

Ecole Nationale Polytechnique



Département d'Electronique

Laboratoire Signal et Communications

## Thèse de Doctorat en Sciences Spécialité : Electronique

Thème

# Analyse Morphologique et Classification Neuro- Génétique d'Anormalités Mammographiques pour l'Aide à la Décision Médicale

Présentée par :

Fatma Zohra REGUIEG

Sous la direction de Mesdames Mhania GUERTI et Nadjia BENBLIDIA Professeurs

Soutenue publiquement le 5 mai 2019

### Composition du Jury :

Président	Cherif LARBES	Professeur	ENP
Directrice de thèse	Mhania GUERTI	Professeur	ENP
Co-Directrice de thèse	Nadjia BENBLIDIA	Professeur	USD-Blida 1
Examineurs	Latifa HAMAMI	Professeur	ENP
	Rachida TOUHAMI	Professeur	ENP
	Abderrezak GUESSOUM	Professeur	USD-Blida 1
	Salim CHITROUB	Professeur	USTHB

**ENP 2019**



Ecole Nationale Polytechnique



Département d'Electronique

Laboratoire Signal et Communications

## Thèse de Doctorat en Sciences Spécialité : Electronique

Thème

# Analyse Morphologique et Classification Neuro- Génétique d'Anormalités Mammographiques pour l'Aide à la Décision Médicale

Présentée par :

Fatma Zohra REGUIEG

Sous la direction de Mesdames Mhania GUERTI et Nadjia BENBLIDIA Professeurs

Soutenue publiquement le 5 mai 2019

### Composition du Jury :

Président	Cherif LARBES	Professeur	ENP
Directrice de thèse	Mhania GUERTI	Professeur	ENP
Co-Directrice de thèse	Nadjia BENBLIDIA	Professeur	USD-Blida 1
Examineurs	Latifa HAMAMI	Professeur	ENP
	Rachida TOUHAMI	Professeur	ENP
	Abderrezak GUESSOUM	Professeur	USD-Blida 1
	Salim CHITROUB	Professeur	USTHB

**ENP 2019**

*“Science sans conscience, n’est que ruine de l’âme” Rabelais*

*‘ La patience est l'art d'espérer. ‘ Luc de Vauvenargues*

*A ceux que j'aime*

# Remerciements

---

Je ne saurais présenter ce prélude à la recherche, sans rendre hommage à qui de droit. Remercier est nécessaire à tout travail qui réussit à voir le jour. Une page de ma vie est en train de se tourner, me laissant face à un avenir prometteur.

Tous ceux que je remercie sincèrement, le savent bien et ceci sans avoir à préciser le moindre nom ou prénom. Mais, il vous serait difficile à vous lecteurs de ce manuscrit, de le savoir, aussi je vais vous y aider, en vous confiant quelques indices.

Je citerai en tout premier, mes doyennes dans la recherche scientifique ; Mme Mhania Guerti, professeur à l'Ecole Nationale Polytechnique d'Alger ainsi que Mme Nadja Benblidia, professeur à l'Université Sâad Dahlab de Blida 1, pour m'avoir orientée vers le thème de l'imagerie médicale et m'avoir fait l'honneur de diriger ce modeste mémoire, avec une confiance absolue. Qu'elles soient dignement et profondément, remerciées.

Cette intronisation dans le monde de l'imagerie médicale, n'aurait pu aboutir sans leur précieuse aide. Pour m'avoir communiqué la passion de la recherche et initiée à la rigueur scientifique, tout en m'offrant leur amitié, je leur rends grâce. Leurs conseils et leurs compétences dans l'ingénierie biomédicale, ont été la source de mon inspiration. Qu'elles trouvent ici, le témoignage de ma reconnaissance et de mon altruisme, gravés à jamais dans mon âme et mon esprit.

Je suis très sensible à l'honneur que M. Cherif Larbes, professeur à l'Ecole Nationale Polytechnique d'Alger me fait, en acceptant de présider le jury de cette thèse.

Je remercie profondément Mme Latifa Hamami, professeur à l'Ecole Nationale Polytechnique d'Alger, pour l'intérêt qu'elle prête à l'examen de cette réflexion doctorale.

Je témoigne ma gratitude à Mme Rachida Touhami, professeur à l'Ecole Nationale Polytechnique d'Alger, pour le temps qu'elle consacre, à la lecture constructive de cette thèse.

J'adresse ma reconnaissance à M. Abderrezak Guessoum, professeur à l'Université de Blida 1, pour la considération qu'il m'octroie, en jugeant ce manuscrit.

Je tiens à présenter mes remerciements à M. Salim Chitroub, professeur à l'université des sciences et de la technologie Houari Boumediene, pour l'égard qu'il témoigne à l'investigation de cette modeste recherche.

Je suis reconnaissante à Mme Isabelle Bloch, professeur à l'école des télécommunications de Paris et M. Frédéric Richard, professeur à l'université d'Aix-Marseille, de m'avoir offert généreusement, l'accès aux bibliothèques numériques de leurs institutions.

Je voue mes profonds remerciements à M. Rangaraj Rangayyan, professeur à l'université de Calgary au Canada, pour la précieuse documentation, fournie avec bienveillance.

Je tiens à exprimer ma gratitude, à M. Mohamed Lakhel, le radiologiste, fondateur de la clinique d'imagerie Lakhel de Blida, qui m'a initiée à l'aspect pathologique, de la sénologie mammaire. Ses conseils ont été très fructueux, quant à l'aspect technique, de ce travail.

J'exprime ma gratitude, à M. Mohamed Aïdja, chargé de cours à l'université de Blida 1, pour avoir cru en mes ambitions.

Je ne peux clôturer cette page, sans adresser mes remerciements, à mes collègues du département d'électronique de l'université de Blida 1, pour leur soutien moral. Enfin merci à ceux dont le nom m'est parfois inconnu et, qui ont contribué efficacement par leur sourire et leur gentillesse, à rendre mes journées plus sereines et à vous, qui êtes en train de lire ce début de manuscrit.

## ملخص

الزيادة المخفية في الإصابة بسرطان الثدي في العالم، ولا سيما في الجزائر، يثير اهتمامنا لتعزيز دعم القرار السريري. هذه الأطروحة تهدف إلى الكشف الشبه التلقائي للأورام الثدي، المأخوذة من التصوير الشعاعي للثدي الرقمي لقاعدة المعطيات MIAS.

نهجنا التقني مستوحاة من الرؤية السريرية، موضوعها يدل على تحليل دقيق للعلامات المرضية والصحية للثدي، لتصنيف موجه نحو نموذجين. يستند النموذج الأول على تمييز الاجسام والتكلسات المستجابة للعلامات الحميدة مقابل الخبيثة، حول الشبكات العصبية الوراثة. يعتمد التفكير الثاني على التمييز بين الأنسجة السليمة والمریضة، عن طريق الفاصل الهامش على نطاق واسع.

يخضع نظام التشخيص الإشعاعي الموقوف إلى ثلاث مراحل رئيسية، قبل اتخاذ القرار. الصور الشعاعية للثدي مطروحة أولا لتحسنها من قبل العمليات المورفولوجية، لتسليط الضوء على العلامات السريرية.

الخطوة الرئيسية للدراسة، هي استخراج المناطق ذات الاهتمام. تخضع هذه العملية لعدة تقنيات لنقاط قوتها. في حالة المنهج لشبكات العصبية، يتم تحديد الأفات وفقا لمجموعات المستويات الكيفية، بحيث التهيئة مقيدة بالمنطقة المشروحة.

في سياق التعلم الإحصائي ل "فابنيك"، يتم الحصول على ترسيم دقيق للهياكل الأنسجة الطبيعية وغير الطبيعية للثدي عن طريق تجميع فضائي غامض بالتعاون مع خوارزمية تشان وفيس، للنهج المتطور الكيفي. يقدر تعقيد علامات الثدي السريرية والصحية، من خلال دمج السمات المورفومترية والنسجية لتوليد توقعين. يتم دمج هذه المعالم القياسية في مسار النظامية النهائي، عند مدخل اثنين المصنفات. البيروبيرون متعدد الطبقات، بحيث يتم تحسين العوامل الاتصالية بواسطة الخوارزميات الجينية، يصنف التشوهات الخبيثة والحميدة. يسمح المنهج الثاني، بوضع العلامات على الأنسجة الصحية والمریضة، وفقا لفاصل الهامش على نطاق واسع. يتم التحقق من النهج المعتمد، باستخدام مجموعة كبيرة من الصور الماموغرافية. يوفر في هذا الإطار، النظام المصمم، دقة موافقه الي 99.25٪ وحساسية بنسبة 100٪. تشهد النتائج المحققة على الإمكانيات القوية لنظام اتخاذ القرارات المقترح، للمساعدة في التشخيص السريري.

**كلمات المفاتيح:** التصوير الشعاعي للثدي الرقمي، النهج المتطور الكيفي، التجميع الفضائي الغامض، النسيج والمورفوميترية، التصنيف العصبي الجيني، الفاصل الهامش الواسع، المساعدة التشخيصية السريرية.

## Abstract

Anxious increase of breast cancer's incidence in the world and particularly in Algeria, arouses our interest, to promote clinical decision support. This thesis is destined for the semi-automatic detection of breast neoplasm taken, from digital mammograms of MIAS database. Our technical approach inspired by a clinical vision, is accomplished by a careful analysis of pathological and healthy mammary signs, for a supervised classification around two models. The first model is based on the neuro-genetic discrimination of masses and calcifications, corresponding to benign versus malignant labels. The second reflexion focuses on the distinction of healthy tissues and abnormalities, by support vector machine. The apprehended radio-diagnostic system browses three key stages, before the decision. The breast's radiographic images are first, enhanced by morphological operators, to highlight the clinical signs. The angular step of the study, is the extraction of regions of interest. It is subjected to several techniques, for their strengths. Therefore, in the case of the connectionist approach, the lesions are localized by an adaptive deformable model, which initialization is guided by, the annotated suspicious zone. In the context of Vapnik's statistical learning, the careful delimitation of normal and abnormal mammary structures, is achieved by a fuzzy spatial clustering and refined, by the Chan and Vese algorithm. The mammograms' complexity is estimated by the fusion of morphometric and textural attributes, to generate their signature. This parametrization is integrated into the ultimate systemic description, through two classifiers. The multilayer perceptron, whose synaptic factors are optimized by the genetic algorithms, categorizes malignant and benign abnormalities. The classification by support vector machine, allows the labelling of healthy and pathological tissues. The validation of the elaborated decisional system, is carried out using a significant set of mammographic images. Its performances provide an accuracy of 99.25% and a sensitivity of 100%. The reached results attest of the high potential of the suggested system, for computer aided-diagnosis.

**Keywords:** Digital Mammogram, Adaptive Deformable Model, Spatial Fuzzy Clustering, Texture and Morphometry, Neuro-Genetic Classification, Support Vector Machine, Computer-Aided Diagnosis.

## Résumé

L'augmentation anxieuse de l'incidence du cancer du sein dans le monde et particulièrement en Algérie, suscite notre intérêt pour promouvoir l'aide à la décision clinique. Cette thèse est destinée à une détection semi-automatique de néoplasmes du sein, prélevés à partir de mammogrammes numériques de la base MIAS. Notre approche technique inspirée par une vision clinique, fait l'objet d'une analyse minutieuse des signes pathologiques et sains du sein, pour une classification dirigée autour de deux modèles. Le premier se fonde sur la discrimination neuro-génétique des masses et des calcifications, répondant aux labels bénins versus malins. La seconde réflexion se focalise sur la distinction des tissus sains et des anomalies, par les séparateurs à vaste marge. Le système de radiodiagnostic appréhendé parcourt trois étapes clés, en amont de la décision. Les clichés radiographiques du sein sont en premier lieu, rehaussés par des opérateurs morphologiques, pour mettre en exergue les signes cliniques. L'étape angulaire de l'étude, est l'extraction des régions d'intérêt. Elle est assujettie à plusieurs techniques, pour leurs atouts. Dans le cas de l'approche connexionniste, les lésions sont localisées suivant les modèles déformables adaptatifs, dont l'initialisation est guidée par la zone annotée. Dans le contexte de l'apprentissage de Vapnik, la délimitation minutieuse des structures tissulaires normales et anormales mammaires, est obtenue par un regroupement spatial flou en coopération avec l'algorithme de Chan et Vese. La complexité des mammogrammes est estimée par la fusion d'attributs morphométriques et texturaux, pour générer leur signature. Cette paramétrisation, est intégrée dans l'ultime parcours systémique, à l'entrée de deux classificateurs. Le perceptron multicouche dont, les facteurs synaptiques sont optimisés par les algorithmes génétiques, catégorise les anomalies malignes et bénignes. Les séparateurs à vaste marge permettent la labellisation des tissus sains et pathologiques. La validation de la démarche adoptée, est réalisée en usant d'un ensemble significatif de clichés mammographiques. Ses performances fournissent une précision de 99.25% et, une sensibilité de 100%. Les résultats atteints, attestent du fort potentiel du système décisionnel proposé, pour une aide au diagnostic clinique.

**Mots-clés :** Mammogramme Numérique, Modèle Déformable Adaptatif, Regroupement Spatial Flou, Texture et Morphométrie, Classification Neuro-Génétique, Séparateurs à Vaste Marge, Aide au Diagnostic.

# Table des matières

<b>Introduction générale.....</b>	<b>14</b>
<b>1 Sémiologie du Sein et Contexte Mammographique.....</b>	<b>18</b>
1.1 Introduction.....	18
1.2 Epidémiologie du néoplasme mammaire.....	18
1.3 Anatomie et physiologie du sein.....	19
1.4 Aspect étiologique du sein et cancérogénèse mammaire.....	21
1.4.1 Cancer du sein non invasif.....	22
1.4.2 Cancer du sein invasif.....	22
1.5 Localisation clinique du cancer mammaire.....	23
1.6 Facteurs de risque de néoplasie mammaire.....	23
1.7 Contexte mammographique.....	23
1.7.1 Découverte des rayons X.....	23
1.7.2 Aperçu historique mammographique.....	24
1.8 Techniques d'examen de la mammographie.....	24
1.8.1 Incidences mammographiques.....	26
1.8.2 Aspect radiologique du sein.....	27
1.8.3 Classification BIRADS de la densité mammaire.....	28
1.8.4 Classification BIRADS des anomalies mammographiques.....	28
1.9 Lésions oncologiques mammographiques.....	29
1.9.1 Masses typiques.....	30
1.9.2 Calcifications néoplasiques.....	32
1.9.3 Asymétrie de densité.....	36
1.10 Limites de la mammographie.....	37
1.11 Rôle des systèmes d'aide à la détection des anomalies mammographiques.....	37
1.12 Conclusion.....	37
<b>2 Systèmes d'Aide au Diagnostic en Imagerie Mammographique.....</b>	<b>40</b>
2.1 Introduction.....	40
2.2 Fonctionnalités de l'imagerie médicale.....	40
2.3 Aperçu historique des systèmes CAD.....	42
2.4 Systèmes CAD commercialisés.....	45
2.5 Utilité des CAD.....	46
2.6 Introduction à la détection assistée par ordinateur dans la mammographie.....	47
2.6.1 Prétraitement.....	48
2.6.2 Segmentation.....	49
2.6.3 Caractérisation des signes symboliques.....	50
2.6.4 Classification des entités symboliques.....	51
2.7 Evaluation des systèmes CAD.....	54
2.8 Systèmes d'aide à la détection assistée par ordinateur dans la mammographie.....	55
2.9 Réflexion technico-médicale d'un système d'aide à la décision clinique de mammogrammes	61
2.10 Conclusion.....	61
<b>3 Correction des Mammogrammes et Localisation Morphologique des Néoplasmes du Sein.....</b>	<b>64</b>
3.1 Introduction.....	64
3.2 Correction du contraste de mammogrammes.....	64
3.2.1 Expansion de la dynamique.....	65
3.2.2 Correction morphologique.....	66
3.3 Détection morphologique des lésions suspectes.....	71
3.3.1 Segmentation par croissance de régions.....	72
3.3.2 Segmentation par les modèles déformables.....	75
3.3.3 Principe de la logique floue.....	86
3.3.4 Regroupement spatial flou.....	87
3.4 Approches adoptées pour la détection des néoplasmes et des tissus sains.....	89
3.4.1 Modèle adaptatif de Chan et Vese.....	89



3.4.2	Approche de croissance de régions associée au modèle de Lankton.....	90
3.4.3	Localisation des tissus mammaires par regroupement spatial flou.....	91
3.5	Conclusion.....	91
<b>4</b>	<b>Signature Morphométrique et Texturale des Mammogrammes.....</b>	<b>93</b>
4.1	Introduction.....	93
4.2	Description sémiologique des anormalités mammaires.....	93
4.3	Analyse de la texture.....	95
4.3.1	Types de texture.....	95
4.3.2	Méthodes d'analyse de texture.....	97
4.4	Description texturale.....	105
4.4.1	Attributs texturaux du modèle neuro-génétique.....	106
4.4.2	Attributs texturaux du modèle de Vapnik.....	107
4.5	Analyse morphométrique.....	108
4.5.1	Descripteurs morphologiques du modèle neuro-génétique.....	108
4.5.2	Descripteurs morphologiques du modèle de Vapnik.....	111
4.6	Conclusion.....	113
<b>5</b>	<b>Classification Neuro-Génétique et par l'Approche de Vapnik des Néoplasmes Mammographiques.....</b>	<b>115</b>
5.1	Introduction.....	115
5.2	Intelligence artificielle et Connexionnisme.....	115
5.2.1	Rôle de l'intelligence artificielle.....	115
5.2.2	Neurobiologie computationnelle.....	116
5.2.3	Notion de connexionnisme.....	116
5.2.4	Paradigme d'apprentissage.....	116
5.3	Réseaux connexionnistes.....	117
5.3.1	Genèse du connexionnisme.....	117
5.3.2	Neurone biologique.....	118
5.3.3	Modélisation d'un neurone.....	118
5.3.4	Fonctions d'activation.....	120
5.3.5	Apprentissage des réseaux de neurones.....	121
5.3.6	Connectivité.....	121
5.3.7	Modèle du perceptron.....	122
5.4	Introduction aux algorithmes génétiques.....	124
5.5	Approche neuro-génétique adoptée.....	125
5.6	Paradigme du classifieur de Vapnik.....	127
5.6.1	Principe.....	128
5.6.2	SVM à marge dure.....	129
5.6.3	SVM à marge souple.....	131
5.6.4	Fonctions à noyaux.....	132
5.6.5	Classification multi-classes.....	134
5.6.6	Classification des mammogrammes par les séparateurs à vaste marge.....	135
5.7	Conclusion.....	136
<b>6</b>	<b>Détection des Anormalités et des Tissus Sains dans les Mammogrammes pour l'Aide à la Décision.....</b>	<b>138</b>
6.1	Introduction.....	138
6.2	Système d'aide au diagnostic des anormalités mammographiques conçu.....	138
6.3	Description de la base de mammogrammes MIAS.....	140
	141	
6.4	Approche neuro-génétique pour la détection morphologique de masses et de calcifications.....	142
6.4.1	Correction morphologique de mammogrammes.....	143
6.4.2	Extraction des signes cliniques du sein.....	145
6.4.3	Caractérisation des signes cliniques du sein.....	149
6.4.4	Catégorisation neuro-génétique.....	154
6.5	Discussion.....	155
6.6	Approche statistique de Vapnick pour la détection morphologique de masses et de tissus sains en mammographie numérique.....	158
6.6.1	Amélioration du contraste des mammogrammes.....	159
6.6.2	Extraction des signes cliniques du sein.....	161

6.6.3	Description morphométrique et de la granularité des régions d'intérêt.....	162
6.6.4	Discrimination des tissus sains et des anomalies mammaires par les SVM.....	163
6.6.5	Interprétation .....	164
6.7	Etude comparative avec les systèmes CAD .....	164
6.8	Conclusion.....	165
	<b>Conclusion générale et perspectives .....</b>	<b>166</b>
	<b>Bibliographie .....</b>	<b>168</b>
	<b>Liste des contributions de l'auteur .....</b>	<b>180</b>

## Liste des tableaux

Tableau 1.1 : Incidence du cancer du sein en 2015 et projection future en Algérie.....	19
Tableau 1.2 : Classification BI-RADS de la densité mammaire.....	28
Tableau 1.3 : Classification BI-RADS des anomalies mammographiques .....	29
Tableau 1.4 : Classification radiologique des calcifications du sein selon Le Gal .....	35
Tableau 6.1 : Signature texturale de masses dans vingt mammogrammes .....	151
Tableau 6.2 : Variance et Signature morphométrique des masses dans vingt mammogrammes .....	152
Tableau 6.3 : Signature texturale des calcifications dans vingt mammogrammes .....	153
Tableau 6.4 : Variance et Signature morphométrique des calcifications dans vingt mammogrammes .....	154
Tableau 6.5 : Paramètres de la classification par les SVM.....	163
Tableau 6.6 : Performances de vingt systèmes CAD mammographiques .....	165

## Liste des figures

Figure 1.1 : Incidence du cancer mammaire en 2014 : cas de l'Algérie .....	19
Figure 1.2 : Incidence du cancer mammaire en 2014 : cas de la France .....	19
Figure 1.3 : Structure anatomique du sein : coupe sagittale.....	20
Figure 1.4 : Unité ducto-lobulaire.....	20
Figure 1.5 : Cancérogénèse des cellules mammaires.....	22
Figure 1.6 : Localisation des cancers canaux dans le sein.....	22
Figure 1.7 : Localisation clinique des néoplasmes mammaires .....	23
Figure 1.8 : Exemple d'un mammographe analogique .....	25
Figure 1.9 : Exemples de mammographes numériques.....	26
Figure 1.10 : Exemples de clichés mammographiques .....	27
Figure 1.11 : Différents types de densité mammaire .....	28
Figure 1.12 : Distinctions morphologiques de masses mammaires .....	30
Figure 1.13 : Morphologie des contours d'une masse .....	31
Figure 1.14 : Différents types de calcifications .....	34
Figure 1.15 : Catégorisation morphologique des calcifications du sein selon Le Gal .....	36
Figure 1.16 : Exemple d'une asymétrie de densité .....	36
Figure 2.1 : Fonctions de l'imagerie médicale.....	42
Figure 2.2 : Masse produite par une asymétrie de densité .....	43
Figure 2.3 : Masse annotée dans un mammogramme conventionnel.....	44
Figure 2.4 : Premier prototype CAD circa .....	45
Figure 2.5 : Mammogrammes annotés via le CAD Image Checker .....	46
Figure 2.6 : Mammogrammes rehaussés via le CAD Power Look .....	46
Figure 2.7 : Système de détection assistée par ordinateur de pathologies en mammographie numérisée...47	
Figure 2.8 : Constitution d'un mammogramme .....	48
Figure 2.9 : Exemple d'une courbe ROC suivant différentes discriminations .....	55
Figure 2.10 : Système d'aide à la décision de pathologies en mammographie numérisée.....	61
Figure 3.1 : Image mammographique rehaussée par expansion de la dynamique .....	65
Figure 3.2 : Eléments structurants plans .....	67
Figure 3.3 : Erosion d'une image mammographique .....	68
Figure 3.4 : Dilatation d'une image mammographique .....	68
Figure 3.5 : Ouverture d'une image mammographique .....	69
Figure 3.6 : Fermeture d'une image mammographique .....	70
Figure 3.7 : Transformation du chapeau haut de forme sur une image mammographique .....	71
Figure 3.8 : Synopsis général du processus de la croissance de régions .....	73
Figure 3.9 : Segmentation d'une image synthétique en niveaux de gris .....	74
Figure 3.10 : Application de la croissance de régions sur les images mammographiques.....	75
Figure 3.11 : Modèle graphique du contour actif.....	76
Figure 3.12 : Propagation d'une courbe selon une force normale.....	79
Figure 3.13 : Représentation de $\psi(x, t)$ et de son intersection avec le plan de niveau zéro .....	80
Figure 3.14 : Evolution d'un contour au gré de l'ensemble de niveau .....	81
Figure 3.15 : Détection de calcifications par les ensembles de niveaux .....	83
Figure 4.1 : Evolution de la malignité des masses en fonction de la complexité du contour .....	94
Figure 4.3 : Exemples de textures déterministes.....	96
Figure 4.4 : Exemples de textures directionnelles .....	96
Figure 4.5 : Exemples de textures aléatoires.....	96
Figure 4.6 : Exemple de construction d'un motif binaire .....	99
Figure 4.7 : Information locale et information globale .....	99
Figure 4.8 : Anormalités mammographiques détectées par la transformée de Gabor.....	101
Figure 4.9 : Plus proches voisins de x selon 4 directions.....	104
Figure 4.10 : Exemples de détermination d'une matrice de cooccurrence.....	105
Figure 4.11 : Boîte englobante .....	111
Figure 4.12 : Représentation de l'ellipse équivalente à une anormalité.....	113
Figure 5.1 : Processus de l'IA.....	116
Figure 5.2 : Neurone biologique .....	118
Figure 5.3 : Modèle d'un neurone artificiel .....	119
Figure 5.4 : Représentation matricielle d'un neurone artificiel.....	120
Figure 5.5 : Exemples de Fonctions d'activation connexionnistes .....	121
Figure 5.6 : Perceptron simple .....	122
Figure 5.7 : Perceptron multicouche .....	123
Figure 5.8 : Exemples de séparation linéaire et non linéaire de données .....	128
Figure 5.9 : Exemple de discrimination linéaire par les SVM.....	129
Figure 5.10 : Configuration transformée des SVM.....	129

Figure 5.11 : Graphe représentatif d'un SVM binaire à marge dure.....	130
Figure 5.12 : Graphe représentatif d'un SVM à marge souple .....	132
Figure 6.1 : Système d'aide au diagnostic des anormalités et des tissus sains mammographiques .....	139
Figure 6.2 : Exemples de mammogrammes bénins de la base MIAS .....	140
Figure 6.3 : Exemples de mammogrammes malins de la base MIAS .....	141
Figure 6.4 : Exemples de mammogrammes avec calcifications bénignes .....	141
Figure 6.5 : Exemples de mammogrammes avec calcifications malignes .....	141
Figure 6.6 : Architecture du système d'aide au diagnostic des anormalités mammographiques .....	143
Figure 6.7 : Correction du contraste des anormalités malignes .....	144
Figure 6.8 : Correction du contraste des anormalités bénignes.....	144
Figure 6.9 : Correction du contraste des calcifications malignes.....	145
Figure 6.10 : Correction du contraste des calcifications bénignes.....	145
Figure 6.11 : Localisations de masses malignes .....	146
Figure 6.12 : Localisation d'une masse bénigne.....	147
Figure 6.13 : Localisation de calcifications malignes.....	147
Figure 6.14 : Détection de calcifications bénignes .....	147
Figure 6.15 : Détection d'une masse bénigne .....	148
Figure 6.16 : Détection d'une masse maligne .....	148
Figure 6.17 : Détection de calcifications bénignes .....	149
Figure 6.18 : Détection de calcifications malignes .....	149
Figure 6.19 : Apprentissage neuro-génétique pour la génération de la meilleure population .....	155
Figure 6.20 : Variation de la compacité en fonction des masses .....	157
Figure 6.21 : Variation de la moyenne des moments de Hu en fonction des calcifications.....	158
Figure 6.22 : Variation de l'entropie en fonction des masses .....	158
Figure 6.23 : Synoptique de détection d'anormalités et de tissus sains du sein.....	159
Figure 6.24 : Correction du contraste d'un mammogramme malin .....	160
Figure 6.25 : Correction du contraste d'une masse bénigne .....	160
Figure 6.26 : Correction du contraste d'un mammogramme normal .....	160
Figure 6.27 : Extraction d'une anormalité maligne dans un mammogramme à tissu dense .....	161
Figure 6.28 : Extraction du tissu sain dans une structure mammographique dense .....	161
Figure 6.29 : Délimitation d'une anormalité bénigne dans un mammogramme à tissu graisseux .....	162
Figure 6.30 : Variation des descripteurs en fonction des mammogrammes sains.....	162
Figure 6.31 : Variation des descripteurs en fonction des mammogrammes pathologiques .....	163

## Liste des algorithmes

Algorithme 3.1	Modèle adaptatif de Chan et Vese pour la détection morphologique des entités sénologiques du sein.....	89
Algorithme 3.2	Algorithme d'agrégation de pixels associée au modèle de Lankton pour la détection des néoplasmes mammaires .....	90
Algorithme 3.3	Coopération entre le regroupement spatial flou et le modèle de Chan et Vese adaptatif pour la localisation des structures pathologiques et saines dans les mammogrammes.....	91
Algorithme 5.1	Apprentissage et Catégorisation neuro-génétique de masses et de calcifications dans les mammogrammes.....	127
Algorithme 5.2	Classification des mammogrammes par la stratégie un contre un.....	136

## Liste des abréviations

<b>ACP</b>	Analyse en Composantes Principales
<b>ACR</b>	American College of Radiology
<b>ADC</b>	Abnormality Detection Classifier (classifieur de détection des anomalies au gré de l'estimation des poids par l'ACP)
<b>ANN</b>	Artificial Neural Network (réseaux de neurones artificiels)
<b>AUC</b>	Area under curve (Aire sous la courbe)
<b>BI-RADS</b>	Breast Imaging Reporting and Data System
<b>CAD</b>	Computer-aided-diagnosis (système d'aide assistée au diagnostic par ordinateur)
<b>CADe</b>	Computer-aided-detection (système d'aide assistée à la détection, détection assistée par ordinateur)
<b>CADx</b>	Computer-aided-diagnosis (système d'aide assistée au diagnostic)
<b>CBN</b>	Classification par Bayes Naïf
<b>CCNN</b>	Cascade correlation neural network (réseau de neurones à corrélation en cascade)
<b>CNN</b>	Classification neuro-génétique
<b>CNNA</b>	Classification neuro-génétique adoptée
<b>DL-CNN</b>	Deep learning convolutional-fully complex-valued relaxation neural network classifieur
<b>FCRN</b>	Fully complex-valued relaxation neural network classifieur (classifieur neuronal de relaxation à valeurs entièrement complexes)
<b>FDA</b>	Food and Drug Administration (administration des aliments et des médicaments)
<b>FN</b>	Faux négatifs
<b>FP</b>	Faux positifs
<b>GEMS</b>	General Electric Medical Systems (systèmes médicaux électriques généraux)
<b>GLCM</b>	Gray-level co-occurrence matrix (matrice de cooccurrence à niveaux de gris)
<b>IA</b>	Intelligence artificielle
<b>MIAS</b>	Mammographic image analysis society
<b>MLO</b>	Medio-Lateral-Oblique (incidence oblique médio-latérale)
<b>MLP</b>	Multi layer perceptron (perceptron multicouche)
<b>PACS</b>	Picture archiving and communication system (système d'archivage et de transmission d'images)
<b>QKCLDA</b>	Quasi-conformal kernel common locality discriminant (classification discriminante à localisation quasi-conforme)
<b>RBNN</b>	Radial basis neural network (classification par réseau de neurones à base radiale)
<b>ROC</b>	Receiver operating characteristic (caractéristique de fonctionnement du récepteur)
<b>ROI</b>	Regions of interest (régions d'intérêt)
<b>Sp</b>	Spécificité
<b>St</b>	Sensibilité
<b>SVM</b>	Support vector machine (séparateurs à vaste marge, ou machine à vecteurs support)
<b>VAO</b>	Vision assistée par ordinateur
<b>VN</b>	Vrais négatifs
<b>VP</b>	Vrais positifs

# Introduction générale

---

*« La science mesure et calcule, en vue de prévoir et d'agir. Elle suppose d'abord, elle constate ensuite que l'univers est régi par des lois mathématiques. » 'H. Bergson'*

Depuis près de deux siècles, la chimie, la physique, puis la technologie, ont servi de base à la plupart des idées nouvelles, qui ont permis à la médecine de progresser. Dans ce triptyque, la technologie voit son rang s'accroître, pour s'allier avec les disciplines médicales et être, l'une des clefs du progrès. La science d'Hippocrate épouse une nouvelle ère, avec l'avènement de l'imagerie médicale. L'emblème de cette science, née des lois mathématiques, physiques et sémiologiques, est acheminé vers un monde insatiable de cohérence et de précision. Notre ambition, est de parcourir cet espace rationnel, pour explorer les interrogations cliniques, autour de la santé humaine...

## 1. Contexte scientifique

Notre recherche aborde, le monde de la vision par ordinateur et se situe plus largement dans le cadre de l'analyse d'images, de la reconnaissance des formes et de l'intelligence artificielle par certains aspects. La vision de bas niveau, contrairement à ce que cette dénomination, pourrait laisser croire ; présente des difficultés très importantes, puisque c'est sur ces premiers traitements (effectués sur l'image), que va reposer la qualité des résultats obtenus, par la vision de haut niveau.

L'interprétation d'images, dans un contexte clinique, reste un problème majeur de l'imagerie médicale, du fait de la multiplicité des modalités y afférentes, ainsi que des fortes variabilités des structures pathologiques à analyser. La vision par ordinateur, répond à cette problématique, pour l'aide au diagnostic, particulièrement lors de l'investigation mammographique, l'objet de notre étude.

## 2. Problématique et objectifs

L'humanité connaît à notre ère, un monde insatiable de progrès technologique, monde dans lequel, les conséquences matérielles (pollution industrielle, alimentation surgelée, etc..), sont transcendées pour nuire à la santé humaine. C'est à cet effet que, d'innombrables pathologies sénologiques prolifèrent, parmi lesquelles le cancer du sein ; l'une des maladies ébranlant psychologiquement et physiquement, la femme. Aujourd'hui encore, le cancer mammaire est le néoplasme qui cause la première mortalité féminine (626 679 décès en 2018) dans le monde [1, 2] et particulièrement en Algérie où l'incidence, est en nette croissance [3-5]. Différentes études [6-8], ont confirmé que c'est la détection à un stade précoce des cancers infracliniques, qui peut en améliorer le pronostic. Dans cette visée, la thérapie éviterait alors, le cycle étiologique métastatique.



C'est dans ce cadre, que les cliniciens recommandent vivement aux femmes, le dépistage prématuré oncologique : la mammographie s'annonce comme l'examen majeur, qui détecte les anomalies du sein, dont les masses et les calcifications.

Ces signes pourraient malheureusement être occultés par le radiologue : les masses se confondraient avec les tissus denses et les calcifications, s'avèreraient faiblement visibles. Dans cette optique, lorsque le clinicien est confronté à une nature suspecte mammaire ou, lorsque la complexité des structures anatomiques du sein, interfère avec la localisation des masses, la conduite exploratoire à préconiser, est de pratiquer la biopsie, un examen ô combien invasif et anxieux pour la patiente. C'est à cet effet, que le diagnostic assisté par ordinateur, intervient pour la clairvoyance médicale. Ainsi, afin de ne pas recourir d'emblée au geste opératoire, éviter l'inter-variabilité des interprétations cliniques et les erreurs de diagnostic dues à l'erreur humaine ; nous orientons la réflexion de cette étude, vers le développement d'un système semi-automatique d'aide à la décision, qui facilite la lecture des clichés radiographiques du sein. Cette lecture fait appel aux techniques de traitement d'images et d'intelligence artificielle, pour assister le praticien dans sa prise de décision.

Le travail que nous appréhendons, dans cette thèse, a pour objectif de concevoir un système de détection et de classification de calcifications, de masses et de cas sains, à partir de mammogrammes numérisés. Comment cette réflexion est-elle instituée ? Ceci relève des opérations du système de la vision par ordinateur. Notre approche technique inspirée par une perception clinique, fait l'objet d'une analyse minutieuse des mammogrammes de la base MIAS, pour une classification dirigée des pathologies et des cas normaux du sein. La méthodologie parcourue, est usée pour diligenter un processus décisionnel, destiné à révéler la présence ou l'absence de signes cliniques. En effet, l'image mammographique est une image complexe, résultant de la superposition de structures variables, qu'il faut dissocier. Il n'y a pas de modèle mammaire unique et chaque sein, nécessite une analyse rigoureuse et méthodique. Ainsi, lorsqu'une zone suspecte est localisée, la deuxième étape consiste à la classer suivant des critères analytiques caractérisant sa forme et sa texture, afin d'affiner le diagnostic étiologique.

### **3. Motivations et contributions**

L'augmentation alarmante de l'incidence du néoplasme mammaire dans le monde d'aujourd'hui et particulièrement en Algérie ; suscite notre perception humaine, pour invoquer la perception artificielle, vers la réalisation d'un système de conduite radiodiagnostique d'entités oncologiques du sein.

Les limites de la mammographie dans la localisation concise des structures pathologiques, issues d'une sémiologie complexe, dirigent notre vision vers deux approches :

- ❖ la première aborde, la détection morphologique des masses et des calcifications pour l'aide au diagnostic via une classification neuro-génétique. A cet effet, l'opération préliminaire, consiste en la correction du contraste des mammogrammes. La seconde phase du système, permet l'extraction des structures pathologiques, suivant les modèles déformables adaptatifs, guidés par la région suspecte annotée par le clinicien.

La troisième étape de l'analyse, caractérise les anomalies sénologiques, par des attributs morphométriques et texturaux, pour générer leur signature.

Cette paramétrisation, est intégrée dans l'ultime parcours systémique, à l'entrée d'un réseau de neurones classificateur : le perceptron multicouche, optimisé par les algorithmes génétiques, pour la catégorisation de masses et de calcifications, répondant aux critères de malignité et de bénignité, en vue d'une aide à la décision.

- ❖ La seconde réflexion, permet la distinction des tissus sains et des anomalies, par les séparateurs à vaste marge, en se focalisant sur la discrimination des différents pixels, qui jouxtent le voisinage de la zone suspecte. L'accent est mis en premier lieu, sur la correction du contraste des zones d'intérêt, selon des transformations morphologiques. L'opération qui suit, permet la localisation des anomalies ou des tissus sains. Elle est instituée par le regroupement spatial flou, pour la délimitation des dissimilarités entre les zones saines et pathologiques. Ce traitement est ponctué par la prise en compte, de l'information spatiale au gré de critères spécifiques, pour extraire les régions d'intérêt sans aucune ambiguïté. Le raffinement de la segmentation, est obtenu par le modèle de Chan et Vese, pour la conservation de la topologie des concavités. L'ultime opération, est exprimée par la stratégie 'un contre un' des SVM, pour la labellisation des néoplasmes malins versus bénins ainsi que, des tissus sains.

#### **4. Organisation du manuscrit**

Hormis l'introduction et la conclusion, cette thèse arbore l'analyse de mammogrammes, au gré de six chapitres, allant de la sémiologie du sein, jusqu'à la vision élaborée, pour aider le clinicien dans son interprétation.

- ❖ Le premier chapitre introduit la sémiologie du sein, tout en analysant l'imagerie mammographique, au gré de ses avantages et ses limites.
- ❖ Le second chapitre, fait l'objet d'un parcours bibliographique de quelques travaux, consacrés aux systèmes d'aide au diagnostic en imagerie mammographique.
- ❖ Le troisième chapitre aborde en premier lieu, les techniques adoptées pour la correction des mammogrammes. En second lieu, sont décrites les approches axées sur la localisation morphologique des anomalies et des tissus sains du sein.
- ❖ Le quatrième chapitre étudie les paramètres, qui estiment la complexité et la régularité de mammogrammes, à travers la morphométrie et la texture.
- ❖ Le cinquième chapitre, s'insère dans le traitement haut niveau de la perception intelligente, pour présenter les processus de classification des mammogrammes, par l'approche neuro-génétique et par les séparateurs à vaste marge.
- ❖ L'ultime chapitre, expose le système d'aide à la décision clinique des anomalies et des tissus sains du sein, réalisé.

---

# **Chapitre 1      Sémiologie du Sein et Contexte Mammographique**

---

---

# 1 Sémiologie du Sein et Contexte Mammographique

---

*« Le dépistage précoce, est le remède clef de la prévention du néoplasme mammaire. » 'F. Z Reguieg'*

Permettons-nous de spolier cette maxime, dans l'espoir d'atténuer l'anxiété humaine.

## 1.1 Introduction

Chaque année, l'organisation mondiale de la santé (OMS), inscrit à ses registres plus de 1.7 millions de nouveaux cas du cancer mammaire [1, 2], sonnante ainsi l'alarme de sa prise en charge précoce. A notre ère encore, ce type de néoplasie, constitue la première cause de décès féminin dans le monde et particulièrement en Algérie, où l'institut de santé publique (INSP) recense 4600 décès, annuellement [9, 10].

La lutte contre le cancer, devient le défi quotidien des oncologues d'aujourd'hui, dont l'unique espoir, est d'alléger le protocole thérapeutique et de réduire la mortalité. C'est dans cette optique, que la détection des premiers signes anxigènes chez la femme, serait indispensable pour améliorer le pronostic clinique. La mammographie est la radiographie principale, usée pour cette tâche.

Ce chapitre vise à parcourir l'aspect sémiologique et étiologique de la glande mammaire ainsi que, l'examen angulaire régentant le diagnostic du clinicien. L'adjectif majeur est de mieux cerner la problématique, pour concevoir notre système d'aide à la détection et au diagnostic des anomalies mammographiques.

## 1.2 Epidémiologie du néoplasme mammaire

L'incidence du cancer du sein est en nette augmentation, depuis plusieurs décennies. L'enjeu principal de l'épidémiologie de cette néoplasie, met l'accent sur l'efficacité de sa prévention et de sa prise en charge, par les décideurs. Dans cette optique, l'organisation mondiale de la santé [9] a arboré le profil de 184 pays, en vue de créer une base de référence, pour la lutte contre cette 'pandémie'. Cette étude montre que 9000 nouveaux cas de sénologie mammaire, sont estimés sur une population totale de 38,482,000 (cf. Figure 1.1) en Algérie, durant l'année 2014. La France, en dénombre 55000 cas, pour 63,937,000 habitants (cf. Figure 1.2).

M. Hamdi Cherif et al [4, 5, 10] relèvent dans leur investigation épidémiologique algérienne, 10910 nouveaux cas observés durant l'année 2015. Le taux standardisé d'incidence pour 100.000 femmes, est de 65.2 (cf. Tableau 1.1). C'est dans cette optique, que l'Algérie a élaboré une vision stratégique de lutte contre le cancer, d'action quadriennale (2015-2019) centrée sur le patient pour, sa prise en charge efficiente dans toutes ses dimensions [10].

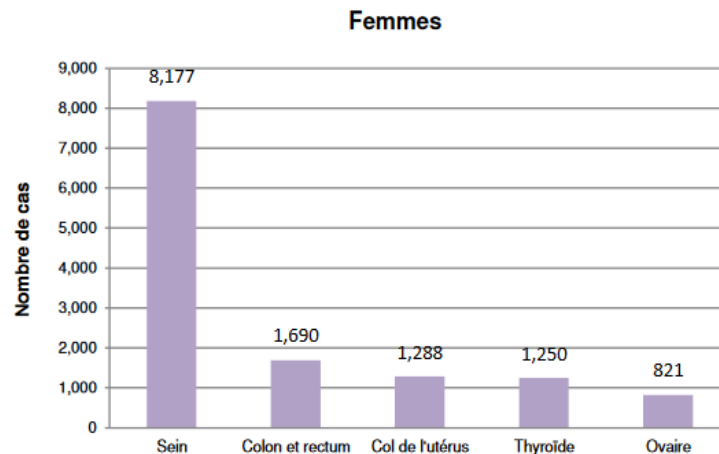


Figure 1.1 : Incidence du cancer mammaire en 2014 : cas de l'Algérie [9]

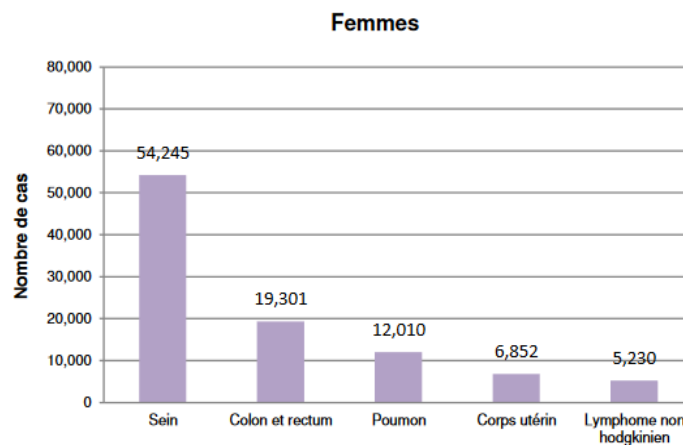


Figure 1.2 : Incidence du cancer mammaire en 2014 : cas de la France [9]

Nombre de cas en 2015	Taux brut en 2015	Taux standardisé en 2015	Incidence brute en 2025	Âge médian
10910	54.4	65.2	18112	47 ans

Tableau 1.1 : Incidence du cancer du sein en 2015 et projection future en Algérie [5,10]

### 1.3 Anatomie et physiologie du sein

L'organe mammaire féminin, parcourt plusieurs phases d'évolution anatomique, de l'étape fœtale à l'âge adulte. C'est au niveau de cette phase, que la complexité du sein, se manifeste. L'organe de l'allaitement (cf. Figure 1.3), est constitué de la glande mammaire, structurée par une arborescence de canaux galactophores où, est sécrété le lait.

Les feuilles de cette arborescence sont diffusées au niveau des unités terminales ducto-lobulaires (cf. Figure 1.4) et le tronc, se déferle vers le mamelon. L'ensemble est entouré d'un tissu de soutien, le tissu conjonctif qui contient la vascularisation et l'innervation du sein. Entre la glande et le muscle pectoral, on trouve le tissu adipeux. Sous la peau, les ligaments de Cooper déterminent des logettes graisseuses [11-14].

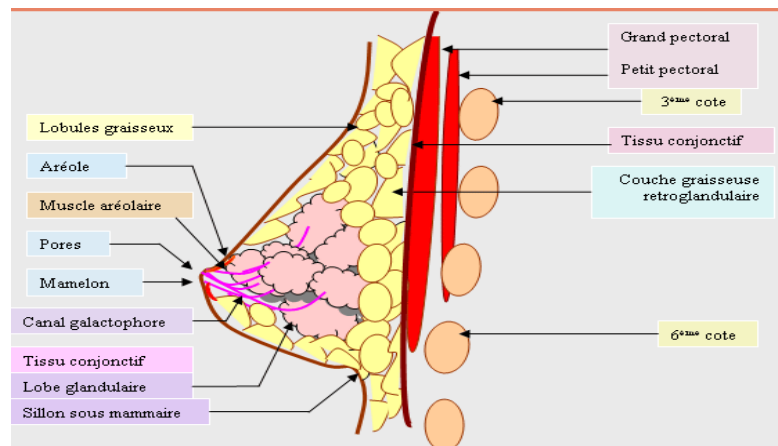


Figure 1.3 : Structure anatomique du sein : coupe sagittale [12]

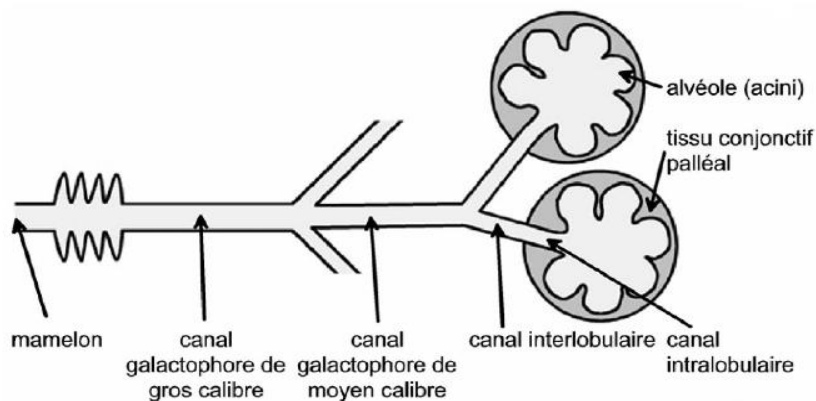


Figure 1.4 : Unité ducto-lobulaire [13]

Le sein joue un rôle important dans la représentation de la féminité. De la puberté où, les adolescentes surveillent avec appréhension le développement de leur organe mammaire, à la ménopause où, sa forme et son volume, évoluent. Le sein est investi d'une valeur symbolique et affective.

Constitué d'un tissu adipeux qui lui offre sa forme et son volume, le sein évolue selon la nature féminine. La glande mammaire est composée d'une vingtaine de lobules et comporte entre 10 et 100 acini, qui sécrètent le lait.

Parcouru par une multitude de vaisseaux sanguins, le sein est maintenu par la peau et des fibres au muscle pectoral, qui recouvrent le thorax [13-14].

Les tissus mammaires baignent dans des liquides lymphatiques, qui permettent d'éliminer les déchets, les cellules mortes et autres débris. Ces liquides circulent dans les vaisseaux lymphatiques, qui se déversent dans les ganglions lymphatiques. On distingue environ 35 ganglions lymphatiques autour de chaque sein, dont la plupart sont situés dans le creux de l'aisselle ou à proximité. Si un néoplasme se développe dans un sein, il s'étend souvent aux ganglions, car la lymphe peut contenir et faire circuler des débris, mais aussi des cellules cancéreuses. La lymphe circule dans toutes les directions, mais les trois-quarts environ des vaisseaux lymphatiques du sein, se déversent dans les ganglions axillaires ; qui deviennent souvent le premier foyer cancéreux, créé à distance de l'organe de l'allaitement.

A l'extrémité des seins les mamelons accueillent en leur centre, une zone pigmentée de couleur marron foncé : l'aréole. Les mamelons peuvent être turgescents ou lisses. Leur taille augmente pendant la gestation simultanément à celle des seins, sous l'effet des œstrogènes et de la progestérone, sécrétées par l'ovaire et le placenta.

La glande hypophysaire située dans le cerveau, provoque la sécrétion de prolactine lors, de la période de l'allaitement maternel. Cette hormone stimule la production de lait dans les lobules mammaires, pour nourrir le nouveau-né. Les canaux galactophores véhiculent le lait, jusqu'à la vingtaine de petits orifices [14].

Au cours du cycle menstruel, la glande mammaire augmente de volume sous l'effet des œstrogènes. A la ménopause, le volume des seins diminue, en raison du déficit en œstrogènes. Le sein de la femme ménopausée, devient essentiellement constitué de tissu adipeux.

#### **1.4 Aspect étiologique du sein et cancérogénèse mammaire**

Le néoplasme mammaire est caractérisé par la prolifération incontrôlée des cellules anormales du sein. Celles-ci se multiplient de façon anarchique, pour former une tumeur. Les cellules cancéreuses peuvent rester confinées dans le sein ou, migrer vers les ganglions avoisinants, voire le reste du corps, se disséminant en métastases [14-16].

Le cancer du sein, naît au niveau des cellules de la partie terminale des canalicules. La cancérogénèse mammaire (cf. Figure 1.5) évolue selon quatre étapes majeures :

- ❖ la première phase est constituée par une simple hyperplasie, c'est-à-dire une augmentation modérée du nombre d'assises cellulaires ;
- ❖ la seconde étape est une hyperplasie atypique où, à l'augmentation nette de l'épaisseur des lésions, s'ajoutent des anomalies cytologiques ;
- ❖ le cancer in situ, une prolifération anarchique avec respect de la membrane basale, se succède à la lésion atypique ;
- ❖ l'ultime phase du néoplasme est celle, du passage du cancer in situ au cancer invasif.

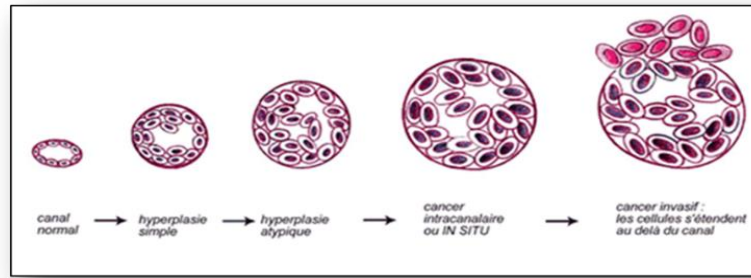


Figure 1.5 : Cancérogénèse des cellules mammaires [16]

Selon le type de cellules à l'origine du néoplasme mammaire, on distingue le cancer canalaire ou lobulaire. Ainsi, un cancer du sein qui touche les cellules qui bordent les canaux, est dit canalaire. Un cancer qui affecte les cellules des lobules, est dit lobulaire.

#### 1.4.1 Cancer du sein non invasif

Un cancer du sein est dit *non invasif* ou *in situ*, lorsque la tumeur reste dans le tissu d'origine et n'envahit pas les tissus voisins (cf. Figure 1.6). Ces cancers sont plus facilement traités que, les cancers qui envahissent les tissus voisins (cancers invasifs). Dans 90 % des cas, les cancers du sein non invasifs sont de type canalaire, les 10 % restants, étant de type lobulaire [16, 17].

#### 1.4.2 Cancer du sein invasif

Le cancer du sein est dit *invasif* ou *infiltrant* (cf. Figure 1.6), lorsque les cellules cancéreuses ne restent pas confinées à leur lieu d'origine, envahissant ainsi les tissus avoisinants : les ganglions locaux (les ganglions axillaires situés sous l'aisselle) [4]. Elles s'accumulent alors, dans les ganglions<sup>1</sup> lymphatiques voisins. On dit qu'il existe un envahissement ganglionnaire. Les cellules cancéreuses ont tendance à migrer dans d'autres organes et, à se disséminer en métastases [16, 17].

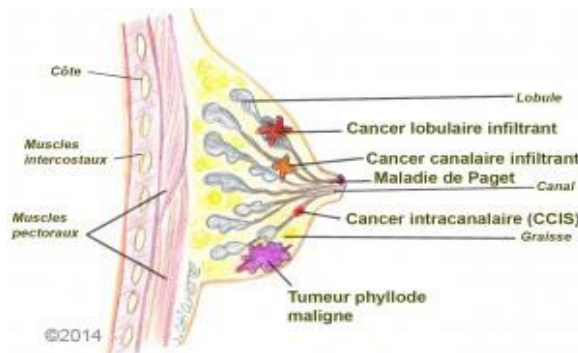


Figure 1.6 : Localisation des cancers canaux dans le sein [17]

<sup>1</sup> **Ganglion** : il s'agit d'un petit renflement réparti le long des vaisseaux lymphatiques. Disposés dans certaines parties du corps, les ganglions sont soit superficiels (dans le cou ou, l'aisselle), soit profonds (dans l'abdomen ou, le thorax). Les ganglions jouent un rôle essentiel dans la protection du corps, contre les infections ou les cellules cancéreuses. Ils mesurent normalement moins d'un centimètre de diamètre. Si leur taille est anormale, on parle d'adénopathie [3].



## 1.5 Localisation clinique du cancer mammaire

Cliniquement, le sein est divisé en quatre quadrants : le supéro-externe, le supéro-interne, l'inféro-externe et l'inféro-interne pour un repérage des lésions (cf. Figure 1.7).

Une forte majorité des cancers prennent naissance dans le tissu glandulaire. En effet, 50 % des cancers du sein se situent dans le quadrant supéro-externe de la glande, 15 % dans le quadrant supéro-interne, 11 % dans le quadrant inféro-externe, 18 % au niveau du mamelon et à peine 6 % dans le quadrant inféro-interne [14].

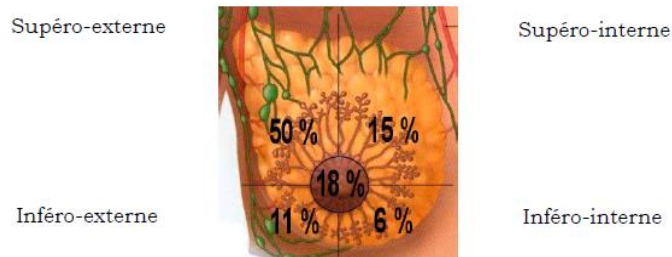


Figure 1.7 : Localisation clinique des néoplasmes mammaires [14]

## 1.6 Facteurs de risque de néoplasie mammaire

La cancérogénèse mammaire est un processus pathologique multifactoriel. Outre le sexe et l'âge, les facteurs de risque les plus connus [18,19], se déferlent ci-dessous :

- facteurs d'origine hormonale et reproductive (âge à la puberté, nombre d'enfants, âge à la première grossesse, sein dense) ;
- facteurs se rapportant à une prédisposition génétique ou, un antécédent personnel de pathologie mammaire ;
- facteurs liés au mode de vie, dont : la prise d'un traitement hormonal substitutif, une alimentation déséquilibrée, un manque d'activité physique, la consommation d'alcool, un surpoids à l'âge adulte et à un sous-poids à la puberté ;
- facteurs liés à l'environnement (exposition aux rayonnements ionisants).

## 1.7 Contexte mammographique

La mammographie née de l'application du radiodiagnostic aux affections du sein, constitue un examen de routine bien codifié, dont les résultats présentent la topographie des lésions mammaires. L'examen en question, est issu du rayonnement X [20-22].

### 1.7.1 Découverte des rayons X

La radiographie, découverte en 1895 par le physicien allemand W. C. Röntgen [20], présente un intérêt diagnostique de premier plan, dans divers axes de la médecine.

Au dix-neuvième siècle, ce physicien étudia le rayonnement cathodique avec des tubes de Crookes. Il s'intéressa plus précisément à la pénétration des rayons dans le verre. Il constata que les rayons cathodiques, pouvaient franchir la paroi du tube et pénétrer de quelques centimètres, dans l'air.

Au cours de ses travaux préliminaires, il décida de recouvrir le tube d'un cache en carton noir, pour mieux apercevoir le phénomène lumineux. Lorsque le hasard dont ne savent se servir que les savants, voulut qu'il y ait à proximité, un flacon rempli de sel de baryum ; Roentgen vit qu'à distance du tube de Crookes et pendant son fonctionnement, ce sel devenait fluorescent. Il démontra en effet que lorsque les rayons cathodiques rencontrent un obstacle, ils lui confèrent la propriété d'émettre, sous la violence du bombardement électronique, une énergie d'un nouveau rayonnement. Il en conclut qu'il vient de découvrir un rayonnement distinct de celui émis par la cathode, très pénétrant puisqu'il est capable de traverser la matière. Ces rayons étant inconnus jusqu'alors, il les nomme « rayons X » du nom de l'inconnue en mathématiques. Le 22 décembre 1895, ce physicien réalisa la toute première radiographie de l'histoire, celle de la main de son épouse, B. Röntgen. Les rayons X suscitèrent immédiatement un vif intérêt, dans diverses applications médicales,

Les premières portes de l'imagerie médicale, ouvrirent au début de l'année 1896. Grâce à cette découverte, Röntgen reçut le premier Prix Nobel de Physique en 1901.

### **1.7.2 Aperçu historique mammographique**

La mammographie a connu ses prémices en 1913, avec la découverte de A. Salomon [14], un chirurgien allemand. Il constate ainsi, qu'il est possible de distinguer différents types du néoplasme mammaire, en radiographiant des échantillons de mastectomie.

Lors de son investigation du rayonnement X, dans les années 1920, W. Vogel [20] un chercheur allemand, a fourni les premières lignes directrices, permettant de déceler les tissus pathologiques et sains du sein.

En 1949, l'uruguayen R. Leborgne [14] insiste sur la nécessité de compresser le sein lors de l'examen radiologique, pour identifier au mieux les signes cliniques. Dans les années 1950, les premiers examens préventifs ont été réalisés, grâce au radiologiste américain J. G. Cohen [14]. C'est à Ch. Gros [14] que revient le mérite du développement du premier mammographe, en 1965 à Strasbourg. Depuis, la mammographie est devenue un outil de diagnostic commun. Dans les années 2000, la mammographie numérique plein champ, a révolutionné l'imagerie du sein [22].

## **1.8 Techniques d'examen de la mammographie**

La mammographie est un type de radiographie, permettant de cartographier la structure interne des seins. Cet examen réalisé grâce à un appareil destiné à cet effet, décèle d'éventuelles lésions oncologiques, affectant la glande mammaire.

La technique en question [20-22], nécessite l'emploi de localisateurs-compresseurs, venant au contact de la peau, pour limiter exactement le faisceau du rayonnement, amoindrir la différence d'épaisseur du sein du sommet à sa base et, améliorer les contrastes dans certaines phases physiologiques.

Ces localisateurs de forme hémisphérique, possèdent une grande base inférieure demi-circulaire dont la surface variable suivant le volume du sein, est formée d'une plaque en matière plastique pour immobiliser et aplatir le sein, sans le blesser.

Ils ménagent entre le foyer de l'ampoule et le cliché, une distance de 50 à 80 cm, suivant la puissance de l'appareillage utilisé.

Une mammographie est pratiquée dans deux circonstances : dans le cadre d'un dépistage ou d'un diagnostic.

- ❖ **Dépistage.** Les femmes se soumettent à une mammographie, en vue du dépistage précoce des premiers signes cliniques oncologiques, ou encore lorsqu'elles présentent un profil à risque. Dans le cadre d'un dépistage organisé, une mammographie est réalisée tous les deux ans. Le dépistage a pour but de découvrir des cancers au stade infraclinique à l'étape des microcalcifications. Il est le meilleur garant d'un excellent pronostic et, d'une thérapie efficace.
- ❖ **Diagnostic.** Il se pratique généralement, pour localiser la présence d'un néoplasme mammaire, après avoir constaté une grosseur ou tout autre signe ou symptôme comme de la douleur, l'épaississement de la peau, un écoulement au mamelon ou la modification de la taille, ou de la forme du sein. Il peut aussi servir d'examen complémentaire, si une mammographie de dépistage révèle une anormalité.

On distingue deux types de mammographies, la mammographie analogique et la mammographie numérique.

#### ❖ Mammographie analogique

La mammographie analogique (cf. Figure 1.8), aussi appelée « classique » ou « conventionnelle » permet d'obtenir une image imprimée, sur un film argentique.



Figure 1.8 : Exemple d'un mammographe analogique [21]

### ❖ Mammographie numérique

La mammographie numérique repose sur deux types de technologies. La première, permet de transformer les clichés d'un film argentique en images numériques, grâce à un scanner ou une caméra CCD (charge coupled device : systèmes à transfert de charges). La seconde catégorie de la mammographie numérique (cf. Figure 1.9), dite « plein champ » ou « numérisation directe », permet quant à elle, de construire et de visualiser le mammogramme, sur l'écran d'un ordinateur grâce à des détecteurs numériques.

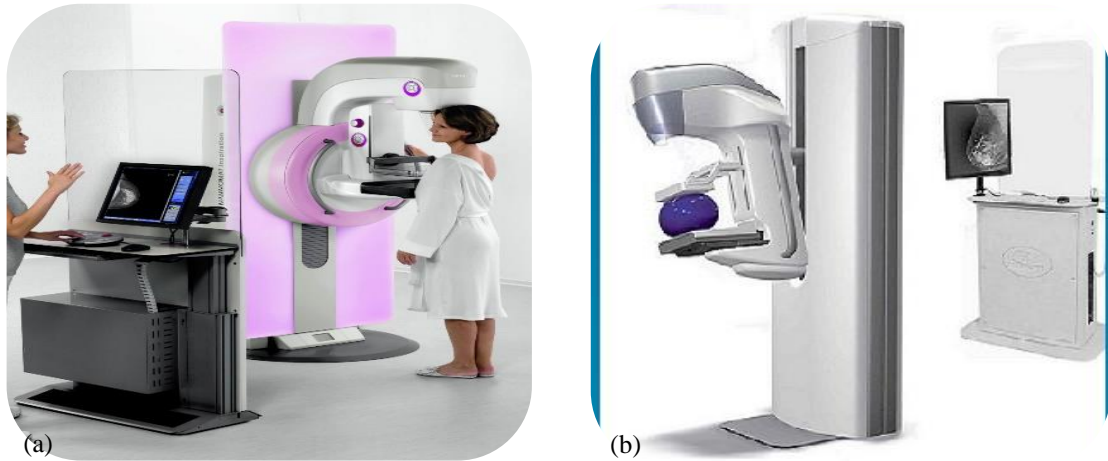


Figure 1.9 : Exemples de mammographes numériques

(a) : Cas d'un mammographe Siemens [23]

(b) : Mammographe Lambda installé à la clinique d'imagerie de Blida [24]

## 1.8.1 Incidences mammographiques

La mammographie, comporte deux incidences d'ensemble [21, 22] ainsi que, des incidences localisées (cf. Figure 1.10).

### 1.8.1.1 Incidence de face

Le sujet est assis en bout de table, les jambes sont glissées sous la table et en relâchement musculaire. La glande mammaire repose et s'étale sur le film. Le bras du côté examiné, prend appui sur la table, où le coude fléchi. La tête est tournée du côté opposé, au sein examiné. Le localisateur vient ainsi, fixer le sein en rasant la paroi thoracique.

### 1.8.1.2 Incidence de profil

Sujet en décubitus latéral du côté à examiner. Le sein repose sur le film surélevé par un billot parallélépipédique en bois. Le bras homologue, est relevé le long de la tête et, l'avant-bras plié. La main du côté opposé, efface le sein non examiné, en le tirant vers le bas pour faciliter la mise en place du localisateur compresseur.

### 1.8.1.3 Incidence localisée

Un localisateur cylindrique à ouverture circulaire de faible diamètre, est utilisé pour l'étude précise des segments suspects de la glande, révélés par l'examen clinique. Les clichés localisés sont pratiqués dans tous les cas et multipliés aussi bien sous l'incidence de face ; que sous l'incidence de profil. Il s'agit de mammogrammes avec une compression sur une zone précise du sein, qui permettent au radiologiste de distinguer si l'anomalie détectée à la mammographie de dépistage ; est réelle ou non.

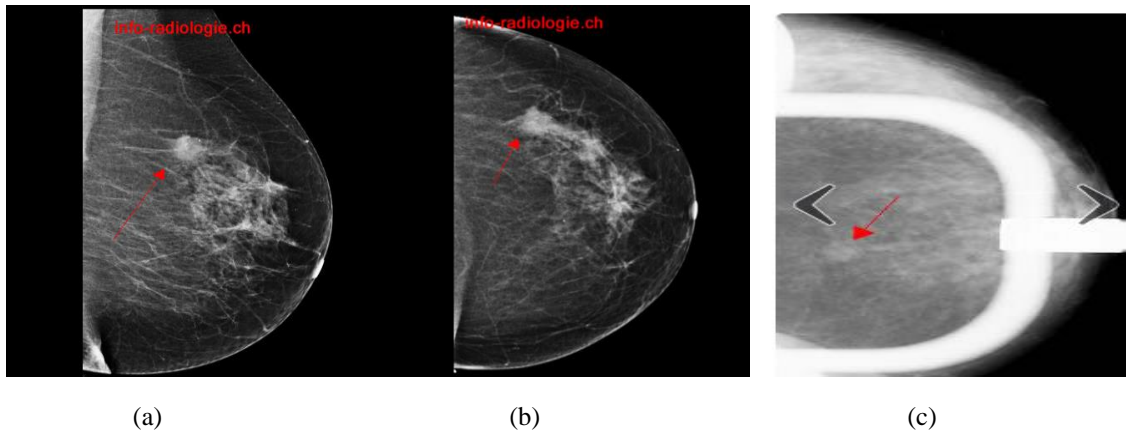


Figure 1.10 : Exemples de clichés mammographiques [22]

- (a): masse spiculée du quadrant supéro-externe du sein gauche (incidence oblique)
- (b): masse spiculée du quadrant supéro-externe du sein gauche (incidence cranio-caudale)
- (c): masse à contours spiculés (incidence localisée).

## 1.8.2 Aspect radiologique du sein

Radiologiquement, le sein a une composante adipeuse radio-transparente et, une composante fibro-glandulaire radio-opaque. Ces différentes composantes ont des coefficients d'atténuation aux rayons X presque identiques et donc, un contraste peu élevé. Dans cette optique, le tissu conjonctif, de densité hydrique, représente l'essentiel des structures opaques en mammographie. Le contraste global observé en un point de l'image est dû à une combinaison, en proportion variable d'un point à l'autre, de l'absorption des tissus conjonctifs et des tissus gras. Les canaux de la glande mammaire ne sont pas mis en relief, en raison de leur contraste différentiel trop proche de celui des tissus environnants et leur diamètre trop faible.

Le sein connaît différentes variations architecturales, au gré de la mammogénèse. La proportion des tissus gras par rapport aux tissus conjonctifs, est dépendante de l'âge et des phases physiologiques. Dans le cas de la jeune femme, on ne constate pratiquement pas de tissu gras sur la mammographie, ce qui rend souvent le diagnostic difficile car, les tumeurs ont un contraste différentiel, proche du conjonctif. Après la ménopause, il se produit une involution graisseuse : la disparition des lobules et des petits canaux. Le sein acquiert ainsi, une structure essentiellement graisseuse. Le mammogramme engendré est radio-transparent [21].

### 1.8.3 Classification BIRADS de la densité mammaire

La plupart des cancers du sein présentent une densité supérieure, ou égale au tissu conjonctivo-glandulaire.

Il est rare (mais non impossible) qu'un cancer du sein, soit de densité inférieure. Les cancers du sein n'incluent pas de tissu adipeux, à moins que l'infiltration par le processus néoplasique n'ait incorporé un éventuel lobule graisseux.

Selon l'analyse visuelle de la répartition des tissus fibro-glandulaires et graisseux, la classification BIRADS de l'ACR [21, 25] catégorise les tissus mammaires, en quatre groupes (cf. Tableau 1.2). Cette distinction estime, le risque d'avoir un cancer en fonction du caractère hétérogène du tissu fibro-glandulaire (cf. Figure 1.11). Ainsi, ce risque est en nette prédominance, lorsque la composante glandulaire est corrélée à la quatrième classe.

Catégorie	Description de la densité mammaire	Terminologie clinique
1	Le sein est presque entièrement graisseux	Tissu radio-transparent
2	Il y a des opacités fibro-glandulaires dispersées	Tissu hypodense
3	Le tissu mammaire est dense et hétérogène	Tissu isodense
4	Le tissu mammaire est extrêmement dense	Tissu hyperdense

Tableau 1.2 : Classification BI-RADS de la densité mammaire [25]

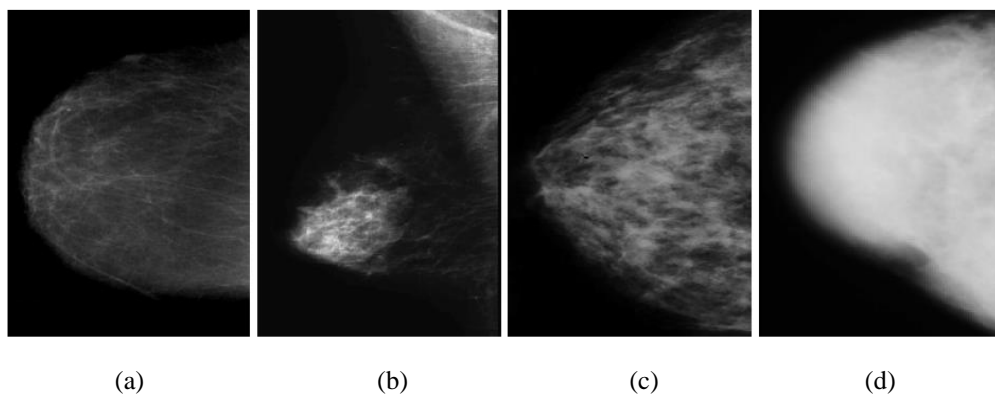


Figure 1.11 : Différents types de densité mammaire [21]

### 1.8.4 Classification BIRADS des anomalies mammographiques

Cette classification permet l'harmonisation des descriptions et, propose une classification prédictive de sévérité des images mammographiques. La classification BIRADS de mammogrammes, comprend six catégories [21, 26]. Ces dernières, ont été définies par l'ACR. Cette distinction radiologique est utilisée par la majorité des professionnels de la santé, à travers le monde (cf. Tableau 1.3).



<b>ACR</b>	<b>Anomalies mammographiques</b>	<b>Interprétation et Attitude</b>
<b>ACR0</b>	C'est une classification d'attente qui s'utilise en situation de dépistage ou dans l'attente d'un second avis.	Un tel classement indique que les résultats, nécessitent des investigations complémentaires.
<b>ACR1</b>	La mammographie n'a révélé aucun problème de santé, la poitrine de la patiente est parfaitement normale.	Aucune surveillance n'est préconisée.
<b>ACR2</b>	L'examen mammographique a révélé des anomalies bénignes.	Sans danger, ce type ne nécessite ni des examens complémentaires, ni de surveillance particulière.
<b>ACR3</b>	Cette classification indique la présence d'une anomalie d'apparence bénigne.	Cette catégorie nécessite tout de même une surveillance à court terme, d'une durée de 3 à 6 mois. Le contrôle pourra prendre la forme de nouveaux examens radiologiques, selon une échéance définie par le radiologue, en association avec l'oncologue.
<b>ACR4</b>	Cette catégorie signifie qu'une lésion douteuse ou indéterminée, a été détectée. La lésion identifiée peut être bénigne, précancéreuse ou indiquer que la patiente souffre d'un cancer.	Une biopsie et/ou une cyto-ponction sont nécessaires, pour apporter un diagnostic précis.
<b>ACR5</b>	L'examen mammographique a relevé la présence d'une lésion évocatrice d'un cancer.	Les patientes présentant ce type de résultat, se voient systématiquement prescrire une biopsie ou une exérèse. Dans la majorité des cas, la classification d'une mammographie en catégorie ACR5 découlera sur l'identification d'un cancer.

Tableau 1.3 : Classification BI-RADS des anomalies mammographiques [26]

## 1.9 Lésions oncologiques mammographiques

La lecture d'une mammographie, peut déceler d'éventuelles modifications physiologiques dues aux anomalies du sein [13, 21, 27, 28]. Celles-ci se regroupent en trois catégories dont :

- ❖ les masses et les distorsions architecturales ;
- ❖ les calcifications ;
- ❖ les densités asymétriques.

### 1.9.1 Masses typiques

Une masse en mammographie est définie par une lésion occupant l'espace visible, sur deux incidences différentes. Etant plus dense que le tissu mammaire adjacent, la masse apparaît blanche sur le mammogramme (cf. Figure 1.12). Elle est décrite suivant sa morphologie (sa taille et sa forme), son contour ainsi que, sa densité.

#### 1.9.1.1 Taille étiologique

Elle est variable de quelques millimètres, à plusieurs centimètres. Cependant elle ne prédit pas le caractère malin, sauf sur des clichés successifs, lorsqu'on la voit, croître régulièrement.

#### 1.9.1.2 Forme étiologique

Une masse mammaire peut être ronde, ovale, lobulée, irrégulière, ou présenter une distorsion architecturale (cf. Figure 1.12).

- ❖ **Forme ronde.** Il s'agit d'une masse sphérique, circulaire ou globuleuse.
- ❖ **Forme ovale.** Elle présente une forme elliptique (ou en forme d'œuf).
- ❖ **Forme lobulée.** La forme de la masse présente une légère ondulation.
- ❖ **Forme irrégulière.** Cette appellation est réservée aux masses, dont la forme est aléatoire.
- ❖ **Distorsion architecturale.** Cette forme inclut des rayons convergeant, vers un point avec une rétraction focale.

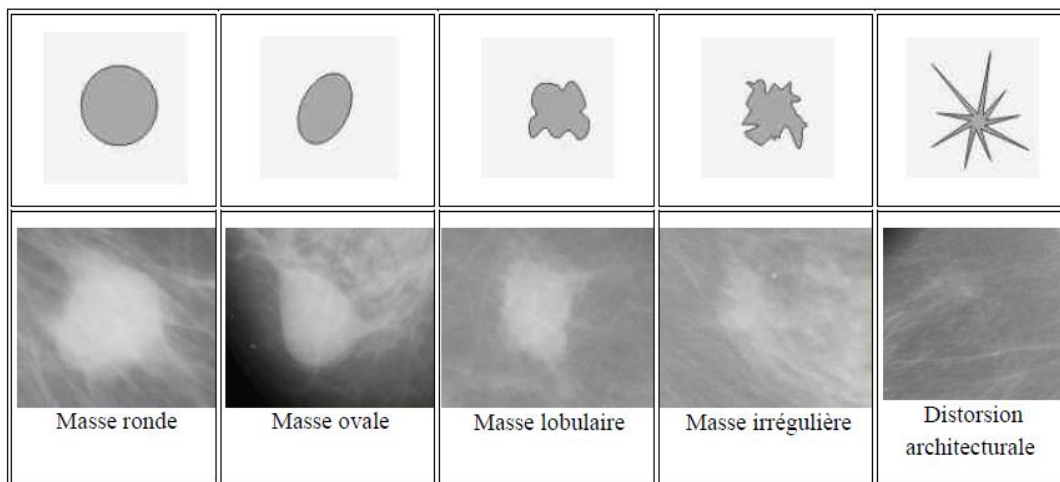


Figure 1.12 : Distinctions morphologiques de masses mammaires [27]

#### 1.9.1.3 Contour

L'étude sémiologique des opacités mammaires, appréhende (cf. Figure 1.13) divers contours, selon le degré de gravité néoplasique.

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



- b. Contour micro-lobulé.** Dans ce cas, de courtes dentelures du contour créent de petites ondulations.
- c. Contour masqué.** Un contour masqué est un contour, qui est caché par le tissu normal adjacent. Ce terme est employé pour caractériser une masse circonscrite, dont une partie du contour est cachée.
- d. Contour indistinct.** Dans ce cas, le contour est mal défini. Ce caractère indistinct (le contraire de circonscrit), peut correspondre à une infiltration.
- e. Contour spiculé.** La masse est caractérisée par des lignes radiaires, prenant naissance, sur le contour de la masse. Ces lignes radiaires, sont appelées les spicules.

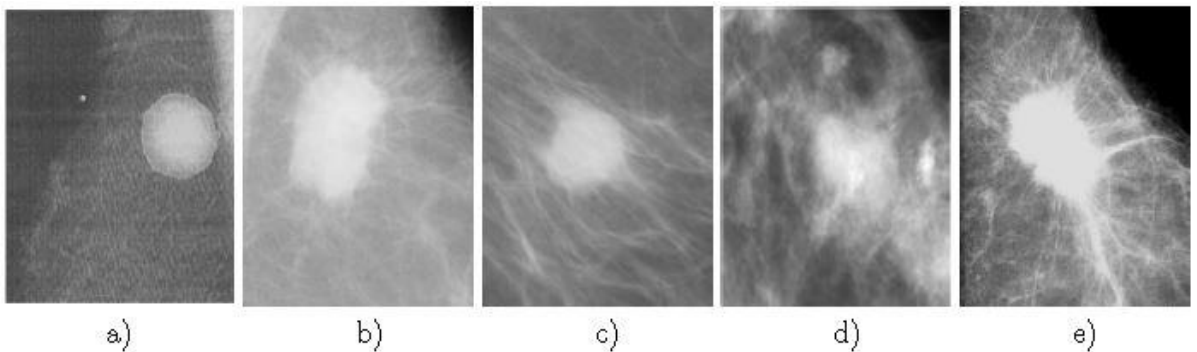


Figure 1.13 : Morphologie des contours d'une masse

a) Contour circonscrit, b) Contour micro-lobulé c) Contour masqué, d) Contour indistinct et e) Contour spiculé [27]

Au gré de leur morphologie, les masses peuvent être régulières ou irrégulières.

#### 1.9.1.4 Masses régulières

Les masses rondes, ovales ou lobulées avec des limites bien définies, sont a priori évocatrices de lésions bénignes. Cependant, cette règle n'est pas absolue et de petits cancers, peuvent être rencontrés.

#### 1.9.1.5 Masses irrégulières

Elles se distinguent par leurs distorsions, dans le tissu mammaire.

##### a. Distorsion architecturale

Une distorsion architecturale, n'est pas une masse en soi, mais une réaction tissulaire traduisant, une interruption focale de la distribution normale du tissu mammaire.

A la mammographie, elle donne l'impression d'une convergence du tissu normal vers un foyer central. On peut retrouver une telle distorsion, aussi bien près de la graisse sous-cutanée ou rétro-mammaire, qu'en plein parenchyme glandulaire.

Une telle distorsion, peut localiser un cancer, mais on peut l'observer en cas de cicatrice chirurgicale, de modification fibro-kystique ou simplement parfois, en cas de superposition de tissus denses. Lorsque le doute clinique s'installe, une telle lésion doit le plus souvent, faire l'objet d'une biopsie.

#### **b. Masse avec des limites en spicules**

Une masse aux bords étoilés, est très suspecte d'un cancer. Les spicules, représentent la modification tissulaire, en rapport avec l'infiltration tumorale.

### **1.9.2 Calcifications néoplasiques**

En termes sénologiques, les signes précoces du néoplasme mammaire, sont signalés par la présence de calcifications au niveau des canaux galactophoriques. Ce sont des dépôts de calcium, qui se forment dans le tissu du sein. Elles n'ont aucun lien avec la quantité de calcium absorbée au cours de l'alimentation ou obtenue, par l'intermédiaire de compléments alimentaires. Les calcifications mammaires sont assez courantes et la plupart, ne sont pas systématiquement associées à un cancer.

Selon l'analyse étiologique des signes pathologiques précoces, le radiologiste étudie leurs descripteurs morphologiques suivant leur disposition sur le mammogramme sur lequel, elles apparaissent sous forme de petits points blancs opaques. Certaines de leurs caractéristiques, comme une forme irrégulière ou en amas, peuvent évoquer une suspicion [14, 21]. Cependant, toutes les calcifications ne reflètent pas systématiquement une tumeur. En effet, on trouve des calcifications artérielles, des calcifications sous-cutanées, ainsi que des kystes de petite taille, contenant du lait calcique. A cet effet, les calcifications d'une taille supérieure au millimètre, de même qu'une microcalcification (inférieure au millimètre) isolée, ne signalent obligatoirement pas un cancer. En fonction du nombre des calcifications, de leur aspect individuel, de la nature de l'amas, on peut donc évaluer la probabilité qu'il y ait une tumeur cancéreuse.

#### **1.9.2.1 Macrocalcifications**

Ce sont des dépôts grossiers de calcium, dans le sein. Elles apparaissent comme de gros points blancs ou des tirets sur une mammographie. Elles sont plus fréquentes chez les femmes âgées, de plus de 50 ans. Elles sont souvent associées à des modifications bénignes qui se produisent, dans le sein, liées à d'anciennes lésions, à une inflammation, ou à des masses telles qu'un, fibroadénome ou parfois, au simple vieillissement de la glande. C'est la raison pour laquelle, lorsque ces macrocalcifications sont découvertes, le radiologue ne recommande pas systématiquement de biopsie.

#### **1.9.2.2 Microcalcifications**

Ce sont les calcifications dont les dimensions, sont inférieures ou égales à 1.5 mm. Ce sont elles, qui posent les problèmes diagnostiques les plus délicats, en raison de leurs multiples étiologies. Celles qui sont d'origine cancéreuse, sont pratiquement toujours intracanalaires.

Elles apparaissent dans des amas de cellules néoplasiques nécrosées, ou parmi des sécrétions mucineuses ou protéiniques. Elles sont constituées de phosphate tricalciques et de cristaux d'hydroxyapatite de calcium.

### 1.9.2.3 Structure clinique des calcifications mammaires

Les calcifications du sein arborent différentes structures (cf. Figure 1.14). Des signes amorphes jusqu'au polymorphes, elles peuvent être rondes, en bâtonnets, ou linéaires.

- a. Calcifications cutanées ou dermiques. Elles présentent typiquement un centre clair. Des clichés en incidence tangentielle sont souvent utilisés, pour confirmer la localisation cutanée de ces microcalcifications.
- b. Calcifications vasculaires. Celles-ci, en rails ou linéaires, sont associées à des structures tubulaires.
- c. Calcifications grossières ou coralliformes. Elles sont de grande taille (au-delà de 2 à 3 mm de diamètre).
- d. Calcifications en bâtonnets. Elles sont généralement associées à une ectasie canalaire (dilatation du canal galactophore) et sont alors, dirigées vers le mamelon. Elles mesurent habituellement plus de 1 mm de large et peuvent présenter un centre clair, si le dépôt calcique se fait dans la paroi du canal.
- e. Calcifications rondes : elles ont une forme ronde et peuvent être de tailles variables. Lorsqu'elles mesurent moins de 0.5mm, elles sont dites punctiformes ou pulvérulentes.
- f. Calcifications à centres clairs. Leur taille peut s'étendre de 1 mm à plus de 1 cm. Elles sont rondes ou ovales, à surface lisse et à centre clair. La paroi calcifiée qui les entoure, est plus épaisse que celle des microcalcifications en coquille d'œuf.
- g. Calcifications en coquille d'œuf ou pariétales. Ces microcalcifications très fines, apparaissent comme des dépôts calciques sur la surface d'une sphère. Ce dépôt mesure généralement, moins de 1 mm.
- h. Calcifications à type de lait calcique. Elles sont sédimentées dans le fond des kystes. Elles sont souvent difficiles à discerner sur l'incidence cranio-caudale. Par contre, l'incidence de profil permet de démontrer leurs formes caractéristiques (semi-lunaires, curvilignes ou linéaires).
- i. Calcifications de suture. Elles correspondent à des dépôts calciques, sur du matériel de suture. Ces microcalcifications sont typiquement linéaires ou tubulaires et, présentent fréquemment des nœuds.
- j. Calcifications dystrophiques. Elles mesurent habituellement plus de 0.5 mm de diamètre et, sont de forme irrégulière. Elles présentent parfois, un centre clair.

Ces microcalcifications sont souvent rencontrées dans un sein irradié ou, après un traumatisme mammaire. Elles représentent la majorité des cas retrouvés en pathologie mammaire.

- k.** Calcifications amorphes ou indistinctes. Elles sont souvent plus ou moins rondes ou, en forme de flocons. Elles sont de petites tailles et généralement, à contours vagues sans forme spécifique.
- l.** Calcifications fines et polymorphes. Elles sont habituellement mieux visibles, que les microcalcifications amorphes. De taille irrégulière et de forme variable, elles mesurent généralement moins de 0.5 mm de diamètre.
- m.** Calcifications linéaires et ramifiées. Elles mesurent moins de 0.5 mm d'épaisseur. Elles sont irrégulières et de formes parfois linéaires, ou curvilignes généralement discontinues, coudées ou branchées.

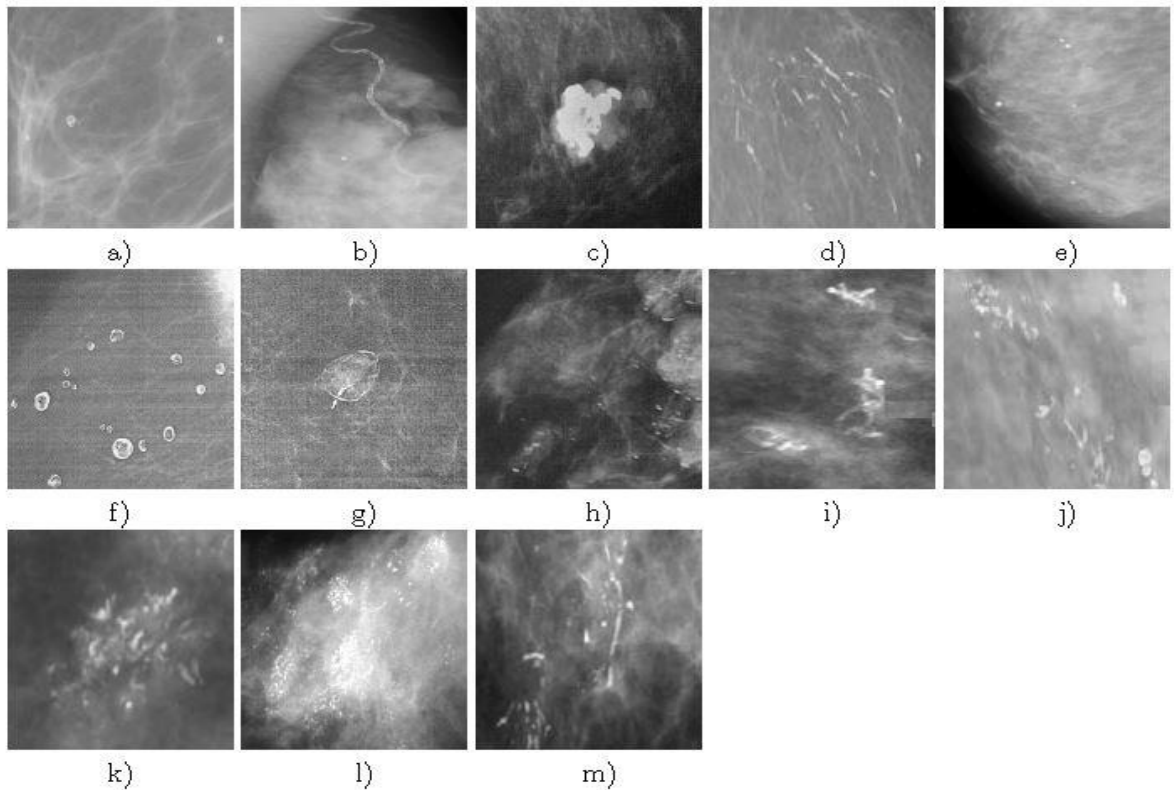


Figure 1.14 : Différents types de calcifications [27]

a) cutanées ou dermiques, b) vasculaires, c) grossières ou coralliformes, d) en bâtonnets, e) rondes, f) à centres clairs, g) en coquille d'œuf ou pariétales, h) à type de lait calcique, i) de suture, j) dystrophiques, k) amorphes ou indistinctes, l) fines et polymorphes, m) linéaires ramifiées.

#### 1.9.2.4 Analyse étiologique des calcifications mammaires

Le repérage des calcifications malignes ou bénignes, est basé sur une nomenclature clinique qui permet l'évaluation du risque. Celle-ci est diligentée essentiellement, par la taille, le nombre et la répartition des premiers signes sénologiques anxiogènes.

- a. **Analyse de la forme.** Les calcifications de forme régulière, rondes ou ovales sont des tumeurs bénignes alors que les calcifications irrégulières, de formes variées, sont souvent associées avec une pathologie maligne.
- b. **Analyse de la taille.** Les calcifications de moins de 0.5 mm, sont souvent associées à des cancers, celles de plus de 2 mm sont des lésions bénignes.
- c. **Analyse du nombre.** Les petites calcifications nombreuses sont suspectes. Un nombre au-dessus de quatre microcalcifications impose, une interprétation plus prudente.
- d. **Analyse de l'orientation.** Les amas en forme triangulaire ou losangique orientée vers le mamelon, indiquent une nature maligne.
- e. **Analyse de la morphologie.** Cette analyse permet le plus souvent, de séparer les microcalcifications bénignes et 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.

#### 1.9.2.5 Classification radiologique des calcifications selon Le Gal

Cette classification a pour but d'isoler les critères distinctifs, qui permettent de reconnaître les lésions néoplasiques à un stade précoce (cf. Figure 1.15). Dans ce cadre, les calcifications mammaires se regroupent au gré de leur morphologie, en cinq types (cf. Tableau 1.4).

Type Le Gal	Aspect mammographique	Risque
Type I	Microcalcifications annulaires rondes à centre clair radio-transparentes.	Aucun (0%).
Type II	Microcalcifications punctiformes, rondes, pleines, radio-opaques, aux contours réguliers et arrondis.	Minime (22%).
Type III	Microcalcifications poussiéreuses, trop fines pour préciser leur forme.	Modéré (36%).
Type IV	Microcalcifications punctiformes irrégulières aux contours anguleux.	Elevé (56%).
Type V	Microcalcifications vermiculaires. Elles ont la forme d'un bâtonnet. Cette forme est souvent irrégulière.	Elevé (96%).

Tableau 1.4 : Classification radiologique des calcifications du sein selon Le Gal [26]

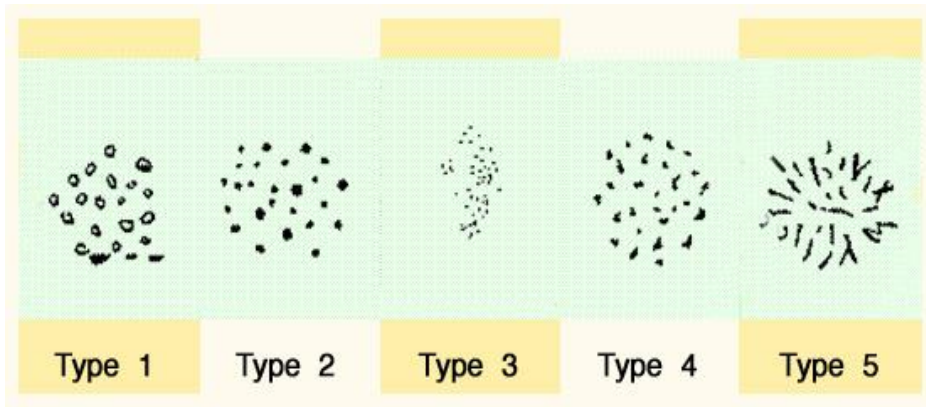


Figure 1.15 : Catégorisation morphologique des calcifications du sein selon Le Gal [26]

### 1.9.3 Asymétrie de densité

L'asymétrie de densité (cf. Figure 1.16), correspond à un aspect asymétrique localisé de la glande mammaire, sans contour définissable.

Une asymétrie peut être la conséquence d'une masse, de zones remaniées par une biopsie antérieure ou d'une mauvaise compression. Une véritable asymétrie de densité, doit se retrouver dans les différents clichés, aussi bien en oblique qu'en incidence de face. Une telle asymétrie, surtout si elle est focalisée, doit être explorée à la recherche d'une éventuelle masse.

Une densité focale asymétrique se caractérise par sa concentration croissante vers son centre, pour former une zone centrale très dense que l'on retrouve sur plusieurs clichés. Elle peut former une masse avec éventuellement des calcifications, retrouvées à la loupe. Les densités bénignes n'ont pas cette concentration centrale. Elles sont en général, assez étendues et diffuses.

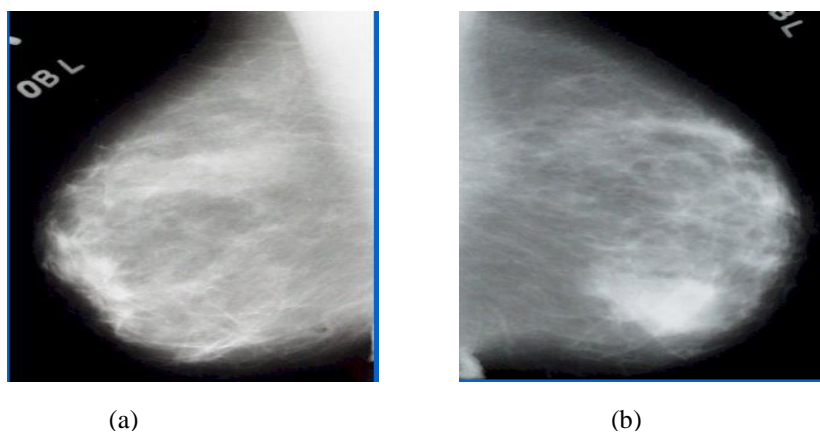


Figure 1.16 : Exemple d'une asymétrie de densité avec une nette différence dans la partie inférieure du sein gauche [28]

(a) Sein droit (b) Sein gauche



## 1.10 Limites de la mammographie

L'enjeu de la mammographie est de déceler les anomalies, avant même qu'elles ne soient palpables, notamment dans le cadre du dépistage des cancers du sein. A cet effet, les cliniciens estiment qu'une participation de 70 % des femmes à un dépistage de masse, augmenterait les chances de guérison de 30 % [21]. Cependant, même si la mammographie est la pierre angulaire du dépistage et du diagnostic des néoplasmes du sein, elle est infaillible lorsque le tissu mammaire est dense. Il est ainsi délicat pour le radiologue, de distinguer les structures saines, des pathologiques. De cette classe tissulaire, un diagnostic incertain, engendre des faux négatifs<sup>2</sup> (10%), rassurant la patiente à tort. La technique radiographique, se voit alors réduire sa sensibilité, par conséquent, sa précision.

La mammographie peut aussi faire l'objet, d'un surdiagnostic. Les risques de faux positifs<sup>3</sup> sont ainsi générés (entre 10 % et 20 %), ce qui entraîne un traitement infondé, chez certaines des femmes concernées.

Dans ce cadre, les limites potentielles de la mammographie, imposent alors aux sénologues la pratique d'examens complémentaires. Parmi ces investigations, figure la biopsie qui permet l'analyse histologique des tissus suspects, pour affiner le diagnostic. Cette technique étant invasive, augmente l'angoisse de la femme.

## 1.11 Rôle des systèmes d'aide à la détection des anomalies mammographiques

La lecture radiologique du sein est diligentée par deux phases : la détection des signes étiologiques et leur catégorisation, suivant les critères recommandés cliniquement. C'est dans ce souci, que la localisation minutieuse de la région suspecte est une étape clef, pour le radiologiste. L'interrogation que se fait alors le clinicien, est comment explorer efficacement un mammogramme, sans se préoccuper de la physiologie de l'organe de l'allaitement. La réponse, est assujettie aux systèmes d'aide à la détection et au diagnostic par ordinateur.

Le rôle du système de détection assistée par ordinateur (CADE) est de déceler, les lésions suspectes dans les mammogrammes. En parcourant la chaîne de la vision par ordinateur, le système de diagnostic assisté par ordinateur (CADx) oriente le radiologiste dans son interprétation, sans pour autant le remplacer.

## 1.12 Conclusion

L'analyse de mammogrammes au négatoscope par le praticien, peut engendrer des diagnostics ambigus. Le recours à l'intelligence artificielle devient incontournable, pour lever cette ambiguïté et assister le clinicien, dans son interprétation. Dans cette optique, les caractères radiologiques suspects du sein, imposent une profonde exploration.

---

<sup>2</sup> Un faux négatif est reconnu comme un cas bénin, alors qu'il représente en réalité une anomalie.

<sup>3</sup> Un faux positif est reconnu comme une lésion, alors qu'il n'indique aucune anomalie.

La réalisation d'un système d'aide au diagnostic pour une analyse minutieuse des clichés mammographiques, répond à cette préoccupation, pour l'extraction des lésions difficilement perceptibles à partir, des structures tissulaires mammaires. Leur détection précise, oriente le clinicien dans sa décision afin, d'assurer une thérapie efficace ; l'espoir de l'humanité de l'ère de l'intelligence artificielle.

Le chapitre qui suit, explore les systèmes d'aide à la décision en imagerie mammographique, pour remédier aux aléas de la lecture radio-clinique et concevoir ainsi, notre réflexion méthodologique.



---

## **Chapitre 2 Systèmes d'Aide au Diagnostic en Imagerie Mammographique**

---

---

## 2 Systèmes d'Aide au Diagnostic en Imagerie Mammographique

---

*« A la source de toute connaissance, il y a une idée, une pensée, puis l'expérience vient confirmer l'idée. » C. Bernard*

### 2.1 Introduction

Le recours à la vision artificielle, née de la philosophie de la perception humaine et, appliquée à l'imagerie mammaire, amoindrit l'inter-variabilité clinique pour guider le praticien, dans son diagnostic et sa décision thérapeutique. Entre la perception humaine et la vision artificielle, s'intègre le monde des connaissances sémiologiques de l'image. C'est de ce monde, que le système d'aide au diagnostic mammographique, est engendré pour assurer une seconde lecture des mammogrammes.

En effet, la quantité de clichés que le radiologue doit interpréter en un temps limité, est considérable. Des erreurs d'interprétation peuvent survenir, pour des raisons anatomiques ou de perception humaine ; ce qui rend la tâche de détection des affections sénologiques du sein, bien ardue en mammographie. L'intervention d'outils d'aide au diagnostic par ordinateur (CAD), devient indispensable pour remédier aux aléas du lecteur, en minimisant en particulier, le nombre de signes radiologiques associés à des lésions malignes, qui pourraient se confondre, avec les structures tissulaires du sein.

La question n'est pas de remplacer le clinicien, mais de lui proposer des systèmes informatiques, qui l'aident dans la lecture morphométrique des clichés mammographiques.

### 2.2 Fonctionnalités de l'imagerie médicale

L'imagerie médicale est l'un des domaines, qui a révolutionné les pratiques de la science d'Hippocrate. Naissant avec la découverte des rayons X par W. Röntgen en 1896 [20], elle assure aujourd'hui, l'exploration in vivo des organes humains. L'enjeu majeur de cette discipline touche ses acteurs principaux, à savoir ; le patient et le clinicien. A cet égard, cette technoscience met en avant un idéal clinique, centré sur la vision artificielle, en vue de la fiabilité diagnostique. Une nouvelle ère des sciences médicales, a vu le jour.

L'intérêt d'une telle découverte pour la médecine, attise la curiosité de A. Bécclère [20, 29], le précurseur de la radiologie en France, chef de service à l'hôpital Tenon à Paris, qui acquiert en 1897 un radioscope. Il y pratique, les premiers dépistages de la tuberculose, grâce aux radiographies des poumons. Il quitte Tenon pour rejoindre l'hôpital Saint-Antoine, en 1898, où il crée le premier centre français de radiologie.

Il y allie l'exercice clinique et la recherche, notamment pour évaluer les nouvelles radiations, leur intensité et leur pénétration dans le corps.

En 1908, il devient membre de l'Académie de médecine. Il fonde en 1909, la société de radiologie médicale de Paris en collaboration avec d'autres confrères, devenue depuis, société française de radiologie.

En 1896, le physicien H. Becquerel [20, 29] découvre par hasard, la radioactivité naturelle, en se rendant compte qu'un sel d'uranium émet spontanément des rayonnements distincts des rayons X. Entre 1900 et 1903, Becquerel, aidé de quelques confrères, constate qu'une exposition de 15 à 20 minutes très près de la source émettrice de rayonnement, peut induire des radiodermes. Il remarque que l'ampoule de verre, qui contient la source émettrice, doit être protégée et qu'une distance de 5 cm est conseillée, pour éviter les effets néfastes de ces rayonnements. C'est le début de la radioprotection.

Cet axe de la médecine intelligente, se développe au cours de la première guerre mondiale, avec notamment la mise en place, par M. Curie [20], de voitures radiologiques qui sillonnent les champs de bataille, pour repérer les éclats d'obus chez les soldats blessés. En 1934, la découverte de la radioactivité artificielle par F. et I. Joliot-Curie [20, 29], ouvre la voie au développement des radioéléments, de brève durée de vie. Ces derniers, associés à des molécules biologiques utilisées comme traceurs, ont permis la découverte du fonctionnement de certains organes. En 1945 a lieu la découverte de la résonance des noyaux des atomes (résonance nucléaire), soumis à un champ magnétique par E. Purcell et F. Bloch [21], tous deux ont reçu le prix Nobel en 1952. La première image IRM sur un animal s'effectue en 1973, par le chimiste américain P. Lauterbur [21]. En 1990, la médecine nucléaire évolue encore, avec la scintigraphie et la tomographie par émission de positons (TEP).

En 1972, les radiologues britanniques A. Mc Cormack et G. N. Hounsfield [20, 29], mettent au point le scanner qui leur valurent le prix Nobel en 1979. Avec l'avènement de cette modalité, l'imagerie numérique a envahi le monde médical, pour voir l'apparition d'autres procédés d'examen de l'organisme, tels que l'angiographie digitale, les ultrasons, la résonance magnétique nucléaire, la tomographie par positons et la mammographie, le processus qui inspire nos travaux dans ce prélude à la recherche. Ce processus radiologiste, a connu son intronisation numérique dans la décennie 2000, envoyant aux oubliettes les films, les loupes et les négatoscopes. Cette technique permet alors, une meilleure détection des néoplasmes et un dépistage précoce des premiers signes anxigènes mammaires. Les balbutiements de cette prospection, ont été initiés par le suédois L. Tabar [30] qui démontre dans son étude en 1985, une diminution de 30% de mortalité pour le cancer mammaire, après cinq ans de dépistage systématique. Aujourd'hui, la mammographie numérique plein champ, a conquis les explorations cliniques pour l'aide au diagnostic.

A l'orée d'un bouleversement vertigineux, l'imagerie médicale se fonde avec l'informatique, pour souffler un vent d'espoir, sur la prise en charge des pathologies. Ses fonctionnalités (cf. Figure 2.1), rendent compte d'une minutieuse investigation au gré des finalités [31] ci-dessous :

- ❖ L'imagerie structurelle recueille des informations sur la morphologie des organes, des tissus ou des cellules, comme la taille, le volume, la localisation, ou encore la forme d'une éventuelle lésion.
- ❖ L'imagerie fonctionnelle, encore appelée imagerie métabolique et moléculaire, vise à étudier le fonctionnement et l'activité physiologique des organes, des tissus ou des cellules par notamment la technologie TEP.
- ❖ L'imagerie interventionnelle correspond à l'ensemble des actes médicaux réalisés sous le contrôle de dispositifs d'imagerie, comme le repérage, le guidage, le contrôle optimal du geste médical, etc.

Types d'imagerie	Modalités	Axes cliniques
Imagerie structurelle	Rayons X (radiologie, angiographie, mammographie, etc...)	Cancérologie
	Champs magnétiques (IRM, etc...)	Neurologie
Imagerie fonctionnelle	Ultrasons (échographie, Doppler, etc...)	Pneumologie
	Rayonnements lumineux (imagerie optique, etc...)	Cardiologie
Imagerie interventionnelle	Rayonnements lumineux (imagerie optique, etc...)	Ophtalmologie
	Imagerie biologique (cytologie, etc...)	Obstétrique
	Radiations (radiothérapie, etc...)	Dentisterie

Figure 2.1 : Fonctions de l'imagerie médicale [31]

Depuis ces découvertes, l'imagerie médicale n'a cessé d'évoluer, de se perfectionner et de proposer des technologies de plus en plus précises, performantes et novatrices. Faisant appel aux techniques du traitement du signal et de l'informatique, l'imagerie médicale associée à la perception artificielle, est devenue aujourd'hui l'alliée inéluctable des médecins et des patients ; car elle offre un meilleur diagnostic et souvent, la détection précoce de certaines pathologies. Le cancer du sein, étant l'exemple le plus connu. Son analyse fait l'objet de cette réflexion doctorale, au gré d'une imagerie structurelle.

### 2.3 Aperçu historique des systèmes CAD

Le concept de systèmes experts informatiques dédiés à l'aide au diagnostic (CAD), émerge depuis plusieurs décennies. Les chercheurs s'y attèlent pour promouvoir l'aide à l'interprétation clinique, des lésions oncologiques mammaires [6-8].

Ces systèmes sont alors destinés à une lecture informatique, dont la visée est d'aider le radiologue dans sa prise de décision.

L'atout principal de l'analyse numérique des mammogrammes, est d'assurer l'augmentation de la sensibilité de la mammographie, tout en réduisant le nombre de faux positifs.

Plusieurs études ont montré que les CAD, pouvaient augmenter de 15 à 20 % le taux de cancers détectés, en situation de dépistage. Il a également été prouvé que les techniques d'intelligence artificielle, permettaient de mieux classer les anomalies détectées, qu'un radiologue non spécialisé en mammographie [32].

L'histoire des systèmes d'aide au diagnostic (CADx), remonte au siècle précédent, dans les années soixante. Le premier, a été promulgué dans les travaux de Winsberg en 1967 [33]. Cette réflexion met l'accent sur l'analyse de mammogrammes conventionnels, en comparant les incidences bilatérales des clichés. Les différences symétriques des structures cliniques des deux images, sont ainsi délimitées. Les fortes variations des intensités dans une zone d'intérêt, marquent la présence d'une pathologie (cf. Figure 2.2).

Ackerman et Gose [34] ponctuent leur analyse sur la caractérisation morphologique des anomalies, en se basant sur leurs propriétés sémiologiques afin, de différencier les lésions malignes versus bénignes (cf. Figure 2.3).

Dans les années 1970, Wee et al [35] se sont focalisés sur des approches permettant d'analyser les foyers des microcalcifications. L'aspect de cette approche, s'articule autour de la taille, de la densité et de la morphologie des calcifications éparses, ainsi que celles en amas.

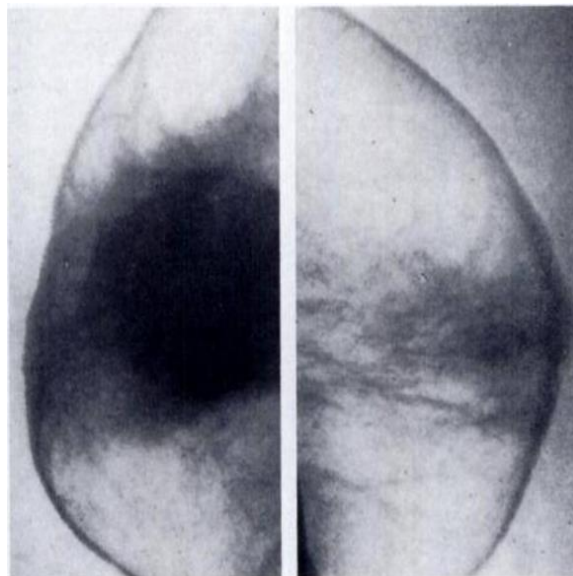


Figure 2.2 : Masse produite par une asymétrie de densité dans une mammographie conventionnelle cranio-caudale [33]

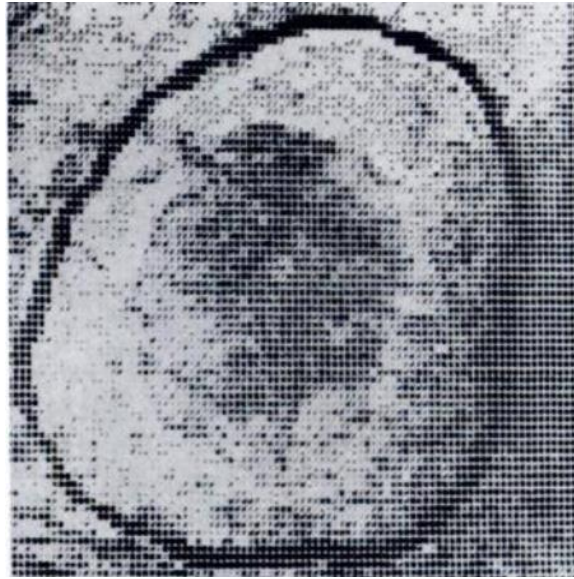


Figure 2.3 : Masse annotée dans un mammogramme conventionnel [34]

En 1975 Tasto et al [36], ont décrit un algorithme de détection de microcalcifications mammographiques, basé sur l'analyse des niveaux de gris. En 1979, Spiesberger [37] a ouvert la voie aux systèmes automatiques d'aide au diagnostic des anomalies mammographiques. Dans sa réflexion, il a institué un système de détection des microcalcifications, via leur caractérisation globale au gré de paramètres statistiques et morphologiques. Les objets candidats sont catégorisés, par un arbre de décision. Ce n'est qu'en 1980, que les CAD ont été lancés au sein des laboratoires K. Rossmann du département de radiologie de l'université de Chicago, pour l'analyse des images médicales [38]. Le premier prototype 'circa' (cf. Figure 2.4) fut développé en 1994 au sein de l'institution de Chicago [38], pour le dépistage du cancer du sein.

Les chercheurs des laboratoires K. Rossmann [38, 39], se sont investis dans l'exploration des CAD en imagerie mammographique, pour la détection précoce des néoplasmes mammaires. Le système en question, reçoit en entrée un cliché radiologique analogique, qui subit la numérisation, pour être analysé automatiquement. L'annotation des mammogrammes visualisés sur un écran d'ordinateur ou imprimé sur un papier thermique, indique la localisation des lésions suspectes. L'adjectif final des laboratoires K. Rossmann, est d'intégrer ainsi, le CAD dans un environnement alliant, l'archivage et la transmission d'images 'PACS' à des fins de diagnostic à distance : la télémédecine est ainsi née.



Figure 2.4 : Premier prototype CAD circa [39]

## 2.4 Systèmes CAD commercialisés

Plusieurs systèmes CAD sont actuellement commercialisés avec un agrément FDA [40-43]. Ils permettent aux radiologues, une aide à l'interprétation clinique. Nous citons quelques systèmes CAD, usités au profit de l'interprétation clinique.

1. **Image Checker** [40]. Cet outil a été agréé par la FDA en juin 1998 pour le dépistage du cancer du sein et en juin 2001, pour la mammographie diagnostique. Ce fut le premier système à être implémenté, sur les mammographes numériques GEMS. Il est conçu pour rechercher les signes, qui peuvent être associés au cancer du sein. Ce système, délivre une détection des microcalcifications avec une sensibilité de 98.5% de sensibilité, alors que la localisation des masses, atteint une sensibilité de 85.7%. Dans cette technologie, les masses sont indiquées par une astérisque et les calcifications sont représentées, par un triangle (cf. Figure 2.5).
2. **Second Look** [41]. Ce système a été agréé par la FDA en 2001 et est actuellement disponible, également sur les mammographes numériques GEMS. Ce CAD a été conçu, pour annoter les régions suspectes dans les mammogrammes. Les masses sont notifiées par des cercles et les calcifications, sont repérées par des rectangles. La sensibilité du système, est de l'ordre de 85%.
3. **MX-Mammo** [42]. Ce système permet d'affiner le diagnostic de la mammographie, avec l'intégration d'examen sénologiques, provenant de l'IRM et de l'échographie.



4. **Power Look Tomo** [43]. Cet outil, présente des avantages significatifs de la technologie révolutionnaire d'intelligence artificielle, pour la tomosynthèse du sein numérique. Il s'agit d'une description analytique mammaire profonde (cf. Figure 2.6). Power Look Tomo permet d'identifier et marquer les zones suspectes précisément, sans obscurcir l'image sous-jacente. Le système en question, réduit le temps de lecture de 29,2%, optimisant ainsi, le flux de travail du radiologue.

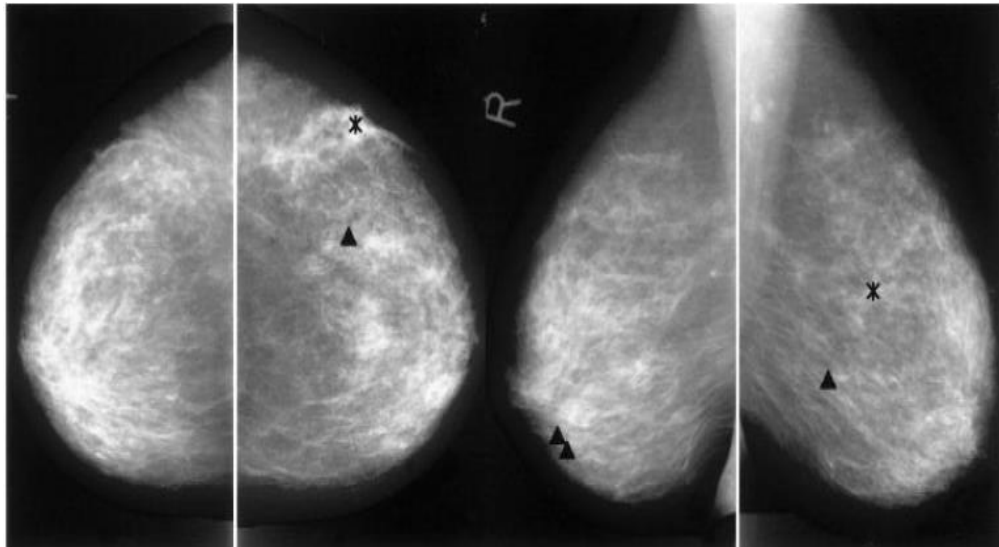


Figure 2.5 : Mammogrammes annotés via le CAD Image Checker [39]

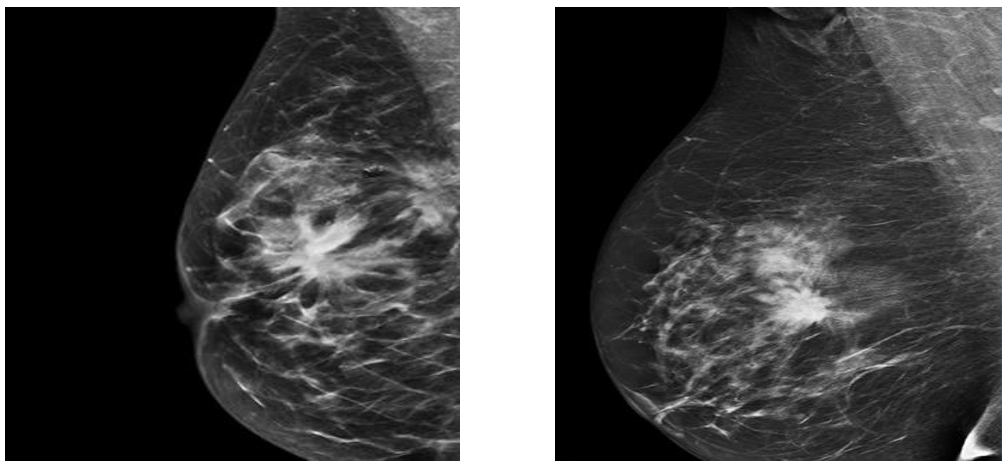


Figure 2.6 : Mammogrammes rehaussés via le CAD Power Look [43]

## 2.5 Utilité des CAD

Les systèmes CAD, remédient à des biopsies inutiles, offrant ainsi, au clinicien une aide précieuse, à l'interprétation des clichés radiographiques du sein. Dans ce cadre, les outils d'aide à la décision, permettent de détecter non seulement, les néoplasmes à un stade précoce, mais aussi de différencier les tissus sains des pathologiques, dans le cas où les structures sont hautement densifiées.



Il est alors possible, d'augmenter la sensibilité de la modalité radiographique du sein, en se soumettant à la vision artificielle. Les systèmes CAD constituent ainsi, une seconde lecture clinique des mammogrammes permettant aux praticiens, d'affiner le diagnostic et de préciser le protocole thérapeutique, à administrer aux patients.

## 2.6 Introduction à la détection assistée par ordinateur dans la mammographie

Le système de détection de pathologies mammographiques, est basé sur la vision par ordinateur (VAO) (appelée aussi vision artificielle) une branche de l'intelligence artificielle dont le principal but, est de permettre à une machine d'analyser, de traiter et d'interpréter une ou plusieurs images, prises par un système d'acquisition.

L'ultime but d'un système CAD, est de figurer la décision du radiologue, en réduisant les faux positifs lors de la détection d'éventuelles anomalies, via l'augmentation de la sensibilité.

Cette manière de procéder peut-être perçue, comme un traitement des données visuelles par le biais de modèles fondés sur les mathématiques, la sémiologie, les statistiques et la théorie d'apprentissage.

Le système de détection assistée par ordinateur (cf. Figure 2.7), parcourt trois phases clés en amont de la décision :

1. Un prétraitement, qui permet la correction du mammogramme, en vue d'éliminer les artefacts ou, de rehausser le contraste. La visée de cette opération, est d'assurer une extraction minutieuse des structures symboliques.
2. Une segmentation, une opération qui localise les signes cliniques, selon des critères spécifiques à la nature de l'image.
3. Une opération de caractérisation des régions d'intérêt, qui pourvoit l'extraction des connaissances sémiologiques mammaires.
4. Une classification qui permet la prise de décision, grâce aux connaissances générées, en utilisant un modèle adéquat.

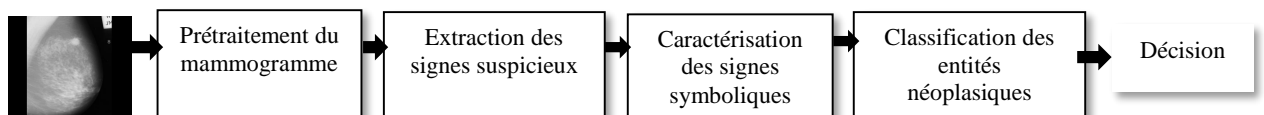


Figure 2.7 : Système de détection assistée par ordinateur de pathologies en mammographie numérisée

### 2.6.1 Prétraitement

Les clichés mammographiques nécessitent une correction, pour la mise en exergue des structures d'intérêt, relativement aux autres composants du sein. En effet, les mammogrammes (cf. Figure 2.8) sont des images difficilement interprétables, étant donné l'enchevêtrement des différentes structures tissulaires et le contraste relativement faible, entre elles. C'est dans ce contexte, que le processus de prétraitement est une opération phare, dans la VAO, pour le rehaussement de la qualité de l'image. Cette amélioration, conditionne l'extraction minutieuse des signes néoplasiques.

Le processus, permet d'user de plusieurs techniques, réduisant les artefacts dus au système d'acquisition ou accentuant le contraste, pour mettre en avant les signes pertinents.

Le prétraitement des images, est obtenu généralement, par des opérations de filtrage. Cependant, leur utilisation pourrait altérer l'information pertinente. C'est pourquoi, leur application, dépend intimement de la structure de l'image. Dans cette optique, les chercheurs se basent sur des approches locales, afin de préserver les détails de l'organe. A cet effet, une pléthore de travaux est accomplie via des approches morphologiques, réalisant ainsi, une correction du cliché du sein, moyennant des paramètres adaptatifs à l'information recherchée. C'est le cas des travaux, de Rick [45], de Mohamed et al [46] ainsi que l'investigation de Boulehmi et al [47].

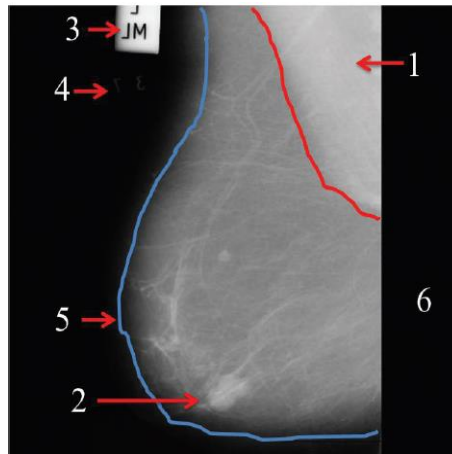


Figure 2.8 : Constitution d'un mammogramme [44]  
1 : Muscle pectoral, 2 : Tumeur, 3 : Etiquette, 4 : Basse intensité  
5 : Contour de la glande mammaire 6 : Fond de la glande

Dans l'exploration de Rick [45], les résidus morphologiques, furent les principaux opérateurs de rehaussement, du contraste des calcifications au gré d'un élément structurant approprié. Dans les travaux de Mohamed et al [46], le prétraitement des mammogrammes, fut abordé en premier lieu, par une égalisation de l'histogramme, suivie d'une accentuation du contraste par les opérateurs de la morphologie mathématique.

Boulehmi et al [47] ont établi leur approche de correction des images mammographiques, en quatre phases : la première consiste à supprimer les raies de numérisation, via des opérations morphologiques. La seconde étape inspirée des travaux de Kumar et al [48], réalise le débruitage du mammogramme, autour des ondelettes. Le traitement qui suit, est l'élimination du muscle pectoral. Un seuillage adaptatif est appliqué sur le mammogramme, mettant en relief tous les pixels appartenant au muscle, pour les écarter de la région d'intérêt. L'ultime étape, est le rehaussement du contraste des structures symboliques. Ceci est accompli, en se focalisant sur l'extraction et la suppression de l'arbre galactophorique, via la transformation du chapeau haut de forme.

L'amélioration de la qualité du mammogramme est obtenue dans les travaux de Kashyap et al [49], par un filtrage anisotropique. Cette technique réduit le bruit dans l'image, tout en préservant les contours. Dans une autre investigation, Kashyap et al [50], ont corrigé le cliché mammaire, grâce à une ouverture morphologique pour retirer les étiquettes et le fond de l'image. Une opération de soustraction entre l'image obtenue et son complément, réalise le rehaussement des régions suspectives.

Dans les travaux de Ferkous [51], le prétraitement des mammogrammes, est réalisé par les modèles déformables, pour ne considérer que la glande mammaire. Le muscle pectoral est alors éliminé.

Dans la réflexion de Hmida et al [52], le contraste du cliché du sein est rehaussé, par une transformation non linéaire, assujettie à une correction gamma.

L'investigation mammographique de Pawar et al [53], est consacrée à l'amélioration des mammogrammes, par une maximisation de l'entropie locale.

### 2.6.2 Segmentation

La segmentation est une opération qui partitionne le cliché mammographique, en des entités d'intérêt. Le traitement en question, est une opération délicate, du fait de la diversité structurale des images. C'est dans cette optique, que le choix d'une technique de segmentation, est imposé par la nature des données analysées.

Les travaux actuels [49, 50, 52-55], se basent principalement sur les modèles dynamiques, la logique floue et les opérateurs morphologiques, afin de ne pas altérer les zones de transition des régions, particulièrement, dans le cas des lésions irrégulières. Dans les travaux de Kashyap et al [49], l'extraction précise des masses est accomplie par les ensembles de niveaux. Dans une seconde réflexion, Kashyap et al [50], ont localisé les masses, autour de l'approche des c-moyennes floues, traitée sur l'image des approximations, obtenue par une décomposition multirésolution.

Dans l'un des travaux de Rangayyan [6], la délimitation des lésions oncologiques du sein, est obtenue par une segmentation en régions adaptative. L'agrégation des pixels d'intérêt, se fait alors, selon des critères géométriques appropriés.

Dans une autre réflexion, de son parcours, Rangayyan se focalise sur la détection des distorsions architecturales, par analyse texturale autour de la transformée de Gabor. Le but, est d'orienter le radiologue, dans son interprétation.

Dans la réflexion de Mohamed et al [46], l'extraction des régions d'intérêt est obtenue, par un seuillage automatique.

Dans les travaux de Hmida et al [52, 54], l'extraction des masses, est obtenue par une hybridation du regroupement spatial flou et, des modèles actifs. L'objectif visé, est d'utiliser du formalisme des contours flous, pour l'obtention d'un contour final optimal délimitant chaque masse, dans la région d'intérêt considérée. Ce formalisme est alors associé à un modèle des contours actifs, pour localiser avec précision, les entités sénologiques.

X. Liu et al [55] ont localisé les contours des masses dans les mammogrammes, suivant les ensembles de niveaux, pour mettre en exergue les différentes formes y afférentes.

Vachier [56] s'est consacrée à la détection des opacités mammaires, par la ligne de partage des eaux, l'un des traitements de la morphologie mathématique.

Dans l'exploration scientifique des travaux d'Eddaoudi [57], la segmentation des anomalies, a été effectuée autour d'une analyse texturale, pour délimiter les zones de forte densité.

F. Liu et al [58] ont localisé les masses dans les mammogrammes, autour des ensembles de niveaux, en intégrant l'information du gradient, pour préserver les zones de transition. Le résultat obtenu, est raffiné par le contour actif géodésique, basé sur les informations locales de l'image.

Dans l'une des investigations présentées dans la synthèse des travaux de Shanmugavadivu et al [59], l'extraction des masses, est accomplie via la ligne de partage des eaux par marquage. Ce résultat est ensuite affiné, grâce aux ensembles de niveaux. Dans une autre exploration, la segmentation des opacités mammaires, est réalisée par une croissance de régions précédée par une opération morphologique permettant la suppression des artefacts et du muscle pectoral.

### **2.6.3 Caractérisation des signes symboliques**

Lorsque la pathologie est localisée, l'étape qui suit, est la description des ROI, en vue d'une extraction des connaissances sémiologiques. Dans ce cadre, le radiologue privilégie la morphométrie, afin d'apprécier les différents changements y afférents. Néanmoins, la tumeur ne varie pas uniquement sur le plan morphologique, son aspect textural est également affecté. Il permet d'estimer le degré de complexité étiologique. A cet égard, divers attributs, sont utilisés dans les travaux d'aide à la détection et au diagnostic du cancer du sein, pour prendre en compte la caractérisation néoplasique [27, 46-58].

Dans les travaux de Rangayyan [6], la caractérisation des entités sénologiques est basée, sur des critères géométriques, fréquentiels et texturaux.

Cheikhrouhou [27] s'est focalisée sur la morphologie sémiologique des masses, pour les caractériser en vue d'évaluer l'ordre de complexité particulièrement, dans les distorsions architecturales.

Mohamed et al [46], ont caractérisé les entités néoplasiques des mammogrammes MIAS, au gré de paramètres texturaux issus de la matrice GLCM ; tels que l'entropie, l'énergie et l'homogénéité.

Boulehmi et al [47] se sont orientés vers la description des masses, autour des paramètres de forme, texturaux ainsi que ceux issus des coefficients d'ondelettes. Parmi lesquels, sont exprimés, la compacité, l'excentricité, la moyenne, l'écart type des coefficients d'approximation, celui des coefficients horizontaux, l'écart-type des coefficients verticaux ainsi que celui des coefficients diagonaux.

Dans l'exploration de Kashyap et al [50], la description des anomalies est effectuée par des descripteurs texturaux édictés par l'entropie, la corrélation et l'homogénéité, évalués sur les quatre directions de  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$ . Des attributs morphologiques révélés par les moments géométriques invariants aux transformations affines, sont également estimés pour caractériser les lésions néoplasiques.

X. Liu et al [55], se sont focalisés sur les descripteurs de Fourier, pour générer les connaissances morphologiques des anomalies traitées. La catégorisation des masses bénignes et malignes, est diligentée par les SVM.

Dans les travaux d'Eddaoudi [57], la paramétrisation des lésions mammaires est arborée autour des descripteurs texturaux, estimés selon deux orientations,  $0^\circ$  et  $90^\circ$ .

Dans les travaux de Hmida et al [52, 54], la caractérisation est exprimée par des paramètres morphologiques, comme le centre de gravité de la lésion et des attributs texturaux, tels que l'entropie. Le but primordial, est d'exprimer le stade d'évolution étiologique de la pathologie.

Dans la synthèse des travaux de Shanmugavadivu et al [59], l'une des réflexions abordées, décrit les anomalies du sein, par les coefficients des ondelettes et les attributs texturaux issus, de la matrice GLCM. La priorité est accordée, à la paramétrisation locale des régions d'intérêt. Dans le même contexte, l'un des travaux cités dans cette synthèse, s'est appuyé sur une caractérisation morphologique et statistique, des lésions du sein.

## **2.6.4 Classification des entités symboliques**

L'ultime phase du système d'aide à la décision, est instituée par une classification des entités symboliques en des labels malins, bénins ou sains. Cette classification est dirigée dans le cas où, les informations sont connues a priori ; ou non supervisée, a contrario.

### **2.6.4.1 Classification non supervisée**

Ces techniques interviennent lorsque l'identité des classes, n'est pas connue. Cela résulte d'un manque de connaissances a priori de la population, à étudier.

Une pléthore d'approches instituant la classification automatique, permettent de créer des regroupements d'individus, ayant des caractères similaires. L'un des modèles les plus connus, est celui de l'algorithme des K-moyennes [60]. Il permet de décomposer une population finie d'éléments, en un nombre K de classes intra-homogènes et inter-hétérogènes. Le modèle en question, conduit à une partition fixe de l'ensemble de données, en optimisant une fonction coût où, chaque individu, est affecté à la classe, dont le centre de gravité en est le plus proche. Avec cette méthode, la partition finale dépend du choix des centres initiaux K.

Une autre méthode de classification non supervisée, est la carte auto-organisatrice de Kohonen connue sous le nom anglais self organizing map (SOM) [61]. Inspirée par le principe neuronal humain, une carte auto-organisatrice est un type de réseau de neurones artificiels dont l'apprentissage, se déroule de manière non supervisée. Son principal rôle, est de faire une projection non linéaire des données de haute dimension, sur un espace de faible dimension.

Une carte auto-organisatrice est composée d'une grille de neurones de faible dimension. Quand la grille est unidimensionnelle, chaque neurone a deux voisins. Quand la grille est bidimensionnelle, l'arrangement des neurones se fait d'une façon rectangulaire où, chaque neurone possède 4 voisins (topologie rectangulaire) ou d'une façon hexagonale où, chaque neurone possède 6 voisins (topologie hexagonale). Les neurones sont reconnus par leur numéro et leur emplacement, sur la grille.

Au cours de l'apprentissage, chaque neurone se spécialise dans la reconnaissance d'un certain type d'entrée. Ainsi, deux observations proches dans l'espace d'entrée, activent deux unités proches sur la carte. Cette méthode est plus robuste aux conditions initiales, que l'algorithme des K-moyennes. Son inconvénient majeur, est le temps de calcul, associé aux itérations qui permettent la construction de la carte auto-organisatrice [61].

#### 2.6.4.2 Classification supervisée

La classification est supervisée ou dirigée, lorsque l'on possède suffisamment d'informations sur la population à traiter. Cette catégorie suppose avoir un groupe d'individus de chaque classe, dont on connaît l'appartenance. Ces individus, forment des échantillons d'apprentissage. Ils sont utilisés, pour entraîner le classifieur. Des échantillons de test, servent à valider la classification, en évaluant sa pertinence à travers la précision, la spécificité, la sensibilité ou d'autres facteurs.

On distingue plusieurs modèles de discrimination dirigée : parmi lesquels, les classifieurs bayésiens naïfs, les réseaux de neurones et les SVM.

La classification naïve bayésienne fut utilisée historiquement, pour la catégorisation de documents et l'élaboration de filtres anti-spam [62]. Aujourd'hui, c'est un algorithme renommé dont les applications peuvent être rencontrées, dans de nombreux domaines, tels que le médical. La méthode de classification naïve bayésienne est un modèle d'apprentissage supervisé, assujetti au théorème de Bayes, dont le rôle est de classifier, les échantillons, ayant des propriétés similaires, mesurées sur des observations.



Les réseaux de neurones artificiels, connus par l'acronyme anglais ANN, sont largement utilisées pour les problèmes de classification. Ils reposent sur la théorie des perceptrons. Un ANN est composé de plusieurs neurones répartis sur une couche d'entrée (désignant les descripteurs), une couche de sortie (désignant le résultat de la classification) et un nombre de couches cachées. Par ailleurs, cette méthode est capable de modéliser, des systèmes non linéaires très complexes. Toutefois, son inconvénient, est le choix du nombre de couches cachées et la lenteur de l'entraînement.

Ainsi, l'utilisateur est amené à faire des essais avec différentes combinaisons du nombre de couches et de neurones afin, d'aboutir au réseau de neurones le plus adapté à son type d'application [27, 61]. La solution, est de réaliser un modèle connexionniste avec une seule couche cachée en association avec un algorithme métaheuristique, pour l'optimisation de l'apprentissage. C'est l'objectif de l'exploration de cette étude.

Face aux réseaux de neurones, les séparateurs à vaste marge, sont initialement conçus pour les problèmes de classification binaires. Ils permettent de séparer linéairement les exemples positifs des négatifs, dans l'ensemble de la base d'apprentissage, par un hyperplan qui garantisse un maximum de marge [63, 64].

Pour les problèmes de non-séparabilité, les SVM, effectuent une transformation non-linéaire des observations d'entrée dans un espace de dimension supérieure, pour se ramener au cas linéaire. En outre, ces modèles peuvent s'adresser également aux problèmes de classification multi-classes.

Les chercheurs en vision artificielle, ont été motivés par la perception supervisée, pour catégoriser les anomalies mammaires. Dans une des investigations de Rangayyan et al [6], la catégorisation des lésions du sein est effectuée, par les séparateurs à vaste marge, pour discerner les masses bénignes versus malignes issues, de l'atlas mammographique de l'université de Calgary (base non publique).

Cheikhrouhou [27] a accompli la classification des masses, autour des réseaux de neurones et des SVM, pour effectuer une comparaison des performances atteintes. L'application radio-diagnostique, est traitée dans ce cas, sur les mammogrammes de la base DDSM [65].

Dans les travaux de Mohamed et al [46], la classification des néoplasmes malins versus bénins de la base MIAS [66], est réalisée par trois approches, dont les SVM, les K plus proches voisins et les réseaux de neurones. L'efficacité du CAD traité, fut obtenue par les SVM avec une précision de 83%.

Dans les travaux de Kashyap et al [49, 50], la catégorisation des masses bénignes versus malignes, est accomplie via les SVM à noyau radial. Le processus est appliqué sur les mammogrammes des bases DDSM et MIAS. Le système génère une sensibilité de 94.63%, une spécificité de 94.72%, une précision de 92.02% et une AUC de 95.33%.

La classification des masses bénignes et malignes, est régentée par les SVM dans les travaux de X. Liu et al [55]. Le processus est appliqué sur 269 clichés mammaires, provenant de la base DDSM. L'efficacité du système étudié, est corroborée par une précision excédant les 90%.

Dans sa réflexion doctorale, Boulehmi [67] s'est intéressée à la discrimination des anomalies mammographiques, en classes malignes et bénignes, autour de deux classificateurs supervisés : le réseau bayésien et le neuro-flou. Le processus est appliqué sur deux bases de mammogrammes : la DDSM ainsi que la MIAS.

Dans les travaux d'Eddaoudi [57], l'ultime phase de la chaîne de vision par ordinateur, se distingue par la classification des masses autour des SVM. Cette distinction discrimine les labels malins versus bénins, des mammogrammes de la base MIAS.

Dans l'investigation scientifique de Hmida et al [52, 54], la classification des anomalies du sein, est effectuée par une méthode basée sur la théorie des possibilités, qui permet de modéliser les ambiguïtés inhérentes aux connaissances exprimées par l'expert.

Dans le contexte de notre étude, les mammogrammes analysés de la base MIAS [67] sont annotés par des experts. A cet effet, la catégorisation des néoplasmes mammaires, est diligentée par des modèles supervisés. Dans le cas de nos travaux, les réseaux neuro-génétiques et les SVM, accomplissent la discrimination des labels traités.

## 2.7 Evaluation des systèmes CAD

Les systèmes CAD peuvent-être évalués au gré de diverses métriques ; dont l'estimation de la sensibilité (cf. 2.1), de la spécificité (cf. 2.2), ainsi que de la précision (cf. 2.3). On considère à cet effet, un échantillon étiologique mammaire, où les catégories étudiées, se réfèrent à des lésions malignes et bénignes.

$$St = \frac{VP}{VP + VN} \quad 2.1$$

$$Sp = \frac{VN}{VN + FP} \quad 2.2$$

$$Pr = \frac{VP + VN}{VP + VN + FP + FN} \quad 2.3$$

Où, le paramètre évaluatif VP correspond aux classes des vrais positifs (lésions malignes détectées en tant que malignes).

La mesure FP représente les classes des faux négatifs (lésions bénignes détectées en tant que malignes),

Le facteur VN est assimilé aux classes des vrais négatifs (lésions bénignes détectées en tant que bénignes).



La métrique FN est assignée aux classes des faux négatifs (lésions malignes détectées en tant que bénignes).

La classification peut être performée aussi, en esquissant la courbe ROC. Celle-ci se schématise par le tracé de l'évolution du taux de vrais positifs (sensibilité), en fonction du taux de faux positifs (1-spécificité). L'exemple de la figure 2.9, illustre le graphique d'une courbe ROC, pour des tests ayant, des capacités de discrimination idéale, moyenne et nulle [68].

L'efficacité d'un système CAD, peut être également, validée par la matrice de confusion [69]. Dans la terminologie de l'apprentissage supervisé, cette matrice est un outil servant à mesurer la qualité d'un système de classification. Chaque colonne de la matrice représente le nombre d'occurrences d'une classe estimée, tandis que chaque ligne correspond au nombre d'occurrences d'une classe réelle (ou de référence).

Les matrices de confusion permettent, d'évaluer l'intensité de la liaison, entre les données de référence et, le résultat d'une classification.

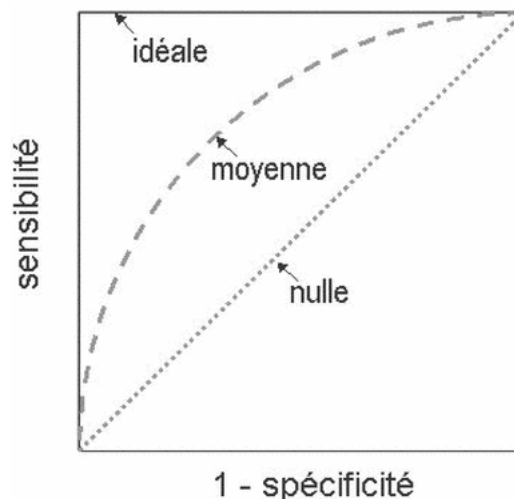


Figure 2.9 : Exemple d'une courbe ROC suivant différentes discriminations [68]

## 2.8 Systèmes d'aide à la détection assistée par ordinateur dans la mammographie

L'analyse des mammogrammes, nécessite de bien connaître leur morphogénèse, pour établir un diagnostic. Leur infinie variété, dépend de la morphologie et de l'âge. Lorsqu'une image mammographique a été repérée, l'étape qui suit consiste à la classer, suivant des critères analytiques tels que sa densité, ses dimensions et son siège et, à rechercher les autres éléments, qui régissent l'interprétation étiologique. A cet égard, le diagnostic assisté par ordinateur, constitue une seconde lecture clinique, pour analyser les mammogrammes et éviter les erreurs de diagnostic dues à la variabilité des seins. Plusieurs approches ont été entreprises à cet effet, au diapason de la perception artificielle.

Nous consacrons cette section, à une esquisse superficielle de quelques sources bibliographiques, qui ont inspiré notre vision d'aide à la décision clinique.

L'une de nos premières sources d'investigation, est basée sur les travaux de Shen et al [70]. A cet effet, les calcifications mammographiques, sont détectées en mettant l'accent sur l'analyse des formes. La localisation de ces premiers signes pathologiques, est obtenue par une segmentation adaptative en croissance de régions.

Ces régions d'intérêt sont caractérisées autour de descripteurs morphologiques, dont la compacité et les moments géométriques. Un modèle connexionniste autour du perceptron multicouche, a catégorisé les calcifications malignes versus bénignes, avec une précision de 97%.

Christoyianni et al [71] ont focalisé leur investigation, sur l'analyse des masses et des calcifications à partir des mammogrammes de la base MIAS. Un prétraitement des clichés, est effectué en premier lieu via une analyse en composantes principales, pour une réduction de la dimensionnalité. Une extraction d'attributs texturaux, tels que la moyenne et la variance sont estimés autour des régions suspectes. Le réseau neuronal à base radiale, génère la distinction entre les anomalies malignes et bénignes. Les performances établies autour de la précision, révèlent un taux de 88.23%.

Panda et al [72], ont basé leur investigation, sur les techniques automatiques du seuillage, pour l'extraction des masses et des calcifications issues de la base MIAS. Une correction des clichés mammaires est réalisée par la transformée en ondelettes, pour une performance des résultats.

Dheeba et Selvi [73] ont consacré leur recherche à la détection des anomalies issues des mammogrammes de la base MIAS, par la transformée de Gabor. La discrimination des structures malignes, bénignes et saines, est engendrée par une classification d'un réseau de neurones à architecture constructive à corrélation en cascade (CCNN). Le système génère d'intéressantes performances, autour d'une précision de 89.23%.

Li et al [74] ont conçu leur système CAD en se consacrant dans la première phase à l'élimination du muscle pectoral et des artefacts, selon une technique de seuillage automatique. La deuxième étape, a été consacrée à l'extraction d'attributs statistiques, tels que la direction du contour, la moyenne, la variance ainsi que l'entropie. L'ultime phase est la classification des structures tumorales du sein, par une analyse discriminante de la localisation commune du noyau quasi conforme (QKCLDA). Dans ce cas, le paramètre du noyau quasi-formel est automatiquement calculé, en optimisant une fonction 'objectif' qui maximise la capacité discriminante des classes. Un taux de reconnaissance de 96.98% est atteint, montrant ainsi la robustesse du système conçu.

Rizzi et al [75] ont ponctué leur recherche, autour d'une synthèse des travaux pour l'analyse des calcifications mammographiques.

L'une des explorations présentées, aborde la correction de mammogrammes, autour de transformations morphologiques. L'extraction des calcifications est conduite par les, K-moyennes. Leur classification en des labels malins et bénins, est accomplie par un réseau neuronal.

Saraswathi and Srinivasan [76], présentent une approche destinée à diagnostiquer le cancer du sein par un classifieur neuronal, à relaxation entièrement complexe. Les mammogrammes segmentés par le regroupement flou des c-moyennes, sont décrits par des paramètres texturaux et morphologiques.

La signature mammaire, est intégrée dans un réseau neuronal complexe (FCRN), pour discriminer les mammogrammes en des entités malignes, bénignes et saines. Des performances intéressantes en découlent : le système délivre une sensibilité de 97% et une précision de 98% catégorisant ainsi, les néoplasmes malins, bénins et sains des mammogrammes MIAS.

Beura et al [77] ont utilisé de la matrice GLCM pour caractériser les structures tissulaires du sein, à partir des coefficients de détails générés, par la transformée en ondelettes discrète 2 D. Les attributs significatifs sont intégrés dans un réseau neuronal, pour distinguer les mammogrammes malins, bénins et normaux. L'approche élaborée, atteste de sa conception efficiente, sur deux bases de données : la DDSM et la MIAS. Ainsi, la précision engendrée pour la distinction des différentes classes des mammogrammes MIAS, est de l'ordre de 98%.

Kaur et Kaur [78] ont développé un système automatique d'aide au diagnostic du cancer du sein. Une première étape, permet l'ajustement du contraste moyennant une fonction linéaire, qui recadre les intensités entre leurs niveaux extrêmes. Des opérations morphologiques, ont permis l'extraction des régions symboliques. Leur signature est accomplie par la supputation d'attributs texturaux. La sévérité des entités néoplasiques, est obtenue par le paradigme de la logique floue.

Sriramkumar et al [79] proposent dans leurs travaux, un système intelligent qui identifie les anomalies dans les mammogrammes, via les réseaux de neurones à base radiale et à rétropropagation ainsi que, les SVM. La localisation des structures d'intérêt est effectuée, en usant des techniques de seuillage optimum. La signature de ces entités, est obtenue via des paramètres texturaux. Le système testé sur les mammogrammes de la base MIAS, génère des résultats intéressants, particulièrement pour les SVM.

Liu et Zeng [80] ont exploré les techniques de la vision artificielle, pour une détection automatique du cancer mammaire, avec une réduction des faux positifs. En premier lieu, les zones suspectives, sont localisées par une croissance de régions et les anomalies sont extraites, par les contours actifs. La signature des mammogrammes est décrite par des facteurs géométriques et texturaux. Les entités sénologiques, sont catégorisées par les SVM, selon les classes malignes versus bénignes. L'approche élaborée, délivre des résultats encourageants.

Prathibha et Sadasivam [81] ont investigué la nature des mammogrammes, par la combinaison des coefficients d'ondelettes et de paramètres spectraux. La sévérité des anomalies de la base MIAS, est déterminée par les SVM. Le système proposé montre d'intéressantes performances révélées, par une précision de 96.26% pour la discrimination entre les échantillons normaux et malins, 92.52% pour les labels normaux versus bénins et, 90.65% pour les classes malignes versus bénignes.

Zyout et al [82], se sont focalisés sur l'extraction de paramètres texturaux multi-échelle et l'optimisation par essaim particulaire, pour la réduction des faux positifs.

Le premier processus, porte sur l'analyse des structures d'intérêt par des attributs texturaux, basés sur les ondelettes et la matrice GLCM, La seconde phase adoptée par les auteurs de cette réflexion, est la sélection des attributs pertinents via l'algorithme d'optimisation par essaim particulaire. Les connaissances mammaires sont soumises à l'apprentissage par les séparateurs à vaste marge, pour la distinction des anomalies et des cas sains. Les résultats issus de la base MIAS, montrent l'efficacité du modèle conçu.

Dans leur synthèse en ingénierie biomédicale, Casti et al [83], ont consacré leur étude, à la détection et la caractérisation des entités sénologiques mammaires. L'accent est mis sur la localisation des signes précoces, des masses et des distorsions architecturales. Dans cette optique, les auteurs se sont concentrés sur la réalisation d'un système d'aide à la détection (CADe) et au diagnostic (CADx), pour l'analyse des néoplasmes du sein. L'une des contributions originales de l'exploration des chercheurs, se focalise sur la mise en évidence des lésions à contours masqués, indépendamment de la segmentation. Ceci se fait par des descripteurs géométriques, qui quantifient la non-stationnarité et la dépendance spatiale de pixels, dans les directions, angulaire et radiale des tumeurs. La catégorisation des anomalies malignes versus bénignes, est accomplie autour d'un réseau neuronal à base radiale. Le modèle décisionnel engendré sur plusieurs bases de mammogrammes (dont la MIAS, montre des résultats efficaces, via une précision de 99%.

Kurt et al [84], se sont intéressés à la conception d'un système d'aide à la détection (CADe) et au diagnostic (CADx) des microcalcifications mammographiques. La première opération, consiste en l'extraction des régions suspectes, via le seuillage automatique d'Otsu. La seconde opération se rapporte à la caractérisation des entités néoplasiques, par des attributs texturaux, tels que, l'énergie, le contraste, la corrélation et la variance. Les connaissances mammaires sont intégrées dans un modèle connexionniste, pour la distinction des anomalies bénignes versus malignes. Les résultats issus de la base MIAS, attestent du synopsis adopté, autour d'une sensibilité de 100%.

Sadad et al [85], ont ponctué leur exploration mammographique, sur l'analyse texturale des signes suspects. Le prétraitement considéré dans cette réflexion, consiste à extraire du cliché du sein, la zone suspecte. La segmentation des structures néoplasiques, s'effectue autour d'une hybridation par l'approche des c-moyennes floues et un partitionnement par croissance en régions.

La caractérisation des lésions étiologiques se fait par l'évaluation de paramètres texturaux tels que : la corrélation, l'homogénéité, le contraste, l'entropie et la variance, estimés sur les directions  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$ . Le processus est acheminé vers l'étape de classification accordée, par plusieurs modèles ; dont les arbres de décision, les SVM et les K plus proches voisins. Les meilleurs résultats engendrés par l'application sur les mammogrammes MIAS, sont obtenus par les arbres décisionnels, avec une précision de 98.2%, une sensibilité de 100% et une spécificité de 97%.

Mata et al [86], ont concentré leur réflexion scientifique, sur la discrimination des tissus malins, bénins et sains des mammogrammes de la base MIAS.

Les régions d'intérêt, sont segmentées via la ligne de partage des eaux morphologique par marquage. Les structures symboliques, sont caractérisées par des descripteurs texturaux, assujettis à la corrélation, l'homogénéité, le contraste et l'entropie. La catégorisation des signes traités, est obtenue par les K-plus proches voisins. Le système CAD conçu, a atteint des résultats satisfaisants, révélés par une précision excédant les 83%.

Ancy et al [87], se sont concentrés sur la description texturale des tissus mammaires provenant de la base MIAS, pour une aide au diagnostic. Le synopsis adopté, se fonde en premier lieu, sur l'accentuation du contraste du cliché du sein, via une transformation gamma. Les mammogrammes rehaussés, sont segmentés selon deux phases : la première consiste à extraire la glande mammaire du fond de l'image et la seconde, est consacrée à l'extraction des signes cliniques par coupe de graphe (graph cut). Les structures néoplasiques et saines, sont caractérisés par l'estimation de quatorze paramètres texturaux. L'ultime analyse systémique, est la classification des anomalies et des tissus normaux, par les séparateurs à vaste marge. Le système CAD conçu, montre des résultats efficaces, générés par une précision de 98.11%, une spécificité de 96.43% et une sensibilité de 98.71%.

Liu et al [88] ont basé leur analyse, sur la fusion de transformations de la morphologie mathématique et, des modèles déformables pour délimiter les masses dans les mammogrammes.

Chan et Vese [89] se sont consacrés dans leur recherche, à la segmentation de l'image, autour de modèles paramétriques des ensembles de niveaux ; pour mettre en exergue, les concavités des contours. Ces modèles ont été appliqués sur les mammogrammes, pour la localisation des masses et des calcifications, dans plusieurs travaux [90-93], qui produisent des résultats fiables.

Martins et al [94] conduisent leur recherche pour la détection de masses mammographiques selon l'algorithme des K-moyennes. L'extraction des connaissances mammaires est régentée par des attributs texturaux, supputés par la matrice de cooccurrence à différentes orientations.

Le système décisionnel appliqué sur des cas réels, offre une efficacité déterminée par une précision de 85%, pour la catégorisation des masses et des mammogrammes sains, autour des SVM.

Li et al [95] ont mis au point une technique hybridant le regroupement spatial flou et les modèles actifs implicites, pour automatiser la détection de pathologies, dans les images médicales, issues des données cardiaques, cellulaires et cérébrales.

Les travaux d'Anitha et Dinesh [96], ont tenu compte du regroupement spatial flou, pour initialiser les ensembles de niveaux et délimiter ainsi, les masses dans les mammogrammes MIAS.

Dans le même ordre d'idée, ces chercheurs [97] ont profité de la propriété du noyau radial gaussien, pour l'introduire dans la fonction optimale du partitionnement flou par les c-moyennes, afin de délimiter les masses au gré d'itérations amoindries.

Les algorithmes évolutionnaires ont été utilisés dans les travaux de, Valenzuela et al [98], pour optimiser la structure neuronale, dans la classification des anomalies mammaires caractérisées globalement suivant ; des descripteurs texturaux et morphologiques. Les performances du réseau optimisé dans cette étude, ont engendré, une précision de 92%, car les opacités, n'ont pas été extraites des autres structures anatomiques mammaires.

Ludwig et al [99] ont optimisé l'apprentissage neuronal, par l'association des algorithmes génétiques et du perceptron multicouche. Cet apprentissage est institué par la maximisation de la marge, pour discriminer efficacement les classes. Les auteurs de cette étude, ont validé leur application sur trois types de bases, dont celle du cancer du sein [100]. Le synopsis adopté, permet la labellisation de ces signaux en deux classes : le sein tumoral et le sein sain.

La réflexion scientifique entreprise par Dheeba et al [101] est soulignée par la détection des anomalies tumorales et des tissus sains du sein via un réseau neuronal d'ondelettes optimisé par les particules d'essaim. L'approche conçue, considère 216 mammogrammes dont les régions suspectes, sont caractérisées par des descripteurs texturaux. L'avantage de cette étude, est de différencier les tissus normaux et les lésions tumorales, pour une aide au diagnostic. Les performances du système élaboré, montrent que l'aire sous la courbe ROC, est de 0,96853 autour d'une sensibilité de 94,167% et une spécificité de 92,105%.

Duraisamy et Emperumal [102] ont exploré les réseaux de neurones convolutifs par apprentissage profond, pour catégoriser les masses et les calcifications. La localisation des lésions du sein, est accomplie par les surfaces de niveaux. Les régions d'intérêt sont modélisées par l'algorithme des FCRN, pour la distinction des classes. Les résultats expérimentaux montrent l'efficacité du CAD, sur les mammogrammes de la base MIAS autour d'une précision de 99%, d'une sensibilité de 98.75% et d'une spécificité de 100%.



Les travaux arborés brièvement, indiquent que les critères majeurs pour la distinction des pathologies oncologiques mammaires ; sont établis autour de leur extraction des autres structures de la glande et de leur description sémiologique. Nous explorons ces travaux, dans un contexte pragmatique, lors de ce modeste prélude à la recherche, dans le monde de l'imagerie mammographique.

## 2.9 Réflexion technico-médicale d'un système d'aide à la décision clinique de mammogrammes

La réflexion émanant de notre vision et du parcours bibliographique, est ponctuée par la localisation de la zone suspecte, suivant des techniques adaptatives, afin d'amoindrir les faux positifs.

La visée attendue, est de confectionner un système (cf. Figure 2.10), permettant la détection précoce des signes cliniques ainsi que, le décèlement de néoplasmes à partir de structures tissulaires complexes. Il ressort donc de l'étude investiguée, que l'analyse concise de mammogrammes, est un enchaînement orchestré d'opérations de la VAO.

L'idée centrale, est de délimiter les régions d'intérêt, en tenant compte des différentes topologies des anomalies. L'objet de notre recherche est alors guidé, par les traitements morphologiques in situ des clichés du sein féminin.

Ces traitements se basent sur la modélisation de la forme des structures néoplasiques ou régulières, au gré de techniques dynamiques, comme les ensembles de niveaux pour l'extraction minutieuse des structures pertinentes. C'est dans cette optique, que notre réflexion s'insère, à la recherche d'approches appropriées, pour la détection des anomalies mammographiques en vue d'une aide à la décision clinique. A cet effet, l'aspect du mammogramme est pris en compte selon ses caractéristiques topologiques, géométriques et sa localisation clinique.

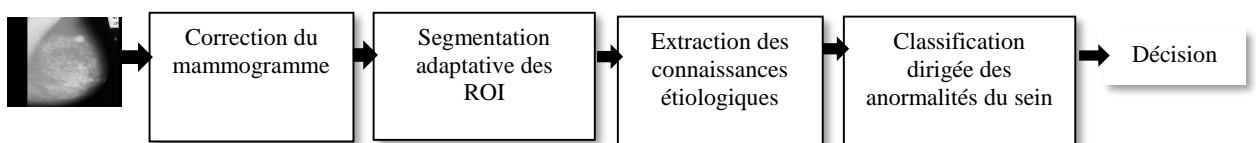


Figure 2.10 : Système d'aide à la décision de pathologies en mammographie numérisée

## 2.10 Conclusion

Il ressort de l'étude fondée par ce chapitre, que l'étape clef de la détection des néoplasmes mammaires, se distingue par la correction des mammogrammes et la localisation concise des ROI.

Dans cet ordre d'idées, notre perception humaine, est orientée vers la description analytique des lésions du sein, en s'appuyant sur des critères morphométriques spécifiques à leur sémiologie.

L'interrogation apposée et imposée par l'aspect complexe de l'organe mammaire, dirige notre vision, vers une hybridation de modèles basés sur la forme, la géométrie et la texture. Ces modèles doivent s'adapter au contenu informatif, des composants d'intérêt. Le chapitre qui suit, parcourt les principales méthodes et techniques, permettant l'amélioration du contraste mammaire et la localisation des pathologies.



---

**Chapitre 3 Correction des  
Mammogrammes et Localisation  
Morphologiques des Néoplasmes du Sein**

---

---

## 3 Correction des Mammogrammes et Localisation Morphologique des Néoplasmes du Sein

---

*« Le mot et l'idée de la recherche sont désormais agréés, même par la multitude, quand il s'agit des entreprises scientifiques » G. Duhamel.*

### 3.1 Introduction

L'aspect radiologique du sein est diligenté par le tissu mammaire, avoisinant les structures pathologiques. En effet, celui-ci s'interférant avec les zones suspectives, engendre des mammogrammes faiblement contrastés. Le clinicien est donc confronté à une intervariabilité de l'interprétation. C'est dans cette optique, que la correction de clichés mammographiques, s'avère indispensable, pour l'amélioration de la qualité de l'image. Cette correction régentera d'une manière concise, la localisation des signes pathologiques pour une aide au diagnostic.

L'adjectif majeur de ce chapitre, est le rehaussement du contraste des zones suspectives, en vue de leur segmentation morphologique, d'une manière concise.

### 3.2 Correction du contraste de mammogrammes

Les anormalités du sein, n'étant pas facilement perceptibles, du fait de la variabilité des structures tissulaires, une correction des mammogrammes destinée à les mettre en évidence, permet subtilement leur extraction. Dans ce cadre, un recours à un prétraitement approprié du cliché mammographique, s'impose pour rehausser son contraste.

Les algorithmes utilisés dans cette optique, sont basés sur une succession de techniques mathématiques (filtrage passe-haut, morphologie mathématique, modification d'histogramme...) dont la principale fonction, est d'améliorer la qualité de l'image [6, 103].

Le problème majeur des algorithmes de rehaussement du contraste [6, 83, 103], réside dans le fait que certaines régions, peuvent ne pas être améliorées convenablement, alors que d'autres, peuvent faire l'objet, d'un excès d'accentuation.

Un faible contraste, peut causer des faux négatifs (FN) générant ainsi, la négligence de plusieurs détails pertinents, qui pourraient correspondre aux premiers signes anxigènes cancérigènes.

Un renforcement excessif du contraste, peut causer des faux positifs (FP), engendrant plusieurs détails inexistantes réellement, s'adjoignant malheureusement, à la lésion. A cet égard, certaines régions normales du tissu mammaire peuvent être considérées comme anormales, ce qui fausserait le diagnostic et engendrerait par conséquent, des biopsies sans aucun intérêt.

Diverses techniques sont usitées, pour accentuer le contraste, parmi lesquelles, la modification globale de l'histogramme [6, 83, 103].

Cette méthode permet de mieux répartir les intensités, sur l'ensemble de la plage des valeurs possibles des niveaux de gris. L'opération en question, peut être réalisée, par l'égalisation d'histogramme ou par l'expansion de la dynamique.

Une autre approche proposée en tant que transformation locale, pour l'amélioration du contraste est basée, sur les opérateurs de la morphologie mathématique [46, 48-50, 83].

Les chercheurs [6, 46-50] [104-107] se sont penchés sur la question, pour rehausser le contraste, sans pour autant altérer l'aspect symbolique du mammogramme. Tous ces travaux, se focalisent sur une correction locale des mammogrammes.

Dans ce contexte, il est judicieux de les appliquer uniquement sur la zone annotée, pour n'accentuer que l'information d'intérêt et mettre en retrait, les autres constituants mammaires. C'est la réflexion visée par notre exploration.

Cette section décrit les principales techniques de mise en avant du contraste, adoptées dans le cadre de cette thèse.

### 3.2.1 Expansion de la dynamique

L'expansion de la dynamique (aussi appelé "linéarisation d'histogramme" ou "étirement de la dynamique") consiste à répartir les fréquences d'apparition des pixels, sur la largeur de l'histogramme (cf. Figure 3.1). Il s'agit ainsi, d'une opération [103-106] qui modifie l'histogramme de telle manière à répartir au mieux, les intensités sur l'échelle des valeurs disponibles. Ceci revient à étendre l'histogramme afin que, l'intensité la plus faible soit à la valeur minimale et que la plus haute, soit à la maximale.

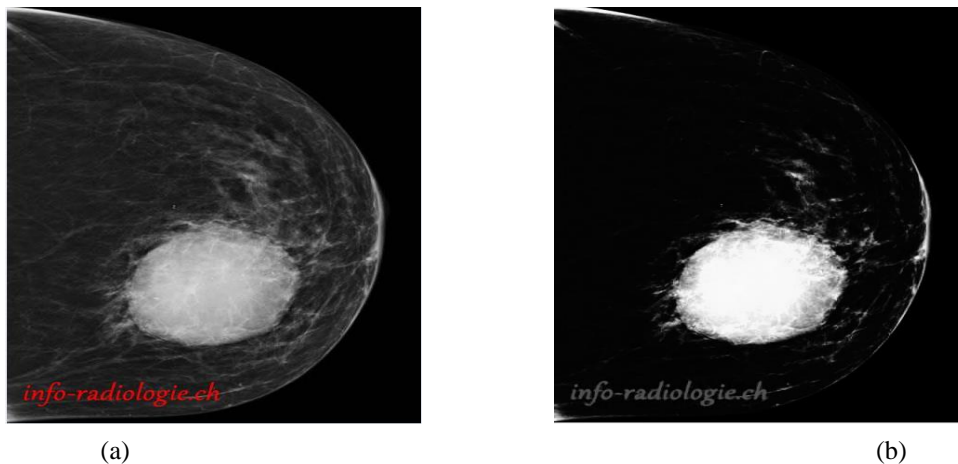


Figure 3.1 : Image mammographique rehaussée par expansion de la dynamique

(a) Image originale [22]

(b) Image rehaussée

De cette façon, si les valeurs de l'histogramme sont très proches les unes des autres, l'étirement fournit une meilleure répartition des intensités, afin de rendre les pixels clairs encore plus clairs et les pixels foncés, proches du noir.

L'expansion de la dynamique, consiste en la normalisation de l'histogramme, suivant une transformation qui affine les niveaux de gris de telle sorte que, l'image utilise toute la dynamique de représentation (cf. Equation 3.1).

La transformation choisie, permet de conserver les contrastes relatifs entre les régions (une région claire, apparaîtra plus claire que le fond, dans l'image transformée). Cette opération, améliore la qualité visuelle de l'image.

$$f_c = (f - N_{\min}) \frac{255}{N_{\max} - N_{\min}} \quad 3.1$$

Où,  $f$  est l'image originale et  $f_c$  l'image améliorée.

$N_{\min}$ , représente, le minimum des intensités de gris et  $N_{\max}$  correspond au maximum des niveaux de gris dans l'image d'origine  $f$ .

### 3.2.2 Correction morphologique

La perception morphologique mammaire, impose une mise en relief des néoplasmes, au gré de leurs formes et leurs natures. Le paradigme mathématique, de Matheron et Serra [108, 109] répond à ces soucis cliniques, en fonction d'aspects techniques. Rechercher une correspondance entre les structures des formes semblables, telle est la visée du concept de Matheron et Serra.

L'analyse d'images par la morphologie mathématique, remonte aux années soixante. Cette théorie a été initialement introduite en sciences des matériaux par J. Serra et G. Matheron pour analyser des objets, au travers de leur texture [108]. Elle a connu durant ces dernières décennies, une grande évolution couvrant ainsi, un large éventail d'applications ; touchant notamment à la robotique et la vision industrielle, l'imagerie médicale ou encore, le multimédia [110].

La morphologie mathématique, est une théorie qui modélise la structure et la forme de l'image. Elle offre un grand nombre d'outils très puissants, pour l'analyse du signal bidimensionnel. Les opérateurs proposés ont été développés au départ, pour traiter des images binaires : c'est la morphologie mathématique ensembliste. Leur utilisation a été ensuite, étendue aux images en niveaux de gris : la morphologie mathématique fonctionnelle est alors née.

#### 3.2.2.1 Élément structurant

L'analyse d'une image par morphologie mathématique consiste à appliquer sur cette image, un ou plusieurs opérateurs morphologiques au moyen d'un masque ou motif géométrique prédéfini de taille et de formes données (carré, rectangle, cercle...). Le motif géométrique (cf. Figure 3.2), appelé élément structurant (ES), est déplacé dans l'image de façon à ce que son centre occupe toutes les positions de l'espace. Pour chacune des positions, on applique l'opérateur choisi. Il en résulte une image transformée, dans laquelle des éléments, peuvent avoir été éliminés, fusionnés ou, amincis. En fonction de la réponse, on construit un ensemble de sortie. On obtient ainsi deux opérations de base, l'érosion et la dilatation à partir desquelles, d'autres transformations sont générées.

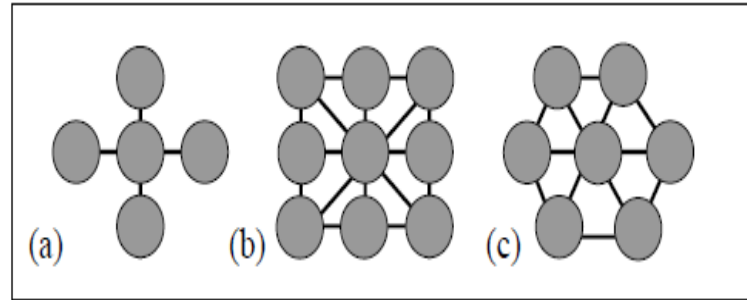


Figure 3.2 : Eléments structurants plans sur des trames carrée (a) en losange (b) et hexagonale (c) [111]

### 3.2.2.2 Erosion binaire

Selon la définition des précurseurs de la morphologie mathématique [111,112], l'érosion consiste à faire déplacer un élément structurant  $B$ , de telle sorte que son centre occupe toutes les positions  $x$  de l'ensemble  $X$ . Ainsi, pour chaque position, la question, est-ce que  $B$  est complètement inclus dans  $X$ , se pose ? L'ensemble des positions  $x$  correspondant à une réponse positive, forme alors, l'érodé de  $X$  par  $B$  (cf. Equation 3.2).

$$Ero = X \ominus B = \{x \text{ dans } X : B \subset X\} \quad 3.2$$

### 3.2.2.3 Erosion fonctionnelle

L'extension aux images à teintes de gris, assimile l'érosion (cf. Figure 3.3) à la valeur minimale de l'image contenue dans la fenêtre définie, par l'élément structurant lorsque celui-ci, est centré en  $x$ .

Soient à cet effet,  $x$  un point d'une image donnée,  $f$  la fonction qui, à  $x$  associe son niveau de gris et  $B_x$ , le voisinage de  $x$  défini par l'élément structurant  $B$  centré en  $x$ . L'érosion fonctionnelle répond à la relation 3.3.

$$Ero(f(x)) = \inf \{f(u), u \in B_x\} \quad 3.3$$

Après une érosion les objets traités, subissent des modifications au gré du masque adopté selon les points ci-dessous :

- la surface des structures de taille supérieure à l'élément structurant, est réduite ;
- les objets de taille inférieure à celle de l'élément structurant disparaissent ;
- 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 inférieure, à celle de l'élément structurant, vont être séparés ;
- l'érosion réduit les pics de niveaux de gris et élargit les vallées. Elle tend donc à assombrir l'image et à étaler, le bord des objets les plus sombres.

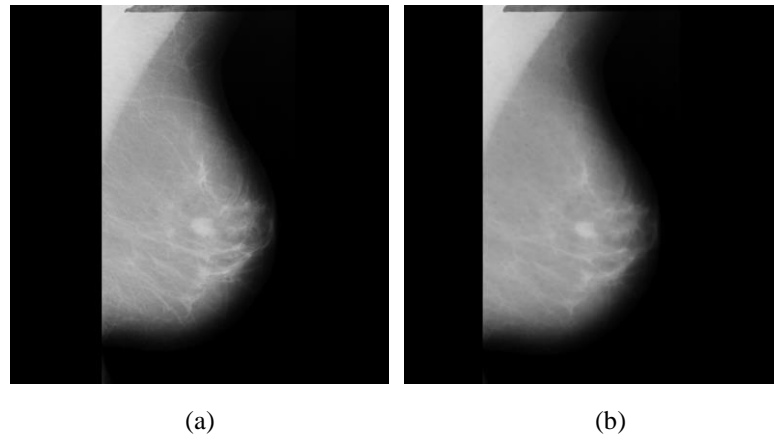


Figure 3.3 : Erosion d'une image mammographique (avec un disque de taille=5)  
 (a) Image originale [66] (b) Image érodée

### 3.2.2.4 Dilatation binaire

L'opération de dilatation est la transformée duale de l'érosion. En prenant le même élément structurant B que pour l'opération précédente, on pose pour chaque point x la question "B touche-t-il l'ensemble X ? C'est à dire, existe-t-il une intersection (cf. Equation 3.4) non vide entre B et X ?

$$Dil = x \oplus B = \{x \in X : B \cap X \neq \emptyset\} \quad 3.4$$

### 3.2.2.5 Dilatation fonctionnelle

L'extension aux images en niveaux de gris, confère la dilatation des points x, aux valeurs maximales de l'image contenues dans la fenêtre définie par l'élément structurant lorsque celui-ci est centré en x.

En se référant aux mêmes hypothèses que pour l'érosion fonctionnelle, son opération duale (cf. Figure 3.4) se formule selon la relation 3.5.

$$Dil(f(x)) = \sup\{f(u), u \in B_x\} \quad 3.5$$

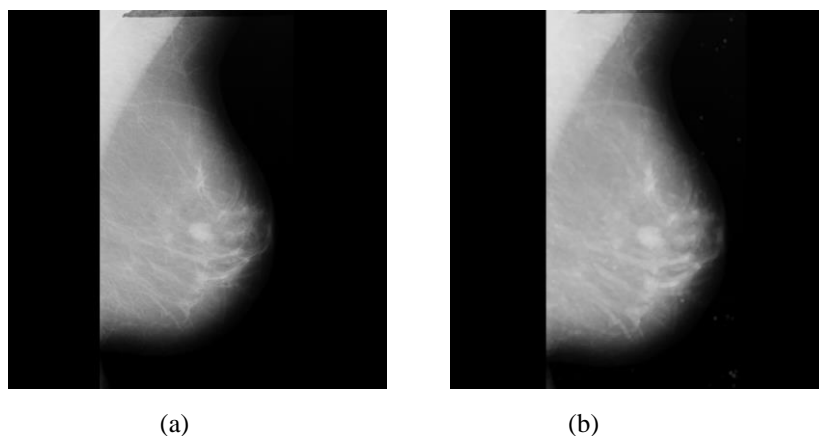


Figure 3.4 : Dilatation d'une image mammographique (avec un disque de taille=5)  
 (a) Image originale [66] (b) Image dilatée

La dilatation produit plusieurs effets caractéristiques :

- les trous de taille inférieure à l'élément structurant sont colmatés ;
- la surface des structures est augmentée ;
- les structures voisines séparées d'une distance inférieure à l'élément structurant, sont fusionnées.

### 3.2.2.6 Ouverture binaire

L'érosion et la dilatation étant 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  $X$  par  $B$  (cf. Equation 3.6).

$$O_B(X) = (X \circ B) \quad 3.6$$

### 3.2.2.7 Ouverture fonctionnelle

L'opération d'ouverture fonctionnelle (cf. Equation 3.7) répond à même définition que dans le cas ensembliste.

$$O_B(f) = Ero(f(x))Dil(f(x)) \quad 3.7$$

L'ouverture supprime les pics, mais préserve les vallées. Elle homogénéise l'image, mais conserve les objets sombres (cf. Figure 3.5).

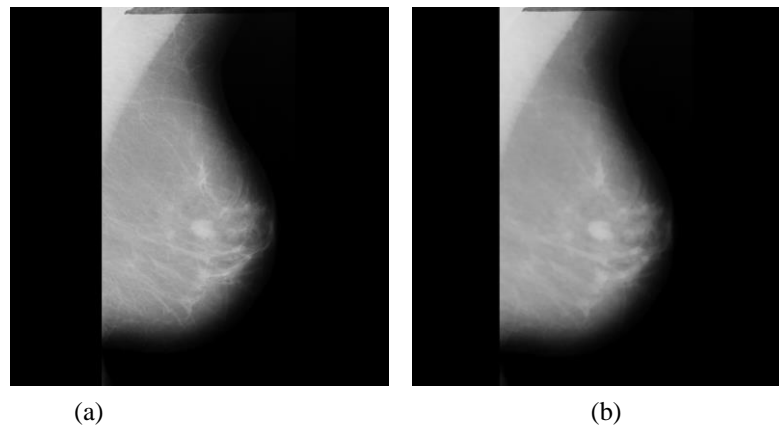


Figure 3.5 : Ouverture d'une image mammographique (avec un disque de taille=5)

(a) Image originale [66]

(b) Ouverture de l'image

### 3.2.2.8 Fermeture binaire

La fermeture est l'opération duale de l'ouverture. Elle est réalisée (cf. Equation 3.8) par une dilatation, suivie d'une érosion (toujours en gardant le même élément structurant). Le résultat est nommé, le fermé de  $X$  par  $B$ .

$$F_B(X) = (X \bullet B) \quad 3.8$$

### 3.2.2.9 Fermeture fonctionnelle

Tout comme la fermeture ensembliste, la fermeture fonctionnelle, est engendrée par une opération de dilatation, suivie d'une érosion.

$$O_B(f) = Dil(f(x))Ero(f(x)) \quad 3.9$$

La fermeture comble les vallées et éclaircit l'image (cf. Figure 3.6).

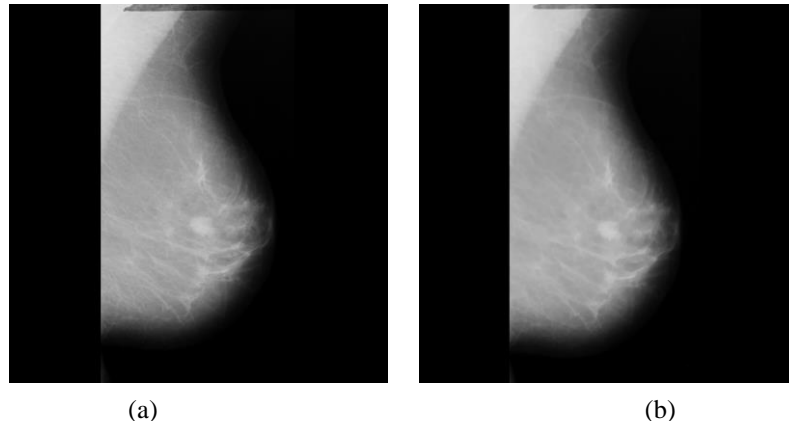


Figure 3.6 : Fermeture d'une image mammographique (avec un disque de taille=5)  
 (a) Image originale [66]      (b) Fermeture de l'image

### 3.2.2.10 Transformation du chapeau haut de forme

Les propriétés de la fermeture et l'ouverture en teintes de gris, permettent de définir des filtres, permettant d'extraire les pics de la fonction  $f$  correspondant aux petites zones claires de l'image ou aux vallées, correspondant aux petites zones sombres [110, 111]. Ces filtres symbolisant les résidus des opérations de base de la morphologie mathématique, sont appelés les chapeaux haut ou bas de forme.

Le chapeau haut de forme [108-111] est une transformation qui permet, de retrouver les informations éliminées par l'ouverture morphologique. On définit la transformation du chapeau haut de forme, comme la différence entre la fonction de départ  $f$  et l'ouvert (cf. Equation 3.10).

$$T_{CHF}(f) = f - foB \quad 3.10$$

$T_{CHF}$  correspond au chapeau haut de forme.  $B$  est l'élément structurant et le symbole  $o$  représente l'ouverture morphologique.

Il s'agit de la transformation morphologique (cf. Figure 3.7) permettant d'extraire sur une image, les structures contrastées et de faible épaisseur. C'est dans ce cadre, que ce filtre est utilisé pour mettre en exergue, les calcifications mammographiques. Plusieurs chercheurs, s'y sont attelés [45, 83].



### 3.2.2.11 Transformation chapeau bas de forme

Pour extraire les vallées, l'opération du chapeau bas de forme (cf. Equation 3.11) y répond au gré de la différence, entre la fermeture de l'image par un élément structurant  $B$  et l'image initiale.

$$T_{CBF}(f) = f \bullet B - f \quad 3.11$$

$T_{CBF}$  correspond au chapeau bas de forme. Le symbole  $\bullet$  dénote la fermeture morphologique.

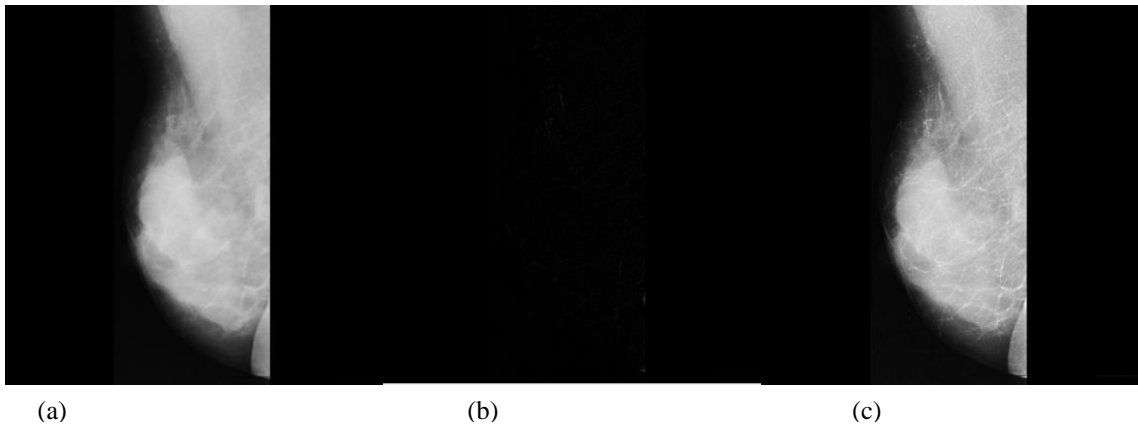


Figure 3.7 : Transformation du chapeau haut de forme sur une image mammographique (avec un disque de taille=5)

(a) Image originale [66] (b) chapeau haut de forme de l'image (c) superposition de (a) et (b)

## 3.3 Détection morphologique des lésions suspectives

La segmentation, est la pierre angulaire de l'analyse de l'image. Cette opération a pour but de séparer différentes zones homogènes d'une image, afin d'organiser les objets en régions ou groupes (clusters) dont les pixels, ont en commun diverses propriétés (intensité, couleur, texture, etc...).

Il s'agit d'un traitement de bas niveau, qui consiste à créer une partition [103, 112] d'une image  $A$  en régions  $R_i$ , telles que :

$$\begin{aligned} \forall i R_i &\neq \phi \\ \forall i, j; i \neq j R_i \cap R_j &= \phi \\ A &= \bigcup_i R_i \end{aligned} \quad 3.12$$

Une région est un ensemble connexe de pixels, ayant des propriétés communes (intensité, texture, couleur...), 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.

A cet effet, 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. Le choix d'une technique est lié à la nature de l'image, aux opérations situées en aval de la segmentation et aux primitives, à extraire. Du fait de cette diversité, il est difficile de définir une méthode de segmentation standard. Cependant, en s'inspirant des systèmes de la vision humaine, la segmentation fait référence, aux notions de différence et de similarité. Il en découle quatre approches courantes : l'approche par frontière, l'approche par région, l'approche par classification et l'approche hybride.

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 discontinuité entre les propriétés de deux ensembles, de points connexes.

L'approche par classification, considère une région comme un ensemble de pixels connexes, appartenant à une même classe. Elle suppose donc, que les pixels qui appartiennent à une même région, possèdent des caractéristiques similaires et forment un nuage de points dans l'espace des attributs.

L'approche coopérative profite des avantages, de deux premières approches. Elle combine à la fois, l'information issue des régions et l'information issue des contours.

L'objectif de cette section, n'est pas de dresser une liste exhaustive des méthodes de segmentation, mais juste celles qui ont éclairé notre approche. L'accent est mis, sur les modèles dynamiques, les techniques de partitionnement par regroupement spatial flou et la croissance de régions.

### **3.3.1 Segmentation par croissance de régions**

Le partitionnement par croissance de régions, est une technique ascendante qui, partant de la représentation de l'image comme un ensemble de pixels, les regroupe selon un double critère d'homogénéité et d'adjacence. La procédure débute par le choix de germes dans l'image, à partir desquels des pixels ou des régions, s'agrègent à condition qu'ils vérifient le critère de similarité établi [103].

Pour segmenter une image selon une croissance de régions, il faudrait se soumettre à certaines conditions notifiées par ; le positionnement des germes initiaux, la sélection du prédicat d'homogénéité et le critère d'arrêt.

#### **3.3.1.1 Principe général du processus de la croissance de régions**

Le synopsis de la croissance de régions (cf. Figure 3.8), se compose d'une phase d'initialisation et d'un processus itératif, pour l'obtention du partitionnement de l'image, en entités symboliques.

##### **1. Phase d'initialisation**

L'étape d'initialisation consiste à sélectionner les germes, à partir des régions les plus significatives de l'image.

En effet, l'emplacement des noyaux initiaux, conditionne fortement le résultat de la segmentation. Ces germes s'accroissent dans la deuxième phase du traitement, en respectant un critère d'homogénéité.

## 2. Phase d'itération

Dans la croissance de régions, la condition d'agglomération implique la définition d'un critère de similarité, entre un point candidat et, la région segmentée.

Ce terme de ressemblance, est utilisé par le prédicat pour décider de l'ajout ou non d'un pixel. A cet effet, le critère implique une mesure d'homogénéité, sur les intensités des points de la région segmentée. Cette évaluation mesure l'homogénéité locale entre le germe initial et un pixel candidat, ou bien globale entre les points constituant, la région et un pixel candidat. Ainsi, pour décider de l'insertion ou non d'un pixel à la région germe, le critère de similarité, est comparée à une valeur seuil. Lorsque le processus converge, la méthode choisit un autre germe parmi les points non segmentés, pour réitérer l'algorithme jusqu'à épuisement des régions susceptibles, d'être fusionnées.

Le choix du seuil a une forte influence, sur le résultat de la segmentation. Une mauvaise sélection de ce paramètre, peut faire croître une région, à l'extérieur de la région à segmenter.

La Figure 3.9 présente un résultat de la segmentation par croissance de régions, sur une image synthétique. Le processus a été initialisé, à partir de deux germes. Le résultat présenté, parcourt différentes itérations, de l'évolution.

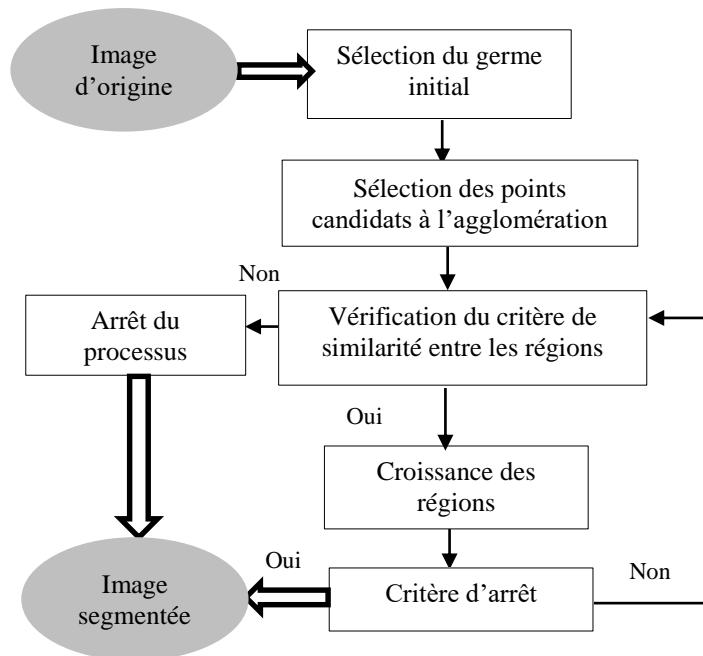


Figure 3.8 : Synopsis général du processus de la croissance de régions [103]

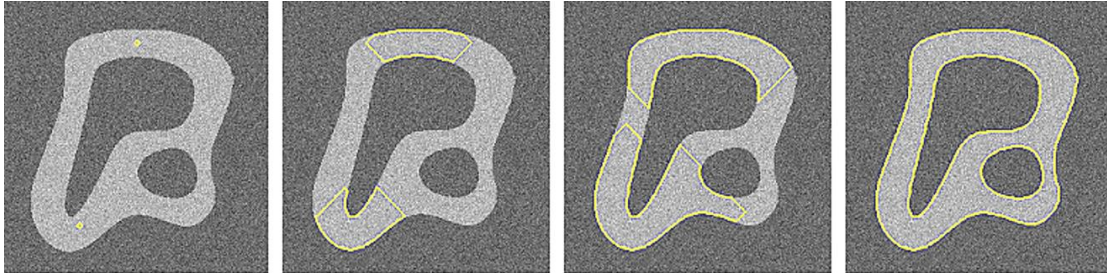


Figure 3.9 : Segmentation d'une image synthétique en niveaux de gris  
(b)-(d) : évolution de la segmentation à partir de deux germes provenant de (a) [103]

### 3.3.1.2 Convergence du processus itératif

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 le prédicat et la partition comprend  $n$  régions ;
- il existe  $n-1$  régions qui vérifient le prédicat, la  $n$ ème région inclut, les points ne le vérifiant pas.

### 3.3.1.3 Propriétés de la croissance de régions

La croissance de régions a souvent fait l'objet, des processus de la segmentation des images mammographiques [6, 83]. Son principal atout, est qu'il s'agisse d'une méthode simple, à mettre en œuvre. Elle présente néanmoins des limitations.

Les principaux avantages et inconvénients de ce type de segmentation [103], peuvent être regroupés ci-dessous :

1. Une préservation de la forme, de chaque région de l'image.
2. Une mauvaise sélection des germes ou un choix du critère de similarité mal adapté, peuvent entraîner des phénomènes de sous-segmentation ou de sur-segmentation.
3. La sensibilité de la croissance de régions, aux points de fuite. Un point de fuite se définit par la connexion de deux régions distinctes aux propriétés statistiques semblables.

L'application de ces méthodes aux images mammographiques, a montré de bonnes performances vis-à-vis de la détection des masses et des calcifications, au gré de prédicats adaptés [6, 83, 113].

La figure 3.10, montre le résultat de cette application, sur des images mammographiques selon les travaux de Punitha et al [113]. L'extraction des masses, est obtenue par une segmentation en croissance de régions améliorée, par les techniques d'optimisation par essaim.

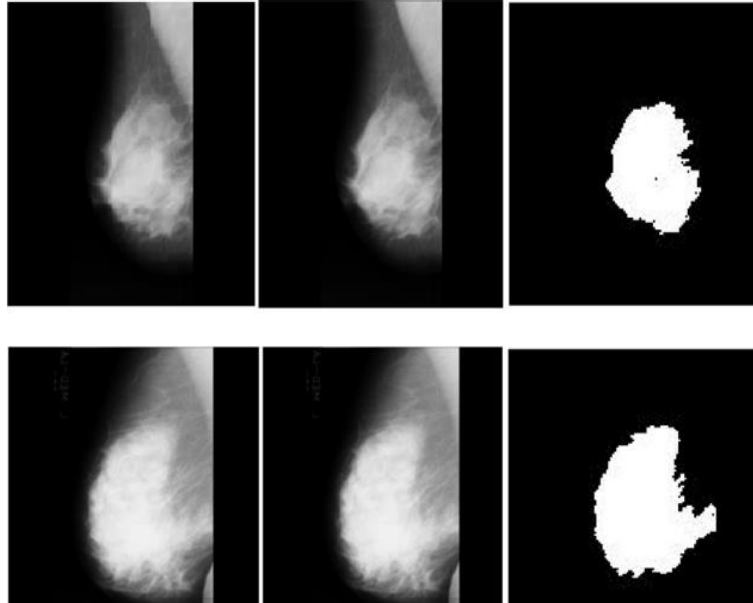


Figure 3.10 : Application de la croissance de régions sur les images mammographiques [113]

Dans le cas de notre prélude à la recherche, le germe initial est institué par le centre de gravité de la zone suspectieuse. Les pixels répondant à la 8-connexité autour du noyau, sont agrégés selon le critère d'homogénéité de la relation 3.13. A cet égard, le processus s'arrête lorsque la différence de l'intensité moyenne entre la région et un pixel candidat, excède la valeur empirique du seuil (S).

$$|I(x, y) - M| \leq S \quad 3.13$$

Où,

$I(x, y)$  correspond à l'intensité du pixel traité.  
 $M$  représente la moyenne de la région.

### 3.3.2 Segmentation par les modèles déformables

Les contours actifs, furent inspirés des modèles élastiques, issus des travaux de Bur et al [114]. Néanmoins, leur véritable insertion dans le monde de l'imagerie, revient à l'équipe de Kass et al [115], pour le partitionnement dynamique de l'image.

Depuis leur apparition vers la fin des années 80, les contours actifs, ont conquis le monde de la segmentation. L'engouement pour ces approches variationnelles, suscite l'intérêt de la communauté scientifique, pour pallier aux limitations causées par les méthodes classiques, notamment en imagerie médicale.

Initialisé proche du contour, le modèle déformable appelé aussi, le 'snake', évolue itérativement, de manière à s'ajuster au contour de la région à détecter. Cette évolution dynamique, est appliquée sous la contrainte de conserver, certaines caractéristiques du contour ; telles que la courbure, ou la répartition des points du contour.

Ces approches variationnelles se scindent en deux grandes classes : les contours actifs paramétriques et les contours actifs géométriques.

Explicitement, les approches paramétriques déplacent un ensemble de points prédéfini, basé sur la minimisation d'une fonctionnelle d'énergie. Tandis que les modèles géométriques (ensembles de niveaux), déplacent implicitement des contours, comme un niveau particulier d'une fonction [116].

La mise en œuvre de ces modèles [115-119], pour une application donnée, répond aux étapes ci-dessous :

- Choix de la représentation continue du contour actif (paramétrique ou implicite).
- Formulation de la fonctionnelle d'énergie. Cette étape indique que la fonctionnelle, doit être construite de telle sorte, qu'un minimum local soit associé, à la frontière de l'objet à détecter.
- Obtention de l'équation d'évolution. Cette phase fait intervenir le calcul de la dérivée variationnelle, de la fonctionnelle (méthode d'Euler Lagrange). Cette équation consiste sur le plan formel, en une équation aux dérivées partielles (EDP).
- Implémentation numérique de l'équation d'évolution. La résolution numérique de cette EDP, est le plus souvent réalisée par des techniques de différences finies.

Tout comme la segmentation d'images classique, les modèles déformables se distinguent ; par les contours actifs basés région [89, 120], ceux orientés contours [114] ainsi que les approches hybrides [112].

L'objet de cette section, n'est pas de faire un état de l'art exhaustif sur les approches variationnelles, mais de se pencher explicitement sur les modèles adoptés.

### 3.3.2.1 Contours actifs paramétriques

Les contours actifs ou les snakes tiennent leur nom, de leur capacité à se déformer comme un serpent. L'approche en question, consiste à faire évoluer un contour  $C$  itérativement au cours du temps, de sa position initiale vers les bords de l'objet à extraire, en minimisant une fonctionnelle d'énergie (cf. Figure 3.11).

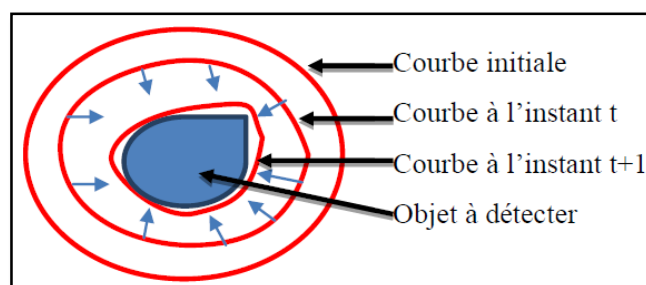


Figure 3.11 : Modèle graphique du contour actif [112]

La courbe initiale se déforme, progressivement selon l'action de plusieurs forces, qui vont la tirer ou la pousser, vers l'objet à extraire.

Ces forces [112, 115-118] 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 souhaite obtenir.

Considérons à cet effet, un contour  $C$  appartenant à une image  $I$ , qui est déformable spatialement et temporellement. Cette courbe est paramétrée par l'abscisse curviligne  $s$  et le temps  $t$ .

Soit  $V(s, t)$  la position d'un point de la courbe  $C$  à l'instant  $t$  (cf. Equation 3.14).

$$V(s, t) = (x(s, t), y(s, t)), s \in [0, 1] \quad 3.14$$

L'énergie  $E$  est alors exprimée par l'équation 3.15 :

$$E(V) = E_{int}(V) + E_{ext}(V) + E_{cont}(V) \quad 3.15$$

### 1. Énergie interne

L'énergie  $E_{int}$  (cf. Equation 3.16), caractérise les propriétés physiques du contour et assure la régularité de la courbe [115, 117].

$$E_{int} = \int_V \left( \alpha(s) |V'(s, t)|^2 + \beta(s) |V''(s, t)|^2 \right) ds \quad 3.16$$

Où,  $v'(s)$  et  $v''(s)$  dénotent respectivement les dérivées, première et seconde de  $v(s)$ .

$ds$  correspond à  $dx dy$ .

Le terme du premier ordre, représentant l'élasticité du contour, est contrôlé par le paramètre  $\alpha(s)$ . Une valeur très grande de ce facteur, limitera l'étirement du contour tandis qu'une discontinuité, pourra être présente dans le cas, où ce paramètre est nul.

Le terme du second ordre, correspondant à la raideur du contour, est contrôlé par le paramètre  $\beta(s)$ . Une valeur très grande de ce paramètre, empêche une forte courbure du contour. Dans le cas où  $\beta$  est nul, la courbe peut être discontinue au second ordre et contenir des coins.

### 2. Énergie externe

L'énergie interne gère la régularisation du contour. L'énergie externe, quant à elle, correspond à l'adéquation aux données.



Elle fait alors, intervenir les caractéristiques de l'image. Comme ce sont les contours recherchés, ce sont donc les points fort du gradient, qui sont pris en compte (cf. Equation 3.17).

$$E_{cont}(V) = - \int_V |\nabla I(V(s,t))|^2 ds \quad 3.17$$

Où,  $\nabla I(V(s,t))$  est le gradient de l'image  $I$  en  $V(s,t)$ .

Dans la plupart des cas, on calcule en fait le gradient de l'image lissée :

$$E_{ext}(V) = - \int_V |\nabla(g_\sigma * I)(V(s,t))|^2 ds \quad 3.18$$

Où  $g_\sigma$  est une gaussienne centrée d'écart type  $\sigma$ .

Un minimum d'énergie sera atteint, si la courbe passe par les points de gradient maximal de l'image. Le signe négatif indique que les points les plus forts du gradient, minimiseront cette énergie et attireront le snake vers le contour de l'objet.

Un résultat efficace de l'évolution de la courbe, dépend de l'initialisation, qui doit être proche de la solution.

### 3. Energie de contexte

L'énergie de contexte permet d'introduire des connaissances a priori, pour l'obtention du snake souhaité. L'énergie la plus connue et aussi la plus utilisée, est celle du 'ballon' introduite dans les travaux de Cohen et al [118]. Elle permet de gonfler ou de rétracter le contour actif, selon le signe choisi.

$$E_{cont} = \pm K_1 \int_C dx dy \quad 3.19$$

Où  $K_1$  est un paramètre de la courbure fixé par l'utilisateur.

Cette énergie présente l'avantage, de transcender les zones à faible gradient, afin de pallier à un des problèmes des contours actifs paramétriques, leur sensibilité au bruit. La convergence vers les frontières fortement marquées, est ainsi atteinte.

La fonctionnelle d'énergie ainsi définie, est ensuite minimisée, en effectuant une descente de gradient, pour aboutir aux équations d'Euler-Lagrange qui déterminent l'évolution du contour [112].

Les contours paramétriques présentent des avantages remarquables ; notamment en termes de structure (courbe fermée) et de qualité, des contours obtenus.



Néanmoins, les performances des snakes, restent très dépendantes de la paramétrisation du contour, qui peut s'avérer délicate. Malgré les forces additionnelles, ils risquent d'être attirés, vers des minimas locaux. A cet égard, les différents termes de la fonctionnelle d'énergie doivent être réglés (paramètres de régularisation), avec soin pour que leurs contributions s'équilibrent. Cela rend donc l'utilisation de cette méthode, difficile dans le cas d'images de mauvaise qualité, ou pour la segmentation de structures, présentant une forte variabilité.

Un des principaux inconvénients des snakes, est leur incapacité à traiter les changements de topologie, en raison de l'approche lagrangienne qui est généralement, utilisée pour leur implémentation [117, 120]. Des approches géométriques révélées par les courbes de niveaux, permettent la segmentation d'objets à topologie variable.

### 3.3.2.2 Contours actifs géométriques

Le principe des contours actifs géométriques, ou implicites connus par les ensembles de niveaux, est inspiré de la théorie de la propagation des courbes subissant une force normale [112, 116], comme cela est illustré sur la figure 3.12.

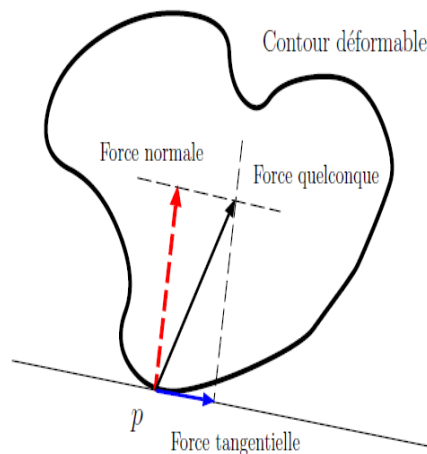


Figure 3.12 : Propagation d'une courbe selon une force normale [112]

Introduites initialement en physique par Osher et Sethian [116], les méthodes des ensembles de niveaux, permettent de résoudre des problèmes de propagation de front d'onde dépendant du temps. Cette conception géométrique, repose sur l'équation de la diffusion de la chaleur [117, 120].

L'idée de la représentation implicite, d'Osher et de Sethian, consiste à insérer le contour, dans un ensemble de dimension supérieure : c'est l'ensemble de niveau, dont la déformation répond au changement de la topologie du contour. Ainsi, dans le cas bidimensionnel, le contour actif est une courbe plane fermée. L'ensemble de dimension supérieure, dans lequel on insère le contour, est une surface. Par conséquent, le contour n'est autre que l'intersection de son plan, avec cette surface (cf. Figure 3.13). L'altitude des points immergés de la surface est négative, alors que celle des points émergés, est positive. Le contour de l'image traitée, est donc composé de pixels de valeur nulle. Plus précisément, si un pixel est à l'extérieur du contour, sa valeur est positive.

Si au moins un de ses voisins, est à l'intérieur de la courbe, sa valeur est négative. Par conséquent, le contour de niveau zéro, se situe entre les deux [120-123].

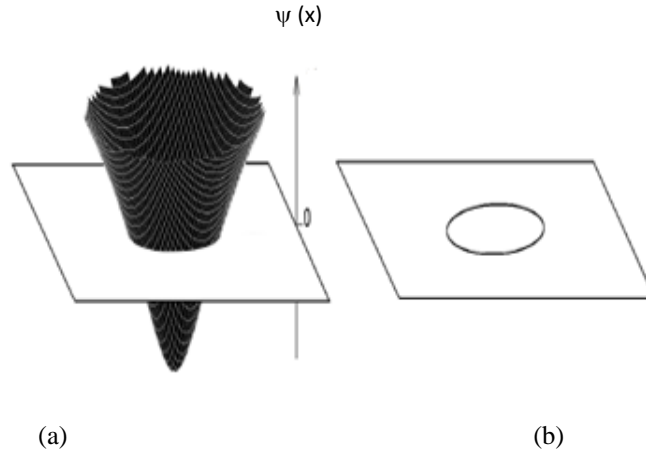


Figure 3.13 : Représentation de  $\psi(x, t)$  et de son intersection avec le plan de niveau zéro [123]  
 (a) fonction d'ensembles de niveaux  $\psi(x, t)$  (b) Contour  $C$  déterminé par l'intersection de  $\psi(x, t)$  et du plan de niveau zéro.

Un ensemble de niveaux établi à partir d'un contour  $C(x, y, t)$ , fermé, régulier et évoluant dans le plan euclidien, est une fonction  $\psi(x, y, t)$  temporelle. Ainsi, la courbe  $C$  est définie comme l'ensemble des points vérifiant la condition ci-dessous :

$$C(t) = \{x / \Psi(x, t) = 0\} \quad 3.20$$

Selon la théorie d'Osher et de Sethian, l'initialisation de  $\psi(x, y, t)$  dans le cas bidimensionnel, correspond à la fonction distance lipschitzienne signée, qui sépare un point  $x$  du contour  $C$ . A cet effet, la fonction des ensembles de niveaux, est assimilée à la relation 3.21.

$$\Psi(x, t) = \begin{cases} 0 & \text{si } x \in C(t) \\ \varepsilon d(x, C(t)) & \text{si } x \in \Omega_{\text{int}} \\ -\varepsilon d(x, C(t)) & \text{si } x \in \Omega_{\text{ext}} \end{cases} \quad 3.21$$

Où,  $\Omega_{\text{int}}$  et  $\Omega_{\text{ext}}$  représentent respectivement les régions, intérieure et extérieure délimitant la courbe  $C$ .

$d(x, C(t))$  représente la distance du point  $x$  à la courbe  $C(t=0)$ . Elle équivaut à  $\inf_{y \in C(t)} \|x - y\|$ .

Le paramètre  $\varepsilon = 1$  ou  $-1$ , selon la convention choisie (dans notre cas,  $\varepsilon = 1$ ).

La relation 3.21 engendre par conséquent, la formulation 3.22 :

$$C(t=0) = \{x \in R^2 / \Psi(x,t) = 0\} \quad 3.22$$

La figure 3.14 exprime bien, le paradigme des ensembles de niveaux. Dans ce cas précis, à l'instant  $t=0$ , la courbe  $C$  est représentée par un cercle unité dans le plan  $xy$  (cf. Figure 3.14.a). Ce cercle est donc considéré, comme l'ensemble de niveau  $\psi = 0$  d'une surface initiale  $z = \Psi(x, y, t=0)$  dans  $R^3$  (cf. Figures 3.14.b, c et d). A cet effet, pour tout point  $x$ , la valeur absolue de  $\psi$ , correspond à la distance minimale au contour.

D'un point de vue géométrique, la forme d'un ensemble de niveau bidimensionnel, signifie que la hauteur de la surface en un point donné, n'est autre que la distance signée relativement au contour initial issu de la projection de ce point ; sur le plan assigné à la fonction  $\psi=0$ .

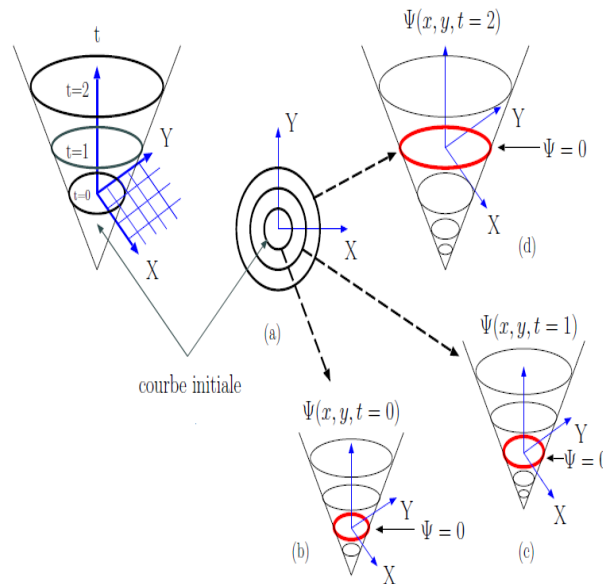


Figure 3.14 : Evolution d'un contour au gré de l'ensemble de niveau  $\psi$  [112]  
 (b) Courbe initiale (b), (c) et (d) Evolution de la courbe

➤ **Evolution numérique des ensembles de niveaux**

L'évolution numérique des ensembles de niveaux a été mise en œuvre, par Caselles et al [119] suivant les équations aux dérivées partielles (EDP). Dans ce contexte, chaque point du contour, se déplace selon la normale (cf. Figure 3.12), à un point  $p$  suivant une vitesse proportionnelle, à la courbure. Ceci se traduit par l'équation 3.23, où la fonction tangentielle n'apparaît pas, car elle n'affecte que la paramétrisation du contour.

$$\frac{\partial C(p,t)}{\partial t} = F\vec{N} \quad 3.23$$

Où  $C$ , est le contour déformable,

$F$  indique la force ou la vitesse d'évolution,

$\vec{N}$  correspond à la normale à la courbe  $C$  au point  $p$  et au temps  $t$ .

L'évolution de la courbe est obtenue, à partir de l'évolution de la carte de distance  $\psi$ . La vitesse d'évolution étant considérée comme normale au contour, l'équation 3.23 se traduit alors, par la relation 3.24.

$$\frac{\partial \Psi}{\partial t} = F |\nabla \Psi| \quad 3.24$$

La relation 3.22 traduit le fait que la courbe  $C$  à l'instant  $t$ , est au niveau 0. La dérivée de cette fonction par rapport à  $t$ , engendre l'équation ci-dessous :

$$\frac{\partial \psi(C, t)}{\partial t} + \nabla \Psi(C, t) \frac{\partial C}{\partial t} = 0 \quad 3.25$$

Nous obtenons ainsi la formulation 3.26 :

$$\frac{\partial \psi(C, t)}{\partial t} = -\nabla \Psi(C, t) F \vec{N} \quad 3.26$$

Sachant que ;  $\vec{N} = \frac{\nabla \Psi}{\|\nabla \Psi\|}$  ; on obtient l'équation 3.27.

$$\frac{\partial \Psi(C, t)}{\partial t} = -F \|\nabla \Psi\| \quad 3.27$$

L'équation correspondant à l'évolution implicite des ensembles de niveaux, se distingue par la formulation 3.28.

$$\frac{\partial \Psi(C, t)}{\partial t} + F \|\nabla \Psi\| = 0 \quad 3.28$$

Dans le cas le plus simple,  $F$  correspond à l'équation 3.29.

$$F = \text{div} \frac{\nabla \Psi}{|\nabla \Psi|} \quad 3.29$$

L'évolution numérique dans le formalisme des ensembles de niveaux, exprime certaines performances parmi lesquelles :

- 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, 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 caractéristiques des ensembles de niveaux, résident dans la gestion automatique de la topologie de la courbe d'une part ainsi que, dans la simplicité de leur mise en œuvre d'autre part.

Plusieurs travaux en imagerie mammographique [90, 122] ont montré les performances des contours implicites, dans la détection des anomalies du sein.

Boujelben et al [90], ont obtenu des résultats intéressants pour le cas d'images mammographiques, dont le contraste est faible. La figure 3.15 illustre le cas de la détection de calcifications, en mammographie numérique.

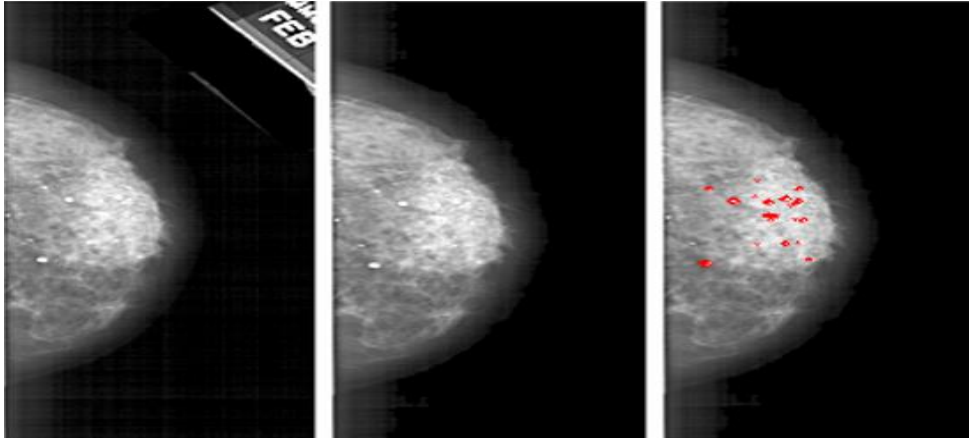


Figure 3.15 : Détection de calcifications par les ensembles de niveaux [90]  
(c) : Image originale (b) : image prétraitée (c) : Image segmentée

Divers modèles déformables [112], réalisent une segmentation par approche régions ou contours. L'objet de notre étude, est l'exploration des contours actifs par régions.

### 3.2.2.3 Contours actifs basés régions

Dans le cadre de l'approche frontière [114, 117], l'information est prise en compte seulement le long du contour. Elle est évaluée via le gradient de l'image, une caractéristique sensible aux artefacts.

La solution adoptée, est alors considérée par les modèles orientés régions, pour lesquels, l'attache aux données dépendant de toute l'information contenue dans l'image, est prise en compte.

Le principe général repose sur l'évolution d'un contour  $C$ , qui segmente l'image en régions, une fois la convergence atteinte. Un seul contour partitionne l'image en deux régions :  $\Omega_{int}$  et  $\Omega_{ext}$ . Dans cette configuration,  $\Omega_{int}$  indique le domaine défini par l'intérieur du contour  $C$  et  $\Omega_{ext}$ , notifie le domaine représenté par l'extérieur de cette courbe. L'attache aux données, est réalisée par une modélisation statistique des régions arborées, par la partition courante.

Les modèles orientés régions, ont été initiés, par la fonctionnelle de Mumford et Shah [122], à l'origine de plusieurs travaux tels que, l'approche de Chan et Vese [87].

### 3.3.2.3 Modèle de Chan et Vese

C'est un modèle dynamique basé sur la fonctionnelle de Mumford-Shah [120]. Qualifié de « modèle sans frontières », l'approche de Chan et Vese [89, 123-127] peut être appliquée sur des images avec des frontières non définies, par le gradient.

Ce modèle a été généré dans sa première version sous la forme biphasée, autrement dit, l'image est segmentée en deux régions. La méthode a été par la suite, étendue aux modes multiphasés, pour segmenter l'image en plusieurs régions.

L'idée fondamentale de ces méthodes évolutives, est la minimisation d'une fonctionnelle d'énergie (3.30) autour de la courbure  $C$ , d'une image  $I$ , afin de déterminer la partition optimale de deux régions, séparées par un contour  $C$ . Dans ce contexte, pour une courbe fermée, le domaine de l'image est partitionné en deux sous-domaines, au gré d'intensités approximativement homogènes. On considère à cet égard, l'image en niveaux de gris  $I$  dans le domaine  $\Omega$ . La courbure évolutive  $C$ , dans l'espace  $\Omega$ , est définie telle, une frontière d'un sous-ensemble de  $\Omega$ . Dans cette optique, la région interne à  $C$  est notifiée par  $\Omega_{int}$  et la région externe à  $C$ , est révélée par  $\Omega_{ext}$ . Ces différentes régions transcrites par les équations  $C_1$  (3.33) et  $C_2$  (3.34), correspondent respectivement, aux intensités moyennes des niveaux de gris, à l'intérieur et à l'extérieur de  $C$ .

$$F(C_1, C_2, C) = \mu \cdot \text{Longueur}(C) + \nu \cdot \text{Aire}(\text{intérieur}(C)) + \lambda_1 \int_{\text{intérieur}(C)} |I(x, y) - C_1|^2 dx dy + \lambda_2 \int_{\text{extérieur}(C)} |I(x, y) - C_2|^2 dx dy \quad 3.30$$

Où,  $\mu \geq 0$ ,  $\nu \geq 0$ ,  $\lambda_1 > 0$  et  $\lambda_2 > 0$  sont les paramètres de régularisation du contour dont, les valeurs varient entre 0 et 1 [124, 125, 126].

Le troisième terme ainsi que le quatrième, issus de la relation 3.30, correspondent aux critères d'attache aux données utilisés, pour guider la courbe vers les frontières de l'objet à segmenter au gré de  $C_1$  et  $C_2$ .

L'extraction de l'information pertinente, est accomplie par la minimisation de la fonctionnelle d'énergie dont l'évolution, répond aux ensembles de niveaux. Dans ce cadre, le modèle de Chan et Vese, est effectué au gré du formalisme, des contours implicites. A cet égard, la courbure  $C \subset \Omega$  est représentée par le niveau zéro de la fonction de Lipschitz  $\phi: \Omega \rightarrow R$  en adéquation avec les conditions (3.31).

$$\begin{aligned} C_1 &= \text{moyenne}(I(x, y)) \text{ pour } \phi(x, y) > 0 \\ C_2 &= \text{moyenne}(I(x, y)) \text{ pour } \phi(x, y) < 0 \end{aligned} \quad 3.31$$

où  $\phi(x, y)$  est la fonction de la distance signée de la courbe  $C$ .

La fonctionnelle d'énergie  $F(C_1, C_2, C)$  est alors, exprimée par  $F(C_1, C_2, \phi)$  dans la relation (3.32). Elle consiste à déterminer la meilleure partition, de l'image en deux régions homogènes séparées par la courbe  $C$ .

$$\begin{aligned}
 F(C_1, C_2, \phi) = & \mu \int_{\Omega} \delta_0(\phi(x, y)) |\nabla \phi(x, y)| dx dy + \nu \int_{\Omega} H(\phi(x, y)) dx dy \\
 & + \lambda_1 \int_{\Omega} |I(x, y) - C_1|^2 H(\phi(x, y)) dx dy \\
 & + \lambda_2 \int_{\Omega} |I(x, y) - C_2|^2 H(1 - \phi(x, y)) dx dy
 \end{aligned} \tag{3.32}$$

Où  $H$  et  $\delta$ , représentent respectivement, les fonctions de Heaviside et de Dirac. Les intensités moyennes  $C_1$  et  $C_2$  sont formulées, par les relations (3.33) et (3.34).

$$C_1 = \frac{\int_{\Omega} I(x, y) H(\phi(x, y)) dx dy}{\int_{\Omega} H(\phi(x, y)) dx dy} \tag{3.33}$$

$$C_2 = \frac{\int_{\Omega} I(x, y) (1 - H(\phi(x, y))) dx dy}{\int_{\Omega} (1 - H(\phi(x, y))) dx dy} \tag{3.34}$$

Ce modèle s'affranchit donc, de toute information de frontière et n'utilise que les moyennes des intensités de niveaux de gris dans les régions intérieure et extérieure ; comme critères d'attache aux données.

### 3.3.2.5 Modèle de Lankton

Il s'agit du modèle de Chan et Vese amélioré. L'approche [128] en question, consiste à faire converger le contour, vers des régions localement homogènes.

Le paradigme étudié, présente un formalisme, qui délimite les objets pertinents, en contrôlant le rayon de localisation  $r$ .

Soit à cet effet, une image  $I$  définie sur un domaine  $\Omega_y$ . On considère  $C$  un contour fermé, représenté comme l'ensemble de niveau 0 d'une fonction  $\phi: C = \{x | \phi(x) = 0\}$  de distance signée.

La courbe  $C$  évolue dans une bande étroite  $\Omega$ , autour de cet ensemble initial, afin de diminuer la complexité calculatoire [123].

Soit à cet égard, la fonction  $\varphi(x)$  qui constitue une version lissée, du delta de Dirac centré en  $\varphi(x) = 0$ , définissant la zone étroite autour de  $C$ .

Autour d'un masque  $B(x, y)$ , circulaire de rayon  $r$ , le paradigme de Lankton est contrôlé par l'équation différentielle ci-dessous :

$$\begin{aligned} \frac{\partial \varphi}{\partial t}(x) = & \lambda \partial \varphi(x) \operatorname{div} \left( \frac{\nabla \varphi(x)}{|\nabla \varphi(x)|} \right) \\ & + \partial \varphi(x) \int_{\Omega_y} B(x, y) \partial \varphi(y) \left( (I(y) - u_x)^2 - (I(y) - v_x)^2 \right) dy \end{aligned} \quad 3.35$$

Où,

$\lambda$  est un paramètre de régularisation de la courbure  $\varphi(x)$ .

$u_x$  et  $v_x$  représentent les moyennes de  $I$  dans le masque  $B$  à l'intérieur et à l'extérieur de  $C$ .

Le deuxième terme est destiné à localiser les régions, comparant les statistiques à l'intérieur et à l'extérieur du contour dans un voisinage de chaque pixel, de la bande étroite  $\partial \varphi(x)$ .

Avec,

$$B(x, y) = \begin{cases} 1 & \text{si } \|x - y\| < r \\ 0, & \text{sinon} \end{cases} \quad 3.36$$

Où,  $x$  et  $y$  correspondent à des variables spatiales indépendantes, chacune d'elles représente un point du domaine  $\Omega$ . Dans ce contexte, les moyennes

$u_x$  et  $v_x$  sont notifiées par les relations 3.37 et 3.38.

$$u_x = \frac{\int_{\Omega_y} B(x, y) H\varphi(y) I(y) dy}{\int_{\Omega_y} B(x, y) H\varphi(y) dy} \quad 3.37$$

$$v_x = \frac{\int_{\Omega_y} B(x, y) (1 - H\varphi(y)) I(y) dy}{\int_{\Omega_y} B(x, y) (1 - H\varphi(y)) dy} \quad 3.38$$

Dans ce modèle, l'extraction des ROI, s'effectue sans altération de l'information. En effet, la segmentation de l'image par l'algorithme de Lankton, permet la prise en compte automatique des changements de topologie.

### 3.3.3 Principe de la logique floue

La logique floue est formalisée par une théorie mathématique des ensembles flous de L. Zadeh [129] qui présente une extension de la théorie des ensembles classiques, aux ensembles définis de façon imprécise. Elle réalise ainsi, la modélisation des imperfections des données, en se rapprochant dans une certaine mesure, de la flexibilité du raisonnement humain.



Le principe consiste à introduire la notion de degré d'appartenance à une décision donnée, pour la vérification d'une condition, lui permettant ainsi, d'interférer avec un état non booléen. De ce fait, la logique floue rend possible, la prise en compte des imprécisions et des incertitudes.

Assimilée au raisonnement humain, la logique floue voit son intervention dans divers domaines, tels que l'imagerie pour l'aide à la décision. Dans ce contexte, l'intérêt des ensembles flous pour le traitement de l'information en imagerie [130], peut se décliner en particulier selon les aspects suivants :

- ❖ La capacité des ensembles flous à représenter l'information spatiale dans les images avec son imprécision, à différents niveaux (local, régional ou global), et sous différentes formes (numérique, symbolique, qualitative ou, quantitative).
- ❖ La possibilité de représenter des informations très hétérogènes, extraites directement des images ou issues de connaissances externes, comme des connaissances expertes ou génériques sur un domaine particulier.
- ❖ La capacité des ensembles flous, à manipuler l'information spatiale.

### 3.3.4 Regroupement spatial flou

La segmentation par l'algorithme des c-moyennes floues (fuzzy c-means) est une approche floue, fondée sur l'optimisation d'un critère quadratique de classification [131-133], où chaque classe est représentée par son centre de gravité. Le processus nécessite de connaître le nombre de classes au préalable et les génère ainsi, par un processus itératif en minimisant une fonction 'objectif' (3.39).

Il permet à cet effet, d'obtenir une partition floue de l'image, en accordant à chaque pixel un degré d'appartenance, à une région donnée.

$$J_m(U, V) = \sum_{i=1}^C \sum_{k=1}^n \mu_{ik}^m \|x_k - v_i\|^2 \quad 3.39$$

Où, n correspond au nombre de pixels dans l'image.

U, notifie la matrice de partition floue dans laquelle,  $\mu_{ik}$  correspond au degré d'appartenance de  $x_k$  dans la classe i.

V est l'ensemble des centres de clusters  $v_i$ .

$\|x_k - v_i\|^2$  estime la distance entre l'objet  $x_k$  et le centre du cluster  $v_i$ .

m représente le facteur de contrôle du regroupement flou ( $m \succ 1$ ).

Dans la segmentation par classification floue, la fonction d'appartenance est réalisée selon les contraintes suivantes :

$$\sum_{i=1}^C \mu_{ik} = 1 \quad \sum_{k=1}^n \mu_{ik} \succ 0 \quad 0 \leq \mu_{ik} \leq 1 \quad 3.40$$

La fonction ‘objectif’ assure son optimisation, selon l’évolution des fonctions d’appartenance  $\mu_{ik}$  (3.41) et des centroïdes  $v_i$  (3.42).

$$\mu_{ik} = \frac{\|x_k - v_i\|^{-2(m-1)}}{\sum_{j=1}^C \|x_k - v_j\|^{-2(m-1)}} \quad 3.41$$

$$v_i = \frac{\sum_{k=1}^n \mu_{ik}^m x_k}{\sum_{k=1}^n \mu_{ik}^m} \quad 3.42$$

L’inconvénient majeur dans l’algorithme des C-moyennes, lors de la segmentation des images, est la négligence de l’information spatiale. A cet effet, la distinction des différents composants mammaires, particulièrement lorsque les pixels du voisinage de la zone d’intérêt sont grandement corrélés, s’effectue par l’intégration de l’information spatiale [95-97] dans la fonction floue d’appartenance  $\mu_{ik}$  (3.43).

$$\mu_{ik} = \frac{\mu_{ik}^p h_{ik}^q}{\sum_{j=1}^C \mu_{jk}^p h_{jk}^q} \quad 3.43$$

Où p et q représentent des paramètres de contrôle du partitionnement.

Dans cette optique, la variable  $h_{ik}$  intègre l’information spatiale via la relation (3.44).

$$h_{ik} = \sum_{j \in N_n} \mu_{nk} \quad 3.44$$

$h_{ik}$  transcrit la fenêtre centrée autour du pixel n, qui génère le cluster correspondant au maximum de la matrice de partition floue.

Dans ce cadre, l’initialisation de la fonction de niveau zéro dans le modèle de Chan et Vese, est adoptée par la relation (3.45) où l’information spatiale, exprime les différents composants mammaires.

$$\phi_0(x, y) = -4e(0.5 - B_k) \quad 3.45$$

Où,  $B_k$  est l’image binaire issue du regroupement spatial flou.

e représente une constante qui ajuste, la fonction de Dirac suivant la relation 3.46.

$$\delta_e(x) = \begin{cases} 0 & |x| > e \\ \frac{1}{2e \left[ 1 + \cos\left(\frac{\pi x}{e}\right) \right]} & |x| \leq e \end{cases} \quad 3.46$$

Où la variable  $x$ , est assignée dans ce cas de figure, à la fonction  $\phi$ .

### 3.4 Approches adoptées pour la détection des néoplasmes et des tissus sains

L'extraction des structures mammaires, est une phase délicate, en raison de la nature complexe des tissus et, de leur variabilité. La conduite à tenir est donc, de concevoir un système offrant la possibilité de se focaliser sur la délimitation des différentes structures pathologiques et saines, au gré de critères adaptatifs. Notre méthodologie aborde à cet égard, trois approches de détection des anomalies du sein. La première localise les lésions, suivant une réflexion de Chan et Vese, en adéquation avec les mammogrammes traités. La seconde se fait selon une croissance de régions associée, au modèle de Lankton. L'ultime approche, s'intéresse particulièrement à la localisation des pathologies et des tissus sains, via les techniques de regroupement spatial flou.

#### 3.4.1 Modèle adaptatif de Chan et Vese

La détection précise des régions d'intérêt (cf. Algorithme 3.1), est assujettie à une organisation hiérarchique de trois phases :

- ✓ *la correction du mammogramme, afin de permettre, la mise en exergue des structures cliniques ;*
- ✓ *la délimitation de la zone annotée par le clinicien, pour générer le masque de la segmentation évolutive. Ce masque, correspond au cercle entourant la zone suspecte ;*
- ✓ *l'initialisation du contour aux frontières de la zone annotée et l'extraction des pathologies mammaires, suivant le modèle adaptatif de Chan et Vese.*

---

Algorithme adaptatif de Chan et Vese pour la détection morphologique des entités sénologiques du sein

---

Début

1. Amélioration du contraste de la zone suspecte mammographique.

2. Initialisation de la fonction  $\phi_0(x, y)$  (3.47) conformément aux coordonnées de la zone annotée dans le mammogramme.

3. Pour  $n=1$  à  $N$  (nombre d'itérations du processus).

3.1 Evaluation des intensités moyennes  $C_1$  et  $C_2$  selon (3.33) et (3.34).

3.2 Evaluation de la fonctionnelle d'énergie  $F(C_1, C_2, \phi)$  selon (3.32).

3.3 Evolution du modèle déformable de la fonctionnelle d'énergie, jusqu'à sa minimisation ou  $n=N$ .

Fin

---

Algorithme 3.1: Modèle adaptatif de Chan et Vese pour la détection morphologique des entités sénologiques du sein

L'opération majeure et délicate du processus, est d'établir la surface de niveau zéro, permettant le partitionnement significatif du mammogramme.

Dans cette optique, le choix de la fonction primaire du modèle adaptatif de Chan et Vese, que nous avons établi dans nos travaux [127], se rapporte à l'équation (3.47) du cercle entourant la zone suspectieuse.

$$\phi_0(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2} - r \quad 3.47$$

Où,  $x_0$  et  $y_0$  correspondent aux coordonnées du centre du cercle, entourant la zone annotée autour du rayon  $r$ .

### 3.4.2 Approche de croissance de régions associée au modèle de Lankton

La méthode suivie pour la détection des anomalies (cf. Algorithme 3.2), 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, pour extraire la zone annotée par le radiologue selon un critère de similarité approprié. La délimitation précise de la masse, est effectuée via le modèle de Lankton à partir du résultat de l'agrégation des ROI.

---

#### Algorithme d'agrégation de pixels associée au modèle de Lankton

---

Début

1. Amélioration du contraste de la zone suspectieuse mammographique.
2. Segmentation par croissance de régions.
  - 2.1 Choix du germe initial (centre du cercle de la zone annotée).
  - 2.2 Agrégation des pixels adjacents selon le critère de similarité 3.13.
  - 2.3 Arrêt du processus lorsque le critère d'homogénéité, n'est plus vérifié.
3. Raffinement du résultat de la croissance de régions par l'approche de Lankton.
  - 3.1 Initialisation de la fonction  $\phi_0(x, y)$  conformément au résultat de l'agrégation de pixels.
  - 3.2 Pour  $n=1$  à  $N$  (nombre d'itérations du processus).
  - 3.3 Evaluation des intensités moyennes  $u_x$  et  $v_x$  selon (3.37) et (3.38).
  - 3.4 Evaluation de l'équation  $\frac{\partial \phi}{\partial t}$  selon (3.35).
  - 3.5 Evolution du modèle déformable, jusqu'à sa minimisation ou  $n=N$ .

Fin

---

Algorithme 3.2 : Algorithme d'agrégation de pixels associée au modèle de Lankton pour la détection des néoplasmes mammaires

Le modèle en question, est fortement dépendant du rayon de localisation  $r$ . Il peut ne pas converger, lorsque ce dernier est inférieur à la distance séparant la courbe, de l'objet d'intérêt. Dans le cas de notre étude, cet aléa est levé, par la considération du rayon de la zone annotée.

La minimisation de l'énergie est obtenue, lorsque la courbure de l'objet est parfaitement approximé par les intensités  $u_x$  et  $v_x$ .

### 3.4.3 Localisation des tissus mammaires par regroupement spatial flou

Le partitionnement des constituants tissulaires du sein, est institué par le regroupement spatial flou, pour générer la fonction initiale du modèle déformable. Le raffinement de la segmentation, est obtenu par le modèle de Chan et Vese.

Le synopsis méthodologique parcourt les étapes explicitées, dans l'algorithme suivant:

---

Algorithme du regroupement spatial flou associé au modèle de Chan et Vese
Début
1. Correction morphologique du mammogramme
2. Application du regroupement spatial flou pour partitionner le mammogramme en des amas adéquats.
2.1 Choix approprié du nombre de clusters 'C' dans le mammogramme (structures tissulaires, zone suspectieuse et autres composants anatomiques).
2.2 Initialisation du degré de flou $m=2$ .
2.3 Initialisation de la matrice de partition floue U selon les contraintes (3.40).
2.4 Minimisation de la fonction 'objectif' J (3.39).
2.5 Evolution de la matrice d'appartenance U (3.41) ainsi que des centroïdes $V_i$ (3.42).
2.6 Intégration de l'information spatiale dans la matrice d'appartenance (3.43) et (3.44).
3. Initialisation de la fonction de niveau zéro $\Phi_0$ (3.45) selon les partitions spatiales floues via la région locale.
4. Pour $It=1 : N_i$ (nombre d'itérations)
4.1 Evaluation des moyennes $C_1$ et $C_2$ selon les relations (3.33) et (3.4).
4.2 Evaluation de la fonctionnelle d'énergie $F(C_1, C_2, \phi)$ au gré de la relation (3.32).
4.3 Evolution de l'objet autour de l'énergie, jusqu'à sa minimisation.
Fin

Algorithme 3.3 : Regroupement spatial flou associé au modèle adaptatif de Chan et Vese pour la localisation des structures pathologiques et saines dans les mammogrammes

## 3.5 Conclusion

Ce chapitre a fait le parcours des principales approches étudiées, dans le cadre de notre préambule à la recherche, pour la correction de mammogrammes ainsi que pour la délimitation des zones suspectieuses.

L'accent est mis sur les modèles implicites, pour la prise en compte de la topologie des anomalies d'une part et sur le regroupement spatial flou, pour la distinction des différentes structures tissulaires mammaires, d'autre part.

L'objet du chapitre qui suit, est l'estimation de la complexité des zones suspectieuses, suivant des descripteurs morphométriques et texturaux.

---

## **Chapitre 4 Signature Morphométrique et Texturale des Mammogrammes**

---

---

## 4 Signature Morphométrique et Texturale des Mammogrammes

---

« Quoique notre esprit ne puisse pas comprendre en soi, le rapport de causalité qui existe entre les phénomènes et la matière, la science exige que nous admettions la nécessité d'un déterminisme dans cette causalité. » C. Bernard.

### 4.1 Introduction

Les connaissances sémiologiques des lésions oncologiques du sein, diligentent le praticien dans son diagnostic. Leur degré de suspicion, est apprécié par l'analyse anatomopathologique ; un examen ô combien invasif et anxieux pour la femme. C'est dans ce cadre, que la description morphométrique et texturale des signes cliniques, intervient pour l'extraction de la signature de mammogrammes. Sa visée est l'estimation de la complexité ou l'uniformité de cette désorganisation. L'examen radiologique des images mammographiques, est alors corrélé à l'analyse de leur granularité et de leur forme, pour assister le clinicien dans son interprétation.

Ce chapitre est consacré à la caractérisation des néoplasmes mammaires, suivant des attributs texturaux et morphologiques. Ce processus est réalisé pour diriger la catégorisation des anomalies tumorales du sein, via une classification supervisée.

### 4.2 Description sémiologique des anomalies mammaires

Le mammogramme suspicieux connaît plusieurs variations, au gré des modifications cliniques (cf. Figure 4.1). La sémiologie de ces néoplasmes, peut-être appréciée par les paramètres qui expriment la malignité, la régularité et la granularité des structures tissulaires du sein. En effet, les masses bénignes ont généralement tendance, à avoir une forme ronde ou ovale et un contour lisse et circonscrit. A contrario, les masses malignes ont une forme irrégulière avec des contours mal définis, flous ou spiculés. A cet égard, la relation, entre la complexité de la forme et le degré de gravité des anomalies du sein, est bien notifiée par la figure 4.1. Dans le cas des calcifications, le degré de malignité est intimement lié à la distribution de ces signes, ainsi qu'à leur morphologie (cf. Figure 4.2).

La conduite à tenir des radiologistes, est une lecture des entités suspicieuses détectées, via un négatoscope. Cette analyse descriptive est ponctuée, par l'étude sémiologique des signes cliniques au gré de leur morphologie. Cependant, cette caractérisation n'est pas suffisante, pour évaluer le degré étiologique des néoplasmes. C'est dans ce cadre, que les processus intelligents modélisés par des théories mathématiques, interviennent pour assister le praticien dans sa décision. A cet égard, la complexité des anomalies mammaires, peut-être instituée par la forme et la texture.

Dans ce contexte, la littérature scientifique [6, 83] présente la caractérisation de la sémiologie des constituants sénologiques du sein, au gré d'une estimation de descripteurs, basés sur des paramètres morphologiques et texturaux ; tels que les moments géométriques, la compacité, l'excentricité ainsi que, les facteurs statistiques, les transformées en ondelettes ou les transformées de Gabor.

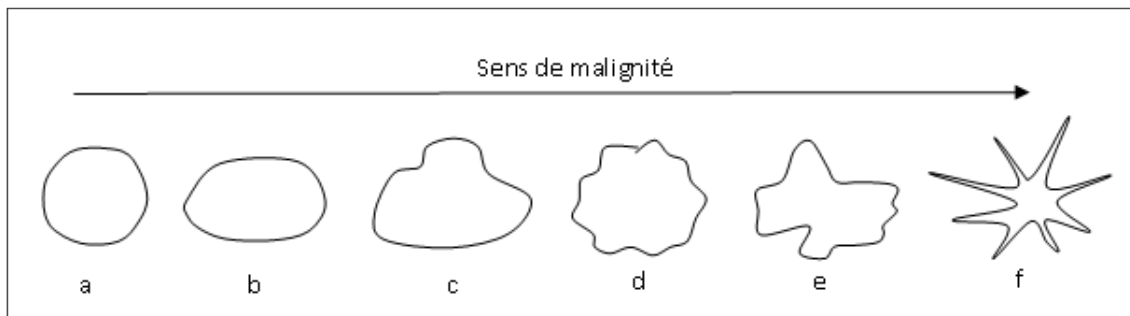


Figure 4.1 : Evolution de la malignité des masses en fonction de la complexité du contour [27]  
 a) Masse circonscrite circulaire, b) Masse circonscrite ovale, c) Masse macro lobulée,  
 d) Masse micro lobulée, e) Masse irrégulière, f) Masse spiculée

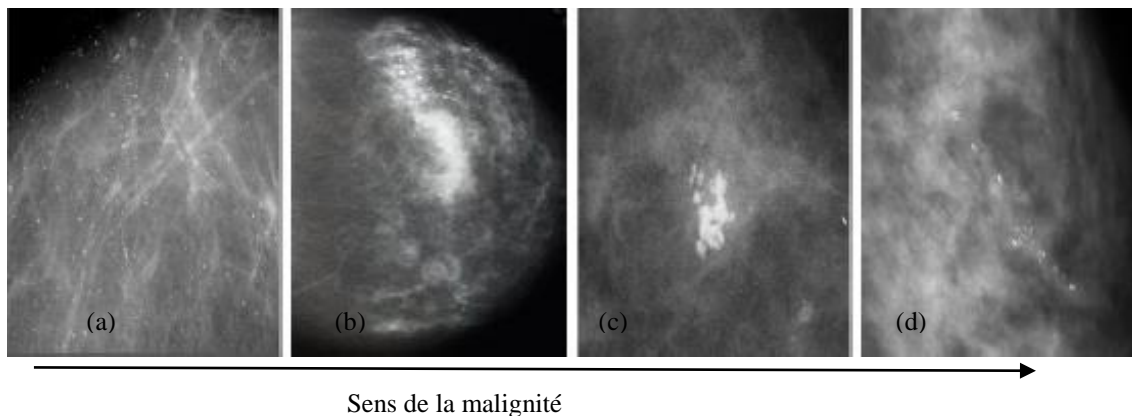


Figure 4.2 : Evolution de la malignité des calcifications en fonction de leur distribution [27]  
 (a) Calcifications Diffuses ou éparées, (b) Répartition régionale,  
 (c) Répartition en cluster (d) Répartition linéaire

A cet effet, dans les travaux de Rangayyan [6], l'orientation des distorsions architecturales et des lésions spiculées dans les mammogrammes, est exprimée par les paramètres de la transformée de Gabor. La distribution de la glande mammaire étant linéaire, elle est dirigée vers le mamelon dans le cas tissulaire sain et vers la zone suspecte, dans le cas pathologique.

Dans l'exploration de Bhateja et al [134], l'estimation de la rugosité néoplasique, se fait par l'approche texturale. Les anomalies de la base MIAS, sont décrites par les quatorze paramètres texturaux dont ; la corrélation, le contraste, le moment angulaire de second ordre, la variance et l'entropie.



L'appréciation morphométrique des signes cliniques du sein, est diligentée par une myriade d'attributs géométriques ; parmi lesquels, le centre de gravité, la surface, les moments géométriques, la convexité et la rectangularité [50, 27, 47, 51].

Dans le parcours bibliographique exploré, les anomalies mammaires sont analysées suivant leur morphologie et leur granularité [6, 83, 84, 135]. En effet, la caractérisation des entités d'intérêt via des descripteurs morphologiques, ne pourrait apprécier que l'aspect géométrique de l'information, sans prendre en compte son degré de complexité. Par conséquent, la fusion de facteurs morphométriques et texturaux, s'avère utile, pour l'extraction minutieuse des connaissances des structures tissulaires saines et sénologiques. C'est l'aspect ciblé, par cette étude.

### **4.3 Analyse de la 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.

Au gré du panel étymologique, la texture est considérée comme un arrangement spatial des éléments constituant une matière. La définition littéraire exprime la texture comme, une répétition spatiale d'un même motif, dans différentes directions de l'espace.

En imagerie, selon Sklansky [136], une région dans une image contient une texture constante, si un ensemble de statistiques locales ou autres propriétés de la fonction image sont constantes, varient faiblement ou sont approximativement périodiques.

La texture est décrite selon Haralick [137], comme étant un ensemble de primitives de taille et de forme variables, présentant une organisation spatiale particulière.

Aux abords de ces définitions, une texture est donc presque de l'ordre de l'abstrait, difficile à définir quantitativement de façon précise. Cependant, elle peut être décrite et quantifiée par des attributs souvent dépendants du domaine exploré (cas d'images médicales, d'images naturelles, de matériaux, etc...) plus faciles à comprendre et à manipuler, au gré des traitements mathématiques.

#### **4.3.1 Types de texture**

Dans la littérature, on trouve diverses familles de textures qu'on peut classer, selon les textures structurées (macro-textures) et les textures aléatoires (micro-structures) [137-140].

##### **4.3.1.1 Textures structurées**

Les textures structurées ont pour caractéristique, la répétition de primitives à intervalles réguliers. Parmi elles, nous pouvons différencier les textures parfaitement périodiques (textures déterministes) et les textures dont, les primitives ont subi quelques transformations (textures directionnelles).

Les textures déterministes (cf. Figure 4.3) se caractérisent par la répétition d'un motif géométrique, appelé "texton". Ce genre de textures, peut donc facilement être synthétisé, à partir d'un motif, si on connaît son orientation et sa taille. Il s'agit de textures périodiques, formée de primitives arrangées d'une façon particulière.

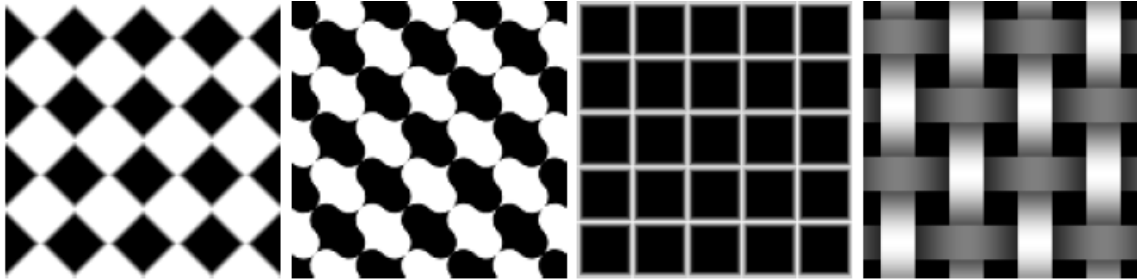


Figure 4.3 : Exemples de textures déterministes [139]

Les textures directionnelles (cf. Figure 4.4), se décrivent comme un mélange, entre une texture stochastique et une texture déterministe. En effet, à l'intérieur de ce type de textures, nous pouvons trouver une répétition de motifs très similaires, mais pas identiques. Ce type de texture ne présente pas de motif de base ; mais il se caractérise par certaines orientations.

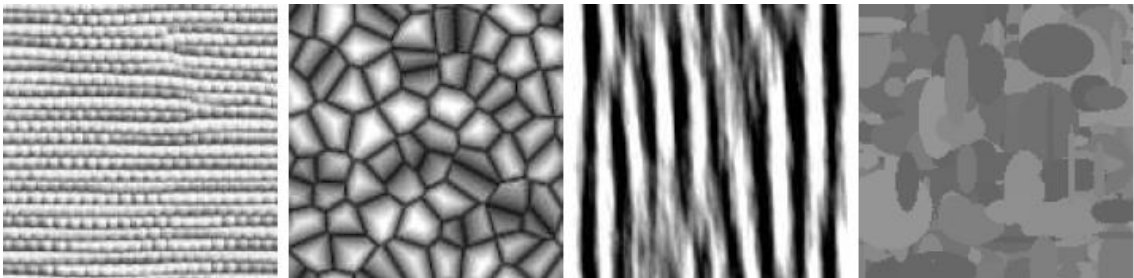


Figure 4.4 : Exemples de textures directionnelles [139]

#### 4.3.1.2. Textures aléatoires

Ces textures sont désordonnées (cf. Figure 4.5) tout en apparaissant, globalement homogènes. Il est impossible de les isoler d'un motif de base ; car la distribution des intensités, n'est l'objet d'aucune régularité.

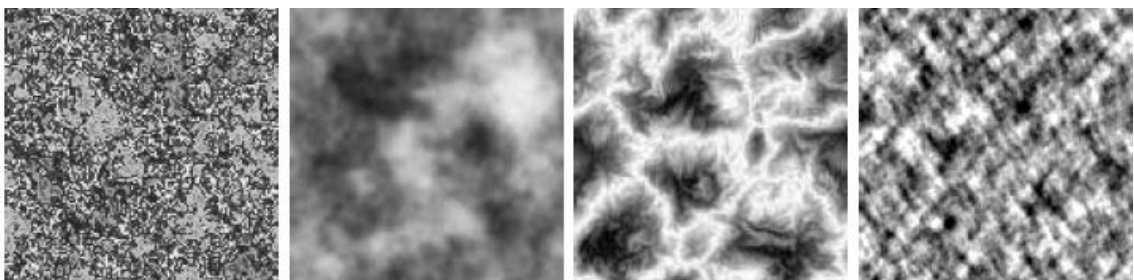


Figure 4.5 : Exemples de textures aléatoires [139]

### 4.3.2 Méthodes d'analyse de texture

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 [54, 138-143] sont regroupées dans quatre familles principales :

- les méthodes structurelles ;
- les méthodes basées sur la modélisation spatiale des textures ;
- les méthodes spatio-fréquentielles ;
- les méthodes statistiques.

#### 4.3.2.1 Méthodes structurelles

Ces méthodes dites aussi géométriques, considèrent la texture comme un agencement d'éléments de base, (primitives selon Haralick [137] ou textons, selon Julesz [142]) qu'il faut extraire, pour en déterminer la règle d'arrangement.

Les dépendances entre les primitives de la texture peuvent être représentées, par des graphes, ce qui permet de synthétiser une texture très régulière.

La description structurelle, peut-être établie, par l'analyse spectrale des textures, ou le motif local binaire [133-136].

#### 1. Méthode de spectre de textures

La méthode de spectre de textures, consiste à décomposer une image en entités, appelées unités de textures [141]. A cet égard, l'image étant divisée en matrices de 3x3 pixels, on considère par la suite, l'ensemble de voisinage  $V = \{V_0, V_1, \dots, V_8\}$  où  $V_0$  correspond au pixel central et  $V_1 \dots V_8$  indiquent les intensités des pixels voisins. Les unités de textures correspondantes, sont représentées par l'ensemble  $E_i = \{E_1, E_2, \dots, E_8\}$  de la relation 4.1.

$$E_i = \begin{cases} 0 & \text{si } V_i < V_0 \\ 1 & \text{si } V_i = V_0 \\ 2 & \text{si } V_i > V_0 \end{cases} \quad 4.1$$

Pour un voisinage de 3\*3 pixels (8 voisins), il est alors, possible d'obtenir  $3^8 = 6561$  unités de textures différentes. A cet effet, des attributs comme la symétrie géométrique, le degré d'orientation ou la symétrie centrale sont estimés, pour caractériser la texture [143].

#### 2. Modèle des motifs binaires

La méthode des motifs binaires locaux (LBP) introduite dans les travaux de Ojala et al [144] a été inspirée par la méthode de spectre de He et Wang [143]. Elle consiste en la génération d'un nombre d'unités de textures, plus restreint.

En considérant un ensemble de voisinage  $V = \{V_0, V_1, \dots, V_8\}$ , on compare les pixels voisins au pixel central, afin de générer un motif binaire. L'affectation binaire est réalisée ainsi : pour  $i = 1 \dots 8$  si  $V_i < V_0$  on assigne au pixel voisin  $i$ , la valeur 0, sinon on lui assigne la valeur 1. Dans l'opération suivante, les valeurs binaires de ce motif, sont multipliées par des poids et sommées pour calculer le code LBP du pixel central (cf. Figure 4.6). L'histogramme des codes LBP, constitue le descripteur caractérisant la texture.

Cette méthode d'analyse, présente une complexité calculatoire faible et permet de différencier, des contenus texturaux. C'est une méthode devenue très populaire et employée jusqu'à nos jours, pour différentes applications en vision par ordinateur [103]. Parmi ces applications, les anomalies mammographiques, ont fait l'objet de l'analyse texturale par les motifs binaires dans les travaux de Pawar et al [53], pour la réduction de faux positifs.

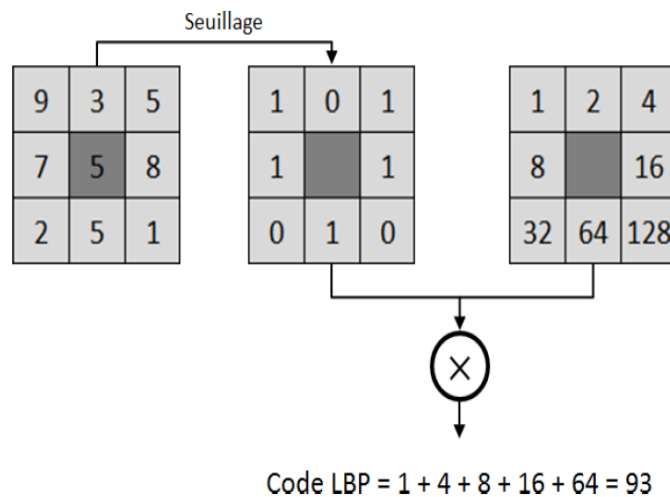


Figure 4.6 : Exemple de construction d'un motif binaire et calcul du code LBP pour un pixel central [139]

#### 4.3.2.1 Méthodes basées sur la modélisation spatiale des textures

Elles se fondent sur la recherche d'un modèle probabiliste, pour décrire ou générer une texture. On peut citer quelques exemples de ces méthodes diligentées, par les champs aléatoires markoviens et les fractales [27, 141].

L'emploi des champs de Markov permet la caractérisation de la dépendance entre pixels, en définissant une structure de voisinage autour de l'image.

La géométrie fractale peut être définie comme une notion complémentaire, à la géométrie euclidienne.

Celle-ci symbolise la dimension d'une figure simple, par un nombre entier. Elle est de 2 dans le plan et notifiée par 3, dans l'espace. La dimension fractale est plus complexe et n'est pas, forcément entière.

Dans ce formalisme, la dimension fractale, symbolise l'irrégularité d'un objet géométrique.

Parmi les nombreuses méthodes permettant d'estimer cette dimension, la plus utilisée, pour analyser la texture d'une image est celle du comptage de boîtes, appelée en anglais 'box counting'. Elle consiste à superposer une grille carrée, sur une image binaire et à comptabiliser, le nombre de cases nécessaires, afin de recouvrir la structure à analyser. La dimension fractale est ensuite obtenue en faisant tendre, la taille des cases de la grille vers 0.

Le modèle fractal a fait l'objet de divers travaux en imagerie mammographique [6, 145]. Dans l'exploration de Seeka et Thomas [145], la caractérisation des tissus pathologiques et sains, est obtenue par des attributs engendrés de la dimension fractale. Ces descripteurs, permettent de prendre en compte le degré d'anisotropie, des signes cliniques.

L'analyse fractale semble intéressante, pour l'investigation des images mammographiques. Néanmoins, lorsque les structures texturales s'enchevêtrent, pour des raisons anatomiques, il devient plus difficile de déceler l'information pertinente.

#### 4.3.2.2 Méthodes spatio-fréquentielles

Dans ces méthodes, les représentations spatio-fréquentielles préservent à la fois les informations globales et locales (cf. Figure 4.7). Elles sont donc bien adaptées, aux signaux quasi périodiques. En effet, les textures sont des signaux quasi périodiques qui ont, une énergie fréquentielle localisée.

Les modèles spatio-fréquentiels, permettent de caractériser la texture à différentes échelles [112, 138].

Le traitement fréquentiel peut-être réalisé, par exemple par : les transformées de Fourier, les filtres de Gabor ou, les transformées en ondelettes.

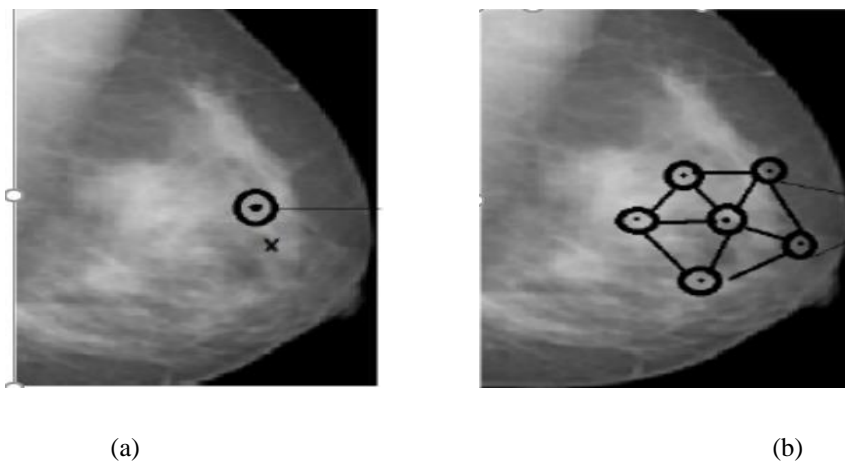


Figure 4.7 : Information locale et information globale [57]

- (a) Information locale extraite autour de chaque pixel de la texture
- (b) Information globale donnée par l'agencement des informations locales

### 1. Transformée de Fourier

Les images peuvent être représentées soit dans le domaine spatial, soit dans le domaine fréquentiel. La transformée de Fourier, permet de passer d'une représentation de l'image dans le domaine spatial, à sa représentation dans le domaine fréquentiel [139].

Il s'agit d'une opération, qui permet de représenter en fréquence (développement sur une base d'exponentielles), des signaux qui ne sont pas périodiques.

Le but de l'utilisation de la transformée de Fourier en imagerie, est de mettre en évidence les caractéristiques fréquentielles d'une texture. C'est le spectre de Fourier (le module de la transformée de Fourier) qui est considéré, en aparté de la phase. En effet, le spectre permet de rendre compte de la distribution énergétique de l'image, de respecter aussi bien la périodicité que l'orientation des motifs de l'image texturale.

Le spectre exhibe l'ensemble des fréquences qui compose un signal mais ne dit rien sur les localisations temporelles de ces fréquences : existent-elles au début du signal, à la fin, ou sur toute la durée ? On dit que la transformation de Fourier perd la localisation. Pour décrire la composition fréquentielle d'un signal sans renoncer, à la localisation temporelle des fréquences qu'il contient, ce sont les représentations temps – fréquence dont il existe de nombreuses variantes, qui assurent cette analyse. Parmi ces variantes, l'analyse en ondelettes et la transformée de Gabor la représentation temps–fréquence la plus simple appelée aussi la transformée de Fourier à fenêtre glissante.

L'une des explorations de la transformée de Fourier réalisée autour des descripteurs normalisés invariants aux transformations affines, est obtenue dans les travaux de Shen et al [70], pour la caractérisation des lésions mammaires.

### 2. Transformée de Gabor

Cette théorie a été définie par Gabor en 1946 [146], partant du fait que la transformée de Fourier permet de mesurer les irrégularités, sans pour autant indiquer leurs localisations. Gabor a ainsi introduit une nouvelle décomposition, en utilisant une fenêtre gaussienne.

Les filtres de Gabor ont la capacité de modéliser la sensibilité fréquentielle et directionnelle caractéristique du fonctionnement du cortex visuel humain, en décomposant le spectre de l'image en images filtrées, selon des plages étroites de fréquences et d'orientations. Ceci permet de fournir des primitives statistiques robustes, très utiles dans la détection des anomalies du sein [6, 83, 147].

Une configuration classique de l'utilisation des filtres de Gabor en analyse de texture, consiste à faire appel à un banc de filtres, chacun centré sur une fréquence et une orientation distinctes, en couvrant au mieux l'ensemble du domaine fréquentiel. Chaque pixel de l'image délivrera ainsi, une réponse pour chaque filtre. La transformée en question, est une fonction gaussienne modulée par une sinusoïde complexe.

Les filtres de Gabor, ont été usés en imagerie mammographique, pour la mise en évidence des anomalies (cf. Figure 4.8), telles que révélées dans les travaux de Raghavendra et al [147].



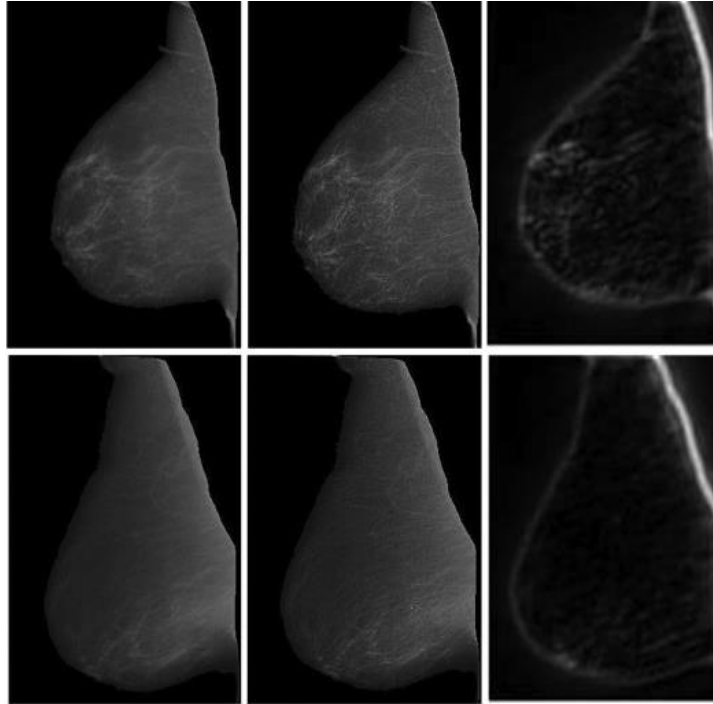


Figure 4.8 : Anormalités mammographiques détectées par la transformée de Gabor [147]

### 3. Transformée en ondelettes

Malgré la capacité remarquable des filtres de Gabor à décomposer le spectre fréquentiel de l'image, les attributs texturaux extraits de ces filtres, peuvent être corrélés en raison de la non-orthogonalité des filtres. Il peut dès lors s'avérer difficile de déterminer si une similarité observée entre échelles d'analyse, est due aux propriétés de l'image ou, à la redondance inhérente à la représentation. En outre, à chaque échelle d'application des filtres de Gabor, les paramètres définissant ces filtres doivent être modifiés. Ces contraintes sont levées par l'utilisation des ondelettes [148]. Celles-ci offrent en effet un cadre d'analyse multi-échelles uniforme (une seule paramétrisation pour toutes les échelles) et permettent de décomposer l'image en sous-bandes orthogonales et indépendantes limitant ainsi, la redondance d'informations.

Apparue au début des années 90, la transformée en ondelettes représente un outil mathématique de décomposition hiérarchique de l'image dans le domaine fréquentiel, tout en préservant le domaine spatial. Elle présente l'avantage d'être efficace en présence de discontinuités, ce qui n'est pas le cas pour les transformées classiques, telles que celle de Fourier.

La transformation en ondelettes, offre une analyse spatio-fréquentielle multirésolution. Elle utilise des translations et des dilatations d'une fonction fixe appelée ondelette mère, ce qui permet de couvrir, pour un signal donné, toutes les fréquences et les localisations temporelles.

Divers travaux recouvrant l'analyse de mammogrammes, ont été réalisés via les transformées en ondelettes pour une analyse multirésolution [6, 149].

Cette analyse consiste à appliquer la transformée en ondelettes, en utilisant deux filtres passe-bas et passe-haut séquentiellement, sur les lignes puis sur les colonnes.

Après chaque filtre, l'image est soumise à un sous échantillonnage, réduisant sa résolution d'un facteur 2 suivant les lignes, ou les colonnes. L'image originale est alors transformée en quatre sous-images notifiées ; par l'image des approximations, les détails horizontaux, les détails verticaux et les détails diagonaux. Une décomposition à n'importe quel ordre, est possible.

#### 4.3.2.4 Méthodes statistiques

Ces méthodes [57, 139] 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 analyser, 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 convertit l'image de la texture (densité spectrale, méthode des extrema locaux, méthodes de transformation de Fourier, ...). Parmi les approches statistiques, qui ont conquis le monde de l'imagerie, on peut citer : la matrice de longueurs de plages, la méthode de différence de niveaux de gris, la méthode de dépendance des régions entourantes, les statistiques du premier ordre ainsi que la matrice de cooccurrence.

##### 1. Matrice de longueurs de plages

La matrice de longueurs de plages (gray level run length matrix GLRLM) se base, sur l'estimation de données statistiques d'ordre supérieur [27, 57].

La méthode des matrices de longueurs de plages (LDP), consiste à représenter les images par des segments de droite (plages) de même luminance puis à extraire, des paramètres traduisant, la répartition statistique de ces différents segments de droite. Ainsi, une longueur de plages, est une suite de points consécutifs et adjacents, ayant le même niveau de gris.

Pour une image donnée, on peut calculer une matrice de longueurs de plages, au gré d'une direction instituée par  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$ . Cette matrice possède n lignes, où n représente le nombre de niveaux de gris de l'image et m colonnes, où m indique la taille maximale de l'image dans la direction choisie.

Parmi les attributs dégagés de ces matrices, on distingue :

- nombre de longueurs de plages (run length number) ;
- proportion de petites plages (short run emphasis) ;
- proportion de grandes plages (long run emphasis) ;
- hétérogénéité des niveaux de gris (gray level no uniformity) ;
- hétérogénéité des longueurs de plage (run length no uniformity).



## 2. Méthode de différence des niveaux de gris

La méthode de différence des niveaux de gris (gray level difference method GLDM) est basée sur le calcul du nombre d'occurrences, d'une différence de niveaux de gris donnée. Cinq caractéristiques sont extraites en utilisant la méthode GLDM : le contraste, la moyenne, l'entropie, le moment de différence inverse et le second moment angulaire [57].

## 3. Méthode de dépendance des régions entourantes

La méthode de dépendance des régions entourantes (surrounding region-dependence matrix SRDM) est une méthode statistique d'analyse de texture. Elle repose sur le calcul de sommes pondérées des régions entourantes dans différentes directions et a été exploitée dans certains travaux, dans le contexte de la détection du cancer du sein [54]. Les principales primitives issues de la matrice SRDM sont :

- la moyenne pondérée horizontale ;
- la moyenne pondérée verticale ;
- la moyenne pondérée diagonale.

## 4. Statistiques du premier ordre

Les statistiques du premier ordre [57], mesurent la probabilité d'observer un niveau de gris à un emplacement aléatoire d'une image. Ces paramètres sont calculés à partir de l'histogramme des niveaux de gris. Parmi les attributs en question, on distingue :

- ❖ **la moyenne** ;
- ❖ **la variance**. Elle représente le moment d'ordre 2 et mesure la répartition des intensités autour de la valeur moyenne ;
- ❖ **l'asymétrie ou le "skewness"**. Ce paramètre correspond au moment d'ordre 3 centré autour de la moyenne, Il mesure l'écart de la distribution des niveaux de gris par rapport à une distribution symétrique ;
- ❖ **l'aplatissement ou le "kurtosis"**. Celui-ci correspond au moment d'ordre 4 centrée autour de la moyenne. Il caractérise la forme du sommet de l'histogramme (plus le kurtosis est faible et plus le sommet de l'histogramme est arrondi).

L'inconvénient majeur des statistiques du premier ordre, basées sur l'histogramme est qu'elles n'exploitent, que l'information en niveau de gris et ne prennent pas en considération, les relations spatiales entre les pixels. C'est dans cette optique, que les attributs statistiques de second ordre, s'avèrent plus efficaces, dans la description des images.

Dans les travaux de Rangayyan [6], la caractérisation des entités anormales du sein telles que les distorsions architecturales, est obtenue par les transformées de Gabor multidirectionnelles.

Dans le cas de cette étude, nous allons nous intéresser aux descripteurs statistiques du second ordre, estimés par la matrice de cooccurrence.

### 5. Matrice de cooccurrence

Les statistiques de premier ordre ne donnent pas d'informations, sur la localisation du pixel. Il est donc nécessaire d'utiliser un ordre supérieur, pour une analyse plus précise. La matrice de cooccurrence, estime des propriétés des images relatives, à des statistiques du second ordre. Une matrice de cooccurrence [137, 140, 141], 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  $\theta$  donnée.

Les directions angulaires  $\theta$  classiquement utilisées, sont  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$  degrés. Les relations de voisinage entre pixels, nécessaires au calcul des matrices, sont illustrées dans la figure (4.9). Par exemple, les plus proches voisins de 'x' selon la direction  $\theta=135^\circ$  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 «  $\theta$  » permet d'évaluer la direction de texture. Cette matrice décrit les régularités observables, dans les niveaux de gris des pixels d'une région. L'exemple de la figure 4.10, montre un exemple de l'élaboration d'une matrice de cooccurrence, pour une distance horizontale valant 1 et une direction de  $0^\circ$ .

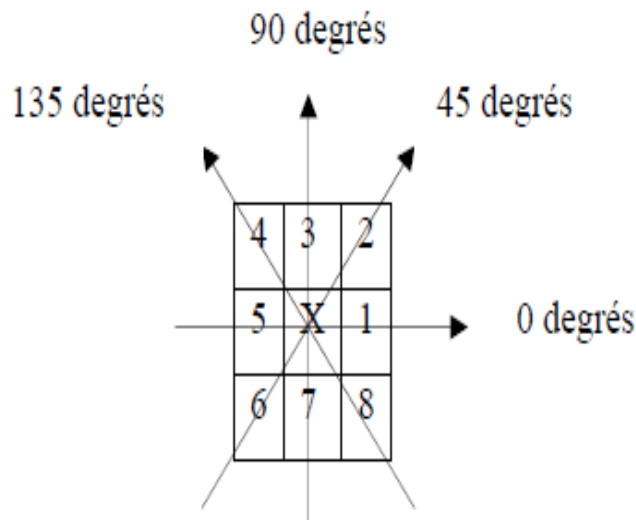


Figure 4.9 : Plus proches voisins de x selon 4 directions [140]

Afin de limiter la complexité calculatoire de la matrice de cooccurrence, une distance  $d=1$  et une orientation de l'angle  $\theta=0^\circ$  sont prises en compte.

1	1	5	6	8
2	3	5	7	1
4	5	7	1	2
8	5	1	2	5

	1	2	3	4	5	6	7	8
1	1	2	0	0	1	0	0	0
2	0	0	1	0	1	0	0	0
3	0	0	0	0	1	0	0	0
4	0	0	0	0	1	0	0	0
5	1	0	0	0	0	1	2	0
6	0	0	0	0	0	0	0	1
7	2	0	0	0	0	0	0	0
8	0	0	0	0	1	0	0	0

(a)
(b)

Figure 4.10 : Exemples de détermination d'une matrice de cooccurrence [140]  
 (a) : matrice initiale sur 8 niveaux (b) : matrice de cooccurrence

Les matrices de cooccurrence contiennent une masse d'information difficilement manipulable, c'est pourquoi elle n'est pas utilisée directement. De ce fait, quatorze indices définis par Haralick [137], correspondant à des caractères descriptifs de la texture, peuvent être estimés, à partir de ces matrices.

#### 4.4 Description texturale

Les connaissances texturales mammaires [54, 67, 127, 149-153], permettent d'orienter la classification des entités sénologiques selon, leur aspect bénin versus malin. Cette discrimination supervisée des masses des calcifications et des tissus sains, s'effectue dans l'ultime étape du système d'aide à la décision.

La granularité des pathologies mammaires est conduite, par une description texturale. A cet égard, la matrice de Haralick, est prise en compte, pour estimer les paramètres statistiques de second ordre de l'image.

Cette matrice est déterminée dans le cas de cette étude, à partir du mammogramme segmenté (image correspondant aux régions d'intérêt) par l'évaluation de l'occurrence, des paires de pixels  $P(i, j, d, \theta)$ . Le premier pixel est localisé au niveau de gris  $i$ , séparé par une distance  $d$  autour de son voisin au niveau de gris  $j$ , selon une direction  $\theta$ .

La matrice de cooccurrence, génère des paramètres statistiques, qui décrivent la complexité de l'image.

Les paramètres considérés, dans cette approche, sont institués par une direction  $\theta = 0^\circ$  et une distance  $d = 1$ .

### 4.4.1 Attributs texturaux du modèle neuro-génétique

Cette section étudie les attributs de la granularité des lésions du sein, du modèle neuro-génétique. La signature étiologique des signes pathologiques explorés, est décrite par six paramètres texturaux.

#### 1. Energie

Ce paramètre (cf. Equation 4.2) mesure l'uniformité de la texture. Il atteint de grandes valeurs, lorsque la distribution des indexes  $i$  et  $j$ , est constante.

$$ENE = \sum_i \sum_j P^2(i, j) \quad 4.2$$

Les facteurs  $P(i, j)$  correspondent aux éléments de la matrice de cooccurrence. Une paire  $P(i, j)$  est assimilée à l'occurrence d'un pixel de niveau de gris  $i$ , séparé par une distance  $d=1$  d'un pixel de niveau de gris  $j$ , autour d'une orientation  $\theta=0^\circ$  dans la carte des néoplasmes. Les niveaux  $i$  et  $j$  sont assignés, aux lignes et aux colonnes dans la matrice de Haralick (indexés).

#### 2. Contraste

Ce facteur indique la concentration des éléments de la matrice de cooccurrence autour de la diagonale. Le contraste (cf. Equation 4.3) est particulièrement élevé, lorsque la texture présente un fort contraste.

$$CST = \sum_i \sum_j (i - j)^2 P(i, j) \quad 4.3$$

#### 3. Entropie

L'entropie (cf. Equation 4.4), évalue le désordre dans une image. Ce paramètre atteint de considérables valeurs, lorsque la texture est complètement aléatoire.

$$ENT = -\sum_i \sum_j P(i, j) \log P(i, j) \quad 4.4$$

#### 4. Variance

La variance (cf. Equation 4.5), estime l'hétérogénéité de la texture. Elle détermine la distribution des indexes  $i, j$ , autour de la moyenne  $m$  (cf. Equation 4.6).

$$\sigma^2 = \sum_i \sum_j (i - m)^2 P(i, j) \quad 4.5$$

$$m = \sum_i \sum_j P(i, j) \quad 4.6$$

## 5. Corrélation

La corrélation (cf. Equation 4.7), évalue la dépendance linéaire entre les indexes  $i$  et  $j$  dans la matrice de cooccurrence.

$$COR = \sum_i \sum_j \left( \frac{(i - m_i)(j - m_j)}{\sigma_i^2 \sigma_j^2} P(i, j) \right) \quad 4.7$$

$$\sigma_i^2 = \sum_i (i - m_i)^2 \sum_j P(i, j) \quad 4.8$$

$$\sigma_j^2 = \sum_j (j - m_j)^2 \sum_i P(i, j) \quad 4.9$$

$$m_i = \sum_i i \sum_j P(i, j) \quad 4.10$$

$$m_j = \sum_j j \sum_i P(i, j) \quad 4.11$$

Où  $\sigma_i^2$ ,  $\sigma_j^2$  (cf. Equations 4.8 et 4.9) représentent respectivement la variance à l'index  $i$  et, la variance à l'index  $j$ .

$m_i$  et  $m_j$  (cf. Equations 4.10 et 4.11) dénotent successivement, la moyenne à l'index  $i$  et, celle à l'index  $j$ , dans la matrice de cooccurrence.

## 6. Homogénéité

Le moment inverse (cf. Equation 4.12) correspond à l'homogénéité de l'image. Ce paramètre est corrélé à une combinaison linéaire de l'énergie et du contraste.

$$HOM = \sum_i \sum_j \left( \frac{P(i, j)}{1 + (i - j)^2} \right) \quad 4.12$$

### 4.4.2 Attributs texturaux du modèle de Vapnik

Les descripteurs de ce modèle, répondent au contraste (cf. Equation 4.3) à l'entropie (cf. Equation 4.4) et à la corrélation (cf. Equation 4.7).

## 4.5 Analyse morphométrique

La notion de la forme est très importante, puisqu'elle nous permet de reconnaître, les objets qui nous entourent.

L'analyse de la morphologie, permet de décrire, les composants d'intérêt, de façon similaire à la perception humaine.

La malignité des anomalies mammaires, est arborée par leur sémiologie. Le degré de malignité des microcalcifications, dépend de leur nombre dans un cluster, de leur contour (régulier, rond, ovale) et de leur forme (polymorphes, annulaires, arborescentes, arciformes, semi-lunaires, rhomboédriques, sédimentées).

Le degré de malignité dans le cas des masses, augmente essentiellement, avec l'élévation de la complexité du contour.

L'analyse morphométrique [6, 83, 127], permet d'apprécier la nature étiologique, des anomalies mammographiques, en évaluant les facteurs les plus pertinents. Cette paramétrisation, met l'accent sur l'étude géométrique des régions suspectes, selon différents attributs. Parmi ces paramètres, la circularité des masses, la rectangularité, l'orientation, l'excentricité, les moments géométriques et la convexité.

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 des facteurs, les plus adaptés à la nature des structures médicales.

Les caractéristiques prises en compte dans le cadre de cette thèse, pour modéliser numériquement les régions d'intérêt mammaires, ont été ratifiées par un expert.

### 4.5.1 Descripteurs morphologiques du modèle neuro-génétique

L'analyse morphométrique apprécie la forme étiologique des mammogrammes par l'évaluation des facteurs les plus pertinents, tels que la compacité et les moments géométriques [8, 27 ; 70, 154] [42, 49 ; 50, 127]. Trois attributs sont choisis dans cette étude, pour décrire la morphologie des structures anormales du sein, suivant leur sémiologie.

#### 1. Compacité

La compacité (cf. Equation 4.13) estime, les irrégularités morphologiques. Ce paramètre présente l'avantage d'être invariant à la rotation, la translation et l'échelle.

$$\text{Comp} = \frac{4\pi A}{\text{Per}^2} \quad 4.13$$

Où  $\text{Per}$  et  $A$  représentent respectivement, le périmètre et la surface de la région d'intérêt. Dans ce contexte, le périmètre correspond, au nombre de pixels du contour de la lésion (longueur de la région).

La compacité permet de distinguer une forme irrégulière d'une forme simple, puisqu'elle attribue à la forme irrégulière une valeur plus élevée.

## 2. Moments invariants de Hu

Ces moments ont fait l'objet de diverses applications, telles que la reconnaissance de formes [155]. Ils permettent de décrire, la forme à l'aide de propriétés statistiques.

Une interprétation physique exprime que le niveau de gris dans une image, est assimilé à une densité de masse.

Les moments géométriques permettent de décrire, une forme à l'aide de propriétés statistiques. Ils représentent les propriétés spatiales de la distribution des pixels, dans l'image. Ils sont facilement calculés et implémentés. Par contre cette approche est très sensible au bruit et aux déformations.

Pour remédier aux aléas de ces facteurs, Hu a proposé un ensemble de sept moments invariants aux translations, rotations et changement d'échelle. Ces attributs offrent un cadre théorique puissant pour caractériser des images, notamment dans le cas de mammogrammes.

La relation (4.14) définit les moments géométriques  $m_{pq}$ .

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad 4.14$$

Où  $f(x, y)$  correspond à l'intensité du pixel de coordonnées  $(x, y)$ .

L'équation (4.15) désigne les moments invariants à la translation considérés par, les moments centraux  $\mu_{pq}$ .

$$\mu_{pq} = \sum_x \sum_y (x - x_g)^p (y - y_g)^q f(x, y) \quad 4.15$$

Où,  $x_g$  et  $y_g$  (4.16) représentent les coordonnées du centre de gravité de l'image.

$$x_g = \frac{m_{10}}{m_{00}}, y_g = \frac{m_{01}}{m_{00}} \quad 4.16$$

Où  $m_{00}$  indique l'aire de la région d'intérêt dans le cas de cette étude.

Lorsque les moments centraux sont normalisés, ils deviennent invariants à l'échelle selon l'expression (4.17).

$$\eta_{pq} = \frac{\mu_{ij}}{\mu_{00}^{(p+q)/2}} \quad 4.17$$

Les descripteurs invariants aux transformations affines générés par Hu, sont déterminés selon les sept facteurs  $\Phi_i$  (4.18).

$$\begin{aligned}
\Phi_1 &= \eta_{20} + \eta_{02} \\
\Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11} \\
\Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \\
\Phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\
\Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
&\quad (3\eta_{21} - \eta_{03})(\eta_{23} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\Phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\Phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12}) - 3(\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{4.18}$$

Les six premiers moments sont invariants aux translations, aux changements d'échelles, aux rotations ainsi qu'aux réflexions. Or, l'invariance aux réflexions peut être problématique quand il s'agit de reconnaître des images « miroirs ». C'est pourquoi Hu exploite aussi dans son système, le 7<sup>ème</sup> moment qui n'est pas invariant aux réflexions. Ce dernier change de signe lorsqu'une telle transformation, est appliquée à l'image et permet donc, de détecter celle-ci.

Les sept descripteurs de Hu sont normalisés, de façon à prendre en compte leur moyenne comme un facteur de forme (4.19). Ce paramètre exprime le degré de diffusion de la densité des lésions, dans la zone suspectieuse.

$$M = \frac{1}{7} \sum_{i=1}^7 \Phi_i \tag{4.19}$$

### 3. Aire

Le paramètre  $A$  (4.20) mesure l'aire de la région d'intérêt.

$$A = \sum_x \sum_y Pix(x, y) \tag{4.20}$$

Où,  $Pix(x, y)$  désigne le pixel à la position  $(x, y)$  dans la région d'intérêt.

L'estimation des attributs morphométriques et texturaux, arbore la signature de chaque mammogramme analysé.

Cette signature est intégrée, dans un classifieur neuro-génétique, autour d'un perceptron multicouche associé aux algorithmes génétiques, pour une aide à la décision.



### 4.5.2 Descripteurs morphologiques du modèle de Vapnik

Nous distinguons dans ce qui suit, les équations de l'excentricité, la convexité et la rectangularité pour la description analytique des formes mammaires [27, 94, 134] [49, 50, 51]. La compacité (cf. Equation 4.13), un paramètre significatif, pour la mise en avant des irrégularités d'une lésion, y est également prise en compte.

#### 1. La convexité

La convexité (cf. Equation 4.21) est la différence relative entre un objet et une coque convexe le contenant. Elle correspond au rapport entre le périmètre de la coque convexe  $P_{cc}$  et le périmètre de l'objet  $P_{ob}$ .

$$Conv = \frac{P_{cc}}{P_{ob}} \quad (4.21)$$

#### 2. La rectangularité

Il s'agit d'un descripteur souvent utilisé, pour décrire le taux d'allongement d'une région [27, 51]. On désigne par  $R$  la région étudiée et  $B_E$  la boîte englobante qui est définie par le plus petit rectangle, contenant la lésion étudiée (cf. Figure 4.11). La mesure de la rectangularité répond à la formulation (4.22).

Le rectangle minimum englobant une forme, appelé aussi la boîte minimale englobante, est en effet le plus petit rectangle contenant tous les points de cette forme.

$$Rect = \frac{Aire(R)}{Aire(B_E)} \quad (4.22)$$

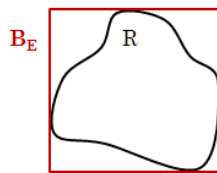


Figure 4.11 : Boîte englobante [27]

La formulation telle que proposée, engendre un descripteur sensible à la rotation. En effet, la rectangularité d'un rectangle droit est égale à 1, alors que celle du même rectangle incliné, est différente de 1 selon la configuration proposée dans la figure 4.11.

Dans les travaux de Cheikhrouhou [27], la prise en compte de l'orientation de l'objet, est exprimée par une configuration ellipsoïdale. La boîte englobante modifiée, symbolise alors, les différentes orientations de la région d'intérêt. Dans ce contexte, on définit l'ellipse équivalente ayant les mêmes moments centrés d'ordre inférieur, ou équivalents à 2, que l'objet à décrire.

Ces moments  $\mu_{02}$ ,  $\mu_{20}$  et  $\mu_{11}$  (cf. Equation 4.15) sont alors calculés, à partir du centre de gravité de l'objet défini par les coordonnées ( $x_g$  et  $y_g$ , cf. Equation 4.16).

L'ellipse considérée ayant les mêmes moments que l'objet à étudier, est définie par ses paramètres distinctifs générés par ; le grand axe  $a_1$ , le petit axe  $a_2$  et l'angle de rotation de l'objet par rapport à l'horizontal (cf. Figure 4.12).

Ces différents paramètres sont déterminés au gré des relations ci-dessous :

$$a_1^2 = \frac{2\left(\mu_{02} + \mu_{20} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\right)}{m_{00}} \quad 4.23$$

$$a_2^2 = \frac{2\left(\mu_{02} + \mu_{20} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\right)}{m_{00}} \quad 4.24$$

Nous rappelons, comme cité plus haut, que  $m_{00}$  le moment d'ordre zéro, représente l'aire de l'objet considéré.

L'angle exprimant l'orientation de la région d'intérêt, est notifié par la relation 4.25.

$$\alpha = \arctan g(t) \quad 4.25$$

Où le paramètre  $t$  est évalué par la formulation 4.26.

$$t = \frac{\mu_{02} - \mu_{20} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}{2\mu_{11}} \quad 4.26$$

La figure 4.12 illustre un exemple d'une masse ainsi, que sa représentation ellipsoïdale. A partir de cette configuration, il est alors aisé de définir (en se basant sur son grand et petit axe), la boîte englobante de même orientation que l'objet considéré.

La rectangularité modifiée conserve la même expression que l'équation 4.22. Ainsi, quelle que soit l'orientation de l'objet, plus la forme de l'objet est rectangulaire, plus le paramètre descriptif, tend vers 1.

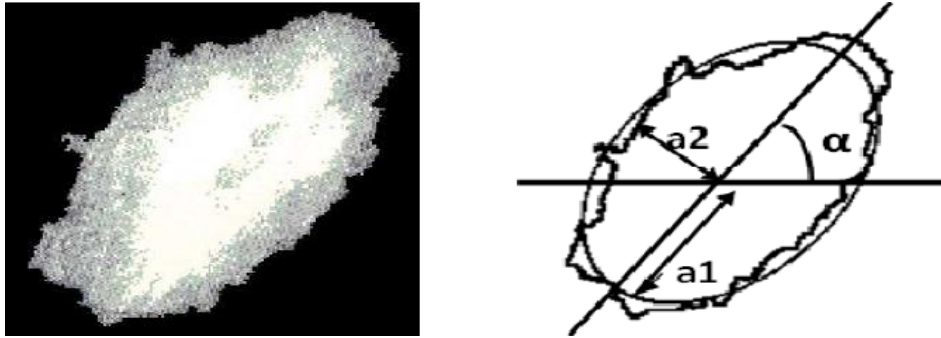


Figure 4.12 : Représentation de l'ellipse équivalente à une anomalie en fonction de ses paramètres usuels [27]

### 3. L'excentricité

L'excentricité (4.27) exprime la distribution de l'objet dans l'espace, tout au long de son axe. Elle représente, le rapport entre la longueur du grand axe et celle, du petit axe de l'objet.

Ce descripteur peut être évalué par la relation 4.27, en fonction des moments centraux et de l'aire de l'objet.

$$Exc = \frac{(\mu_{02} - \mu_{20})^2 + 4\mu_{11}}{A} \quad 4.27$$

Où, A est l'aire de l'objet,

$\mu_{pq}$  indique le moment central supputé par la relation 4.15.

## 4.6 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 neuro-génétique et par les SVM. Le chapitre suivant, est consacré à la classification des anomalies et des tissus sains du sein, autour des modèles connexionnistes et de Vapnik.

---

**Chapitre 5 Classification Neuro-  
Génétique et par l'Approche de Vapnik  
des Néoplasmes Mammographiques**

---

---

## **5 Classification Neuro-Génétique et par l'Approche de Vapnik des Néoplasmes Mammographiques**

---

« Le neurone est l'instigateur, du comportement humain. Le cerveau en est, sa conscience. » F.Z Reguieg

### **5.1 Introduction**

Les connaissances texturales et morphométriques mammaires, permettent d'orienter la classification des entités sénologiques selon, leur aspect bénin ou normal, versus malin. Cette discrimination supervisée, s'effectue dans l'ultime étape du système d'aide à la décision, au gré de deux approches. La première, catégorise les masses et les calcifications, suivant un modèle neuro-génétique. La seconde, permet la labellisation des tissus sains et des anomalies, au gré des supports à vaste marge.

La visée du système décisionnel est de guider le radiologiste, dans son interprétation sans pour autant le remplacer.

Ce chapitre, expose le modèle connexionniste et les séparateurs à vaste marge, en spécifiant les différents paramètres et fonctions de contrôle, y afférents.

### **5.2 Intelligence artificielle et Connexionnisme**

De la philosophie des esprits au cognitivisme, l'intelligence artificielle, a conquis cette ère. Elle est née pour simuler le comportement et le raisonnement du cerveau humain, pour une aide à la décision. Depuis sa création [156], elle ne cesse de révolutionner le monde de la scientificité. Engendrée de l'union entre la psychologie expérimentale et l'informatique, l'intelligence artificielle fait appel à diverses disciplines dont, la neurobiologie computationnelle, le principal vecteur du connexionnisme.

#### **5.2.1 Rôle de l'intelligence artificielle**

Le rôle clef de l'intelligence artificielle, est d'égaliser la perception humaine à travers l'automatisation de ses activités. S'inspirant du processus mental humain, l'IA allie l'esprit et la machine, dans la robotique, la vision et l'apprentissage automatique, impactant sur plusieurs secteurs, dont celui de la santé.

La figure 5.1 illustre le processus de l'IA. Ainsi, c'est à partir des données, que l'IA transmet les informations qui sont exploitées et transformées, en des connaissances. L'ultime étape est la prise de décision, suivant la considération des connaissances extraites. L'unité décisionnelle, se fonde sur le raisonnement et l'apprentissage.

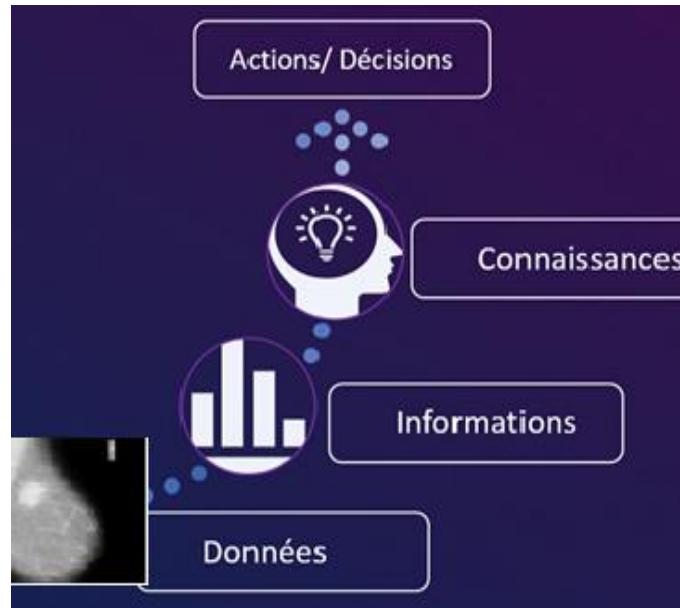


Figure 5.1 : Processus de l'IA [156]

### 5.2.2 Neurobiologie computationnelle

La neurobiologie computationnelle simule les mécanismes de la pensée, à travers la modélisation neuromimétique des neurones. Dans ce dernier cadre, il s'agit de modéliser les fonctionnalités du cerveau, autrement dit de rendre compte de certaines capacités (à apprendre, à s'orienter et se déplacer, etc...) en termes d'algorithmes ou calculs, effectués par une architecture et une dynamique neuronale. Pour désigner ces calculs qui n'ont a priori rien de numérique, on emploie le terme de computation (mot d'origine latine fort justement repris, par la langue anglaise). Les sciences expérimentales sur lesquelles s'appuient ces travaux théoriques, sont les neurosciences. Parmi ces disciplines, on cite en particulier les neurosciences computationnelles, qui s'intéressent au lien entre une architecture neuronale et sa fonction cognitive [60, 61].

### 5.2.3 Notion de connexionnisme

Dans le sillage du cognitivisme [61], le connexionnisme élabore des modèles informatiques, qui visent à simuler des phénomènes d'apprentissage par des réseaux de neurones artificiels. L'un des objectifs ultimes de cette réflexion, est de rivaliser avec la complexité du cerveau humain.

Le cognitivisme voit la pensée comme une manipulation de symboles, combinés par des règles abstraites d'inférence. Selon le connexionnisme, la pensée émerge de l'activité de plusieurs processeurs élémentaires, conçus autour de la neurobiologie humaine. Leur fonctionnement, découle sur la reconnaissance des formes ou des motifs.

### 5.2.4 Paradigme d'apprentissage

L'apprentissage [60, 62] est souvent considéré, comme la caractéristique principale de l'intelligence. De nombreuses études en IA, ont donc été consacrées à l'apprentissage et aux méthodes qui permettent à une machine, d'étendre ses connaissances.

Les informations captées, servent à prendre une décision. Cette perception artificielle inclut à la fois des méthodes d'apprentissage supervisé, non supervisé et par renforcement.

### **1. Apprentissage supervisé**

La formulation du problème de l'apprentissage supervisé est simple : on dispose d'un nombre fini d'exemples d'une tâche à réaliser, sous forme de paires (entrée et sortie désirée). On souhaite à cet effet, obtenir d'une manière automatique, un système capable de trouver de façon relativement fiable, la sortie correspondant à toute nouvelle entrée qui pourrait lui être présentée.

### **2. Apprentissage non supervisé**

Dans l'apprentissage non supervisé il n'y a pas de notion de sortie désirée, on dispose seulement d'un nombre fini de données d'apprentissage, constituées "d'entrées", sans qu'aucun label n'y soit rattaché.

### **3. Apprentissage par renforcement**

Cette méthode vise l'apprentissage du système d'intelligence artificielle en le soumettant à un environnement intégrant diverses situations. Le système est alors récompensé pour son comportement cumulatif, ce qui le guide dans son apprentissage. Contrairement à l'apprentissage supervisé, le système ne reçoit pas d'indications directes sur le résultat souhaité pour une situation particulière. Cette méthode permet donc au système, d'apprendre à traiter des situations nouvelles pour lesquelles, il n'a pas été entraîné au préalable. Il s'agit d'une approche, communément utilisée en robotique.

## **5.3 Réseaux connexionnistes**

De la biologie au connexionnisme, l'objectif de cette partie, est d'étudier les réseaux de neurones, pour établir notre choix concernant la classification des pathologies mammaires.

### **5.3.1 Genèse du connexionnisme**

A l'ère de la psychanalyse, le connexionnisme de S. Freud a retenu notre attention. A partir de ses réflexions psychanalytiques, les prémices de la plasticité neuronale, reposent sur deux états. Le premier est assimilé à des neurones perméables, servant à la perception et le second, considère des neurones imperméables, porteurs de mémoire, appartenant à des processus psychiques.

A l'ère de la cybernétique, l'histoire connexionniste est née en 1943, lorsque les neurologues McCulloch et Pitts [61] inventèrent le neurone formel, un produit scalaire entre un vecteur d'entrée et un vecteur poids, suivi d'un élément à seuil répondant au 0 ou au 1, lorsque le résultat du produit scalaire est supérieur ou non au seuil. Elle se poursuit ainsi, puis dans les années 1950, à l'époque où Hebb [61], proposait le principe de la plasticité synaptique comme modèle de la mémoire, Rosenblatt inventa le perceptron, identique au neurone formel, mais doué de la faculté d'apprentissage.

Il s'agit alors, d'un séparateur linéaire, qui permet d'apprendre à fournir la bonne décision, dans des opérations de classification. Vinrent ensuite Minsky et Papert en 1960, qui démontrèrent qu'avec un perceptron à une couche de neurones, on ne pouvait discriminer que des données linéairement séparables [61, 62]. Dans cet ordre d'idées, il a été conclu que pour des opérations plus complexes de la classification, il faudrait assembler plusieurs couches et effectuer un apprentissage au gré des facteurs synaptiques. Le perceptron multicouche pour la classification de données non linéaires, a ainsi vu le jour.

### 5.3.2 Neurone biologique

Un réseau connexionniste, est en fait un réseau de neurones artificiels basé sur un modèle simplifié de neurone. Ce modèle permet certaines fonctions du cerveau, comme la mémorisation associative, l'apprentissage par l'exemple, le travail en parallèle. Cependant, le neurone artificiel est loin de posséder toutes les capacités du neurone biologique. En effet, les réseaux de neurones biologiques (cf. Figure 5.2) sont ainsi beaucoup plus compliqués que les modèles mathématiques et informatiques [60, 61]. En effet, le neurone biologique possède trois principales composantes : les dendrites, le corps cellulaire et l'axone. Les dendrites forment un maillage de récepteurs nerveux qui permettent d'acheminer vers le corps du neurone, des signaux électriques en provenance d'autres neurones. Ce maillage agit comme un intégrateur en accumulant des charges électriques. Lorsque le neurone devient suffisamment excité (lorsque la charge accumulée dépasse un certain seuil), par un processus électrochimique, il engendre un potentiel électrique qui se propage à travers son axone, pour éventuellement venir exciter d'autres neurones. Le point de contact entre l'axone d'un neurone et le dendrite d'un autre neurone, s'appelle la synapse. L'arrangement spatial des neurones et de leur axone, ainsi que la qualité des connexions synaptiques individuelles, déterminent la fonction précise d'un réseau de neurones biologique. C'est en se basant sur ces connaissances, que le modèle mathématique a été défini.

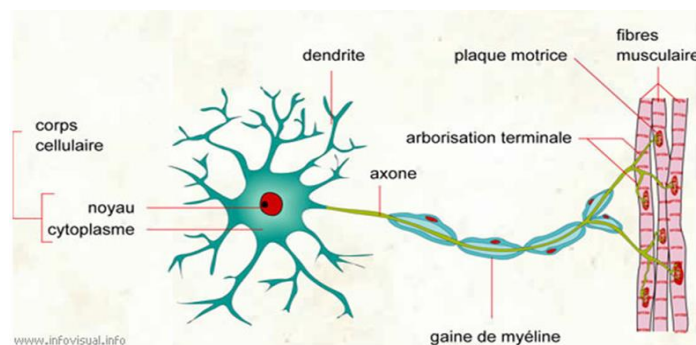


Figure 5.2 : Neurone biologique [61]

### 5.3.3 Modélisation d'un neurone

Le modèle mathématique d'un neurone artificiel, est illustré par la figure 5.3.



Un neurone est essentiellement constitué d'un intégrateur, qui effectue la somme pondérée de ses entrées. Le résultat  $n$  de cette somme est ensuite transformé, par une fonction de transfert  $f$  qui produit la sortie  $a$  du neurone [157].

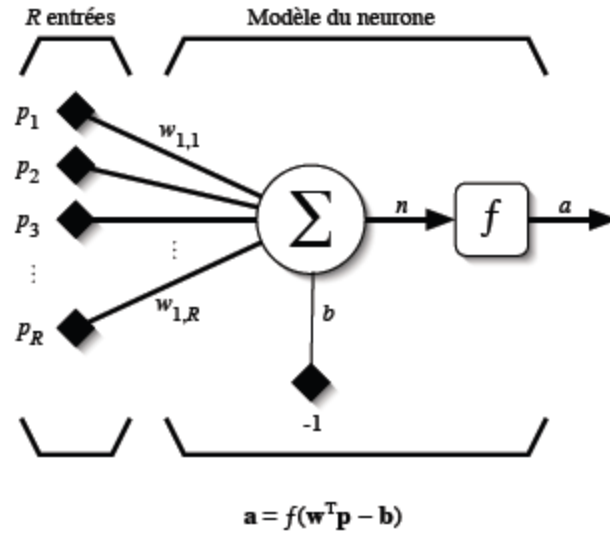


Figure 5.3 : Modèle d'un neurone artificiel [157]

Les  $R$  entrées du neurone correspondent au vecteur  $p = [p_1, p_2, \dots, p_n]^T$  alors que  $w = [w_{1,1}, w_{1,2}, \dots, w_{1,R}]^T$ , représente le vecteur des poids du neurone. La sortie  $n$  de l'intégrateur est donnée par l'équation suivante :

$$n = \sum_{j=1}^R w_{1,j} p_j - b = w_{1,1} p_1 + w_{1,2} p_2 + \dots + w_{1,R} p_R - b \quad 5.1$$

L'équation transcrite ci-dessus, peut aussi s'écrire sous forme matricielle :

$$n = \mathbf{w}^T \mathbf{p} - b \quad 5.2$$

Cette sortie correspond à une somme pondérée des poids et des entrées, moins ce qu'on nomme le biais  $b$  du neurone. Le résultat  $n$  de la somme pondérée, est le niveau d'activation du neurone. Le biais  $b$ , correspond au seuil d'activation du neurone. Lorsque le niveau d'activation atteint ou dépasse le seuil  $b$ , alors l'argument de  $f$  devient positif (ou nul). Sinon, il est négatif.

Un poids d'un neurone artificiel représente donc, l'efficacité d'une connexion synaptique. Un poids négatif vient inhiber une entrée, alors qu'un poids positif vient l'accentuer.

Lorsqu'on ajoute au modèle formulé par l'équation 5.2 la fonction d'activation  $f$ , on obtient la sortie du neurone (cf. Equation 5.3).

5.3

$$a = f(n) = f(W^T p - b)$$

En remplaçant  $w^T$  par une matrice  $W = w^T$  d'une seule ligne, on obtient la forme suivante :

$$a = f(W_p - b) \quad 5.4$$

L'équation 5.4 permet d'introduire un modèle plus compact, que celui de la figure 5.3. La figure 5.4 illustre celui-ci. On y représente les  $R$  entrées, comme un rectangle noir (le nombre d'entrées est indiqué sous le rectangle). De ce rectangle, sort le vecteur  $p$  dont la dimension matricielle, est  $R \times 1$ . Ce vecteur est multiplié par une matrice  $W$ , qui contient les poids (synaptiques) du neurone. Dans le cas d'un neurone simple, cette matrice possède la dimension  $1 \times R$ . Le résultat de la multiplication correspond au niveau d'activation, qui est ensuite comparé au seuil  $b$  par soustraction. Finalement, la sortie du neurone est calculée par la fonction d'activation  $f$ .

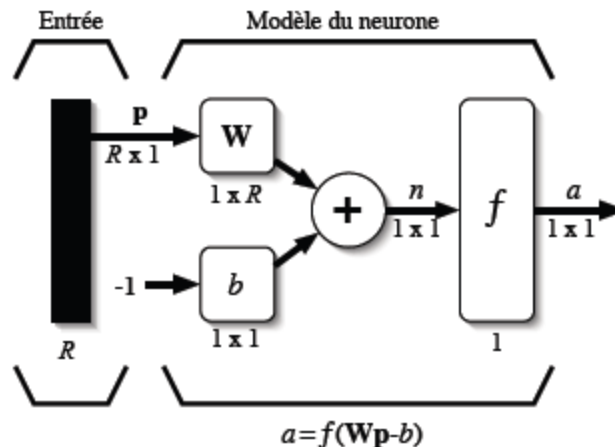


Figure 5.4 : Représentation matricielle d'un neurone artificiel [157]

### 5.3.4 Fonctions d'activation

Les fonctions d'activation des réseaux connexionnistes, permettent de transformer les signaux émis par les neurones de la couche d'entrée à l'aide, d'une fonction mathématique. Cette fonction  $f$ , peut profondément influencer la performance du réseau. Il est donc important, de bien choisir ce type de fonction.

Les fonctions d'activation les plus courantes sont les fonctions, linéaire, seuil, sigmoïde et tangente hyperbolique. Ainsi, selon les fonctions esquissées (cf. Figure 5.5), chacune d'elle, permet l'activation neuronale au gré de ses spécificités. Dans ce cas, comme son nom l'indique, la fonction 'seuil' applique un seuil sur son entrée. Plus précisément, une entrée négative ne passe pas le seuil, la fonction retourne alors la valeur 0 (on peut interpréter ce 0, comme signifiant 'faux'). Une entrée positive ou nulle, dépasse le seuil et la fonction retourne 1 (vrai).

La fonction linéaire est très simple, elle affecte directement son entrée à sa sortie ( $a=n$ ). On remarque alors que le biais  $b$  détermine l'emplacement du seuil, sur l'axe  $W^T p$ , où la fonction passe de 0 à 1.

Dans ce cas, la sortie du neurone correspond à son niveau d'activation, dont le passage à zéro, se produit lorsque  $W^T p = b$ .

La sigmoïde est un compromis intéressant entre les deux précédentes. Elle ressemble soit à la fonction seuil, soit à la fonction linéaire, selon que l'on est loin ou près de  $b$ . Son équation est générée, par la relation 5.5.

$$a = \frac{1}{1 + \exp^{-n}} \quad 5.5$$

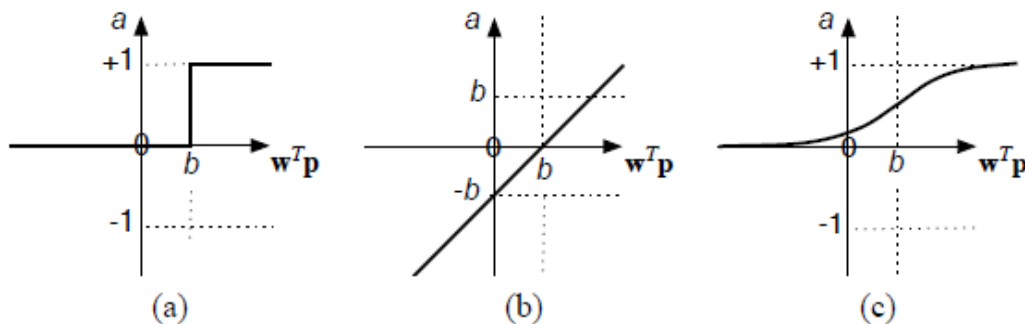


Figure 5.5 : Exemples de Fonctions d'activation connexionnistes [157]  
 (a) Fonction 'seuil' (b) Fonction linéaire (c) Fonction sigmoïde

### 5.3.5 Apprentissage des réseaux de neurones

L'apprentissage, pour les réseaux de neurones formels, consiste à calculer les paramètres de telle manière que les sorties du réseau de neurones soient, pour les exemples utilisés lors de l'apprentissage, aussi proches que possible des sorties désirées. Elles peuvent être le code de la classe à laquelle appartient la forme que l'on veut classer, la valeur de la fonction que l'on veut approcher ou celle de la sortie du processus que l'on veut modéliser, ou encore la sortie souhaitée du processus à commander.

Les techniques d'apprentissage des réseaux de neurones formels, sont des algorithmes d'optimisation. Ils cherchent à minimiser l'écart entre les réponses réelles du réseau et les réponses désirées, en modifiant les paramètres par étapes (appelées « itérations ») successives. La sortie du réseau de neurones s'adapte de mieux en mieux aux données au fur et à mesure, que l'apprentissage se déroule [61].

### 5.3.6 Connectivité

La connectivité des réseaux est la manière dont les neurones, sont connectés entre eux.

Elle peut être totale (tous les neurones sont connectés entre eux) ou organisée par couche (les neurones d'une couche ne sont connectés qu'à la couche précédente, en entrée et à la couche suivante, en sortie). Il existe des réseaux monocouches et multicouches [61].

### 5.3.7 Modèle du perceptron

Le Perceptron, un processus calculatoire connexionniste, né pour imiter les capacités du cerveau, a commencé à poindre au milieu du XX<sup>ème</sup> siècle. C'est un modèle inspiré des théories cognitives de F. Hayek et de D. Hebb [61]. Il s'agit d'un neurone formel muni d'une règle d'apprentissage qui permet de déterminer automatiquement les poids synaptiques de manière à séparer un problème d'apprentissage supervisé. Si le problème est linéairement séparable, un théorème assure que la règle du perceptron, permet de trouver une séparatrice entre les deux classes. Les prémices de l'intelligence artificielle, voient le jour.

#### 5.3.7.1 Perceptron simple

Le perceptron simple, permet de classifier des éléments compris dans deux catégories différentes ; à condition que qu'elles soient linéairement séparables. Par la suite, il est également possible de prédire, à quelle catégorie appartient un élément. Dans sa version la plus simple, le perceptron n'a qu'une seule sortie 'Y' à laquelle toutes les entrées 'Xi' sont connectées (cf. Figure 5.6).

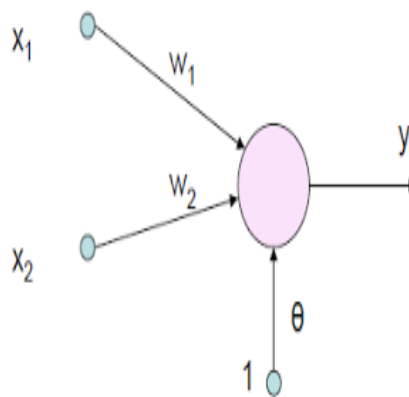


Figure 5.6: Perceptron simple [61]

#### 5.3.7.2 Perceptron multicouche

Le perceptron multicouche, est sans doute le plus connu des réseaux de neurones, adressé à la classification des données non linéaires. Il s'agit d'un classifieur neuronal formel organisé en plusieurs couches, au sein desquelles une information, circule de la couche d'entrée vers la couche de sortie. Chaque couche, est constituée d'un nombre variable de neurones. La structure du MLP, est représentée par la figure 5.7.

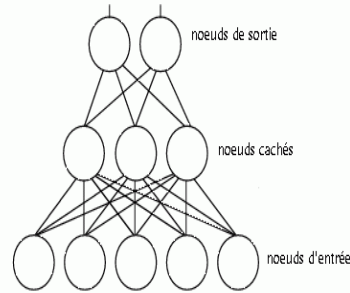


Figure 5.7 : Perceptron multicouche [61]

### 1. Les nœuds 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 mammographique, cette couche recevra les caractéristiques des signes cliniques. Sa taille est donc directement déterminée, par le nombre de variables d'entrées.

### 2. Les nœuds cachés

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. Cependant, c'est dans cette configuration, que sont générées les facteurs synaptiques, via une fonction d'activation non linéaire.

Le choix de la taille des couches cachées, n'est pas implicite et doit être, ajusté au gré des résultats attendus. Il sera alors préférable, pour obtenir de bon résultats, d'expérimenter le plus de tailles possibles.

### 3. Les nœuds de sortie

La troisième 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.

#### 5.3.7.3 Phase d'apprentissage du perceptron multicouche

L'apprentissage du perceptron multicouche, consiste à supprimer les connexions synaptiques et à les ajuster, pour s'approcher de la solution souhaitée. C'est la rétropropagation du gradient, qui estime les poids (appelés aussi, poids synaptiques du nom des synapses, termes désignant la connexion biologique entre deux neurones) pour le réseau à apprentissage supervisé ; qui consiste, à minimiser l'erreur quadratique de la sortie (somme des carrés de l'erreur de chaque composante entre la sortie réelle et la sortie désirée).

L'apprentissage est alors réalisé par l'ajustement des poids, obtenu directement à partir de l'erreur quadratique moyenne calculée, entre la sortie obtenue et la sortie désirée [157].

Cependant, la modélisation connexionniste, est une opération délicate, en raison du choix du nombre de neurones cachées et de la lenteur de l'entraînement.

Dans le cas de l'approche traitée, nous remédions à ces aléas, pour convenir d'une hybridation neuro-génétique, afin d'affiner la discrimination des anormalités mammaires.

## 5.4 Introduction aux algorithmes génétiques

Cette famille de techniques s'inspire de la théorie Darwinienne de l'évolution pour résoudre des problèmes d'optimisation. La théorie de l'évolution présentée par Charles Darwin [158] en 1859 dans son livre "l'origine des espèces", repose sur trois principes clefs : l'hérédité, la variation et la sélection. Ainsi, lorsqu'un organisme vivant se reproduit, il transmet ses caractéristiques à ses descendants, à travers ses gènes. Cependant l'hérédité des caractères n'est pas parfaite, les gènes peuvent subir des mutations et les descendants peuvent donc, présenter des variations de caractères. Par conséquent, les différents individus appartenant à une population d'organismes vivants ne sont pas tous identiques, et certains d'entre eux peuvent avoir des variations qui leur permettent de mieux survivre et de se reproduire davantage dans un certain environnement. Ces individus ont donc un avantage sélectif, plus de chances de transmettre leurs gènes à la génération future. À travers ce processus, les organismes s'adaptent à leur environnement, au cours des générations. Les algorithmes génétiques, et les algorithmes évolutionnaires, reposent sur ces principes.

Comment fonctionnent-ils ? Tout d'abord, on cherche à représenter les solutions possibles du problème d'optimisation sous la forme d'un génome, puis on génère une population de solutions potentielles au problème donné, on sélectionne ensuite les solutions les plus performantes vis à vis de la tâche à optimiser. On crée alors, une nouvelle population en copiant à l'identique les solutions sélectionnées, enfin on applique des opérateurs de variation aux génomes des individus de la nouvelle population, afin de créer des solutions différentes.

La population d'enfants devient à son tour la population parentale et on itère la même procédure, jusqu'à ce qu'une solution satisfaisante soit trouvée. En pratique il existe un grand nombre de façons de représenter les solutions potentielles à un problème, les codages les plus courants sont par exemple les vecteurs de nombres (binaires, entiers ou réels) et les graphes (par exemple des structures analogues aux arbres de décision). De même, il existe différentes méthodes de sélection (tournoi, par rang, uniforme...) et de variation (mutations ponctuelles, croisement) et le bon choix de ces différents aspects est crucial, pour obtenir des résultats pertinents.

Ces algorithmes permettent d'explorer l'espace de solutions possibles, de manière non exhaustive, afin d'obtenir une solution satisfaisante. Par conséquent, ils sont particulièrement utiles dans le cas d'espaces de très grande taille, présentant des optima locaux, difficiles à explorer avec des algorithmes déterministes d'optimisation.

De plus, ils s'adaptent facilement à des espaces de données qui changent dans le temps. C'est cette propriété, qui est usée dans la classification connexionniste adoptée, pour optimiser la phase d'apprentissage suivant des paramètres appropriés.

### 5.5 Approche neuro-génétique adoptée

Le modèle du perceptron multicouche adopté, est configuré par trois couches : une couche d'entrée, une couche cachée et, une couche de sortie pour la distinction des anomalies traitées (masses malignes et bénignes, calcifications malignes et bénignes). Le réseau neuronal apprend les similitudes des différents cas, à partir de leurs connaissances dans l'ensemble des données signées.

Le principe fondamental du classifieur choisi [99], est parcouru par l'entraînement du réseau connexionniste associé aux algorithmes génétiques. Pendant la phase d'apprentissage, les deux modèles d'entrée et de sortie sont présentés au perceptron afin, de lui permettre d'apprendre la signature de ces deux vecteurs. Les résultats corrects sont fournis au réseau, si bien que celui-ci puisse ajuster ses connexions synaptiques ( $W_i, B_i$ ) par la propriété d'optimisation des algorithmes génétiques. Après l'apprentissage, le classifieur est testé, en lui indiquant seulement les vecteurs d'entrée sans les sorties désirées et, en observant si la classe estimée, est proche de la classe désirée.

Dans un réseau neuronal classique [61], on intègre les entrées dans la couche primaire et, l'apprentissage se fait avec un certain nombre d'itérations et, de tests. L'arrêt s'effectue, lorsque l'erreur quadratique moyenne (cf. Equation 5.6) est faible ce qui, engendre un taux de classification performant. Cependant, l'inconvénient majeur de ce réseau, est la minimisation approximative de l'erreur quadratique, ainsi que la lenteur de l'apprentissage, particulièrement lorsque les classes sont difficilement séparables, ce qui produit un processus aléatoire.

La classification neuro-génétique, permet de remédier à cet aléa, en recherchant la fonction 'objectif' (cf. Equation 5.9), qui ajuste les poids ( $W_i$ ) et les biais ( $B_i$ ) de chaque catégorie mammaire, au gré de la génération optimale de la population. Au départ de l'algorithme, lorsqu'on spécifie le nombre de neurones, le réseau est créé, suivant la configuration accordée. Les biais et les poids sont initialisés aléatoirement, selon une distribution uniforme [99].

$$E = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2 \tag{5.6}$$

- Où, E représente l'erreur moyenne quadratique.
- N, est le nombre d'échantillons dans la base d'entraînement.
- $y_i$ , correspond à la classe estimée par le MLP pour le vecteur d'entrées  $x_i$ .
- $\hat{y}$ , est la classe désirée.

L'activation des couches neuronales, se fait par la fonction sigmoïde (cf. Equation 5.5), suivant les relations (5.7) et (5.8).

$$y_h = \varphi(w_1 \cdot x + b_1) \tag{5.7}$$

$$\hat{y} = w_2^T y_h + b_2 \tag{5.8}$$

Où,  $w_1$ ,  $w_2$ ,  $b_1$ ,  $b_2$  sont les connexions neuronales du perceptron multicouche, permettant de déterminer l'état des sorties de chacune des couches.

$\phi$  est la fonction d'activation du réseau de neurones, adoptée par la sigmoïde.

$Y_h$ , est la sortie de la couche cachée.

$w_1$  la matrice dont les poids synaptiques, assurent la connexion du vecteur d'entrée vers la couche cachée.

$w_2$  la matrice dont les poids synaptiques, assurent la connexion de la couche cachée vers la couche de sortie.

Ludwing [99], a consacré ses travaux, à l'amélioration de l'apprentissage du perceptron multicouche, en s'appuyant sur l'insertion de facteurs synaptiques dans la fonction E (cf. Equation 5.6), pour optimiser l'entraînement.

La relation (5.6), est modifiée, de façon à réduire l'espace d'entraînement, suivant la minimisation des valeurs propres des facteurs synaptiques, pour une catégorisation plus concise des classes. Elle se traduit alors, par l'équation (5.9).

$$F = V_{\min} + V_{\max} + \frac{c}{N} \sum_{i=1}^N H(y_i, \hat{y}) \tag{5.9}$$

Où,  $V_{\min}$  arbore la valeur minimale de  $w_1 w_1^T$ .

$V_{\max}$  esquisse la valeur maximale de  $w_1 w_1^T$ .

$C$ , indique, le paramètre de régularisation.

$H(y_i, \hat{y})$  représente le facteur maximal des fausses classes des données d'entrée.  $X_i$  régulé, par le paramètre  $C$ .

$\hat{y}$  est la classe désirée.

$y_i$  indique la classe estimée.

Les étapes d'apprentissage du réseau de neurones, en hybridation avec les algorithmes génétiques, sont ponctuées par l'algorithme ci-dessous :



---

Algorithme d'apprentissage neuro-génétique

---

X, matrice de la signature des classes mammaires pathologiques suivant un vecteur à N échantillons.

Y, modèles des classes (-1 pour les masses, 1 pour les calcifications).

Nnc: nombre de neurones cachés.

C: paramètre de régularisation.

maxgener: nombre maximal de générations.

Pops: taille de la population.

Sorties: w1, w2, b1, b2 (paramètres neuronaux NN).

1. Générer un ensemble de chromosomes {Cr}, pour une taille Pops, pour la population initiale, en prenant en compte, le nombre d'éléments d'entrée et le nombre de neurones cachés. Par conséquent, chaque chromosome est un vecteur  $C_r = [w_1, \dots, w_n; b_1, \dots, b_n]$ , contenant tous les poids et les biais générés aléatoirement.
  2. Pour génération = 1: maxgener faire: pour PS = 1 à Pops, réorganiser les gènes  $C_r^{ps}$  afin de composer les paramètres  $w_1, w_2, b_1$ , et  $b_2$ .
  3. Pour  $i = 1: N$  calculer  $\hat{y}_i$  des individus 'ind' selon l'équation (5.8), en usant des poids et biais.
  4. Calculer  $F$  selon la relation (5.9).
  5. Stocker la fonction objective des individus 'ind'.
  6. Classer les individus selon leur fitness  $F_{ps}$ .
  7. Stocker les gènes du meilleur individu dans  $C_r^{best}$ .
  8. Réarranger les gènes du meilleur individu  $C_r^{best}$  pour composer les facteurs synaptiques NN ( $w_1, w_2$ , et  $b_1$  et  $b_2$ ).
  9. Introduire les facteurs synaptiques optimisés dans le perceptron multicouche.
  10. Classifier un nouveau mammogramme référencé par sa signature.
- 

Algorithme 5.1 : Apprentissage et Catégorisation neuro-génétique de masses et de calcifications dans les mammogrammes numériques

## 5.6 Paradigme du classifieur de Vapnik

Introduits par Vapnik [63, 64] en 1995, les machines à vecteurs de support, ou les séparateurs à vaste marge, sont un ensemble de techniques d'apprentissage supervisé, destinées à résoudre des problèmes de discrimination et de régression. Elles sont utilisées dans la reconnaissance de formes, la catégorisation de texte ou encore, le diagnostic médical, l'objet de notre investigation. Les SVM reposent sur deux notions : celle de marge maximale et, celle de fonction noyau. Le principe de base des SVM, consiste de ramener le problème de la discrimination à celui, de la recherche d'un hyperplan optimal. Deux idées permettent d'atteindre cet objectif :

- ❖ La première consiste à définir l'hyperplan, comme solution d'un problème d'optimisation, sous contraintes dont la fonction 'objectif' ne s'exprime qu'à l'aide de produits scalaires, entre vecteurs et dans lequel le nombre de vecteurs supports contrôle, la complexité du modèle.
- ❖ Le passage à la recherche de surfaces séparatrices, non linéaires est obtenu par l'introduction d'une fonction noyau (kernel) dans le produit scalaire induisant, implicitement une transformation non linéaire des données vers un espace intermédiaire, de plus grande dimension.

- ❖ Ce concept exprime bien, l'appellation couramment rencontrée de machine à noyau (kernel machine). Sur le plan théorique, la fonction noyau définit un espace hilbertien, dit auto-reproduisant et isométrique par la transformation non linéaire de l'espace initial et dans lequel, est résolu le problème linéaire (cf. Figure 5.8).

La marge est la distance, entre la frontière de séparation et, les échantillons les plus proches appelés vecteurs support. Dans un problème linéairement séparable, les SVM trouvent une séparatrice qui maximise, cette marge. Dans le cas d'un problème non linéaire, on utilise une fonction noyau pour projeter les données dans un espace, de plus grande dimension où, elles seront linéairement séparables (cf. Figure 5.8).

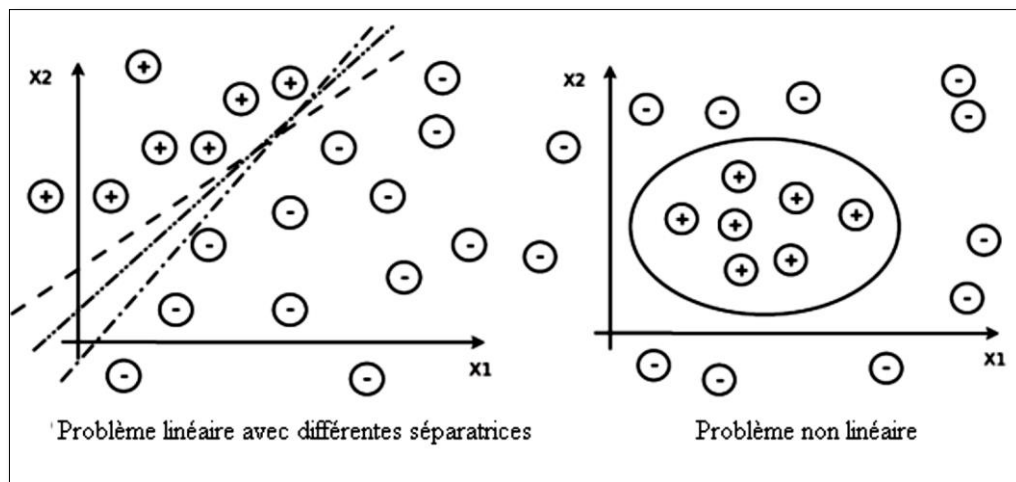


Figure 5.8 : Exemples de séparation linéaire et non linéaire de données [63]

### 5.6.1 Principe

La technique des séparateurs à vaste marge [159], repose sur le concept de plans de décision définissant, des frontières de décision. Dans cette optique, un plan de décision, est un plan qui permet de séparer un ensemble d'objets, appartenant à des classes différentes.

L'illustration de la figure 5.9 constitue, un exemple représentatif. Dans cet exemple, les objets appartiennent soit à la classe verte, soit à la classe rouge. La droite de séparation définit une frontière, à droite de laquelle, tous les objets sont verts et à gauche de laquelle tous les objets, sont rouges. Tout nouvel objet (matérialisé par un cercle blanc), apparaissant à gauche, sera affecté à la classe rouge (ou au contraire affecté à la classe verte, s'il se trouve à droite de la ligne de séparation).

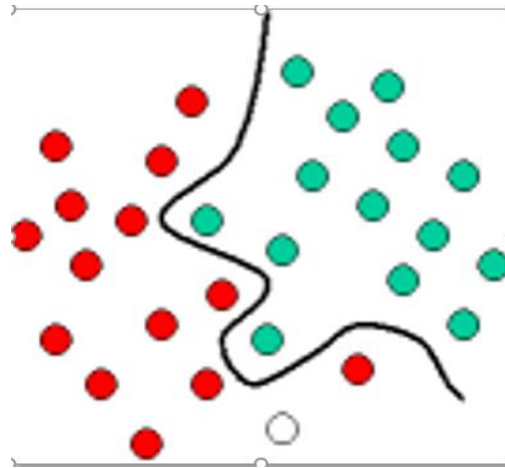


Figure 5.9 : Exemple de discrimination linéaire par les SVM [159]

L'exemple classique illustré ci-dessus, est un modèle de classification linéaire, permettant de répartir un ensemble d'objets, dans leurs groupes respectifs (vert et rouge dans le cas présent) à l'aide d'une droite.

Dans la configuration de la figure 5.10, nous pouvons séparer les objets projetés de la figure 5.9 par une ligne droite et par conséquent, au lieu de construire une courbe complexe (comme à gauche du schéma), nous pouvons nous contenter de trouver une droite, qui va séparer de façon optimale les objets verts des objets rouges [159].

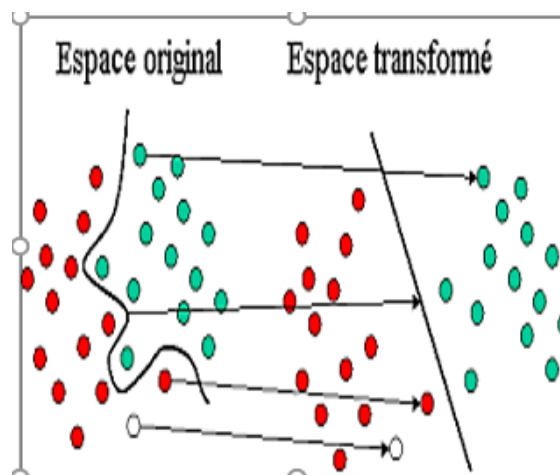


Figure 5. 10 : Configuration transformée des SVM [159]

### 5.6.2 SVM à marge dure

Si les données sont linéairement séparables, on parle d'une machine à vecteur support à marge dure (cf. Figure 5.11). L'hyperplan séparateur est représenté par l'équation (cf. Equation 5.10).

$$H(x) = W^T x + b \tag{5.10}$$

Où  $H$  est l'hyperplan séparateur,  
 $X$  représente le vecteur d'entrée,  
 $W$  correspond au vecteur poids de dimension  $m$ ,  
 $B$  notifie le biais.

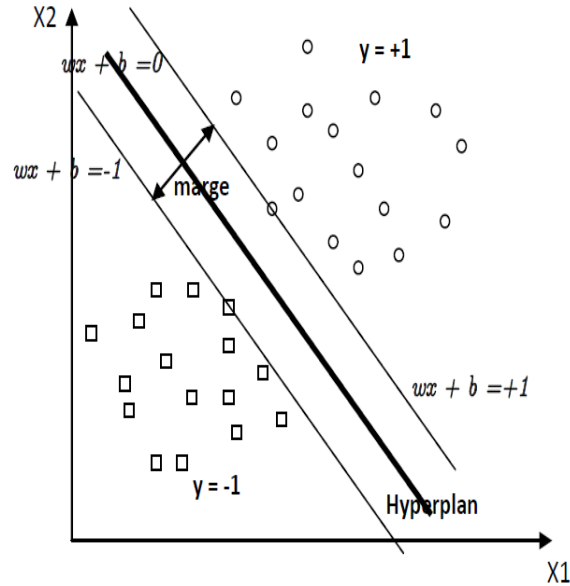


Figure 5.11 : Graphe représentatif d'un SVM binaire à marge dure [160]

La fonction de décision, pour un exemple  $x$ , est alors exprimée par la formulation exprimée dans les relations suivantes :

$$\begin{cases} \text{Classe} = 1, \text{ si } H(x) > 0 \\ \text{Classe} = -1, \text{ si } H(x) < 0 \end{cases} \quad 5.11$$

Les deux classes étant linéairement séparables, il n'existe aucun exemple qui se situe sur l'hyperplan ( $H(x) = 0$ ). Il convient alors d'utiliser la fonction de décision 5.12.

$$\begin{cases} \text{Classe} = 1, \text{ si } H(x) > 1 \\ \text{Classe} = -1, \text{ si } H(x) < -1 \end{cases} \quad 5.12$$

La région qui se trouve entre les deux hyperplans des relations 5.13 et 5.14, est appelée la région de généralisation de la machine d'apprentissage. Plus cette région est importante, plus est grande, la capacité de généralisation de la machine. La maximisation de cette région est l'objectif de la phase d'entraînement qui consiste, à rechercher l'hyperplan qui maximise la marge. L'hyperplan de marge maximale va certainement maximiser la capacité de généralisation de la machine d'apprentissage [63, 160].

La détermination de l'hyperplan optimal, passe par la détermination de la distance euclidienne minimale entre l'hyperplan et l'exemple, le plus proche des deux classes. Puisque le vecteur  $w$  est orthogonal sur l'hyperplan séparateur, la droite parallèle à  $w$  et reliant un exemple  $x$  à l'hyperplan, est notifiée par l'équation 5.13.

$$\frac{aW}{\|W\|} + x = 0 \tag{5.13}$$

Où,  $a$  représente la distance entre  $x$  et l'hyperplan.

La relation 5.13 découle sur l'équation 5.14 pour sa résolution.

$$a = -\frac{W^T x + b}{\|W\|} \tag{5.14}$$

La distance de tout exemple de l'hyperplan doit être supérieure, ou égale à la marge  $\delta$ .

$$\frac{y_i(W^T x_i + b)}{\|W\|} \geq \delta \tag{5.15}$$

A cette relation, la contrainte (5. 16) est donnée.

$$\|W\|\delta \geq 1 \tag{5.16}$$

L'hyperplan séparateur optimal, est atteint par la résolution de l'équation 5.17.

$$\left\{ \begin{array}{l} \text{Minimiser } \frac{1}{2}\|W\|^2 \\ \text{sous contrainte s} \\ y_i(W^T x_i + b) \geq 1 \end{array} \right. \tag{5.17}$$

Les données qui vérifient l'égalité de la contrainte, représentent les vecteurs supports.

### 5.6.3 SVM à marge souple

Dans le cas où les données ne sont pas linéairement séparables [159, 160], il est possible de suivre la même démarche adoptée dans le cas séparable avec, l'introduction de variables d'écart. L'idée est de modéliser les erreurs potentielles, par des variables d'écart positives  $\xi$  associées à chacune des observations  $(x_i, y_i), i \in [1, n]$ . A cet effet, deux cas se présentent selon les relations (5.18) et (5.19).

$$y_i(W^T x_i + b) \geq 1 \Rightarrow \xi_i = 0 \tag{5.18}$$

$$y_i(W^T x_i + b) < 1 \Rightarrow \xi_i = 1 - y_i(W^T x_i + b) > 0 \tag{5.19}$$

Si  $\xi_i < 1$ , la donnée  $x_i$  ne respecte pas la marge mais reste bien classée, sinon  $x_i$  est mal classée par l'hyperplan. Dans ce cas, le problème consiste alors à maximiser la marge et minimiser aussi, la somme des erreurs permises  $Q(W) = \sum_{i=1}^n \xi_i$  simultanément.

On constate que si toutes les variables d'écart  $\xi_i = 0$ , on retrouve le problème des SVM linéairement séparables. La résolution du problème précédent peut s'effectuer grâce à l'introduction d'un terme d'équilibrage  $C > 0$  fixé qui permet de poser le problème sous la forme suivante :

$$\left\{ \begin{array}{l} \text{Minimiser } \frac{1}{2} \|W\|^2 + C \sum_{i=1}^n \xi_i \\ \text{sous les contraintes} \\ y_i (W^T x_i + b) \geq 1 - \xi_i \quad i = 1..n \\ \xi_i \geq 0 \end{array} \right. \quad 5.20$$

Où  $C$ , est un paramètre positif qui assure un équilibre, entre la maximisation de la marge et la minimisation de l'erreur de classification.

Dans l'exemple de la figure 5.12, tous les écarts sont nuls sauf un, celui du point bleu mal classé. Cet écart mesure la distance du point à la marge numérique, de l'hyperplan séparateur.

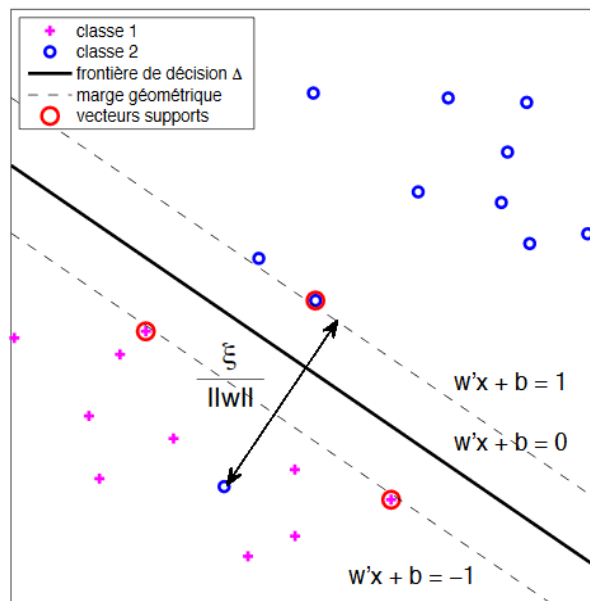


Figure 5.12 : Graphe représentatif d'un SVM à marge souple [160]

### 5.6.4 Fonctions à noyaux

La tâche de classification, n'est malheureusement pas aussi simple et il faut souvent recourir à des structures plus complexes, pour trouver une séparation optimale, c'est-à-dire classer correctement de nouveaux objets (les observations de test) sur la base des exemples disponibles (observations d'apprentissage). C'est le cas dans l'illustration de la figure 5.10.

Par rapport au schéma 5.9, il est évident que nous avons besoin d'une courbe (qui est plus complexe qu'une droite), pour séparer totalement les objets verts des objets rouges. Les tâches de classification qui utilisent des lignes de séparation, pour faire la distinction entre des objets appartenant à des classes distinctes, sont connues sous le nom de modèles de classification par hyperplans de séparation. Les SVM sont particulièrement bien adaptés pour traiter ce type de tâches. Dans le cas esquissé par la figure 5.10, nous pouvons voir que les objets originaux (situés à gauche du schéma) sont transformés, c'est-à-dire réorganisés, à l'aide d'un ensemble de fonctions mathématiques appelées noyaux. Le processus de réorganisation des objets, est appelé projection (ou transformation) [64, 159-160].

La transformation d'espace, est souvent réalisée à l'aide d'une fonction scalaire  $F = \{\phi(x) | x \in X\}$ . Le problème revient alors, à optimiser l'équation 5.21.

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m \alpha_i \alpha_j y_i y_j \langle \phi(x_i), \phi(x_j) \rangle \quad 5.21$$

Où  $\langle \phi(x_i), \phi(x_j) \rangle$  est le produit scalaire des deux images des vecteurs  $x_i$  et  $x_j$  dans le nouvel espace et dont le résultat est un scalaire.  
 $\alpha_i$  et  $\alpha_j$  correspondent aux coefficients multiplicateurs de Lagrange.

Les multiplicateurs de Lagrange, sont assujettis aux conditions de Karush-Kuhn-Tucker [105], pour la recherche de l'hyperplan optimum.

Le produit scalaire  $\langle \phi(x_i), \phi(x_j) \rangle$  est alors déterminé par un noyau, au gré d'une fonction symétrique à deux variables (5.22).

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle \quad 5.22$$

La fonction noyau, doit respecter les conditions de Mercer [159]. Ceci insinue que cette fonction doit être symétrique, sans aucune valeur propre négative.

L'équation 5.21 devient alors,

$$Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad 5.23$$

La fonction de décision est alors arborée par la relation 5.24.

$$H(x) = \sum_{i \in S} \alpha_i y_i K(x_i, x_j) + b \quad 5.24$$

Où S représente l'ensemble des vecteurs supports.

Parmi les fonctions à noyau les plus usités, on distingue :

❖ **Noyau linéaire**

Ce noyau (5.25) intervient dans la fonction de décision, si les données sont linéairement séparables. Il n'est pas alors nécessaire de changer d'espace.

$$K(x_i, x_j) = x_i x_j \tag{5.25}$$

❖ **Noyau radial gaussien**

Le noyau gaussien (5.26) réalise la distinction des classes, en assurant la maximisation de la marge selon le paramètre  $\sigma$ .

$$k(x_i, x_j) = \varphi(x_i) \varphi(x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{5.26}$$

Où  $\sigma$  est un réel positif qui représente la largeur de bande du noyau.

**5.6.5 Classification multi-classes**

Les SVM ont été conçus, pour une classification binaire. Cependant, dans la plupart des contextes, nous cherchons à résoudre des problèmes multi-classes. Le principe de base, consiste à une composition de plusieurs hyperplans binaires, permettant de tracer les frontières de décision entre les différentes classes. Les modèles les plus connues, sont considérés par les stratégies 'un contre tous' et 'un contre un' [159-161].

**1. Un contre tous**

La méthode un contre tous, consiste à construire K classifieurs binaires en attribuant le label 1, aux échantillons de l'une des classes et le label -1, à toutes les autres. En phase de test, le classifieur donnant la valeur de confiance (la marge) la plus élevée, remporte le vote.

La stratégie observée, détermine alors pour chaque classe K, l'hyperplan  $H_K(W_K, b_K)$ . Celui-ci est identifié par la relation 5.27.

$$H_k(x) = \text{signe}(\langle W_k, x \rangle + b_k) = \begin{cases} +1 & \text{si } f_k(x) > 0 \\ 0 & \text{sinon} \end{cases} \tag{5.27}$$

Où  $f_k(x)$  notifie la fonction de décision de la relation 5.24.

La valeur de l'hyperplan, permet de classer l'échantillon x. Dans le cas où, le label n'est pas précisé, il faudrait estimer l'expression de la fonctionnelle 5.28.

$$k^* = \text{Arg}_{1 \leq k \leq K} \text{Max}(H_k(x)) \tag{5.28}$$

Dans certains cas, la relation ci-dessus, peut engendrer des faux positifs. Ce qui engendre, des ambiguïtés. A cet égard, la classe retenue, est celle qui maximise l'expression,  $f_k(x) = \langle W_k, x \rangle + b_k$  au diapason de la relation 5.29.



$$k^* = \text{Arg}_{1 \leq k \leq K} \text{Max}(\langle W_k, x \rangle + b_k) \quad 5.29$$

Tout nouvel exemple  $x$  est affecté à la classe dont l'hyperplan est le plus loin de  $x$ , parmi les classes ayant  $H(x) = 1$ .

## 2. Un contre un

La méthode 'un contre un' consiste, à construire  $K(K-1)/2$  SVM binaires. La classe gagnante est obtenue soit par un vote majoritaire, soit grâce à l'estimation de probabilités a posteriori. Pour chaque paire de classes  $(k, s)$ , on établit une fonction de décision binaire (5.29),  $H_{ks} : \mathbb{R} \rightarrow \{-1, +1\}$ .

$$H_{ks} = \text{signe}(f_{ks}(x)) \begin{cases} +1 & \text{si } f_{ks}(x) > 0 \\ 0 & \text{sinon} \end{cases} \quad 5.30$$

Pour  $K(K-1)/2$  classes,  $K$  fonctions de décision sont définies (5.31).

$$H_k(x) = \sum_{s=1}^m H_{ks}(x) \quad 5.31$$

La règle de classification d'un nouvel échantillon, est exprimée par la formulation 5.32.

$$k^* = \text{Arg}_{1 \leq k \leq K} (\text{Max } H_k(x)) \quad 5.32$$

La stratégie 'un contre un', présente l'avantage d'être rapide et efficace. Le problème d'ambigüité est levé, du fait du traitement de chaque couple de classes.

### 5.6.6 Classification des mammogrammes par les séparateurs à vaste marge

La stratégie "un contre un", dans le cas de notre approche se fait suivant un noyau gaussien (5.26), pour la distinction des classes en maximisant la distance les séparant. Le processus pris en compte dans cette approche, est diligenté par l'algorithme 5.2.

L'estimation de la probabilité a posteriori, permet de discriminer les classes néoplasiques en se référant aux relations qui suivent.

En se basant sur la relation 5.33, qui s'appuie sur le fait que la fonction de discrimination idéale, adopte le même signe que celle de  $f(x)+b$  ; il est possible alors, d'obtenir un estimateur de la probabilité a posteriori selon la relation 5.34.

$$\log \frac{\hat{P}(Y = 1/x)}{\hat{P}(Y = -1/x)} = a_1 f(x) + a_2 \quad 5.33$$

$$\hat{P}(Y = 1/x) = 1 - \frac{1}{1 + e^{a_1 f(x) + a_2}} \quad 5.34$$

Où  $a_1$  et  $a_2$  sont des paramètres, estimés par le maximum de vraisemblance.

Cette condition implique que les SVM, sont efficaces autour d'un noyau adéquat.

---

**Algorithme de classification de mammogrammes par les SVM**

---

$X$ , matrice de la signature des classes mammaires pathologiques suivant un vecteur à  $n$  échantillons.

$Y$ , modèles des classes (-1 pour les anormalités bénignes, 1 pour les lésions malignes et 0 pour les cas sains).

$C$  : paramètre de régularisation.

Sorties : anormalité bénigne, lésion maligne ou tissu sain.

1. Choix de la fonction scalaire  $k(x_i, x_j)$  selon le noyau gaussien (5.26)
  2. Détermination des vecteurs poids  $W$ , par optimisation selon les contraintes de la relation (5.20).
  4. Elaboration de la fonction de décision, permettant la maximisation de la marge, pour chaque paire de classes (5.31).
  5. Evaluation des probabilités a postériori pour chaque catégorie  $\hat{P}(Y = 1/x)$  selon 5.34.
  6. Obtention de la base d'apprentissage par la distinction des vecteurs supports  $S$ .
  7. Test d'un nouveau mammogramme par le calcul à nouveau, de la fonction de décision suivant le noyau gaussien et les vecteurs supports.
  7. Prise de décision au gré du maximum a postériori, selon la relation 5.34.
- 

Algorithme 5.2 : Classification des mammogrammes par la stratégie un contre un

## 5.7 Conclusion

Ce chapitre a fait l'objet d'une étude des classifieurs adoptés, au gré des structures analysées. Les paramètres usés, permettent la régulation des modèles neuro-génétique et de Vapnik, pour minimiser les faux positifs. L'ultime chapitre, présente le système de radiodiagnostic élaboré, pour une aide à l'interprétation des mammogrammes complexes et uniformes.

---

## **Chapitre 6 Détection des Anormalités et des Tissus sains dans les Mammogrammes pour l'Aide à la Décision**

---

---

## **6 Détection des Anormalités et des Tissus Sains dans les Mammogrammes pour l'Aide à la Décision**

---

« Le fait suggère l'idée, l'idée dirige l'expérience, l'expérience juge l'idée. » C. Bernard

### **6.1 Introduction**

La discrimination radiologique des anomalies du sein, se focalise sur leur morphogénèse, pour une interprétation clinique. Cependant, lorsque les structures tissulaires du sein sont denses, la délimitation de lésions éventuelles s'avère délicate, pour une prise de décision. La mammographie révèle ses limites.

Le recours à la vision par ordinateur, devient un élément clé, pour la correction de la lecture radiologiste de mammogrammes.

L'objectif majeur de ce chapitre, est d'exposer le système d'aide à la détection et au diagnostic des anomalies du sein, réalisé. C'est dans cette visée à double sens, que le CAD conçu dans nos travaux, aborde une conduite radio-diagnostique. Le système en question est diligenté, par une hybridation de méthodes pour la localisation minutieuse des régions d'intérêt ainsi que, pour leur description sémiologique. L'ultime étape de la réflexion appréhendée, est la classification supervisée des mammogrammes. Une classification neuro-génétique, catégorise les masses et les calcifications. Une seconde discrimination est arborée par les séparateurs à vaste marge, pour labelliser les tissus sains et les lésions.

### **6.2 Système d'aide au diagnostic des anomalies mammographiques conçu**

Notre approche technique inspirée par une vision clinique, a fait l'objet d'une analyse minutieuse des mammogrammes MIAS, pour une classification dirigée des signes pathologiques du sein. Le système de radiodiagnostic appréhendé (cf. Figure 6.1), parcourt trois étapes clés, en amont de la décision.

1. Une phase de correction de mammogrammes, autour d'opérateurs morphologiques.
2. Une phase de délimitation des régions suspectes, suivant un partitionnement adaptatif arboré, par les modèles déformables, une agrégation de pixels homogènes et un regroupement spatial, suivant l'approche de la logique floue.
3. Une phase de description des signes cliniques, suivant une hybridation de la texture et de la morphométrie, pour établir la signature mammaire.
4. Création de la base de connaissances mammaires, autour de deux classifieurs.

Le perceptron multicouche associé aux algorithmes génétiques, catégorise les masses et les calcifications. Les supports à vaste marge, distinguent les anomalies des tissus sains mammaires.

Les classes obtenues par l'approche neuro-génétique, répondent aux critères de malignité et de bénignité, des masses et des calcifications. Les catégories engendrées par les SVM, différencient les tissus sains des anomalies malignes et bénignes.

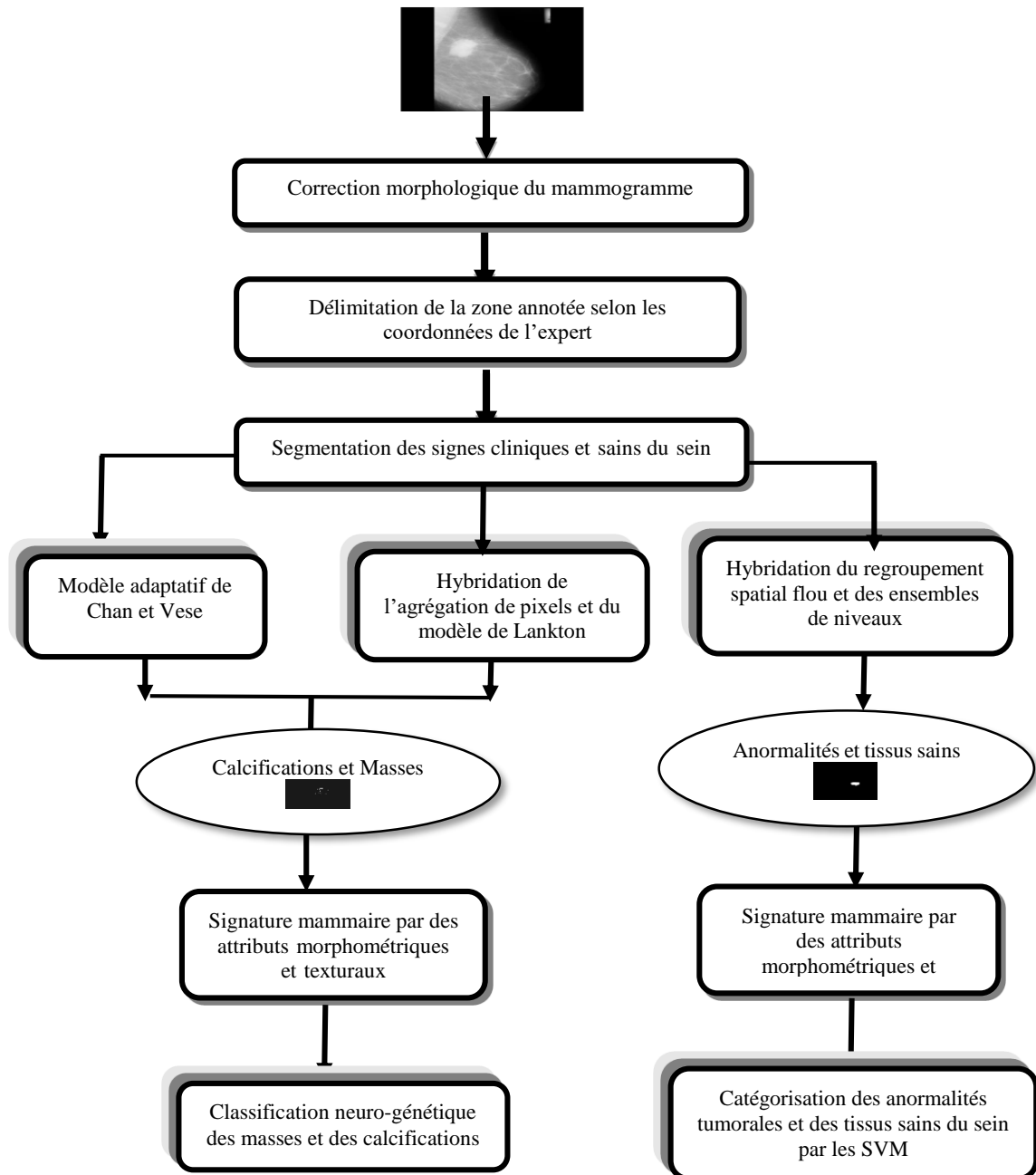


Figure 6.1 : Système d'aide au diagnostic des anomalies et des tissus sains mammographiques

### 6.3 Description de la base de mammogrammes MIAS

Le système d'aide à la décision adopté est appliqué sur des mammographies réelles générées, à partir de la base de données MIAS (Mammographic Image Analysis Society) [106]. Ces images de type MLO (incidence oblique médio-latérale), numérisées avec une résolution de 50 microns par pixel avec 8 bits, ont des dimensions spatiales de 1024 x 1024 pixels. La base de données MIAS contient 322 images, correspondant aux radiographies normales, bénignes et malignes du sein.

Les mammogrammes de la base MIAS (cf. Figures 6.2, 6.3, 6.4 et 6.5), comprennent 207 clichés normaux, 25 images révélant des calcifications et 90 cas incluant des anormalités (masses, distorsions architecturales et asymétries de densité). Les clichés explorés, arborent trois types de structures tissulaires (graisseuses, glandulaires et denses).

La base MIAS est annotée par des experts. Les régions suspectieuses sont ainsi, repérées par des coordonnées radiales spécifiques (centre et le rayon du cercle entourant la zone suspecte), à chaque cas.

Les images considérées dans l'approche décisionnelle élaborée, se répartissent en deux sous-ensembles. Le premier prend en compte, 90 images contenant des masses et 25 mammogrammes comprenant des calcifications ; répondant aux néoplasmes bénins et malins.

Le second ensemble met l'accent sur les 207 mammogrammes normaux de la base et, les 115 anormalités des deux catégories (bénignes et malignes).

L'investigation de cette réflexion scientifique, se concentre sur l'analyse de la zone clinique annotée, afin de minimiser les faux positifs.

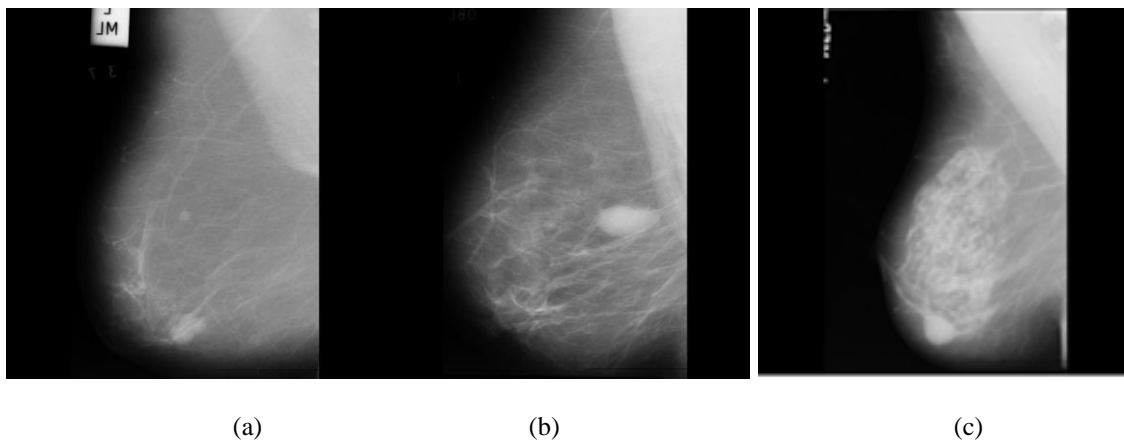
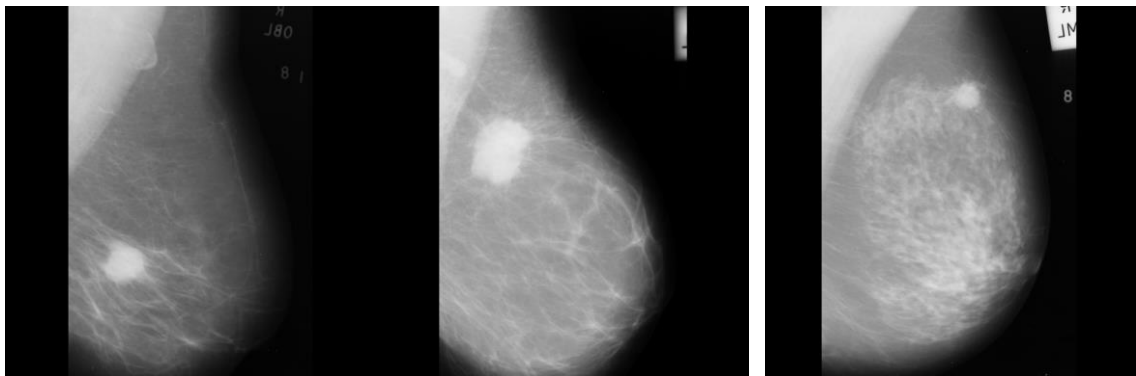


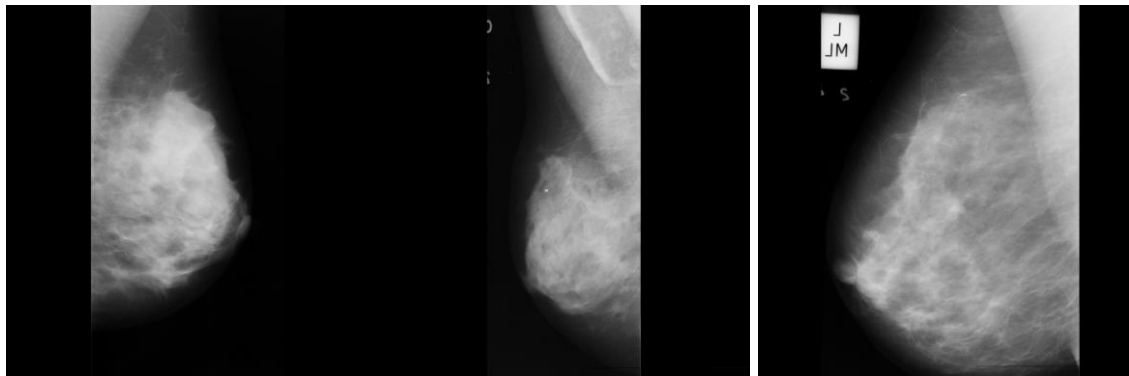
Figure 6.2 : Exemples de mammogrammes bénins de la base MIAS [66]

(a) mdb005 (b) mdb025 (c) mdb021



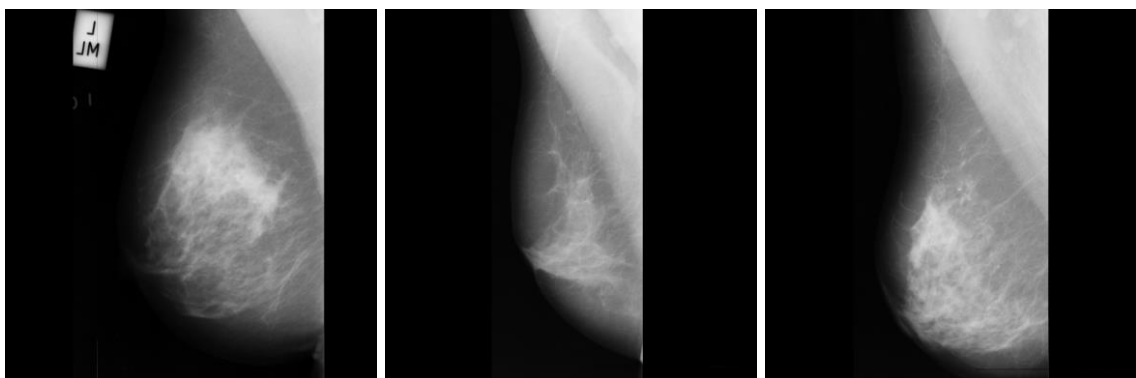
(a) (b) (c)

Figure 6.3 : Exemples de mammogrammes malins de la base MIAS [66]  
mdb028 (b) mdb184 (c) mdb202



(a) (b) (c)

Figure 6.4 : Exemples de mammogrammes avec calcifications bénignes [66]  
(a) mdb226 (b) mdb223 (c) mdb219



(a) (b) (c)

Figure 6.5 : Exemples de mammogrammes avec calcifications malignes [66]  
(a) mdb209 (b) mdb213 (c) mdb249

## 6.4 Approche neuro-génétique pour la détection morphologique de masses et de calcifications

Notre première contribution aborde la détection de masses et de calcifications, autour d'un système semi-automatique d'aide à la décision.

Cette réflexion se distingue, par deux ensembles. Le premier met en exergue les signes recherchés, par le modèle adaptatif de Chan et Vese. Le second, détecte les anormalités du sein, en se basant sur une coopération séquentielle entre une croissance de régions et, la segmentation évolutive de Lankton.

Le système d'aide au diagnostic conçu, parcourt la détection des signes cliniques mammaires, selon trois phases. Les images sont en premier lieu, corrigées par des techniques morphologiques, afin de rehausser le contraste, pour mettre en évidence les structures recherchées. L'opération angulaire du système, arborée en second lieu, est la segmentation des images. Elle est basée sur les ensembles de niveaux, qui considèrent ainsi, l'information locale au gré de critères spécifiques, de manière à localiser minutieusement les régions d'intérêt.

En troisième lieu, les régions obtenues, sont caractérisées suivant des critères de malignité et de bénignité, établis par la texture et la morphométrie.

L'ultime phase de l'analyse systémique, est une classification neuro-génétique des masses et des calcifications, selon des catégories malignes versus bénignes.

La méthodologie conçue selon la première vision (cf. Figure 6.6), a fait l'objet d'une publication dans un journal international d'ingénierie biomédicale et technologique [127]. L'étude en question, parcourt la détection des signes cliniques mammaires, autour de la philosophie de la VAO. Les images sont en premier lieu, corrigées par des techniques morphologiques, afin de rehausser le contraste, pour mettre en évidence les structures recherchées. L'extraction concise des signes cliniques, est instituée par le modèle adaptatif de Chan et Vese, moyennant une initialisation judicieuse.

En troisième lieu, les régions obtenues, sont caractérisées suivant des critères de malignité et de bénignité, établis par l'estimation de la granularité et l'appréciation de la forme. L'ultime phase de l'analyse systémique, est une classification neuro-génétique des masses et des calcifications, selon des catégories malignes versus bénignes.



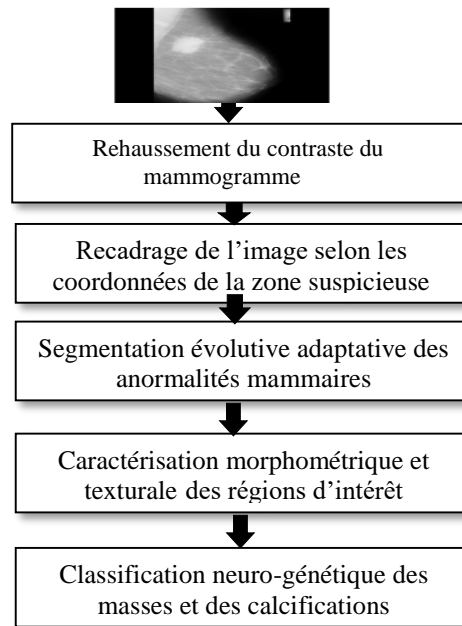


Figure 6.6 : Architecture du système d'aide au diagnostic des anomalies mammographiques

#### 6.4.1 Correction morphologique de mammogrammes

La complexité tissulaire du sein, rend souvent délicat la mise en avant des zones suspectieuses. A cet effet, la superposition des signes cliniques avec les composants glandulaires mammaires, engendre des mammogrammes faiblement contrastés. C'est dans cette optique, que la correction de clichés mammographiques, s'inscrit afin de régenter d'une manière concise, la localisation morphologique des régions d'intérêt pour une aide au diagnostic.

Le premier ensemble use de deux types de prétraitement. Le procédé primaire met en relief les masses (cf. Figures 6.7 et 6.8), suivant un ajustement de la dynamique des intensités de l'image. La seconde correction, rehausse le contraste des calcifications, au gré de résidus de la morphologie mathématique. Une superposition du chapeau haut de forme sur l'image d'origine, est réalisée en adoptant un disque de rayon approprié (cf. Figures 6.9 et 6.10).

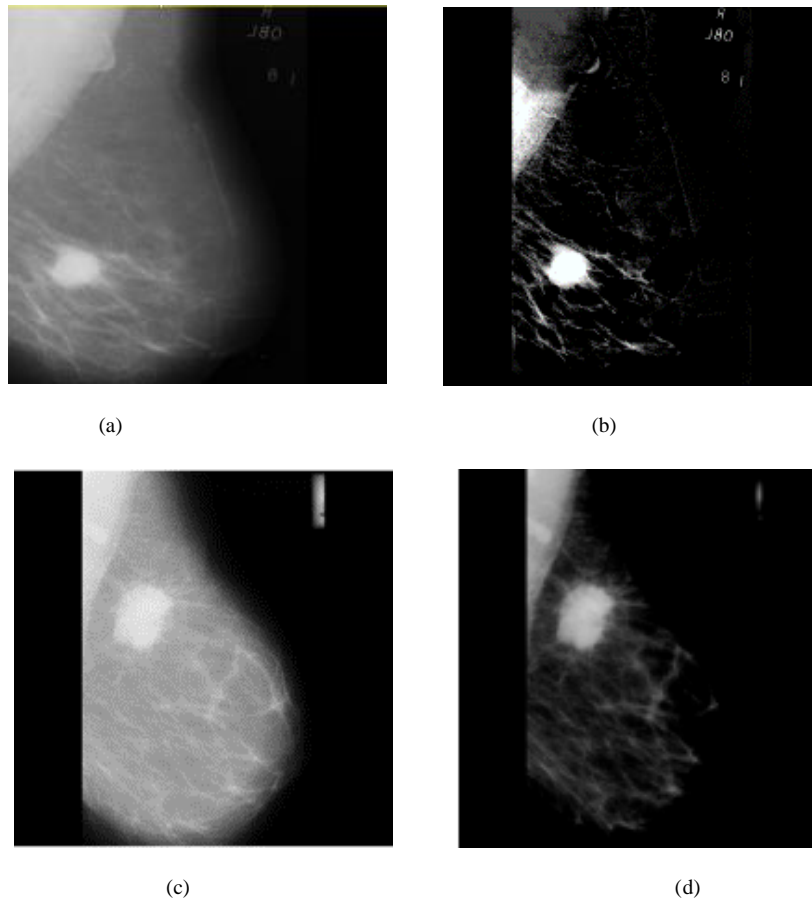


Figure 6.7 : Correction du contraste des anomalies malignes  
(a) Image d'origine mdb028 (b) Mammogramme mdb028 rehaussé  
(c) Mammogramme mdb184 (d) Image mdb184 corrigée

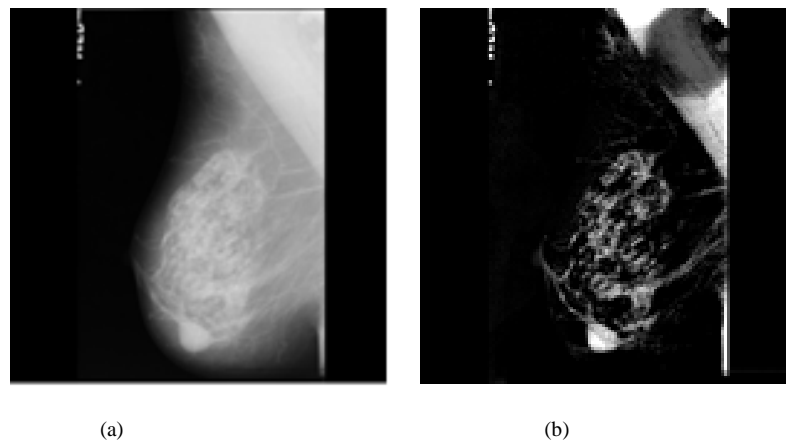


Figure 6.8 : Correction du contraste des anomalies bénignes  
(a) Mammogramme d'origine mdb021 (b) Image améliorée

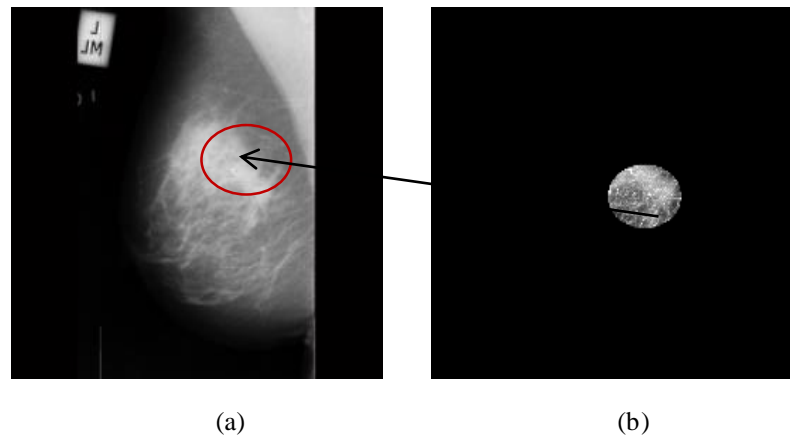


Figure 6.9 : Correction du contraste des calcifications malignes  
(a) Mammogramme original mdb209 (b) Contraste amélioré de la zone suspectieuse  
(Rayon du disque= 5 pixels)

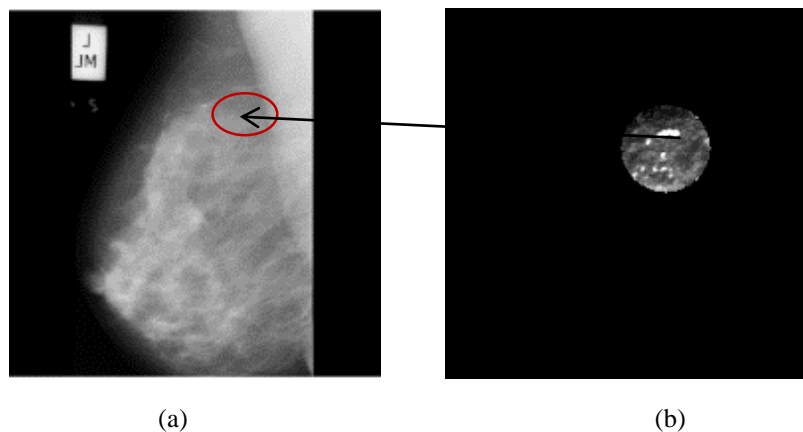


Figure 6.10 : Correction du contraste des calcifications bénignes  
(a) Mammogramme original mdb219 (b) Contraste amélioré de la zone suspectieuse  
(Rayon du disque= 5 pixels)

#### 6.4.2 Extraction des signes cliniques du sein

L'extraction des structures mammaires, est une phase délicate, en raison de la nature complexe des tissus et, leur variabilité. A cet effet, plusieurs travaux y ont été consacrés à partir desquels, nous avons puisé notre méthodologie, pour se focaliser sur les modèles déformables, suivant l'étude de Chan et Vese ainsi que de Lankton, pour retenir les signes pertinents.

La détection précise des régions d'intérêt, est assujettie à une organisation hiérarchique de trois phases :

- ✓ la correction du mammogramme, afin de permettre, la mise en exergue des structures cliniques ;
- ✓ la délimitation de la zone annotée par le clinicien, pour générer le masque de la segmentation évolutive. Ce masque, correspond au cercle entourant la zone suspectieuse ;
- ✓ l'initialisation du contour aux frontières de la zone annotée et l'extraction des pathologies mammaires, suivant le modèle adaptatif de Chan et Vese (cf. Algorithme 3.1) ou celui de Lankton (cf. Algorithme 3.2).

### 1. Extraction des signes cliniques selon le modèle adaptatif de Chan et Vese

Le processus (cf. Algorithme 3.1) évolue implicitement, depuis la surface initiale  $\phi_0(x, y)$  vers l'objet à extraire, en minimisant la fonctionnelle d'énergie. Celle-ci, se compose de deux termes : une énergie interne, qui permet de mesurer l'adéquation avec la forme et, une énergie externe, qui estime l'adéquation avec l'apparence. La régularisation et le lissage de la courbure sont réalisés au cours du traitement, via les attributs  $\lambda_1$  et  $\lambda_2$ . La localisation minutieuse des masses (signes  $>1\text{cm}$ ) et des calcifications, dépend intimement du paramètre  $\mu$  ( cf. Figures 6.11, 6.12, 6.13et 6.14).

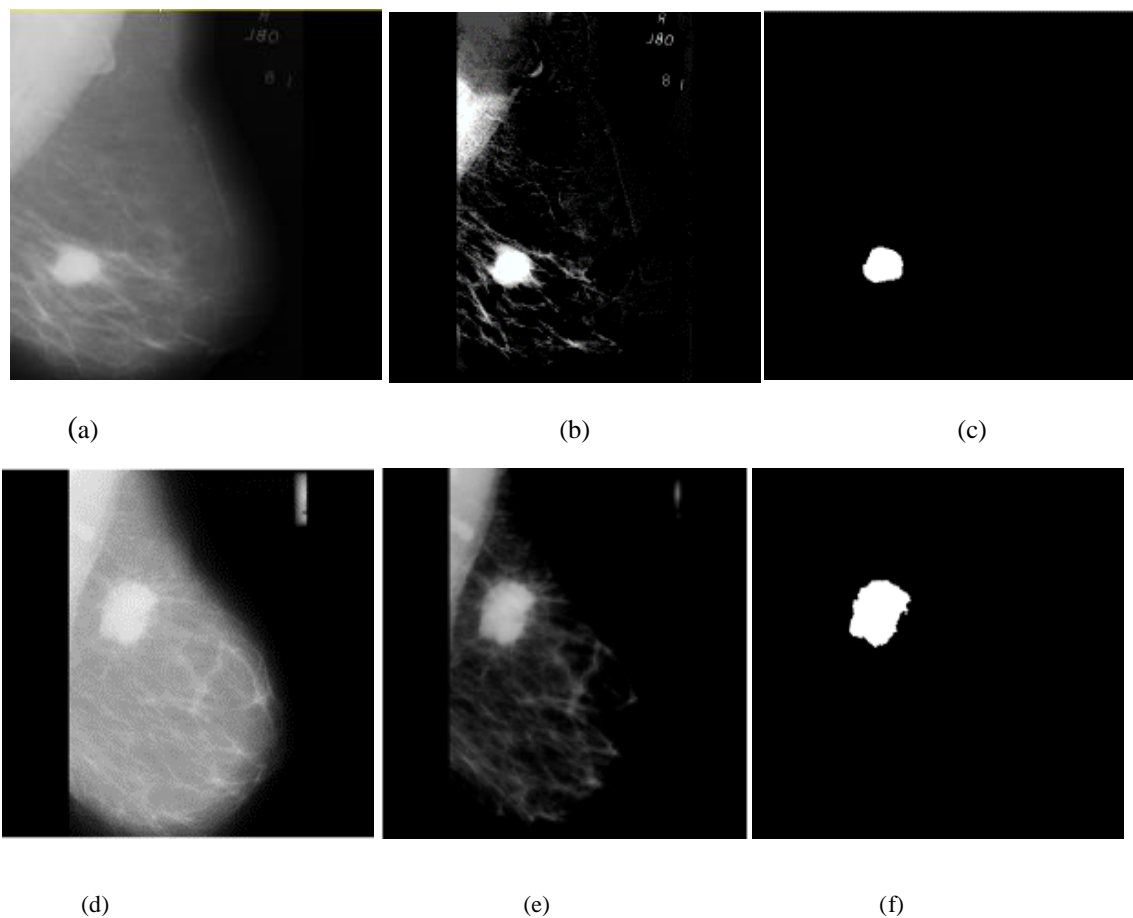


Figure 6.11 : Localisations de masses malignes

- (a) Image d'origine mdb028 (b) Mammogramme mdb028 corrigé (c) Extraction de la masse (nombre d'itérations=270) (d) Image originale mdb184 (e) Image améliorée (f) Extraction de la masse (nombre d'itérations= 500)

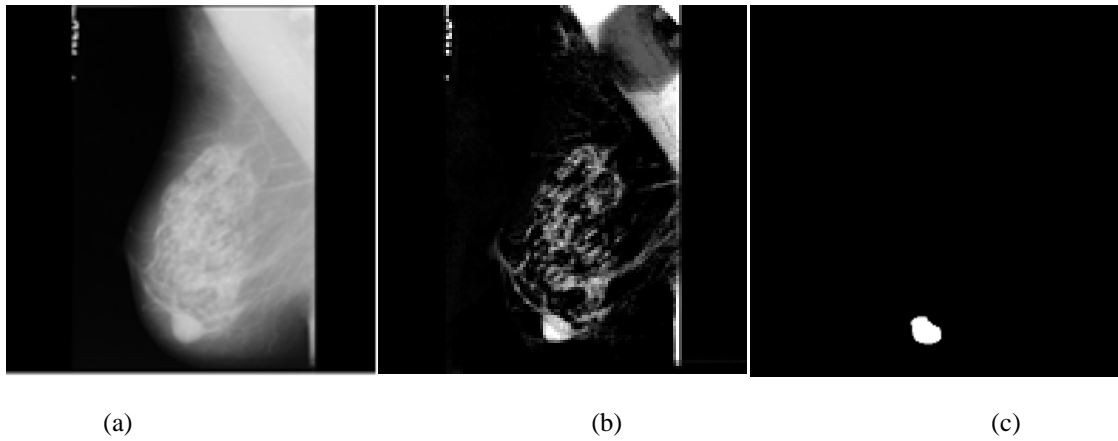


Figure 6.12 : Localisation d'une masse bénigne  
(a) Image d'origine mdb021 (b) Contraste amélioré  
(c) Délimitation de la masse (nombre d'itérations= 200)

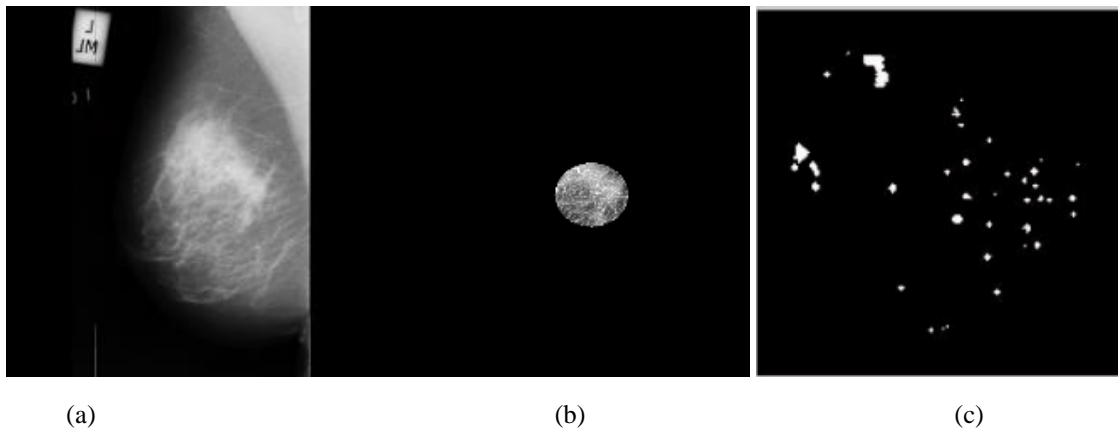


Figure 6.13 : Localisation de calcifications malignes  
(a) Mammogramme original mdb209 (b) Contraste amélioré (rayon du disque= 5 pixels)  
(c) Extraction des calcifications (nombre d'itérations= 1000)

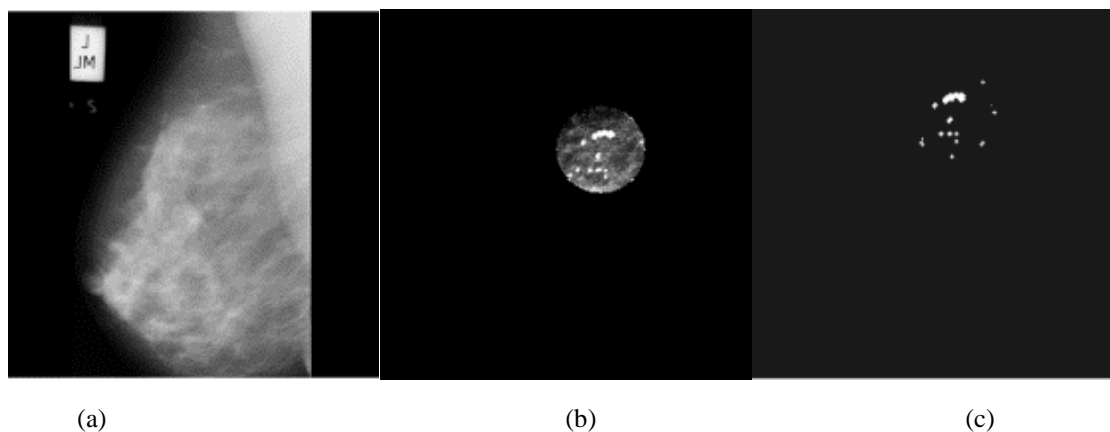


Figure 6.14 : Détection de calcifications bénignes  
(a) Image originale mdb219 (b) Contraste amélioré (rayon du disque= 5 pixels)  
(c) Extraction des calcifications (nombre d'itérations= 800)

## 2. Extraction des signes cliniques selon le modèle de Lankton

La détection des pathologies mammaires est faite dans ce cas, par la segmentation coopérative moyennant, l'agrégation des pixels et le modèle de Lankton (cf. Algorithme 3.2).

Le choix du germe initial dans la croissance de régions, est institué par le centre de gravité de la zone suspectieuse annotée, par le radiologiste. Ceci permet d'éviter la sur-segmentation. La taille de la masse identifiée, par un expert, a une importante influence sur le choix du nombre d'itérations. Les figures suivantes, illustrent respectivement les résultats obtenus.

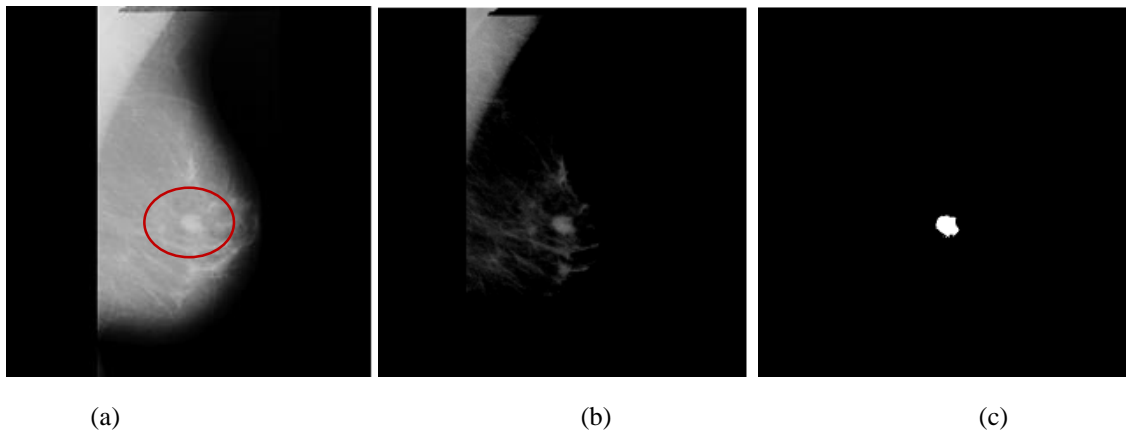


Figure 6.15 : Détection d'une masse bénigne  
(a) Image originale mdb010 (b) Contraste amélioré  
(c) Extraction de la masse (nombre d'itérations= 70)

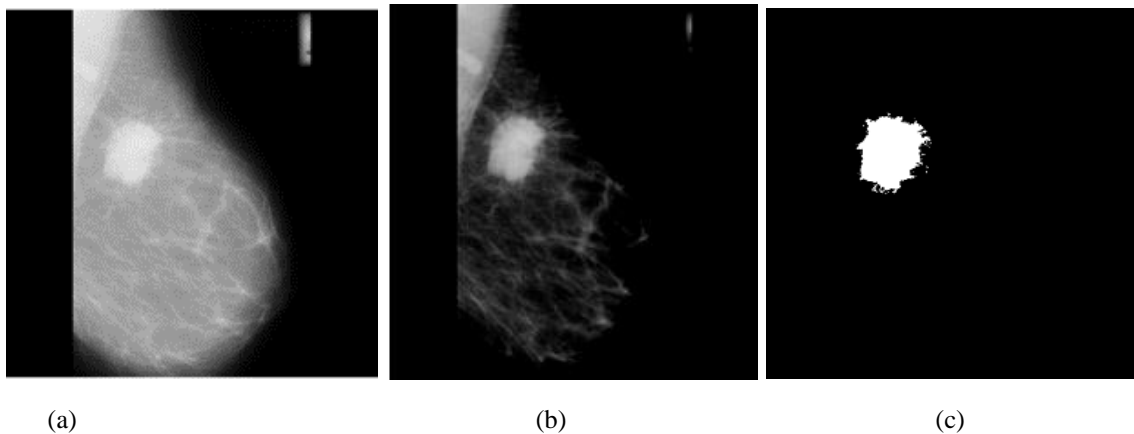


Figure 6.16 : Détection d'une masse maligne  
(a) Image originale mdb184 (b) Contraste amélioré  
(c) Extraction de la lésion (nombre d'itérations= 250)

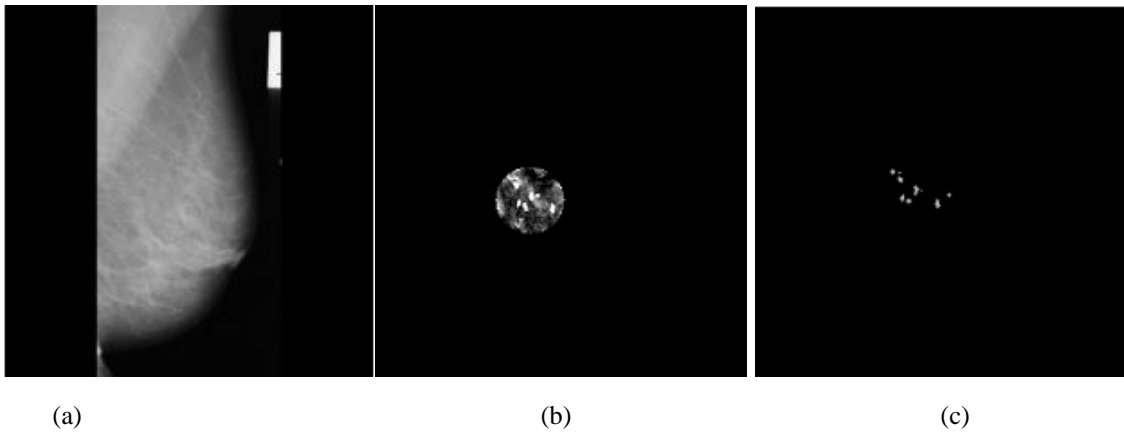


Figure 6.17 : Détection de calcifications bénignes dans le mammogramme mdb252  
(a) Image originale mdb252 (b) Contraste amélioré  
(c) Extraction des calcifications (nombre d'itérations= 600)

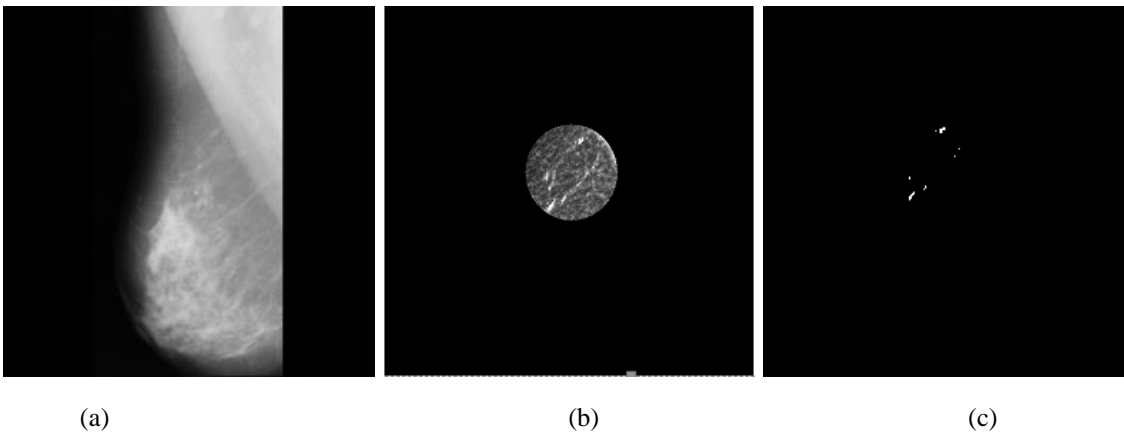


Figure 6.18 : Détection de calcifications malignes  
(a) Image originale mdb249 (b) Contraste amélioré  
(c) Extraction des calcifications (nombre d'itérations= 900)

La segmentation en croissance de régions est initiée, par le germe défini par le centre de la zone suspectieuse. Le seuil obtenu expérimentalement, varie entre 10 et 40, selon le type de l'image. Il apparait des cas où, le seuil est fixé en dessous de 10, à cause d'une masse difficilement visible.

Le nombre d'itérations est intimement corrélé à la taille de la surface du néoplasme. Il augmente lorsque l'anormalité est prépondérante et, vice-versa.

### 6.4.3 Caractérisation des signes cliniques du sein

Les connaissances texturales et morphométriques mammaires, permettent d'orienter la classification des entités sénologiques selon, leur aspect bénin versus malin. Cette discrimination supervisée des masses et des calcifications, s'effectue dans l'ultime étape du système d'aide à la décision, au gré du modèle neuro-génétique.

Les mammogrammes de la base MIAS considérés, sont répartis suivant quatre classes. Les néoplasmes traités, répondent aux masses et aux calcifications bénignes versus malignes.

La signature étiologique des signes pathologiques explorés (cf. Tableaux 6.1, 6.2, 6.3, 6.4 et 6.5), est décrite par six paramètres texturaux (l'énergie, le contraste, l'entropie, la variance, la corrélation et l'homogénéité) et trois attributs morphométriques (la compacité, les moments invariants de Hu et la surface). L'adjectif primordial de cette caractérisation, est l'extraction d'une base de connaissances mammaires, pour assister le praticien dans sa décision.

Image	Entropie	Energie	Contraste	Homogénéité	Corrélation
Mdb184 (M.mass)	0.3035	0.9514	0.0116	0.9989	0.9936
Mdb021 (B.mass)	0.0599	0.9915	0.0039	0.9997	0.9840
Mdb028 (M.mass)	0.1057	0.9843	0.0050	0.9997	0.9905
Mdb141 (M. mass)	0.0127	0.9984	0.0012	0.9999	0.9679
Mdb001 (B. mass)	0.6020	0.8957	0.0218	0.9983	0.9931
Mdb002 (B.mass)	0.1475	0.9774	0.0086	0.9996	0.9892
Mdb010 (B. mass)	0.0414	0.9943	0.0031	0.9983	0.9806
Mdb012 (B.mass)	0.0338	0.9953	0.0023	0.9999	0.9799
Mdb058 (M.mass)	0.0283	0.9963	0.0027	0.9998	0.9739
Mdb069 (B.mass)	0.0544	0.9926	0.0022	0.9997	0.9838
Mdb095 (M.mass)	0.0278	0.9961	0.0034	0.9999	0.9747
Mdb092 (M.mass)	0.0683	0.9899	0.0040	0.9999	0.9839
Mdb111 (M. mass)	0.3179	0.9476	0.0117	0.9992	0.9935
Mdb115 (M.mass)	0.3863	0.9397	0.0112	0.9990	0.9941
Mdb130	0.0117	0.9984	0.0017	0.9999	0.9573



(M.mass)					
Mdb132	0.0313	0.9958	0.0023	0.9999	0.9798
(B.mass)					
mdb121	0.1123	0.9830	0.0093	0.9996	0.9843
(B. mass)					
Mdb150	0.1012	0.9855	0.0051	0.9995	0.9862
(B. mass)					
Mdb181	0.0558	0.9923	0.0044	0.9998	0.9786
(M.mass)					
Mdb142	0.0258	0.9965	0.0024	0.9999	0.9715
(B.mass)					

Tableau 6.1 : Signature texturale de masses dans vingt mammogrammes

Image	Variance	Compacité	Surface en pixels	Moment normalisé de Hu
Mdb184	2.1274e+21	0.7583	25696	0.0278
(M. mass)				
Mdb021	2.3122e+21	0.8857	4368	0.0336
(B. mass)				
Mdb028	2.2784e+21	0.9342	8173	0.0290
(M.mass)				
Mdb141	2.3444e+21	0.9076	811	0.0410
(M. mass)				
Mdb001	1.8800e+21	0.3905	56479	0.0349
(B.mass)				
Mdb002	2.2461e+21	0.8833	11828	0.0294
(B.mass)				
Mdb010	2.3256e+21	0.9514	2881	0.030
(B.mass)				
Mdb012	2.3297e+21	0.9615	2432	0.0362
(B.mass)				
Mdb058	2.3344e+21	0.9658	1906	0.0320
(M.mass)				
Mdb069	2.3172e+21	0.9748	3814	0.0376
(B.mass)				
Mdb095	2.3337e+21	0.8307	1983	0.0296
(M. mass)				

Mdb092	2.3144e+21	0.8129	5173	0.0337
(M.mass)				
Mdb111	2.1092e+21	0.6893	27864	0.0291
(M.mass)				
Mdb115	2.0745e+21	0.8532	32103	0.0292
(M. mass)				
Mdb130	2.3445e+21	0.7783	793	0.0344
(M.mass)				
Mdb132	2.3324e+21	0.8085	2134	0.0367
(B. mass)				
Mdb121	2.2722e+21	0.4685	8836	0.0354
(B.mass)				
Mdb150	2.2842e+21	0.6377	7509	0.0360
(B.mass)				
Mdb181	2.3155e+21	0.4484	3976	0.0379
(M.mass)				
Mdb142	2.3357e+21	0.7419	1763	0.0337
(B.mass)				

Tableau 6.2 : Variance et Signature morphométrique des masses dans vingt mammogrammes

Image	Entropie	Energie	Contraste	Homogénéité	Corrélation
Mdb209	0.0031	0.9995	0.0059	0.9998	0.4625
(M.Cals)					
Mdb219	0.0032	0.9997	0.0024	0.9999	0.6930
(B.Cals)					
Mdb213	0.0012	0.9998	0.0013	0.9998	0.5808
(M.Cals)					
Mdb239	0.0008	0.9999	0.0020	0.9998	0.3580
(M.Cals)					
Mdb211	0.0002	0.9999	0.0004	0.9998	0.3091
(M. Cals)					
Mdb241	0.0004	0.9999	0.0005	0.9999	0.2513
(M. Cals)					
Mdb218	0.0013	0.9998	0.0017	0.9999	0.5253
(B; Cals)					
Mdb223a	0.0008	0.9999	0.0014	0.9999	0.3482
(B. Cals)					

Mdb227	0.0005	0.9999	0.0009	0.9999	0.4601
(B. Cals)					
Mdb231	0.0007	0.9999	0.0008	0.9999	0.4463
(M. Cals)					
Mdb238	0.0015	0.9998	0.0017	0.9999	0.5855
(M. Cals)					
Mdb252	0.0013	0.9998	0.0018	0.9999	0.5177
(B. Cals)					
Mdb253	0.0010	0.9998	0.0021	0.9999	0.4574
(M. Cals)					
Mdb249b	0.0014	0.9998	0.0019	0.9999	0.4642
(M. Cals)					
Mdb249a	0.0020	0.9997	0.0030	0.9999	0.4970
(M.Cals)					
Mdb248	0.0013	0.9998	0.0013	0.9999	0.7112
(B. Cals)					
Mdb240	0.0016	0.9998	0.0023	0.9999	0.6514
(B. Cals)					
Mdb236	0.0001	0.9999	0.0004	0.9999	0.0005
(B. Cals)					
Mdb222	0.0002	0.9999	0.0003	0.9999	0.3750
(B. Cals)					
Mdb223b	0.0002	0.9999	0.0003	0.9999	0.3356
(B. Cals)					

Tableau 6.3 : Signature texturale des calcifications dans vingt mammogrammes

Image	Variance	Compacité	Surface en pixels)	Moment normalisé de Hu
Mdb209	2.3495e+21	0.0949	167	3.4610
(M.Cals)				
Mdb219	2.3498e+21	0.2129	103	0.4442
(B.Cals)				
Mdb213	2.3511e+21	0.2366	61	0.5138
(M Cals)				
Mdb239	2.3513e+21	0.2296	45	0.0927
(M. Cals)				
Mdb211	2.3518e+21	0.0597	11	0.0097

(M. Cals)				
Mdb241	2.3516e+21	0.3529	21	0.0798
(M. Cals)				
Mdb218	2.3510e+21	0.1525	71	0.0016
(B. Cals)				
Mdb223a	2.3513e+21	0.3762	45	0.1350
(B. Cals)				
Mdb227	2.3516e+21	0.3939	27	0.0191
(B. Cals)				
Mdb231	2.3514e+21	0.3202	40	0.6749
(M. Cals)				
Mdb238	2.3509e+21	0.1734	77	0.0136
(M. Cals)				
Mdb252	2.3510e+21	0.1976	69	0.4001
(B. Cals)				
Mdb253 (M. Cals)	2.3512e+21	0.3771	54	0.1500
Mdb249b (M.Cals)	2.3509e+21	0.2064	74	1.1360
Mdb249a (M.Cals)	2.3504e+21	0.1386	106	0.3292
Mdb248 (B.Cals)	2.3511e+21	0.3974	68	0.0001
Mdb240 (B. Cals)	2.3508e+21	0.1903	91	0.3000
Mdb236 (B. Cals)	2.3519e+21	0.3317	6	0.0122
Mdb222 (B. Cals)	2.3519e+21	0.2555	8	0.0071
Mdb223b (B. Cals)	2.3518e+21	0.0616	13	0.0009

Tableau 6.4 : Variance et Signature morphométrique des calcifications dans vingt mammogrammes

#### 6.4.4 Catégorisation neuro-génétique des masses et des calcifications

Le modèle du classifieur neuro-génétique adopté (cf. Algorithme 5.1), est configuré par trois couches : une couche d'entrée, une couche cachée et, une couche de sortie pour la distinction des anormalités traitées (masses malignes et bénignes, calcifications malignes et bénignes). Le réseau neuronal apprend les similitudes des différents cas, à partir de leurs connaissances dans l'ensemble des données signées.

A la première itération, une comparaison des entrées avec les sorties est réalisée, pour vérifier l'ajustement des biais et des poids. Comme ils ont été initialisés aléatoirement, ils nécessitent d'être réglés. C'est dans ce contexte, que les algorithmes génétiques prennent leur importance, avec l'évaluation des individus ( $W_i, B_i$ ). Le processus de sélection, du croisement et de la mutation, est alors institué pour estimer les facteurs synaptiques du perceptron multicouche. A partir de là, une nouvelle population est créée, et à nouveau évaluée. Le réseau est modélisé suivant les opérations génétiques, jusqu'à atteindre la stabilité de la population (cf. Figure 6.19). A cet effet, lorsque la fonction d'évaluation, devient constante, l'optimisation des connexions synaptiques est alors réalisée, sans avoir à faire une multitude d'entraînements, sans intérêt. Le modèle connexionniste approprié est alors conçu, au gré d'un apprentissage optimum.

L'approche de la classification adoptée, régent la discrimination de deux classes bien distinctes, les masses et les calcifications. Chaque classe, est répartie en deux sous classes, indiquant ainsi, les catégories bénignes et malignes, après ordonnancement des poids et des biais.

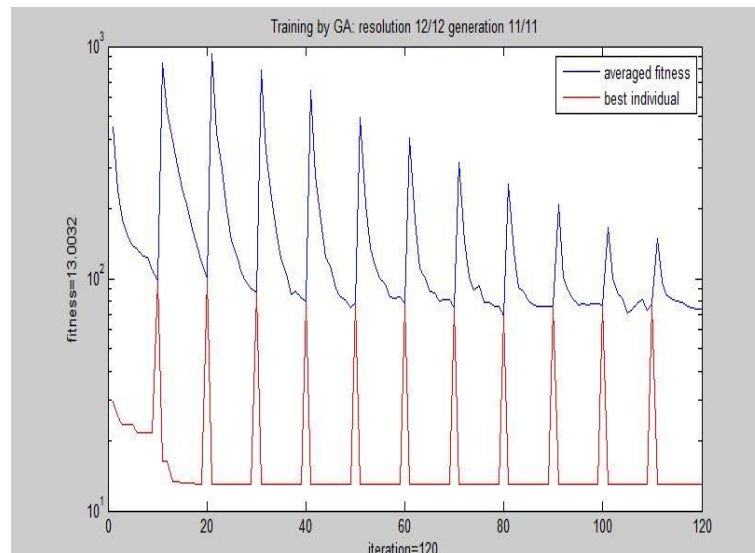


Figure 6.19 : Apprentissage neuro-génétique pour la génération de la meilleure population ( $w_1, w_2, b_1, b_2$ ) dans le cas de masses et de calcifications

## 6.5 Discussion

Un système d'aide à la détection des anomalies mammaires, est élaboré dans cette réflexion, autour d'une fusion de techniques de traitement d'images et d'intelligence artificielle, pour minimiser les erreurs de diagnostic.

En amont de la segmentation des mammogrammes, une phase de prétraitement est accomplie, afin de rehausser le contraste des signes cliniques. A cet égard, les masses sont mises en avant des autres constituants du sein, par l'expansion de la dynamique des intensités dans les clichés radiographiques (cf. Figures 6.7 et 6.8).

Le contraste des calcifications, (cf. Figures 6.9 et 6.10), est amélioré, par la superposition du mammogramme original sur le résultat du chapeau haut de forme.

Cette opération, considère un disque dont le rayon correspond à la taille d'une calcification (5 pixels, dans les cas étudiés, ce qui correspond à une dimension de 2.5 mm). La localisation des masses et des calcifications, est entreprise suivant le modèle de Chan et Vese, adapté aux mammogrammes, grâce à l'initialisation, considérée par un masque circulaire, dont le centre, est celui de la zone suspectieuse annotée par le radiologiste. En effet, le choix judicieux du contour initial, permet l'évolution dynamique du modèle, de l'extérieur vers l'intérieur de la courbe C paramétrée, jusqu'à extraire les lésions tumorales, sans les faux positifs. Dans nos expérimentations, la régularisation des régions, est obtenue au gré des paramètres  $\lambda_1 = 1$  et  $\lambda_2 = 1$ . Le facteur d'échelle  $\mu$  est choisi empiriquement, dépendamment des pathologies à retenir. A cet effet, il équivaut à 0.01 pour la détection des calcifications et à 0.06 pour les masses. Les régions d'intérêt sont délimitées par la minimisation de la fonctionnelle d'énergie, guidée par un nombre d'itérations adéquat. Nous remarquons ainsi, que ce paramètre, varie en fonction du nombre de régions suspectieuses et des catégories des signes cliniques, à retenir. Les figures 6.11 allant à 6.18, illustrant les résultats de cette approche, à l'effigie des mammogrammes de la base MIAS, attestent de ce constat. A cet effet, l'extraction de la masse maligne du cliché mdb184 (cf. Figures 6.11.f et 6.16), dont l'aire de la zone suspectieuse converge vers 163229.76 pixels, est délimitée après 500 itérations. La détection de la masse bénigne de l'image mdb021 (cf. Figure 6.12), dont la surface de la zone clinique révèle 30156.56 pixels, est obtenue après 200 itérations. Un nombre d'itérations considérable de l'ordre de 1000, en raison du nombre de régions à mettre en exergue, ponctue la localisation des calcifications du mammogramme mdb209 (cf. Figure 6.13). Ces résultats montrent la liaison conjointe entre le nombre d'itérations, l'aire de la zone suspectieuse et la sémiologie des signes pathologiques diffusés, dans le tissu mammaire.

Le principal atout, du modèle de Chan et Vese adopté, est la préservation de la topologie des régions. Cette propriété, intervient, dans la précision de la caractérisation étiologique mammaire. Les signes oncologiques retenus, sont soumis à l'analyse quantitative hybride, générant ainsi, leur signature, autour de paramètres texturaux et morphométriques, pour diriger le praticien dans son interprétation. Nous remarquons à cet effet, que la compacité évaluée pour les masses (cf. Figure 6.20), excède la valeur de 0.9 atteignant le 1, lorsque la forme est circulaire ratifiant ainsi, les irrégularités de la pathologie.

Le moment normalisé de Hu, paramètre d'inertie, arbore l'analyse des calcifications (cf. Figure 6.21), suivant l'invariance des transformations affines et, croit en fonction de la complexité de la tumeur.

L'entropie (cf. Figure 6.22), qui exprime le désordre dans une image, est ponctuée par des valeurs significativement considérables, pour les masses diffusées amplement dans le tissu du sein. Ce descripteur évolue proportionnellement, à l'aire des structures oncologiques.

Les attributs déterminés constituent la signature étiologique de chaque signe clinique du sein. Ces connaissances sont intégrées à l'entrée du perceptron multicouche en hybridation avec les algorithmes génétiques, afin de générer la classe appropriée via, un entraînement évolutionnaire, pour l'optimisation de l'espace des données. L'algorithme original élaboré par Ludwig, est un classifieur binaire.

L'un des points majeurs de notre contribution, est la réalisation d'une discrimination neuro-génétique multi-classe, après ordonnancement des facteurs synaptiques. Dans cette optique, cent quinze mammogrammes de la base MIAS sont considérés pour diligenter le processus décisionnel, dans ce prélude à la recherche. Soixante-dix-sept échantillons sont ainsi, assujettis à l'apprentissage et, trente-huit font l'objet du test.

Lors de l'entraînement, le modèle connexionniste optimum est engendré, par neuf neurones dans l'architecture de la couche cachée, pour une meilleure population des attributs synaptiques. La stabilité de la population est atteinte pour 11 générations (cf. Figure 6.19) au gré de 120 itérations. La régulation des meilleurs individus, est déclinée par un facteur de pénalisation de 80, pour une discrimination concise des anormalités. Le système d'aide au diagnostic élaboré, reconnaît les néoplasmes mammaires avec une précision de 99.25%, une sensibilité de 100% et une spécificité de 98%.

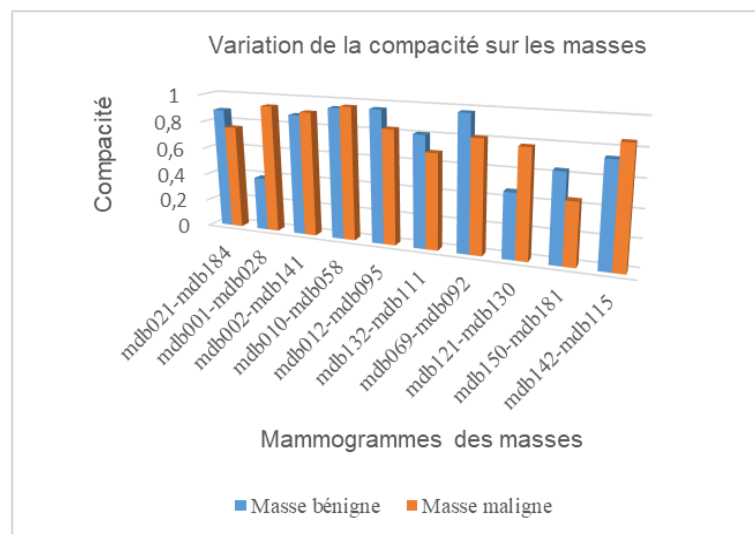


Figure 6.20 : Variation de la compacité en fonction des masses

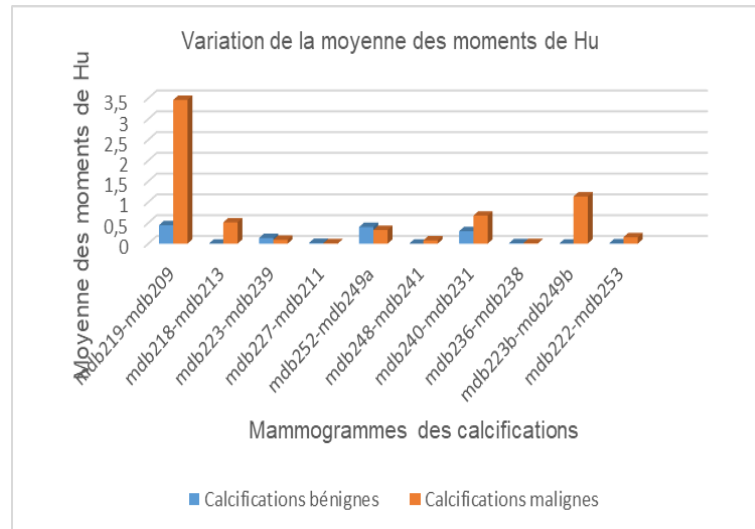


Figure 6.21 : Variation de la moyenne des moments de Hu en fonction des calcifications

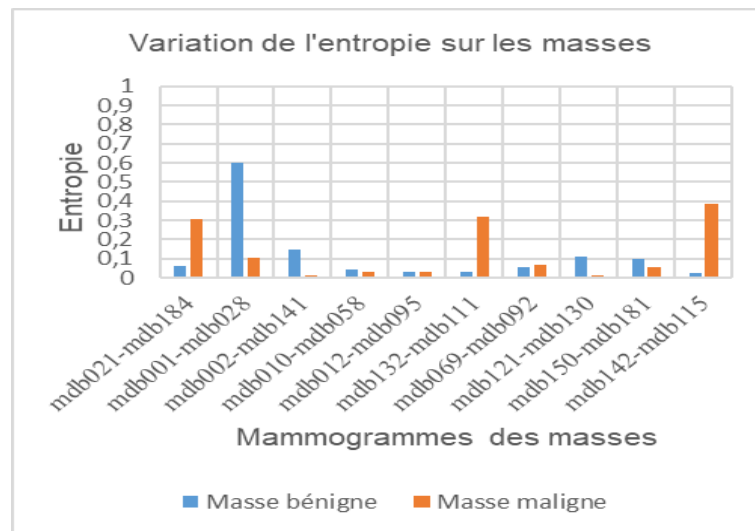


Figure 6.22 : Variation de l'entropie en fonction des masses

## 6.6 Approche statistique de Vapnick pour la détection morphologique de masses et de tissus sains en mammographie numérique

L'objectif de notre seconde contribution (cf. Figure 6.23), est de réaliser un système d'aide à la décision par la fusion de la morphométrie et de la texture, pour la détection de néoplasmes et de tissus sains du sein. Le synopsis méthodologique supervise un processus décisionnel, destiné à révéler, la présence ou l'absence d'une pathologie. A cet effet, le mammogramme est une image complexe résultant de la superposition de structures variables qui doivent être dissociées, pour la distinction concise des zones suspectieuses.



Ce processus est accompli par l'hybridation du regroupement spatial flou et, des modèles déformables. La segmentation dynamique est initialisée par le résultat de la délimitation spatiale, guidée par l'aire clinique annotée. Il n'existe pas de modèle unique du sein ; chaque cas requiert une analyse rigoureuse. Dans ce contexte, lorsqu'une lésion est repérée, l'étape qui lui succède est de la décrire selon des critères analytiques, caractérisant sa forme et sa texture. Cette description permet l'estimation de la complexité ou, la régularité des mammogrammes. Ces connaissances sont incorporées dans le classifieur de Vapnik multi-classes, pour la distinction des anormalités et des normalités mammaires.

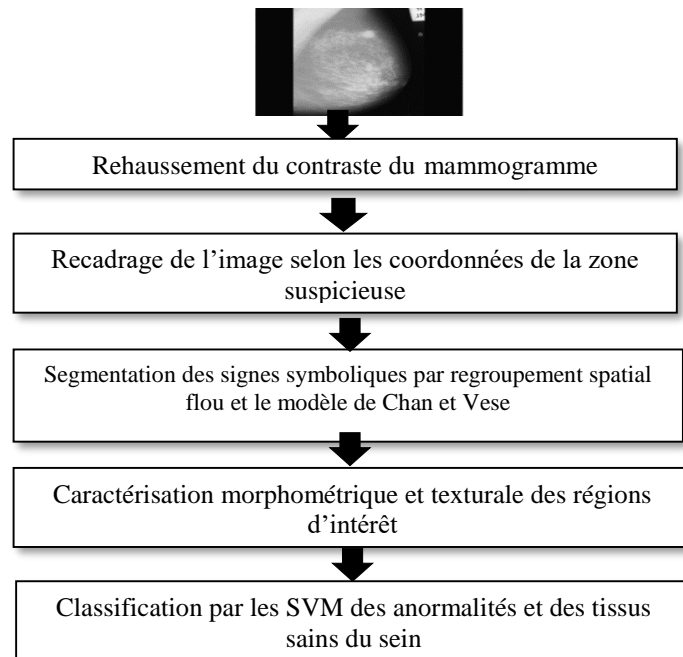


Figure 6.23 : Synoptique de détection d'anormalités et de tissus sains du sein

### 6.6.1 Amélioration du contraste des mammogrammes

Le prétraitement dans le cas de cette approche, consiste en premier lieu à appliquer sur l'image originale une opération du chapeau haut de forme, pour faire apparaître les structures claires de l'image. Une superposition de ce résultat sur l'image originale, améliore sa qualité visuelle.

L'opération du chapeau haut de forme, est basée sur un élément structurant autour d'un disque, avec une taille qui prend en compte le rayon de la zone annotée. Il varie entre 60 et 120 dans le cas de mammogrammes sains selon le type du tissu mammaire. Le prétraitement de cette réflexion, se focalise sur les transformations morphologiques, pour accentuer le contraste étologique des mammogrammes. Le processus en question met en relief les régions d'intérêt (cf. Figures 6.24, 6.25 et 6.26), par la superposition du résultat du chapeau haut de forme, sur le cliché d'origine en adoptant un disque dont le rayon, répond à celui de la zone suspectieuse.

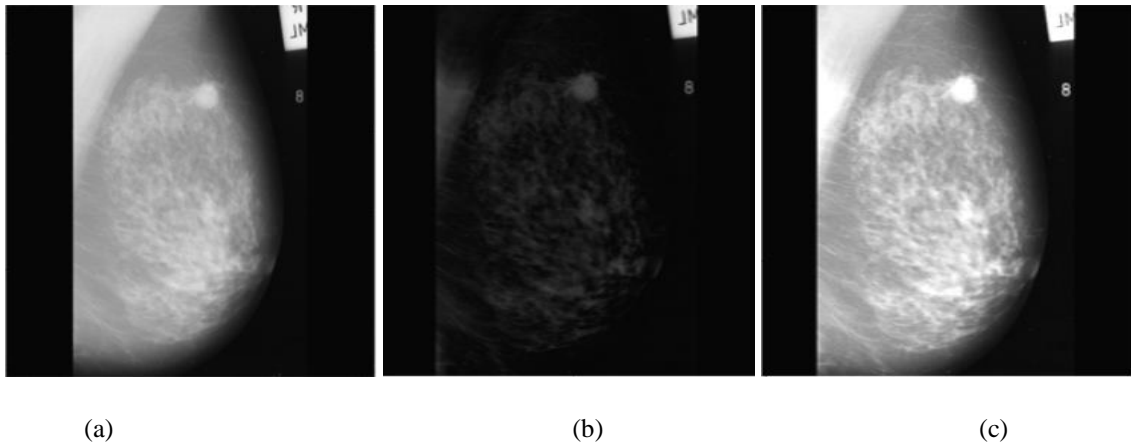


Figure 6.24 : Correction du contraste d'un mammogramme malin  
(Élément structurant utilisé : disque d'un rayon de 37 pixels correspondant au rayon de la zone suspectieuse)  
(a) Image d'origine mdb202 (b) Résultat du CHF (c) Image (a) superposée au résultat du CHF :  
contraste rehaussé

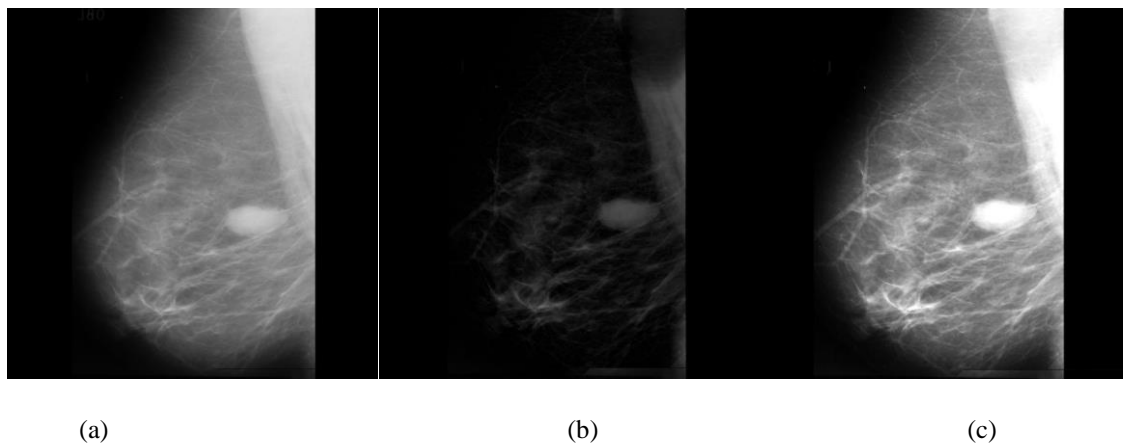


Figure 6.25 : Correction du contraste d'une masse bénigne  
(Élément structurant utilisé: disque d'un rayon de 79 pixels)  
(a) Image d'origine mdb025 (b) Résultat du CHF  
(c) Image (b) superposée au résultat du CHF : contraste rehaussé

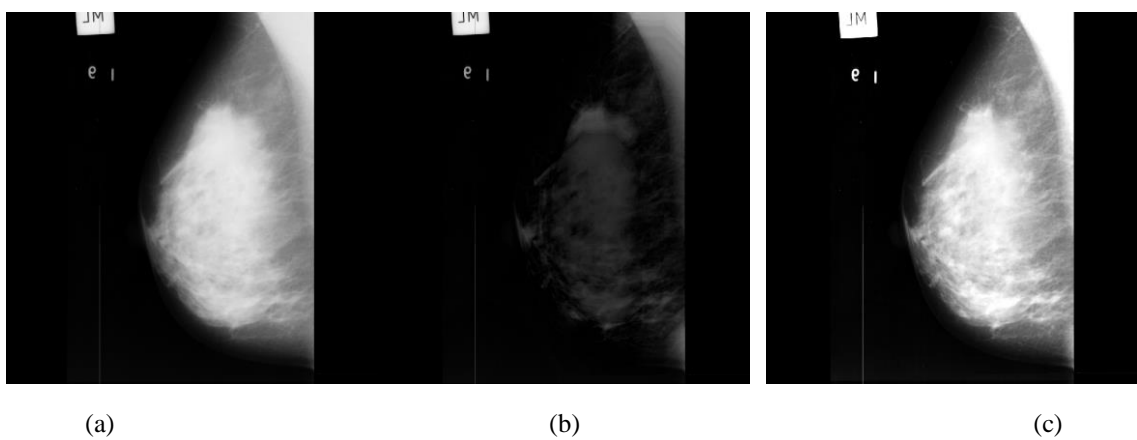


Figure 6.26 : Correction du contraste d'un mammogramme normal  
(Élément structurant utilisé: disque d'un rayon de 100 pixels )  
(a) Image d'origine mdb003 (b) Résultat du CHF  
(c) Image (b) superposée au résultat du CHF : contraste rehaussé

### 6.6.2 Extraction des signes cliniques du sein

Le partitionnement des constituants tissulaires du sein, est institué par le regroupement spatial flou, pour générer la fonction initiale du modèle déformable. Le raffinement de la segmentation, est obtenu par le modèle de Chan et Vese. L'approche développée, répond aux étapes de l'algorithme 3.3.

L'inconvénient majeur dans l'algorithme des C-moyennes floues, lors de la segmentation des images, est la négligence de l'information spatiale. A cet effet, la distinction des différents composants mammaires, particulièrement lorsque les pixels du voisinage de la zone d'intérêt sont grandement corrélés, s'effectue par l'intégration de l'information spatiale dans la fonction floue d'appartenance  $\mu_{ik}$ . La pathologie est ainsi, totalement dissociée des composants sains du sein. Les figures 6.27, 6.28 et 6.29 attestent de la coopération du regroupement flou et du modèle déformable de Chan et Vese, pour délimiter les structures mammaires, au gré des tissus denses et gras.

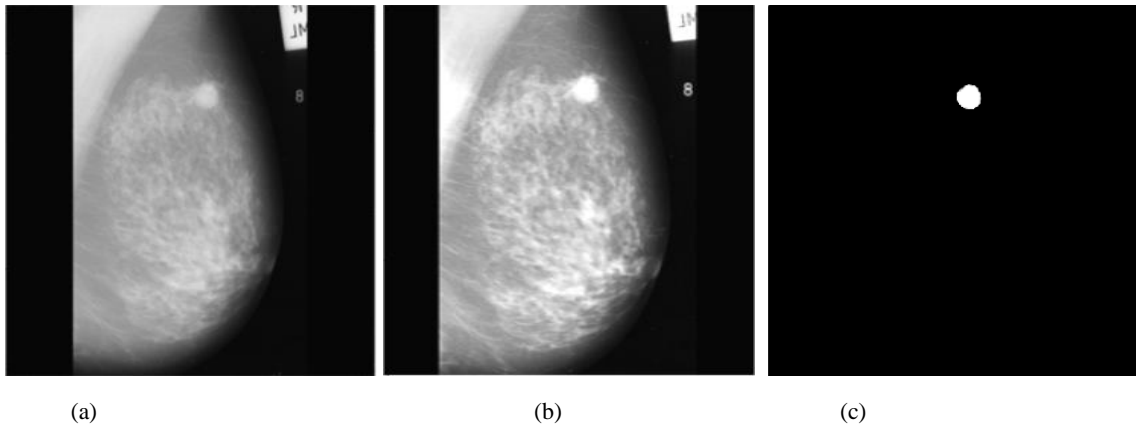


Figure 6.27 : Extraction d'une anomalie maligne dans un mammogramme à tissu dense  
(a) Image d'origine mdb202 (b) Mammogramme amélioré (c) anomalie extraite  
(Nombre de clusters : 3, nombre d'itérations 80)

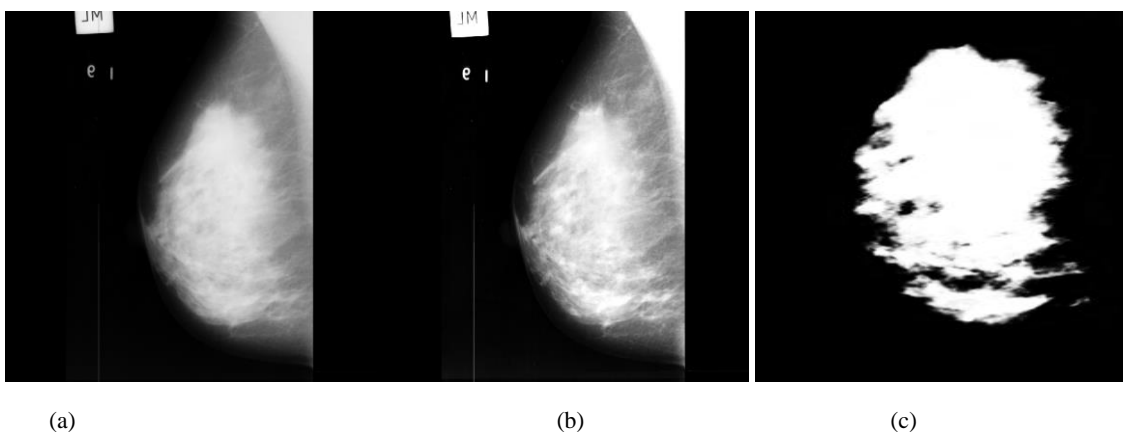


Figure 6.28 : Extraction du tissu sain dans une structure mammographique dense  
(a) Image d'origine mdb003 (b) Localisation du tissu sain  
(Nombre de clusters =4 et nombre d'itérations =100)

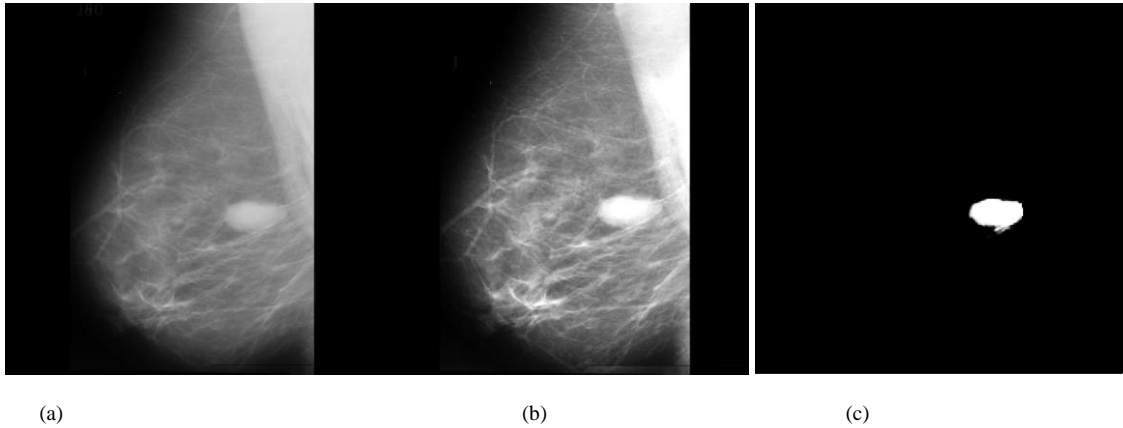


Figure 6.29 : Délimitation d'une anomalie bénigne dans un mammogramme à tissu gras  
(a) Image d'origine mdb025 (b) Rehaussement du tissu mammaire (c) Localisation du tissu sain  
(Nombre de clusters =4 et nombre d'itérations =120)

### 6.6.3 Description morphométrique et de la granularité des régions d'intérêt

La signature des anomalies et des tissus sains, est établie suivant l'évaluation, de descripteurs texturaux (l'entropie, le contraste et la corrélation) et morphométriques (la compacité, l'excentricité, la rectangularité et la convexité), afin d'apprécier la rugosité du tissu mammaire ou son uniformité.

Au cours des expérimentations menées, nous remarquons que dans le cas d'images saines, les paramètres descriptifs, sont du même ordre, alors que pour les cas pathologiques, ils varient suivant le type de tissu mammaire et la lésion. Les graphes suivants, transcrivent ce constat (cf. Figures 6.30 et 6.31).

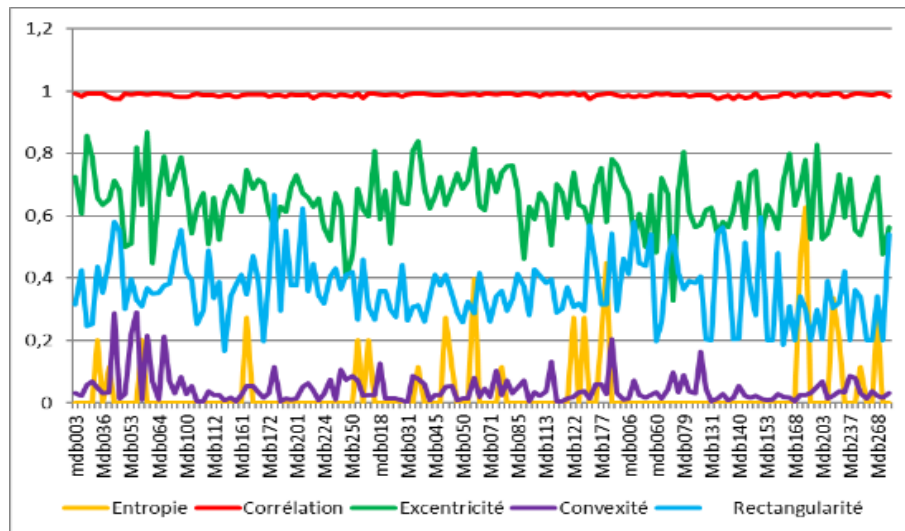


Figure 6.30 : Variation des descripteurs en fonction des mammogrammes sains

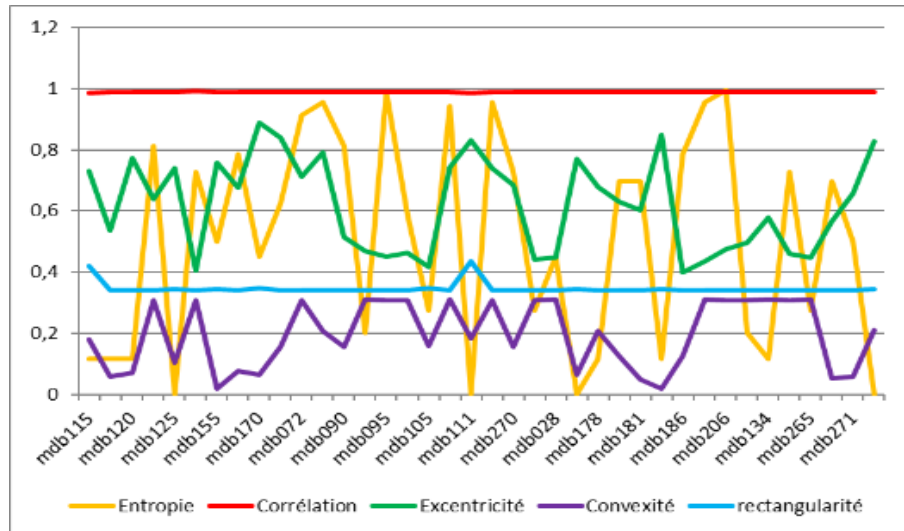


Figure 6.31 : Variation des descripteurs en fonction des mammogrammes pathologiques

#### 6.6.4 Discrimination des tissus sains et des anomalies mammaires par les SVM

Dans cette approche, nous avons sélectionné les 2/3 de la base MIAS (214 mammogrammes) pour l'apprentissage et 1/3 pour la phase d'exploitation (autour de 108 clichés). L'apprentissage, se fait via la stratégie "un contre un", suivant un noyau gaussien, pour la distinction des classes en maximisant la distance les séparant.

L'apprentissage vise à traduire les connaissances issues des entités du sein, fournies à l'entrée des SVM, en des labels sains, bénins et malins. Ces attributs fournis à l'entrée du classifieur sous forme de vecteurs, traduisent les signaux diffusés par le sein. La sortie du modèle engendre trois catégories, où on peut qualifier la tumeur de bénigne si c'est -1, de maligne si c'est 1, de tissu sain si c'est 0. Le tableau 6.5, présente les paramètres et les performances du système, engendrés pour chaque catégorie.

Dans le cadre applicatif de cette labellisation, l'estimation des probabilités a postériori de chaque classe est prise en compte, pour catégoriser une nouvelle observation.

Le système obtenu, reconnaît les anomalies et les tissus réguliers, avec une précision de 99% au gré d'un taux d'erreur de 0.0119.

Classes	Matrice des poids	Probabilité a postériori	Précision
Anormalité bénigne (-1)	0.702335329341	0.139700598802395	100%
Anormalité maligne (1)	0.850299401197	0.197664670658683	98%
Tissu sain (0)	0.317365269461	0.682634730538922	99%

Tableau 6.5 : Paramètres de la classification par les SVM et performances du CAD proposé

### 6.6.5 Interprétation

La seconde approche aborde l'analyse des mammogrammes, pour discerner les anomalies et les tissus sains. Les signes physiologiques pertinents, ont été mis en exergue par les opérateurs de la morphologie mathématique, où l'élément structurant adopté, est un disque dont le rayon équivaut à la zone suspecte. Un regroupement spatial flou, partitionne le mammogramme en décorrélant les pixels dissemblables de la zone annotée, pour ne prendre en compte que la classe d'intérêt. Le modèle de Chan et Vese, affine la délimitation de ces régions, en conservant la courbure des contours. Le choix du nombre d'itérations a été pris en compte judicieusement, puisque lors des tests, nous avons remarqué que la taille de la zone symbolique, influe sur le nombre d'itérations considéré. Ce nombre croît proportionnellement aux anomalies. L'adjectif majeur de cette segmentation est de décrire les tissus mammaires, au gré de paramètres morphométriques et texturaux. Cette description exprime la complexité et, la régularité (cas sains), ou l'irrégularité (cas de distorsions) des clichés radiologiques du sein, afin d'établir un dictionnaire de connaissances étiologiques. Cette signature intégrée dans un classifieur SVM, autour d'une maximisation de la marge par un noyau gaussien radial et, un facteur de pénalité de 70, a engendré des résultats encourageants.

## 6.7 Etude comparative avec les systèmes CAD

L'efficacité de l'approche développée dans le cadre de cette thèse, est attestée par une étude comparative instituée, par la synthèse de vingt systèmes d'aide au diagnostic mammographique, accomplis par les chercheurs soucieux de la santé, durant la période, de 2002 à 2018 (cf. Tableau 6.6).

L'accent est particulièrement mis sur les travaux de Duraiamy et Emperumal [102], où le CAD exploré au gré des réseaux connexionnistes convolutifs, a engendré les mêmes performances que notre système décisionnel.

La réflexion visée lors de l'élaboration de la problématique référée, par les limites de l'examen mammographique (difficulté de délimiter les structures oncologiques à cause de la complexité anatomique), a donc donné naissance à des résultats encourageants.

Classification	Auteurs	Base usée	Précision	Performances
SVM	(Mohamed et al 2014) [46]	MIAS	83%	Facteurs non évalués
MLP	(Arafi et al, 2013) [152]	MIAS	95.49%	Facteurs non évalués
CCNN	(Dheeba et al 2012) [73]	MIAS	89.25%	St=85.2% Sp est non évaluée
QKCLDA	(Li et al, 2013) [74]	MIAS	96.98%	Facteurs non évalués
CNN	(Valenzuela et al 2005) [98]	MIAS	92%	Facteurs non évalués
CBN	(Lavanya et al 2012) [150]	MIAS	91.67%	St=85% Sp=100%
RBNN	Christoyianni et al 2002) [71]	MIAS	88.23%	Facteurs non évalués

<b>ANN</b>	Mohammed et al 2010) [151]	MIAS	Non évaluée	St=90.91% Sp=83.87%
<b>MLP</b>	(Rizzi et al, 2012) [75]	MIAS	Non évaluée	St=98% Sp est non évaluée
<b>ANN</b>	(Beura et al, 2015) [77]	MIAS/DDS	98%	AUC=94.2%
<b>FCRN</b>	(Saraswathi et Srinivasan, 2014) [76]	MIAS	98%	St=97%
<b>SVM</b>	(Prathibha et Sadasivam, 2012) [81]	MIAS	96.25%	Facteurs non évalués
<b>ADC</b>	(Jen et al 2015) [126]	MIAS	Non évaluée	St=88% Sp est non évaluée
<b>SVM</b>	(Zyout et al, 2015) [82]		97%	AUC=96%
<b>SVM</b>	(Kashyap et al, 2018) [49, 50]	MIAS		AUC=95.33%, St=94.63% et Sp=94.72%
<b>DL-CNN</b>	(Duraismy, et Emperumal, 2017) [102]	MIAS	99%	AUC=98%, St=98.75% et Sp=100%
<b>SVM</b>	(Ancy et al, 2018) [87]	MIAS	98.11%	St=98.71% Sp=96.43%
<b>Arbres de décision</b>	(Sadad et al, 2018) [85]	MIAS	98.2%	St=97% Sp=100%
<b>ANN</b>	(Boulehmi et al, 2016) [47]	MIAS	99%	Facteurs non évalués
<b>RBNN</b>	(Casti et al, 2017) [83]	MIAS	99%	St=100%
<b>CNNA</b>	Notre proposition, 2018 [127]		99.25%	St=100% Sp=98%

Tableau 6.6 : Performances de vingt systèmes CAD mammographiques

## 6.8 Conclusion

Détecter une anomalie en mammographie, repose tout d'abord sur la correction du contraste de l'image. Cette étape est primordiale, pour mettre en avant les structures pertinentes, relativement aux autres constituants du sein. L'extraction des connaissances étiologiques, à travers la phase d'analyse guidée par l'expertise, est soulignée par des modèles mathématiques. Le choix de ces modèles, est ponctué par la prise en compte de paramètres spécifiques à la localisation des masses ou des calcifications (facteurs de régularisation dans les modèles déformables). Cette démarche hiérarchisée, lève le voile sur les incohérences cliniques et découle, sur des résultats concis. Une anomalie est ainsi accentuée, localisée précisément (en tenant compte de la complexité topologique) et classée selon son degré de suspicion (obtenu par une supputation morphométrique).

Le système d'aide au diagnostic mis au point, génère des performances exploitables dans le monde radiologique, puisqu'il permet la discrimination des anomalies et des tissus sains, avec un taux de reconnaissance efficace.

Notre approche ratifie ainsi, la mise en œuvre des modèles usés pour l'analyse des signes sénologiques et sains mammaires, suivant un choix judicieux des attributs de caractérisation.



## Conclusion générale et perspectives

---

*Mieux vaut se mettre sérieusement à quelque chose de médiocre que de rêver éternellement au parfait » 'P. Palau'*

Puiser dans le cœur de la technicité de l'imagerie, pour assouvir notre objectif de développer un système d'aide au diagnostic radiologiste ; telles sont les préoccupations de nos travaux ; établis autour de la détection de signes oncologiques mammaires, dont l'interprétation s'est alliée, à celle d'un spécialiste de la santé.

Les travaux réalisés lors de ce prélude, ont donné naissance à un système intelligent, pour l'extraction de structures oncologiques et sains du sein. Son application dans l'imagerie mammographique, permet de réduire les erreurs de diagnostic clinique et les confusions dues à la complexité tissulaire du sein, auxquelles, l'examen radiologique est confronté à chaque fois.

En effet, la mammographie est un domaine en évolution croissante, depuis plusieurs décennies, avec l'insertion du champ numérique. Dans cette optique, les systèmes numériques ont été introduits, afin d'améliorer les performances cliniques, de réduire la durée de la procédure médicale, de régenter la précision du geste et de se soucier du confort de la patiente. Cependant, certaines limitations de ce type d'imagerie, sont toujours d'actualité, comme la superposition de tissus pouvant potentiellement cacher des lésions. En effet, certaines études montrent, que 76% des lésions qui sont manquées par les radiologues, se trouvent dans des seins denses. Ainsi, la superposition des tissus serait la principale cause de la réduction de visibilité, des signes radiologiques. Dans ce contexte, la détection automatique de cancers mammaires, permet de fournir au radiologue des outils très utiles, dans sa tâche de dépistage, ou pour une analyse concise de tumeurs suspectes.

Le synopsis méthodologique semi-automatique d'aide au diagnostic élaboré, répond à cette anxiété, en se focalisant sur l'extraction des connaissances sémiologiques mammaires ; selon les critères morphométriques et texturales adaptés à la malignité, la bénignité et la régularité des mammogrammes.

L'approche développée est robuste vis-à-vis de la principale difficulté rencontrée par les contours actifs, qui est le choix du premier niveau. Dans cette réflexion, l'initialisation est diligentée, par l'équation du cercle qui entoure la région annotée, par le radiologiste. L'opération est alors réduite à cliquer sur le centre du cercle, dont le rayon est conforme à la zone suspecte. Par conséquent, cette forme circulaire est automatiquement créée pour orienter le processus de la segmentation dynamique, dans le but d'une extraction précise des pathologies mammaires. Les cliniciens ont validé les résultats obtenus, par le modèle déformable guidé, qui gère automatiquement les changements topologiques.



Il a été démontré que l'analyse hybride entre les attributs morphométriques et texturaux, est un moyen pratique de caractériser les néoplasmes du sein.

Cette description fournit des mesures quantitatives, de la rugosité des anomalies et de leur complexité. Dans ce contexte, il a été constaté qu'une masse spinale, présente une entropie plus élevée que celle d'une masse lisse.

L'institution décisionnelle confectionnée suivant les réseaux connexionnistes d'une part et les SVM d'autre part, a atteint un taux de reconnaissance efficace. Elle permet ainsi de assister le radiologue, dans l'interprétation des clichés mammographiques ; réduisant ainsi, la variabilité des diagnostics, assurant la prévention et le contrôle thérapeutique des signes sénologiques.

Le système proposé, a un double rôle. Il aide le radiologue, dans la détection précoce du cancer du sein (CADe) et fournit une aide au diagnostic (CADx). Il peut constituer ainsi, un outil essentiel pour la télémédecine, dans les régions enclavées telles que, le Sahara maghrébin.

Le processus de radiodiagnostic conçu peut subtilement s'ouvrir, au champ des perspectives, de l'analyse d'images mammaires, en corrélant les données mammographiques, aux images cytologiques du sein, pour affiner davantage la description étiologique. La lecture technique des clichés mammographiques, promulguée par la présente étude, s'intègre alors dans une large heuristique couplée à un système expert, aux données de l'anamnèse et à d'autres attributs, afin d'en déduire une conduite diagnostique.

La construction d'une ontologie pour la détection des néoplasmes du sein, réduirait le fossé sémantique. Cette conception viserait à fournir un examen complémentaire, à la médecine, pour le suivi et le contrôle thérapeutique. Le radiologue ne sera donc plus confronté, à recourir d'emblée à la chirurgie, dès le moindre doute.

## Bibliographie

---

*"Il est des livres dont il faut seulement goûter, d'autres qu'il faut dévorer, d'autres, enfin, mais en petit nombre, qu'il faut pour ainsi dire mâcher et digérer". F. Bacon.*

La réflexion technico-clinique adoptée dans le cadre de cette thèse, pour la réalisation d'un système d'aide au diagnostic de mammogrammes, est ponctuée par le déferlement bibliographique, arboré par ordre de parution dans le texte selon les manuels, mémoires, thèses, rapports, articles de journaux et de conférences référencés ci-dessous :

- [1] Organisation mondiale de la santé (OMS). Dernières données mondiales sur le cancer, [https://www.iarc.fr/fr/media-centre/pr/2018/pdfs/pr263\\_F.pdf](https://www.iarc.fr/fr/media-centre/pr/2018/pdfs/pr263_F.pdf), 2018, date de consultation, septembre 2018.
- [2] Organisation mondiale de la santé (OMS). Dernières statistiques mondiales sur le cancer, [http://www.iarc.fr/fr/media-centre/pr/2013/pdfs/pr223\\_F.pdf](http://www.iarc.fr/fr/media-centre/pr/2013/pdfs/pr223_F.pdf), 2013, date de consultation, septembre 2014.
- [3] L. Abid. Épidémiologie des cancers en Algérie : problématique des registres des cancers. *Journal Africain du cancer*, Springer, Vol.1, Num.2, pages 98-103, 2009.
- [4] M.H. Cherif, D. Serraino, A. Mahnane, S. Laouamri, Z. Zaidi, H. Boukharouba, D. Cherka, M. Rakeb, L. Kara, A. Ayat, S. Birri, S. Virdone, P. De Paoli, E. Bidol. Time trends of cancer incidence in Setif, Algeria, 1986–2010: an observational study. *BMC Cancer*, Vol. 637, pages 1-8, <http://www.biomedcentral.com/1471-2407/14/637>, 2014, date de consultation, Octobre 2016.
- [5] M.H. Cherif, D. Serraino, A. Mahnane, S. Laouamri, Z. Zaidi, H. Boukharouba, D. Cherka, M. Rakeb, L. Kara, A. Ayat, S. Birri, S. Virdone, P. De Paoli, E. Bidol. Cancer estimation of incidence and survival in Algeria 2014. *Journal of Cancer Research & Therapy*, Vol. 3, Num. 9, pages 100-104, 2015.
- [6] R.M. Rangayyan, Recent advances in breast imaging, mammography, and computer-aided diagnosis of breast cancer, Editor, J. S. Suri; SPIE Press Book, 2006.
- [7] J. Tang, R.M Rangayyan, J. Xu, I. El Naqa, Y. Yang. Computer-Aided Detection and Diagnosis of Breast Cancer with Mammography: Recent Advances, *IEEE Transactions on Information Technology in Biomedicine*, Vol. 13, Num. 2, pages 236-251, 2009.
- [8] L. Yanfeng, C. Houjin, C. Lin, M. Jinyuan. A Survey of Computer-aided Detection of Breast Cancer with Mammography, *Journal of Health & Medical Informatics*, Vol. 7, Num.4, <http://dx.doi.org/10.4172/2157-7420.1000238>, pages 1-6, 2016.
- [9] Organisation mondiale de la santé (OMS). Cancer : profils des pays en 2014, <http://www.who.int/cancer/country-profiles/fr/>, date de consultation, juin 2016.

- [10] Ministère algérien de la santé. Plan national du cancer du sein 2015–2019, [http : //www.santé.gov.dz/images/plan%20cancer/plan\\_national\\_cancer.pdf](http://www.santé.gov.dz/images/plan%20cancer/plan_national_cancer.pdf), 2014, date de consultation, avril 2017.
- [11] Institut national français du cancer. Anatomie du sein. <http://www.e-cancer.fr/patients-et-proches/les-cancers/cancer-du-sein/anatomie-du-sein>, date de consultation, février 2017.
- [12] S. Oulmane, Cancer du sein, intérêt du dépistage. Cabinet médical "Ibnou Sina, <http://www.dr-somia.oulmane.com/sein.html>, date de consultation, avril 2017.
- [13] 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.
- [14] R. Ghys. Les maladies du sein, Edition Québec, Canada, 1989.
- [15] Société canadienne du cancer. Cancer du sein. <http://www.cancer.ca/fr-ca/cancer-information/cancer-type/breast/breast-cancer/the-breasts/?region=qc>, date de consultation, mars 2017.
- [16] J. Peter. Les cancers in situ. [https://www.gyneobs.com/gyneobs/In\\_situ.html](https://www.gyneobs.com/gyneobs/In_situ.html), 2017, date de consultation, janvier 2017.
- [17] Hôpital du Saint-Sacrement du CHU de Québec, Programme Québécois de dépistage du cancer du sein, Carcinome infiltrant, <http://www.depistagesein.ca/carcinome-infiltrant/>, date de consultation, mars 2017.
- [18] Vidal, Eurekasanté. Les facteurs de risque du cancer du sein. <https://eurekasante.vidal.fr/maladies/cancers/cancer-sein.html?pb=facteurs-risque>, 2017, date de consultation, septembre 2017.
- [19] B. Séradour, Le dépistage du cancer du sein : un enjeu de santé publique. Edition, Springer-Verlag Paris, 2007.
- [20]. A. Trial et A. Rescaniers. Guide pratique de techniques du radiodiagnostic. Edition, Vigot frères, 1976.
- [21] H. Tristant, M. Benmussa et J. Bokobsa. Mammographie et échographie mammaire, de l'image au diagnostic, Edition Flammarion 1988.
- [22] F.-X. Slosman. Mammographie <https://www.info-radiologie.ch/mammographie.php>, 2017, date de consultation, mars 2017.
- [23] Image In Monaco. Matériel. <https://www.image-in.mc/materiel/>, Monaco, date de consultation, mars 2017.
- [24] Centre d'imagerie Dr M. Lekhal. Mammographie, <https://cilm-dz.com/examens/imagerie-de-la-femme/mammographie/mammographie.html>, Blida, date de consultation, mars 2017.
- [25] J.P. Alunni, M.H. Marty, V. Feillel. Classification BI-RADS : densité mammaire, la lettre du sénologue, Num. 40, 2008.

- [26] Haute Autorité de Santé (HAS). Classification des anomalies mammographiques de BI-RADS, <http://www.aly-abbara.com/echographie/biometrie/scores/acr.html>, 2010, date de consultation, mars 2017.
- [27] I. Cheikhrouhou. Description et classification des masses mammaires pour le diagnostic du cancer du sein, thèse de doctorat, Université d'Evry-Val d'Essonne, France, 2012.
- [28] J.F. Heron. Interprétation mammographique, [http://www.oncoprof.net/Generale2000/g04\\_Diagnostic/Mammographie/g04\\_mm11.html](http://www.oncoprof.net/Generale2000/g04_Diagnostic/Mammographie/g04_mm11.html), 2017, date de consultation, septembre 2017.
- [29] Commissariat à l'énergie atomique et aux énergies alternatives. Imagerie médicale, <http://www.cea.fr/comprendre/Pages/sante-sciences-du-vivant/essentiel-sur-imagerie-medecale.aspx>, 2017, date de consultation, juin 2017.
- [30] C. Dillenbourg-Wertz. Le dépistage du cancer du sein en 2016, Périodique Trimestriel, Info-Cancer, Num. 84, [www.cancer.lu/sites/cancer/files/files/IC84\\_p4-7\\_stopouencore\\_historique.pdf](http://www.cancer.lu/sites/cancer/files/files/IC84_p4-7_stopouencore_historique.pdf), mars 2016, date de consultation, juin 2016.
- [31] H. Allias-Denis, B. Baroin. Imagerie médicale du futur, Edition, N. Merle-Lamoot, G. Pannetier, France, 2013.
- [32] C. Balleyguier, B. Boyer, A. Athanasiou, D. Vanel, R. Sigal. Comprendre et utiliser le CAD en mammographie, Journal de Radiologie, Vol. 8, Num.1, page 29-35, 2005.
- [33] F. Winsberg, M.D. Milton Elkin, J. Macy, V. Bordaz, W. Weymouth. Detection of Radiographic Abnormalities in Mammograms by Means of Optical Scanning and Computer Analysis, Journal of Radiology Vol. 89, Num.2, pages 211-215, 1967.
- [34] L. V. Ackerman, E. E. Gose. Breast lesion classification by computer and xeroradiography, Journal of Cancer, Vol. 30, pages 1025-1035, 1972.
- [35] W. G. Wee, M. Moskowitz, N-C. Chang, Y-C. Ting, S. Pemmeraju. Evaluation of mammographic calcifications using a computer program, Journal of Radiology, Vol. 116, pages 717-720, 1975.
- [36] G. L. Masala. Computer Aided Detection on Mammography, International Journal of Medical and Health Science, Vol.2, Num.3, pages 131-136, 2008.
- [37] W. Spiesberger. Mammogram inspection by computer, IEEE Transactions on Biomedical Engineering, Vol. 26, pages 213-219, 1979.
- [38] K. Doi. Computer-aided diagnosis in medical imaging: historical review, current status and future potential, Computer Medical Imaging Graphics, Vol.31, Num. 4, pages 198-211, 2007.
- [39] M. L. Giger, H. P. Chan, J. Boone. Anniversary Paper: History and status of CAD and quantitative image analysis: The role of medical physics, Journal of Medical Physics, Vol. 35, Num.12, pages, 5799-5820, 2008.
- [40] Image Checker. <http://www.r2tech.com>, date de consultation, juin 2016.
- [41] Second Look. <http://www.icadmed.com>, date de consultation, septembre 2018.

- [42] Carestream. Vue mammo. <https://www.carestream.com/fr/fr/medical/products/vue-healthit/radiology/carestream-vue-mammo>, 2018, date de consultation, février 2018.
- [43] Power Look Tomo. <https://www.icadmed.com/powerlook-tomo-detection.html>, date de consultation, juin 2018.
- [44] Z. Zeng, C. J. Hou, Q. H. Hu, Y. Liu, C. Wang, R. Wei, X. Ming Fan. Mammography and ultrasound effective features in differentiating basal-like and normal-like subtypes of triple negative breast cancer, *International Journal of Oncotarget*, Vol. 8, Num.45, pages 79670-79679, 2017.
- [45] A. Rick. Représentation de la variabilité dans le traitement d'images flou : Application à la mammographie numérique, thèse de doctorat, Université de Paris 6, France, 1999.
- [46] H. Mohamed, M.S. Mabrouk, A. Sharawy. Computer aided detection system for microcalcifications in digital mammograms. *Computer methods and programs in biomedicine*, Vol. 116, pages 226-235, 2014.
- [47] H. Boulehmi, H. Mahersia, K. Hamrouni. A new CAD system for breast microcalcifications diagnosis, *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 4, pages 133-143, 2016.
- [48] N. Kumar, S. Amutha, D. R. Ramesh Babu. Enhancement of mammographic images using morphology and wavelet transform, *Computer Technology and Applications*, Vol. 3 Num.1, pages 192-198, 2012.
- [49] K.L. Kashyap, M. K. Bajpai, P. Khanna. Globally supported radial basis function based collocation method for evolution of level set in mass segmentation using mammograms, *Computers in Biology and Medicine*, Vol. 87, Pages 22-37, 2017.
- [50] K. L. Kashyap, M. K. Bajpai, P. Khanna. An efficient algorithm for mass detection and shape analysis of different masses present in digital mammograms, *Multimedia Tools and Applications*, Vol. 77, Num. 8, pages 9249–9269, 2018.
- [51] C. Ferkous. Extraction des connaissances pour la segmentation d'images mammographiques, thèse de doctorat, Université d'Annaba, Algérie, 2017.
- [52] M. Hmida, K. Hamrouni, B. Solaiman, S. Boussetta. An efficient method for breast mass segmentation and classification in mammographic images, *International Journal of Advanced Computer Science and Applications*, Vol. 8 Num.11, pages 256-262, 2017.
- [53] M. M. Pawar, S. N. Talbar, A. Dudhane. Local binary patterns descriptor based on sparse curvelet coefficients for false-positive reduction in mammograms, *Journal of Healthcare Engineering*, Vol. 2018, pages 1-16, 2018.
- [54] M. Hmida. Reconnaissance de formes basée sur l'approche possibiliste dans les images mammographiques, thèse de doctorat, Université de Bretagne Loire, France, 2017.

- [55] X. Liu, X. Xu, J. Liu, J. Tang. Mass classification with level set and shape analysis for breast cancer diagnosis using mammography. Dans : DS. Huang, Y. Gan, P. Gupta, M.M. Gromiha. Advanced intelligent computing theories and applications with aspects of artificial intelligence. ICIC 2011, Lecture Notes in Computer Science, Vol. 68, Num. 39, pages 630-637, Springer, Berlin, Heidelberg, 2011.
- [56] C. Vachier. Extraction de caractéristiques, Segmentation d'image et Morphologie mathématique, thèse de doctorat, Ecole Nationale Supérieure des Mines de Paris, France, 1995.
- [57] F. Eddaoudi. Aide à la détection des pathologies du sein par les techniques de segmentation et de codage de textures des images mammographiques, thèse de doctorat, Université de Rabat, Maroc, 2012.
- [58] F. Liu, Z. Gong, Y. Chen, Y. Gu. Segmentation of masses in mammograms by a novel integrated active contour method, International Journal of Computational Science and Engineering, Vol. 11, Num.2, pages 207-215, 2015.
- [59] P. Shanmugavadivu, A. Krishnaveni, V. Sivakumar, S.G. Lakshmi Narayanan. Digital mammogram segmentation and feature extraction : a review, International Journal on Computer Science and Engineering, Vol. 9 Num.12, pages 701-705, 2017.
- [60] F. Diener, M. Diener. Classification automatique, laboratoire J.-A. Dieudonné, université de Nice - Sophia Antipolis, <https://math.unice.fr/~diener/MAB06/cluster.pdf>, date de consultation, juin 2016.
- [61] G. Dreyfus, M. Samuelides, J.-M. Martinez, M. B. Gordon, F. Badran, S. Thiria, L. Hérault. Réseaux de neurones (méthodologies et applications), Eyrolles, 2004.
- [62] P. Vincent. Fondements de l'apprentissage machine, cours, Université de Montréal, [https://www.iro.umontreal.ca/~vincentp/ift3395/cours/5\\_classifieur\\_bayes.pdf](https://www.iro.umontreal.ca/~vincentp/ift3395/cours/5_classifieur_bayes.pdf), date de consultation, juin 2016.
- [63] V. N. Vapnik. The nature of statistical learning theory, Springer-Verlag, New York, 1995.
- [64] C. Cortes, V. N. Vapnik. Support-Vector Networks, Machine Learning, Vol. 20, Num.3, pages 273-297, 1995.
- [65] D. Kopans, R. Moore, K. Bowyer. DDSM: digital database for screening mammography, <http://marathon.csee.usf.edu/Mammography>, date de consultation, juin 2016.
- [66] J. Suckling, S. Astley, D. Betal, N. Cerneaz, D. R Dance, S. L Kok, J. Parker, I. Ricketts, J. Savage, E. Stamatakis, P. Taylor. The mammographic image analysis Society digital mammogram database, <http://www.peipa.essex.ac.uk/ipa/pix/mias/>, 1994, date de consultation, septembre 2004.
- [67] H. Boulehmi. Analyse d'images mammographiques par analyse de texture : Application à la détection précoce du cancer, thèse de doctorat, Université de Tunis El Manar, Tunisie, 2016.
- [68] T. Perneger, A. Perrier. Analyse d'un test diagnostique, Revue des Maladies Respiratoires, Vol. 21, Num. 2 pages 398-401, 2004.

- [69] F. Carrat, A. Mallet, V. Morice. Evaluation de l'intérêt diagnostique des informations médicales, <http://www.chups.jussieu.fr/poly/biostats/poly/Poly.ch.5.2.html>, Université de Jussieu, France, 2014, date de consultation, juin 2016.
- [70] L. Shen, R. M. Rangayyan. Application of shape analysis to mammographic calcifications, *IEEE transactions on Medical Imaging*, Vol. 13, Num. 2, pages 263-274, 1994.
- [71] I. Christoyianni, A. Koutras, E. Dermatas, G. Kokkinakis. Computer aided diagnosis of breast cancer in digitized mammograms. *Journal of Computerized medical imaging and graphics*, Vol.26, pages: 309-319, 2002.
- [72] R.N. Panda, B. K. Panigrahi and M. R. Patro. Feature extraction for classification of microcalcifications and mass lesions in mammograms, *IJCSNS, International Journal, of Computer Science and Network Security*, Vol. 9, Num. 5, pages 255-265, 2009.
- [73] J. Dheeba, S.T. Selvi. Cascade correlation neural network based microcalcification detection in mammographic images. Dans: S.C. Satapathy, P.S. Avadhani, A. Abraham. *Proceedings of the international conference on information systems design and intelligent applications*, India, 2012. *Advances in Intelligent and Soft Computing*, Vol. 132. Springer, Berlin, Heidelberg, pages 153-160, 2012.
- [74] J.B. Li, Y. Peng, D. Liu. Quasi-conformal kernel common locality discriminant analysis with application to breast cancer diagnosis. *Journal of Information Sciences*, Vol. 223, pages 256-269, 2013.
- [75] M. Rizzi, M. D'Aloia, B. Castagnolo. Review: health care CAD systems for breast microcalcification cluster detection. *Journal of Medical and Biological Engineering*, Springer, Vol. 32, Num.3, pages 147-156, 2012.
- [76] D. Saraswathi, E. Srinivasan. An ensemble approach to diagnose breast cancer using fully complex-valued relaxation neural network classifier, *International Journal of Biomedical Engineering and Technology*, Vol. 15, Num. 3, pages 243-260, 2014.
- [77] S. Beura, B. Majhi, R. Dash. Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer, *Journal of Neurocomputing* Vol. 154, pages 1–14, 2015.
- [78] J. Kaur, M. Kaur. Automatic Cancer Detection in Mammographic Images, *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 5, Num. 7, pages 43-476, 2016.
- [79] D. Sriramkumar, R. Malmathanraj, R. Mohan S. Umamaheswari. Mammogram tumour classification using modified segmentation techniques, *International Journal of Biomedical Engineering and Technology*, Vol. 13, No. 3, pages 218-239, 2013.
- [80] X. Liu, Z. Zeng. A new automatic mass detection method for breast cancer with false positive reduction, *Journal of Neurocomputing* Vol. 152, pages 388–402, 2015.
- [81] B.N. Prathibha, V. Sadasivam. Hybrid transforms domain-based mammogram analysis using C-SVM classifier, *International Journal of Medical Engineering and Informatics*, Vol. 4, Num. 2, pages 146-156, 2012.

- [82] I. Zyout, J. Czajkowska, M. Grzegorzek. Multi-scale textural feature extraction and particle swarm optimization based model selection for false positive reduction in mammography, *Computerized Medical Imaging and Graphics*, Vol. 46, Num. 2, pages 95–107, 2015.
- [83] P. Casti, A. Mencattini, M. Salmeri, R. M. Rangayyan. *Computerized analysis of mammographic images for detection and characterization of breast cancer, synthesis lectures on biomedical engineering*, Morgan & CLaypool Publishers, 2017.
- [84] B. Kurt, V. Nabiyev, K. Turhan. An automated computer-aided detection (CADe) and diagnosis (CADx) system for breast microcalcifications in mammograms, *Journal of Engineering, Science and Technology*, Vol.6, Num.3, pp. 355-376, 2018.
- [85] T. Sadad, A. Munir, T. Saba, A. Hussain. Fuzzy C-means and region growing based classification of tumor from mammograms using hybrid texture feature, *Journal of Computational Science*, Vol. 29, pages 34–45, 2018.
- [86] B.N. Beena Ullala Mata, M. Meenakshi. Mammogram Image Segmentation by Watershed Algorithm and Classification through k-NN Classifier, *International Journal of Advances in Image Processing*, Vol. 8, Num.1, pages 1-7, 2018.
- [87] C. A. Ancy, L. S. Nair. Tumour Classification in Graph-Cut Segmented Mammograms Using GLCM Features-Fed SVM. Dans : V. Bhateja, C. Coello, S. Satapathy, P. Pattnaik, *Intelligent Engineering Informatics. Advances in Intelligent Systems and Computing*, Vol. 695, pages 197-208, Springer, Singapore, 2018.
- [88] J. Liu, J. Chen, L. Chun, J. Tang, Y. Deng. Mass segmentation using a combined method for cancer detection, *BMC Systems Biology*, Vol.3, pages 1-10, 2011.
- [89] T. F Chan, L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, Vol. 10, Num. 2, pages 266-277, 2001.
- [90] 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*, Vol. 3, No 4, pages 1-14, 2011.
- [91] B. C. Patel, G. R. Sinha. Energy and region based detection and segmentation of breast cancer mammographic. *International Journal Image, Graphics and Signal Processing*, Vol. 6, pages 44-51, 2012.
- [92] P. Sudharsan, C.L. Yashwanthi Sivakumari. A comparative analysis of segmentation techniques in mammogram images. *International Journal of Engineering Research and Technology*, Vol. 2, pages 1006-1010, 2013.
- [93] A. Lavanya, S. Narasimhalu. Inverse maximum likelihood-based edge detection for segmentation of breast lesion using active contour, *International Journal of Biomedical Engineering and Technology (IJBET)*, Volume 22, No. 3, pages 272- 283, 2016.
- [94] L. de Oliveira Martins, G. Braz Junior, A. Correa Silva, A. C. de Paiva, M. Gattass. Detection of Masses in Digital Mammograms using K-means and Support Vector Machine. *Electronic Letters on Computer Vision and Image Analysis*, Vol 8, Num. 2 pages 39-50, 2009.



- [95] B.N. Li, C. K. Chui, S. Chang, S.H. Ong. Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation, *Computers in Biology and Medicine*, Vol. 41, pages 1-10, 2011.
- [96] J. Anitha, J. Dinesh Peter. A Spatial Fuzzy based Level Set Method for Mammogram Mass Segmentation, *International Conference on Electronics and Communication System*, pages 1-6, 2015.
- [97] J. Anitha, J. Dinesh Peter. Mass segmentation in mammograms using a kernel-based fuzzy level set method, *International Journal of Biomedical Engineering and Technology*, Vol. 19, Num. 2, pages 133-153, 2015.
- [98] O. Valenzuela, I. Rojas, F. Rojas, L. Marquez. Predicting and classification breast cancer in digital mammography using soft computing techniques, *IMACS: International Association for Mathematics and Computers in Simulation*, <http://metronu.ulb.ac.be/imacs>, date de consultation, juin 2012.
- [99] O. Ludwig, U. Nunes, R. Araujo. Eigenvalue decay: a new method for neural network regularization. *International Journal of Neurocomputing*, Vol. 124, pages 33-42, 2014.
- [100] L. Aha. UCI, Center for Machine Learning and Intelligent Systems, <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer>, 1987, date de consultation, décembre, 2016.
- [101] J. Dheeba, N. A Singh, S.T Selvi. Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. *Journal of Biomedical Informatics*, Vol.49, pages 45-52, 2014.
- [102] S. Duraisamy, S. Emperumal. Computer-aided diagnosis system using deep learning convolutional fully complex-valued relaxation network classifier, *IET Computer Vision*, Vol.11, Num.8, 2017.
- [103] J.P. Cocquerez, S. Philipp. *Analyse d'images : filtrage et segmentation*, édition Masson, Paris, 1995.
- [104] 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, *Innovation and Research in Biomedical engineering*, Vol.26, Num.5, pages 347–356, 2005.
- [105] H. Cheng, X. Shi, R. Min, L. Hu, X. Cai and H. Du: Approaches for automated detection and classification of masses in mammograms, *Journal of Pattern Recognition*, Vol. 39, Num.4, pages 646-668, 2006.
- [106] F. Mohanty, S. Rup, B. Dash. An Improved CAD Framework for Digital Mammogram Classification Using Compound Local Binary Pattern and Chaotic Whale Optimization-Based Kernel Extreme Learning Machine. Dans : V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, I. Maglogiannis, *Artificial Neural Networks and Machine Learning, ICANN 2018. Lecture Notes in Computer Science*, Vol. 11140, pages 14-23, Springer, 2018.

- [107] V. Magudeeswaran, K. Balasubramanian. Fuzzy weighted histogram equalisation for contrast enhancement of mammogram images, International Journal of Biomedical Engineering and Technology, Vol.28, Num.3, pages 232-242, 2018.
- [108] S. Beucher, Segmentation d'images et morphologie mathématique, thèse de doctorat, Ecole Nationale Supérieure des Mines de Paris, France, 1990.
- [109] N. Stikov. Traitement d'images : approche morphologique, cours en ligne, école polytechnique de Montréal, [http://moodle.polymtl.ca/pluginfile.php/322302/mod\\_page/content/11/12\\_Morphologie.pdf](http://moodle.polymtl.ca/pluginfile.php/322302/mod_page/content/11/12_Morphologie.pdf), 2016, date de consultation, juin 2017.
- [110] P. Soille. Morphological analysis image, Springer-Verlag Berlin Heidelberg, 2004.
- [111] V. Risson, Application de la morphologie mathématique à l'analyse des conditions d'éclairage des images couleur, thèse de doctorat, École Nationale Supérieure des Mines de Paris, France, 2001.
- [112] M. Bergounioux. Introduction au traitement mathématique des images : méthodes déterministes, Collection Mathématiques et Vision, Vol.76, Springer, Berlin, Heidelberg, 2015.
- [113] S. Punitha, A. Amuthan, K. Suresh Joseph. Benign and malignant breast cancer segmentation using optimized region growing technique, Future Computing and Informatics Journal, <http://www.journals.elsevier.com/future-computing-and-informatics-journal/>, article en édition, novembre 2018.
- [114] D. J. Burr. Elastic machine of line drawings, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 3, Num. 6, pages 708-713, 1981.
- [115] M. Kass, A. Witkin, D. Terzopoulos. Snakes. Active Contour Models, International Journal of Computer Vision, Vol. 1, Num. 4, pages 321-331, 1988.
- [116] S. Osher and J.A. Sethian. Front propagating with curvature dependent speed : algorithms based on hamilton-jacobi formulations, Journal of Computational Physics Vol. 78, pages 12-49, 1988.
- [117] J. J. Rousselle, les contours actifs, méthode de segmentation : application à l'imagerie médicale, thèse de doctorat, Université F. Rabelais de Tours, France, 2003.
- [118] J. J. Cohen, L. D. Cohen and I. Cohen. Finite element methods for active contour models and balloons for 2D and 3D Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, pages 1131-1147, 1991.
- [119] V. Caselles, F. Catte, T. Coll, F. Dibos. A geometric model for active contours in image processing, Numerische Mathematik, Vol. 66, pages 1-31, 1993.
- [120] D. Mumford, J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems, Communications on Pure and Applied Mathematics, Vol. 42, pages 577-684, 1989.
- [121] C. Ciofolo. Segmentation de formes guidée par des modèles en neuro-imagerie, thèse de doctorat, Université de Rennes, France, 2005.

- [122] F. Boutaouche. Analyse et interprétation des images mammographiques, thèse de doctorat, Université des Sciences et Technologie d'Oran, Algérie, 2018.
- [123] J. Olivier. Méthodes d'accélération et approches supervisées pour les contours actifs : Application à la segmentation d'images 2D, 3D et texturées, thèse de doctorat, Université de Tours, France, 2009.
- [124] R. Crandall, Image segmentation using the Chan-Vese algorithm, ECE532 project, [http://math.arizona.edu/~rcrandall/ECE532\\_ProjectPaper.pdf](http://math.arizona.edu/~rcrandall/ECE532_ProjectPaper.pdf), 2009, date d'accès, janvier 2015.
- [125] X. F. Wang, S. D. Huang, H. XU. An efficient local Chan-Vese model for image segmentation, International journal of Pattern recognition, Vol. 43, pages 603-618, 2010.
- [126] C.C. Jen, S.S. Yu. Automatic detection of abnormal mammograms in mammographic images, International journal of Expert system with applications, Vol. 42, pages 3048–3055, 2015.
- [127] F. Z Reguieg, N. Benblidia, M. Guerti. Morphological Detection and Neuro-Genetic Classification of Masses and Calcifications in Mammograms for Computer-Aided Diagnosis, International Journal of Biomedical Engineering and Technology, Vol. 28, Num.3, pages 203-231, 2018.
- [128] S. Lankton, A. Tannenbaum. Localizing Region-Based Active Contours, IEEE Transactions on Image Processing, Vol. 17, Num. 11, pages 2029–2039, 2009.
- [129] L. Zadeh. Fuzzy sets. Journal of Information and Control, Vol. 8, Num. 3, pages 338-353, 1965.
- [130] I. Bloch. Flou en traitement d'images et vision, photocopié de cours, [https://perso.telecom-paristech.fr/bloch/AIV/flou\\_tdi.pdf](https://perso.telecom-paristech.fr/bloch/AIV/flou_tdi.pdf), Ecole Nationale Supérieure des Télécommunications, Paris, France, date de consultation, septembre 2016.
- [131] J. C. Dunn. A Fuzzy Relative of the isodata process and its use in detecting compact well-separated clusters, Journal of Cybernetics Vol. 3, pages 32-57, 1973.
- [132] J. C. Bezdek. Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.
- [133] M. Rastgarpour, J. Shanbehzadeh, H. Soltanian Zadeh. A hybrid method based on fuzzy clustering and local region-based level set for segmentation of inhomogeneous medical images, Journal of Medical System Vol. 38, Num. 8, 2014.
- [134] V. Bhateja, A. Tiwari, A. Gautam. Haralick Features-Based Classification of Mammograms Using SVM. Dans : V. Bhateja, N. Nguyen, S. Satapathy, D.N Le. Information Systems Design and Intelligent Applications. Advances in Intelligent Systems and Computing, Vol. 672, pages 787-795, Springer, Singapore, 2018.
- [135] X. Liu, J. Tang. Mass Classification in Mammograms Using Selected Geometry and Texture Features, and a New SVM-Based Feature Selection Method, IEEE Systems Journal, Vol. 8, Num.3, pages 910-920, 2014.
- [136] J. Sklansky. Image segmentation and feature extraction. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 8, pages 237–247, 1978.

- [137] R. M Haralick. Statistical and structural approaches to texture. Proceedings of the IEEE, Vol.67, Num. 5, pages 786–804, 1979.
- [138] M. Unser. Description statistique de la texture. Application à l’inspection automatique. Thèse de doctorat, Ecole Polytechnique Fédérale de Lausanne, Suisse, 1984.
- [139] L. Paulhac. Outils et méthodes d’analyse d’images 3D texturées : Application à la segmentation des images échographiques, thèse de doctorat, Université de Tours, France, 2009.
- [140] 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, France, 1999.
- [141] M. Tuceryan, A. K. Jain. Texture Analysis, pages 235–276, Handbook of Pattern Recognition and Computer Vision, 1993.
- [142] B. Julesz. Textons the elements of texture perception and their interactions. Nature, Vol. 290, Num.5802, pages 91–97, 1981.
- [143] D.C. He, L. Wang. Texture unit, texture spectrum, and texture analysis. IEEE Transactions on Geoscience and Remote Sensing, Vol. 28, Num. 4, pages 509-512, 1990.
- [144] T. Ojala, M. Pietikainen, T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, Num.7, pages 971 –987, 2002.
- [145] D. Sankar, T. Thomas. Classification of Mammograms into Normal, Benign and Malignant based on Fractal Features, International Journal of Image, Graphics and Signal Processing, Vol. 3, Num. 5, pages 36-44, 2016.
- [146] D. Gabor. Theory of communication, J. Inst. Elect. Eng, Vol. 93, pages 429- 457, 1946.
- [147] U. Raghavendra, U. Rajendra Acharya, H. Fujita, A. Gudigar, J. Hong Tan, S. Chokkadi. Application of Gabor wavelet and Locality Sensitive Discriminant Analysis for automated identification of breast cancer using digitized mammogram images. Applied Soft Computing, Vol. 46, pages 151–161, 2016.
- [148] S. MALLAT. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, Num. 7, pages 674-693, 1989.
- [149] N. Gedik. A new feature extraction method based on multi-resolution representations of mammograms, Applied Soft Computing Vol. 44, pages 128–133, 2016.
- [150] G. Lavanya, D. Susdarvizhi. Breast tumour detection and classification using naïve Bayes classifier algorithm. International Journal of Emerging Trends in Engineering and Development, Vol. 3, Num. 1, pages 79-86, 2012.
- [151] J.I. Mohammed, A. Majid, A.S Maher. An efficient automatic mass classification method in digitized mammograms using artificial neural network. International Journal of Artificial Intelligence and Applications, Vol.1, Num.3, pages 1-13, 2010.

- [152] A. Arafi, Y. Safi, R. Fajr, A. Bouroumi. Classification of mammographic images using Artificial neural networks, *Applied Mathematical Sciences*, Vol. 7, Num. 89, pages 4415 - 4423, 2013.
- [153] D. Sehrawat, A. Sehrawat, D. Jaiswal, A. Sen. Detection and Classification of Tumor in Mammograms using Discrete Wavelet Transform and Support Vector Machine, *International Research Journal of Engineering and Technology*, Vol. 4, Num.3, Pages 1328-1334, 2017.
- [154] F.Z. Reguieg, N. Benblidia, M. Guerti, Caractérisation morphométrique des calcifications mammographiques par analyse multirésolution Congrès International en Informatique Appliquée CIIA05, Université de Bordj Bou Arréridj, Algérie, Novembre 2005.
- [155] M.K Hu. Visual pattern recognition by moment invariants, *Transactions on Information Theory*, Vol. 8, Num. 2, pages 179-187, 1962.
- [156] A. M. Turing. Computing machinery and intelligence. *Mind: a quarterly review of psychology and philosophy*, Oxford University Press, Vol. 59, No. 236 pages 433-460, 1950.
- [157] C. Gagné. Perceptron multicouche, polycopié de cours, Université Laval, Canada, 2016.
- [158] P. Jauffret. Introduction aux algorithmes génétiques, polycopié de cours, Université de Strasbourg, [https://masterchemoinfo.u-strasbg.fr/Documents/Fouille\\_de\\_donnees/Cours/algo-genetique.pdf](https://masterchemoinfo.u-strasbg.fr/Documents/Fouille_de_donnees/Cours/algo-genetique.pdf), date de consultation, septembre 2016.
- [159] Wikistat. Machines à vecteurs supports, <https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-m-app-svm.pdf>, date de consultation, septembre 2016.
- [160] A. Djeflal. Utilisation des méthodes Support Vecteur Machine (SVM) dans l'analyse des bases de données, thèse de doctorat, Université de Biskra, Algérie, 2012.
- [161] J. Weston, C. Watkins. Multi classes support vector machine, rapport technique, CSD TR 98-04, Université Royale Holloway de Londres, 1998.

## Liste des contributions de l'auteur

---

- [1] **F.Z Reguieg**, N. Benblidia, M. Guerti. Caractérisation morphométrique des calcifications mammographiques par analyse multirésolution Congrès International en Informatique Appliquée CIIA05, Université de Bordj Bou Arréridj, Novembre 2005.
- [2] **F.Z Reguieg**, N. Benblidia, Classification automatique d'images mammaires par corrélation d'histogrammes locaux, 1<sup>ère</sup> conférence internationale sur le génie électrique, ICEE2000, Novembre 2000, université de Boumerdès, Algérie, Novembre 2000.
- [3] **F.Z Reguieg**, N. Benblidia, M. Guerti, A. Guessoum. Morphometric Appreciation of Mammary Breast Calcifications Around Multiresolution Analysis by Inertia and Frequentiel Factors, Conférence Arabe Internationale sur la technologie, ACIT2005, Jordanie, [www.acit2k.org/ACIT](http://www.acit2k.org/ACIT), 2005.
- [4] S. Belhadi, N. Benblidia, **F.Z Reguieg**, A. Guessoum. Détection automatique des microanévrismes sur des angiographies rétinienne de patients diabétiques. 2<sup>èmes</sup> Journées d'Études Algéro-françaises en Imagerie Médicale (JETIM'06), USTHB, 2006.
- [5] L. Aoudjit, N. Benblidia, **F.Z Reguieg**. Indexation et Recherche d'Images par le Contenu Visuel, Application au Diagnostic de la Rétinopathie Diabétique, Séminaire National en Informatique, SNIB2010, Université de Biskra, Algérie, Novembre 2010.
- [6] S. Belhadi, N. Benblidia, **F. Z Reguieg**. Automated Detection and Quantification of Microaneurysms and Blood Vessels in Diabetic Retinopathy Images. 2<sup>nd</sup> Neuromed School. Université de Blida 1, Algérie, 2011.
- [7] **F.Z Reguieg**, N. Benblidia, M. Guerti. Morphological Detection and Neuro-Genetic Classification of Masses and Calcifications in Mammograms for Computer-Aided Diagnosis, International Journal of biomedical engineering and technology, IJBET, Vol. 28, Num.3, pages 203-231, 2018.