(s) connue(s) et calcification(s) sur matériel de suture.
 - Macrocalcifications sans opacité (adénofibrome, kyste, adiponécrose, ectasie canalaire sécrétante, calcifications vasculaires, etc).
 - Microcalcifications annulaires ou arciformes, semi-lunaires, sédimentées, rhomboédriques.
 - Calcifications cutanées et calcifications punctiformes régulières diffuses.
- **ACR 3** : **Il existe une anomalie probablement bénigne pour laquelle une surveillance à court terme est conseillée.** Cette classification est réalisée pour les cas suivants :
 - Microcalcifications rondes ou punctiformes régulières ou pulvérulentes, peu nombreuses, en petit amas rond isolé.
 - Petit(s) amas rond(s) ou ovale(s) de calcifications amorphes, peu nombreuses, évoquant un début de calcification d'adénofibrome.

- Opacité(s) bien circonscrite(s), ronde(s), ovale(s) ou discrètement polycyclique(s) sans microlobulation, non calcifiée(s), non liquidiennes en échographie.
- Asymétrie focale de densité à limites concaves et/ou mélangée à de la graisse.
- **ACR 4: Il existe une anomalie indéterminée ou suspecte qui indique une vérification histologique.** Ceci se fait dans les cas qui suivent.
 - Microcalcifications punctiformes régulières nombreuses et/ou groupées en amas aux contours ni ronds, ni ovales.
 - Microcalcifications pulvérulentes groupées et nombreuses.
 - Microcalcifications irrégulières, polymorphes ou granulaires, peu nombreuses.
 - Image(s) spiculée(s) sans centre dense.
 - Opacité(s) non liquidienne(s) ronde(s) ou ovale(s) aux contours lobulés, ou masqués, ou ayant augmenté de volume.
 - Distorsion architecturale en dehors d'une cicatrice connue et stable.
- **ACR 5 : Il existe une anomalie évocatrice d'un cancer.** Ceci est observé pour les cas suivants :
 - Microcalcifications vermiculaires, arborescentes ou microcalcifications irrégulières, polymorphes ou granulaires, nombreuses et groupées.
 - Groupement de microcalcifications quelle que soit leur morphologie.
 - Microcalcifications associées à une anomalie architecturale ou à une opacité.
 - Microcalcifications groupées ayant augmenté en nombre ou microcalcifications dont la morphologie et la distribution sont devenues plus suspectes.
 - Opacité mal circonscrite aux contours flous et irréguliers.
 - Opacité spiculée à centre dense.

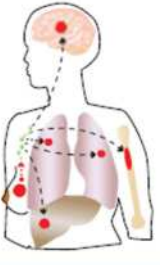
				
Stade 0	Stade I	Stade IIA/IIB	Stade IIIA/IIIB	Stade IV
Cancer du sein non invasif ou in situ (aussi appelé précancéreux)	Cancer invasif localisé à l'intérieur du sein seulement	Cancer de plus grande taille ou qui a atteint des ganglions	Cancer du sein localement plus avancé	Cancer du sein avancé ou métastatique
Les tumeurs sont uniquement à l'intérieur des canaux	Les tumeurs mesurent moins de 2 cm	Les tumeurs mesurent entre 2 et 5 cm ou les cellules cancéreuses ont envahi de 1 à 3 ganglions	Les tumeurs mesurent plus de 5 cm ou les cellules cancéreuses ont atteint 4 ganglions et plus	Présence de métastases (foie, os, cerveau, poumons ou autres endroits)

Figure 1.9 : classification des images mammographiques [10]

1.7 Conclusion

Ce chapitre a été consacré au contexte mammographique; où nous avons présenté quelques notions sur l'anatomie du sein, les différentes anomalies mammaires ainsi que, leurs classifications. Nous connaissons bien que la pathologie du sein est très complexe et ne pourrait être détaillée totalement dans ce chapitre.

La mammographie, reste le principal outil de dépistage et de diagnostic du cancer du sein. Mais il est possible que dans le futur (quelques dizaines d'années ?), la mammographie soit supplantée par d'autres techniques plus sensibles et, non irradiantes.

Les glandes mammaires peuvent être le siège, de tumeurs bénignes ou malignes, dont la lecture clinique demande, une précision du diagnostic. Cette dénomination regroupe des entités multiples, de pronostics différents, dont la connaissance, permet d'appréhender au mieux le diagnostic, le bilan d'extension préopératoire et la surveillance post-thérapeutique. Le plus souvent, des explorations complémentaires à la mammographie, sont indispensables

pour affiner l'examen. Le traitement d'image, devient un outil crucial, pour aider les experts dans la détection de ces anomalies, de façon plus précise.

Le chapitre suivant, est consacré à une introduction aux techniques de prétraitement et de segmentation en imagerie, dont l'étude, nous aide à opter pour une méthodologie appropriée à la détection de pathologies mammographiques dans le cadre de notre projet.

2.1 Introduction

L'imagerie mammographique regroupe les moyens d'acquisition et de restitution d'images, du sein à partir l'absorption de rayons X établie par la mammographie. Elle a révolutionné la médecine en permettant, à des fins diagnostiques, de visualiser de manière non invasive et dynamique; l'anatomie, la physiologie et le métabolisme mammaire. Cette visualisation indirecte nécessite l'emploi de techniques de traitement d'images, permettant de filtrer, de détecter, de quantifier et de, caractériser les données numériques acquises.

Le traitement d'images, par sa diversité même, est parfois très déroutant pour le chercheur débutant. Les informations que l'on souhaite extraire d'une image, sont si variées qu'il est difficile, lorsque l'on doit traiter une nouvelle application, de savoir à quel modèle se raccrocher et quelle méthodologie suivre [11]. Pour cela l'imagerie fait appel à des outils, des algorithmes, qui permettent d'agir sur l'image numérisée. L'un des processus fondamentaux dans la chaîne de traitement d'image est la segmentation, une étape incontournable dans tout processus d'analyse d'image.

Dans ce qui suit, nous allons étudier les techniques de prétraitement adoptées dans le cadre de notre initiation à la recherche, pour le rehaussement du contraste de certaines structures pathologiques, dans l'image mammographique, afin de faciliter la tâche de la segmentation. Nous parcourons aussi, les approches de la segmentation, qui se basent, sur les méthodes par approche région et les modèles déformables.

2.2 Système de détection automatique

De manière globale un système de détection automatique de cancers en mammographie se compose de deux branches: une dédiée à la détection des microcalcifications et l'autre à la détection des opacités. Chacun de ces modules peut se décomposer comme une étape de marquage suivie d'une prise de décision.

Le marquage peut selon les cas être composé d'une détection rapide suivie d'une segmentation. La prise de décision se compose quant à elle d'une étape d'extraction de caractéristiques suivie d'une étape de classification. Certains de ces éléments peuvent apparaître de manière plus ou moins implicite.

La figure 2.1 illustre la décomposition de haut niveau du processus de détection automatique des lésions mammographiques. Dans certains cas, un prétraitement des données permettant de mettre en évidence les signes recherchés, doit être réalisé, pour le rehaussement du contraste [12].

Les étapes de détection automatique de lésions mammographiques se basent sur le système de la figure 2.1.

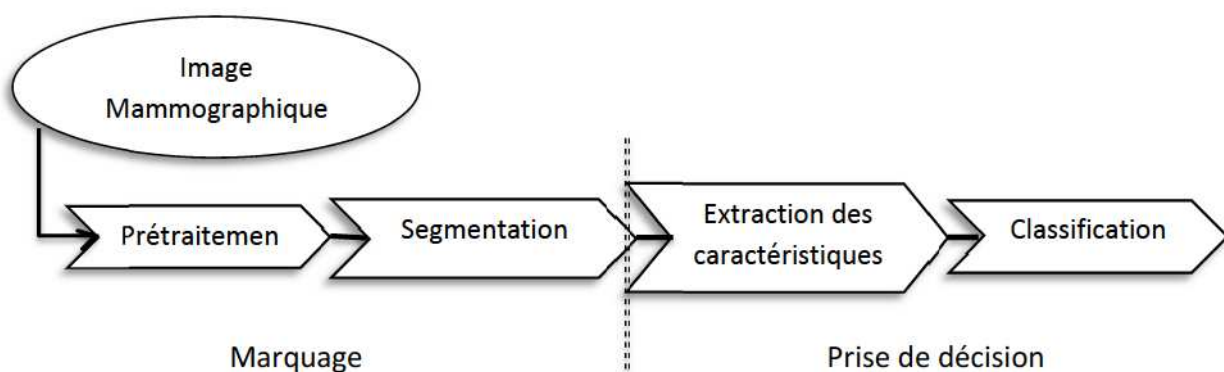


Figure 2.1 : Schéma général d'une chaîne de détection de microcalcifications ou d'opacités en mammographie [12]

- L'étape de prétraitement, sert à améliorer la qualité de l'image avant toute manipulation.
- L'étape de segmentation permet de détecter les lésions mammographiques.

- L'étape d'extraction des caractéristiques a pour but de caractériser les lésions à travers des attributs spécifiques à la forme et, la texture dans notre cas.
- Une dernière étape est celle de la classification et de la prise de décision.

2.3 Prétraitement

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 pour faciliter leur détection.

Comme on connaît bien que pour les images mammographiques, la détection de certaines lésions est un peu difficile du à la variabilité des tissus mammaires; ce qui nécessite un prétraitement spécifique pour le rehaussement du contraste [13].

Il existe une approche couramment utilisée en traitement d'image consiste à travailler sur l'histogramme dans le but de définir une fonction de transfert, sur les niveaux de gris permettant de mettre en évidence les détails présents dans l'image.

Une approche utilisée pour l'amélioration du contraste, est la modification globale de l'histogramme. Cette méthode consiste à réaffecter les valeurs d'intensités des pixels afin de rendre la nouvelle répartition de ces derniers plus uniforme ; ceci peut être réalisé par l'égalisation d'histogramme ou par l'étirement d'histogramme.

L'idée d'un prétraitement, bien qu'intuitive, peut poser quelques problèmes. En effet, dans l'idéal on voudrait mettre en évidence seulement les zones potentiellement suspectes, dans le but de faciliter leur détection ultérieurement. Or pour pouvoir accomplir cette tâche, il faudrait savoir quelles zones de l'image on doit améliorer, c'est-à-dire connaître les zones suspectes, ce qui est difficile sans l'aide de l'expert. D'autre part, l'étape de prétraitement peut aussi modifier certaines structures et les rendre faussement suspectes, ce qui peut être problématique pour l'étape de détection. De manière similaire, un prétraitement peut modifier substantiellement les propriétés de l'image, rendant la modélisation de l'étape de détection délicate [13]. Pour ce fait, il est nécessaire de mettre en valeurs des approches locales, permettant de préparer le cliché mammographique, pour une bonne segmentation. Dans notre cas, on réalise le rehaussement du contraste des masses et des calcifications mammographiques, en se basant sur les opérateurs de la morphologie mathématique.

2.4 Introduction à la morphologie mathématique [15][16]

La morphologie mathématique est une science de la forme et de la structure. Le principe de base de cette méthode est de comparer un objet qu'on analyse à un autre objet de référence, connu par sa taille et sa forme, appelé élément structurant.

On appelle morphologie mathématique *ensembliste*, quand le traitement se porte sur des images binaires et, quand l'utilisation se fait sur des images en niveaux de gris, on fait référence à la morphologie mathématique *fonctionnelle*.

Les opérations de base, sont l'érosion et la dilatation. Celles-ci, sont à la base des autres opérateurs plus complexes tels que l'ouverture, la fermeture et, le chapeau haut de forme.

2.4.1 Opérateurs de base de la morphologie mathématique

a. Élément structurant

Un élément structurant B est un ensemble qui présente, une forme géométrique (carrée, circulaire, linéique,...) et une taille spécifique.

b. Erosion

Soit X l'ensemble à éroder et, B un élément structurant de géométrie simple tel qu'un cercle. L'érosion consiste, à faire déplacer B de telle sorte que son centre occupe toutes les positions X de l'espace. Pour chaque position, on pose la question: est-ce que B est complètement inclus dans X ?

L'ensemble des positions x correspondant à une réponse positive, forme le nouvel ensemble Y, appelé érodé de X par B, noté $X \ominus B$.

Cet ensemble satisfait l'équation suivante :

$$X \ominus B = \{x \text{ dans } X : B \subset X\} \quad 2.1$$



a) Image mammographique initiale b) Image mammographique érodée

Figure 2.2: érosion d'une image mammographique

Le traitement obtenu par l'érosion indique les points suivants:

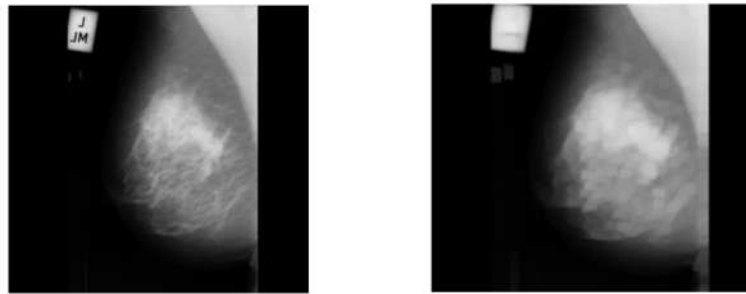
- Les objets de taille inférieure à celle de l'élément structurant vont disparaître.
- Les autres seront « amputés » d'une partie correspondant à la taille de l'élément structurant.
- S'il existe des trous dans les objets, c'est-à-dire des morceaux de fond à l'intérieur des objets ils seront accentués.
- Les objets reliés entre eux par un élément de taille inférieur à celle de l'élément structurant vont être séparés.

c. Dilatation

Cette opération se définit de manière analogue à l'érosion. En prenant le même élément structurant B , on pose pour chaque point x la question « Bx touche-t-il l'ensemble X ? » ; c'est-à-dire y a-t-il une intersection non vide entre Bx et X ?

L'ensemble des points de l'image correspondant aux réponses positives forme le nouvel ensemble Y des dilatés de X , noté comme suit :

$$y = x \oplus B = \{ \text{dans } 1: B \cap X \neq \emptyset \} \quad 2.2$$



a) Image mammographique initiale b) Image mammographique dilatée

Figure 2.3: dilatation d'une image mammographique

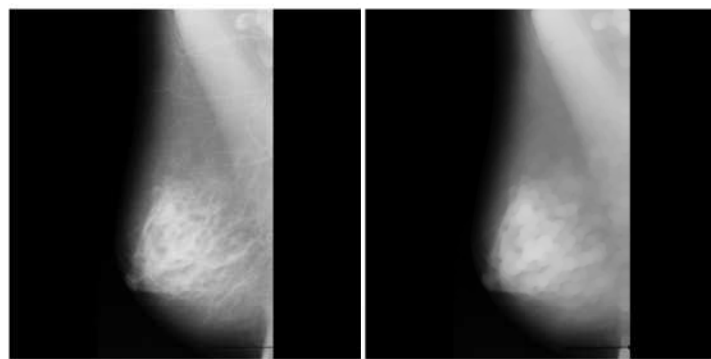
2.4.2 Opérateurs morphologiques transformés

a. Ouverture

Puisque l'érosion et la dilatation sont des transformations itératives, il est possible d'effectuer sur un ensemble X une érosion, puis de dilater l'ensemble érodé par un même élément structurant B . Le résultat est nommé « l'ouvert » de X de B , noté comme suit :

$$OB(X) = (X \ominus B) \oplus B \quad 2.3$$

Après une ouverture, on ne retrouve pas l'ensemble de départ ; l'ensemble ouvert est plus régulier et moins riche en détails que l'ensemble X initial. La transformation par ouverture adoucit les contours, coupe les isthmes étroits, supprime les petites îles et les caps étroits.



a) Image mammographique initiale b) Image mammographique avec ouverture

Figure 2.4 : ouverture d'une image mammographique par un élément structurant circulaire

L'opération d'ouverture de l'image selon la figure ci-dessus, permet :

- De lisser les formes.
- D'éliminer les composantes connexes plus petites que B.
- De conserver souvent la taille et la forme, de ne pas conserver nécessairement la topologie.

b. Fermeture

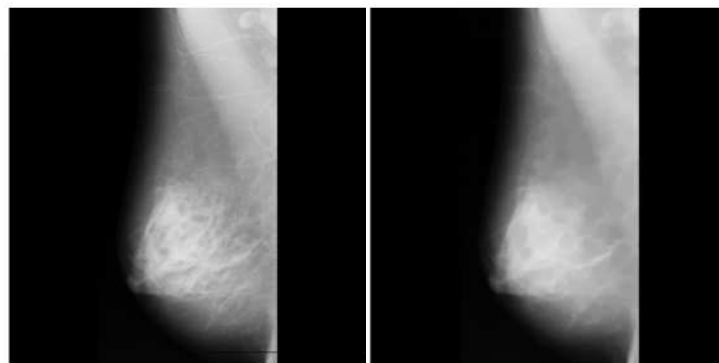
Cette étape représente l'opération inverse de l'ouverture, c'est-à-dire que l'on applique tout d'abord une dilatation puis une érosion (toujours en gardant le même élément structurant).

Le résultat est nommé « le fermé » de X par B, noté comme suit :

$$F^B(X) = (X \oplus B) \ominus B \quad 2.4$$

Un ensemble fermé, est également moins riche en détails que l'ensemble initial (figure 2.5).

La transformation par fermeture, bouche les canaux étroits, supprime les petits lacs et les golfs étroits.



a) Image mammographique initiale b) Image mammographique avec fermeture

Figure 2.5 : fermeture d'une image mammographique par un élément structurant circulaire

L'opération de fermeture suivant le résultat de la figure ci-dessus permet de:

- Boucher les trous plus petits que B.
- Conserver souvent la taille et la forme.
- Ne pas conserver nécessairement la topologie.
- Souder les formes proches en particulier.

c. *Chapeau haut de forme et chapeau bas de forme*

Soient $x = (i, j)$ un point d'une image, f la fonction qui à x associe son niveau de gris et B , le voisinage de x défini par l'élément structurant B centré en x [16].

Les propriétés de la fermeture et l'ouverture en teintes de gris, permettent de définir des filtres pouvant extraire les pics de la fonction $f(x)$ correspondant, aux petites zones claires de l'image ou les vallées, correspondant aux petites zones sombres. Ces filtres, sont appelés **Chapeaux Haut de Forme** et, consistent à effectuer une simple différence entre l'image d'origine et l'ouverture (ou la fermeture). L'appellation Chapeau se justifie par le fait que ces zones, sont représentées sur le graphe de la fonction $f(x)$ par des pics (resp. des vallées étroites).

En effectuant la différence, entre l'image initiale et son ouverture par un élément structurant B de taille n , (Bn), on extrait les pics dont l'épaisseur est inférieure à n .

On appelle cette transformation **Chapeau Haut de Forme Blanc (White Top Hat)**. Cette transformation est établie par l'équation 2.5.

$$\boxed{WTH(x) = f(x) - f_{Bn}(x)} \quad [11] \quad 2.5$$

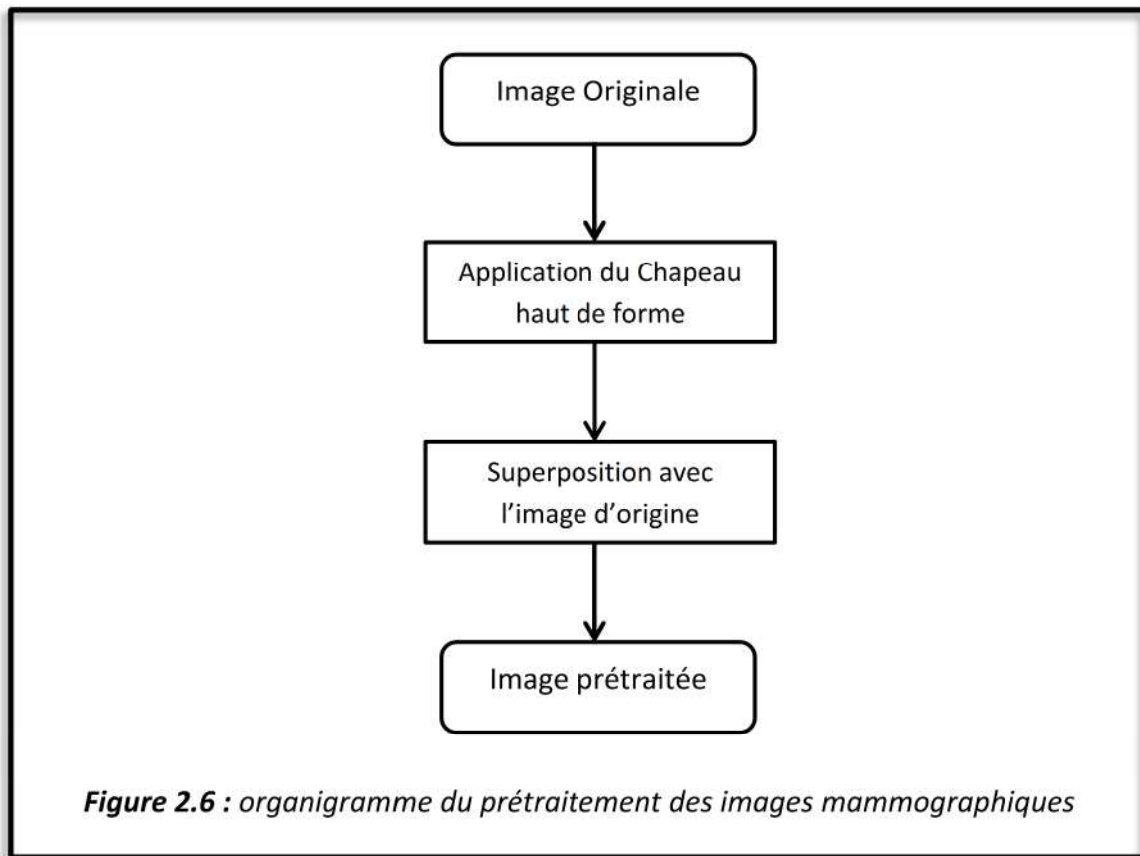
Pour extraire les vallées d'épaisseur inférieures à n , on effectue la différence entre la fermeture de l'image par un élément structurant B de taille n et l'image initiale. Cette transformation est appelée **Chapeau Haut de Forme Noir (Black Top Hat)**. Cette transformation est établie par l'équation 2.6.

$$\boxed{BTH(x) = f_{Bn}(x) - f(x)} \quad 2.6$$

Les chapeaux haut de forme blanc et noir, sont des filtres dits morphologiques et sont adaptés pour l'extraction d'objets linéaires fins, d'une certaine épaisseur, qu'il soient sombres ou clairs localisant; ainsi des structures dans l'image qui sont plus petites que l'élément structurant.

Nous adoptons, la transformation du chapeau haut de forme, pour rehausser le contraste dans les images mammographiques, en utilisant un élément structurant de la forme d'un

disque. L'organigramme de la figure 2.6, illustre les étapes de prétraitement des images mammographiques.



2.5 Segmentation

2.5.1 Introduction à la segmentation

La segmentation des images constitue le cœur de tout système de vision. C'est une étape importante dans le processus d'analyse des images. La segmentation est un traitement de bas niveau, qui consiste à créer une partition d'une image **A** en sous-ensembles **R_i**, appelés régions [11], tels que:

$$\left. \begin{array}{l} \forall i R_i \neq \emptyset \\ \forall i, j; i \neq j R_i \cap R_j = \emptyset \\ A = \bigcup_i R_i \end{array} \right\} \quad 2.7$$

Une région est un ensemble connexe de pixels, ayant des propriétés communes (**intensité, texture, ...**) qui les différencient des pixels des régions voisines. Lorsque le nombre de régions est limité à deux, la segmentation prend le nom de **binarisation** et, se résume alors à séparer un objet, du fond. Dans certains cas, le nombre de régions est connu d'avance ainsi que, certaines caractéristiques de ces zones. Pour ces situations, la segmentation prend le nom de classification et, consiste à associer, chaque pixel de l'image à une de ces régions.

Il n'y a pas de méthode unique de segmentation. Selon **Cocquerez [11]**, le choix d'une technique est lié à la nature de l'image, aux opérations situées en aval de la segmentation, aux primitives à extraire, aux contraintes d'exploitation. Du fait de cette diversité, il est difficile de définir une méthode universelle de segmentation qui convienne à tous les cas. Cependant, en s'inspirant des systèmes de la vision humaine, la segmentation fait référence, aux notions de différence et de similarité. Il en découle deux approches courantes: l'approche « **frontière** » et, l'approche « **région** ». L'approche région, s'attache à faire apparaître des régions homogènes selon un critère d'homogénéité (niveaux de gris, couleur ou texture). C'est une approche, qui fait référence à des regroupements de régions, ayant des caractéristiques similaires; alors que, l'approche frontière tente de trouver des contours ou frontières de régions, présentant une variation rapide du même critère, ou, une discontinuité entre les propriétés de deux ensembles, de points connexes. Ces deux approches, sont duales en ce sens, qu'une région définit une ligne par son contour et, qu'un contour fermé, définit une région. Elles amènent, cependant à des algorithmes différents et, ne fournissent pas, en général, les mêmes résultats. Certaines méthodes, récentes tentent de faire coopérer Certaines méthodes, récentes tentent de faire coopérer plusieurs méthodes de segmentation, afin de combiner les avantages de chaque méthode ou d'exploiter leurs complémentarités.

Un algorithme de segmentation [11] s'appuie donc sur :

1. la recherche de discontinuités afin de mettre en évidence les contours;
2. la recherche d'homogénéité locale pour définir les régions;
3. ou encore, sur la coopération des deux principes (**frontière-région**).

Bien qu'il existe une multitude d'algorithmes de segmentation selon le domaine et les contraintes étudiées, il en découle quatre grandes approches qui peuvent se distinguer par:

- l'approche région,
- l'approche contour,
- segmentation fondée sur la classification ou le seuillage des pixels en fonction de leur intensité,
- l'approche coopérative.

2.5.1 Approche région [11]

Les méthodes de cette approche, s'intéressent au contenu de la région (des informations locales ou globales). Elles regroupent les pixels vérifiant des propriétés communes (niveau de gris, écart-type,...). Il existe plusieurs méthodes telles que, la segmentation par croissance de régions, par division de régions et, par fusion de régions que nous présentons ci-dessous.

a. Segmentation par division

La division consiste à partitionner l'image, en régions homogènes selon un critère donné. 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 de classes homogènes.

b. Segmentation par fusion de régions

Les techniques de fusion (*region merging*) sont des méthodes ascendantes où tous les pixels sont considérés. Pour chaque voisinage de pixel, un prédicat P est testé. S'il est vérifié les pixels correspondants sont regroupés, dans une région. Après le parcours de toute l'image, les groupes de voisinages se voient appliquer le même test et sont regroupés, si P est vérifié. Le processus est itéré jusqu'à satisfaction d'un critère d'arrêt.

c. Segmentation par division/fusion

Son principe est de combiner les deux méthodes présentées afin de pallier à leurs inconvénients (division de régions et fusion de régions) de la manière suivante : une première étape de division donne comme résultat, une image divisée en plusieurs régions. Par la suite, une étape de fusion intervient afin de corriger le résultat obtenu par la première étape, en regroupant les régions similaires. Ce procédé est répété jusqu'à l'obtention d'une segmentation intéressante.

d. Segmentation par croissance de régions

L'idée derrière cette méthode est de se fixer un point de départ dans l'image, que l'on nommera « *germe* » (*point d'amorçage*) de la région. Ces germes désignent les points ou régions de départ au sein de l'image à segmenter, selon un critère d'homogénéité pour la région. Par une procédure récursive, chaque germe est comparé à son voisinage spatial immédiat et ; en fonction de ces critères, les régions similaires adjacentes sont fusionnées aux germes de départ, donnant lieu à des régions de plus grandes tailles. De cette manière, on fait croître la région tant que le critère est respecté et on obtient en bout de ligne, une nouvelle région connexe. Le processus est ensuite, réitéré jusqu'à l'épuisement des régions susceptibles d'être fusionnées.

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 pixellique.

➤ Les paramètres de la croissance de région [17][18]

L'utilisation de la segmentation par croissance de régions, nécessite certains paramètres à prendre en compte, tels que le choix des germes initiaux, la mesure d'homogénéité et, le critère d'arrêt.

a- Les germes initiaux

La détermination des germes initiaux, est une phase importante pour que les régions puissent croître facilement, dans toutes les directions. Leur emplacement a une influence, sur le résultat de la segmentation. Le germe initial, doit faire partie de la zone de recherche au risque d'avoir une mauvaise segmentation au pire totalement erronée.

b- Mesure d'homogénéité

La segmentation par croissance de régions, est basée sur l'évaluation des pixels à ajouter. Cette évaluation mesure l'homogénéité locale entre le germe initial et, un pixel candidat ou bien global entre les points constituant la région et, un pixel candidat. En pratique, cette mesure est souvent une distance entre la valeur de l'intensité ou une autre grandeur utilisée (niveau de gris, couleur...). Si cette distance est inférieure à un seuil, ou bien cette distance est la plus petite des distances obtenues avec tout autre pixel candidat à l'agglomération; alors le point est ajouté à la région. Le choix du seuil, a une forte influence sur le résultat de la segmentation. Un mauvais réglage de ce paramètre, fait qu'une région croît, à l'extérieur de la région à segmenter.

c- Processus itératif

Suite au choix des germes initiaux et le critère d'homogénéité, un processus itératif de déformation ajoute progressivement des points situés à la périphérie de la région en train de croître; s'ils respectent la mesure d'homogénéité. Ces nouveaux points, sont regroupés dans l'ensemble que l'on note $C^{[n]}$.

Considérons une croissance de régions ne faisant croître, qu'une seule région. Étant donnée, une région initiale $R^{[0]}$, la région suivante $R^{[n+1]}$ est obtenue à partir de la région courante $R^{[n]}$ et de $C^{[n]}$ par la réunion suivante:

$$R^{[n+1]} = R^{[n]} \cup C^{[n]} \quad 2.8$$

d- Convergence, critère d'arrêt

Comme l'itération porte sur une région croissante et bornée, la convergence est assurée au bout d'un nombre fini d'itérations. Le processus de croissance, peut s'arrêter selon deux scénarios :

- Toutes les régions satisfont la mesure d'homogénéité et, la segmentation comprend n régions.
- Il existe $n-1$ régions qui vérifient la mesure d'homogénéité, la $n^{\text{ième}}$ région comprend les points ne vérifiant pas la mesure d'homogénéité.

➤ **Les avantages et les inconvénients de la segmentation par croissance de région**

Parmi les avantages de cette technique, nous pouvons citer:

- ✓ la simplicité et la rapidité de la méthode;
- ✓ la segmentation d'objet à topologie complexe;
- ✓ la préservation de la forme de chaque région de l'image.

Cependant, il existe plusieurs inconvénients comme:

- ✓ L'influence du choix des germes initiaux et, du critère d'homogénéité sur le résultat de la segmentation.
- ✓ Il peut y avoir des pixels qui ne peuvent pas être classés.

L'application de ces méthodes aux images mammographiques, a montré de bonnes performances vis-à-vis de la détection des masses tissulaires et des calcifications[19] [20]. Un inconvénient majeur de cette approche, réside dans le fait que le résultat de la segmentation dépende, de l'ordre de regroupement des régions élémentaires [20]. Tenir compte de l'homogénéité globale, donne un algorithme sensible à l'ordre de parcours des points et, au bruit.

La figure suivante montre le résultat de cette application sur des images mammographiques selon les travaux de H. Cheng, et al [21].

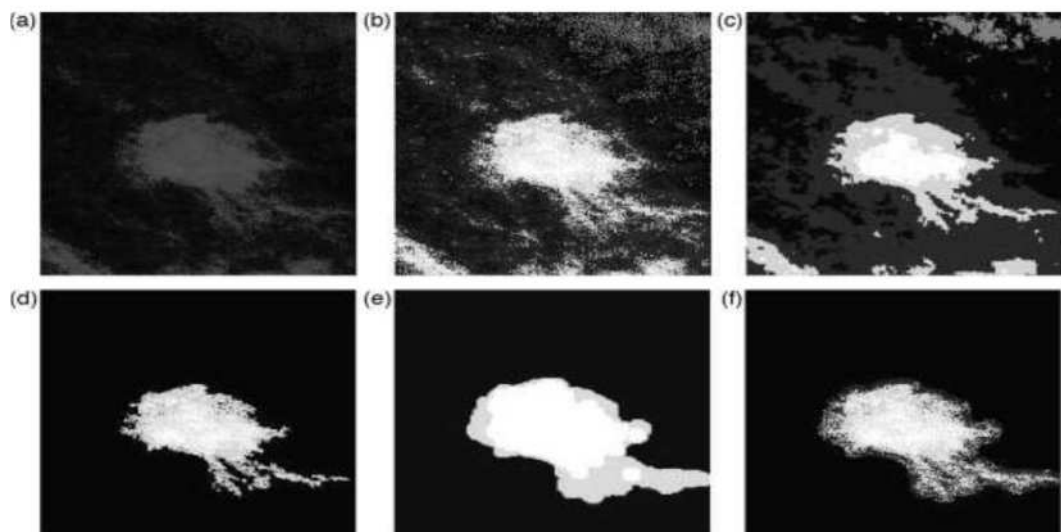


Figure 2.7 : application de l'algorithme de croissance de régions sur les images mammographiques [21]

L'approche que nous avons choisie, consiste à prendre le germe comme étant le centre de la tumeur expertisé par le médecin radiologue, et de faire croître la région en comparant tous les pixels voisins non affecté à la région. La différence entre la valeur de l'intensité $I(x, y)$ et la moyenne de la région M , est utilisée en tant que mesure de similitude.

Le processus s'arrête lorsque la différence d'intensité entre la région et le nouveau pixel, devient plus grande que la valeur choisie du seuil.

$$|I(x, y) - M| \leq t \quad 2.9$$

2.5.2 Approche Contour

L'approche contour [11] consiste à identifier, les changements entre les régions. En général, un élément de contours est un point de l'image appartenant à la frontière de deux ou plusieurs objets, ayant des niveaux de gris différents. En effet, une frontière est définie comme un endroit de l'image où le changement en niveaux de gris, est le plus important. Ces frontières constituent le contour des objets.

Ces méthodes sont basées, sur la recherche de variations locales significatives. Largement étudiée pour les images en niveaux de gris, la recherche de contours repose principalement sur le calcul des dérivées.

Les méthodes de segmentation basées sur l'approche contour, ont pour objectif de trouver les lieux de fortes variations du niveau de gris. Une multitude de méthodes ont été développées. Ces dernières s'appuient sur la détection des discontinuités dans l'image et peuvent être divisées en trois classes :

- Les méthodes dérivatives: les contours sont assimilés au point fort du gradient ou aux passages par zéro de la dérivée seconde.
- Les méthodes analytiques.
- Les modèles déformable (Snakes, levels sets).

Nous nous basons dans notre étude, sur les modèles déformables.

2. 6 Les modèles déformables [22][23]

Les algorithmes de segmentation fondés sur les modèles déformables ont l'avantage, de fournir des contours ou surfaces fermés. Parmi ces méthodes, on trouve les contours actifs et les ensembles de niveaux.

2.6.1 Les contours actifs (*Snakes*)

Les contours actifs sont définis, par une courbe paramétrique pouvant être fermée ou non. Un *snake* consiste à placer aux alentours de la forme à détecter une ligne initiale de contour. Cette ligne va se déformer, progressivement selon l'action de plusieurs forces qui vont la tirer ou la pousser vers la forme.

Ces forces sont représentées par trois énergies associées au snake:

- Une énergie propre, due uniquement à la forme du contour, dite énergie *interne*.
- Une énergie potentielle imposée par l'image dite énergie *externe*. C'est elle qui va attirer la ligne du snake vers les contours réels présents sur l'image.
- Une énergie de *contexte* qui exprime certaines contraintes supplémentaires qui peuvent être imposées par l'utilisateur vu le snake qu'il veut obtenir.

Donc, le contour actif est formé d'une série de points mobiles et répartis sur une courbe en deux dimensions. La courbe est placée dans la zone d'intérêt de l'image ou autour d'un objet (figure 2.8).

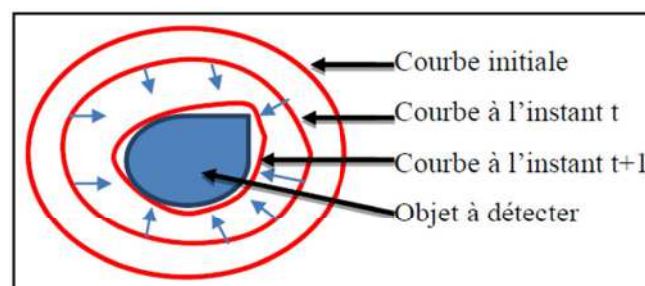


Figure 2.8 : Principe du contour actif [22]

Le principe de cette méthode consiste, à initialiser un contour grossièrement autour de l'objet à détecter et, à le faire déplacer par un processus d'itération afin d'épouser au mieux, les contours de la région recherchée.

Le processus d'itération, est appliqué de façon à conserver certaines caractéristiques; telles que la courbure, la répartition des points, ou d'autres contraintes liées à la disposition des points. Ceci est explicité par un exemple simple donné par la figure 2.9.

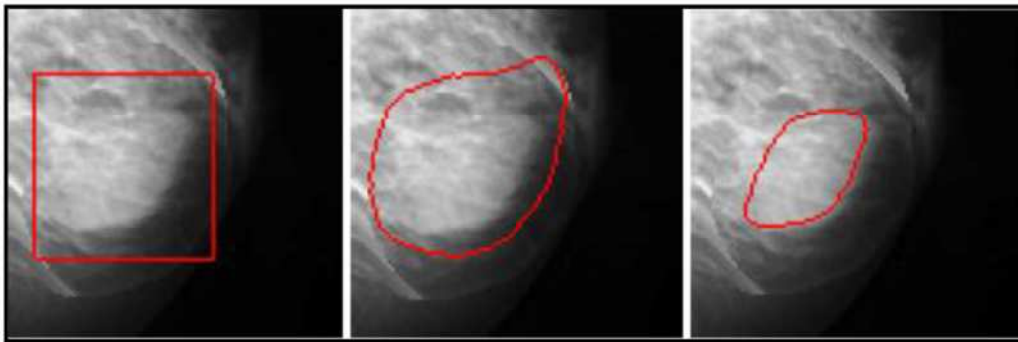


Figure 2.9 : Segmentation (contour actif) d'une masse circonscrite [24].

Cette approche a montré des avantages remarquables:

- En termes de structure (courbe fermée),
- De qualité des contours obtenus,
- De rapidité de convergence de l'algorithme.

H. Cheng et al, ont obtenu des résultats intéressants pour le cas d'image mammographiques, dont le contraste est faible. Néanmoins, ces méthodes présentent des inconvénients :

- À l'initialisation, le contour doit être le plus proche possible de l'objet à détecter au risque de ne pas être attiré par l'objet à cause de la portée réduite du gradient de l'image.
- Les contours actifs sont connus pour leur sensibilité à l'initialisation et leurs temps de convergence longs.
- Les contours actifs basés région possèdent une meilleure capacité de segmentation, mais celle-ci reste fortement dépendante du descripteur de régions choisi pour guider la courbe.

2.6.2 Les ensembles de niveaux (*Level sets*)

La méthode des ensembles de niveaux (figure 2.10) est une méthode de simulation numérique utilisée pour l'évolution des courbes et des surfaces, dans les domaines discrets.

C'est une représentation variationnelle de contours qui évoluent dans le temps, mais différemment des contours actifs [23]. En effet, les ensembles de niveaux peuvent changer de topologie si, les contours l'imposent (un contour simple peut par exemple donner naissance à plusieurs contours fils qui vont évoluer séparément).

Ainsi, un contour simple peut évoluer en deux contours séparés, ou à l'inverse, deux contours séparés se réunissent, en un seul contour. Pour permettre son évolution ultérieure, la fonction ψ dépend du temps :

$$\psi: D \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$(x, t) \rightarrow \psi(x, t)$$

A un instant donné t , la surface Γ est définie comme l'ensemble des points x vérifiant la condition :

$$\Gamma = \{x \mid \psi(x, t) = 0\} \quad 2.10$$

Un choix habituel pour l'initialisation de $\psi(x)$ est la distance euclidienne signée qui sépare x du contour Γ . Le signe dépend de la position du point x à l'intérieur ou à l'extérieur de Γ . Sur la figure 2.10, la région R_{in} est l'ensemble des points pour lesquels ψ est positive :

$$R_{in} = \{x \mid \psi(x, t) \geq 0\} \quad 2.11$$

En utilisant la distance euclidienne signée, la fonction d'ensembles de niveaux peut donc être initialisée comme suit :

$$\psi(x, t) = \begin{cases} 0 & \text{si } x \in \Gamma \\ \inf_{y \in D} \|x - y\| & \text{si } x \in R_{in} \\ -\inf_{y \in D} \|x - y\| & \text{si } x \notin R_{in} \end{cases} \quad 2.12$$

Les variables qui vont être modifiées par la méthode d'évolution, sont les niveaux de la fonction ψ en chaque point de D .

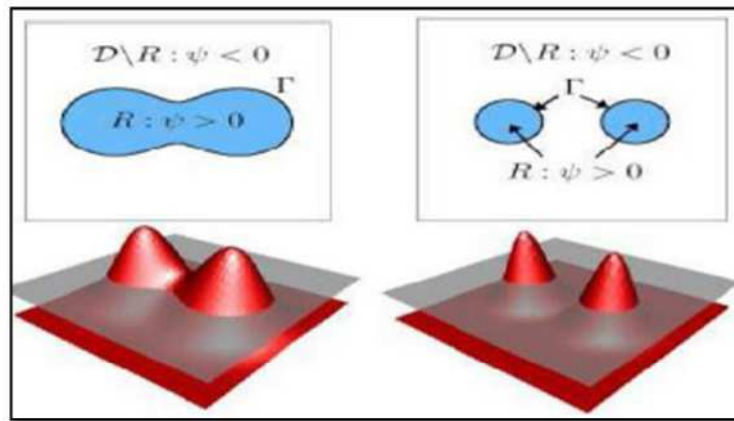


Figure 2.10 : Fonction d'ensembles de niveaux [22].

La méthode des ensembles de niveaux présente de nombreux avantages:

- la courbe peut se diviser ou fusionner au cours de l'évolution.
- les propriétés géométriques de la courbe, telles que la courbure ou le vecteur normal qui sont importants pour les contours actifs, peuvent être simplement déterminés.
- la formulation n'est pas limitée, aux courbes en deux dimensions et peut être facilement étendue, aux dimensions plus élevées.

Les principales limites des méthodes de détection de contour sont les suivantes :

- Les contours extraits selon les méthodes classiques souvent ne correspondent pas nécessairement à la limite des objets.
- Les techniques de détection de contour dépendent de l'information contenue dans le voisinage local de l'image. Il n'y a pas d'information globale.
- Le processus de fermeture des contours peut parfois conduire à des discontinuités et des lacunes dans l'image.
- Il est souvent difficile d'identifier et de classer les contours parasites.

2.6.3 Le modèle Chan et Vese [25] [26]

C'est modèle dynamique basé sur la fonctionnelle de Mumford-Shah. En effet, qualifié de « modèle sans frontières », le modèle de Chan et Vese peut être appliqué sur des images avec des frontières non définies, par le gradient. Ce modèle a été exposé dans sa première version sous la forme biphasé, autrement dit, l'image est segmentée en deux régions. La méthode a été par la suite étendue à plusieurs phases, pour segmenter l'image en plusieurs

régions. Ce modèle est formulé, à l'aide d'une équation aux dérivées partielles dont, la résolution consomme un temps de calcul appréciable.

Soit I , l'image à segmenter en deux régions. Cette image est définie, dans un domaine spatial Ω , telle que $I = \{I(x, y), (x, y) \in \Omega\}$ et soit $\Gamma(t)$, la courbe d'évolution qui définit le contour des objets. La courbe initiale nécessaire, au déroulement de l'algorithme est souvent un cercle de rayon quelconque et centré, en un point quelconque de l'image. La courbe $\Gamma(t)$, divise le domaine Ω en deux régions Ω_0 et Ω_1 qui sont respectivement les régions intérieure et extérieure à $\Gamma(t)$. Cette courbe $\Gamma(t)$ est représentée implicitement par la courbe de niveau zéro d'une fonction auxiliaire $\Phi(x, y, t)$ appelée fonction level set (équation 2.12).

$$\begin{cases} \Phi(x, y, t) > 0 \text{ pour } (x, y) \in \Omega_1 \\ \Phi(x, y, t) = 0 \text{ pour } (x, y) \in \Gamma \\ \Phi(x, y, t) < 0 \text{ pour } (x, y) \in \Omega_0 \end{cases} \quad 2.13$$

La fonction $\Phi(x, y, t)$ est habituellement prise comme la fonction distance, signée du point (x, y) à l'interface $\Gamma(t)$.

La fonction level set, est associée une fonction Heaviside $H(\Phi)$ qui prend la valeur 1 pour $\Phi \geq 0$ et vaut 0 ailleurs. Pour chaque région Ω_k ($k=0$ ou 1) et en liaison avec cette fonction Heaviside, une fonction d'appartenance g_k est définie telle que $g_0 = 1-H$ et $g_1 = H$. Chaque région ou plus exactement chaque classe Ω_k est caractérisée par une fonction lisse I_k qui est l'intensité moyenne de la région exprimée par l'expression 2.13 suivante :

$$I_k = \frac{\int_{\Omega_k} I(x, y) dx dy}{\int_{\Omega_k} dx dy} = \frac{\int_{\Omega} I(x, y) g_k(x, y) dx dy}{\int_{\Omega} g_k(x, y) dx dy}, \quad k \in \{0, 1\} \quad 2.14$$

Dans le cas d'une image de taille $M \times N$ pixels, l'expression numérique de l'intensité moyenne de la région Ω_k est :

$$I_k = \frac{\sum_{i=1}^M \sum_{j=1}^N I(i, j) g_k(i, j)}{\sum_{i=1}^M \sum_{j=1}^N g_k(i, j)}, \quad k \in \{0, 1\} \quad 2.15$$

La segmentation par modèle Chan et Vese consiste, à partitionner l'image en régions d'intensités homogènes en faisant évoluer la courbe $\Gamma(t)$, donc la fonction $\Phi(x, y, t)$ jusqu'à ce qu'il y ait coïncidence avec les frontières des régions.

L'évolution de la courbe $\Gamma(t)$ est réalisée, par la minimisation de la fonctionnelle énergie exprimée par l'équation 2.16.

$$E(\Phi, I_k) = \sum_{k=0}^1 \int_{\Omega} \lambda_k [(I(x, y) - I_k)^2 g_k(x, y) dx dy] + \mu \int_{\Omega} |\nabla H(\Phi)| dx dy \quad 2.16$$

Le second terme de cette expression, représente la longueur de la courbe Γ pondérée par le paramètre positif μ fixé par l'utilisateur et les λ_k sont des constantes positives prises égales à 1.

La voie classique pour solutionner ce problème de minimisation, est d'utiliser l'équation correspondante d'Euler-Lagrange. Cela conduit à une équation d'évolution qui s'exprime directement à l'aide de la fonction level set Φ :

$$\frac{\partial \Phi}{\partial t} = \delta_{\varepsilon}(\Phi) [\mu \operatorname{div} \left(\frac{\nabla \Phi}{|\nabla \Phi|} \right) - \sum_{k=0}^1 \lambda_k (I - I_k)^2 \frac{\partial g_k}{\partial H}] \quad 2.17$$

δ_{ε} est une version régularisée de la fonction de Dirac qui est la dérivée de la fonction de Heaviside normalisée. Ces deux fonctions sont données par les expressions suivantes :

$$\begin{cases} H_{\varepsilon}(\Phi) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan \left(\frac{\Phi}{\varepsilon} \right) \right) \end{cases} \quad 2.18$$

$$\begin{cases} \delta_{\varepsilon} = H'_{\varepsilon}(\Phi) = \frac{1}{\pi \varepsilon} \left(\frac{1}{1 + \left(\frac{\Phi}{\varepsilon} \right)^2} \right) \end{cases} \quad 2.19$$

Lorsque les constants λ_k sont prises égales à 1, l'expression 2.15 devient :

$$\Delta \Phi = \Delta t. \delta_{\varepsilon}(\Phi) [\mu \operatorname{div} \left(\frac{\nabla \Phi}{|\nabla \Phi|} \right) + (I - I_0)^2 - (I - I_1)^2] \quad 2.20$$

L'équation 2.19 est résolue, par une méthode numérique itérative. A chaque itération, les valeurs de I_k , g_k et donc la fonction Φ sont mises à jour jusqu'à atteindre la convergence ou le nombre d'itérations, fixé par l'utilisateur. La courbe Γ finale obtenue, représente le contour des objets contenus dans l'image. Elle est représentée par le niveau zéro de la fonction level set. L'image segmentée résultante, désignée par $U(x, y)$ est telle que :

$$U(x, y) = \sum_k I_k. g_k(x, y) \quad 2.21$$

2.6.4 Modèle de Lankton [27]

Il s'agit du modèle de Chan et Vese amélioré, puisqu'il permet de segmenter les régions non homogènes. L'approche consiste à faire converger le contour vers des régions localement homogènes suivant les niveaux de gris. Toutefois, la méthode est sensible à l'initialisation. Le critère de l'énergie est:

$$E(\phi) = \int_{\Omega} \delta(\phi(\vec{x})) \int_{\Omega} B(\vec{x}, \vec{y}) F(\phi) d\vec{x} d\vec{y} + \lambda \oint_{\delta\Omega} ds \quad 2.22$$

D'où ;

$$B(\vec{x}, \vec{y}) = \begin{cases} 1 & \text{if } \|\vec{x} - \vec{y}\| < r \\ 0 & \text{if } \|\vec{x} - \vec{y}\| > r \end{cases} \quad 2.23$$

Et :

$$F(\vec{x}, \vec{y}) = (I(\vec{y}) - \mu_{in}(\vec{x}))^2 H(\phi(\vec{y})) + (I(\vec{y}) - \mu_{out}(\vec{x}))^2 H(-\phi(\vec{y})) \quad 2.24$$

La fonction F est une mesure de l'énergie interne générique utilisée pour représenter l'adhésion locale, à un modèle donné, à chaque point du contour.

$$E(\phi) = \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} B(x, y) \cdot F(I(y), \phi(y)) dy dx \quad 2.25$$

Dans le calcul de E, on ne considère que les contributions des points, proches du contour, pour extraire les régions d'intérêt sans altération de l'information.

Le lissage de la courbe, est assuré par le terme de régularisation λ . Parmi les avantages de ce modèle, on cite :

-  Adaptation au changement de topologie : plusieurs objets peuvent être segmentés à la fois;
-  prise en compte des changements de topologie automatique.

2.7 Approche coopérative [28]

La coopération entre les méthodes de segmentation par régions et par contours donne une meilleure segmentation, car elle prend en compte, les caractéristiques des entités de l'image. En effet, les méthodes coopératives prennent avantage de la nature complémentaire de l'information, sur la région et sur le contour. Elles combinent les techniques de segmentation basées sur les régions et celles sur les contours. Ainsi, une segmentation par coopération régions-contours peut être exprimée comme une coopération entre ces deux concepts, afin d'améliorer le résultat final de segmentation [29].

La plupart des travaux de coopération sont orientés coopération régions/contours. La motivation de cette coopération, est due principalement à la dualité des deux concepts et à la possibilité de corriger les inconvénients, d'une approche de segmentation par une autre et inversement.

Cette dualité a donné naissance à la coopération dans la segmentation d'image. Selon la manière de faire coopérer deux processus de segmentation région et contour, on peut distinguer trois approches différentes: la coopération séquentielle, la coopération par fusion de résultats et la coopération mutuelle.

2.7.1 Coopération séquentielle

Le principe général de la coopération séquentielle (figure 2.11), est que l'une des techniques, par région ou par contour, est exécutée en premier lieu. Son résultat est exploité par la deuxième technique, pour améliorer les critères ou les paramètres de la segmentation.

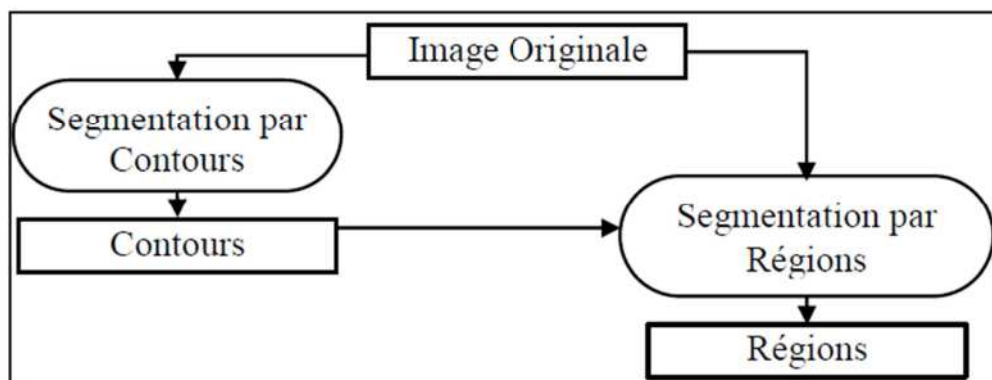


Figure 2.11 : Principe de la coopération séquentielle [28]

2.7.2 Coopération des résultats

Dans la coopération des résultats (figure 2.12), les deux types de segmentation sont exécutés parallèlement et indépendamment. La coopération est faite au niveau des résultats respectifs. Ils sont intégrés dans le but d'atteindre une meilleure segmentation que celle obtenue par une seule des approches.

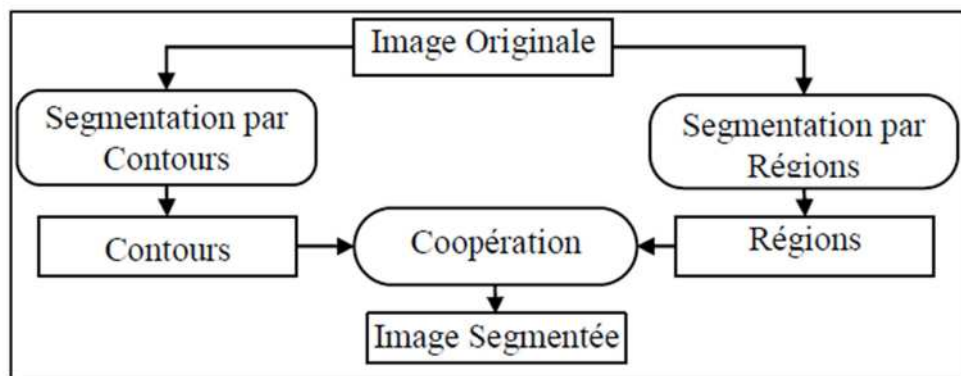


Figure 2.12 : Principe de coopération des résultats [28]

2.7.3 Coopération mutuelle

Dans l'approche de coopération mutuelle (figure 2.13), les différentes techniques de segmentation sont exécutées en parallèle, tout en échangeant mutuellement des informations. La coopération permet de prendre des décisions plus sûres. Si le détecteur de contour et le détecteur de région n'arrivent pas à agréger un pixel, on peut diminuer le seuil de similarité par exemple.

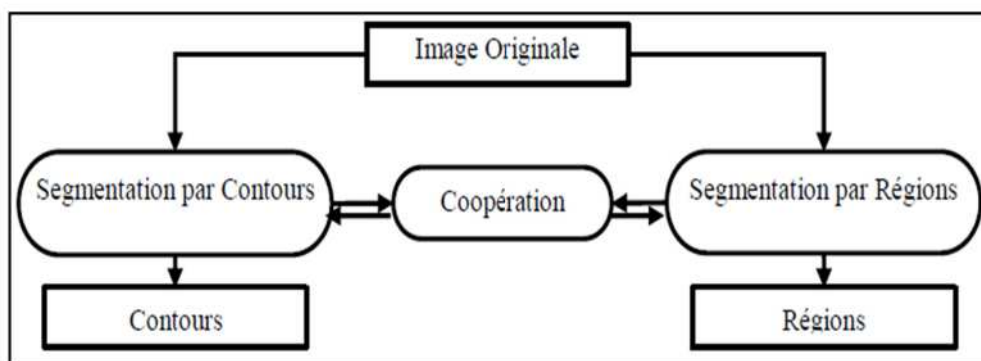


Figure 2.13 : Principe de coopération mutuelle [28]

2.8 Parcours préliminaire aux travaux du laboratoire LATSI

L'équipe d'imagerie du laboratoire LATSI, a mis en œuvre plusieurs systèmes d'aide à la décision pour l'extraction et la reconnaissance de pathologies mammographiques. Ses travaux sont consacrés, à l'analyse topologique et texturale pour la réalisation d'un système d'aide au diagnostic, suivant une classification connexionniste [29], par les arbres de décision [30] ainsi que par les SVM [31] [32] et les algorithmes génétiques en hybridation avec les réseaux de neurones [33][34] [35]. L'objectif de ce travail, est de réaliser, une classification par, le modèle du perceptron multicouche, pour la discrimination des signes malins et bénins des masses et des calcifications mammaires.

2.9 Méthodologie adoptée pour les masses et les calcifications

2.9.1 Méthodologie adoptée pour les masses

La méthode suivie pour la détection des masses (figure 2.14), est une coopération séquentielle, entre la segmentation par la croissance de régions et, le modèle de Lankton.

Un prétraitement est appliqué en premier sur l'image mammographique, pour rehausser le contraste des régions d'intérêt. La deuxième étape, consiste en l'application de la croissance en régions, en tenant compte de la zone expertisée par le radiologue. La délimitation précise de la masse, est effectuée via le modèle de Lankton.

On résume ces étapes par l'organigramme suivant :

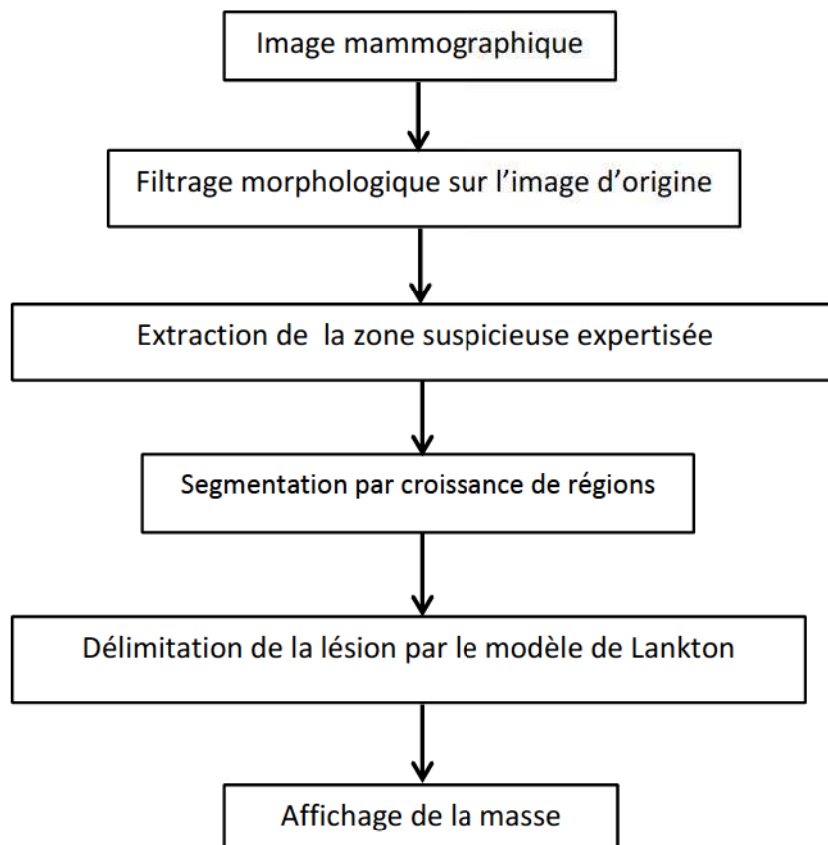


Figure 2.14 : *Processus de détection des masses*

2.9.2 Méthodologie adoptée pour calcifications

La détection des calcifications est appliquée sur la zone expertisée par le radiologue, en respectant les étapes de l'organigramme de la figure 2. L'image, subit un rehaussement du contraste, suivant l'opération du chapeau haut de forme, que l'on superpose à l'image d'origine, à partir de laquelle, on extrait les calcifications, suivant le modèle de Lankton; suivant le choix de paramètres adéquats.

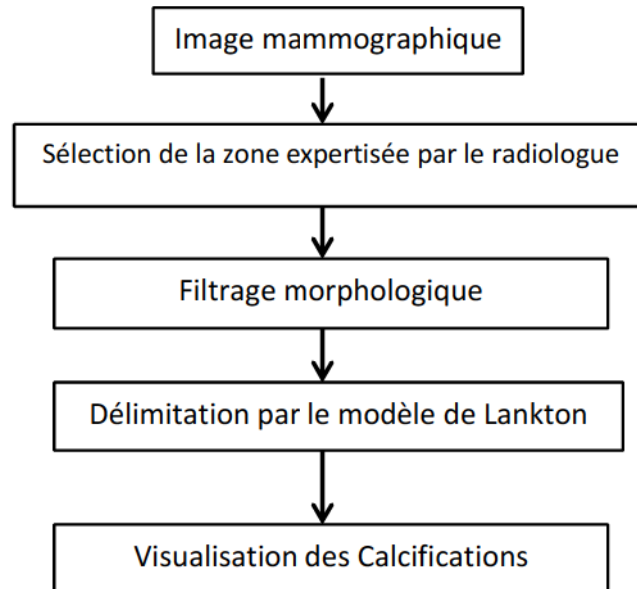


Figure 2.15 : *Processus de détection des calcifications*

2.10 Conclusion

Dans ce chapitre, nous avons parcouru les principales méthodes de prétraitement adoptées par les opérateurs de la morphologie mathématique ainsi que celles de la segmentation effectuées par la croissance de régions et les modèles déformables; pour concevoir notre système de détection des masses et des calcifications. Nous nous sommes intéressés, plus précisément aux modèles déformables, pour une précision des résultats de la segmentation. La segmentation d'image est une étape incontournable et fondamentale dans le traitement de l'image médicale. Ses résultats conditionnent fortement l'étape de l'interprétation si bien, qu'une bonne segmentation assure une bonne caractérisation. Le chapitre qui suit, est consacré à la caractérisation en se basant sur les descripteurs texturaux et morphologiques, ainsi qu'à la classification connexionniste des pathologies mammaires.

Chapitre 3 Caractérisation et classification connexionniste d'images mammographiques

3.1 Introduction

Ce chapitre introduit un point très important de notre travail: l'étude de la caractérisation des images. Par ce terme, nous entendons l'ensemble des méthodes permettant d'extraire des informations relatives à la texture et à la morphologie, par conséquent au contenu structurel de l'image. Cette caractérisation nous dirige vers la classification connexionniste, des masses et des calcifications mammographiques.

3.2 Descripteurs de texture [36][37] [38]

3.2.1 Notion de texture

La notion de texture en traitement d'image, est d'autant plus délicate à aborder que les définitions que l'on en donne, sont multiples. Ces définitions, sont souvent liées à un aspect particulier mais sont rarement génériques et, la quantité importante d'approches utilisées pour l'analyse des textures, témoigne de l'absence d'une définition précise. L'objet de ce chapitre préliminaire, est tout d'abord de faire une synthèse des différentes interprétations, de la notion de texture. Bien que la notion de texture soit naturelle pour l'être humain, elle résiste depuis Longtemps à toute tentative de définition. On peut s'en approcher en disant qu'une texture est une zone de l'image qui présente certaines caractéristiques d'homogénéité qui la font apparaître comme une zone unique. Nous pouvons aussi la décrire comme étant un ensemble de primitives de taille et de forme variables, présentant une organisation spatiale particulière.

3.2.2 Types de texture

a. Texture aléatoire

Ces textures sont désordonnées tout en apparaissant, globalement homogènes. Il est impossible d'isoler ces textures d'un motif de base; car la distribution d'intensité n'est l'objet d'aucune régularité.

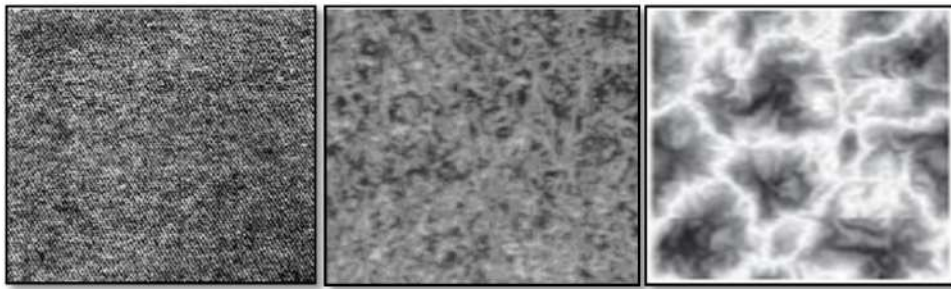


Figure 3.1 : exemples de texture aléatoire [36]

b. Texture périodique

Ce genre de texture est caractérisé par la répétition spatiale d'un motif textural appelé « **texton** ». Il peut être facilement synthétisé à partir du motif de base, de son orientation et de sa taille. Une texture périodique est formée de primitives arrangées d'une façon particulière.

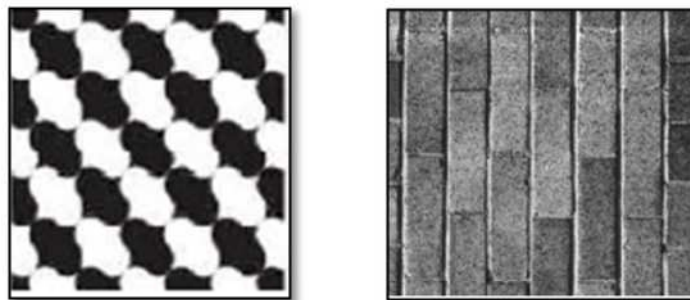


Figure 3.2 : exemples de texture périodique [36]

c. Texture directionnelle

Ce type de texture n'est pas totalement aléatoire et, ne présente pas de motif de base ; mais il se caractérise par certaines orientations.

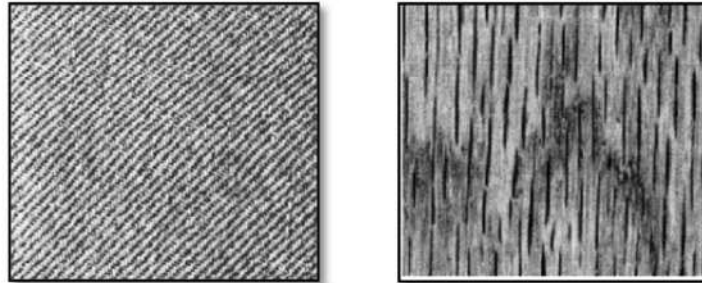


Figure 3.3 : exemples de texture directionnelle [36]

3.2.3 Méthode d'analyse de texture [37]

L'objectif principal de l'analyse de la texture, est la détermination de signatures texturales permettant une meilleure caractérisation de l'image. Autrement dit, c'est la formalisation des descriptifs de la texture par des paramètres mathématiques, qui serviraient à l'identifier. Le choix d'une méthode de caractérisation de textures, est étroitement lié à l'application visée et à la nature, de la texture à discriminer. Ces méthodes peuvent être regroupées en trois catégories.

a. Méthodes statistiques

Ces méthodes sont utilisées pour des structures plus fines, sans régularité apparente. Des paramètres statistiques sont estimés pour chaque pixel de l'image; d'où l'idée de cette méthode est d'étudier, la relation entre un pixel et ses voisins. Suivant la modalité des images à étudier, la signature la plus discriminante de la texture, est à rechercher soit dans des méthodes qui exploitent directement les propriétés statistiques, de la texture (*Matrice de cooccurrence, Matrice de longueur, Matrice de voisinage, ...*) ou bien dans les méthodes qui exploitent les propriétés statistiques, à partir d'un plan transformé dans lequel on réécrit l'image de la texture (*densité spectrale, méthode des extrema locaux, méthodes de transformation de Fourier, ...*).

b. Méthodes à base de modèle

Cette catégorie de méthodes, estime un modèle paramétrique en fonction de la distribution d'intensité des descripteurs de texture calculés. On peut citer quelques exemples de ces méthodes: Les modèles fractales et, les champs aléatoires Markoviens [36].

c. Méthodes fréquentielles

Elles sont basées sur l'analyse des fréquences de l'image, par exemple: les filtres de Gabor, les transformés en ondelettes,

Dans le cas du notre projet, nous allons intéresser aux descripteurs statistiques, estimés par la *matrice de cooccurrence*.

3.2.4 Matrice de cooccurrence [38]

Elle estime des propriétés des images relatives, à des statistiques du second ordre. Une matrice de cooccurrence, mesure la probabilité d'apparition des paires de valeurs de pixels situés à une certaine distance, dans l'image. Elle est basée sur le calcul de la probabilité $P(i, j, d, \theta)$ qui représente le nombre de fois où un pixel de niveau de couleur i , apparaît à une distance relative d , d'un pixel de niveau de couleur j et selon une orientation θ donnée. Les directions angulaires θ classiquement utilisées, sont 0, 45, 90 et 135 degrés. Les relations de voisinage entre pixels, nécessaires au calcul des matrices, sont illustrées dans la figure 3.4. Par exemple, les plus proches voisins de 'x' selon la direction $\theta=135$ degrés, sont les pixels 4 et 8. La distance « d » permet d'avoir une description significative de la périodicité de la texture et l'angle « θ » permet d'évaluer la direction de texture. Cette matrice décrit les régularités observables dans les niveaux de gris des pixels d'une région.

Afin de limiter le nombre de calculs, on prend fréquemment, pour une distance $d=1$, une orientation de l'angle $\theta=0^\circ$.

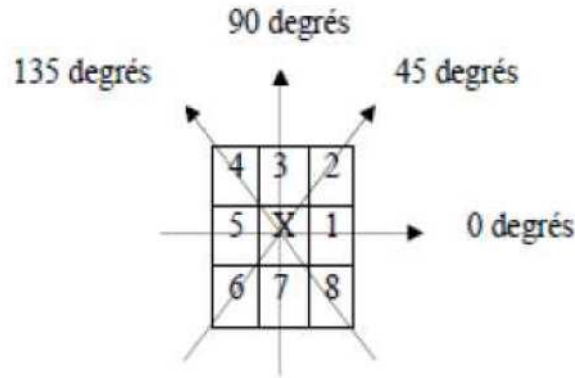


Figure 3.4 : plus proches voisins de x selon 4 directions [36]

Les matrices de cooccurrence contiennent une masse d'information difficilement manipulable, c'est pour cela qu'elle n'est pas utilisée directement. De ce fait, quatorze indices définis par Haralick [38], qui correspondent à des caractères descriptifs de la texture peuvent être calculés, à partir de ces matrices.

Nous présentons les 6 paramètres texturaux considérés, comme les plus pertinents, dans le cas de notre travail.

- **L'entropie**

Ce paramètre (équation 3.1) fournit un indicateur du désordre que présente, une texture. L'entropie atteint de fortes valeurs lorsque la texture, est complètement aléatoire (sans structure apparente).

$$ENT = - \sum_i \sum_j (\log P_{ij}(d, \theta) P_{ij}(d, \theta)) \quad 3.1$$

- **Le contraste**

Il peut être définie, comme le rapport entre les parties les foncées et les parties les plus claires d'une image. La valeur de ce paramètre (équation 3.2), est d'autant plus élevée, que la texture, présente un fort contraste.

$$CST = \sum_i \sum_j ((i, j)^2 P_{ij}(d, \theta)) \quad 3.2$$

- **La variance**

La variance (équation 3.3) mesure l'hétérogénéité de la texture. Elle augmente lorsque les niveaux de gris différent, de leur moyenne μ . La variance est indépendante du contraste.

$$VAR = \sum_i \sum_j ((i - \mu)^2 P_{ij}(d, \theta)) \quad 3.3$$

- **L'énergie**

Ce paramètre (équation 3.4), mesure l'uniformité de la texture. Il atteint de fortes valeurs, lorsque la distribution des niveaux de gris est constante, ou de forme périodique.

Les valeurs élevées d'énergie, sont obtenues pour les matrices $P(d, \theta)$ lorsque (d, θ) correspond à la période.

$$ENE = \sum_i \sum_j P_{ij}(d, \theta) \quad 3.4$$

- **La corrélation**

Elle est définie (équation 3.5), comme le rapport entre les parties les plus foncées et les parties les plus claires, d'une image. Mesurer le contraste revient à évaluer la dispersion des niveaux de gris d'une image.

$$COR = \sum_i \sum_j \left(\frac{(i-\mu)(j-\mu)P_{ij}(d, \theta)}{\sigma^2} \right) \quad 3.5$$

- **Homogénéité**

Plus on retrouve le même couple de pixels, plus cet indice est élevé, par exemple image uniforme, ou texture périodique dans le sens de la translation. Ce paramètre est corrélé à une combinaison linéaire de l'énergie et du contraste.

$$Hom = \sum_i \sum_j \left(\frac{P_{ij}(d, \theta)}{1+(i+j)^2} \right) \quad 3.6$$

3.3 Descripteurs de forme [29][39][40][41]

Dans la littérature, des dizaines de paramètres de forme, permettent de caractériser la morphologie des particules. Une des difficultés de l'analyse morphologique, réside dans le choix du paramètre de forme les plus adaptés à la nature des structures médicales.

Les caractéristiques que nous avons adoptées, ce sont celles qui nous paraissent les plus pertinentes, pour modéliser numériquement les régions d'intérêt mammaires. Elles permettent ainsi, d'identifier les pathologies mammaires, en se basant sur la compacité, la rectangularité, la convexité et les facteurs de Gupta.

- **La compacité**

Cette mesure décrit la géométrie des motifs. La compacité vaut 1 lorsque la région est un cercle et diminue, à mesure que la région est allongée.

$$COM = \frac{4\pi A}{P^2} \quad 3.7$$

Où A représenté la surface de la région d'intérêt et P, le périmètre correspondant.

- **La convexité**

Ce paramètre mesure le rapport entre la surface de la région et, celle de son enveloppe convexe. Elle donne une valeur de 1 si elle épouse parfaitement l'enveloppe convexe et si elle est inférieure à 1 elle est faiblement convexe.

$$CONV = \frac{S_{région}}{S_{enveloppe\ convexe}} \quad 3.8$$

- **La rectangularité**

Ce descripteur de forme est connu, comme rectangularité ou boîte englobante. Il est souvent utilisé pour décrire le taux de rectangularité et d'allongement d'une région. On désigne par R la région et BE la boîte englobante qui est définie par le rectangle, contenant la lésion étudiée. La mesure de rectangularité notée $Rect$ peut alors se définir comme suit :

$$Rect = \frac{aire(R)}{aire(Be)} \quad 3.9$$

Selon cette formule, plus l'objet à décrire est rectangulaire plus la valeur de rectangularité est proche de 1.

- **Facteur de Gupta [29][39]**

L'extraction de ce type de descripteur, consiste à décomposer la séquence de N pixels du contour de la région d'intérêt, dans une base de fonction.

Ce type de primitives, permet l'invariance via la rotation, la translation, le changement d'échelle ainsi que la réflexion.

Soit à cet effet, les coordonnées de N pixels du contour C décrit par $C_i(x_i, y_i)$ avec $i^{\text{ème}}$ point allant de 1 à N.

Dans ce cas le $P^{\text{ième}}$ moment est défini par:

$$m_p = \frac{1}{N} \sum_{n=0}^{N-1} [d(n)]^p \quad 3.10$$

Telle que la distance euclidienne correspond à la relation suivante:

$$d(n) = \sqrt{[x(n) - \bar{x}]^2 + [y(n) - \bar{y}]^2} \quad 3.11$$

Le $P^{\text{ième}}$ moment central est défini comme :

$$M_p = \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^p \quad 3.12$$

Ces moments engendrent des moments normalisés, invariants relativement à la translation, la rotation, à l'homothétie ainsi qu'à la réflexion. Ils se définissent comme suit :

$$\overline{m_p} = \frac{m_p}{(M_2)^{p/2}} = \frac{\frac{1}{N} \sum_{n=0}^{N-1} [d(n)]^p}{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^2 \right\}^{p/2}} \quad 3.13$$

$$\overline{M_p} = \frac{M_p}{(M_2)^{p/2}} = \frac{\frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^p}{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^2 \right\}^{p/2}} \quad 3.14$$

De fait de leur sensibilité au bruit, Gupta [39] sélectionne des moments de faible ordre pour établir les facteurs suivants:

$$F_1 = \frac{(M_2)^{1/2}}{m_1} = \frac{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^2 \right\}^{1/2}}{m_1} \quad 3.15$$

$$F_2 = \frac{M_3}{(M_2)^{3/2}} = \frac{\frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^3}{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^2 \right\}^{3/2}} \quad 3.16$$

$$F_3 = \frac{M_4}{(M_2)^2} = \frac{\frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^4}{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^2 \right\}^2} \quad 3.17$$

Les variations des facteurs F_2 et F_3 sont faibles et ne montrent guère, une progression d'une catégorie de forme à une autre. De plus, F_2 varie significativement au sien de chaque

échantillon; ce qui produit un mauvais indicateur de forme. Donc un nouvel ensemble de caractéristiques est proposé selon Gupta :

$$F_1 = \frac{(M_2)^{1/2}}{m_1} = \frac{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^2 \right\}^{1/2}}{m_1} = \hat{F}_1 \quad 3.18$$

$$\hat{F}_2 = \frac{(M_3)^{1/3}}{m_1} = \frac{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^3 \right\}^{1/3}}{m_1} \quad 3.19$$

$$\hat{F}_3 = \frac{(M_4)^{1/4}}{m_1} = \frac{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^4 \right\}^{1/4}}{m_1} \quad 3.20$$

Le moment normalisé ($mf = \hat{F}_3 - \hat{F}_1$) est donné par la relation 3.21.

$$m = \frac{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^4 \right\}^{1/4} - \left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - m_1]^2 \right\}^{1/2}}{m_1} \quad 3.21$$

C'est ce facteur qui sera considéré, pour caractériser, selon Gupta, les lésions mammographiques.

3.4 Classification connexionniste [29][42][43][44]

La classification est la dernière étape, d'un système de diagnostic assisté par ordinateur. Elle exploite le résultat de la description, (qui lui-même exploite le résultat de la segmentation) pour pouvoir décider, de la nature pathologique des masses et des calcifications.

La notion de classification signifie, l'affectation d'une étiquette à des échantillons d'une base de données, en utilisant un certain nombre de caractéristiques. Ces caractéristiques doivent bien évidemment représentatifs de, chaque échantillon.

Dans notre étude, la classification adoptée, est celle du perceptron multicouche. On présente dans ce qui suit quelques classifieurs connexionnistes, avant d'aborder le réseau choisi dans notre étude.

3.4.1 Les réseaux de neurones

Un réseau neuronal est l'association, en un graphe plus ou moins complexe, d'objets élémentaires, les neurones formels. Les principaux réseaux se distinguent par l'organisation du graphe, de leur architecture suivant le nombre de neurones, la présence ou non de boucles de rétroaction dans le réseau, le type des neurones (leurs fonctions de transition ou d'activation) et enfin par l'objectif visé: apprentissage supervisé ou non.

1. Neurone biologique

De façon très réductrice, un neurone biologique (figure 3.5) est une cellule qui se caractérise par:

- des synapses, les points de connexion avec les autres neurones, les fibres nerveuses ou musculaires;
- des dendrites, les "entrées" des neurones; qui servent à recevoir les signaux provenant des autres neurones en amont.
- l'axone, la "sortie" du neurone vers d'autres neurones ou des fibres musculaires;
- le noyau qui active la sortie en fonction des stimulations en entrée.

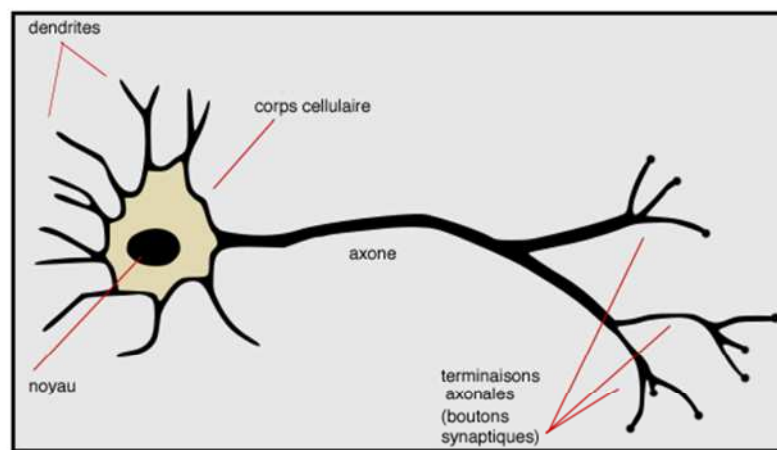


Figure 3.5 : schéma d'un neurone biologique [42]

C'est à partir de l'hypothèse que le comportement intelligent humain et de la structure et des éléments de bases du système nerveux central (que sont les neurones); que l'on a développé les réseaux de neurones artificiels. Cette inspiration à partir du modèle

biologique provient du fait que le cerveau humain, est un système apprenant basé sur une structure contenant environ 100 milliards de neurones reliés, entre eux par un million de milliards de synapses.

2. Les neurones artificiels

Les RNA (réseaux de neurones artificiels) sont des modèles d'entrées-sorties basés sur les neurones biologiques. Le but initial de cette modélisation, est de reproduire les capacités du cerveau humain à interpoler ou à classifier des entités.

Selon Haykin [43], un RNA (figure 3.6) est un processus distribué de manière massivement parallèle, qui a une capacité naturelle de mémoriser des connaissances de façon expérimentale et, de les rendre disponibles pour l'utilisation.

Il ressemble au cerveau en deux points:

- la connaissance est acquise à travers, un processus d'apprentissage;
- les poids des connexions entre les neurones, sont utilisés pour mémoriser la connaissance.

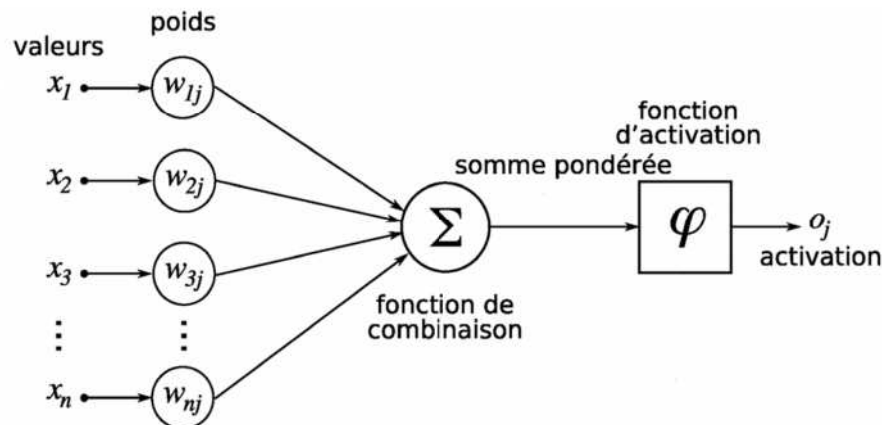


Figure 3.6 : structure d'un neurone artificiel [42]

Les entrées du réseau, notées $x_1, \dots, x_n$ (ce sont les variables explicatives du modèle), sont liées aux cellules de la couche suivante, par des poids synaptiques w_{nj} déterminant; l'effet du signal par le neurone n sur le neurone j .

Chaque cellule de la première couche cachée (deuxième couche dans le réseau) va évaluer l'intensité de l'information qui ; lui parvient de la couche d'entrée en calculant la somme pondérée (3.22).

$$O_j = \sum w_{nj} x_n \quad 3.22$$

La cellule en question, va traiter l'information, puis la transmet à la cellule suivante à l'aide d'une fonction d'activation ζ qui anime le neurone en déterminant son activation donnée par O_j équivalente de la sortie de neurone. Elle est égale à:

$$O_j = \zeta (\sum w_{nj} x_n + \theta_j) \quad 3.23$$

Il existe différents types de fonctions de transfert pour le neurone artificiel (figure 3.7).

- a : fonction à seuil (S, la valeur du seuil),
- b : linéaire par morceaux,
- c : sigmoïde.

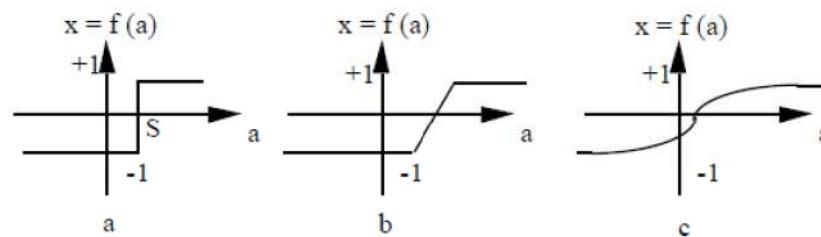


Figure 3.7 : Différents types de fonctions de transfert pour le neurone artificiel [42]

3.4.2 Architecture [42][43]

On peut classer les RNA en deux grandes catégories.

1. Les réseaux "feed-back"

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

a) Les cartes auto-organisatrices de Kohonen

Ce sont des réseaux à apprentissage non-supervisé qui établissent une carte discrète, ordonnée topologiquement, en fonction de couches 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.

b) 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.

2. Les réseaux "feed-forward"

Ils sont 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.

a) Les réseaux à fonction radiale

Ce sont les réseaux que l'on nomme aussi RBF ("*Radial Basic Function*"). L'architecture est la même que pour les perceptrons multicouche (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. 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) Le perceptron

Le perceptron est l'un des réseaux de neurones les plus utilisés pour des problèmes d'approximation, de classification et de prédiction; c'est un classifieur linéaire. Ce type de réseau neuronal ne contient aucun cycle (*feed-forward* neural network).

❖ Le perceptron monocouche

C'est historiquement le premier RNA, c'est le Perceptron de *Rosenblatt* [42]. C'est un réseau simple, puisque 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 (ou selon la règle de Hebb [42]).

❖ 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. Ce réseau, peut résoudre des problèmes non-linéairement séparables et des problèmes logiques plus compliqués, notamment le fameux problème du XOR. Il suit aussi un apprentissage supervisé, selon la règle de correction de l'erreur.

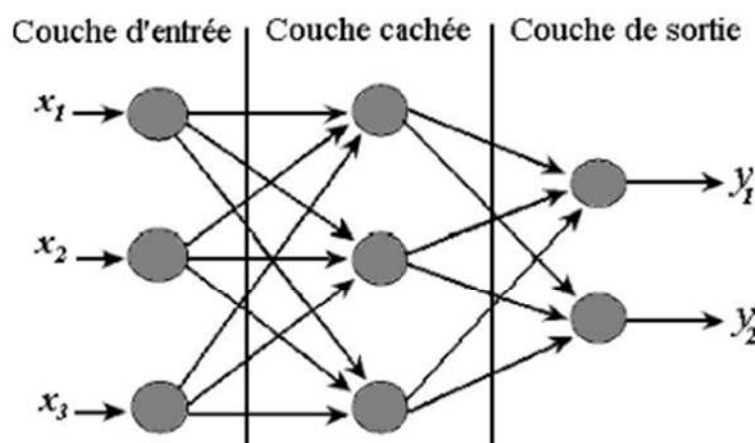


Figure 3.8 : Schéma d'un PMC [42][43]

- **La couche d'entrée**

La première couche, est appelée couche d'entrée. Elle recevra les données source, que l'on veut utiliser pour l'analyse. Dans le cas de l'aide au diagnostic médical, cette couche recevra les symptômes. Sa taille est donc directement déterminée, par le nombre de variables d'entrées [43].

- **La couche cachée**

La seconde couche est une couche cachée, en ce sens qu'elle n'a qu'une utilité intrinsèque pour le réseau de neurones et, n'a pas de contact direct avec l'extérieur. Les fonctions d'activations, sont en général non linéaires sur cette couche, mais il n'y a pas de règle à respecter. Le choix de sa taille n'est pas implicite et doit être, ajusté. En général, on peut commencer par une taille moyenne des couches d'entrée et de sortie mais ce n'est pas toujours, le meilleur choix. Il sera souvent préférable, pour obtenir de bon résultats, d'essayer le plus de tailles possibles.

- **La couche de sortie**

La dernière couche, est appelée couche de sortie. Elle donne le résultat obtenu, après compilation par le réseau des données de l'entrée, dans la première couche. Dans le cas de l'aide au diagnostic médical, cette couche indique le type de pathologie. La taille de cette couche, est directement déterminée par le nombre de variables que l'on veut, en sortie [44].

➤ **Fonction d'activation du PMC [42]**

Les neurones d'un réseau possèdent des fonctions d'activation (figure 3.7), qui vont transformer les signaux émis par les neurones de la couche précédente à l'aide, d'une fonction mathématique. Cette fonction est la fonction d'activation f elle-même, qui peut profondément influencer sur la performance du réseau. Il est donc important de bien choisir ce type de fonction d'activation, dans un réseau de neurones. Les neurones d'entrée, ne possèdent généralement aucune fonction d'activation.

La fonction f , peut être paramétrée de façon quelconque, mais deux types de paramétrage sont fréquemment utilisés:

- Les paramètres sont attachés aux entrées du neurone; la sortie est une fonction non linéaire d'une combinaison des entrées, pondérées par les paramètres.

- Les paramètres sont attachés à la non-linéarité du neurone et ils interviennent directement, dans la fonction d'activation.

➤ Utilisation du PMC [42]

L'algorithme le plus simple utilisant un perceptron, est l'algorithme par correction d'erreur. Il s'agit de présenter au perceptron, différents vecteurs d'entrées associées à une variable booléenne. On appelle vecteur d'entrée, un vecteur qui initialise chaque entrée du perceptron à 0 ou 1. C'est l'ensemble des vecteurs d'entrées qu'on appelle échantillon. Le booléen représente la classe associée au vecteur. Tous les poids sont initialisés à une valeur aléatoire ; cependant au moins un poids doit avoir une valeur non nulle pour que, la sortie ne soit pas toujours égale à zéro.

L'algorithme consisté à présenter au perceptron, un certain nombre de fois ces vecteurs d'entrée. Lorsqu'un vecteur est présenté au perceptron, les poids sont modifiés à chaque fois que la sortie est différente, de la valeur de la classe. L'apprentissage est terminé une fois que tous les vecteurs ont été présentés, sans modification des poids.

Il convient de remarquer qu'il s'agit effectivement d'un algorithme, par correction d'erreur car les poids ne sont modifiés que si la sortie calculée, est différente de la sortie attendue. La règle de modification des poids est la suivante:

$$W_i \leq W_i + (C-O).X_i \quad 3.24$$

où :

W_i : le poids de l'entrée ;

C : la classe ;

O : la sortie calculée ;

X_i : l'entrée i .

➤ Limites du PMC

Ce modèle a rapidement, montré ces limites. En effet, l'algorithme exige un échantillon linéairement séparable, ce qui est rarement le cas des données que l'on veut faire apprendre artificiellement. Dans le cas où l'échantillon n'est pas linéairement séparable, l'algorithme ne

se terminera pas; c'est pourquoi les réseaux de neurones multicouches, ont été mis en place.

Cette contrainte ne permettra pas l'apprentissage du XOR (ou exclusif) par l'algorithme basé sur la notion du perceptron. Il apparaît évidemment impossible, de séparer linéairement XOR en deux classes distinctes.

Pour améliorer l'apprentissage du perceptron, Widrow et Hoff [42] ont proposé une nouvelle règle d'apprentissage plus proche des mécanismes d'optimisation. Il s'agit non seulement de trouver la frontière séparant les classes lorsqu'elle existe; mais aussi de trouver la meilleure solution qui minimise, l'erreur lorsque cette séparation linéaire, n'est pas possible.

3.4.3 L'apprentissage des réseaux de neurones [29][42][43][44]

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 des RNA, peut se faire selon différentes règles.

1. Types d'apprentissage [42][43]

a) Apprentissage supervisé

Dans ce type d'apprentissage, le réseau s'adapte par comparaison entre le résultat qu'il calcule, 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.

b) Apprentissage 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 des catégories, en attribuant et en optimisant une valeur de qualité, aux catégories reconnues.

2. Règles d'apprentissage

Dans les systèmes experts, les connaissances de l'expert ont une forme énumérée: elles sont exprimées sous forme de règles. Dans le cas des réseaux de neurones, les connaissances ont une forme distribuée: elles sont codées dans les poids des connexions, la topologie du réseau, les fonctions de transfert de chaque neurone, le seuil de ces fonctions et, la méthode d'apprentissage utilisée. Il existe un certain nombre de méthodes d'apprentissage.

a) Règle de Hebb

Cette règle, basée sur des données biologiques, modélise le fait que si des neurones, de part et d'autre d'une synapse, sont activés de façon synchrone et répétée, la force de la connexion synaptique augmente. Il est à noter ici que l'apprentissage est localisé, c'est-à-dire que la modification d'un poids synaptique w_{ij} ne dépend que, de l'activation d'un neurone i et d'un autre neurone j .

b) Rétro-propagation du gradient de l'erreur [29][44]

Cet algorithme est utilisé dans les réseaux de type feed-forward. Ce sont des réseaux de neurones à couche, ayant une couche d'entrée, une couche de sortie et au moins, une couche cachée. Il n'y a pas de récursivité dans les connexions et, pas de connexions entre les neurones de la même couche. Le principe de la rétro-propagation consiste à présenter au réseau un vecteur d'entrées, de procéder au calcul de la sortie par propagation à travers les couches, de la couche d'entrées vers la couche de sortie. La sortie obtenue, est comparée à la sortie désirée, une erreur est alors obtenue. A partir de cette erreur, est calculé le gradient de l'erreur qui est à son tour propagé, de la couche de sortie vers la couche d'entrée, d'où le terme de rétro-propagation.

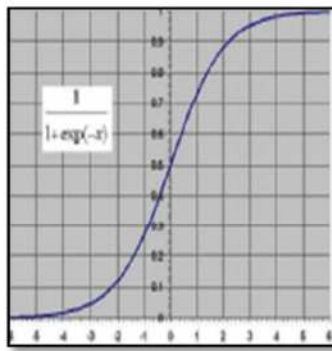
Cela permet la modification des poids du réseau et réalise, donc l'apprentissage. L'opération est réitérée pour chaque vecteur d'entrée et cela, jusqu'à ce que le critère d'arrêt soit vérifié.

L'algorithme de rétro-propagation, consiste à déterminer l'erreur effectuée par chaque neurone puis à modifier les poids, pour minimiser cette erreur.

Mais il est nécessaire d'appliquer cet algorithme jusqu'à ce que l'erreur quadratique (la somme des carrés des différences entre l'entrée obtenue et l'entrée désirée), soit inférieure à un certain seuil (équation 3.25).

$$E = \sum (S_i - Y_i)^2 \quad 3.25$$

Le réseau est représenté par trois couches de neurones. Chaque neurone d'une couche étant relié à tous les neurones de la couche suivante, par une liaison pondérée. Chaque neurone intègre les données grâce à une fonction non linéaire, qui permet l'apprentissage de classe non linéairement séparable; en utilisant la fonction sigmoïde (figure 3.9 et équation 3.26).



$$\left. \begin{aligned} s(x) &= \frac{1}{1 + e^{-x}} \\ ds(x) &= s(x)(1 - s(x)) \end{aligned} \right\} 3.26$$

Figure 3.9 : représentation de la fonction sigmoïde [42]

On effectue tout d'abord une propagation de l'entrée à travers le réseau. Considérons dans ce cas, **S** le vecteur de sortie obtenue après la propagation de l'entrée **X** à travers le réseau, et **Y** le vecteur des sorties que l'on voudrait, obtenir. Pour la couche de sortie on considère le neurone *i* ; on obtient alors l'erreur pour ce neurone: $S[i] - Y[i]$.

On nomme *p* la taille de la couche cachée. Le neurone *i* reçoit alors les *p* sorties, de neurones de la couche cachée modulées de leurs poids respectifs. Le gradient de l'erreur permet la convergence du réseau, suivant la modification des poids. On obtient la formule suivante :

La règle de modification des poids pour un exemple *k* ;

$$W_{ij} = W_{ij}(k-1) - pas * D_i * O_j * ds(I_i) \quad 3.27$$

Pour la couche de sortie ; $D_i = (S_i - Y_i)$

Avec :

W_{ij} : le poids entre le neurone i d'une couche et le neurone j de la couche précédente ;

Pas : pas d'apprentissage ;

O_j : sortie du neurone j ;

I_i : entrée du neurone i avec : $I_i = \sum W_{ij} O_j$

S_i : sortie obtenue ;

Y_i : sortie recherchée ;

ds : la fonction sigmoïde.

La méthode consiste à remarquer que l'erreur effectuée, par un neurone de la couche cachée, est la somme des erreurs des neurones de la couche de sortie, modulée par les poids respectifs des liaisons entre les deux couches. L'erreur pour le neurone i de la couche cachée devient :

$$\sum_h W_{hi} * (S[h] - Y[h]) \quad 3.28$$

$S(h)$: sortie calculée.

On obtient alors la formule de modification des poids suivante :

$$W_{ij}(k) = W_{ij}(k - 1) - pas * D_i * O_j * ds(I_i) \quad 3.29$$

Pour la couche cachée :

$$D_i = \sum_h D_h * W_{hi} \quad 3.30$$

3.4.4 Avantages et inconvénients des réseaux de neurones [42]

Les réseaux de neurones présentent des avantages et des inconvénients que nous citons ci-dessous.

1. Avantages

- ✓ Précision du classifieur (si bien paramétré).
- ✓ Faculté d'apprentissage à partir d'exemples représentatifs, par « rétro-propagation des erreurs.
- ✓ Apprentissage automatique des poids.
- ✓ possibilité de faire le parallélisme (les éléments de chaque couche peuvent fonctionner en parallèle);

- ✓ résistance aux pannes (si un neurone ne fonctionne plus, le réseau ne se perturbe pas).

2. Inconvénients

Les inconvénients se résument suivant :

- ✓ la complexité de l'architecture du réseau.
- ✓ Difficulté du paramétrage, surtout pour le nombre de neurones, dans la couche cachée.
- ✓ Le problème du sur-apprentissage (apprentissage au détriment de la généralisation).

3.5 Conclusion

Ce chapitre a fait l'objet d'une analyse texturale, qui se base sur quelques descripteurs statistiques tels que l'entropie et la variance, qui nous informent sur la distribution des niveaux de gris ainsi que sur l'aspect des pathologies mammaires. Une description morphologique, a également été prise en compte suivant la compacité, la surface et d'autres attributs, afin d'identifier le changement de la forme, des lésions mammographiques. La discrimination de ces lésions, est réalisée à partir des vecteurs d'attributs caractéristiques par une classification suivant le perceptron multicouche.

Le chapitre suivant, présente quelques résultats engendrés par notre méthodologie, appliquée sur des images mammographiques réelles. Ces résultats sont suivis d'une discussion.

Chapitre 4 Application “Mammoconnex” sur les images mammographiques

4.1 Introduction

L'idée principale de ce chapitre est de présenter l'application développée “Mammoconnex”, ainsi que tout traitement appliqué, sur les images mammographiques pour l'extraction et la classification connexionniste des pathologies mammaires issues de la base de données MIAS (Mammographic Image Analysis Society) [45]. Ces images de type MLO (incidence latérale oblique) ont une résolution spatiale de 1024 x 1024. Les images que nous avons étudiées comprennent des masses et des calcifications, réparties selon des cas malins et bénins.

4.2 Environnement de travail

Le développement du notre système est fait, sous environnement Windows 7 en utilisant un outil de programmation (Matlab version 8.4 R2014b).

4.2.1 Matériel utilisé

Notre système a été implémenté à l'aide d'un PC portable HP Pavilion G6 Notebook PC, avec un processeur Intel Core i3 CPU M380 @ 2.53GHz 2.53 GHz et une mémoire (RAM) de 4.00 Go.

4.2.2 Présentation de l'outil Matlab

Le logiciel Matlab [46] est un logiciel de manipulation de données numériques et de programmation dont le champ d'application, est essentiellement les sciences appliquées. Son objectif, par rapport aux autres langages, est de simplifier au maximum la transcription en langage informatique d'un problème mathématique, en utilisant une écriture la plus proche possible, du langage naturel scientifique.

Le nom Matlab vient de Matrix Laboratory. En Matlab les objets sont tous par défaut des matrices. Une variable réelle est donc vu par Matlab comme une matrice 1x1. Le produit est donc par défaut un produit matriciel.

Pour lancer Matlab, il suffit de lancer la commande **Matlab** dans un terminal, où s'ouvre alors la fenêtre principale de **Matlab**. On peut y lancer directement des lignes de commande mais la plupart du temps on passera par l'éditeur de Matlab, qui permet de réaliser des scripts et des fonctions.

L'intérêt de **Matlab** tient, d'une part, à sa simplicité d'utilisation: pas de compilation, déclaration implicite des variables utilisées et, d'autre part, à sa richesse fonctionnelle: arithmétique matricielle et nombreuses fonctions de haut niveau dans divers domaines (analyse numérique, statistique, commande optimale, représentation graphique, ...).

On peut utiliser **Matlab**, soit en mode *en ligne*, c'est-à-dire saisir des commandes dans la fenêtre d'exécution au fur et à mesure, soit en mode *programmation*, en écrivant dans des fichiers séparés (*.m) l'enchaînement des commandes. Ces fichiers s'appellent des *scripts* et on les construit à l'aide de n'importe quel éditeur de texte (par exemple *emacs*, ...). Le mode en ligne permet d'obtenir des résultats simples qui ne sont pas sauvegardés. Le mode programmation, quant à lui, permet de développer des applications très complexes.

a) Fonctionnalités du Matlab

Pour débiter avec Matlab, il est mieux d'utiliser l'aide en ligne. Comme Il existe plusieurs commandes et fonctions de MATLAB que l'on pourra trouver dans les différentes boîtes à outils (toolbox) qui viennent compléter, le logiciel de base.

➤ **Les fonctions d'aide :**

- **help** : génère l'aide d'une commande ;
- **info** : fournit des informations sur la version de Matlab ;
- **lookfor** : recherche du mot-clé indiqué dans toutes les entrées d'aide ;
- **demo** : lance la démonstration.

➤ **Les fonctions liées à l'espace de travail :**

- **clear** : supprime toutes les variables du Workspace ;
- **who, whos** : fait la liste des variables de l'espace de travail ;
- **load** : restaure l'espace de travail à partir du disque ;
- **save** : sauvegarde l'espace de travail sur le disque.

➤ **Les fonctions liées à la gestion des fichiers :**

- **path** : visualise ou change le chemin de recherche du répertoire de Matlab ;
- **mkdir** : réalise la création d'un répertoire.
- **pwd** : indique le répertoire courant.

b) Présentation du GUI

Les interfaces graphiques (ou interfaces homme-machine) sont appelées GUI (pour Graphical User Interface) sous MATLAB. Elles permettent à l'utilisateur d'interagir avec un programme informatique, grâce à différents objets graphiques (boutons, menus, cases à cocher...). Avec une interface de ce type, l'utilisateur n'a pas besoin de connaître un langage ni de saisir des commandes pour se servir d'une application. L'interface graphique contient généralement des commandes telles que des menus, des barres d'outils, des boutons et des curseurs.

➤ **Création de l'interface**

A l'aide de l'éditeur de mise en page GUIDE, on peut créer notre interface (figure 4.1) utilisateur de manière graphique. GUIDE génère alors automatiquement le code Matlab correspondant, qui nous pouvons ensuite modifier afin de définir le comportement de notre application.

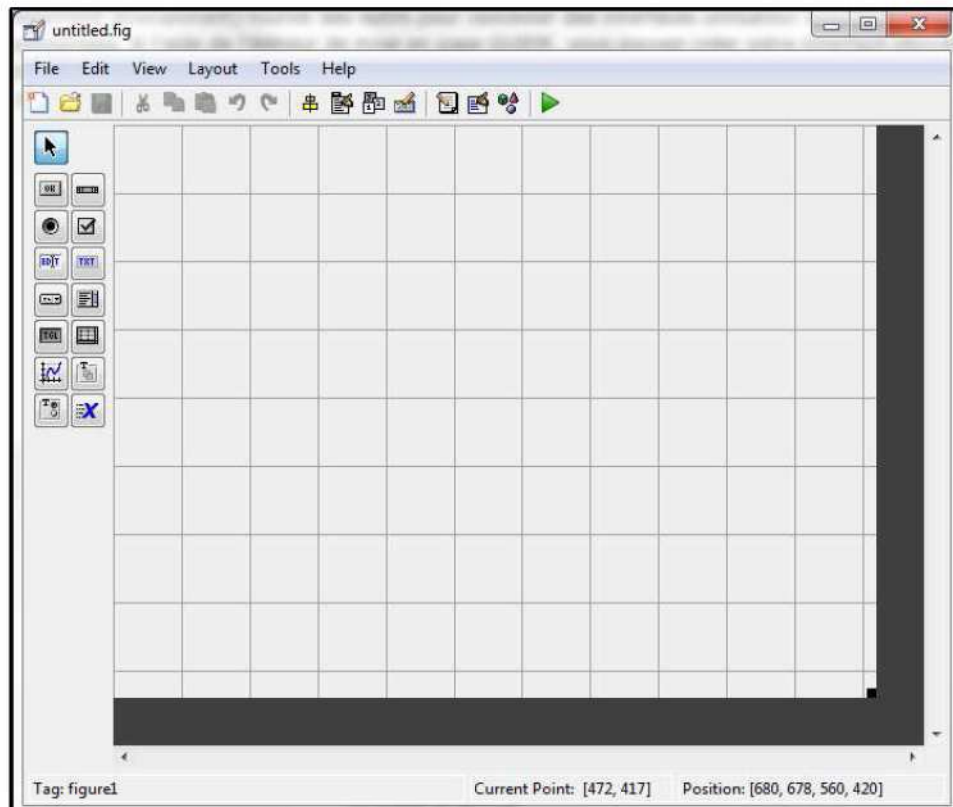


Figure 4.1 : fenêtre principale du GUI

Quelques fonctions MATLAB permettent de gérer certains identifiants particuliers :

- **gca** : récupère l'identifiant de l'objet Axes courant ;
- **gcbf** : récupère l'identifiant de l'objet Figure où se trouve l'objet graphique dont l'action est en cours ;
- **gcbo** : récupère l'identifiant de l'objet graphique dont l'action est en cours ;
- **gcf** : récupère l'identifiant de l'objet Figure courant ;
- **gco** : récupère l'identifiant de l'objet graphique courant.

4.3 Présentation du système développé

Notre système (figure 4.2) permet une classification des pathologies mammaires suivant le perceptron multicouches après avoir appliqué en premier lieu un rehaussement du contraste, puis une segmentation en région pour bien extraire les zones d'intérêt, suivie

d'une caractérisation texturale et morphologique de chaque image, en finissant par la classification à l'aide du perceptron multicouches. On peut présenter tous ces étapes dans cet organigramme :

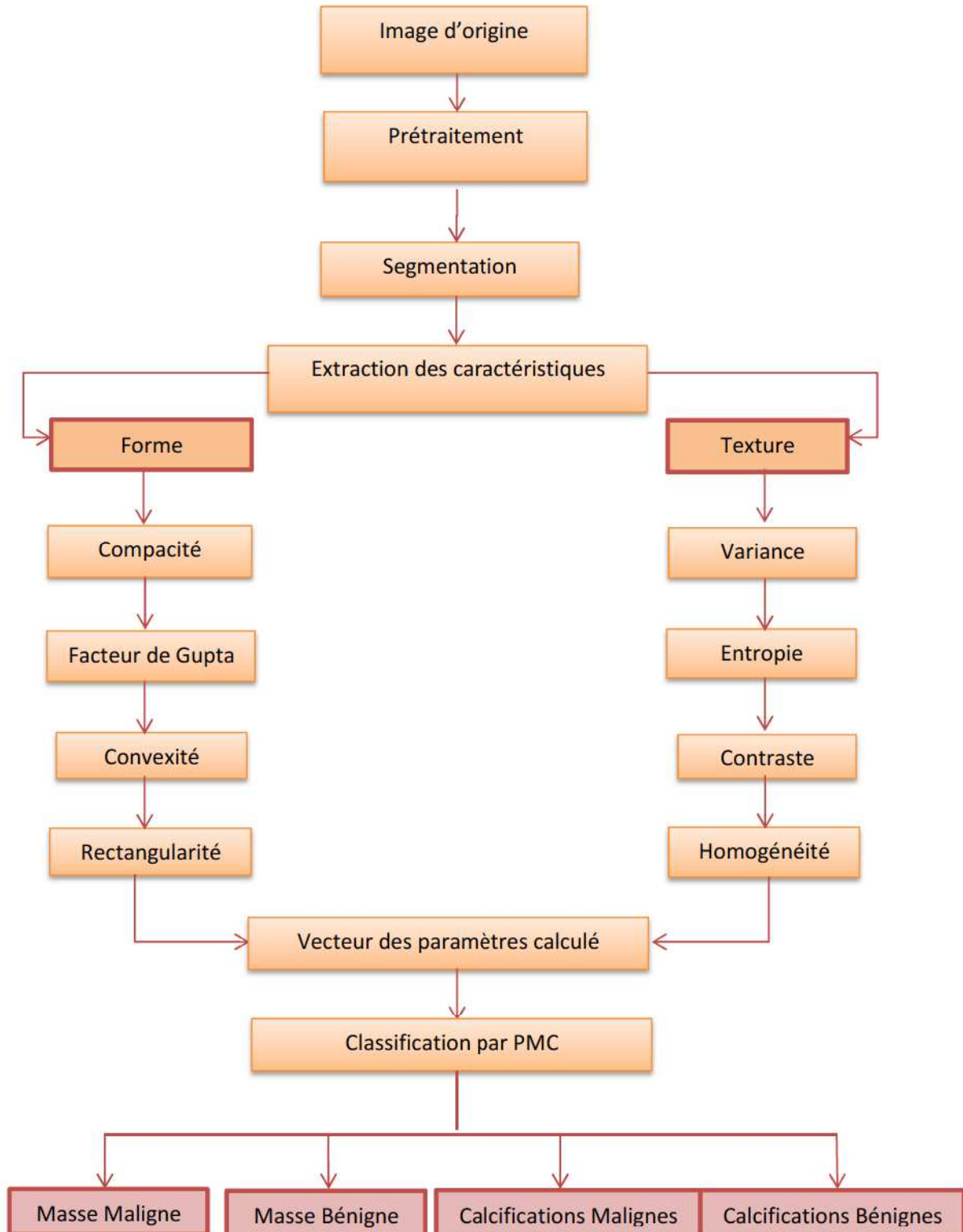


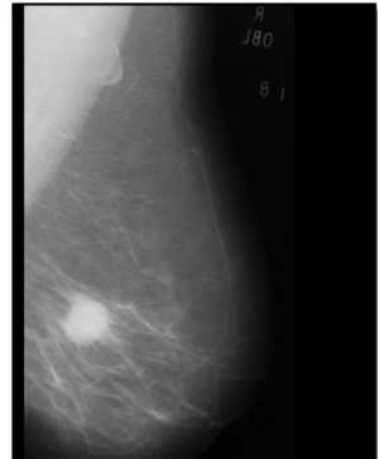
Figure 4.2 : synoptique du Système d'aide à la décision mammographique "Mammoconnex"

4.3.1 Images utilisées

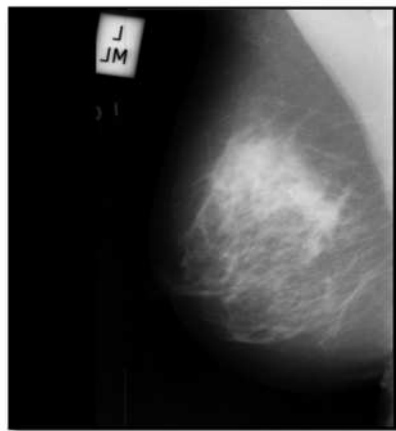
On applique notre système d'aide à la décision sur la base de données MIAS. Cette base des images mammographiques numériques, représente des masses bénignes/Malignes et des calcifications bénignes/Malignes, ainsi que des images des cas sains. La figure 4.3 présente quelques exemples.



(a) mdb021



(b) mdb028



(c) mdb209

Figure 4.3: images de la base MIAS, (a) Masse bénigne, (b) Masse maligne, (c) Calcifications malignes [44].

4.4 Résultats et interprétation

L'application "Mammoconnex", effectue l'analyse texturale et morphologique des différentes pathologies mammaires, en vue de leur classification par le perceptron multicouche.

4.4.1 Prétraitement

a) Cas des masses

Afin de bien visualiser les masses; on suit un prétraitement qui consiste en premier lieu à appliquer une transformation du chapeau haut de forme sur l'image originale avec un disque de 100. Une superposition du résultat de chapeau haut de forme avec l'image d'origine, permet d'améliorer la qualité visuelle de l'image.

Les figures suivantes illustrent des résultats du prétraitement

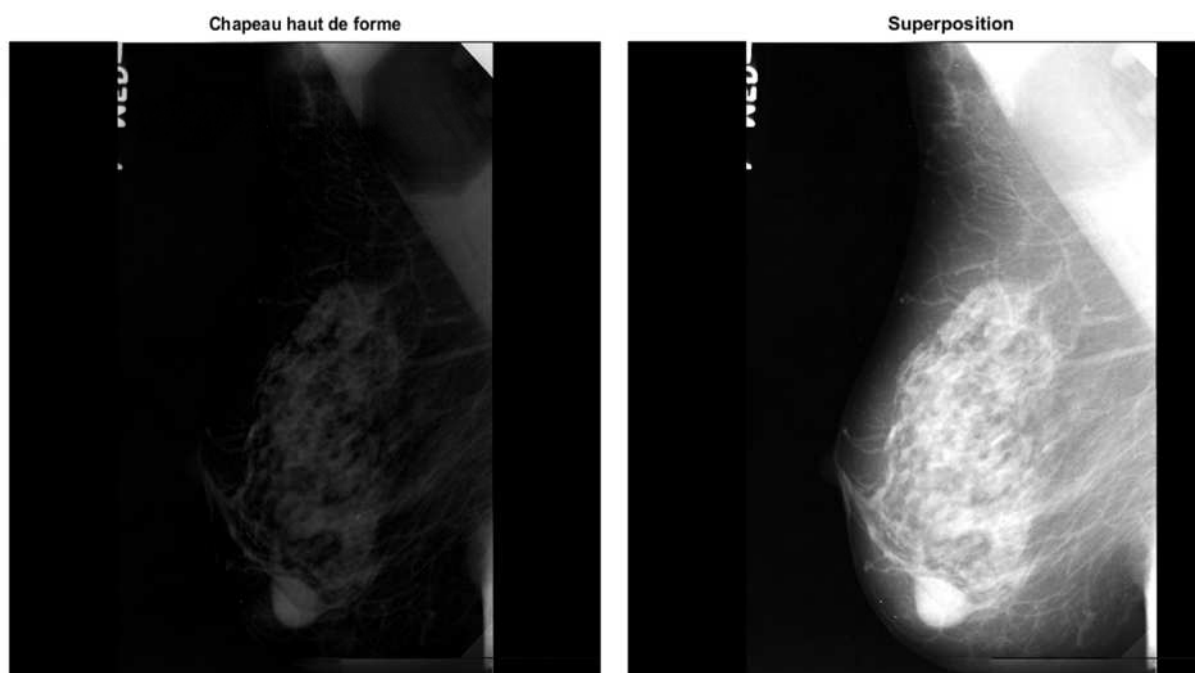


Figure 4.4 : Prétraitement d'une masse bénigne [mdb021], (a) Chapeau haut de forme, (b) Superposition

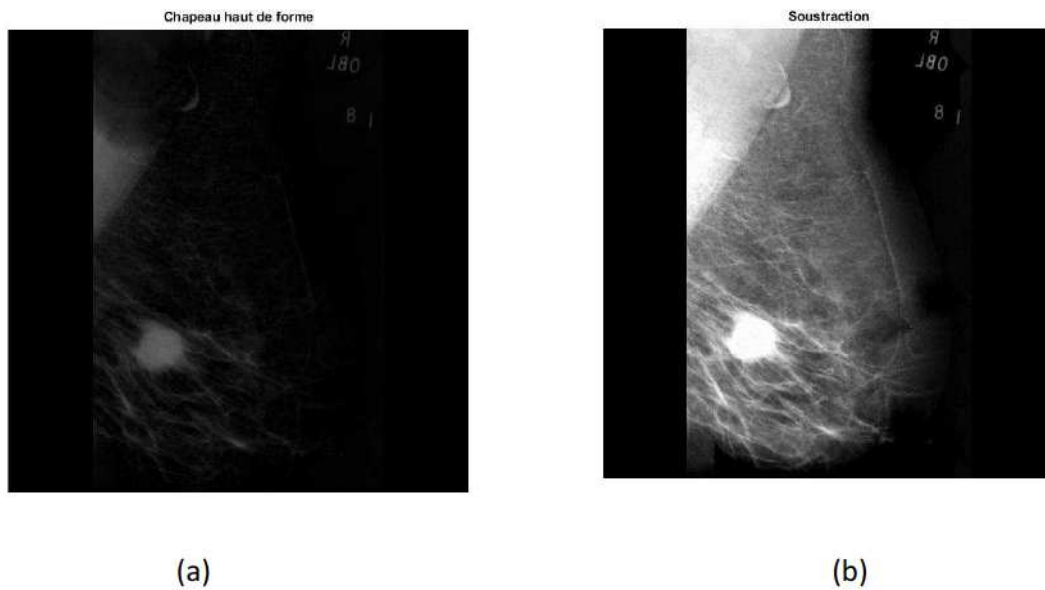


Figure 4.5 : Prétraitement d'une masse maligne [mdb028] (a) Chapeau haut de forme (disque de 100), (b) Superposition

b) Cas des calcifications

Le prétraitement concernant les calcifications, consiste à sélectionner la zone d'intérêt ; puis à effectuer une transformation du chapeau haut de forme avec un disque de taille variant entre 5 et 7. On superpose, ce résultat sur l'image originale, pour bien visualiser les calcifications (figures 4.6).

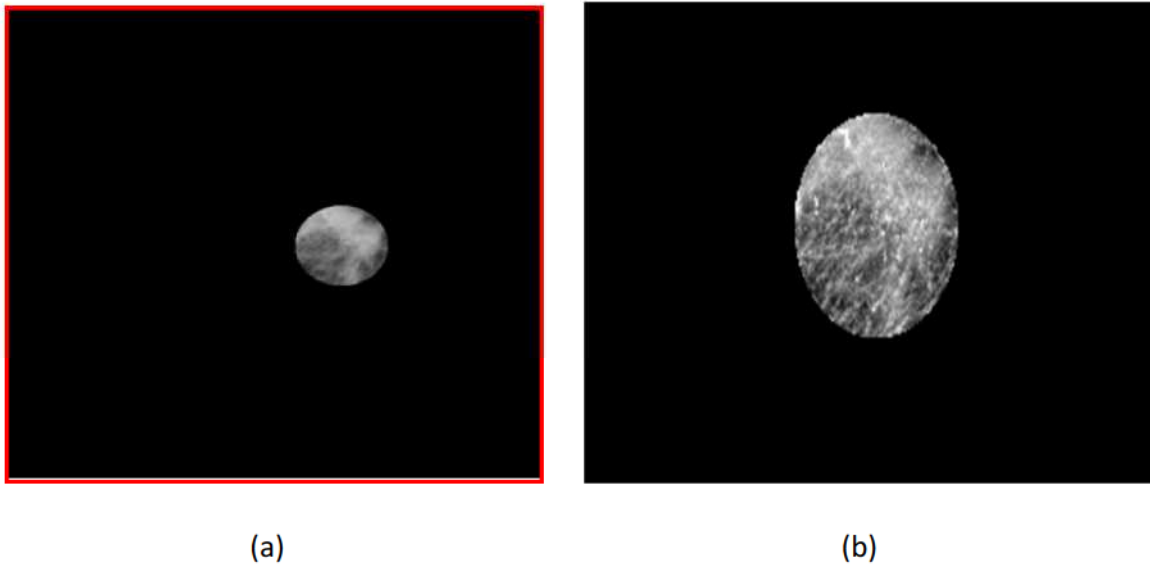


Figure 4.6 : Prétraitement des calcifications malignes [mdb209], (a) Découpage de la zone d'intérêt
(b) Superposition avec Chapeau haut de forme (disque=5)

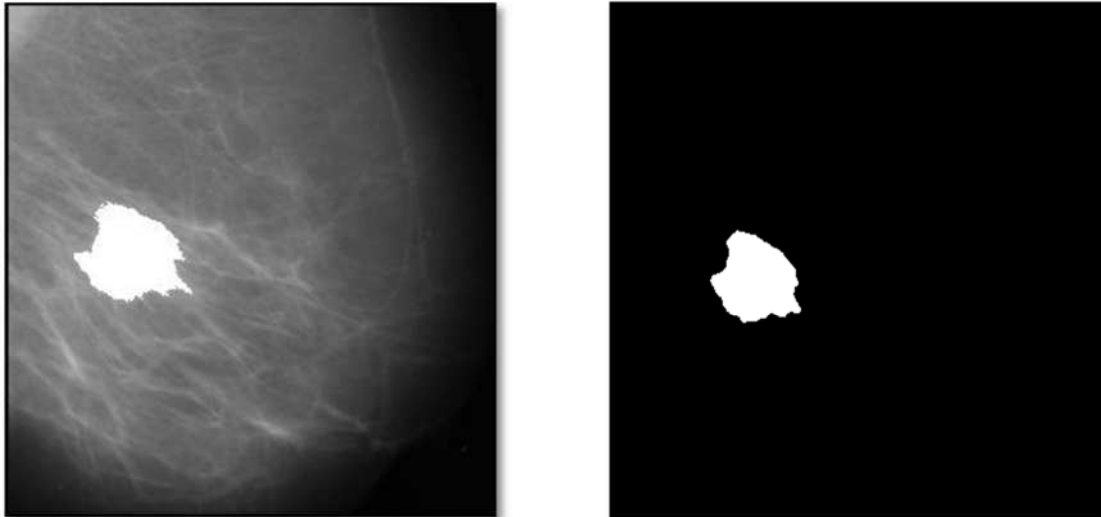
4.4.2 Extraction des régions d'intérêt

Dans notre application, la détection des pathologies mammaires est faite par la segmentation coopérative moyennant, la segmentation en régions et les contours actifs pour les masses. Les calcifications sont extraites, grâce au modèle de Lankton.

a) Cas des masses

Dans le but de délimiter la région d'intérêt, on a choisi la technique de segmentation de régions en coopération avec les contours actifs, ce qui nous donne une bonne délimitation du contour de la masse.

Pour l'étape de la segmentation en croissance de régions, le germe initial est défini par le centre de la tumeur; qui est donné par des médecins radiologues. Le seuil t obtenu expérimentalement, varie entre 10 et 40, selon le type de l'image. Il apparaît des cas où, le seuil est fixé en dessous de 10 à cause d'une masse difficilement, visible. Le nombre d'itérations pour le cas des modèle déformables (modèle de Lankton), varie selon le type de masse appréhendée.



(a) Image segmentée [seuil=40] b) Résultat du contour actif [60 itérations]

Figure 4.7 : détection d'une masse maligne « mdb028 »

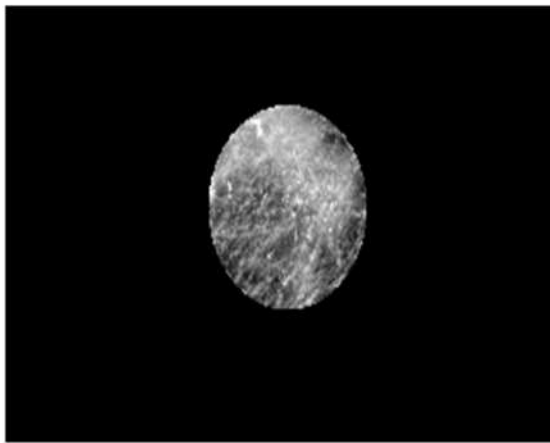


(a) Image segmentée [seuil=40] (b) Résultat du contour actif [10 itérations]

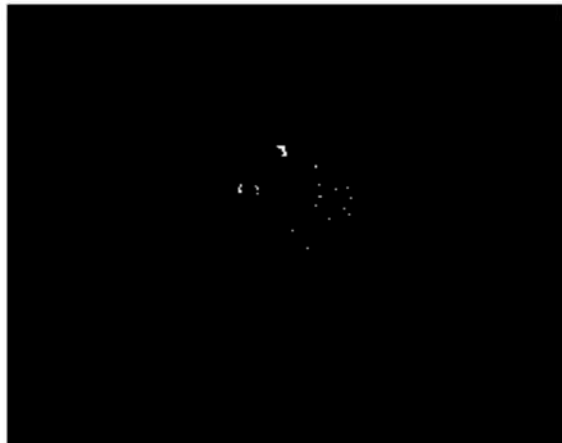
Figure 4.8 : détection d'une masse bénigne « mdb021 »

b) Cas de calcifications

La détection des calcifications, se fait par le modèle de Lankton après un prétraitement de l'image effectué par un chapeau haut de forme (élément structurant autour d'un disque de taille 5) superposé à l'image originale. Les figures suivantes illustrent quelques résultats.



(a) Résultat du prétraitement



(b) Résultat du contour actif [2000 itérations]

Figure 4.9: Détection des calcifications maligne « mdb209 »

4.4.3 Extraction des caractéristiques

La base d'images mammographiques MIAS, a été analysée en explorant des descripteurs texturaux et morphologiques. Pour notre cas nous avons choisi, quatre paramètres texturaux (variance, entropie, contraste, homogénéité) et quatre pour la forme (compacité, facteur de Gupta, convexité, rectangularité).

Les tableaux suivants, représentent l'évaluation des paramètres utilisés pour la caractérisation, des masses et des calcifications.

	Variance	Entropie	Contraste	Homogénéité	Compacité	Facteur de Gupta	Convexité	Rectangularité
mdb001	1.3562e+023	0.1669	7.5605e-004	0.99962	0.23392	0.052443	0.82674	0.73728
mdb010	1.4969e+23	0.013722	0.00014892	0.99993	0.39824	0.085182	0.7646	0.67945
mdb021	1.4727e+23	0.047394	1.6801e-004	0.99992	0.85941	0.030663	0.95864	0.76472
mdb127	1.4931e+023	0.0198	1.3174e-004	0.99993	0.44574	0.070923	0.72717	0.64118
mdb126	1.4991e+023	0.0105	1.1455e-004	0.99994	0.50774	0.062979	0.89336	0.66303
mdb069	1.4856e+23	0.030453	1.6419e-004	0.9999	0.49821	0.064695	0.82409	0.68966
mdb081	1.4363e+23	0.086563	7.1214e-004	0.99964	0.19309	0.066287	0.67763	0.70811
mdb290	1.4813e+023	0.0362	1.7183e-004	0.99991	0.72902	0.027295	0.91369	0.73932
mdb104	1.4817e+23	0.035297	2.1574e-004	0.99989	0.47066	0.054487	0.88813	0.75077
mdb083	1.4911e+23	0.022787	1.2983e-004	0.9999	0.60425	0.046653	0.8883	0.73423
mdb145	1.4974e+023	0.0126	1.7756e-004	0.99991	0.29568	0.06989	0.67046	0.63855
mdb150	1.4542e+023	0.0692	2.6920e-004	0.99987	0.40451	0.056824	0.82395	0.70898
mdb167	1.4979e+023	0.012758	8.4005e-05	0.99996	0.74123	0.037214	0.90235	0.75436
mdb198	1.4524e+023	0.06922	5.6704e-004	0.99972	0.18585	0.10176	0.62082	0.62177
mdb314	1.4923e+023	0.0208	1.5846e-004	0.99992	0.50977	0.038775	0.89953	0.76004

Tableau 4.1 : paramètres des masses bénignes

	Variance	Entropie	Contraste	Homogénéité	Compacité	Facteur de Gupta	Convexité	Rectangularité
mdb023	1.4895e+023	0.022307	1.3746e-004	0.99993	0.68356	0.041135	0.90055	0.75737
mdb028	1.4468e+23	0.077936	2.2911e-004	0.99991	0.70687	0.034912	0.94434	0.76674
mdb058	1.5014e+23	0.0069166	6.8732e-05	0.99997	0.76846	0.033671	0.90673	0.75681
mdb090	1.4996e+23	0.0094637	1.3937e-004	0.99993	0.24439	0.094431	0.05312	0.47108
mdb095	1.4974e+023	0.0131	1.3364e-004	0.9999	0.30258	0.10295	0.74589	0.60345
mdb102	1.487e+23	0.026593	0.00038184	0.99981	0.2052	0.07686	0.6984	0.59578
mdb110	1.4717e+023	0.0481	2.4438e-004	0.99988	0.6257	0.043114	0.94433	0.75311
mdb125	1.4951e+23	0.015715	8.2096e-005	0.99988	0.30655	0.079844	0.61736	0.51941
mdb141	1.4983e+23	0.012215	7.0641e-005	0.99996	0.73847	0.044391	0.91896	0.74434
mdb184	1.4112e+023	0.11532	0.00037611	0.9998	0.62446	0.029137	0.90683	0.76768
mdb186	1.501e+023	0.0062533	0.00019856	0.9999	0.071156	0.073851	0.19235	0.28466
mdb206	1.5024e+023	0.0055	4.5821e-005	0.99998	0.8854	0.024756	0.95772	0.77188
mdb265	1.4931e+023	0.020024	0.00011646	0.99994	0.638	0.062703	0.8167	0.70388
mdb270	1.5047e+023	7.3964e-4	4.5821e-005	0.99998	0.19164	0.149	0.045131	0.40361
mdb271	1.4579e+023	0.0647	2.7684e-004	0.99986	0.57893	0.081526	0.89023	0.76603
mdb158	1.4975e+023	0.0131	1.1837e-004	0.99994	0.57746	0.027576	0.88307	0.74303

Tableau 4.2 : Paramètres des masses Malignes

	Variance	Entropie	Contraste	Homogénéité	Compacité	Facteur de Gupta	Convexité	Rectangularité
mdb209	1.5033e+23	0.0025	1.5465e-04	0.9999	0.1273	0.13719	5.7837e-05	0.5372
mdb211	1.5043e+23	0.0017878	3.6275e-05	0.99998	0.62904	0.0055217	0.0022727	0.69175
mdb213	1.5047e+23	0.0008079	4.0093e-05	0.99998	0.25125	0.013312	0.0016181	0.58837
mdb231	1.505e+23	0.00037111	2.2911e-05	0.99999	0.71918	0.0003622	0.0014006	0.70457
mdb238	1.5049e+23	0.00037111	3.0547e-05	0.99998	0.81506	0.0029308	0.0011561	0.52116
mdb241	1.505e+23	0.00026615	2.482e-05	0.99999	2.4972	0.00052382	0.00061387	0.65854
mdb256	1.505e+23	0.00029651	2.2911e-05	0.99999	1.3945	0.00044791	0.00032457	0.71428

Tableau 4.3 : Paramètres des calcifications Malignes

	Variance	Entropie	Contraste	Homogénéité	Compacité	Facteur de Gupta	Convexité	Rectangularité
mdb218	1.5051e+23	0.00015684	1.3364e-05	0.99999	8.9439	0.002162	0.0063291	0.75
mdb219	1.5045e+23	0.0011013	5.9186e-05	0.99997	0.27003	0.0086415	0.001199	0.59796
mdb222	1.505e+23	0.00029651	2.8638e-05	0.99999	1.9842	0.0024558	0.0011468	0.64311
mdb223(a)	1.5038e+23	0.0020455	0.00010692	0.14636	0.043406	0.043406	0.00078094	0.65729
mdb227	1.5051e+23	0.00021996	1.1455e-05	0.99999	1.0846	0.002551	0.63158	0.63277
mdb240	1.5041e+23	0.0017878	6.1095e-05	0.99997	0.27031	0.40478	0.00073368	0.63068
mdb248	1.5047e+23	0.0009158	3.6275e-05	0.99998	0.32112	0.0014157	0.0042735	0.53262

Tableau 4.4 : Paramètres des calcifications Bénignes

L'extraction de paramètres texturaux et morphologiques, nous permet de constituer des vecteurs caractéristiques de chacune des images, pour une classification connexionniste.

À cet effet, nous montrons, l'influence de l'entropie, le contraste et, la compacité sur les masses et les calcifications, suivant les graphes des figures ci-dessous.

➤ **Cas des masses**

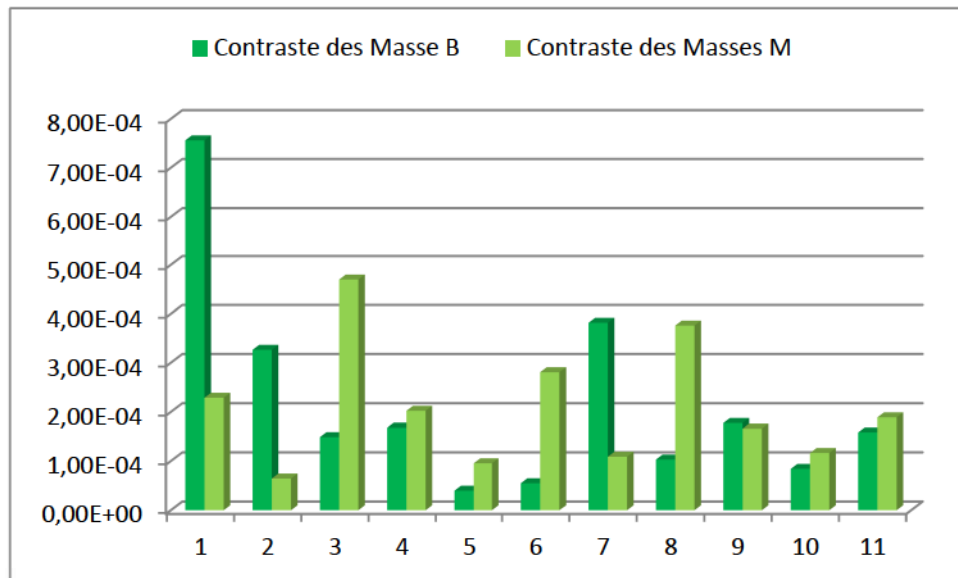


Figure 4.10 : influence du contraste sur les masses.

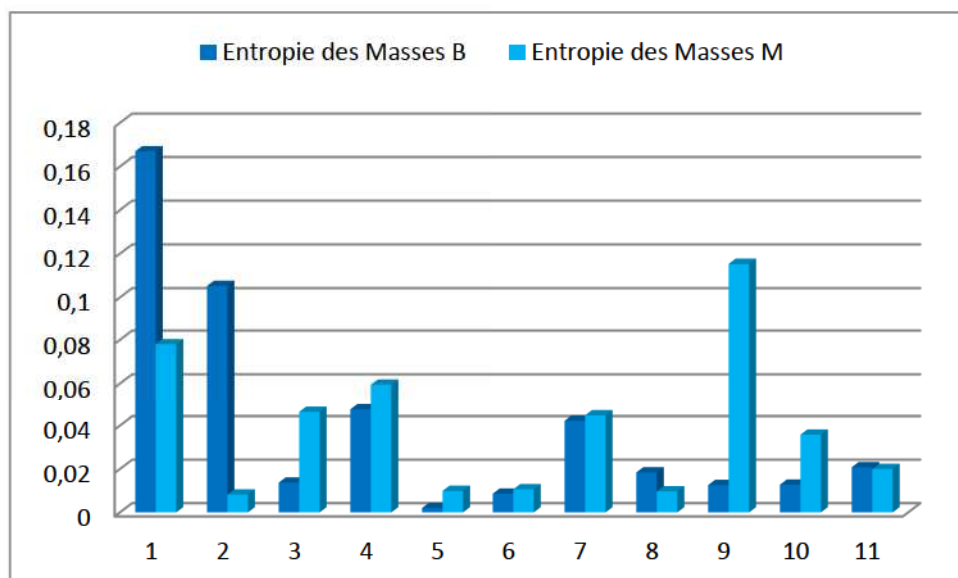


Figure 4.11 : influence de l'entropie sur les masses.

➤ Cas des calcifications

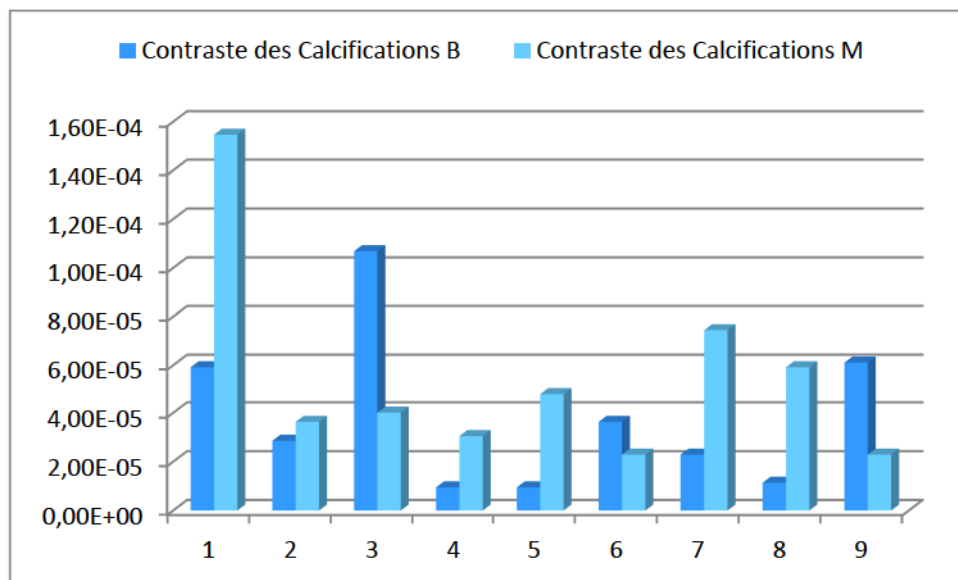
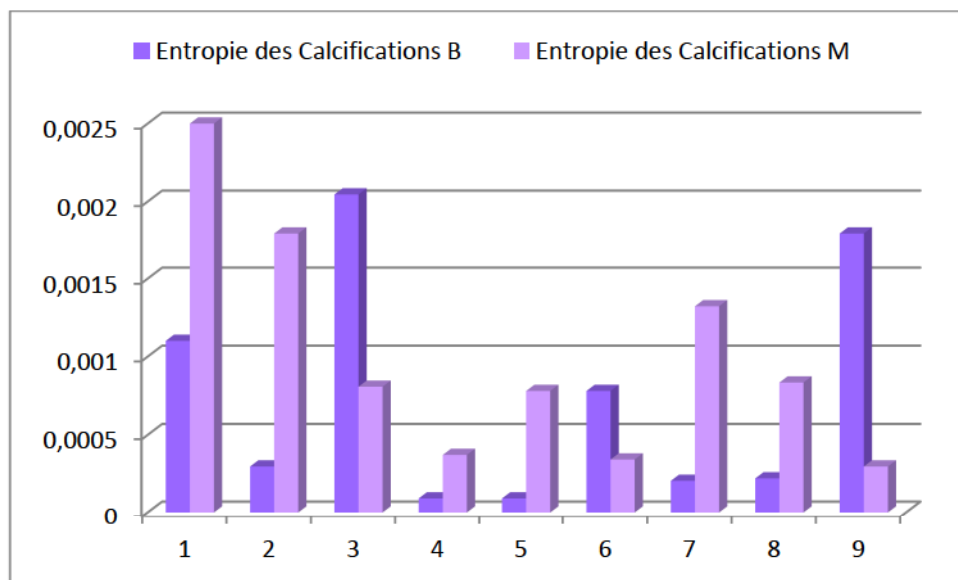


Figure 4.12: influence du contraste sur les calcifications.



Figures 4.13 : influence de l'entropie sur les calcifications.

On remarque que l'entropie et le contraste, ont un impact important sur le type de la tumeur. Lorsque la lésion est maligne, l'entropie (idem pour le contraste) est plus grande que dans le cas bénin.

4.4.4 La classification

Le but de cette partie est de faire une classification, qui nous permet de distinguer entre les masses malignes et les masses bénignes ainsi que les calcifications malignes et bénignes. Cela se fait par le biais du perceptron multicouche.

L'apprentissage (figure 4.14) est l'étape essentielle dans les systèmes de classification. Les résultats corrects (c'est-à-dire les valeurs que l'on désire que le réseau obtienne en sortie) sont fournis au réseau, si bien que celui-ci ajuste ses poids de connexions, pour les obtenir.

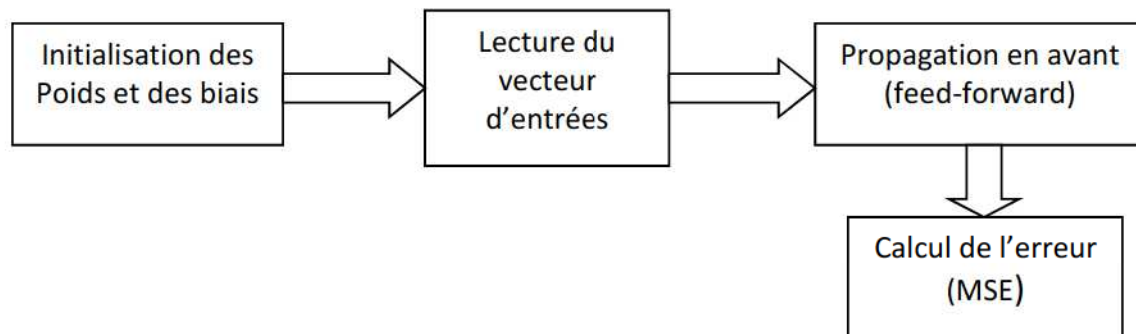


Figure 4.14 : principales étapes de l'apprentissage

L'initialisation des poids se fait d'une manière aléatoire, ils sont initialisés aléatoirement mais dans un intervalle prédéfini, les valeurs des poids et des biais sont comprises entre $[-1, +1]$. Cette technique est utilisée afin de permettre une convergence rapide au réseau.

Après l'apprentissage, le réseau est testé en lui donnant seulement les valeurs d'entrée mais pas les sorties désirées, et en regardant si le résultat obtenu est proche du résultat désiré.

On considère, les paramètres extraits comme une couche d'entrée, une couche cachée, et les classes étudiées comme la couche de sortie.

L'apprentissage a été effectué avec 75% de la base (200 images), le test avec 10% et 15% pour la validation. On a varié le nombre de neurones de la couche cachée, entre 8 et 30, mais les résultats, intéressants ont été obtenus, avec le nombre 25.

Après le lancement de l'entraînement, avec un nombre d'itérations égales à 241, nous avons obtenu ces résultats: une performance égale à 0.1408.

Le graphe ci-dessous représente les performances du réseau de neurone. La courbe en bleu représente l'apprentissage, celle en rouge, pour le résultat du test, le vert pour la validation.

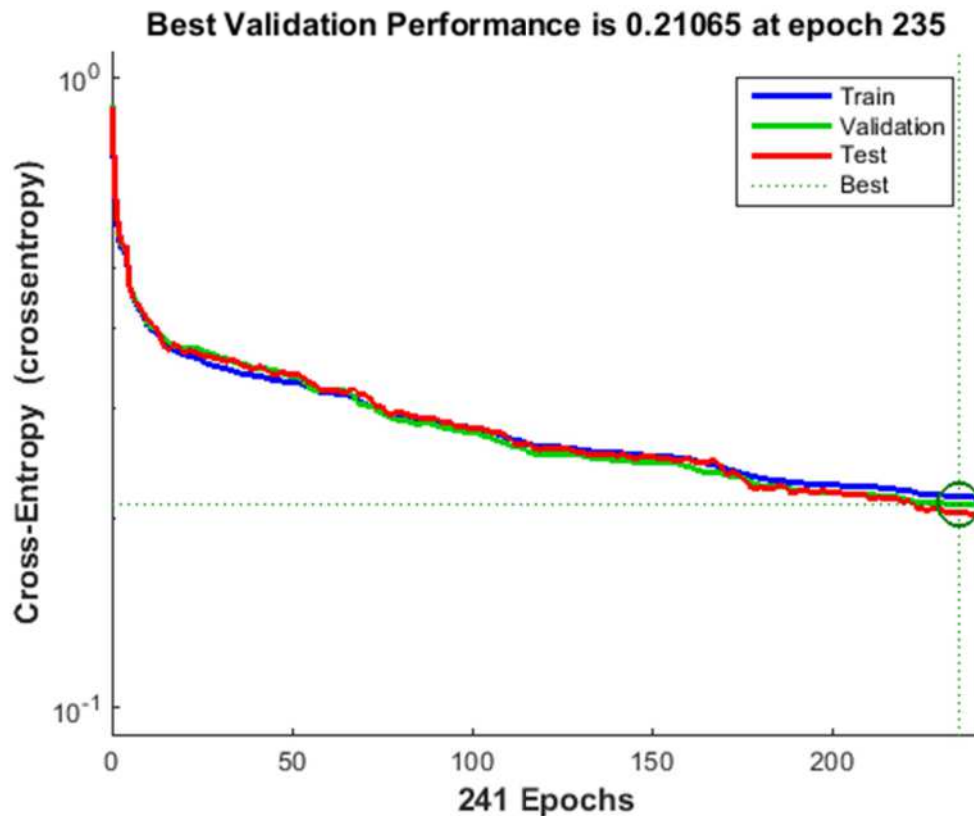


Figure 4.15 : Performances du réseau.

4.4. 5 Evaluation des résultats

Dans cette étape nous faisons le test pour chaque classe. Nous avons évalué le taux de reconnaissance, suivant la matrice de confusion, qui est un moyen d'évaluer l'efficacité du classifieur utilisé. Ce taux approche les 92% de reconnaissance des pathologies étudiées.

	MM	MB	CM	CB	Taux total
MM	80%	10%	0%	0%	80%
MB	0%	94%	0%	0%	94%
CM	0%	0%	94%	6%	94%
CB	0%	0%	0%	100%	100%
Taux total	100%	100%	100%	100%	92%

Tableau 4.5 : matrice de confusion pour la catégorie des tests

(MM : masse maligne, MB : masse bénigne, CM : calcifications malignes, CB : calcifications bénignes)

4.5 Conclusion

La détection des tumeurs mammaires, est toujours difficile à cause de la variété des types de cancers (masses, calcifications) et la variabilité des tissus. Le système Mammoconnex, offre une classification neuronale intéressante des masses et des calcifications. Ce qui permet d'établir une interprétation sur les différents résultats obtenus. Dans le même cadre, la segmentation, coopérative entre la croissance de régions et le modèle de Lankton, détecte les masses, suivant les paramètres adéquats choisis expérimentalement. Nous arrivons aussi, à mieux distinguer les différentes calcifications dans les images traitées, grâce à l'application du modèle de Lankton sur l'image prétraitée, avec un nombre important d'itérations. La classification par le perceptron multicouche, a bien différencié les masses et les calcifications, après avoir effectué un entraînement des données avec différents de nombres de neurones dans la couche cachée.

Conclusion générale

Dans ce mémoire, nous nous sommes intéressés à l'étude des systèmes de détection de pathologies mammographiques, pour une aide à la décision.

En analyse d'images médicales, la segmentation est sans aucun doute la tâche la plus importante. Elle nécessite des connaissances a priori sur le type et la qualité d'images à traiter et, la structure à segmenter. Nous avons fait une étude des différents types de segmentation, pour opter vers une méthode de coopération entre la croissance en régions et un modèle du level set, afin de délimiter les masses d'une façon concise. Dans le même cadre, l'extraction des calcifications, est faite par le modèle de Lankton, pris en compte par un nombre d'itérations adéquat.

En premier lieu, nous avons appliqué un prétraitement sur les images originales pour avoir un rehaussement du contraste afin d'améliorer les résultats ultérieurs de la segmentation des lésions mammaires. La prochaine étape a été consacrée à la caractérisation des masses et des calcifications, avec un choix pertinent des paramètres de texture et de forme, en vue d'une classification pour une prise de décision. A la fin de ce mémoire, on se trouve à la partie de classification des anomalies mammaires, suivant le modèle du perceptron multicouche.

Notre classifieur a donné des performances intéressantes, puisqu'il a pu classer les quatre types de classes (masse maligne, masse bénigne, calcifications malignes et calcifications bénignes) avec un taux de reconnaissance de 92%.

Notre système d'aide à la décision, engendre, un taux de reconnaissance intéressant, mais des améliorations sont nécessaires pour l'enrichir. D'autres techniques peuvent être étudiées, dans le cadre de la segmentation. Nous pouvons aussi extraire d'autres types de caractéristiques texturales et morphologiques de l'image, en vue de caractériser davantage de pathologies. L'entreprise d'une étude de classification par un autre type de réseaux de neurones d'une part et par, les modèles Bayésiens ou les supports à vecteur machine d'autre part, serait envisageable pour une comparaison des classifieurs. Une base de données, plus importante améliorerait le taux de reconnaissance du système.

Notre système développé constitue un outil d'aide à la décision, qui peut donner aux médecins radiologues une seconde opinion, dans la détection des anomalies mammaires.

Bibliographie

- [1] A.Bendib , N . Benzidane H.Guendouz , A.Abdelouahab, Dépistage du cancer du sein, Journée scientifique sur le cancer du sein, association d'aide aux malades atteints de cancer, El-Badr, CHU, F.Fanon, Blida, Février 2010.
- [2] Eureka Santé, symptômes et évolution du cancer du sein, http://www.eurekasante.fr/maladie/cancers/cancer_des_seins, novembre 2011, date d'accès, septembre 2014.
- [3] les seins et les glandes mammaires (notions avancées), [http://bv.alloprof.qc.ca/science-et-technologie/l'univers-vivant/le-maintien-de-la-vie/les-tissus,-les-organes-et-les-systemes-biologiques/les-systemes-biologiques/le-systeme-reproducteur-feminin/l'anatomie-du-systeme-reproducteur-de-la-femme/les-seins-et-les-glandes-mammaires-\(notions-avancees\).aspx](http://bv.alloprof.qc.ca/science-et-technologie/l'univers-vivant/le-maintien-de-la-vie/les-tissus,-les-organes-et-les-systemes-biologiques/les-systemes-biologiques/le-systeme-reproducteur-feminin/l'anatomie-du-systeme-reproducteur-de-la-femme/les-seins-et-les-glandes-mammaires-(notions-avancees).aspx), 1996, date d'accès, septembre 2014.
- [4] Anatomie et physiologie du sein, Société canadienne du cancer, <http://www.cancer.ca/fr-ca/cancer-information/cancer-type/breast/anatomy-and-physiology/?region=qc>, date d'accès, septembre 2014.
- [5] Cancer du sein, institut National du Cancer, France, <http://www.e-cancer.fr/cancerinfo/les-cancers/cancer-du-sein/le-sein>, 2010, date d'accès, septembre 2014.
- [6] Comprendre le cancer du sein, Centre de Sénologie des Coteaux, <http://www.centre-des-coteaux.be/site.php?r=comprendre>, 2012, date d'accès, septembre 2014.

- [7] N. Evrard, la mammographie, <http://sante-az.aufeminin.com/w/sante/s671/maladies/mammographie.html>, octobre 2013, date d'accès, novembre 2014.
- [8] J.F. HERON, les calcifications mammaires cancérologie générale, Faculté de Médecine de Caen France, http://www.oncoprof.net/Generale2000/g04_Diagnostic/Mammographie/g04_mm10.html, décembre 2011, date d'accès, novembre 2014.
- [9] P. Kestener, Analyse multifractale 2D et 3D à l'aide de la transformation en ondelettes: application en mammographie et en turbulence développée, Thèse de doctorat, Université de Bordeaux, 2003
- [10] Classification en six catégories des images mammographiques en fonction du degré de suspicion de leur caractère pathologique, correspondance avec le système BIRADS de l'*American College of Radiology*(ACR), ANAES / Service des recommandations et références professionnelles / Février 2002.
- [11] J.P. Cocquerez, S. Philipp ; Analyse d'images- Filtrage et segmentation, Masson, Paris, 1995.
- [12] A. Rick, Représentation de La Variabilité dans Le Traitement d'Images Flou, Thèse de doctorat, université Paris VI, 1999.
- [13] G. Kom , A. Tiedeu , M. Kom, C. Nguemgne, J. Gonsu , Détection automatique des opacités dans les mammographies par la méthode de minimisation de la somme de l'inertie, ITBM-RBM, Elsevier, pp 347–356, 2005.
- [14] O. Abdelli, Segmentation d'image par seuillage d'histogramme bidimensionnel, mémoire de Magister, Université de Tizi Ouzou, 2011
- [15] S. Beucher, Segmentation d'images et morphologie mathématique, thèse de doctorat, Ecole Nat. Sup. des Mines de Paris, 1990.
- [16] M.Coster et J.L.Cherman, « Traitement d'image par morphologie mathématique application au domaine médical et industriel », Thèse de doctorat, université B. Pascal, 1987.

- [17] ROSE, Croissance de régions variationnelles et contraintes géométriques tridimensionnelles pour la segmentation d'image, Thèse de doctorat, Institut National des Sciences Appliquées de Lyon, 2008
- [18] C.-H. Wei*, S.Y. Chen, and X. Liu, "Mammogram retrieval on similar mass lesions," *Journal of Computer Methods and Programs in Biomedicine*, vol. 106, no. 3, pp. 234-248, 2012.
- [19]: 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.
- [20]: L.Shen, R. M. Rangayyan, J.E L.Desautels, Detection and classification of mammographic calcifications, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol 7, N° 6, pp 1403-1415, June 1993.
- [21]: H. Cheng, X. Shi, R. Min, L. Hu, X. Cai and H. Du: Approaches for automated detection and classification of masses in mammograms, Elsevier, vol. 39, no.4, pp. 646-668, 2006.
- [22]: J. J. Rousselle, les contours actifs, méthode de segmentation: application à l'imagerie médicale, thèse de Doctorat, Université F. Rabelais de Tours, juillet 2003
- [23]: J.M. Rendon Mancha, " régions actives morphologiques: application à la vision par ordinateur ", thèse de Doctorat, Université R. Descartes, Paris V, 2002.
- [24] 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.
- [25]T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image. Process.* vol. 10, no. 2, pp. 266–277, Feb. 2001.
- [26] R. Crandall, Image Segmentation Using the Chan-Vese Algorithm, ECE532 project, http://math.arizona.edu/~rcrandall/ECE532_ProjectPaper.pdf, 2009, date d'accès, janvier 2015.
- [27] S.Lankton, A. Tannenbaum, Localizing Region-Based Active Contours, *IEEE Trans Image Process*, 17(11): 2029–2039, 2009.

- [28] M. Meliani, segmentation d'Image par coopération régions-contours, mémoire de Magister en informatique, Ecole Doctorale STIC, Alger, 2010.
- [29] F. Boudierba, A. Y. Brahim, analyse topologique d'images mammographiques, mémoire d'ingénieur, Département d'électronique, Université S. Dahlab de Blida, Septembre 2009.
- [30] K. F. Cissé, caractérisation texturale des images mammographiques, mathématique, mémoire de Master, Département d'électronique, Université S. Dahlab de Blida, Juillet 2011.
- [31] S. Haoulia, Détection morphologique et texturale de pathologies mammographiques pour l'aide à la décision au diagnostic, mémoire de Master, Département d'électronique, Université de Blida 1, Juillet 2013.
- [32] N. Rendja, H. Belaïd, détection morphométrique d'opacités et de microcalcifications mammographiques pour l'aide à la décision, mémoire de Master, Département d'électronique, Université de Blida 1, Juin 2014.
- [33] A. S. Aïdi, Analyse connexionniste de pathologies mammographiques pour l'aide à la décision, mémoire de Master, Département d'électronique, Université de Blida 1, Juin 2014.
- [34] M. Djebbari, A. Kerkache, Détection et classification connexionniste de structures pathologiques mammographiques, mémoire de Master, Département d'électronique, Université de Blida 1, Juillet 2013.
- [35] N. Zaarour, Détection morphométrique de pathologies tumorales pour l'aide à la décision, mémoire de Master, Département d'électronique, Université de Blida 1, septembre 2013.
- [36] S. Herlidou, caractérisation tissulaire en IRM par l'analyse de texture : étude du tissu musculaire et de tumeurs intracrâniennes, Thèse de doctorat, Université de Rennes, 1999
- [37] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification", *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. SMC-3, No.6, November 1973, pp.610-621.

- [38] A. M. Khuzi, R. B. Beng, Z. W. Wan and N. A. Mengsc: Identification of masses in digital mammogram using gray level co-occurrence matrices, Biomedical Imaging and Intervention Journal, vol. 5, no. 3, pp. e17, 2009.
- [39] : K.Hassan, Exploitation des grandes bases de données d'image sur la grille de calcul, Mémoire de DEA, Université Lumière Lyon 2, France, Juillet 2003.
- [40]: F.Z.Regueig, N.Benblidia, M.Guerti, Morphometric appreciation of mammary breast calcifications around multiresolution analysis by inertia and frequential factors, ACIT2005, Amman, Jordanie, Décembre 2005 .
- [41] I. Cheikhrouhou, Description et classification des masses mammaires pour le diagnostic du cancer du sein, thèse de doctorat, Université d'Evry-Val d'Essonne, 2012.
- [42] C. Bouveyron, modélisation et classification des données de grande dimension application à l'analyse d'images, thèse de Doctorat, université J. Fourier, 2006
- [43] T.J. Haykin, Neural network, a comprehensive foundation, Prentice-Hall, 1994.
- [44] G. Dreyfus, M. Samuelides, J.-M. Martinez, M. B. Gordon, F. Badran, S. Thiria, L. Hérault, réseaux de neurones (méthodologies et applications), Eyrolles, 2004 .
- [45] Medical Image/Video Engineering Group, MIAS: Mammographic Image Analysis Society, <http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html>.
- [46]: R. C. Gonzalez, Digital Image processing using Matlab, Pearson publication, 2005.