NNT : 2016SACLE045

Thèse de doctorat
de
l'Université Paris-Saclay
préparée à
l'université d'Evry Val d'Essonne

Ecole doctorale n°580
Sciences et Technologies de l'Information et de la Communication
Spécialité de doctorat : Traitement du Signal et des Images
par

M. Sègbédji Rethice Théophile Junior Goubalan

Titre de la thèse
Contributions à l'analyse d'images médicales pour la reconnaissance du cancer du sein

Thèse présentée et soutenue à Evry, le 09 décembre 2016.

Composition du Jury :

M.	Yves GOUSSARD	Professeur Ecole Polytechnique de Montréal	(Président)
M.	William PUECH	Professeur Université de Montpellier	(Rapporteur)
M.	Rachid JENNANE	Professeur Université d'Orléans	(Rapporteur)
M.	Mouloud ADEL	Professeur Université d'Aix-Marseille	(Examinateur)
M.	Hichem MAAREF	Professeur Université d'Evry Val d'Essonne	(Directeur de thèse)
M.	Khalifa DJEMAL	Maître de conférences Université d'Evry Val d'Essonne	(Co-directeur de thèse)

Dédicaces

Il n'y a que deux guerriers qui soient toujours vainqueurs, le temps et la patience.

Léon Tolstoï, *Guerre et Paix*

A mon père *Serge*
A ma mère *Bernadette*
A *Fortune Précieuse*

Remerciements

J'adresse mes sincères remerciements à mon directeur de thèse, M. Hichem MAAREF, Professeur de l'Université d'Evry Val d'Essonne, pour son encadrement, ses conseils et surtout pour la confiance qu'il a placé en ma personne. Cette confiance m'a permis d'explorer différentes pistes scientifiques afin de mener à bien ce projet de thèse.

Mes remerciements vont également à l'endroit de mon encadrant, M. Khalifa DJEMAL, maître de conférences de l'Université d'Evry Val d'Essonne, pour son soutien et ses conseils.

Mes chaleureux remerciements vont à l'endroit de M. Yves GOUSSARD, Professeur de l'Ecole Polytechnique de Montréal, pour m'avoir fait l'honneur de prendre part à mon jury de soutenance de thèse. Je le remercie aussi pour son encadrement et ses conseils lors de mon séjour à l'Ecole Polytechnique de Montréal dans le cadre de la bourse Mitacs-Campus France et bien après. J'ai beaucoup appris à ses côtés (pas uniquement son expertise des champs de Markov, mais également que le fait d'être une référence dans un domaine n'empêche pas d'être humble) et j'ai vraiment apprécié le fait qu'il ait toujours su me consacrer de son temps malgré son emploi du temps très chargé.

Je tiens à exprimer toute ma gratitude à M. William PUECH, Professeur de l'Université de Montpellier ainsi qu'à M. Rachid JENNANE, Professeur de l'Université d'Orléans pour avoir accepté juger mes travaux en tant que rapporteurs. Je souhaite aussi les remercier pour leurs apports constructifs lors de l'évaluation de ces travaux. Mes remerciements les plus chaleureux vont également à l'endroit de M. Mouloud ADEL, Professeur de l'Université d'Aix-Marseille qui a accepté prendre part à mon jury de soutenance de thèse.

Je remercie toutes les personnes avec lesquelles j'ai travaillé et surtout collaboré pendant ma thèse, les enseignants-chercheurs qui m'ont accordés des centaines heures d' enseignements à l'université, en particulier les sieurs Jean-Yves DIDIER, Alain MOLLICA et Malick MALLEM pour ne citer que ceux-là. Grâce à vous j'ai développé plusieurs aptitudes en matière d'enseignement au niveau universitaire, licence et master inclus.

J'aimerais également exprimer toute ma sympathie au personnel administratif de l'Université d'Evry Val d'Essonne pour leur disponibilité en particulier Sabine, Aissatou et Gislin. Je souhaite remercier les étudiants qui ont ravivé chez moi l'envie d'enseigner au cours de ces 3 années de partage.

Je remercie tous les membres des laboratoires IBISC et NeuroPoly avec une pensée particulière pour tous les doctorants, post-doctorants, ingénieurs de recherche anciens et nouveaux avec lesquels nous avons discuté de la recherche, de la science et de nombreux autres sujets. Je veux citer Alia, Dmitri, Houssem, Rima, Khouloud, Fida, Elie, Aldo, Tanguy, Benjamin, Ariane, Mai, Jean-clément, Maxime, Hakim et tous les autres, merci pour la bonne ambiance de travail.

J'aimerais finir en remerciant mes frères Marcel et Rodolphe pour leur soutien infaillible, ainsi que leurs prières qui m'accompagnent. Je ne pourrai finir sans dire merci à mes amis qui de près ou de loin m'ont toujours soutenu, je veux citer Côme, Florent, Wilfried, Cédric, Murielle, Evelyne, Mireille, Afi, Deo-Gratias, Guevara, Riad, Loic, Pietro, Benedicta, Expédit, Gildas et tous ceux dont je n'ai pas pu mentionner le nom mais qui ont toujours été là pour moi.

Table des matières

1 Généralités sur le cancer du sein **7**

1.1 Introduction . . . 7
1.2 Anatomie du sein chez la femme . . . 7
1.3 Le cancer du sein . . . 8
1.4 Dépistage du cancer du sein . . . 10
1.5 L'imagerie médicale et le cancer du sein . . . 11
1.5.1 L'échographie . . . 12
1.5.2 L'imagerie par résonance magnétique . . . 13
1.5.3 La mammographie . . . 14
1.5.4 La tomosynthèse du sein . . . 16
1.6 Description des masses . . . 18
1.6.1 La forme . . . 19
1.6.2 Le contour . . . 19
1.6.3 La densité . . . 20
1.7 Classification des pathologies du sein . . . 21
1.8 La détection ou le diagnostic assisté par ordinateur . . . 22
1.8.1 Importance des systèmes de diagnostic assisté par ordinateur . . . 22
1.8.2 La détection assistée par ordinateur . . . 23
1.8.3 Le diagnostic assisté par ordinateur . . . 24
1.9 Les bases de mammographies . . . 25
1.9.1 Les exigences . . . 25
1.9.2 Mammographic Image Analysis Society Digital Mammogram Database . . . 26
1.9.3 Digital Database for Screening Mammography . . . 26
1.9.4 BancoWeb LAPIMO . . . 26
1.9.5 INbreast . . . 27
1.10 Problématique . . . 27
1.11 Conclusion . . . 29

2 Pré-traitement des mammographies **32**

2.1 Travaux existants sur le pré-traitement en mammographie . . . 33
2.2 Méthode proposée . . . 34
2.2.1 Extraction des patchs et pré-traitement . . . 34
2.2.2 Apprentissage du dictionnaire parcimonieux et algorithme du K-SVD . . . 35
2.3 Réduction de la taille du dictionnaire . . . 37
2.4 Suppression du bruit . . . 38
2.5 Implémentation numérique . . . 39
2.6 Résultats et discussion . . . 40
2.6.1 Application aux ROIs . . . 40
2.7 Extension aux mammographies entières . . . 43
2.8 Conclusion . . . 44

3 Segmentation des masses et extraction des spicules **47**
3.1 Segmentation des masses . . . 47
3.2 Etat de l'art des méthodes de segmentation . . . 48
3.2.1 Segmentation basée région . . . 48
3.2.2 Segmentation basée modèle . . . 54
3.3 Segmentation supervisée ou non-supervisée . . . 55
3.4 Description du problème de la segmentation . . . 56
3.4.1 Pré-traitement des données . . . 59
3.5 Méthodologie de validation . . . 64
3.5.1 Matériel . . . 64
3.5.2 Evaluation de la performance de la segmentation . . . 65
3.5.3 Approches retenues pour comparaison . . . 66
3.5.4 Implémentation numérique . . . 67
3.6 Segmentation des masses : résultats et discussion . . . 68
3.7 Extraction des spicules . . . 71
3.8 Les motifs de convergence en mammographie . . . 74
3.9 Travaux existants sur les lésions stellaires . . . 74
3.10 Approche proposée pour l'annotation des spicules . . . 77
3.10.1 Extraction des structures allongées . . . 78
3.10.2 Détection des segments convergents . . . 80
3.11 Extraction des spicules : résultats et discussion . . . 81
3.12 Conclusion . . . 84

4 Description des masses et aide à la décision **86**
4.1 Introduction . . . 86
4.2 Description de la forme . . . 87
4.2.1 Descripteurs standards . . . 88
4.2.2 Descripteurs spécifiques . . . 88
4.3 Description de la texture . . . 88
4.4 Evaluation de la performance de la classification . . . 89
4.4.1 La matrice de confusion . . . 89
4.4.2 La courbe ROC . . . 89
4.4.3 La mesure de la performance . . . 89
4.5 Classification des masses mammaires . . . 90
4.6 Classification de la densité mammaire . . . 92
4.7 Les systèmes de DAOx basés CBIR . . . 93
4.8 Le fossé sémantique . . . 95
4.9 Les réseaux de neurones artificiels . . . 96
4.9.1 Les réseaux de neurones à convolution . . . 96
4.9.2 Apprentissage d'un réseau artificiel de neurones . . . 98
4.10 Apprentissage de représentation profonde par *Deep learning* . . . 104
4.10.1 L'apprentissage et composition du AlexNet . . . 104
4.11 La solution proposée . . . 104
4.11.1 Implémentation numérique . . . 106

4.11.2 Résultats . . . 107
4.11.3 Discussion . . . 107
4.12 Les défis et les tendances du *Deep Learning* . . . 113
4.12.1 Compréhension des principes théoriques . . . 113
4.12.2 La vision à l'échelle humain . . . 114
4.12.3 Apprentissage avec des données limitées . . . 115
4.12.4 La complexité en temps de calcul . . . 115
4.12.5 Des modèles plus puissants . . . 116
4.13 Conclusion . . . 117

A Les descripteurs de forme **123**
A.1 Les descripteurs standards . . . 123
A.1.1 Les moments de Hu . . . 123
A.1.2 Les moments de Zernike . . . 124
A.1.3 Les moments de Krawtchouk . . . 124
A.2 Les descripteurs spécifiques . . . 125

B Les descripteurs standards de texture **129**
B.1 Les filtres dyadiques de Gabor . . . 129
B.2 Les ondelettes . . . 129

C Approches de sélection des descripteurs **131**
C.1 La sélection par mérite individuel . . . 131
C.2 La sélection séquentielle . . . 131
C.3 La sélection par algorithmes génétiques . . . 132

Bibliographie **136**

Table des figures

1.1 Anatomie du sein chez la femme : (a) structure anatomique du sein et (b) vascularisation du sein. 8
1.2 Tendances chronologiques de l'incidence et de la mortalité liées au cancer du sein chez la femme de 1980 à 2012. Image tirée de [Binder-Foucard et al., 2013]. . . . 9
1.3 Répartition du cancer du sein chez la femme en 2012, en fonction de leur âge. Image tirée de [Binder-Foucard et al., 2013]. 10
1.4 Echographie mammaire d'une lésion maligne : (a) image en niveau de gris et (b) échographie doppler-couleur montrant une masse maligne au contour spiculé. . . 12
1.5 IRM mammaire : (a) image du sein sans injection de produit de contraste et (b) image du sein avec injection de gadolinium après 45 minutes. 13
1.6 La compostion d'un mammographe. 14
1.7 Illustrations des incidences principales en mammographie : (a) incidence cranio caudale (CC), (b) incidence médio latérale oblique (MLO) et (c) incidence de profil. Image tirée de [Heron, 2009]. 15
1.8 Illustration du principe de la tomosynthèse 17
1.9 Les formes des opacités selon le BI-RADS : (a) ronde, (b) ovale, (c) lobulaire et (d) irrégulière. 19
1.10 Les différents types de contour des opacités selon le BI-RADS : (a) contour circonsrit, (b) contour masqué, (c) contour micro-lobulé, (d) contour mal défini et (e) contour spiculé. 20
1.11 Exemple de la forme d'une distorsion architecturale 20
1.12 Types de densité des masses selon le BI-RADS : (a) B1 (densité $< 25\%$), (b) B2 (densité comprise entre 25 et 50%), (c) B3 (densité comprise entre 50 et 75%) et (d) B4 (densité $> 75\%$). 21
1.13 Bases de mammographies : (a) image issue de la MIAS, (b) image provenant de la DDSM, (c) image de la BancoWeb LAPIMO, (d) image de la INbreast, (e) amas de microcalcifications de la INbreast et (f) masse spiculée de la INbreast. 28
1.14 Schéma des différentes étapes d'un système de DAOx. 30

2.1 Extraction des patchs et pré-traitement : (a) image originale, (b) ROI, (c) extraction de patchs sur (b) avec une insertion d'un zoom x5 et(d) patchs normalisés . 34
2.2 Apprentissage de dictionnaire parcimonieux : (a) dictionnaire appris sur une base de patchs et (b) courbe d'erreur d'apprentissage. Notez la décroissance monotone de l'erreur d'apprentissage en (b). 37
2.3 Pourcentages cumulés des valeurs propres. 40
2.4 Comparaison de quatre méthodes de suppression de bruit sur une image extrait de la MIAS : (a) ROI, (b) GN, (c) K-SVD suivie d'une normalisation de l'histogramme, (d) K-SVD-ACP suivie d'une normalisation de l'histogramme, (e) ondelettes de Daubechies et (f) *curvelet*. 41

2.5 Comparaison de quatre méthodes de suppression de bruit sur une image extrait de la INbreast : (a) ROI, (b) GN, (c) K-SVD suivie d'une normalisation de l'histogramme, (d) K-SVD-ACP suivie d'une normalisation de l'histogramme, (e) ondelettes de Daubechies et (f) *curvelet*. 42
2.6 Suppression du bruit et amélioration de contraste sur l'image du haut provenant de la MIAS et celle du bas provenant du INbreast : (a) images originales non-bruitées et (b) résultats après application du K-SVD-ACP. 44
2.7 Suppression du bruit et amélioration de contraste sur l'image du haut provenant de la MIAS et celle du bas provenant du INbreast : (a) images originales bruitées et (b) résultats après application du K-SVD-ACP. 44

3.1 Amélioration de contraste : (a) image originale, (b) étirement d'histogramme et (c) transformation logarithmique de l'image originale. 60
3.2 Segmentation des masses malignes en mammographie : (a) ROIs, (b) segmentation automatique par log-PRF, (c) image binaire de (b) et (d) vérité terrain. 71
3.3 Segmentation des masses malignes en mammographie : (a) ROIs, (b) FH, (c) KM, (d) PRI et (e) vérité terrain. 72
3.4 Segmentation des masses malignes en mammographie : les cas d'échec. (a) ROIs, (b) segmentation automatique par log-PRF et (c) vérité terrain. 72
3.5 Exemple d'une lésion spiculée (a) et d'une distorsion architecturale (b) en mammographie. 74
3.6 Extraction des spicules : (a) Région d'intérêt (ROI), (b) suppression du bruit et amélioration de contraste, (c) segmentation de la masse, (d) version binaire de l'image en (c), (e) discrétisation des structures curvilignes, (f) spicules annotées, (g) superposition des images (a, f) et (h) résultat après réduction des faux positifs. 83

4.1 Relation connue entre la complexité du contour et la malignité. (a) masse circonscrite circulaire, (b) masse circonscrite ovale, (c) masse macro lobulée, (d) masse micro lobulée, (e) masse irrégulière, (f) masse spiculée. 87
4.2 Reconnaissance de visages avec un CNN. Image tirée de [Jones et al., 2014]. . . . 97
4.3 Illustration des différents composants d'un CNN : (a) architecture générale d'un CNN, (b) fonctionnement d'une couche de convolution, (c) couche de sous- échantillonnage et (d) les couches *fully-connected* (FC). 99
4.4 Illustrations des images intervenant dans le processus d'aide à la décision : (a) ROI, (b) image binaire obtenue après segmentation et (c) les spicules. 105
4.5 Illustration du déroulement de la phase *off-line*. 106
4.6 Classification des masses en mammographie selon leur morphologie : (a) ROIs bénigne et maligne, (b) diagnostic fourni par notre système. 108
4.7 Classification des masses en mammographie selon la densité de leur centre : (a) ROIs bénigne et maligne, (b) diagnostic fourni par notre système. 108
4.8 Illustration du déroulement de la phase *on-line*. 109
4.9 Matrice de confusion de la classification des masses dans les classes bénigne et maligne. 110
4.10 Courbe ROC des descripteurs extraits par AlexNet. 111

4.11 Matrice de confusion de la classification des masses dans les classes graisseuse et dense. 112

Liste des tableaux

1.1 La classification BI-RADS de l'ACR. 22
1.2 Interprétation et conduite à tenir (CAT) en fonction des classes du BI-RADS. . . 23
1.3 Synthèse des attributs des différentes bases d'images les plus utilisées pour la conception et la validation des systèmes de DAOx. 28

2.1 Les valeurs des moyenne et écart-type de l'EQM obtenues sur les bases de ROIs de la MIAS et du INbreast après exécution des ondelettes de Daubechies, du K-SVD, K-SVD-ACP et des *curvelet* sur des images numériquement bruitées. 42
2.2 Les valeurs des moyenne et écart-type du PSNR obtenues sur les bases de ROIs de la MIAS et du INbreast après exécution des ondelettes de Daubechies, du K-SVD, K-SVD-ACP et des *curvelet* sur des images numériquement bruitées. 43
2.3 Les valeurs des moyenne et écart-type du temps de calcul (CTe) obtenues sur les bases de ROIs de la MIAS et du INbreast après exécution des ondelettes de Daubechies, du K-SVD, K-SVD-ACP et des *curvelet* sur des images numériquement bruitées. 43

3.1 Statistiques descriptives de la répartition des valeurs de AOR basées sur le log-PRF 70
3.2 Segmentation des masses mammaires : Evaluation de la performance des méthodes de l'état-de-l'art en mammographie 70
3.3 Evaluation de la performance des meilleures méthodes de segmentation non-supervisées de l'état-de-l'art 71
3.4 Evaluation du temps de calcul 73
3.5 Comparaison de quelques travaux de recherche sur la base des métriques utilisées 82

4.1 Performance des différents descripteurs de forme en terme de l'aire sous la courbe ROC. 111

Introduction générale

Contexte et motivations

Le cancer du sein représente l'un des enjeux majeurs de la santé publique, en raison du fait qu'il est le cancer le plus fréquent chez la femme et la première cause de mortalité par cancer en France et dans le monde. En effet, près d'une femme sur sept sera concernée par cette pathologie durant son existence, le risque augmentant avec l'âge. De plus, des études réalisées dans le monde en 2012, font état de 522000 décès dus au cancer du sein au cours de la même année civile, ce qui représente une augmentation de 14% par rapport à 2008. En France, l'Agence Nationale de Santé Publique dont le rôle est de surveiller et d'observer de façon permanente l'état de santé de la population, a étudié en 2013 les tendances évolutives de l'incidence et de la mortalité liées au cancer sur la période 1980-2012. Les conclusions indiquent que l'incidence du cancer du sein qui a beaucoup augmenté entre 1980 et 2000 est en diminution progressive depuis 2005. En effet, le taux d'incidence standardisé à la population mondiale a augmenté de 1,4% par an en moyenne entre 1980 et 2012, mais on note une diminution un peu plus importante de 1,5% par an si on prend en compte uniquement la période allant de 2005 à 2012. Quant à la mortalité, elle est restée pratiquement stable jusqu'en 1995 et ce en dépit de la forte augmentation de l'incidence sur la même période ; avant de connaître une chute significative en 2012. En effet, on observe une dimunition moyenne de la mortalité qui est passée de 0,6% par an entre 1980 et 2012 à 1,5% par an entre 2005 et 2012. A l'inverse de certaines régions du monde, le nombre de décès liés au cancer du sein en France ne cesse de diminuer, et ce malgré l'apparition de 48763 nouveaux cas en 2012. Cela s'explique en partie par l'amélioration de la prise en charge thérapeutique et au diagnostic plus précoce grâce aux campagnes de dépistage. Dans l'objectif d'assurer le dépistage précoce de la tumeur, les radiologues ont augmenté la fréquence des mammographies – qui est l'examen de référence utilisé dans le cadre du dépistage –, notamment pour la tranche d'âge la plus concernée, à savoir 50 à 74 ans en France, ce qui représente une quantité considérable de données, rendant fastidieuse et chronophage le processus d'interprétation des images par les radiologues.

Au regard de l'augmentation du nombre de mammographies ces dernières années, différents travaux de recherche ont été menés soit pour détecter automatiquement les lésions – *c-à-d.* micro-calcifications ou masses – dans les images à l'aide des systèmes de détection assistée par ordinateur soit pour fournir un second avis à propos de la lésion détectée à travers des systèmes de diagnostic assisté par ordinateur. Ceci a conduit à la conception de plusieurs méthodes de traitement d'images mammographiques ([Oliver et al., 2008], [Tao et al., 2010], [Palma et al., 2014]) toutes s'inspirant de la classification *Breast Imaging Reporting and Data System* (BI-RADS) de l' *American College of Radiology* (ACR) pour classifier les masses en deux catégories à savoir bénignes et malignes.

Le cas qui nous occupe dans ce travail est la conception des systèmes de diagnostic assisté par ordinateur et se décline en cinq étapes que sont le pré-traitement, la segmentation, l'extraction des spicules et d'autres descripteurs, l'apprentissage d'un modèle et la classification. La différence qui existe entre un système de détection et un système de diagnostic, c'est que ce dernier intègre

l'extraction de descripteurs et leur classification afin de fournir de l'aide au diagnostic. Toutefois, la qualité et la pertinence du second avis fourni par tout système de diagnostic dépendent fortement de la qualité de l'image, de la précision de la segmentation des masses et dans notre cas, de la détection des structures d'intérêt à savoir les spicules. C'est ce qui justifie le fait que nous nous intéressons dans cette thèse à toutes ces étapes. Contrairement à ce qui s'est toujours fait dans l'état-de-l'art, où les chercheurs ont systématiquement cherché à classifier les masses comme bénignes ou malignes, ce travail met surtout l'accent sur la distinction entre les masses spiculées et les distorsions architecturales notamment en raison de la difficulté éprouvée par les radiologues à les différencier sur des images de superposition obtenues à l'issue de la mammographie ; ce qui représente un vrai défi en vision par ordinateur. En effet, selon le BI-RADS, les masses bénignes ont une forme ronde ou ovale et un contour circonscrit ou microlobulé, tandis que les masses malignes ont d'habitude une forme lobulée ou irrégulière et un contour indistinct ou spiculé. On constate donc que la morphologie des masses peut s'avérer suffisante pour séparer les masses malignes de celles bénignes. Ceci n'est évidemment pas suffisant pour distinguer les masses malignes entre elles et nécessite la combinaison de plusieurs autres critères pour espérer y parvenir.

Dans ce travail, nous avons conçu un système de diagnostic assisté par ordinateur entièrement automatique et dédié à la fois à l'extraction des masses spiculées et des distorsions architecturales. Pour y parvenir, nous avons proposé une méthode de pré-traitement des images afin de supprimer le bruit et d'améliorer le contraste entre les différentes structures anatomiques ; ensuite nous avons mis en place une approche de segmentation non-supervisée des masses qui est à la fois robuste et efficace. Puis, nous avons procédé à l'extraction des spicules de l'image par l'intermédiaire d'une approche originale de segmentation de structures d'intérêt. Les trois méthodes proposées ont été par la suite comparées avec les approches les plus efficaces de l'état-de-l'art afin de prouver leur efficacité. Les différentes stratégies de comparaison mises en place ont permis de prouver que nos méthodes sont à la fois bien plus rapides et plus efficaces que celles proposées dans la littérature. A notre connaissance, à ce jour, aucune étude n'existe sur la distinction entre les masses malignes spiculées et les distorsions architecturales au travers de l'extraction des spicules ; ce travail peut donc être vu comme le premier du genre. Pour finir, nous avons procédé à une extraction non-supervisée des descripteurs à l'aide d'une méthode issue du *Deep learning* à savoir les réseaux de neurones artificiels et plus précisément les réseaux de neurones à convolution. Les descripteurs extraits sont ensuite appris pour concevoir un modèle dont nous nous servons pour fournir de l'aide au diagnostic, soit sur la morphologie de la masse ou soit sur la nature dense ou graisseuse du centre de cette dernière.

Contributions et plan du manuscrit

Les recherches menées dans le cadre de cette thèse portent sur l'étude de la nature maligne des masses mammaires, l'étude des différentes étapes d'un système de diagnostic assisté par ordinateur, la comparaison des différentes stratégies d'aide à la décision en imagerie du sein et l'éventuel apport et les inconvénients relatifs au *deep learning*. L'originalité de notre travail réside fondamentalement dans : (a) la proposition d'une approche de réduction de la taille d'un dictionnaire parcimonieux appris sur une base d'images mammographiques pour la réduction du

bruit présent dans les images, (b) la conception d'une méthode de segmentation non- supervisée qui est à la fois robuste, efficace, rapide et ne nécessite qu'un seul paramètre à préciser et (c) le développement d'une technique originale d'extraction des spicules de l'image et (d) la construction de deux modèles offrant une aide au diagnostic à la fois sur le caractère bénin/malin d'une masse en se fondant sur sa forme, de même que sur la nature dense/graisseuse du centre de la masse. Ce document se compose de quatre chapitres organisés comme suit :

Dans le premier chapitre, l'anatomie du sein a d'abord été étudiée, puis le cancer du sein est présenté afin de justifier l'intérêt que nous lui portons. Ensuite, une présentation des examens radiologiques que sont le dépistage et le diagnostic du cancer du sein a permis de mettre en lumière le besoin d'automatiser l'interprétation des mammographies par les radiologues. Nous avons ensuite évoqué les outils d'imagerie médicale utilisés dans le cadre du dépistage et du diagnostic tout en mettant l'accent sur les avantages et les inconvénients de chacun d'eux. Puis, nous exposons le lien qui existe entre la composition des structures anatomiques du sein et leur apparence sur une image après la réalisation de la mammographie. Ensuite, nous présentons les caractéristiques – *c-à-d.* la forme, le contour et la densité – des masses afin de différencier non seulement les masses bénignes des masses malignes mais aussi et surtout pour identifier les différences entre les masses malignes spiculées et les distorsions architecturales. La classification dont nous nous sommes servis pour cette caractérisation nous a également permis d'évoquer les principaux lexiques exitants en mammographie comme la classification de Legal et celle du BI-RADS. Dans cette thèse, c'est la classification BI-RADS de l'ACR qui a été adoptée car c'est celle qui est à la fois la plus complète et la plus utilisée à travers le monde. De plus, le risque associé à chaque classe du BI-RADS et la conduite à tenir par les radiologues ont été exposés. Par la suite, on définit les systèmes de détection et de diagnostic assisté(e) par ordinateur de même que les différentes étapes qui les constituent. Cela nous a permis de mettre l'accent sur la problématique de cette thèse à savoir la distinction entre les masses spiculées et les distorsions architecturales. Finalement, les différentes bases d'images qui existent et qui sont utilisées pour concevoir les systèmes de détection ou de diagnostic ont été largement évoquées.

Dans le deuxième chapitre, après avoir mentionné les différentes méthodes standards de pré-traitement utilisées en traitement d'images, nous avons évoqué les inconvénients qu'elles présentent pour des images mammographiques. Cela nous a permis d'exposer les méthodes de pré-traitement spécifiques et sophistiquées conçues par les chercheurs en vision par ordinateur pour pallier les limites des approches classiques. Nous avons par la suite identifié les problèmes du choix de façon empirique des paramètres indispensables au bon fonctionnement de ces approches et le temps de calcul relativement important qu'elles peuvent demander pour parvenir à des résultats de qualité moyenne. Ceci nous a servi de transition pour évoquer une piste d'amélioration des approches de pré-traitement, qui allie à la fois la robustesse, la rapidité numérique et supprime toute interaction – qui peut s'avérer contre productive – avec l'utilisateur. Notre méthode repose sur l'apprentissage d'un dictionnaire parcimonieux sur notre base d'images – *c-à-d.* une représentation condensée des images introduisant un grand nombre de valeurs nulles pour ne retenir que l'information la plus pertinente ; laquelle information est obtenue par le fait que seul un petit nombre d'éléments du dictionnaire intervient dans la décomposition – à l'aide d'une technique de *machine learning* qui s'appelle le K-*Singular Value Decomposition* (K-SVD). Le problème lié à l'apprentissage d'un dictionnaire c'est la taille de ce dernier, *c-à-d.* qu'elle est la taille offrant les performances optimales ? D'après la littérature, plus la taille est grande

et meilleures sont les performances, mais cela implique un temps de calcul important aussi bien pour l'apprentissage que pour la reconstruction. Nous avons donc proposé de combiner le K-SVD avec une technique de réduction de dimension très connue à savoir l'analyse en composantes principales (ACP). Cette combinaison nous a permis de trouver une taille adéquate pour le dictionnaire appris et offrant des performances optimales. Une comparaison avec les méthodes les plus efficaces de l'état-de-l'art nous a permis de démontrer la pertinence de notre approche.

Le troisième chapitre porte sur l'étape de la segmentation des masses et de l'extraction des spicules. Dans un premier temps, nous exposons les méthodes de segmentation les plus utilisées en imagerie du sein après une classification de ces approches en techniques supervisées et non-supervisées. Cela nous permet de déboucher sur les avantages et les limites de chacune des méthodes. L'accent est ensuite mis sur les méthodes de segmentation utilisant des formulations variationnelles et les difficultés rencontrées dans l'optimisation de celles-ci en raison du fait qu'elles convergent vers un minimum local très éloigné de l'optimum global sont évoquées. Dans ce contexte, les approches d'optimisation par coupure de graphes ont été abordées. Nous avons par la suite évoqué les méthodes de segmentation par champs de Markov (MRF) en raison de leur aptitude à prendre en compte les interactions locales entre les pixels dans un voisinage donné. Une étude minutieuse des techniques de segmentation basées MRF permet de constater que la quasi totalité de ces méthodes est supervisée – *c-à-d.* nécessite une interaction humaine ou la prise en compte d'information a priori –, ce qui engendre un temps de calcul exponentiel et la spécification de plusieurs hyper-paramètres. Un exposé des différences entre les approches supervisées et celles non-supervisées est ensuite fait et il montre que les méthodes supervisées nécessitent une base d'apprentissage contenant la vérité terrain, ce qui ne nous convient pas en raison du manque de données labélisées. Ceci nous conduit à la conception d'une méthode de segmentation non-supervisée basée sur les champs de Markov qui est capable de donner de bonnes performances en temps réel tout en ne demandant la spécification que d'un seul paramètre. L'approche proposée est fondée sur l'utilisation d'un champ de Markov bien particulier à savoir les champs de Pickard et exploite une méthode d'optimisation basée sur l'utilisation des chaînes de Markov télégraphique et d'une estimation des paramètres par le biais de l'algorithme d'Espérance-Maximisation (EM). Les performances obtenues démontrent la supériorité de notre approche comparée à toutes les méthodes de segmentation proposées dans l'état-de-l'art et ce sur les aspects suivants : temps de calcul, robustesse, nombre d'hyper-paramètres à fournir.

Toujours dans ce troisième chapitre, nous avons abordé l'extraction ou la segmentation des spicules. La présentation des méthodes de détection des masses spiculées a permis de se rendre compte que la plupart des approches développées sont fondées à la fois sur la détection de centre dense et de calcul de descripteurs pour l'extraction des spicules ou des masses spiculées. Les performances obtenues varient à la fois en fonction des descripteurs utilisés et du jeu de données ayant servi à la validation de l'approche. De plus, il a été montré qu'en dehors de très peu de travaux ([Kobatake et Yoshinaga, 1996], [Zwiggelaar et al., 1999] [Muralidhar et al., 2010]) les méthodes existantes n'ont jamais tenté de segmenter les structures anatomiques afin de s'en servir pour justifier la nature maligne ou bénigne de la masse. Etant donné l'importance que revêt la segmentation des spicules aussi bien pour la distinction entre les masses malignes et bénignes que pour la caractérisation des distorsions architecturales, nous avons proposé une méthode originale s'inspirant des techniques utilisées en télédétection, pour extraire convenablement les spicules. L'approche conçue repose sur la discrétisation des structures curvilignes de l'image à l'aide

d'une transformée locale de Radon ; ensuite à l'aide du champ de Markov prenant en compte des contraintes géométriques sur la nature des structures anatomiques, nous recomposons les structures curvilignes de l'image de départ. Pour finir, l'utilisation d'un modèle a contrario nous permet de retenir uniquement les structures curvilignes qui sont des spicules. Ce qui permet à la fois de caractériser les masses spiculées et les distorsions architecturales. Une étude comparative avec les méthodes utilisées dans l'état-de-l'art permet de prouver la qualité de notre approche.

Au quatrième chapitre, nous évoquons l'extraction non-supervisée des descripteurs à l'aide d'une approche issue du *Deep learning* et la construction d'un modèle d'aide à la décision soit sur la forme ou sur la densité basé sur les descripteurs extraits. L'approche proposée se distingue de celles utilisées dans l'état-de-l'art ; en effet, historiquement, afin de classer les masses comme bénignes ou malignes, les chercheurs en vision par ordinateur se sont consacrés à la conception des descripteurs de forme, de contour, de densité et de texture et à leur classification afin de juger de la nature de la masse. Les performances obtenues se sont révélées intéressantes et n'ont cessées d'être améliorées. Toutefois, les systèmes conçus sur de telles approches souffrent de plusieurs limitations qui justifient leur absence en milieu médical. Parmi ces limitations, nous pouvons évoquer le fait que les performances obtenues varient en fonction des descripteurs utilisés et du jeu de données ayant servi pour la validation. De plus, la combinaison de certains descripteurs peut dégrader les performances de ces systèmes, ce qui oblige à sélectionner de façon minutieuse les descripteurs les plus pertinents. Pour y rémédier, nous proposons une approche fondée sur l'extraction non-supervisée des descripteurs combinée au classifieur SVM pour apprendre un modèle efficace à la fois pour la forme et pour la densité des masses.

A la fin, la conclusion générale présente un résumé des contributions apportées ainsi que des pistes portant aussi bien sur les probables solutions pour améliorer les méthodes basées sur le *Deep learning* que sur les avantages et les limites liés à la tomosynthèse – qui est une modalité qui peut s'avérer bien utile dans les campagnes de dépistage – pour aider à mieux identifier les distorsions architecturales.

Chapitre 1

Généralités sur le cancer du sein

1.1 Introduction

La compréhension de la procédure d'interprétation de l'imagerie mammaire par les radiologues est cruciale dans le processus d'identification des verrous qui semblent représenter un frein à un dépistage efficace du cancer du sein par les experts. Cela passe par une étude de l'aspect médical du cancer du sein, dont le but est de nous aider à nous approprier quelque peu l'expertise développée en imagerie du sein par les radiologues. Ainsi, dans l'idée d'aider les radiologues à améliorer la précision de leur diagnostic tout en réduisant le nombre de biopsies inutiles, les chercheurs en vision par ordinateur ont proposé des systèmes de diagnostic assisté par ordinateur dont le but est d'identifier les lésions mammaires et de les classifier de façon automatique. Malgré la réelle difficulté posée par la qualité des images, la conception de ces systèmes ne cesse de connaître de grandes avancées depuis plus d'une vingtaine d'années. Notamment, en raison de leur capacité à fournir un résultat objectif et reproductible, donc fiable ; ce qui n'est toujours pas le cas des radiologues.

Dans ce chapitre, nous présenterons brièvement l'anatomie du sein chez la femme, ce qui va nous permettre d'aborder la notion du cancer du sein et son impact sur la population féminine. Ensuite, nous exposerons la procédure du dépistage du cancer du sein, puis nous continuerons avec les principes généraux d'acquisition en imagerie du sein. A ces éléments viendront s'ajouter l'analyse des caractéristiques des lésions mammaires, dont l'étude de la différence entre lésions malignes de stades d'évolution différents servira de socle pour introduire les systèmes de diagnostic assisté par ordinateur. Cette introduction permettra de comprendre l'articulation des diverses étapes qui composent le système de diagnostic que nous proposons et qui sera détaillé tout le long des autres chapitres. Nous finirons par une présentation des bases d'images publiques qui nous permettent de concevoir notre système.

1.2 Anatomie du sein chez la femme

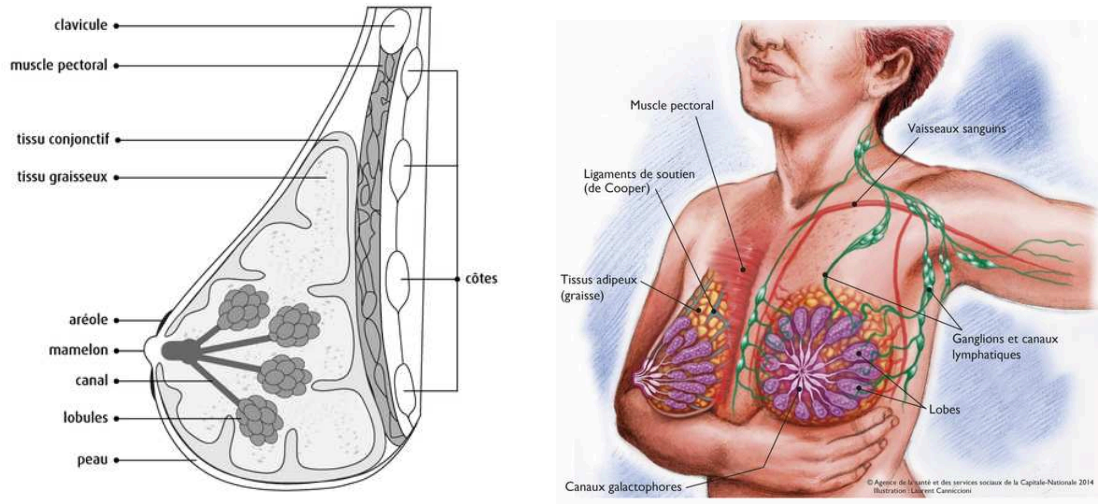
Le sein est une masse de tissu adipeux, glandulaire et conjonctif qui couvre les muscles importants de la cage thoracique, comme le muscle pectoral qui est fixé à la clavicule, le sternum et la plupart des cartilages des côtes. Le sein de la femme a pour rôle la production, le stockage et la sécrétion du lait pour l'allaitement du nourrisson.

La glande mammaire présente dans chaque sein, est une structure de densité variable adaptée à la secrètion du lait. Elle est formée d'une vingtaine de lobes ; chaque lobe étant composé de vingt à quarante lobules, elles aussi constituées à leur tour de plusieurs alvéoles. L'alvéole est une cavité arrondie en forme de cul de sac qui représente la partie sécrétrice de la glande. Chaque alvéole est drainée par un canal intralobulaire ; les alvéoles et les canaux intralobulaires forment

un lobule qui se draine par un canal interlobulaire. Plusieurs lobules se réunissent pour former un lobe glandulaire qui se draine par un canal galactophore. Le rôle des canaux galactophores est de transporter le lait vers le mamelon, qui est la région d'où le lait est expulsé. Le sein est également composé de plusieurs vaisseaux lymphatiques dont la plupart se déversent dans un réseau de ganglions lymphatiques qui sont situés sur les bords de la poitrine ou dans les tissus à proximité de l'aisselle et de la clavicule.

La taille et le poids du sein varient durant la vie d'une femme. Ils sont principalement influencés par la quantité de tissu adipeux présent dans le sein. Le développement du sein chez la femme se produit en deux étapes sous l'influence des hormones : à l'approche de la puberté, le volume du sein augmente en réponse à l'action de l'oestrogène sur les canaux galactophores, la deuxième phase survient sous l'influence de la production de progestérone quand débute l'ovulation. La croissance du sein est essentiellement terminée environ un an et demi après la fin de la période des premières menstruations, tandis que le développement des alvéoles continue en réponse à une stimulation hormonale. Il est à noter que les seins s'atrophient progressivement à la ménopause.

La figure 1.1 ci-dessous donne un aperçu de la disposition des structures anatomiques importantes du sein.



(a) (b)

FIGURE 1.1 – Anatomie du sein chez la femme : (a) structure anatomique du sein et (b) vascularisation du sein.

1.3 Le cancer du sein

Par définition, le cancer est un groupe de maladies qui se caractérisent par la prolifération et la propagation anarchiques de cellules anormales, dont l'évolution a pour conséquence le décès de la personne concernée. Il débute quand de nombreuses mutations génétiques se sont accumulées dans une cellule sans être réparées ; devenue cancéreuse, la cellule va se multiplier

de façon incontrôlée pour donner naissance à un amas de cellules appelé tumeur maligne et qui se constitue dans l'organe de départ du cancer. Au fil des multiplications successives, certaines cellules cancéreuses peuvent acquérir la capacité de quitter leur tissu d'origine par l'intermédiaire des vaisseaux sanguins ou canaux lymphatiques ; elles peuvent alors coloniser d'autres organes et y former des tumeurs secondaires appelées métastases.

Le cancer du sein quant à lui se développe généralement dans les canaux galactophores et les lobules ; il présente divers symptômes au nombre desquels on peut citer : la présence d'une boule dans le sein, l'apparition de ganglions dures au niveau de l'aisselle et diverses modifications de la peau du sein et du mamelon. Actuellement, les causes de l'apparition d'un cancer du sein ne sont pas claires ; certains facteurs de risque – élément qui augmente la probabilité d'apparition du cancer – ont néanmoins été identifiés. Il s'agit de l'âge, la prédisposition génétique liée le plus souvent à une altération des gènes BRCA1 ou BRCA2, d'un antécédent personnel de cancer du sein, ou d'un antécédent personnel d'irradiation thoracique.

D'un point de vue épidémiologique, le cancer du sein est le cancer le plus fréquent chez la femme et la première cause de mortalité par cancer en France et dans le monde. En effet, près d'une femme sur sept sera concernée par cette pathologie durant son existence, le risque augmentant avec l'âge ; de plus, des études réalisées dans le monde en 2012, font état de 522000 décès dus au cancer du sein au cours de cette année civile, ce qui représente une augmentation de 14% par rapport à 2008 [IARC, 2013]. En France, l'Institut National de Veille Sanitaire – devenue Agence Nationale de Santé Publique en 2016 – dont le rôle est de surveiller et d'observer de façon permanente l'état de santé de la population, a étudié en 2013 les tendances évolutives de l'incidence et de la mortalité liées au cancer sur la période 1980-2012 [Binder-Foucard et al., 2013]. Dans cette étude, l'incidence et la mortalité sont respectivement, le nombre de nouveaux cas de cancer du sein apparu chaque année et le nombre de décès estimé par an.

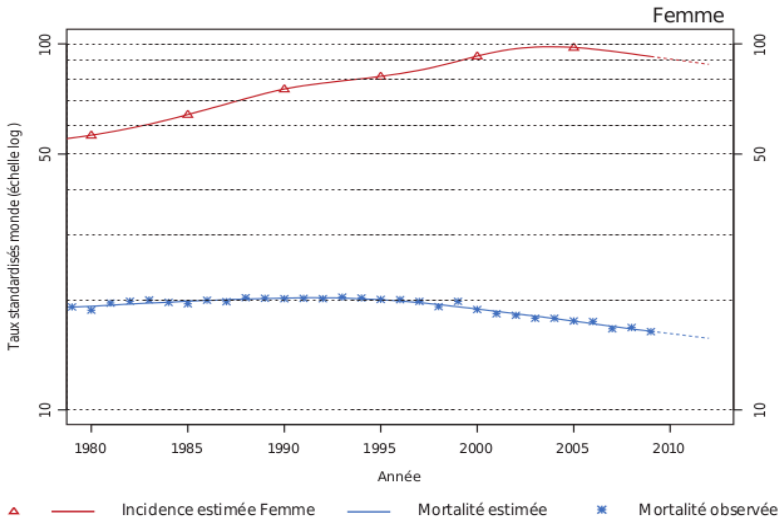


FIGURE 1.2 – Tendances chronologiques de l'incidence et de la mortalité liées au cancer du sein chez la femme de 1980 à 2012. Image tirée de [Binder-Foucard et al., 2013].

Au regard de la figure 1.2, on remarque que l'incidence de ce cancer qui a beaucoup augmenté

entre 1980 et 2000 est en diminution progressive depuis 2005. En effet, le taux d'incidence standardisé à la population mondiale a augmenté de 1,4% par an en moyenne entre 1980 et 2012, mais on note une diminution un peu plus importante de 1,5% par an si on prend en compte uniquement la période allant de 2005 à 2012. Quant à la mortalité, elle est restée pratiquement stable jusqu'en 1995 et ce en dépit de la forte augmentation de l'incidence sur la même période ; avant de connaître une chute significative en 2012. En effet, on observe une dimunition moyenne de la mortalité qui est passée de 0,6% par an entre 1980 et 2012 à 1,5% par an entre 2005 et 2012.

Regardons de plus près l'année 2012, à travers la figure 1.3 ci-dessous. On se rend compte que 22% des cancers du sein étaient observés chez les femmes de moins de 50 ans, 54% chez celles agées de 50 à 74 ans et 24% chez celles agées de plus de 74 ans. En ce qui concerne la mortalité, 8, 24 et 40% sont les chiffres notés pour les mêmes tranches d'ages.

En France, le nombre de décès liés au cancer du sein ne cesse de diminuer, à l'inverse de certaines régions du monde et ce malgré l'apparition de 48763 nouveaux cas en 2012 ; cela s'explique en partie par l'amélioration de la prise en charge thérapeutique et au diagnostic plus précoce grâce aux campagnes de dépistage.



FIGURE 1.3 – Répartition du cancer du sein chez la femme en 2012, en fonction de leur âge. Image tirée de [Binder-Foucard et al., 2013].

1.4 Dépistage du cancer du sein

Le dépistage du cancer consiste à réaliser des examens de surveillance d'une personne supposée en bonne santé, dans le but d'arriver à détecter des anomalies pouvant être des signes avant coureur d'un cancer, bien avant l'apparition des premiers symptômes. C'est une démarche qui s'avère très efficace en ce sens qu'elle permet de réduire de façon significative le taux de mortalité lié au cancer du sein en améliorant les chances de guérison. En effet, s'il est détecté tôt, le cancer du sein peut être guéri dans neuf cas sur dix, ce qui participe à une réduction de 15 à 21% du

taux de mortalité lié à ce cancer. En France, au dépistage *individuel*, s'est ajouté le dépistage *organisé*, depuis 2004. Les campagnes de dépistage *organisé* sont gratuitement destinées aux femmes ayant un âge compris entre 50 et 74 ans et qui ne présentent ni symptômes, ni facteur de risque autre que leur âge. Le dépistage du cancer du sein est caractérisé par trois phases à savoir : un interrogatoire suivi d'un examen clinique et un examen radiologique complété éventuellement par un examen biologique. Le but de l'examen radiologique est de vérifier la présence d'une anomalie radiologique afin de déterminer la conduite adéquate à tenir.

Malgré l'existence de plusieurs modalités d'imagerie médicale, la mammographie, qui est une radiographie des seins, est l'examen de référence utilisé lors des campagnes de dépistage du cancer du sein en raison de sa capacité à détecter les cancers de petite taille bien avant que les femmes ne soient en mesure de les palper ou que les premiers symptômes ne se manifestent. Ainsi, elle permet de détecter, plus de 90% des cancers. Dans les programmes de dépistage, il est réalisé sur chaque femme quatre mammographies, à raison d'une mammographie de face et de profil pour chaque sein. Bien que cela ne soit pas systématique, il peut arriver que le radiologue demande des examens complémentaires tels que l'échographie ou l'imagerie par résonance magnétique lorsqu'il est en présence de seins denses rendant d'éventuelles lésions plus difficiles à détecter. Si après l'examen des mammographies, le radiologique déclare que la mammographie est normale, les clichés de ces mammographies sont envoyés en seconde lecture, pour qu'un autre expert puisse se prononcer sur les mêmes images, afin de ne pas manquer un signe radiologique suspect.

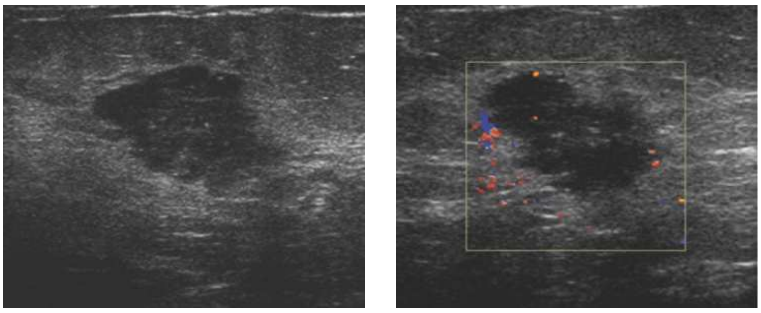
En dehors des images de dépistage, la mammographie diagnostique est réalisée, lorqu'une femme ou son médecin a remarqué au cours de l'examen du sein une anomalie parmi celles énumérées plus haut. La mammographie diagnostique peut être unilatérale ou bilatérale selon les signes observés ; son but est de rechercher un signe radiologique dans une zone suspecte ou d'analyser de façon plus précise une anomalie détectée cliniquement. Contrairement à la mammographie de dépistage qui se fait à intervalle de deux ans pour chacune des femmes se trouvant entre 50 et 74 ans, il n'y a pas d'intervalle de temps requis pour la mammographie diagnostique. Une mammographie de contrôle peut donc être réalisée quatre à six mois après la mammographie diagnostique en fonction de l'anomalie observée.

1.5 L'imagerie médicale et le cancer du sein

En France et dans plusieurs pays du monde, la démarche diagnostique ou dépistage liée au cancer du sein est caractérisée par trois phases à savoir : un interrogatoire suivi d'un examen clinique, un examen radiologique complété éventuellement par un examen biologique. Le but de l'examen radiologique est de vérifier la présence d'une anomalie radiologique afin de déterminer la conduite adéquate à tenir. Dans le cadre des campagnes de dépistage individuel ou collectif, une ou plusieurs modalités d'imagerie médicale peuvent être utilisées afin de compléter l'examen clinique. La mammographie à rayon X est l'examen de référence systématiquement proposé aux femmes pour le dépistage, suivie au besoin de l'échographie, de l'imagerie par résonance magnétique ou de la tomosynthèse afin de caractériser plus précisément la nature d'une lésion repérée à la mammographie. Dans la suite, nous allons décrire le principe de fonctionnement de chacune de ces modalités, tout en mettant l'accent sur la technique d'acquisition de l'image mammaire.

1.5.1 L'échographie

L'échographie est une technique d'imagerie utilisant des ondes de fréquences élevées (supérieur à 20 KHz) appelées ultrasons pour visualiser les tissus. Les ultrasons sont émis et reçus à l'aide d'une sonde de fréquence variable que le radiologue déplace sur la peau recouverte de gel de l'organe à explorer. Le rôle du gel est de permettre un bon contact entre la sonde et la peau afin d'assurer une bonne propagation des ondes. A chaque fois que l'onde ultrasonore propagée dans le corps rencontre des interfaces de tissu d'impédances différentes, elle est réfléchie et retournée à la sonde sous forme de signal, lequel signal sera analysé et restitué en temps réel sur l'écran d'un ordinateur. L'échographie est de plus en plus associée au doppler-couleur, afin d'obtenir une meilleure caractérisation des structures vasculaires et des lésions. Sur la figure 1.4 ci-dessous, on présente deux images échographiques, toutes deux du sein d'une femme.



(a) (b)

FIGURE 1.4 – Echographie mammaire d'une lésion maligne : (a) image en niveau de gris et (b) échographie doppler-couleur montrant une masse maligne au contour spiculé.

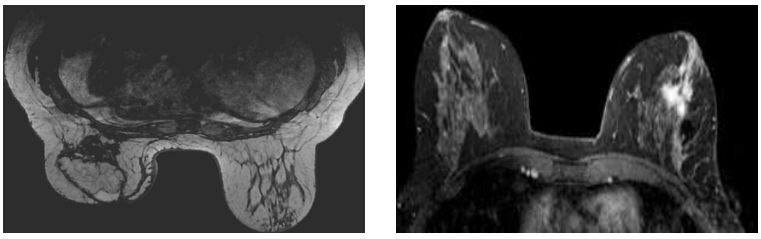
En médecine, l'échographie est un examen de première intention associé à la radiographie car elle est peu coûteuse et non invasive. Ce qui présente un réel intérêt, car elle est sans danger pour les femmes enceintes, pour qui les rayons X sont contre-indiqués. L'échographie mammaire est un examen qui vient en complément de la mammographie et est très utile pour évaluer la nature liquide ou solide des masses palpées ou découvertes sur les mammographies. A cela s'ajoute l'aspect temps-réel de l'échographie qui permet à la fois de comparer ce que l'on sent avec les doigts au cours de la palpation à ce que l'on observe sur l'écran et de guider le radiologue lorsqu'il réalise un prélèvement. De plus, l'échographie est indispensable lorsque la densité des seins ne permet pas d'avoir une mammographie de bonne qualité. Toutefois, elle présente plusieurs inconvénients qui font qu'elle demeure une technique complémentaire : la qualité des échographies est altérée par un bruit appelé *speckle* qui est le résultat d'une interférence entre les ondes, ce qui rend difficile la lecture des échographies, sans compter le fait que les microcalcifications sont difficiles à détecter à l'échographie en raison de leur petite taille ; pour finir, l'échographie est un examen dépendant de l'opérateur, ce qui signifie que le nombre, la qualité et l'interprétation des coupes acquises durant cet examen dépendent de la dextérité et de l'expérience du radiologue, d'où les résultats peuvent varier d'un radiologue à un autre.

1.5.2 L'imagerie par résonance magnétique

L'imagerie par résonance magnétique (IRM) est une technique d'imagerie médicale qui exploite les propriétés magnétiques des atomes placés dans un champ magnétique. En effet, chaque atome dispose de propriétés magnétiques intrinsèques caractérisées par un vecteur de magnétisation qui a une orientation aléatoire dans l'espace. Lorsque les atomes sont placés dans un champ magnétique élevé, ils ont une orientation parallèle à ce dernier et se maintiennent ainsi dans une position d'équilibre. Lorsque les atomes sont excités par une onde de radiofréquence à une certaine fréquence – qui varie en fonction du type d'atome et du champ magnétique –, ils basculent leur vecteur de magnétisation d'un certain angle, c'est le phénomène de résonance. A l'arrêt de l'onde de radiofréquence, le vecteur revient à sa position d'équilibre tout en émettant un signal qui est capté par les antennes réceptrices. C'est ce phénomène qui permet l'exploration du corps humain à travers l'acquisition des images.

Dans le corps humain, c'est le proton H^+ qui est soumis à l'excitation et à la relaxation, en raison de la présence de l'eau – dont l'hydrogène est le constituant majeur – en quantité plus ou moins importante dans l'ensemble des tissus humain. Les signaux émis par ces protons sont alors recueillis par les antennes présentes dans l'appareil d'IRM et traités numériquement dans le but de reconstruire une image en coupes. L'IRM est donc une technique qui permet l'acquisition d'un volume d'image 2D.

Plus particulièrement pour l'acquisition d'une image du sein par IRM, on procède par injection d'un produit de contraste appelé gadolinium, dont le but est de permettre la détection d'un petit cancer qui a plus de vaisseaux que le reste du sein et ce à travers l'enrégistrement de la circulation du produit de contraste dans le sein. La figure 1.5 ci-dessous donne un aperçu du résultat obtenu après un IRM avec ou sans produit de contraste. L'IRM mammaire est parfois



(a) (b)

FIGURE 1.5 – IRM mammaire : (a) image du sein sans injection de produit de contraste et (b) image du sein avec injection de gadolinium après 45 minutes.

employée pour faire la différence entre une lésion bénigne et une lésion maligne au cours du dépistage des femmes à très haut risque de cancer du sein. Elle peut être utilisée lorsque l'imagerie standard n'offre aucune certitude sur l'absence de malignité. Cependant, l'IRM est très onéreuse et présente également des contre-indications à savoir les pace-maker, les corps métalliques dans l'oeil, les implants cochléaires, etc $\cdots$ qui en font un examen risqué pour la plupart des femmes concernées par le programme de dépistage collectif.

1.5.3 La mammographie

1.5.3.1 Le principe

La mammographie est l'examen radiologique du sein de la femme à l'aide des rayons X, dont l'objectif est de détecter les lésions de petites tailles, le plus souvent trop petites pour être palpées. Elle est à ce titre, l'examen de référence utilisé à travers le monde, pour le dépistage du cancer du sein, mais aussi pour le diagnostic de certains symptômes touchant le sein comme les douleurs, la modification de la peau, l'inflammation et l'écoulement du mamelon. Par ailleurs, elle permet une exploration totale du tissu mammaire en deux incidences.

Le mammographe est l'appareil utilisé pour l'acquisition des mammographies. Une illustration d'un mammographe est présentée à la figure 1.6 ci-dessous. Le mammographe dispose d'un tube radiogène – qui génère des rayons X de faible énergie, *c-à-d.* entre 20 et 50 keV –, d'un compresseur et d'une cassette. Dans un premier temps, les deux seins sont comprimés à tour de rôle afin d'étaler le tissu mammaire, ce qui offre une meilleure visualisation des tissus du sein et une réduction de la dose de rayons X délivrée. Ensuite, le manipulateur radio procède à l'exposition du sein aux rayons X ; ce qui aboutit à la projection du sein sur le film contenu dans la cassette ou une acquisition digitale de grande qualité par des capteurs le cas échéant. Les plages de densité visibles sur l'image et le contraste entre les structures anatomiques sont obtenues grâce aux différences de l'atténuation des différents types de tissu présents dans le sein.

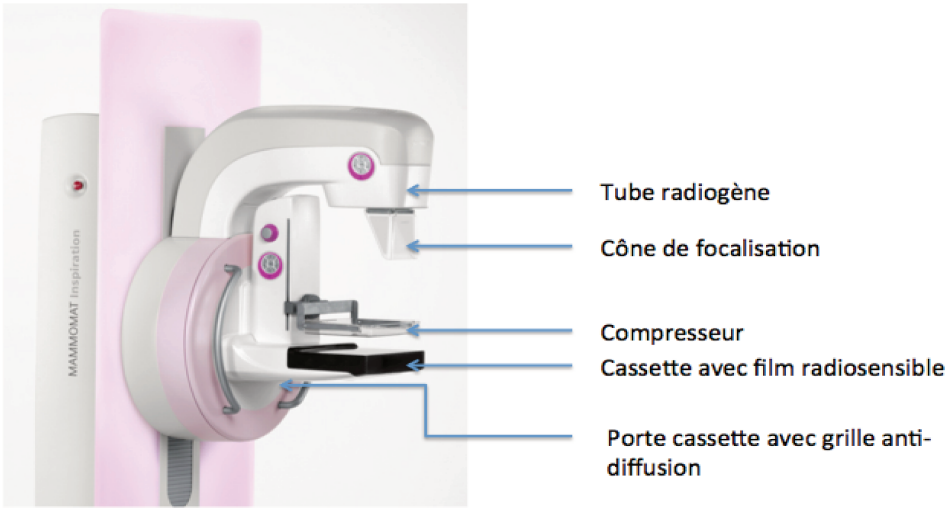


FIGURE 1.6 – La compostion d'un mammographe.

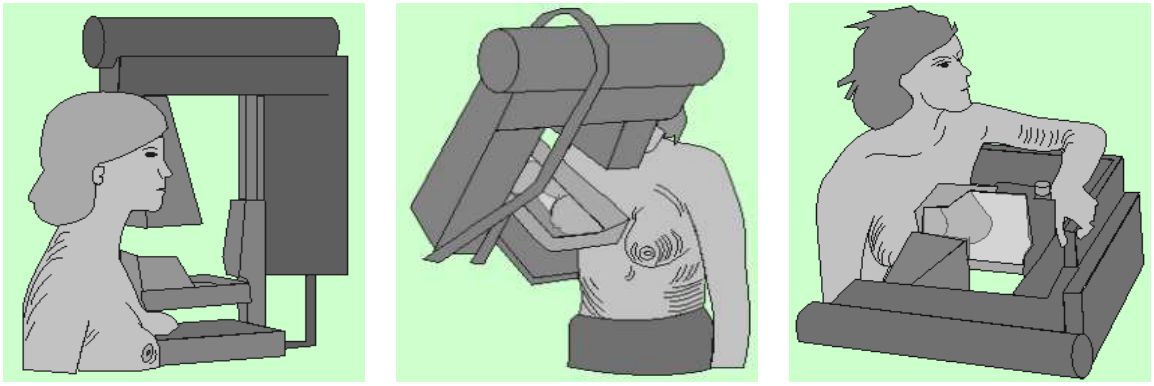
1.5.3.2 Relation entre l'anatomie et les images mammographiques

L'image obtenue après la mammographie est la conséquence de l'atténuation des faisceaux de rayons X ayant traversés les différents tissus mammaires au cours de l'examen. L'atténuation du faisceau de rayons X se fait en fonction de la nature du tissu traversé. Au nombre de ces tissus, on peut citer la graisse, qui est considérée comme une zone radio transparente étant donné qu'elle

possède une densité physique très légère. La conséquence, c'est qu'elle présente une apparence sombre sur une image mammographique. Par contre, le tissu fibroglandulaire et le calcium – qui est un composant principal des lésions mammaires – sont des éléments radio opaques qui apparaissent claires sur un cliché de mammographie. Pour ce qui est de l'eau, c'est une matière radio opaque qui apparait claire sur l'image. A la lumière des informations relatives à l'anatomie et à la radio transparence, on peut affirmer que l'aspect général d'une mammographie est sombre tandis que les zones contenant des microcalcifications ou les masses sont plus claires.

1.5.3.3 Les incidences en mammographie

En raison de la complexité de l'anatomie du sein, la mammographie est d'habitude réalisée sous différents angles appelés incidences. L'objectif d'une bonne incidence est de visualiser la majeure partie du tissu mammaire en l'étalant le plus possible sur la plaque radiographique. En fonction de la partie du sein qu'on désire observer, diverses incidences sont observées. Les incidences les plus utilisées dans la pratique sont : celle de face ou la cranio caudale (CC), l'incidence médio latérale oblique (MLO) (encore appelée oblique externe) et l'incidence de profil. Chacune de ces incidences permet d'explorer une région bien précise : l'incidence CC – dont la réalisation implique que le tube radiogène et le détecteur soient horizontaux – offre une bonne exploration des régions centrale et interne du sein ; pour ce qui est de la MLO, elle requiert une inclinaison de 40 à 60° du détecteur – en fonction de la morphologie de la patiente de façon à ce que l'orientation des fibres du grand pectoral soit parrallèle à ce dernier –, ce qui permet d'explorer la totalité de la glande mammaire sur une seule incidence. Quant à l'incidence de profil, sa réalisation nécessite que le détecteur soit en position verticale et disposé entre les deux seins. Cette incidence, à l'image de celle de face (CC) explore la région centrale et interne du sein. La figure 1.7 ci-dessous illustre le positionnement du tube radiogène et du détecteur pour les différentes incidences.



(a) (b) (c)

FIGURE 1.7 – Illustrations des incidences principales en mammographie : (a) incidence cranio caudale (CC), (b) incidence médio latérale oblique (MLO) et (c) incidence de profil. Image tirée de [Heron, 2009].

1.5.4 La tomosynthèse du sein

La tomosynthèse du sein ou *Digital Breast Tomosynthesis* (DBT) est une nouvelle modalité d'imagerie qui a le potentiel de pallier plusieurs limites de la mammographie conventionnelle en produisant une représentation 3D du sein. Après une synthèse des limitations de la mammographie 2D et des autres techniques d'acquisition 3D, nous évoquerons le principe général de la tomosynthèse avant de finir par une description de ses avantages et inconvénients.

1.5.4.1 Motivation

La mammographie conventionnelle est un examen de projection durant lequel une structure 3D en l'occurence le sein de la femme est projetée sur une surface plane – le détecteur –, résultant en une perte potentielle d'information. Dans ce cas de figure, deux problèmes majeurs peuvent être rencontrés. Le premier est la non détection d'une lésion qui est dissimulée par les tissus qui se sont retrouvés sur la trajectoire du rayon passant par cette dernière. Cette situation est naturellement très inquiétante, car ne pas détecter une lésion implique de ne pas la traiter au plus tôt. La seconde difficulté qu'on peut rencontrer est la détection d'un signe radiologique issu de l'accumulation de diverses structures anatomiques normales et indépendantes se retrouvant sur une trajectoire de rayons X, et formant une fois projetées, une structure laissant penser à la présence d'une lésion. Bien que ce type de fausse alerte ait des conséquences moins graves que la non détection, le fait de devoir subir des examens complémentaires et dans certains cas une biopsie, pour infirmer la présence d'un cancer, est une situation stressante pour la patiente, qu'il est préférable d'éviter.

Les problèmes rencontrés sont donc majoritairement causés par la projection de l'information durant la formation d'une image de mammographie. La solution la plus adéquate pour régler ces problèmes est alors de travailler sur une représentation tridimensionnelle du sein. Les méthodes de formation d'image 3D sont régulièrement utilisées en médecine, comme le scanner par tomographie ; mais son utilisation pour le sein pourrait engendrer des problèmes liés à la dose de rayons X délivrée à la patiente [John et Ewen, 1989]. En effet, pour être capable d'acquérir une image de cette région du corps humain, il faut que les rayons X soient capables de traverser cette région de part et d'autre, ce qui nécessite une dose beaucoup plus importante que celle utilisée pour une mammographie conventionnelle [Sibala et al., 1981]. Par ailleurs, il existe d'autres types d'acquisition s'appuyant sur d'autres modalités ; c'est le cas par exemple de l'IRM du sein, qui rappelons le est employée dans le cadre du dépistage. Toutefois, pour des raisons de coût et de disponibilité, elle est souvent réservée aux patients présentant un risque certain [Saslow et al., 2007].

Parmi toutes les méthodes permettant d'obtenir une représentation 3D de la structure du sein, la tomosynthèse semble donc être en mesure de jouer un rôle important à la fois dans le cadre du diagnostic et du dépistage, en se positionnant comme une technique complémentaire crédible à la mammographie.

1.5.4.2 Principe général

Le principe de la tomosynthèse est de reconstruire l'information tridimensionnelle d'un objet à partir de plusieurs radiographies à faible dose de radiation acquises sous différents

angles [Dobbins III et Godfrey, 2003]. Comme on peut le remarquer sur la figure 1.8, en prenant différents clichés à des positions différentes d'un même objet, la localisation relative des structures qui le constituent, une fois projetées sur le détecteur change. Pour disposer d'une information 3D de l'objet dont on vient de réaliser l'acquisition, ces différentes projections sont passées en entrée d'un algorithme de reconstruction ; la phase de reconstruction fournit un ensemble de coupes parallèles au détecteur à différentes profondeur par rapport à ce dernier, qui représentent l'atténuation radiologique des structures qui composent l'objet ou l'organe imagé.

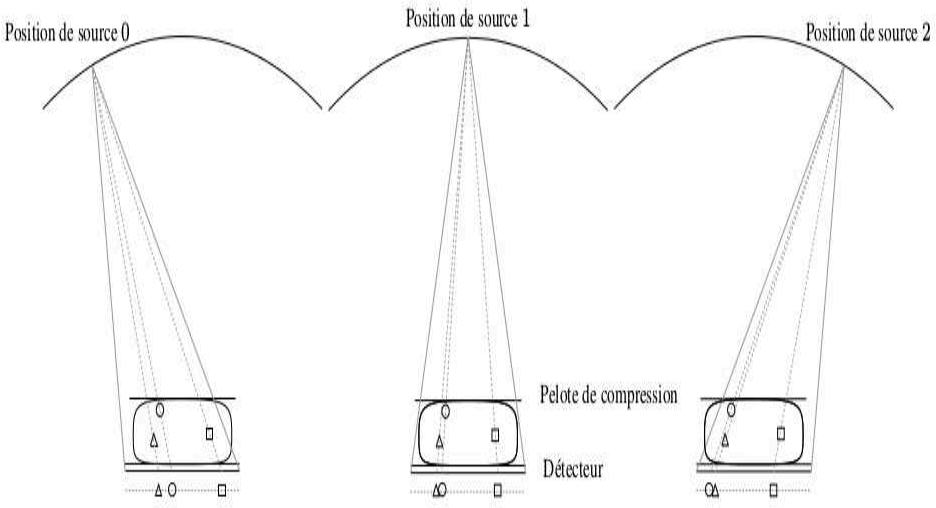


FIGURE 1.8 – Illustration du principe de la tomosynthèse

Globalement, l'ossature des systèmes de tomosynthèse du sein ressemble beaucoup à celle des appareils de mammographie. Le sein de la patiente est placé en contact avec le support patient et comprimé à l'aide d'une pelote de compression. La position du tube évolue selon une trajectoire souvent circulaire, dont le centre de rotation se trouve au niveau du détecteur. Puis, une série de projections 2D à faible dose est acquise sous différentes orientations du tube à rayons X autour du sein. Le nombre de projections peut varier entre 9 et 45, pour une ouverture angulaire de 30 à 90 degrés. En terme de radiation, la dose allouée à chaque projection est la même ou deux fois la dose délivrée pendant une mammographie standard *c-à-d.* l'équivalent de la dose totale de radiation délivrée après la réalisation des incidences CC et MLO ; c'est ce qui explique le fait que les images passées en entrée des algorithmes de reconstruction soient de qualité médiocre en comparaison à celles obtenues en mammographie conventionnelle. Mais après l'étape de reconstruction, on aboutit à des images de bien meilleure qualité.

1.5.4.3 Les limites et avantages de la tomosynthèse

L'avantage fondamental de la DBT, c'est qu'elle permet au radiologue de mieux visualiser le parenchyme mammaire en s'affranchissant des superpositions inhérentes à la mammographie 2D ; ce qui rend les signes radiologiques associés à une lésion suspecte (les distorsions architecturales, par exemple) plus faciles à discerner [Gennaro et al., 2008]. La conséquence est une diminution du taux d'images douteuses et la réduction des clichés supplémentaires ; à cela s'ajoute une

rapidité à préciser le caractère malin ou bénin d'une lésion suspecte et un meilleur taux de détection particulièrement pour les seins denses.

Toutefois, la création de volume de tomosynthèse peut être obtenue de plusieurs manières. En effet, l'algorithme de reconstruction joue un rôle important sur l'aspect des données reconstruites ; ainsi, le type de filtre utilisé dans les approches de retroprojection filtrée ou le nombre d'itérations de même que la technique de mise à jour dans les méthodes itératives ont une influence sur le bruit et le flou contenu dans les coupes reconstruites [Zhang et al., 2006]. Une autre limite est la présence d'artéfacts de reconstruction – qui résultent de l'impossibilité de faire le tour de la patiente durant l'acquisition des images en tomosynthèse – ce qui engendre un manque d'information utile pour une reconstruction parfaite de l'image. De façon plus concrète, cela se manisfeste par un étalement ponctuel ayant l'allure d'une étoile allongée dans la direction de la hauteur [Blessing et al., 2006], l'inconvénient c'est que les objets apparaissent déformés dans la même direction.

Pour finir, le pas d'échantillonnage qui peut être important en fonction de la hauteur des coupes reconstruites est une limite non négligeable qui peut aboutir au fait que les microcalcifications se retrouvent entre deux coupes reconstruites. Dans cette situation, la visibilité des microcalcifications peut être affectée ; il a toutefois été montré qu'une ouverture angulaire moins importante pouvait atténuer ce problème [Lau et al., 2008]. Cela découle du fait qu'une faible ouverture angulaire entraine une augmentation de l'étalement des microcalcifications selon l'épaisseur des coupes, favorisant donc une éventuelle intersection avec les coupes reconstruites. Il faut aussi noter que la tomosynthèse émet plus de radiations qu'une mammographie conventionnelle sans toutefois dépasser les normes recommandées ; sans oublier le temps de lecture plus long pour les radiologues.

En dépit de ce qu'elle présente comme avantages, la tomosynthèse est encore une technologie prometteuse non utilisable à grande échelle. Néanmoins, elle devrait pouvoir être introduite en milieu médical d'ici cinq ans en France, surtout pour pallier le faux diagnostic concernant les distorsions architecturales. En attendant que son utilisation en milieu médical devienne une réalité, la tomosynthèse fait l'objet de plusieurs travaux de recherche et ce au détriment de la mammographie conventionnelle. Les approches que nous avons proposé dans cette thèse restent néanmoins applicables à la tomosynthèse en raison de ce que les difficultés rencontrées avec cette modalité sont également les mêmes qu'en mammographie 2D.

1.6 Description des masses

En imagerie du sein, on distingue deux grands types de lésions ; on a d'un côté les microcalcifications et de l'autre les masses. Etant donné que notre sujet de thèse concerne les masses mammaires, nous ne ferons ici que la description de ces dernières. Par définition, une masse est une opacité importante occupant un espace et vue sur deux incidences différentes. Dans le cas où, la masse est observée uniquement sur une seule incidence, elle est appelée asymétrie jusqu'à ce que sa nature tridimensionnelle soit confirmée. Différents attributs permettent aux médecins de décrire les masses en vue de déterminer leur nature ; ce sont : la forme, le contour et la densité.

1.6.1 La forme

D'après la classification BI-RADS de l'ACR [ACR et al., 2003], les opacités mammaires peuvent avoir une forme ronde, ovale, lobulée ou irrégulière.

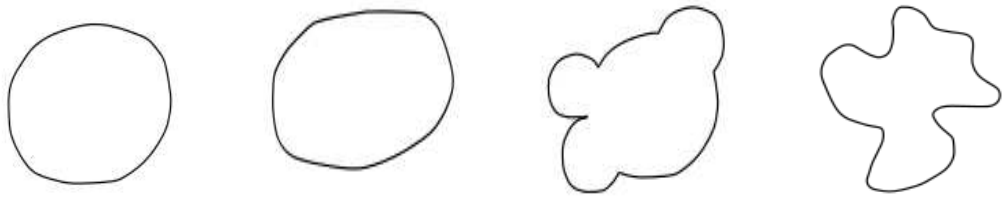
(a) **ronde** : c'est l'aspect d'une masse sphérique, circulaire ou globuleuse,

(b) **ovale** : de telles masses présentent une forme d'ellipse,

(c) **lobulée** : la forme de la masse est légèrement ondulée,

(d) **irrégulière** : c'est le terme utilisé pour désigner les masses qui présentent une forme aléatoire et ne peuvent être classées dans les trois classes précédentes.

Ces différentes formes sont illustrées de manière schématique à la figure 1.9.



(a) (b) (c) (d)

FIGURE 1.9 – Les formes des opacités selon le BI-RADS : (a) ronde, (b) ovale, (c) lobulaire et (d) irrégulière.

1.6.2 Le contour

En dehors de leur forme, les opacités varient également au niveau de leur contour ; le BI-RADS dénombre cinq classes de contour que sont les contours : circonsrit, masqué, micro-lobulé, mal-défini et spiculé. La figure 1.10 présente un exemple de chacune de ces cinq catégories de contour.

(a) **circonscrit** : ce sont des contours bien définis où la frontière entre la masse et le fond de l'image est nette et qui correspondent en général à des lésions bénignes,

(b) **masqué** : ce sont pour la plupart les masses dont le contour est dissimulé par la superposition de tissus normaux, faisant penser à une masse circonscrite dont une partie du contour est cachée,

(c) **micro-lobulé** : c'est un contour qui contient de petites ondulations,

(d) **mal-défini** : c'est un contour qui peut laisser envisager la présence d'une infiltration,

(e) **spiculé** : c'est un contour comportant des structures filiformes qui rayonnent en s'éloignant du centre de la masse et qui suggèrent fortement une malignité.

Hormis les microcalcifications et les masses, il existe une troisième forme de signes radiologiques qui traduisent la présence d'un cancer et qui sont appelées les distorsions architecturales. Elles ne comportent pas de centre dense, comme c'est le cas des opacités décrites plus haut. Les distorsions architecturales sont des signes hautement suspects qui se traduisent par des structures linéaires ou des spicules qui convergent vers une même zone focale. Un exemple schématique illustrant les distorsions architecturales est présenté à la figure 1.14 ci-dessous.



(a) (b) (c) (d) e)

FIGURE 1.10 – Les différents types de contour des opacités selon le BI-RADS : (a) contour circonsrit, (b) contour masqué, (c) contour micro-lobulé, (d) contour mal défini et (e) contour spiculé.

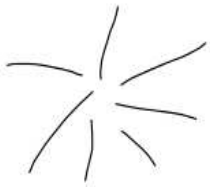


FIGURE 1.11 – Exemple de la forme d'une distorsion architecturale

1.6.3 La densité

La composition du sein varie souvent d'une femme à une autre ; l'un des aspects visibles de cette grande variabilité est la densité radiologique du tissu mammaire. Des études récentes ont suggéré que la densité mammaire est un facteur de risque important, en raison de ce que les pathologies peuvent être dissimulées à l'intérieur d'une région très dense [Eccles et al., 2013]. Dans le but d'aider à une meilleure compréhension de la composition de la densité du sein, plusieurs travaux et modèles ont été proposés. [Wolfe, 1976] fut l'un des premiers à aborder le sujet ; il a classé de façon empirique la densité mammaire en quatre catégories – allant du sein graisseux au sein fibroglandulaire dense – ; ces travaux n'ont cependant pas fait l'unanimité dans la communauté scientifique en raison du fait que les mammographies utilisées dans cette étude sont issues du premier essai du programme de dépistage aléatoire et contrôlé [Sickles, 2007] et des conditions d'acquisition limitées des appareils de mammographie (*c-à-d.* positionnement et compression du sein) du début des années 70 [Bailar, 1976]. Fortement influencés par les travaux de Wolfe *et al.*, [Tabár et Dean, 1982], ont proposé un modèle basé sur le mélange de quatre *briques* qui constituent l'anatomie normale du sein ; ce qui leur a permis par la suite de subdiviser les images en cinq catégories de risque en tenant compte de la distribution des briques du sein [Tabár et al., 2005]. L'avantage du modèle de Tabar *et al.*, c'est qu'il est basé sur les corrélations de l'anatomie mammaire plutôt que sur une simple lecture des mammographies, comme l'a fait Wolfe. De plus, le modèle d'évaluation du risque mammaire en fonction de la densité que Tabar *et al.* ont proposé est plus discriminant que celui de Wolfe, permettant ainsi une précision reproductible de la classification. Quant à [Boyd et al., 1995], ils ont conçu une méthode utilisant un outil de détection assistée par ordinateur pour estimer le pourcentage de

région dense dans le sein ; leur travaux ont servi à infirmer certaines conclusions de [Wolfe, 1976] à propos du risque de cancer lié aux quatre groupes proposés. Face aux différents modèles existants et dans le but de standardiser les comptes rendus des mammographies, l'ACR a mis au point une classification BI-RADS définissant quatre classes relatives à la composition du sein.

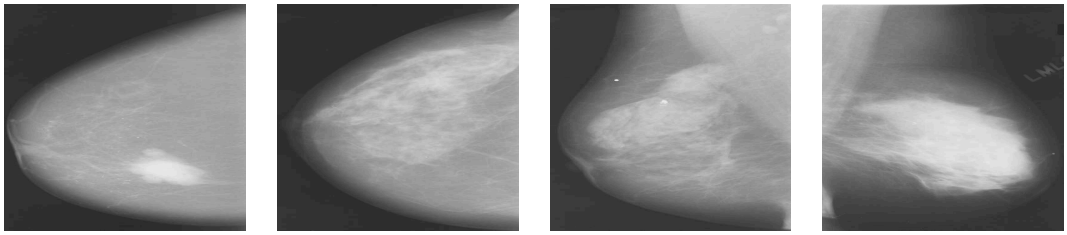
(a) **B1** : le sein est presque entièrement graisseux et homogène, radio transparent et facile à interpréter pour les radiologues (densité occupe moins de 25% de la glande mammaire),

(b) **B2** : le sein est graisseux et hétérogène, et on note la présence d'opacités fibroglandulaires dispersées dans le parenchyme (la densité occupe entre 25 et 50% de la glande mammaire),

(c) **B3** : le sein est dense et hétérogène (la densité occupe entre 50 et 75% de la glande mammaire),

(d) **B4** : le sein est hyper dense et homogène ; l'interprétation de la mammographie est donc très difficile car la densité peut dissimuler une masse (la densité occupe plus de 75% de la glande mammaire),

La figure 1.12 ci-dessous donne un aperçu de la texture des différentes classes de densité évoquées plus haut.



(a) (b) (c) (d)

FIGURE 1.12 – Types de densité des masses selon le BI-RADS : (a) B1 (densité $< 25\%$), (b) B2 (densité comprise entre 25 et 50%), (c) B3 (densité comprise entre 50 et 75%) et (d) B4 (densité $> 75\%$).

1.7 Classification des pathologies du sein

Dans l'idée de permettre aux radiologues d'avoir un canevas qui leur permet d'interpréter les mammographies et d'étudier convenablement les lésions mammaires et ce dans le but de faciliter une seconde lecture par leur collègue, plusieurs lexiques standards et des classifications ont été proposées. On peut citer celle de [Gal et al., 1984], celle de [Lanyi, 1985] et le BI-RADS [ACR et al., 2003]. A ce jour, les classifications les plus répandues et les plus utilisées par les radiologues à travers le monde, sont celles de Le Gal et le BI-RADS. Dans cette partie, nous nous intéresserons uniquement au BI-RADS car c'est elle qui nous sert de référence pour concevoir notre système.

La classification de Le Gal proposée en 1984 avait l'avantage d'être simple, mais présentait d'importants défauts. En effet, cette classification se base uniquement sur les microcalcifications et ne prend pas en compte la morphologie des masses, les distorsions architecturales et la conduite à tenir quand l'expert se trouve en présence d'une lésion. Dans l'objectif de résoudre

les problèmes posés par la classification de Le Gal, l'ACR a conçu un guide mammographique standardisé et complet qui présente l'avantage de prendre en compte les divers cas de malignité. Ce système propose une répartion des lésions en six catégories en fonction du niveau de suspicion que semble présenter la lésion. Ainsi, le BI-RADS prend en compte à la fois la forme et le contour des différentes masses, les microcalcifications, leur texture, sans oublier les distorsions architecturales, de même que leur niveau de malignité. Dès lors que l'expert arrive à attribuer une classe à la lésion concernée, il sait la conduite à tenir grâce aux directives associées à cette classe. Tout ceci explique la popularité de cette classification, dont l'utilisation par les radiologues a permis d'améliorer le taux de détection des lésions. Les tableaux 1.1 et 1.2 ci-dessous donnent respectivement un aperçu de la description des masses associées à chacune des classes du BI-RADS et de la conduite à tenir en présence des différents cas de figure.

A travers la présentation des différentes lésions du sein dans les sections précédentes, nous avons pu constater la complexité de la détection de ces pathologies et toutes les compétences et qualité que requièrent une bonne interprétation des mammographies, à savoir l'expérience, l'attention, la présence d'esprit et la connaissance du BI-RADS par l'expert. En résumé, ce dernier doit faire preuve de subtilité, car les conclusions de son interprétation des mammographies sont cruciales pour la survie de la patiente. Face à tous ces facteurs qui ne sont pas de nature à assurer une bonne analyse des images et une décision objective et absolument fiable d'un radiologue à un autre, beaucoup d'équipes de recherche se sont lancées dans la conception de systèmes pouvant faciliter la tâche de détection des lésions et de prise de décision des radiologues.

Table 1.1 – La classification BI-RADS de l'ACR.

Classes	Description
$\mathcal{BI}-\mathcal{RADS}$ **0**	Cliché insuffisant pour vérifier la présence ou non de lésions
$\mathcal{BI}-\mathcal{RADS}$ **I**	Absence de masses suspectes
$\mathcal{BI}-\mathcal{RADS}$ **II**	– Opacité ronde avec macrocalcifications – Ganglion intra mammaire – Image de densité graisseuse ou mixte – Opacité ronde correspondant à un kyste en échographie
$\mathcal{BI}-\mathcal{RADS}$ **III**	Masse circonscrite ronde ou ovale sans microlobulations non liquidienne à l'échographie
$\mathcal{BI}-\mathcal{RADS}$ **IV**	– Masse spiculée sans centre dense – Masse ronde ou ovale à contour lobulé ou masqué
$\mathcal{BI}-\mathcal{RADS}$ **V**	– Masse spiculée à centre dense – Masse à contour flou et irrégulier

1.8 La détection ou le diagnostic assisté par ordinateur

1.8.1 Importance des systèmes de diagnostic assisté par ordinateur

Au départ, la motivation qui sous-tendait la conception des systèmes automatiques de traitement des images issues de la mammographie, était d'améliorer la qualité des images, tout en

TABLE 1.2 – Interprétation et conduite à tenir (CAT) en fonction des classes du BI-RADS.

Classes	Interprétation	Risque	CAT
$\mathcal{BI}-\mathcal{RADS}$ **0**	Investigation incomplète		Examens
$\mathcal{BI}-\mathcal{RADS}$ **I**	Image normale	0%	Dépistage
$\mathcal{BI}-\mathcal{RADS}$ **II**	Lésion bénigne	0%	Dépistage
$\mathcal{BI}-\mathcal{RADS}$ **III**	Lésion probablement bénigne	$< 2\%$	Suivi
$\mathcal{BI}-\mathcal{RADS}$ **IV**	Lésion suspecte demandant une biopsie	> 2 et $< 90\%$	Biopsie
$\mathcal{BI}-\mathcal{RADS}$ **V**	Lésion fortement suspecte d'un cancer	$> 90\%$	Biopsie

mettant à la disposition des radiologues des outils leur permettant de détecter et de localiser les lésions dans les images mammographiques, de tels systèmes sont connus sous le nom de système de détection assistée par ordinateur (DAOe). Ces systèmes se sont révélés très efficaces et sont disponibles sur la quasi totalité des mammographes modernes. Mais avec l'augmentation exponentielle de la quantité d'images à interpréter du fait des campagnes de dépistage, on a très vite constaté que l'analyse des images mammographiques par les radiologues entraine un taux d'erreur de diagnostic assez important et des examens complémentaires onéreux voire inutiles. Les chercheurs en vision par ordinateur ont donc tenté de proposer divers systèmes pouvant servir de second avis aux radiologues au cours de la phase d'interprétation. Connus sous le nom de système de diagnostic assisté par ordinateur (DAOx), ce sont des outils offrant un pipeline complet, allant de la détection des masses à leur classification comme bénigne ou maligne. Bien qu'à ce jour aucun système de DAOx ne soit commercialisé et utilisé en milieu clinique, les performances des systèmes de DAOx ne cessent de s'améliorer et sont de plus en plus en mesure de fournir une aide précieuse pour le diagnostic, notamment ce qui concerne la distinction entre les masses malignes et bénignes.

1.8.2 La détection assistée par ordinateur

De nombreuses études ont montré que 10 à 30% des cancers sont manqués par les radiologues durant l'interprétation des images [Elter et Horsch, 2009] et que ce chiffre à tendance à diminuer grâce à une seconde lecture réalisée par un autre expert, en raison de l'influence de l'expérience des radiologues sur leur diagnostic. Cependant, il n'est pas toujours possible de disposer d'un second avis dans certaines localités et ce même dans les pays développés. Dans ce cas de figure, les systèmes de DAOe peuvent guider l'expert dans son diagnostic, en se focalisant sur les régions suspectes.

Actuellement, les systèmes de DAOe dédiés à la mammographie sont présents sur les mammographes modernes et sont utilisés en mammographie diagnostique ; néanmoins, leur rôle en mammographie de dépistage reste encore flou. [Ciatto et al., 2003] ont tenté d'apporter un début de réponse à cette question à travers une étude menée dans le cadre d'un test de dépistage. Après avoir comparé le résultat de lecture d'images mammographiques par les radiologues à celui obtenu par un système de DAOe, ils sont parvenus à la conclusion selon laquelle, la performance d'une seule lecture du système de DAOe est similaire à celle de deux radiologues. Toutefois, ces résultats sont à nuancer, car les systèmes de DAOe sont loin d'être d'une grande utilité pour

les cas les plus compliqués parmi les lésions mammaires, à savoir les masses spiculées et les distorsions architecturales. En effet, [Baker et al., 2003] ont réalisé une étude visant à évaluer la sensibilité de deux systèmes de DAOe commercialisés – que sont ImageChecker ® M1000, version 2.5 (R2 *Technology*, Sunnyvale, CA) et SecondLook®, version 4.0 (*CADx Medical Systems*, Laval, QC, Canada) – et utilisés en milieu hospitalier sur la détection des distorsions architecturales dans le cadre d'un dépistage. Pour cela, ils ont pris un échantillon de 45 cas pour 80 images sur lesquelles sont présentes des distorsions architecturales. Après lecture, le R2 ImageChecker a réussi à détecter 22 des 45 cas de distorsions architecturales sur 30 des 80 images, ce qui représente respectivement une performance de 49 et 38%. Quant à SecondLook, c'est respectivement 33 et 21%. Ces résultats viennent montrer du doigt la difficulté qu'il existe à détecter certaines lésions en mammographie.

Par contre, pour ce qui concerne les microcalcifications, les performances sont toutes autres. En effet, [Krupinski et Nishikawa, 1997] ont remarqué – après avoir comparé les performances d'un système de DAOe avec celles de six radiologues pour la détection des microcalcifications sur un jeu de données de 80 images – que 10% des amas de microcalcifications ont été détectés par le système mais ont été ratés par les six experts. Néanmoins, 11% des amas ont été manqués par le système mais ont été repérés par un expert au moins. Pour finir, ils ont constaté que seulement 5% des amas ont été manqués à la fois par les radiologues et le système de DAOe. Ce qui prouve l'efficacité de ces systèmes en milieu médicale pour la détection des microcalcifications.

1.8.3 Le diagnostic assisté par ordinateur

Le diagnostic des radiologues à propos des lésions mammaires est une tâche subjective qui peut donc être entâchée d'erreur dont les conséquences peuvent s'avérer mortelles pour les patientes. Plusieurs raisons comme la fatigue, le manque d'expérience en imagerie du sein et la quantité d'images à analyser chaque jour rendent cette tâche fastidieuse et justifient les erreurs de diagnostic. A l'inverse, un système automatique de DAOx conçu sur un modèle qui prend en compte des descripteurs pertinents de la lésion mammaire fera toujours preuve de robustesse, ce qui pourrait aider à réduire le nombre de biopsies inutiles. En effet, il a été prouvé qu'après une biopsie, seulement moins du tiers des images mammographiques suspectes sont réellement des mammographies de lésions cancéreuses [Meyer et al., 1990]. Aussi, la conception des systèmes de DAOx pourrait aider les radiologues à diminuer le nombre de biopsies et éviter à la patiente une situation de stress inutile.

Un système automatique de DAOx employé dans un contexte de dépistage ou de diagnostic servira donc de second avis au radiologue et n'a absolument pas vocation à se substituer à ce dernier. De façon générale, le fonctionnement d'un système de DAOx est une suite de plusieurs étapes toutes aussi importantes les unes que les autres. Quand le système reçoit en entrée une image, il procède d'abord à un pré-traitement de celle-ci, *c-à-d.* élimination du bruit d'acquisition ou *speckle* et amélioration du contraste de l'image, afin d'obtenir la meilleure qualité d'image possible. Cette étape est suivie de la détection des anomalies (*p. ex.* microcalcifications, opacité ou distorsions architecturales). Sur l'anomalie détectée sont ensuite extraits plusieurs attributs mathématiques pouvant aider à décrire au mieux la nature de la lésion. Pour finir, les descripteurs extraits sont ensuite classifiés en s'appuyant sur des algorithmes adaptés à cette tâche, afin de décider de la nature de la masse.

1.9 Les bases de mammographies

Il existe plusieurs bases d'images utilisées par les chercheurs en imagerie du sein – dans le but de concevoir des systèmes de DAOx –, dont certaines sont publiques et d'autres sont privées. Toutefois, toutes ne respectent pas les exigences que doivent satisfaire une base dédiée à la conception d'un système de DAOx. Dans un premier temps, nous allons évoquer brièvement les exigences relatives à la création d'une base de mammographies, puis nous finirons en détaillant quelques unes des bases disponibles tout en mettant l'accent sur leur forces et leurs faiblesses.

1.9.1 Les exigences

Dans des études menées entre 1996 et 1998, [Nishikawa, 1998] a énuméré certaines exigences qui doivent être respectées dans la création de base de mammographies numériques. Certaines d'entre elles sont abordées ci-dessous.

1.9.1.1 La séléction des cas

Toute base de mammographies doit contenir divers cas avec des images de sein normal, de même que celles de toutes les pathologies rencontrées en imagerie du sein et ce quelque soit la densité mammaire. Les images de sein normal mais contenant des structures qui peuvent prêter à confusion doivent être maintenues dans la base, car elles sont importantes pour rendre les classifieurs plus robustes. Les cas doivent être collectés par des radiologues ayant une réelle expérience en mammographie ; chaque cas doit contenir quatre images obtenues à l'aide des incidences standards (MLO et CC), à moins que la femme n'ait subit auparavant une mastectomie du sein. Quant à l'acquisition des images, elle doit être réalisée dans les bonnes conditions, en veillant à ce que le positionnement de la femme et l'exposition aux rayons-X soient conformes à la procédure standard, de manière à eviter tout flou cinétique sur les images.

1.9.1.2 La vérité terrain

La preuve de la réalisation de la biopsie pour tous les cas évoqués doit être disponible. Par ailleurs, l'annotation doit inclure la vérité terrain concernant la cytologie /histologie de tous les cas, et les limites du contour de la lésion obtenue par la segmentation manuelle, réalisée par les radiologues experts en imagerie du sein.

1.9.1.3 Les informations associées à chaque cas

Les informations relatives à l'âge, aux antécédents personnels et familiaux de la patiente peuvent être incluses dans la mesure où elles peuvent s'avérer utiles pour améliorer les systèmes de DAOx. Quant à la classification de la densité du sein selon l'ACR et le BI-RADS, elles sont obligatoires.

1.9.1.4 L'organisation de la base et sa distribution

Il n'y a pas de format de fichier spécifique, toutefois, la plupart des images médicales sont sauvegardées au format DICOM (*digital imaging and communications in medicine*). Une par-

tition de ces images en sous base d'apprentissage et de test est également conseillée en vue de faciliter la comparaison des différents systèmes de DAOe ou de DAOx. Enfin, les bases doivent être accessibles à tous et de préférence via internet, sans oublier qu'une maintenance permanente doit être obligatoirement assurée par les propriétaires de la base.

1.9.2 Mammographic Image Analysis Society Digital Mammogram Database

Mammographic Image Analysis Society (MIAS) [Suckling et al., 1994] est la plus ancienne base d'images publiquement disponible ; elle a été conçue au Royaume-Uni en 1994 et est encore largement utilisée dans l'état-de-l'art. La MIAS contient 161 cas pour un total de 322 images numérisées obtenues en incidence MLO, sur lesquelles sont présentes toutes les pathologies à savoir les lésions bénignes et malignes, mais également les images normales. Cette base dispose d'un nombre important de masses spiculées, ainsi que d'une information relative à la densité du sein ; toutefois, la classification des masses réalisée par les radiologues ayant évalué cette base ne respecte pas les standards de l'ACR. Cependant, avec l'augmentation de l'usage de la classification BI-RADS de l'ACR, certains auteurs se sont essayés à une classification des masses de la MIAS afin qu'elle puisse correspondre aux standards en vigueur en imagerie du sein [Oliver et al., 2010b].

Un autre défaut de la MIAS est l'annotation de ces images, qui consiste à indiquer le centre et le rayon de la région d'intérêt, c'est-à-dire la région où se situe la pathologie. Ce genre d'annotation est considérée comme insuffisant pour certaines études comme la segmentation des masses [Oliver et al., 2010a], où l'on souhaite que toutes les lésions circonscrites ou spiculées soient manuellement segmentées afin de faire une comparaison très précise avec les méthodes automatiques. Pour finir, l'autre inconvénient de la MIAS, c'est la résolution à laquelle les images sont numérisées et qui fait que cette base ne convient pas pour des expériences liées à la détection des micro-calcifications [Domınguez et Nandi, 2007].

1.9.3 Digital Database for Screening Mammography

Constituée en 1999 aux Etats-Unis (USA), la *Digital Database for Screening Mammography* (DDSM) [Heath et al., 1998] est la plus grande base publique de mammographies. Elle contient 2620 cas incluant deux images (incidences MLO et CC) de chaque sein, pour un total de 10480 images où se retrouvent toutes les pathologies, des images normales à celles contenant des masses bénignes et malignes. Pour chacune de ces images, les informations relatives à l'âge de la patiente, la densité du sein et une classification respectant le BI-RADS sont fournies.

A la différence de la MIAS, les annotations concernant les pathologies sont faites manuellement par les radiologues, qui encerclent de façon un peu grossière la région suspecte. Toutefois, des travaux de recherche ont souligné que la DDSM n'était pas adéquate pour la validation des algorithmes de segmentation, en raison du manque de précision de la segmentation manuelle [Song et al., 2010].

1.9.4 BancoWeb LAPIMO

Plus récente que les deux précédentes bases, la *BancoWeb LAPIMO* est une base créée en 2010 au Brésil [Matheus et Schiabel, 2011]. Elle dispose de 320 cas avec 1473 images (MLO et CC)

contenant toutes les pathologies. Les antécédents de la patiente, de même qu'une classification BI-RADS sont fournis. En ce qui concerne les masses, seules certaines images sont annotées sous forme de ROI ; mais toutes les pathologies disposent d'une description textuelle.

1.9.5 INbreast

La INbreast [Moreira et al., 2012] est une base d'images acquises au Portugal entre avril 2008 et juillet 2010 au *Centro Hospitalar* de S. Joao (CHSJ) et qui vient combler l'attente des chercheurs en imagerie du sein, en ce sens qu'elle ne souffre pas des inconvénients liés aux autres bases évoquées plus haut. Les images acquises grâce à l'équipement *MammoNovation* de Siemens® sont des FFDM de 70 μm par pixel avec une résolution de 14 bit. Les images de cette base sont de 3328×4084 ou 2560×3328 pixels en fonction de la plaque de compression utilisée durant l'acquisition de la mammographie ; plaque dont les dimensions tiennent compte de la taille du sein de la patiente.

Chaque image est sauvegardée sur le format DICOM et contient les informations relatives à l'âge de la patiente, l'heure d'acquisition, les antécédents familiaux de la patiente, l'évaluation de la densité du sein en fonction de l'échelle de l'ACR et une classification de la masse selon le BI-RADS. Les images de la INbreast sont des mammographies de dépistage, de diagnostique et de suivie de patientes. En tout, 115 cas y sont collectés, dont 90 possèdent deux images (MLO et CC) de chaque sein, tandis que les 25 autres cas restants sont des femmes ayant subi une mastectomie et pour lesquelles il n'y a que deux images d'un seul sein, pour un total de 410 images en tout. Comme les bases précédentes, cette base contient toutes les pathologies liées au sein, de même que les mammographies normales.

Les annotations des images de la INbreast sont dans un premier temps réalisées par un radiologue, puis, elles sont ensuite validées par un autre radiologue entre avril et décembre 2010. Le logiciel utilisé pour la réalisation de l'annotation est Osirix® - un logiciel libre d'imagerie médicale permettant de visualiser et de traiter les images au format DICOM et qui ne tourne que sur Mac OS X. Pour les asymétries, les microcalcifications, les masses, les distorsions architecturales, le muscle pectoral, un contour bien précis est fourni par les radiologues. De plus, une ellipse englobant les amas de microcalcifications est utilisée quand il y a plusieurs microcalcifications regroupées au même endroit, tandis que pour les masses spiculées, une ellipse entourant toutes les spicules est utilisée.

Une illustration des images provenant des quatre bases de mammographies est présentée à la figure 1.13 ci-dessous. Par ailleurs, pour une meilleure comparaison des caractéristiques de ces bases, nous avons fait dans le tableau 1.3 ci-dessous, une synthèse de leurs forces et faiblesses.

1.10 Problématique

A travers les éléments présentés ci-dessus, on se rend compte de l'importance du cancer du sein en tant que problème prioritaire de santé publique du fait du taux de mortalité et de l'incidence qui lui sont associés. Des données qui en font la première cause de mortalité par cancer chez les femmes en France. L'utilisation des systèmes de DAOe/DAOx dédiés au cancer du sein en milieu clinique constitue une piste pour la réduction du taux de mortalité lié à ce cancer. Mais comme nous l'avons démontré plus haut, les systèmes conçues à ce jour sont tous

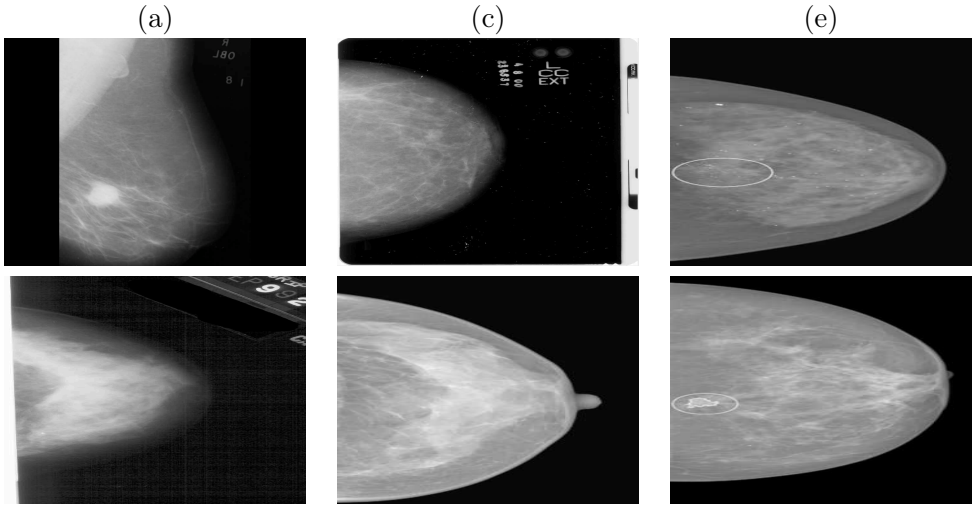


FIGURE 1.13 – Bases de mammographies : (a) image issue de la MIAS, (b) image provenant de la DDSM, (c) image de la BancoWeb LAPIMO, (d) image de la INbreast, (e) amas de microcalcifications de la INbreast et (f) masse spiculée de la INbreast.

TABLE 1.3 – Synthèse des attributs des différentes bases d'images les plus utilisées pour la conception et la validation des systèmes de DAOx.

Description	**MIAS**	**DDSM**	**BrancoWeb**	**INbreast**
Origine	G-B	E-U	Brésil	Portugal
Année	1994	1999	2010	2010
Nombre de cas	161	2620	320	115
Incidences	MLO	MLO & CC	MLO & CC	MLO & CC
Nombre d'images	322	10480	1400	410
Mode d'acquisition	Films	Films	Films	FFDM
Résolution	8 bits/pixel	8 ou 16 bits/pixel	12 bits/pixel	14 bits/pixel
Type de lésions	Tous les types	Tous les types	Tous les types	Tous les types
Vérité terrain	Centre de la masse	Grossière	Inexistante	Parfaite segmentation manuelle
BI-RADS	Non	Oui	Oui	Oui
Densité du sein	Oui (pas ACR)	Oui (ACR)	Oui (pas ACR)	Oui (ACR)
Système de recherche	Non	Pas fonctionnel	Oui	Oui
Accès public	Oui	Oui	Oui	Oui
Maintenance	Non	Non	Oui	Oui

destinés à distinguer les masses malignes de celles bénignes ; ils ignorent donc la troisième forme de signe radiologique traduisant le cancer, à savoir les distorsions architecturales. Les rares

tentatives menées dans l'évaluation de la performance des systèmes de DAOe pour la détection des distorsions architecturales ont prouvé que ces outils sont encore loin de produire des résultats fiables en milieu clinique [Baker et al., 2003].

Partant de ce constat, cette thèse s'est fixée pour objectif d'arriver à discriminer toutes les masses et plus spécifiquement les masses malignes entre elles, à savoir les masses classées BI-RADS IV – qui sont dans une grande proportion des distorsions architecturales – et les masses classées BI-RADS V – qui sont des masses malignes spiculées. Les motivations qui sous-tendent cet objectif sont doubles : d'une part la conception d'un système de DAOe permettrait de détecter efficacement les cas de distorsions architecturales et les masses spiculées et d'autre part, il offrirait la possibilité de réduire de façon importante le taux de faux négatifs sur ce type de lésions mammaires, étant donné que ce sont les distorsions architecturales qui sont les lésions les plus manquées par les radiologues – entre 30 et 45% [Baker et al., 2003]. En effet, à la lecture du champ risque de cancer du tableau 1.2, on se rend compte que le risque de cancer associé à une masse classée V est largement supérieur au risque encouru par une patiente dont la masse est de type IV sur l'échelle du BI-RADS, soit un risque $> 90\%$ contre un risque compris entre $[2, 90\%[$. Cela nous permet d'affirmer qu'une masse détectée dans sa phase IV peut encore être traitée à temps sans risquer d'avoir recours à des solutions invasives (radio et chimio-thérapie, voire mastectomie) ; de plus une détection précoce de cette masse limiterait les risques de métastases. A l'inverse, le fait de manquer une lésion de type IV conduit inévitablement au pire des scénarios car le prochain stade de développement de la distorsion architecturale est la classe V, *c-à-d.* une masse maligne spiculée à centre dense. La classification BI-RADS de l'ACR 1.1 nous permet d'observer que les masses de classe IV et V sont toutes des masses au contour irrégulier et/ou spiculé ; la seule différence apparente entre ces deux lésions semble être la nature dense du centre des masses de classe V contre le centre non-dense des masses de type IV.

La question qui se pose est donc celle-ci : *comment caractériser les masses au sens de la forme, de la densité et des spicules voire fusionner ces informations afin d'obtenir d'excellents résultats de classification pour les masses en général et en particulier pour les masses de classes IV et V ?.*

Pour tenter de répondre à cette interrogation, nous avons proposé un système de DAOx focalisé sur la caractérisation du contour des masses, l'extraction des spicules. Enfin, nous servant des résultats de ces différentes opérations, nous avons extrait de façon non-supervisée des descripteurs – grâce à une méthode issue du *Deep learning* – permettant une classification des masses d'une part en fonction de leur forme/contour et d'autre part en fonction de leur texture/densité par l'intermédiaire d'un classifieur SVM. Une présentation détaillée de chaque étape de la conception du système de DAOx est exposée dans la suite. Ci-dessous un schéma qui illustre le système de DAOx que nous avons proposé dans ce travail de thèse.

1.11 Conclusion

Dans ce chapitre ont été abordées des notions de généralité en rapport avec le cancer du sein et les systèmes d'acquisition en imagerie du sein. Après avoir présenté de façon concise l'anatomie du sein, nous avons exposé les données statistiques qui prouvent que le cancer du sein est l'un des principaux problèmes de santé publique à l'échelle mondiale. Par la suite, nous nous sommes

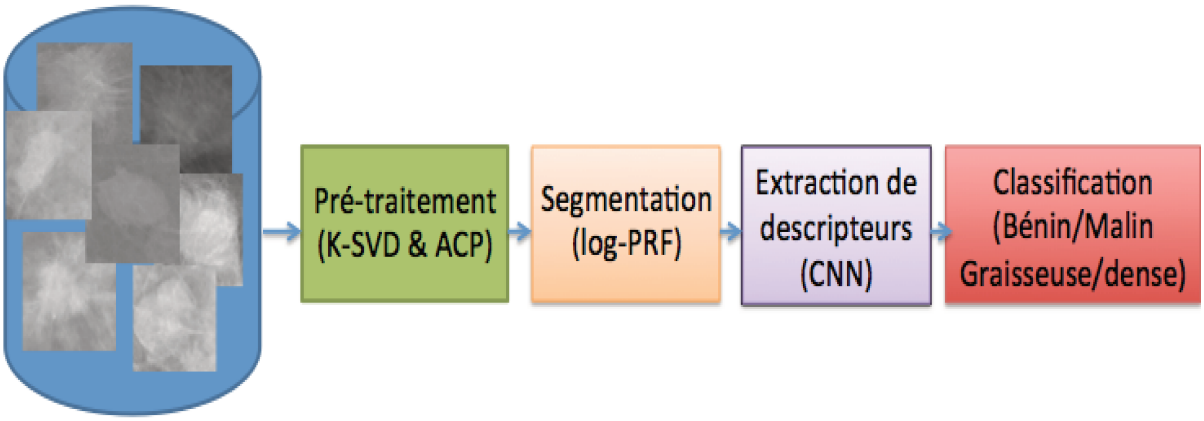


FIGURE 1.14 – Schéma des différentes étapes d'un système de DAOx.

intéressés aux manifestations du cancer du sein et aux modalités d'imagerie médicale permettant de l'identifier. Nous avons par la même occasion évoquer la classification BI-RADS proposée par l'ACR comme guide de l'interprétation des pathologies mammaires, tout en mettant en avant le caractère subjectif, fastidieux et chronophage de l'interprétation des images par les radiologues. L'objectif qui sous-tend la présentation de ces éléments, c'est de mettre en lumière les difficultés rencontrées par les radiologues dans l'analyse des mammographies et d'arriver ainsi à convaincre de la nécessité de développer des systèmes de DAOx surtout pour les cas – les masses malignes de type IV et V – jugés les plus complexes. Dans le prochain chapitre, nous allons exposer la première étape de la conception du système de DAOx que nous proposons, à savoir le pré-traitement des images. Cette étape est vraiment importante pour améliorer la qualité des mammographies afin de nous permettre d'obtenir les meilleures performances possibles de notre DAOx.

Chapitre 2

Pré-traitement des mammographies

Les mammographies sont des images qui n'offrent pas un bon contraste entre les tissus mammaires sain et malin ; ceci s'explique par le fait que l'atténuation des rayons X entre ces tissus varie peu, surtout chez les jeunes femmes qui ont un parenchyme mammaire très dense [Scharcanski et Jung, 2006]. La conséquence, c'est que les signes de suspicion d'un cancer du sein tel que, les masses et les micro-calcifications s'avèrent compliqués à repérer dans le tissu mammaire environnant, surtout quand ils sont de petites tailles. Ceci justifie l'usage des techniques de pré-traitement des mammographies, *c-à-d.* réduction du bruit et ou rehaussement du contraste, dans le but d'augmenter le contraste entre les différentes structures anatomiques et le reste de l'image, afin de faciliter à la fois l'interprétation de la mammographie par le radiologue et la réalisation des autres étapes indispensables pour une meilleure performance des systèmes de DAOx.

Les techniques standard d'amélioration de contraste et de réduction du bruit ont d'abord été appliquées aux mammographies, puis, on s'est rendu compte que, hormis l'amélioration apportée à l'image, ces approches contribuaient à amplifier de façon significative le bruit présent dans l'image [Chan et al., 1987]. Des méthodes un peu plus spécialisées ont alors été proposées et ont globalement aidé à améliorer la qualité des mammographies, comme l'a relevé [Ram, 1982], qui a également insisté sur le fait que l'utilisation de ces techniques en milieu clinique peut réduire de moitié la dose de radiation délivrée durant l'acquisition des mammographies.

Cependant, le réel problème des algorithmes de pré-traitement se situe dans le fait que certaines régions peuvent être traitées convenablement, tandis que d'autres le sont plus ou moins. Un manque de pré-traitement peut engendrer la négligence de plusieurs détails concernant la lésion ; dans ce cas de figure, certaines tumeurs peuvent ne pas être détéctées, causant ainsi des faux négatifs. A contrario, l'excès de rehaussement de contraste peut rajouter des détails ou artefacts à l'image, donnant ainsi au radiologue l'impression d'être en présence d'une tumeur, ce qui aboutit à des biopsies inutiles et une augmentation des faux positifs. A cela s'ajoute le fait que les approches locales et multi-échelles nécessitent la spécification de plusieurs paramètres (*p. ex.* la taille et le type de filtre, les seuils et les différentes résolutions) qui ont un grand impact sur la qualité des résultats obtenus.

Après avoir exposé de façon concise les différentes classes de méthodes de pré-traitement, nous allons présenter en détails la solution que nous proposons – qui est basée sur une combinaison efficace des algorithmes de K-SVD et d'ACP – pour résoudre convenablement ce problème. Nous finirons par une illustration de la méthode sur des mammographies ainsi qu'une étude comparative approfondie avec les techniques les plus utilisées et offrant de meilleurs résultats en mammographie.

2.1 Travaux existants sur le pré-traitement en mammographie

La restauration des images a depuis toujours, fait l'objet d'une attention particulière, car le bruit a un impact néfaste sur la qualité des traitements qu'on peut apporter à l'image. Ceci est particulièrement vrai en mammographie où le bruit rend difficile la détection des structures fines et petites de l'image. Globalement, on peut subdiviser les techniques de pré-traitement des images en trois classes : les approches globales, celles locales et celles qui sont basées sur la multi-résolution. En nous inspirant des résultats obtenus par [Chan et al., 1987], on peut dire que les transformations globales sont des techniques non adaptées à l'information locale contenue dans les mammographies. En effet, il existe une variabilité importante dans la taille, la forme, le contour et la densité des lésions, qui fait que les techniques standards fournissent des résultats loin d'être satisfaisants. Il est donc judicieux de développer des techniques locales d'amélioration du contraste adaptées au contenu des mammographies.

C'est dans ce cadre que [Gordon et Rangayyan, 1984] ont proposé une technique de rehaussement de contraste des mammographies, basée sur un voisinage du pixel courant, nommée *Adaptive Neighborhood Contrast Enhancement* (ANCE). Dans cette approche, les auteurs considèrent chaque pixel courant comme un pixel germe dans une méthode de croissance de régions. La procédure de croissance de régions identifie l'ensemble des pixels qui ont la même intensité de niveau de gris que le pixel germe et reliés à ce dernier, de même qu'une zone de trois pixels d'épaisseur entourant la région à laquelle appartient le pixel germe. La nouvelle valeur du pixel germe après rehaussement, est obtenue en faisant la différence entre le fond de l'image et la région à laquelle appartient le pixel germe. Pour sa part, [Karssemeijer, 1993] a montré que le bruit avait tendance à augmenter avec l'intensité des niveaux de gris des pixels dans les images où le contraste local et l'intensité de l'image sont inter-dépendants, ce qui est le cas des mammographies. Il a alors proposé une méthode d'égalisation du bruit afin d'obtenir des images où le contraste local est approximativement le même pour toutes les intensités de l'image. [McLoughlin et al., 2004] ont amélioré cette technique dans le but d'avoir une meilleure estimation du bruit ; méthode qui sera par la suite approfondie par [Scharcanski et Jung, 2006]. Une approche Bayésienne pour l'amélioration de contraste, faisant l'hypothèse que le bruit serait a priori un bruit Gaussien additif, a été proposée par [Simoncelli et Adelson, 1996] ; c'est un algorithme de suppression de bruit semi-aveugle, qui est basé sur une décomposition de l'image sous forme de pyramide.

La dernière catégorie des techniques de suppression de bruit est basée sur les transformations d'ondelettes [Laine et al., 1994]. Ces méthodes varient selon l'ondelette mère, la fonction et les coefficients utlisés, et opèrent comme un filtre passe-bande ; ce qui leur permet d'améliorer le contraste des lésions de façon indépendante [Mencattini et al., 2008]. Dans un régistre similaire, [Mihcak et al., 1999] ont proposé un modèle statistique spatialement adaptatif pour la suppression du bruit ; dans leur approche, les coefficients des ondelettes ont été modélisés comme des variables aléatoires Gaussiennes ayant une forte corrélation locale. Par la suite, les *curvelet* ont été proposés comme une amélioration des ondelettes, qui code l'information directionnelle avec une bien meilleure précision au fur et à mesure que la résolution s'affine [Candes et al., 2006].

Toutefois, certaines de ces méthodes, tout en améliorant la visibilité des structures anatomiques, participent à la distorsion de leur apparence et peuvent engendrer un mauvais diagnostic. Dans l'idée d'éviter cette situation, nous avons mis en place une approche qui est basée sur la

combinaison efficace des algorithmes de K-SVD [Aharon et al., 2006] et de l'analyse en composantes principales (ACP) [Kirby et Sirovich, 1990].

2.2 Méthode proposée

2.2.1 Extraction des patchs et pré-traitement

Premièrement, nous procédons à l'extraction aléatoire de plusieurs régions carrée appelées *patchs* et notées p, de chacune des images contenues dans nos bases ; chacun des patchs a une taille de $s \times s$ pixels, pour lesquels une intersection est autorisée. Ces patchs sont ensuite centrés (2.1) et normalisés (2.2) comme montré à la figure 2.1, avec $p(x, y)$ la valeur de l'intensité du pixel situé dans le patch p et dont la position est donnée par les coordonnées x et y.

$$p'(x,y) = p(x,y) - \bar{p}, \quad \text{avec } \bar{p} = \frac{1}{s^2}\sum_{x=1}^{s}\sum_{y=1}^{s} p(x,y) \ . \tag{2.1}$$

$$p''(x,y) = \frac{p'(x,y)}{||\ p'\ ||_2}, \quad \text{avec } ||\ p'\ ||_2 = \sqrt{\sum_{x=1}^{s}\sum_{y=1}^{s} p'(x,y)^2} \ . \tag{2.2}$$

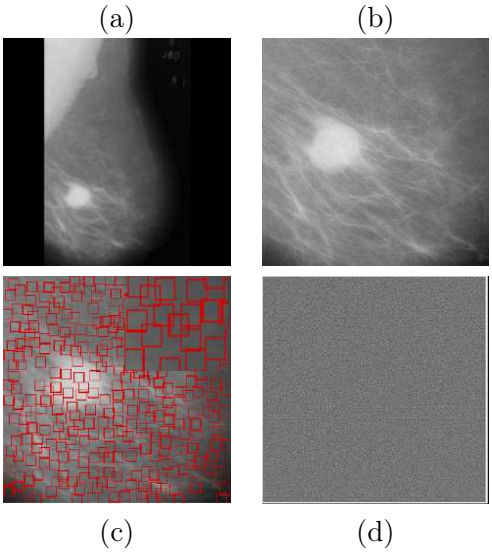


FIGURE 2.1 – Extraction des patchs et pré-traitement : (a) image originale, (b) ROI, (c) extraction de patchs sur (b) avec une insertion d'un zoom x5 et(d) patchs normalisés

2.2.2 Apprentissage du dictionnaire parcimonieux et algorithme du K-SVD

Le K-SVD est un algorithme qui permet l'apprentissage des mots visuels les plus fréquents de l'image et leur représentation sous forme de dictionnaire dans le but d'aboutir à une représentation parcimonieuse de l'image. Etant donné une base contenant des images dédiées à l'entraînement, le but est de chercher le dictionnaire qui offre une meilleure représentation de chaque image de cette base tout en imposant des contraintes de parcimonie strictes. Cet algorithme peut être perçu comme une généralisation du processus de clustering lié au K-means [Aharon et al., 2006]. L'approche que nous proposons pour la suppression de bruit, s'inspire du K-SVD en raison de sa robustesse, du faible taux d'erreur engendré par l'apprentissage mais également parce qu'il nous permet de supprimer toute spécification de paramètres, une fois le bon niveau de parcimonie trouvé.

Soit N le nombre de patchs extraits de toute la base d'entraînement $\{y_i\}_{i\in\{1;N\}}$ où chaque patch est un vecteur colonne tel que : $y_i \in \mathbb{R}^n$ avec $n = s^2$ et $\mathbf{Y}_{n\times N}$ la matrice contenant tous les patchs. On va chercher à représenter $\mathbf{Y}$ comme une combinaison linéaire $\mathbf{X}$ d'éléments de base notés $\mathbf{D}$:

$$\mathbf{Y} \approx \mathbf{DX} \ . \tag{2.3}$$

$\mathbf{D}_{n\times K}$ est une matrice dont les vecteurs colonne sont les K éléments du dictionnaire qu'on cherche à apprendre. $\mathbf{X}_{K\times N}$ est la matrice dont les vecteurs lignes $\{x_k\}_{k\in\{1;N\}}$ contiennent toutes les projections de $\mathbf{Y}$ sur $\mathbf{D}$. Dans le but d'obtenir la meilleure représentation des données d'entrée $\mathbf{Y}$, le problème consiste à trouver $\mathbf{D}$ et $\mathbf{X}$ minimisant le critère ci-dessous :

$$\min_{\mathbf{D},\mathbf{X}}\{|| \mathbf{Y} - \mathbf{DX} ||_2^2\} \ . \tag{2.4}$$

avec $|| \mathbf{X} ||_2$ la norme L_2 de $\mathbf{X}$ définit comme $\sqrt{\sum_{ij} \mathbf{X}_{ij}^2}$. Si n < K, la matrice $\mathbf{D}$ est de rang plein, ce qui signifie qu'il existe une infinité de solutions ; ce qui nécessite d'imposer des contraintes pour limiter l'espace de recherche. Notamment les contraintes restreignant la projection de y_i sur $\mathbf{D}$ en utilisant une approche parcimonieuse ont été beaucoup étudiées [Chen et al., 1989]. Ce qui nous ramène à résoudre le problème suivant :

$$\min_{\mathbf{D},\mathbf{X}}\{|| \mathbf{Y} - \mathbf{DX} ||_2^2\}, \quad \text{tel que} \ : \ \forall i \ || \ x_i \ ||_0 \leq T_0 \ . \tag{2.5}$$

où $|| \ x_i \ ||_0$ est la norme L_0 du vecteur x_i, c'est-à-dire le nombre de composantes non-nulles et T_0 le niveau de parcimonie, c'est-à-dire le nombre d'entrées non-nulles. Pour résoudre le problème posé en (2.5), nous utilisons l'algorithme optimisé du K-SVD [Aharon et al., 2006], [Rubinstein et al., 2010] qui consiste en une optimisation alternée. On résoud itérativement un problème en $\mathbf{X}$ à $\mathbf{D}$ fixé : c'est l'étape du *codage parcimonieux* à laquelle nous résolvons :

$$\min_{\mathbf{X}}\{|| \mathbf{Y} - \mathbf{DX} ||_2^2\}, \quad \text{tel que} \ : \ \forall i \in \{1; N\} \ || \ x_i \ ||_0 \leq T_0 \ . \tag{2.6}$$

Plusieurs techniques existent pour résoudre l'équation (2.6) ; dans notre cas, nous utilisons *orthogonal matching pursuit (OMP)* [Tropp, 2004] qui est un algorithme glouton et non optimal, permettant d'approximer la solution de l'équation. Il possède des propriétés intéressantes au regard de la réduction d'erreur entre les dictionnaires parcimonieux et le dictionnaire original.

La seconde étape du K-SVD consiste en une optimisation de $\mathbf{D}$ avec $\mathbf{X}$ fixé :

$$\min_{\mathbf{D}}\{|| \mathbf{Y} - \mathbf{DX} ||_2^2\} \ . \tag{2.7}$$

Ceci implique une optimisation de chaque atome d_k du dictionnaire de façon indépendante des autres éléments supposés fixes. La fonction coût de l'équation (2.7) peut être écrite comme suit :

$$\begin{aligned} f(\mathbf{D}) = || \mathbf{Y} - \mathbf{DX} ||_2^2 = \left\| \mathbf{Y} - \sum_{j=1}^{K} d_j x_T^j \right\|_2^2 &= \left\| \left(\mathbf{Y} - \sum_{j \neq k}^{K} d_j x_T^j \right) - d_k x_T^k \right\|_2^2 \\ &= \left\| E_k - d_k x_T^k \right\|_2^2 \ . \end{aligned} \tag{2.8}$$

avec d_k une colonne de $\mathbf{D}$ et les coefficients x_T^k correspondant à la k^{th} ligne de $\mathbf{X}$. Il est à noter que x_T^k n'est pas équivalent à x_k, le dernier étant la k^{th} colonne de $\mathbf{X}$. Chaque produit $d_j x_T^j$ est une matrice de taille $(n \times N)$. Le terme E_k dans l'équation (2.8) est l'erreur de reconstruction par rapport aux données d'entrée, pour une optimisation exécutée sur K-1 éléments du dictionnaire $\{d_j\}_{j \neq k}$.

Etant donné que ces éléments sont supposés fixés, on cherche à résoudre $f(d_k)$. On peut y arriver en utilisant la méthode des moindres carrées ; toutefois, on souhaite garantir que la mise à jour de la solution permet une décomposition parcimonieuse. Afin de garantir cette propriété, des poids spécifiques ont été définis :

$$w_k = \{i | 1 \leq i \leq K, x_T^k(i) \neq 0\} \ . \tag{2.9}$$

où w_k est l'ensemble des indices pour les données d'entrées qui utilisent d_k. La méthode utilisée pour optimiser d_k est la suivante :

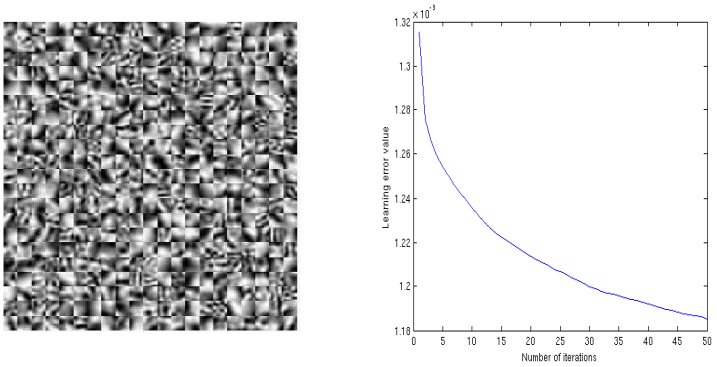
1. Restreindre E_k en choisissant uniquement les colonnes correspondant à w_k et déterminer E_k^R. On note $\mathbf{\Omega_k}$ la matrice de taille $N \times \mid w_k \mid$ qui correspond au fait de recopier N fois le vecteur ligne w_k, ce qui donne $E_k^R = E_k \mathbf{\Omega_k}$ et $x_R^k = x_T^k \mathbf{\Omega_k}$,

2. La solution consiste à minimiser de $\left\| E_k^R - d_k x_R^k \right\|_2^2$ à la fois par rapport à d_k et x_R^k en garantissant que le support de x_T^k (c'est-à-dire les w_k non-nuls) reste inchangés. Pour calculer $\underset{d_k, x_R^k}{min} \left\| E_k^R - d_k x_R^k \right\|_2^2$ on réalise une décomposition en valeurs singulières sur la fonction coût $E_k^R = \mathbf{U \Delta V^T}$. La solution de d_k et x_R^k est :

   (a) d_k est la première colonne de $\mathbf{U}$ (matrice unité),

   (b) x_R^k est la première colonne de $\mathbf{V}$ (matrice unité) multipliée par $\mathbf{\Delta}(1,1)$ (matrice diagonale).

Cet algorithme a été appliqué sur les patchs extraits de la base d'images comme décrit à la section 2.2.1. La figure 2.2 donne un aperçu des résultats obtenus sur une des bases d'images utilisée.



(a) (b)

FIGURE 2.2 – Apprentissage de dictionnaire parcimonieux : (a) dictionnaire appris sur une base de patchs et (b) courbe d'erreur d'apprentissage. Notez la décroissance monotone de l'erreur d'apprentissage en (b).

2.3 Réduction de la taille du dictionnaire

Afin de réduire de façon efficace la taille du dictionnaire appris plus tôt dans l'idée d'optimiser le temps de calcul tout en retenant une très grande partie de l'information contenue dans le dictionnaire, nous avons décidé d'utiliser l'ACP [Kirby et Sirovich, 1990]. L'ACP est un algorithme largement utilisé pour la réduction de dimension et l'extraction de descripteurs, dont le but est de trouver un sous-espace linéaire de dimension inférieure à celui de l'espace de départ des descripteurs, de manière à ce que les éléments du sous-espace soient ceux ayant la plus grande variance. Dans notre cas, la taille du dictionnaire appris varie en fonction du bon-vouloir de l'utilisateur et il n'y a aucune façon de la contrôler, ce qui peut avoir diverses conséquences. Ainsi, un dictionnaire de grande taille implique un temps de calcul élevé. Aussi nous avons réfléchi à une technique pour réduire le temps de calcul en cherchant les atomes du dictionnaire qui contiennent la plus grande quantité de l'information apprise.

Considérons le dictionnaire $\mathbf{D}\{d_k\}_{k\in\{1;L\}}$ où L est la taille du dictionnaire parcimonieux et d_k le patch de taille $s \times s$ qui est considéré comme un vecteur colonne de dimension s^2. Ainsi, l'ensemble des patchs correspond à une collection de points dans un grand espace. Les patchs extraits sur notre base d'images, sont similaires dans une large mesure, en raison des motifs répétitifs qu'on retrouve en mammographie et souvent désignés comme les *blocs de Tabar* [Tabár et al., 2005]. Ces motifs ne sont pas uniformément distribués dans l'espace de départ et peuvent donc être décrits par un sous-espace de plus faible dimension. L'idée principale de l'ACP est de trouver le sous-ensemble de patchs qui représente de façon significative la majorité des atomes du dictionnaire.

Le dictionnaire appris est donc soumis à l'ACP qui cherche un ensemble de M vecteurs orthogonaux ($M < L$) qui décrivent au mieux la distribution des données. On suppose que la projection est notée comme suit : $\mathbf{Z} = \mathbf{AD}$ où $\mathbf{A} = [u_1^T, \cdots, u_{\mathrm{M}}^T]$ et $u_i^T u_i = 1$ pour $i \in \{1; M\}$. La variance de $\mathbf{Z}\{z_k\}_{k\in\{1;L\}}$ qui est la trace de la matrice de covariance de $\mathbf{Z}$ est à minimiser.

Ainsi, la fonction objective qui en résulte est la suivante :

$$\mathbf{Z}^* = \underset{\mathbf{Z}}{\text{argmax}} \text{ tr } (\mathbf{C_Z}) \ . \tag{2.10}$$

où

$$\mathbf{C_Z} = \frac{1}{L}\sum_{k=1}^{L}(z_k - \bar{\mathbf{Z}})(z_k - \bar{\mathbf{Z}})^T \ . \tag{2.11}$$

et

$$\bar{\mathbf{Z}} = \frac{1}{L}\sum_{k=1}^{L} d_k \ . \tag{2.12}$$

Soit $\mathbf{C_D}$ la matrice de covariance de $\mathbf{D}$. Donc, $\text{tr}(\mathbf{C_Z}) = \text{tr}(\mathbf{A}\mathbf{C_Z}\mathbf{A}^T)$. En utilisant les multiplicateurs de Lagrange, désignés par λ_i, l'expression suivante sera évaluée par chaque i :

$$u_i^T \mathbf{C_Z} u_i + \lambda_i(1 - u_i^T u_i) \ . \tag{2.13}$$

En mettant la dérivée par rapport à u_i à zéro, on peut observer que l'équation (2.13) admettra un point stationnaire si :

$$\mathbf{C_Z} u_i = \lambda_i u_i \ . \tag{2.14}$$

Ce qui signifie que u_i est une valeur propre de $\mathbf{C_D}$. Si nous multiplions chaque membre de l'équation (2.14) par u_i^T et utilisons la propriété remarquable suivante $u_i^T u_i = 1$, on observe que la variance est donnée par :

$$u_i^T \mathbf{C_Z} u_i = \lambda_i \ . \tag{2.15}$$

Ainsi, d_k peut être représenté comme :

$$d_k = \sum_{k=1}^{L}(d_k^T u_i)u_i \ . \tag{2.16}$$

Notez que d_k peut être également approximé par :

$$\hat{d}_k = \sum_{k=1}^{M}(d_k^T u_i)u_i \ . \tag{2.17}$$

où u_i est le vecteur propre de $\mathbf{C_D}$ correspondant à la i^{th} plus grande valeur propre.

2.4 Suppression du bruit

Une fois le dictionnaire appris et la sélection de sa taille optimisée, il est utilisé pour supprimer le bruit présent dans chacun des patchs extraits de l'image bruitée. La méthode de suppression de bruit consiste au décodage de chaque patch bruité y_i sur le dictionnaire, c'est-à-dire à la détermination de la projection optimale $\hat{x}$ de façon à minimiser l'équation ci-dessous :

$$\hat{x} = \min_{x_i} || \ x_i \ ||_0, \quad \text{tel que} \ : \ || \ \mathbf{Y} - \mathbf{D}\mathbf{X} \ ||_2^2 \leq \Gamma_0 \ . \tag{2.18}$$

Contrairement au cas du codage parcimonieux, ici, on ne cherche pas à reconstruire exactement le patch y_i parce qu'il est bruité. A la place, on cherche une erreur de reconstruction ($\leq \Gamma_0$) raisonable.

2.5 Implémentation numérique

Les images de la MIAS et de la INbreast ont été utilisées afin d'apprendre sur chacune de ces bases un dictionnaire parcimonieux. Le choix de la MIAS se justifie par le fait que c'est une base de mammographies numérisées et donc que chacune de ces images est affectée par le bruit de numérisation qu'on va tenter de réduire dans la suite. Quant à la INbreast, elle est la base de mammographies la plus récente, sur laquelle nous souhaitons étudier la réduction du bruit d'acquisition à l'aide de notre méthode.

Pour chacune des images de la MIAS et de la INbreast, on extrait une ROI de taille 300×300 pixels en vue d'extraire le même nombre de patchs sur tous les ROIs. La mise à l'échelle de chaque ROI est faite grâce à une interpolation bi-cubique. Ensuite, 20000 patchs de taille $s \times s$ sont extraits sur chacune des bases, avec s=14 et δ=6, l'intersection maximale autorisée entre deux patchs. Pour apprendre un dictionnaire sur la MIAS et la INbreast, nous avons utilisé les paramètres suivants : taille du dictionnaire = 500, niveau de parcimonie = 10 et le nombre d'itérations = 50.

En vue de tester l'efficacité de cette approche, nous avons ajouté numériquement du bruit gaussien additif (GN) aux images, avec un $\sigma = 20$. Cette valeur a été estimée en fonction des paramètres généralement utilisés pour l'acquisition des mammographies. L'incertitude liée au bruit quantique a été considérée pour deux situations : premièrement, une acquisition de mammographie standard avec un air kerma (*kinetic energy released per unit mass*) – grandeur physique utilisée pour la mesure des faisceaux physiques sans charge – de 50 μGy à la surface du détecteur et une efficacité quantique de détection (DQE) de 0.35 ; puis une acquisition de mammographie à faible dose, avec un air kerma de 35 μGy et un DQE de 0.5 comme suggéré par [Monnin et al., 2007]. L'atténuation exponentielle de 17.5 keV de photons sur un tissu de sein compressé de 4 cm [Poulos et al., 2003] et une plaque de compression équivalent à un total de 2 cm de tissu ont été considérés. L'écart-type relatif aux valeurs des pixels a été estimé dans les mêmes conditions, ce qui aboutit à 15% dans les deux cas de figure. Cela correspond à un $\sigma = 20$ pour des images de niveau de gris codés sur 8 bits, où la radio-densité la plus élevée du sein est représentée par une valeur d'intensité de pixel de 133.

Dans le but de nous comparer aux ondelettes de Daubechies, nous avons utilisé la *wavelet toolbox* de MATLAB® (2014a) avec les paramètres suivants : wname=db10, niveaux de décomposition=3, seuils (coefficient horizontal/diagonal/vertical)=4/4/8 (les mêmes seuils pour les trois niveaux de décomposition). Pour la comparaison avec les *curvelet*, nous avons eu recours à la toolbox fournie par les auteurs des *curvelet* [Candes, 2012], dont l'utilisation consiste à appliquer sur chaque image une transformée de Fourier espacée de façon inégale (USFFT), par le biais d'une grille rectangulaire décimée et inclinée le long de la direction principale de chaque *curvelet*.

Dans l'idée d'obtenir la meilleure reconstruction possible, nous avons utilisé un seuil issu de la formule suivante : $\Gamma_0 = s \times \sigma \times \texttt{bias}$, avec $\texttt{bias} = 1.15$; ce dernier ayant été trouvé empiriquement après de multiples essais.

Afin de montrer de façon objective l'efficacité de cette technique, nous avons décidé de mesurer la qualité de reconstruction de notre méthode par l'intermédiaire du *peak signal to noise ratio* (PSNR) et de l'erreur quadratique moyenne (EQM). Soit I_s l'image originale, I_r l'image reconstruite, N_r, M_c et d respectivement le nombre de lignes, de colonnes et la dynamique de

l'image.

$$\mathrm{PSNR}(I_\mathrm{r}, I_\mathrm{s}) = 10 \log_{10} \left(\frac{d^2}{\mathrm{EQM}(I_\mathrm{r}, I_\mathrm{s})} \right) \quad . \tag{2.19}$$

avec

$$\mathrm{EQM}(I_\mathrm{r}, I_\mathrm{s}) = \frac{1}{N_\mathrm{r} \times M_\mathrm{c}} \sum_{i=1}^{N_\mathrm{r}} \sum_{j=1}^{M_\mathrm{c}} [I_\mathrm{r}(i,j) - I_\mathrm{s}(i,j)]^2 \quad . \tag{2.20}$$

Plus grande est la valeur du PSNR et meilleure est la qualité de l'image reconstruite. Malgré la très forte corrélation entre le PSNR et l'EQM, les valeurs liées à ces deux mesures ont été exposées dans la suite par souci de clarté.

2.6 Résultats et discussion

2.6.1 Application aux ROIs

L'utilisation de l'ACP a permis de réduire la taille du dictionnaire de 500 à 40 éléments ; ce qui représente une réduction de 12.5 de la taille initiale du dictionnaire. La figure 2.3 ci-dessous donne une vue globale de la distribution des valeurs propres calculées grâce à l'ACP. Elle démontre que seul 8% des données du dictionnaire de départ contiennent plus de 99% de l'information disponible dans notre base d'images. Ce résultat vient confirmer notre hypothèse de départ fondée sur le caractère similaire des patchs extraits sur une mammographie.

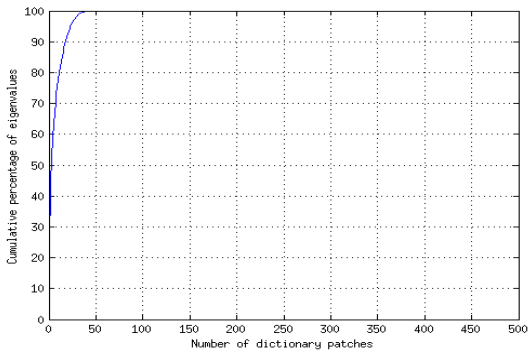


FIGURE 2.3 – Pourcentages cumulés des valeurs propres.

Pour évaluer la qualité de notre méthode, nous avons effectué une comparaison avec les ondelettes de Daubechies [Daubechies, 1992], qui représentent l'une des techniques les plus utilisées pour la suppression de bruit. Par ailleurs, nous nous sommes également comparé aux *curvelet* [Candes et al., 2006] qui partagent les mêmes concepts théoriques que les ondelettes, tout en étant une version très améliorée de celles-ci. De façon générale, les ondelettes et les *curvelet* sont deux méthodes très répandues parmi les techniques de suppression de bruit et sont basées sur le choix des niveaux de décomposition d'une image de même que sur le choix du seuil pour chacun de ces niveaux. Pour finir, notre approche basée sur la combinaison du K-SVD et de l'ACP (K-SVD-ACP) a été comparée au K-SVD conventionnel. Les résultats obtenus sont montrés sur les figures 2.4 et 2.5 ci-dessous. Les différences entre les images bruitées avant et après

suppression du bruit semblent être évidentes : sur les ROIs, on peut noter que les images (c), (d) et (f) de la figure 2.4 et (c), (d) et (f) de la figure 2.5 sont floutées comparées à la ROI sans ajout de bruit Gaussien additif. Cependant, sur les figures 2.4 (e) et 2.5 (e) on peut remarquer qu'une importante quantité du bruit numériquement ajouté subsiste toujours dans l'image filtrée par les ondelettes de Daubechies. A l'inverse, les *curvelet* suppriment effectivement le bruit au prix d'un ajout de flou et d'artefacts qui sont plus ou moins importants dans différentes régions de l'image sur les figures 2.4 (f) et 2.5 (f). De la même manière, on observe que le K-SVD – montré sur les images 2.4 (c) et 2.5 (c) – de même que le K-SVD-ACP – illustrées par les images 2.4 (d) et 2.5 (d) – semblent souffrir de l'ajout d'une certaine quantité de flou sur l'image ; toutefois, elles suppriment tout aussi bien le bruit que les *curvelet*. De plus, on peut remarquer que les méthodes de suppression de bruit basées sur l'apprentissage d'un dictionnaire parcimonieux semblent répondrent correctement à une normalisation de l'histogramme comme le montrent ces figures, ce qui n'est pas le cas des images obtenues après utilisation des *curvelet*.

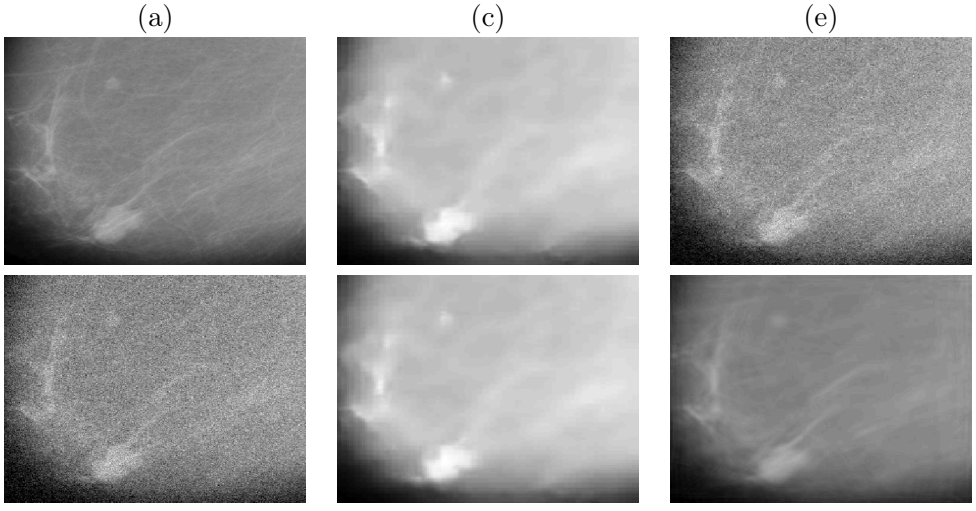


FIGURE 2.4 – Comparaison de quatre méthodes de suppression de bruit sur une image extrait de la MIAS : (a) ROI, (b) GN, (c) K-SVD suivie d'une normalisation de l'histogramme, (d) K-SVD-ACP suivie d'une normalisation de l'histogramme, (e) ondelettes de Daubechies et (f) *curvelet*.

Ensuite, le PSNR a été utilisé pour confirmer quantitativement les résultats montrés sur les figures ci-dessus. Les valeurs du PSNR de (24.61±0.05) dB contre (38.80±1.5) dB ont été obtenues pour la MIAS et (24.62±0.04) dB vs (39.01±1.39) dB pour la INbreast en utilisant respectivement les ondelettes de Daubechies et les *curvelet*. Ces résultats disposés dans les tableaux 2.1 et 2.2, de même que les valeurs de l'EQM, montrent que le K-SVD-ACP à l'instar du K-SVD et des *curvelet* est plus efficace que les ondelettes de Daubechies pour la suppression de bruit sur les mammographies avec les paramètres utilisés dans cette étude. Toutefois, il est important de signifier que les *curvelet* au même titre que les ondelettes de Daubechies nécessitent une sélection rigoureuse des niveaux de décomposition de l'image et des seuils associés, tandis que le K-SVD et l'approche que nous proposons s'affranchissent de cette tâche. Par ailleurs, les

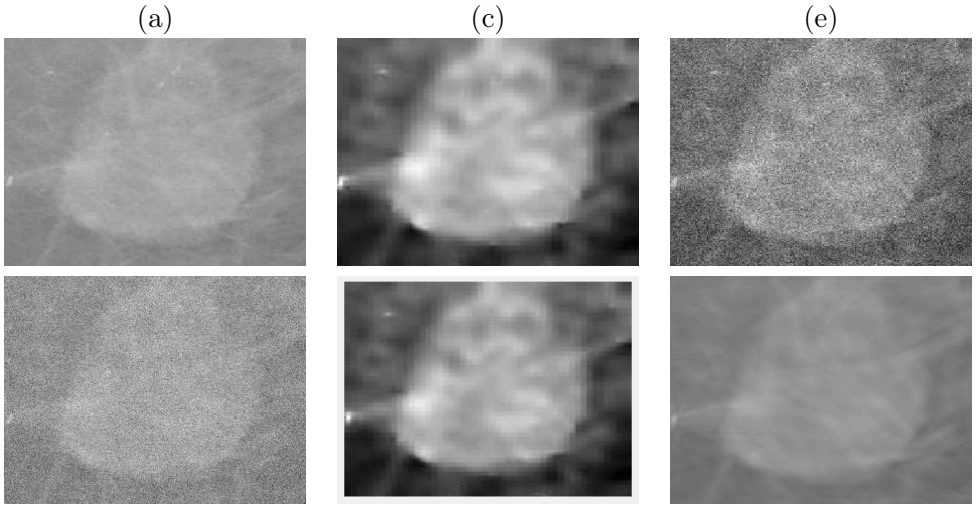


FIGURE 2.5 – Comparaison de quatre méthodes de suppression de bruit sur une image extrait de la INbreast : (a) ROI, (b) GN, (c) K-SVD suivie d'une normalisation de l'histogramme, (d) K-SVD-ACP suivie d'une normalisation de l'histogramme, (e) ondelettes de Daubechies et (f) *curvelet*.

TABLE 2.1 – Les valeurs des moyenne et écart-type de l'EQM obtenues sur les bases de ROIs de la MIAS et du INbreast après exécution des ondelettes de Daubechies, du K-SVD, K-SVD-ACP et des *curvelet* sur des images numériquement bruitées.

	Daubechies	LSD	K-SVD-ACP	*curvelet*
	EQM	EQM	EQM	EQM
MIAS	62.10±0.40	5.70±2.87	5.68±2.75	**4.01±1.26**
INbreast	61.91±0.53	6.08±2.80	6.08±2.78	**4.02±1.38**

valeurs moyennes de PSNR montrent que le K-SVD, le K-SVD-ACP et les *curvelet* montrent des performances similaires en terme de reconstruction de l'image, avec un léger avantage pour les *curvelet*. De façon spécifique, la valeur absolue de la différence des moyennes de PSNR entre les *curvelet* et les techniques d'apprentissage de dictionnaire parcimonieux est de 1.17 dB. Il est important d'ajouter que, considérant la dynamique d'une image codée sur 8 bits, la valeur maximale de PSNR qu'on peut obtenir est de 48.13 dB. L'efficacité des méthodes basées sur l'apprentissage d'un dictionnaire parcimonieux est liée au fait qu'elles supposent que le bruit présent dans l'image empêche une décomposition parcimonieuse en raison de sa nature aléatoire. Dans cette étude, nous sommes parvenus à trouver un bon niveau de parcimonie pour chacune des bases d'images utilisées, ce qui nous a permis d'améliorer la qualité des images de la MIAS et de la INbreast. Ainsi, une fois que ces paramètres sont fixés, notre méthode peut être vue comme une approche de suppression de bruit entièrement automatique, ce qui présente l'intérêt d'éviter une perte de temps considérable liée à la sélection de plusieurs paramètres comme le niveau de décomposition et les seuils pour chaque image de la base.

TABLE 2.2 – Les valeurs des moyenne et écart-type du PSNR obtenues sur les bases de ROIs de la MIAS et du INbreast après exécution des ondelettes de Daubechies, du K-SVD, K-SVD-ACP et des *curvelet* sur des images numériquement bruitées.

	Daubechies	K-SVD	K-SVD-ACP	*curvelet*
	PSNR (dB)	PSNR (dB)	PSNR (dB)	PSNR (dB)
MIAS	24.61±0.05	37.87±1.97	37.91±1.73	**38.80±1.5**
INbreast	24.62±0.04	37.55±2.06	37.56±1.96	**39.01±1.39**

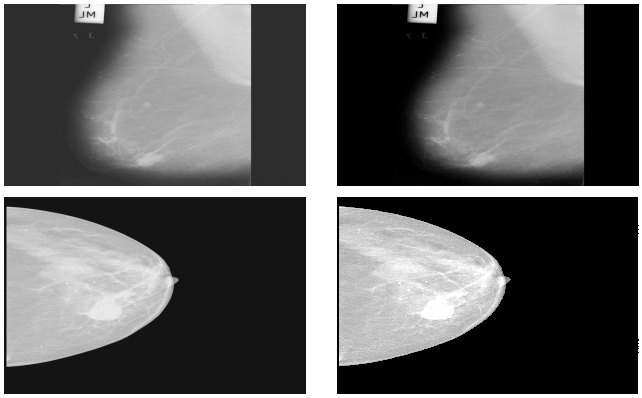
TABLE 2.3 – Les valeurs des moyenne et écart-type du temps de calcul (CTe) obtenues sur les bases de ROIs de la MIAS et du INbreast après exécution des ondelettes de Daubechies, du K-SVD, K-SVD-ACP et des *curvelet* sur des images numériquement bruitées.

	Daubechies	K-SVD	K-SVD-ACP	*curvelet*
	CTe (s)	CTe (s)	CTe (s)	CTe (s)
MIAS	0.15±0.06	3.70±0.23	**1.65±0.19**	3.76±0.15
INbreast	0.26±0.10	3.84±0.21	**1.77±0.18**	3.76±0.19

Pour finir, nous avons également évalué le temps de calcul et les résultats obtenus sont affichés dans le tableau 2.3. Pour les ondelettes de Daubechies, le temps de calcul est de (0.15±0.06) s et (0.26±0.10) s respectivement pour la MIAS et la INbreast. Tandis que pour les *curvelet*, le temps de calcul est de (3.76±0.15) s et (3.76±0.19) s pour ces mêmes bases d'images. Quant à l'éxécution du K-SVD conventionnel, elle se fait en (3.70±0.23) s et (3.84±0.21) s respectivement pour la MIAS et la INbreast. On remarque que ces deux dernières valeurs sont largement au-dessus des temps de calcul des ondelettes de Daubechies. Pour ce qui est de la méthode que nous avons proposé, elle est 2× plus rapide que les *curvelet* et le K-SVD, avec des temps de calcul de (1.65±0.19) s et (1.77±0.18) s pour les mêmes bases de mammographies. Cependant, le temps de calcul de notre technique est bien en deça de celui des ondelettes de Daubechies. Globalement, on observe que notre méthode est capable d'atteindre des performances de suppression de bruit similaires à celles obtenues avec les meilleurs algorithmes de l'état-de-l'art et ce deux fois plus rapidement.

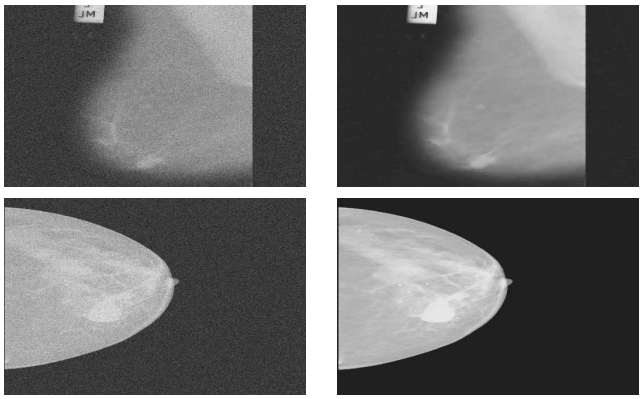
2.7 Extension aux mammographies entières

Le dictionnaire appris sur chaque base d'images a été ensuite utilisé pour tenter de supprimer le bruit dans chaque mammographie. Les résultats de l'application de notre approche sur des images non bruitées de façon numérique sont montrées à la figure 2.6. Sur cette figure, on peut distinguer une amélioration du contraste due à la normalisation de l'histogramme de chaque image et une réduction du bruit liée au système d'acquisition. La figure 2.7 montre les résultats obtenus après ajout d'un bruit Gaussien additif aux mammographies.



a) b)

FIGURE 2.6 – Suppression du bruit et amélioration de contraste sur l'image du haut provenant de la MIAS et celle du bas provenant du INbreast : (a) images originales non- bruitées et (b) résultats après application du K-SVD-ACP.



a) b)

FIGURE 2.7 – Suppression du bruit et amélioration de contraste sur l'image du haut provenant de la MIAS et celle du bas provenant du INbreast : (a) images originales bruitées et (b) résultats après application du K-SVD-ACP.

2.8 Conclusion

Dans ce chapitre, nous avons présenté une technique de suppression de bruit basée sur le K-SVD et l'ACP. Notre contribution dans cette partie est basée sur la réduction de la taille du dictionnaire appris sur une base d'entraînement. L'approche proposée (K-SVD-ACP) a montré une excellente qualité visuelle et une amélioration de contraste uniforme tout en conservant l'apparence des structures anatomiques présentes dans l'image, surtout sur les mammographies entières. Cet algorithme a surpassé la suppression du bruit basée sur les ondelettes de Daubechies, mais requiert un temps de calcul plus grand. Les performances notées à l'aide du K-SVD-ACP en ce qui concerne la qualité d'image sont quasiment similaires à celles obtenues avec le K-SVD classique et les *curvelet*. Cependant, le K-SVD-ACP a besoin d'un temps de reconstruction deux

fois moins long que le K-SVD et les *curvelet*. Par ailleurs, on a montré qu'on peut supprimer le bruit et améliorer le contraste de toutes les images contenues dans différentes bases d'images, en utilisant les mêmes paramètres ; ce qui résoud le problème lié au choix de paramètres tels que la taille des filtres et les valeurs des seuils, qui peuvent subvenir durant l'utilisation des ondelettes, des *curvelet* et des filtres passe-bas. Cela nous permet d'affirmer que nous disposons d'un outil entièrement automatique permettant donc d'éliminer toute interaction avec l'utilisateur, mais également très rapide pouvant donc être utilisé sans problème dans un contexte médical où toute perte de temps peu avoir des conséquences significatives sur les autres tâches.

Chapitre 3

Segmentation des masses et extraction des spicules

3.1 Segmentation des masses

La segmentation est un problème omni-présent en traitement d'images et en vision par ordinateur dont le but est d'isoler un objet du fond de l'image. Dans le cadre de la conception d'un système de DAOx, la segmentation joue un rôle central en ce sens que les autres étapes à savoir la description de la masse, la détection des structures d'intérêt comme les spicules et la classification des différentes lésions en dépendent fortement. Ainsi, plus la segmentation est précise, meilleure sera la description de la nature du contour, de la texture et de la forme de la masse, ce qui aura pour conséquence une bonne classification de la tumeur. La particularité de la segmentation réside dans le fait qu'il n'existe pas une méthode adaptée à toutes les applications. A ce titre, elle a fait l'objet de plusieurs publications, dans lesquelles les chercheurs ont tenté d'apporter des solutions diverses et variées qui tiennent souvent compte de la nature des images traitées.

De façon générale, les méthodes de segmentation sont regroupées en deux catégories que sont les approches supervisées et non-supervisées. Les approches non-supervisées réalisent une partition de l'image en plusieurs régions distinctes qui satisfont des propriétés liées au niveau de gris des pixels et à la texture. Ces approches sont également subdivisées en trois groupes que sont les méthodes basées régions, celles basées contour et celles basées *clustering*. Pour leur part, les approches supervisées – *c-à-d.* nécessitant une intervention humaine – sont fondées sur l'utilisation d'une information a priori sur la nature de l'objet à extraire et éventuellement le fond de l'image.

La plupart des méthodes de segmentation les plus efficaces proposent des formulations variationnelles dont le but est d'optimiser une fonction coût qui contient des expressions prenant en compte les descriptions des régions de l'image et leur contour. Parmi les formulations continues, les plus efficaces sont celles qui minimisent des fonctionnelles de courbes actives par le biais des *level sets* à travers une descente de gradient. Toutefois, étant donné que la descente de gradient converge vers un minimum local, les résultats obtenus ne sont pas toujours convaincants. A cela s'ajoute le fait que ces méthodes requièrent un temps de calcul assez important sans pour autant donner l'assurance d'avoir un résultat correct, *c-à-d.* tendre vers un minimum local proche de l'optimum global. Pour régler ce problème, une formulation discrète – qui conçoit l'image comme des fonctions discrètes sur une grille – a été proposée dans bon nombre de travaux de recherche. Parmi ces approches discrètes, les méthodes d'optimisation combinatoire qui font appelent aux algorithmes de coupe minimale sont les plus efficaces, car elles s'avèrent bien plus rapides et plus efficaces que la formulation continue, du fait de leur capacité à trouver un minimum local

qui est très proche du minimum global. Cependant, les méthodes d'optimisation par coupe minimale, malgré leur efficacité sur plusieurs types d'images, échouent à segmenter correctement certaines images médicales dont les mammographies, en raison de la complexité de celles-ci. Afin de gérer convenablement ces cas de figure, nous avons proposé une méthode de segmentation non-supervisée basée sur les champs de Markov et qui simpifie énormément les calculs réalisés dans le cadre de la convergence de notre fonction d'énergie, tout en faisant preuve d'une très grande rapidité et efficacité en comparaison des méthodes d'optimisation par coupure de graphe.

Après avoir exposé de façon précise les approches de segmentation les plus populaires – dans le cadre de l'imagerie du sein – tout en mettant l'accent sur leurs forces et faiblesses, nous mettrons en évidence les difficultés liées à l'exploitation des méthodes de segmentation existantes aussi bien d'un point de vue méthodologique, que sur le plan des performances attendues, ce qui nous permettra de déboucher sur nos attentes concernant une méthode de segmentation adaptée à la mammographie, surtout aux cas les plus compliqués. Suite à cela, nous détailleront la solution que nous avons proposée pour satisfaire ces attentes, tout en démontrant sa pertinence aussi bien sur la méthode que sur l'efficacité en mammographie. Nous finirons par une illustration de la méthode sur des mammographies pour lesquelles notre approche échoue à fournir une bonne segmentation ainsi qu'une discussion sur les avantages et les améliorations à envisager pour parfaire cette approche.

3.2 Etat de l'art des méthodes de segmentation

3.2.1 Segmentation basée région

Les méthodes de segmentation basées régions ont pour but de diviser une image en plusieurs régions homogènes disjointes et connectées ; elles reposent sur le principe d'homogénéité qui signifie qu'il doit y avoir au moins un descripteur qui, dans une certaine mesure reste uniforme pour tous les pixels appartenant à une même région de l'image. Les techniques basées régions peuvent être divisées en deux grands groupes que sont : les méthodes de croissance de régions et celles de division et fusion de régions.

3.2.1.1 Croissance de régions

Ces méthodes sont basées sur la propagation d'un point germe initial selon un critère spécifique d'homogénéité dans l'idée de faire croître itérativement une région de l'image représentant l'objet à segmenter, en agrégeant des pixels voisins qui ont des propriétés semblables au point germe. Ces techniques ont été beaucoup utilisées pour la segmentation des masses en mammographie, l'objectif étant d'extraire la masse du fond de l'image. En 1995, [Huo et al., 1995] ont développé un algorithme semi-automatique dans lequel l'étape de croissance de régions est automatiquement exécutée après que le radiologue ait manuellement sélectionné le point germe. Un peu plus tard, deux versions améliorées de cette technique ont été proposées par [Kupinski et Giger, 1998]. La première version fait intervenir le gradient d'indice radial qui est une mesure de la quantité moyenne de gradient radialement dirigé vers l'extérieur ; quant à la seconde, elle est batie sur une approche probabiliste qui consiste à modéliser la probabilité qu'un pixel appartienne à une région par une loi non Gaussienne tandis que celle d'appar-

tenir au fond de l'image est traduite par une loi de probabilité uniforme. Les auteurs ont ensuite montré à travers des expériences que l'approche probabiliste donne les meilleurs résultats. Dans le même ordre d'idée, [Martí et al., 2003] ont exploité les régions actives pour extraire la masse ; pour cela, ils ont modélisé la masse comme une fonction Gaussienne à laquelle ils ont incorporé des descripteurs de contour, de texture et de forme, tout en associant une distribution uniforme au fond de l'image. Un raisonnement différent a été mené par [Guliato et al., 1998] qui a proposé un algorithme de croissance de région basé sur la logique floue et qui tient compte du caractère incertain du contour des régions suspectes, en vue de préserver la transition entre la masse et le tissu sain. Une alternative à cette méthode a été proposée par [Petrick et al., 2002] qui a introduit l'information relative au gradient dans la croissance de régions, dans le but de réduire la fusion entre les structures adjacentes et celles qui se chevauchent. L'identification de l'ensemble optimal des points germe initiaux afin d'améliorer l'algorithme de croissance de régions a été explorée dans plusieurs travaux. [Zheng et al., 1995] ont proposé d'utiliser l'image des contours – obtenue en faisant la différence de deux images floutées issues de la convolution de l'image initiale par un filtre Gaussien – comme germe de départ. Quant à [Qi et Snyder, 1998], ils ont utilisé les courbes de Bézier pour interpoler les histogrammes à partir desquels ils ont déduit les valeurs des seuils comme étant des maxima locaux. [Zhang et al., 2004] et [Zhang et Say-Wei, 2006] ont de leur côté divisé le sein en plusieurs sous régions à l'intérieur desquelles les pixels d'intensité de niveau de gris maximale sont utilisés comme des points germes initiaux. Une autre piste consistait à rechercher le maximum local dans l'image à l'aide des opérateurs de morphologie mathématique comme préconisé par [Hejazi et Ho, 2005].

Pour ce qui est de l'algorithme fondé sur la division et la fusion [Chen et Pavlidis, 1979], il se divise en deux étapes. Premièrement, l'image de départ est itérativement décomposée en plusieurs sous régions jusqu'à ce que le critère d'homogénéité soit respecté, puis, tous les couples de régions voisines sont explorés de façon à fusionner ceux dont l'union respecte le critère d'homogénéité. [Rangayyan et al., 1997], se servant d'une région contenant la masse et sélectionnée manuellement, a utilisé cette technique pour réaliser une approximation du contour de cette région.

La ligne de partage des eaux (LPE) (*watershed*) [Beucher et Lantuéjoul, 1979] est une autre approche de segmentation basée régions, qui s'inspire des outils développés en morphologie mathématique. Cette notion est définie par analogie géographique comme le complémentaire des bassins versants, un bassin versant étant la zone associée à un minimum régional, tel qu'une goutte d'eau qui tombe dans cette région et suit la ligne de plus grande pente s'arrêtera dans ce minimum. La transposition de cette notion en terme mathématique pose de nombreux problèmes, comme la définition d'une ligne de plus grande pente, les plateaux et bien d'autres. La majeure partie des définitions sont algorithmiques, cela signifie que la LPE est définie d'après le moyen qui permet de la construire. L'une des méthodes les plus connues est l'algorithme *d'immersion* qui consiste à remplir les bassins versants pour déterminer leur limite [Vincent et Soille, 1991], [Vincent, 1992], [Soille, 2013]. Malgré son efficacité, la LPE engendre une sur-segmentation de l'image, ce qui en imagerie du sein a pour conséquence la dégradation du contour des masses surtout sur des échographies (qui sont caractérisées par du bruit, le *speckle* et des textures irrégulières). La solution la plus sophistiquée utilisée pour résoudre ce problème consiste à effectuer un marquage puis une reconstruction éventuellement précédée de pré-traitements ; dans ce cadre, les marqueurs sont définis comme des composants connectés de l'image dont les pixels doivent satis-

faire une condition de similarité ; ils doivent être déterminés à l'intérieur et à l'extérieur des objets que l'on veut segmenter. L'étape de reconstruction permet ensuite de retrouver chaque objet à partir de son marqueur. Elle doit satisfaire deux conditions : préserver le nombre de composantes connexes et rester dans l'espace de départ. Ainsi, en mammographie, les travaux de recherche qui ont été menés aussi bien pour la détection que la segmentation des masses ont tous utilisé cette alternative : [Huang et Chen, 2004] utilisent le filtre passe-bas de Wiener afin de supprimer le bruit et de générer des marqueurs ayant une bonne précision ; [Gulsrud et al., 2006] ont d'abord amélioré le contraste de l'image en réalisant une égalisation d'histogramme [Pizer et al., 1987] sur chacune des fenêtres glissantes extraites de l'image à segmenter ; puis, ils procèdent à un lissage morphologique de l'image précédente par l'utilisation d'un filtre alterné séquentiel, en réalisant une ouverture morphologique suivie d'une fermeture, dans le but d'atténuer les pics étroits et les vallées étroites d'une certaine taille. Enfin, sur cette dernière image, ils calculent le gradient morphologique, dont ils se servent pour définir un marqueur interne, qui ici est le centre de la masse et un marqueur externe placé dans le fond de l'image englobant la masse. [Gomez et al., 2010] procèdent dans un premier temps à une convolution de l'image par huit filtres de Newton, ce qui aboutit à une image de gradient qui améliore le contour de l'image ; ensuite, un seuillage de l'image initiale est réalisé en vue d'obtenir des régions binaires sur lesquelles les marqueurs interne et externe sont calculés par l'intermédiaire de l'érosion et de la dilatation ; enfin une reconstruction morphologique est réalisée sur les deux types d'images obtenues.

La difficulté liée à la croissance de régions réside dans la définition d'un critère d'arrêt pertinent permettant une segmentation optimale des régions à segmenter. En effet, si le point germe initial est selectionné dans une région non homogène, cette méthode aura tendance à produire de fortes variations, ce qui aboutira à un arrêt rapide de la croissance de la région. De plus, l'intervention du radiologue pour sélectionner le point germe pour chacun des objets à segmenter, est un inconvénient majeur de ces algorithmes. Quant au concept de la division et de la fusion, il souffre de deux limites majeures que sont : le manque de précision sur la forme des objets segmentés en raison de la forme carrée introduite par la segmentation sur le contour des objets, et sa sensibilité au parcours des régions. Enfin, pour ce qui est de la LPE, le choix des méthodes de pré-traitement et des bons marqueurs est indispensable pour éviter la sur-segmentation ; à cela s'ajoute le manque de régularité, qui peut être résolu dans une certaine mesure par une inondation faite de manière *visqueuse* [Vachier et Meyer, 2005].

3.2.1.2 Segmentation basée contour

Historiquement, les approches les plus connues de détection des contours se basent sur le gradient pour détecter les transitions entre les structures à segmenter et le fond de l'image. Le contour est alors exprimé comme le lieu des maxima du gradient dans la direction du gradient. Parmi ces approches classiques, celles utilisant les détecteurs de gradient par masques comme les opérateurs de Roberts [Roberts, 1963], Prewitt [Prewitt, 1970] et Sobel [Sobel, 1970] sont les plus utilisées. En dépit de la simplicité de mise en oeuvre et la rapidité de calcul liées à ces masques, la qualité des contours qu'ils fournissent sur des images médicales est très médiocre. En effet, les contours extraits ne forment pas nécessairement un ensemble de courbes connectées et fermées qui entourent des régions distinctes. Ce qui implique l'utilisation de méthodes de poursuite ou de fermeture de contour via une connaissance a priori sur la forme de l'objet à

segmenter ou par interpolation afin de relier les points du contour. A la fin des années 70, les travaux de [Marr, 1976] sur la conception d'un détecteur de contour inspiré des systèmes biologiques ont constitué une réelle avancée ; ces méthodes fondées sur le passage par zéro du Laplacien permettent de bien mettre en évidence les extrémum de la dérivée, ce qui garantit l'obtention d'un réseau de lignes fermées [Marr et Hildreth, 1980]. S'inspirant de ces méthodes, les approches analytiques issues des travaux de [Canny, 1986] et [Deriche, 1987], fondées sur les critères de bonne détection, de bonne localisation et d'unicité du contour ont permis d'avoir des détecteurs de très bonne qualité. Toutefois, l'identification des régions de l'image sur la base de l'information contenue sur leur contour est loin d'être triviale en raison de l'absence de l'incorporation par ces algorithmes de l'expérience et de l'information contextuelle utilisées par la vision humaine.

Les premiers travaux sur la détection des contours des masses sont basés sur l'utilisation d'un détecteur *Laplacien of Gaussian* suivi d'une amélioration locale du contraste [Petrick et al., 1995], [Petrick et al., 1996]. [Parr et al., 1994] ont utilisé les filtres de Gabor pour localiser les lésions stellaires.

Dans l'idée d'éviter les effets indésirables liés au filtrage des images, certains travaux ont été menés par plusieurs équipes de recherche. C'est ainsi qu'une représentation topographique de la mammographie a été proposée par [Hong et Brady, 2003] dans le but de détecter les régions suspectes de l'image. En seuillant différents niveaux de topographie, ils ont été capables de trouver divers types de régions comme la densité mammaire ou les masses. Une approche similaire a été proposée par [Yin et al., 2003] qui ont utilisé un maillage intelligent pour retrouver les masses. [Raso et al., 2004], [Fauci et al., 2005] et [Cascio et al., 2006] ont recherché les contours des masses en se servant d'un algorithme itératif dans lequel pour chaque maximum local, un seuil est selectionné pour dessiner un contour iso-dense. La valeur du seuil étant fournie par l'utilisateur sur la base de l'analyse de l'histogramme de l'image.

Dans certains travaux, l'information du contour a été utilisée pour améliorer les résultats de la segmentation initiale ; ce fut le cas de [Kobatake et al., 1999], [Sahiner et al., 2001b] et [Sahiner et al., 2001a] qui se sont servis des contours actifs (*snakes*) comme dernière étape de leur algorithme. Quant à [Nakagawa et al., 2004], ils se sont servis de deux images pour calculer les forces de leur *snakes*, l'une relative à l'intensité du contour et l'autre basée sur l'intensité des niveaux de gris. [Allen et al., 2003] et plus récemment [Yuan et al., 2007], [Shi et al., 2008] ont exploité les lignes de niveaux (*level sets*) pour trouver avec précision le contour des masses. [Timp et Karssemeijer, 2004] ont réussi à trouver un bon contour de la masse en exploitant une technique d'optimisation basée sur la programmation dynamique ; leur approche utilise à la fois l'information relative au contour de la masse, de même qu'une connaissance a priori sur la distribution des valeurs de niveaux de gris des pixels appartenant à la ROI centrée sur la masse à segmenter. Ils ont ensuite prouvé la supériorité de leur méthode comparée au travail de [Kupinski et Giger, 1998] et celui fondé sur un modèle de contour discret [te Brake et Karssemeijer, 1999].

3.2.1.3 Segmentation basée *clustering*

Les techniques de *clustering* [Jain et al., 1999] sont des méthodes courantes en segmentation d'image et peuvent être divisées en deux catégories à savoir les algorithmes de *partition* et ceux *hiérarchiques*. La principale différence est que les méthodes hiérarchiques produisent plusieurs

régions imbriquées les unes dans les autres, tandis que les méthodes de partition produisent une seule région.

L'algorithme de partition le plus connu est celui des K-Means [MacQueen et al., 1967] qui est à la fois facile à implémenter et possède une faible complexité. En mammographie, il a été utilisé dans le but de générer une segmentation initiale [Sahiner et al., 1996b], [Sahiner et al., 1998a]. Le résultat de cette segmentation ayant été affiné plus tard à l'aide des informations sur le contour des masses. [Li et al., 1997b], [Li et al., 1999] et [Li et al., 2002] ont pour leur part utilisé une généralisation des K-Means qui incorpore l'information spatiale [Pappas, 1992] afin d'améliorer la détection faite un peu plus tôt à l'aide des ondelettes dans un cadre de multi-résolution. [Qian et al., 1998b], [Qian et al., 1998a], [Qian et al., 1999] et [Qian et al., 2001] ont étudié l'influence des pré-traitements sur l'algorithme de seuillage proposé par Li *et al.* et leur conséquences sur la classification.

L'algorithme du *Fuzzy* C-Means (FCM) proposé par [Bezdek, 2013] et qui est une extension des K-Means, permet à chaque objet de l'image d'être associé à un cluster en utilisant une relation d'appartenance floue. En imagerie du sein, [Velthuizen, 2000] a utilisé le FCM afin de regrouper les pixels qui ont des valeurs de niveaux de gris similaires, tandis que [Chen et Lee, 1997] l'ont utilisé sur un ensemble de descripteurs extraits par l'intermédiaire des ondelettes après segmentation par champs de Markov (*MRF*). La force des MRFs est leur capacité à modéliser les probabilités jointes comme des interactions spatiales locales, introduisant de ce fait dans l'algorithme l'information relative aux dépendances locales entre les pixels [Bishop, 2006]. Cette information a été exploitée par [Rogova et al., 1999] qui a proposé un algorithme de relaxation stochastique contrainte par une fonction de disparité, qui estime la similarité entre deux clusters de pixels dans l'espace des descripteurs. Contrairement à Rogova *et al.*, [Cao et al., 2004a] et [Cao et al., 2004b] ont proposé deux méthodes de segmentation des masses par clustering basées sur des outils de la théorie de l'information ; la première technique est le recuit simulé déterministe [Rose, 1998] qui a pour particularité d'incorporer l'aléatoire dans la fonction coût à minimiser. Dans la seconde méthode, les auteurs ont combiné le FCM et le recuit simulé déterministe afin d'améliorer l'algorithme précédent. [Suliga et al., 2008] ont quant à eux proposé une méthode de segmentation s'inspirant des MRFs et basée cluster ; dans ce papier, ils ont défini une nouvelle énergie de vraisemblance qui est la somme des potentiels de vraisemblance, où le potentiel de vraisemblance est la distance euclidienne entre un pixel et sa classe d'affectation après clustering. Ils ont ensuite incorporé les potentiels des cliques, histoire de se servir de la cohérence spatiale qui existe entre les pixels. Pour finir, deux nouveaux paramètres ont été proposés par les auteurs afin de normaliser les potentiels et d'améliorer l'homégénéité intra-classe. Bien que dans un tout autre régistre, nous avons jugé utile de mentionner le travail de [Brzakovic et al., 1990] qui constitue l'un des premiers travaux sur la détection et la segmentation des masses en imagerie du sein ; la méthode proposée par les auteurs est basée sur un concept de multi-résolution utilisant une pyramide de fusion floue. Cette pyramide représente une structure de données dans laquelle l'image de départ est la base à l'aide de laquelle chaque couche située au niveau supérieur mais de résolution inférieure à la base est construite. Les liaisons entre les couches filles et parents sont propagées grâce à une fonction floue ; aini la relation spatiale est prise en compte.

En nous servant des travaux de [Fu et Mui, 1981] comme repère, nous avons décidé de classer la segmentation par seuillage dans la catégorie des méthodes de *clustering*. La plupart des méthodes sont basées sur les intensités de niveaux de gris des pixels d'une mammographie ; c'est

ainsi que [Abdel-Dayem et El-Sakka, 2005] ont trouvé le meilleur seuil pour détecter les masses en minimisant une fonction d'entropie floue de l'image. Pour leur part, [Matsubara et al., 1997] et [Matsubara et al., 1998] ont utilisé différentes valeurs de seuils en fonction des tissus à segmenter dans le sein ; lesquelles valeurs étaient déterminées suite à une analyse de l'histogramme. Les résultats de ce travail ont par la suite été améliorés par [Goto et al., 1998]. [Özekes et al., 2005], se sont eux servis d'un seuillage directionnel pour estimer la forme de la masse ; pour cela, ils ont balayé les pixels de l'image suivant huit directions en utilisant plusieurs seuils. Enfin, à l'aide d'un *template* d'une masse circulaire, ils ont pu trier les masses parmi les objets détectés. [Sameti et Ward, 1996] et [Sameti et al., 1997] ont eux introduit dans l'algorithme de seuillage une relation spatiale entre les pixels voisins ; pour y arriver, ils ont dans un premier temps séparé l'image en plusieurs ROIs, puis une fonction d'appartenance floue a été attribuée à chaque pixel contenu dans une ROI. Pour finir, à chaque itération, ils calculent une valeur d'erreur dont ils se servent pour effectuer une mise à jour de la valeur d'appartenance.

Dans certains travaux, les auteurs ont jugé utile de réaliser le seuillage sur l'image pré-traitée plutôt que sur l'image originale ; c'est le cas de [Kobatake et Murakami, 1996] qui après avoir utilisé le filtre d'*iris* conçu pour améliorer le contraste des opacités de forme circulaire, ont réalisé un seuillage adaptatif pour détecter les masses. Une approche similaire a été utilisée par [Varela et al., 2007]. Cependant, d'autre auteurs ont préféré extraire dans un premier temps des descripteurs de l'image, sur lesquels un seuillage est ensuite réalisé. C'est le cas de [Heath et Bowyer, 2000] qui ont developpé une méthode de détection des masses basée sur un filtre nommé *average fraction under minimum* (AFUM) qu'ils ont conçu pour vérifier la diminution en intensité de niveaux de gris des pixels d'une fenêtre centrée en un point de l'image. L'image obtenue après application de ce filtre est enfin filtrée dans le but d'identifier des masses. Une autre vision des choses consiste à réaliser un seuillage sur le résultat de la soustraction entre l'image originale et celle obtenue après pré-traitement. A titre d'exemple, [Polakowski et al., 1997] a réalisé la soustraction de deux versions filtrées – par deux filtres Gaussiens d'écart-type différents – de l'image originale, suivie d'un seuillage sur l'image issue de la soustraction. [Catarious Jr et al., 2004] ont quant à eux proposé l'utilisation d'une procédure itérative pour affiner le résultat de la soustraction. Plus tard, ils ont étudié l'impact des choix des différentes valeurs d'écart-type à la fois sur le taux de vrais et de faux positifs dans l'image mais également sur la description des régions suspectes détectées [Catarious Jr et al., 2006]. Pour leur part, [Kom et al., 2007] ont préféré soustraire l'image originale de celle obtenue après une transformation linéaire du contraste de l'image de départ.

Néanmoins, les méthodes de partition présentent deux inconvénients majeurs : elles se doivent de connaître le nombre de régions de l'image et elles ne prennent pas en compte la cohérence spatiale qui existe entre les pixels de l'image. Pour ce qui est des approches de seuillage, l'inconvénient majeur est le choix de la valeur du seuil ou de l'intervalle de seuillage ; certes il existe des méthodes de détermination automatique du seuil [Otsu, 1975], [Reddi et al., 1984] mais elles échouent presque systématiquement dans plusieurs applications [Whatmough, 1991] notamment en mammographie ou il n'existe pas une franche séparation entre les différentes classes d'intensité de niveaux de gris des pixels qui composent les divers tissus du parenchyme mammaire.

3.2.2 Segmentation basée modèle

La segmentation basée modèle se définit comme une méthode qui possède une phase d'apprentissage durant laquelle les objets à détecter sont appris de façon automatique par le système de DAOx. Ce qui permet ensuite au système d'être en mesure de détecter ou de classer de nouvelles images en tenant compte de la présence ou non d'objets similaires à ceux observés durant la phase d'entraînement. Durant cette étape d'apprentissage, le système apprend sur les images contenant des masses, leur probable situation ainsi que la variation en terme de forme et de taille des masses. Sur les images de la base d'entraînement où ne figurent pas des lésions, le système peut apprendre des descripteurs qui représentent les tissus sains.

Au nombre des méthodes de segmentation basées modèles, le *pattern matching* fait partie des plus utilisées ; dans cette approche, la phase d'apprentissage se fait sur des images qui contiennent les objets à détecter. Elle a été exploitée dans bon nombre de travaux [Lai et al., 1989] et [Constantinidis et al., 2001], où les auteurs ont utilisé la distance de la corrélation croisée normalisée (NCC) comme mesure de similarité. Une autre mesure de similarité – à savoir, l'information mutuelle – permettant de vérifier si une ROI contient une vraie masse, a également été utilisée [Tourassi et al., 2003]. De leur côté, [Freixenet et al., 2008] ont proposé d'utiliser un modèle probabiliste de *template matching* afin de détecter les masses. Dans ce travail, la forme et les déformations de l'atlas déformable sont apprises sur des exemples de vraies masses ; à la suite de quoi, un schéma Bayésien a été exploité pour recaler l'atlas déformable appris sur les réels contours des objets contenus dans une mammographie. D'autre part, une technique similaire a été utilisée par [Hatanaka et al., 2001] pour détecter les masses avec une perte partielle de certaines régions ; c'est une situation qui est observée pour les masses se situant entre le muscle pectoral et le parenchyme mammaire. [Chang et al., 1996a], [Chang et al., 1996b] ont pour leur part établit cinq règles qui leur permettent de détecter les régions suspectes ; cette technique est ensuite combinée avec la méthode de [Zheng et al., 1995] qui fait appel aux opérateurs logiques.

Dans un tout autre registre, il existe enfin des approches basées sur l'utilisation des réseaux de neurones artificiels comme classifieurs. Dans ces genres de méthodes, le problème de la segmentation est formulé comme la classification d'une ROI comme bénigne ou maligne ; les descripteurs impliqués sont d'habitude ceux liés à l'intensité ou à la texture d'un ensemble de ROIs contenant des masses et des échantillons de tissus normaux. Par exemple, [Stathaki et Constantinides, 1994] ont entrainé un classifieur *artificial neural networks* (ANN) en se servant de descripteurs extraits d'un modèle auto-régressif bi-dimensionnel. [Wei et al., 1995] a pour sa part utilisé un réseau de neurones à convolution (voir chapitre sur la description des masses et l'aide à la décision) directement sur l'image initiale au lieu de se servir d'un ANN qui prend en entrée des descripteurs extraits manuellement comme c'est le cas de [Diahi et al., 1996]. Avec l'essor des techniques d'apprentissage profond ou *deep learning*, certains chercheurs se sont intéressés à leur l'utilisation en mammographie. Ainsi, [Dhungel et al., 2015c] et [Dhungel et al., 2015b] ont proposé une approche de segmentation des masses combinant à la fois le *structured support vector machine* (SSVM) [Tsochantaridis et al., 2005], [Szummer et al., 2008] et le *deep learning*. Dans leur travail, le SSVM apprend une forme structurée, qui est censée correspondre à la segmentation de la masse contenue dans l'image d'entrée. Pour se faire, les auteurs ont défini trois énergies ; la première est basée sur l'information a priori – à propos de la forme, de la situation et de la taille des masses – présente dans la base d'entraînement contenant la vérité terrain ;

la deuxième qui utilise un mélange de Gaussiennes (GMM) et la dernière qui est calculée à partir du *deep belief networks* (DBN) [Hinton et Salakhutdinov, 2006]. Dans le même régistre, les mêmes auteurs ont proposé une méthode de détection des masses en mammographie qui est basée sur un apprentissage par cascade combinant le *deep learning* et le *random forests*. La technique proposée en [Dhungel et al., 2015a], [Dhungel et al., 2015d] consiste à faire une première détection à l'aide du DBN ; les pixels des régions obtenues sont ensuite classifiés à l'aide du GMM pour réduire les faux positifs. Sur les régions restantes, ils appliquent une cascade de réseau de neurones à convolution dont les descripteurs sont appris par un classifieur SVM, afin de réduire d'avantage le nombre de faux positifs. Pour finir, ils utilisent des descripteurs extraits manuellement sur les régions obtenues en sortie de l'étape précédente afin de réaliser un apprentissage par cascade de *random forests* dont le but est de réduire de façon significative le nombre de faux positifs.

Le réel inconvénient des approches de segmentation basées modèles, est le fait qu'elles soient incapables de prendre en compte la grande variation qu'on peut noter dans les formes des masses en mammographie, comme [Carneiro et al., 2015] et [Dhungel et al., 2015d] l'ont eux mêmes reconnus. De plus, leur performance et leur robustesse dépendent fortement de la taille et du contenu du jeu de données qui doit être le plus grand et le plus varié possible ; chose difficile à obtenir en imagerie médicale. La conséquence, c'est que la qualité de la segmentation des contours que ces approches fournissent est très grossière et ne permet pas d'évaluer de façon objective la nature de la masse. Ce qui évidemment ne nous convient pas. A cela s'ajoute le temps de calcul énorme nécessaire pour segmenter une image et le nombre exponentiel d'hyper-paramètres à trouver de façon convenable, entre 5 et 60 millions de paramètres en fonction du réseau de neurones choisi.

3.3 Segmentation supervisée ou non-supervisée

Dans notre cas, le problème de la segmentation se pose de façon soutenue du fait de son utilité pour la suite des traitements ultérieurs entrant dans le cadre de la conception du système de DAOx. Il urge donc de trouver une approche qui puisse permettre d'extraire de façon correcte les contours des masses. Considérant les difficultés évoquées pour chacune des approches de segmentation décrites plus haut, nous avons décidé de proposer une méthode de segmentation convenable à notre application ; pour cela, nous avons fait le choix de nous placer dans un cadre non-supervisé (ou presque). Dans la suite de cette section, nous allons tenter de justifier notre position.

Le contexte dans lequel nous nous plaçons alors, à savoir celui de la segmentation non-supervisée semble être selon nous, la meilleure des pistes pour plusieurs raisons. (***i***) En effet, contrairement aux bases d'images naturelles et des objets de la vie courante (*p. ex.* Caltech01, STL-10, CIFAR, MNIST, etc $\cdots$) qui sont disponibles en grande quantité et largement accessibles à tous, les images médicales sont régies par des codes de déontologies et des règles liées au respect de la vie privée des patients, qui font que leur diffusion est extrêmement réglémentée. Ceci explique le fait qu'il y ait peu de base d'images à accès publique et beaucoup de bases privées. (***ii***) A cela s'ajoute le fait que, dans le cas des mammographies, l'identification des structures d'intérêt, comme les masses et les spicules requiert une expertise que seuls les radio-

logues spécialisés en imagerie du sein détiennent ; ce qui n'est absolument pas le cas des objets du quotidien, pour lesquels n'importe qui peut fournir une vérité terrain sans risque de se tromper. (***iii***) La difficulté est encore bien plus grande en imagerie du sein et plus précisément en mammographie où 10 à 35% des cancers sont manqués par les spécialistes en raison de la pauvre qualité de l'image et surtout du manque d'expérience de certains radiologues. L'argument de l'expérience prend tout son sens, quand on se rend compte de la grande variabilité qui existe entre les différentes mammographies au vue de la forme, du contour des masses, de la répartition des micro-calcifications, des différentes textures de l'image et aussi des structures anatomiques qui peuvent induire en erreur même les spécialistes les plus expérimentés. C'est justement la raison pour laquelle les mammographies réalisées dans les programmes de dépistage sont suivies par deux spécialistes, et sans doute, ce qui explique l'absence totale d'une segmentation manuelle (hormis les coordonnées du centre de la masse) dans la MIAS, et une segmentation peu précise dans la DDSM. Il convient d'ajouter que la INBreast vient en quelque sorte régler le problème de la mise à disposition d'une vérité terrain très précise ; pour autant, elle contient très peu d'images pour être utilisée dans la conception d'un modèle supervisé capable de s'affranchir de la grande variabilité des masses.

Au vue de tous ces éléments, il parait judicieux de se tourner vers une approche non-supervisée qui respecte plusieurs conditions afin d'être une alternative crédible aussi bien dans un cadre de recherche, pour le développement d'un système de DAOx efficace, que dans des milieux hospitaliers où les compétences d'un radiologue spécialisé en imagerie du sein ferait défaut. Au nombre des conditions que nous jugeons indispensables pour une bonne méthode de segmentation en imagerie du sein, on a : le caractère non-supervisé de la méthode ; sa robustesse face aux cas les plus compliqués, qui impliquent une très grande densité du parenchyme mammaire et donc un faible contraste entre la masse et le reste du tissu mammaire ; une interaction quasi inexistante de l'utilisateur avec le système de segmentation, *c-à-d.* un seul (ou quasiment aucun) paramètre à spécifier, ce qui permettra d'être à la fois robuste et efficace quelque soit l'image à segmenter ; et enfin, le dernier et non le moindre, une méthode temps réel, à savoir une segmentation réalisée en moins d'une seconde.

3.4 Description du problème de la segmentation

Le choix d'une modélisation d'image doublement stochastique – c'est-à-dire que la composante markovienne est *non-observable* (cela signifie que les interactions entre les pixels ne se traduisent pas par des mesures calculables par des systèmes d'acquisition et observables sur une image) – dont au moins l'une des composantes est markovienne, se justifie par la capacité de la composante markovienne à prendre en compte les interactions spatiales locales entre les pixels voisins qui composent les différentes parties de l'image à travers la spécification des énergies locales ; ce qui se traduit par l'uniformité d'une région ou le caractère continu du contour d'un objet de l'image. La conséquence, c'est que les approches markoviennes fonctionnent généralement bien pour diverses applications ; ainsi, depuis plus de deux décennies, plusieurs travaux exploitant divers MRFs pour la segmentation et la détection des masses ont été publiés.

De toute évidence, le point critique dans le développement des méthodes de segmentation basées MRF est la détermination d'un compromis approprié entre le côté sophistiqué du modèle et

la praticité de l'algorithme qui en résulte. A ce jour, les méthodes de segmentation par champs de Markov proposées en mammographie se servent de divers modèles d'images bien connues comme le champ logistique à plusieurs niveaux (MLL) [Dubes et al., 1990] utilisé par [Li et al., 1995] et [Szekely et al., 2006] ou le modèle de Potts exploité à titre d'exemple par [Suliga et al., 2008]. Ensuite, conformément à la nature probabiliste des champs de Markov, le maximum a *posteriori* (MAP) est généralement utilisé pour réaliser la segmentation. Malheureusement, la maximisation de la vraisemblance a *posteriori* est rendue difficile par la structure non-linéaire et la nature discrète de certaines variables inconnues. En vue de réaliser une réelle optimisation, on peut se tourner vers le recuit simulé [Geman et Geman, 1984] (SA) – à l'instar de [Comer et al., 1996] –, qui a l'avantage de converger vers un minimum global (néanmoins, [Greig et al., 1989] ont prouvé qu'en pratique, le SA donne des résultats très éloingés du minimum global), mais implique généralement des calculs importants et complexes. C'est la raison pour laquelle, certains auteurs ([Li et al., 1995], [Szekely et al., 2006] et [Suliga et al., 2008]) se contentent d'une approche simple comme les modes conditionnels itérés (ICM) [Besag, 1986], qui est une méthode sous-optimale qui converge vers un minimum local. Cependant, le minimum local peut se retrouver très éloigné du minimum global, ce qui aboutit à des résultats qui ne respectent aucune propriété de l'image initiale. Donc, quand les résultats de la segmentation ne sont pas satisfaisants, cela peut être le fait d'un mauvais choix de la fonction d'énergie ou d'une solution éloignée de l'optimum global. A cela s'ajoute le fait que les techniques d'optimisation locale sont sensibles à l'initialisation.

De façon un peu plus détaillée, [Li et al., 1995] ont – afin d'affiner leur algorithme de détection des masses conçu dans un cadre multi-échelle – proposé une technique qui modélise la formation d'une région de l'image par un champ MLL. Enfin, pour l'estimation du MAP de leur MRF, ils ont utilisé l'ICM en raison de sa rapidité. Quant à [Szekely et al., 2006], après avoir utilisé un arbre de décision pour classifier comme bénigne ou maligne des fenêtres glissantes extraites un peu plus tôt sur une mammographie, ont procédé de la même façon que Li *et al.* pour améliorer leur résultat. Après avoir procédé à un rehaussement du contraste de l'image à l'aide d'un filtre morphologique, [Li et al., 1996], [Li et al., 1997a] et [Li et al., 2001] ont modélisé l'histogramme de la mammographie à l'aide d'un mélange de Gaussiennes (*finite generalized Gaussian mixture*, FGMM) avant de se servir de l'algorithme d'Espérance-Maximisation (EM) [Dempster et al., 1977] pour déterminer à la fois les paramètres et la forme du noyau du FGMM. Ils ont ensuite évalué les performances de ces diverses valeurs à l'aide de la relaxation globale de l'entropie (GRE). Par ailleurs, dans le souci de trouver automatiquement le nombre de régions de l'image, ils ont utilisé deux outils de la théorie de l'information à savoir le critère d'information d'*Akaike*(AIC) [Akaike, 1974] et le *minimum description length*(MDL) [Rissanen, 1978]. Pour finir, ils se sont servis de l'algorithme d'étiquetage de relaxation Bayésienne [Hummel et Zucker, 1983] – qui prend en compte les interactions locales entre les pixels voisins dans le processus d'étiquetage – afin d'améliorer leur segmentation. Pour leur part, [Zhen et Chan, 2001] ont conçu une approche de multi-résolution basée sur les MRFs (MMRFs) qui combine plusieurs techniques pour la détection des masses. Dans un premier temps, ils ont localisé les régions suspectes de l'image à l'aide d'une analyse fractale ; puis, ils ont utilisé des ondelettes à trois niveaux, afin d'extraire des descripteurs de l'image. Ensuite, ils ont utilisé une version modifiée du modèle de Potts, pour laquelle l'estimation du MAP est réalisée grâce à l'algorithme de l'ICM. [Suliga et al., 2008] ont proposé une méthode de *clustering*

basée sur les MRFs ; pour cela, ils ont défini une nouvelle énergie de vraisemblance comme la somme des potentiels de vraisemblance (LP), où chaque LP équivaut à la distance euclidienne entre un pixel et sa classe après le *clustering*. Puis, ils ont incorporé les potentiels des cliques à la manière du modèle de Potts et ce de façon similaire à l'approche de [Zhen et Chan, 2001], dans le but d'exploiter la cohérence spatiale qui existe entre les pixels. Pour finir, ils ont introduit deux paramètres dans la fonction d'énergie en vue de normaliser les potentiels et d'améliorer l'homogénéité intra-cluster ; pour ce qui est de l'optimisation il suit le même principe que l'ICM. [Comer et al., 1996] ont utilisé un échantillonneur de Gibbs pour générer un MRF, pour lequel l'algorithme EM est utilisé pour estimer les paramètres du maximum de vraisemblance. Par la suite, ils ont utilisé le *maximizer of the posterior marginals* (MPM) [Marroquin et al., 1987] pour affecter les pixels aux différentes classes. Dans un cas, comme dans l'autre, aucune de ces deux techniques – SA et ICM – n'apporte une réponse au problème capital de l'estimation des hyper-paramètres d'un modèle markovien ; car la détermination des paramètres du modèle est soit difficile et chronophage ou numériquement impossible à faire.

Dans l'objectif de résoudre de façon efficace le problème de l'optimisation de l'énergie, [Boykov et al., 2001] a considéré plusieurs cas de fonctions d'énergies qui préservent les discontinuités présentes dans une image ; parmi lesquelles les fonctions d'énergie utilisées dans le cadre d'un étiquetage Bayésien des champs de Markov de premier ordre. Pour y arriver, ils ont proposé deux techniques d'optimisation d'énergie par coupes minimales qui ont l'avantage de fournir un minimum local proche de l'optimum global et qui reposent sur deux grands mouvements que les auteurs appelent l'$\alpha-\beta-$*expansion* et $\alpha-$*swap*. Les deux mouvements en question peuvent changer simultanément les étiquettes des pixels de l'image ; ce qui n'est pas le cas du SA et de l'ICM, qui se caractérisent par un petit déplacement durant lequel seul un pixel change d'étiquette à la fois. Comme indiqué plus haut, la recherche de ces minima locaux n'est pas aisée et peut nécessiter des itérations exponentielles d'*expansion* et de *swap* avant d'aboutir à une solution correcte. C'est pour résoudre ce problème que les auteurs ont proposés d'utiliser les techniques de coupes minimales, afin de trouver les mouvements optimaux d'*expansion* et de *swap*. En dehors de la convergence vers un minimum local proche du minimum global, les deux algorithmes proposés par [Boykov et al., 2001] sont beaucoup plus rapides que le SA et permettent des mouvements beaucoup plus important dans le paysage énergétique. Cette approche a permis de développer des modèles de segmentation non-supervisée par coupure minimale surtout grâce au fait que le terme d'attache aux données peut être écrit sous la forme exigée par les algorithmes de coupure de graphe, et ce en utilisant un modèle constant par morceaux ou sa généralisation Gaussienne. Toutefois, en dépit de leur utilité, ces modèles ne sont pas toujours adaptés au contenu réel des images, où peuvent cohabiter des distributions Gaussiennes, de Gamma, ou de Whishart [Goudail et Réfrégier, 2004], comme c'est le cas avec les images médicales. Il est toujours possible de pallier à ce problème en autorisant une interaction de l'utilisateur avec le système ; *c-à-d.* une segmentation supervisée (*p. ex.* [Boykov et Jolly, 2001], [Blake et al., 2004], [Rother et al., 2004] ou l'algorithme *random walks* de [Grady, 2006]), mais la conséquence c'est l'obtention d'une modélisation à la fois difficile et chronophage.

Ces éléments indiquent que la plupart des méthodes de segmentation évoquées un peu plus tôt sont loin d'être adaptées à une utilisation dans un environnement médical, à cause de leur manque de simplicité, de robustesse, d'efficacité et également parce qu'elles sont généralement supervisées. Donc, bien que les méthodes à base de MRF sont potentiellement très puissantes

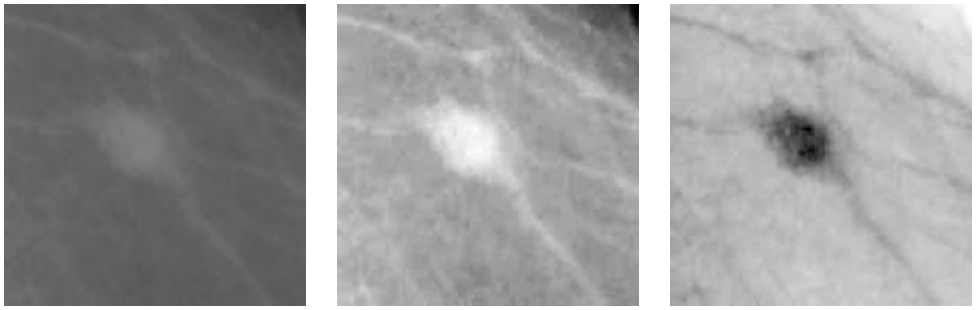
elles impliquent des calculs d'une grande complexité, tout en étant chronophages. Il urge donc de disposer d'une méthode simple et rapide fondée sur des principes statistiques solides comme l'ont également préconisés les auteurs de l'ICM dans l'un de leurs travaux [Geman et al., 1987], où ils ont proposé une formulation très générale du problème de l'estimation non-supervisée des paramètres du modèle markovien et une discussion sur sa complexité. L'utilisation de l'approche que nous proposons peut donc être vue comme une réponse efficace apportée au souhait formulés par ces auteurs. Dans ce chapitre, nous avons tenté de résoudre certaines de ces limitations en utilisant un MRF particulier connu sous le nom de champs aléatoires de Pickard (PRF) [Pickard, 1977] et [Pickard, 1980], qui à notre connaissance n'a jamais été utilisé pour la segmentation des masses en imagerie du sein. De plus, comme on le verra dans la suite, le modèle PRF est adapté à la mise en place de méthodes de segmentation non-supervisée, de sorte que le seul paramètre à fournir est le nombre de niveaux identifiés dans l'image ; par ailleurs, à l'inverse des autres méthodes de segmentation, la procédure d'initialisation est simple et, pour finir, le coût de calcul s'en trouve significativement réduit comparé aux autres approches basées MRFs existantes.

3.4.1 Pré-traitement des données

Les mammographies sont souvent caractérisées par un faible contraste et un contenu relativement bruité en raison de la projection des structures anatomiques 3D dans un plan 2D. Ces défauts ont pour inconvénient de masquer la lésion ; c'est pour cette raison que nous avons décidé de réaliser un simple rehaussement de contraste dans le but de normaliser les mammographies qui sont sujettes à une forte variabilité de leur contenu. Le pré-traitement proposé se décompose comme suit : dans un premier temps, nous avons fait un étirement d'histogramme, de manière à ce que la nouvelle dynamique de notre mammographie occupe l'intervalle $[c_{\min}, c_{\max}]$, où $c_{\min}$ est le 5eme centile et $c_{\max}$ le 95eme centile des intensités de pixels de l'image initiale. Ces quantités sont déterminées de façon empirique et nous permettent d'éliminer les valeurs extrêmes d'intensité de pixels qui ont une faible occurence – par rapport aux autres pixels de l'image – et qui affectent de façon significative l'opération d'étirement d'histogramme. Ensuite, une transformation logarithmique est appliquée sur les valeurs d'intensité de pixels étirées obtenues à l'étape précédente. L'expression exacte de la transformation logarithmique prend la forme suivante :

$$\mathbf{Z} = a \log_{10}(b + c\mathbf{Y}) \tag{3.1}$$

où $\mathbf{Y}$ et $\mathbf{Z}$ sont respectivement l'image originale et celle obtenue après rehaussement de contraste et où a, b et c sont des constantes de valeurs respectives 1, 10 et -9. Ces valeurs sont choisies de façon à passer de la dynamique $[0, 255]$ à $[0, 1]$ tout en effectuant un rehaussement de contraste en particulier sur les pixels de forte intensité qui sont susceptibles de correspondre aux régions denses du sein. Le fait d'effectuer l'inversion avant la transformation logarithmique nous permet d'étirer l'histogramme pour les grandes valeurs d'intensité de $\mathbf{Y}$, ce qui entraîne une amélioration de contraste des structures d'intérêt. Un exemple du résultat du pré-traitement est montré à la figure 3.1 ci-dessous.



(a) (b) (c)

FIGURE 3.1 – Amélioration de contraste : (a) image originale, (b) étirement d'histogramme et (c) transformation logarithmique de l'image originale.

3.4.1.1 Approche de segmentation proposée

Le type de MRF et des modèles d'observation ont un impact critique aussi bien sur la complexité des calculs impliqués dans l'optimisation que sur la qualité de la segmentation. Ici, notre choix est de sélectionner un MRF et des modèles d'observation simples et non triviaux, dans l'idée de réaliser une segmentation robuste, fiable et non-supervisée des mammographies. Pour le modèle d'observation, nous avons choisi un modèle Gaussien simple de type un-à-un ayant la forme ci-dessous :

$$\begin{array}{rcl} P(\mathbf{Z}|\mathbf{X}) & = & \prod_{i,j} P(Z_{i,j}|X_{i,j}) \ , \\ P(Z_{i,j}|X_{i,j}=n) & = & \mathcal{N}(m_n, r_n) \ . \end{array} \tag{3.2}$$

où $\mathcal{N}(m,r)$ désigne une distribution Gaussienne de moyenne m et de variance r et où i, j et k sont respectivement, l'indice des lignes, des colonnes et l'état de $X_{i,j}$. Le modèle (3.2) signifie essentiellement que les pixels de l'image observée sont tirés indépendamment des distributions Gaussiennes dont la moyenne et la variance caractérisent chaque région. Pour ce qui est du modèle du MRF, nous avons sélectionné le PRF qui a été étudié par Pickard [Pickard, 1977], [Pickard, 1980] et qui s'est révélé très utile pour la simplification du problème de la segmentation [Devijver et Dekesel, 1988]. Plus précisément, de tels champs sont stationnaires et leur distribution jointe est définie par une mesure homogène τ sur une cellule élémentaire de quatre pixels qui doivent satisfaire plusieurs conditions de symétrie et d'indépendance [Pickard, 1980]. [Champagnat et al., 1998] ont démontré que les MRFs stationnaires sur un réseau cartésien peuvent être caractérisés par leur distribution marginale sur une cellule élémentaire de quatre pixels, et dans certains cas importants – *c-à-d.* les champs Gaussien et symétrique –, les seuls champs stationnaires sont les PRFs. Pour réaliser la segmentation on se sert du MAP qui consiste à maximiser pour chaque pixel $\mathbf{X}_{i,j}$, la vraisemblance $P(X_{i,j}|\mathbf{Z})$. L'idée de la mise en oeuvre rapide d'une méthode de segmentation non-supervisée numériquement efficace, repose sur l'approximation suivante proposée par [Devijver et Dekesel, 1988] :

$$P(X_{i,j}|\mathbf{Z}) \approx P(X_{i,j}|Z_{i,.}, Z_{j,.}) \ , \tag{3.3}$$

où $Z_{i,.}$ et $Z_{j,.}$ désignent respectivement la i-ième ligne et la j-ième colonne de $\mathbf{Z}$. En faisant usage de la règle de Bayes sur l'équation (3.3), les propriétés d'indépendance conditionnelles entre les

pixels des champs de Pickard conduisent à l'expression simplifiée suivante :

$$P(\mathbf{X}_{i,j}|\mathbf{Z}_{i,.},\mathbf{Z}_{j,.}) \propto P(\mathbf{Z}_{i,.}|\mathbf{X}_{i,j})P(\mathbf{Z}_{.,j}|\mathbf{X}_{i,j})P(\mathbf{X}_{i,j}) \ . \tag{3.4}$$

La distribution marginale de chaque ligne et colonne de $\mathbf{X}$ possède la structure d'une chaîne de Markov stationnaire et réversible, et sous de faibles approximations, cette propriété peut être utilisée pour evaluer de façon efficace les deux premiers termes du membre de droite de l'équation (3.4), à l'aide d'algorithmes *forward-backward* [Devijver et Dekesel, 1988]. Il convient de souligner que le type de paramétrisation des chaînes de Markov a une grande influence sur le coût de calcul des procédures du *forward-backward*.

L'estimation du PRF et des paramètres du modèle d'observation $\mathbf{\Theta}$ qui contrôle la distribution et les mesures des chaînes de Markov représente un énorme défi. Cependant, vu le cadre probabiliste général adopté ici, *c-à-d.* (3.2) il est naturel d'avoir recours au *maximum de vraisemblance* (ML) qui possède de bonnes propriétés asymptotiques. L'estimée des paramètres à l'aide du ML est donc défini par $\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} P(\mathbf{Z},\mathbf{\Theta})$; toutefois, le problème de maximisation n'admet pas de solution explicite, ce qui nous contraint à utiliser des méthodes numériques d'optimisation. Grâce à la structure des chaînes de Markov sur les lignes et colonnes de $\mathbf{X}$, il est possible d'avoir recours à l'algorithme EM qui possède l'avantage d'allier les propriétés de convergence satisfaisantes et de mise en oeuvre raisonnable.

Partant d'une valeur initiale quelconque $\mathbf{\Theta}^0$, la méthode EM génère une suite d'estimées $\hat{\mathbf{\Theta}}^k$ en alternant les deux étapes suivantes :

$$\text{Espérance (E) : Exprimer } Q(\mathbf{\Theta},\hat{\mathbf{\Theta}}^k;\mathbf{Z}) \ , \tag{3.5}$$

$$\text{Maximisation (M) : } \hat{\mathbf{\Theta}}^{k+1} = \arg\max_{\mathbf{\Theta}} Q(\mathbf{\Theta},\hat{\mathbf{\Theta}}^k;\mathbf{Z}) \ , \tag{3.6}$$

où la fonction Q est définie comme :

$$\begin{array}{rcl} Q(\mathbf{\Theta},\mathbf{\Theta}';\mathbf{Z}) & \triangleq & E[J(\mathbf{X},\mathbf{Z};\mathbf{\Theta})|\mathbf{Z};\mathbf{\Theta}'] \ , \\ J(\mathbf{X},\mathbf{Z};\mathbf{\Theta}) & \triangleq & ln\ P(\mathbf{Z}|\mathbf{X};\mathbf{\Theta}) + ln\ P(\mathbf{X};\mathbf{\Theta}) \ . \end{array} \tag{3.7}$$

L'algorithme EM est une procédure itérative qui augmente la vraisemblance d'un vecteur de paramètres $\mathbf{\Theta}$ pour des observations $\mathbf{Z}$ à chaque itération, jusqu'à ce qu'il converge vers un point critique de la fonction de vraisemblance. Il faut noter que dans [Devijver et Dekesel, 1988], les distributions de $X_{i,.}$ et $X_{.,j}$ sont paramétrisées d'une manière standard par les probabilités initiales et de transition. Par conséquent, la représentation non-paramétrique choisie ne garantit pas la stationnarité et la réversibilité de chaque ligne et colonne de $\mathbf{X}$, et la mesure τ ne satisfait généralement pas les conditions nécessaires pour générer un champ de Pickard. Par ailleurs, le nombre de paramètres à estimer est de l'ordre de $O(n^2)$, ce qui alourdit de façon exponentielle les calculs et entraîne des difficultés de convergence de l'algorithme lorsque le nombre d'états augmente.

Dans le but d'apporter une solution efficace à ces difficultés théoriques et pratiques, Goussard *et al.* ont proposé une nouvelle paramétrisation *télégraphique* (TM) des champs de Pickard sur les lignes et les colonnes de $\mathbf{X}$ ainsi que sur les formules d'apprentissage correspondantes. L'apport de cette paramétrisation est double : dans un premier temps, elle garantit que

certaines propriétés théoriques du champ sont vérifiées ; puis, les techniques qui en découlent offrent une complexité numérique plus faible et une convergence plus sûre que l'approche décrite en [Devijver et Dekesel, 1988].

3.4.1.2 Paramétrisation *Télégraphique* du modèle

La paramétrisation TM adoptée par Goussard *et al.* est une généralisation d'une classe de chaînes de Markov à valeurs discretes détaillée en [Godfrey et al., 1980] pour la segmentation des signaux sismiques. Le vecteur des probabilités initiales $\boldsymbol{p}$ et la matrice des probabilités de transition $\boldsymbol{P}$ de chaque ligne et colonne sont paramétrisés par deux vecteurs $\boldsymbol{\lambda}$ et $\boldsymbol{\mu}$ tel que :

$$\boldsymbol{P} = \boldsymbol{\Lambda} + (\mathbf{1} - \boldsymbol{\lambda})\boldsymbol{\mu}^t,\ \boldsymbol{p} = (\boldsymbol{I} - \boldsymbol{\Lambda} + \boldsymbol{\mu}\boldsymbol{\lambda}^t)^{-1}\boldsymbol{\mu}\ , \tag{3.8}$$

avec $\boldsymbol{\lambda} \triangleq vect(\lambda_k)$, $\boldsymbol{\Lambda} \triangleq diag(\lambda_k)$, $\mathbf{1} = (1, ..., 1)^t$ et $\boldsymbol{I}$ la matrice d'identité. On peut vérifier que la matrice $diag(\boldsymbol{p})(\boldsymbol{P})$ est symétrique et que les conditions pour que $\boldsymbol{p}$ et $\boldsymbol{P}$ soient respectivement un vecteur de probabilités et une matrice des probabilités de transition valides sont données par :

$$\forall k = 1, ..., n, \mu_k > 0; \sum_{k=1}^{n} \mu_k = 1;\ \forall k, -\frac{\mu_k}{1 - \mu_k} \leq \lambda_k \leq 1\ . \tag{3.9}$$

Donc, tant que le vecteur de probabilités de l'état initial est égal à $\boldsymbol{p}$ et que les contraintes définies en (3.9) sont satisfaites, (3.8) définie une chaîne de Markov stationnaire et réversible, qui est paramétrisée avec $\boldsymbol{\theta} = \{\boldsymbol{\lambda}, \boldsymbol{\mu}\}$. L'étape de la segmentation est réalisée de la même manière qu'en [Devijver et Dekesel, 1988]. L'intérêt de la paramétrisation TM repose sur une simplification de l'algorithme *forward-backward* employé pour estimer les valeurs des marginales de vraisemblance $P(\mathbf{X}_{i,j}|Z_{i,.}, Z_{j,.})$. Chaque itération de l'algorithme nécessite le calcul des probabilités conditionnelles de la forme $P(X_{i,j}|X_{i-1,j})$ et $P(X_{i,j}|X_{i,j-1})$, qui dépendent de $\boldsymbol{P}$. Le fait d'exprimer $\boldsymbol{P}$ comme une fonction de vecteurs $\boldsymbol{\lambda}$ et $\boldsymbol{\mu}$ permet de faire passer la complexité du calcul de $O(n^2)$ à $O(2n)$.

Par ailleurs, la maximisation de la vraisemblance exacte $P(\mathbf{Z}|\boldsymbol{\Theta})$ ne donne pas une expression traitable des estimées des paramètres du modèle. Aussi, Goussard *et al.* ont approximé la vraisemblance exacte par une pseudo-vraisemblance $\hat{P}(\mathbf{Z}|\boldsymbol{\Theta})$ definie comme :

$$\hat{P}(\mathbf{Z}|\boldsymbol{\Theta}) \triangleq \prod_i P(\mathbf{Z}_{i,.}|\boldsymbol{\Theta}) \prod_j P(\mathbf{Z}_{.,j}|\boldsymbol{\Theta})\ . \tag{3.10}$$

Ainsi, un algorithme EM pour une maximisation de façon itérative de $\hat{P}$ peut être déduite. La difficulté principale se trouve à l'étape M qui, selon (3.5)-(3.7), consiste à maximiser la quantité Q definie par :

$$Q(\boldsymbol{\Theta}, \boldsymbol{\Theta}') \triangleq \sum_{\mathbf{X}_{(.)}} lnP(\mathbf{X}_{(.)}, \mathbf{Z}_{(.)}; \boldsymbol{\Theta}) P(\mathbf{X}_{(.)}|\mathbf{Z}_{(.)}; \boldsymbol{\Theta}')\ , \tag{3.11}$$

où $X_{(.)}$ et $Z_{(.)}$ sont respectivement une ligne et une colonne quelconque de $\mathbf{X}$ et $\mathbf{Z}$. En nous servant toujours des hypothèses faites plus haut, la maximisation de Q conduit à deux problèmes indépendants d'optimisation. Le premier s'applique aux paramètres de $\mathcal{N}(m_n, r_n)$ et, aussi longtemps que les conditions peu orthodoxes tiennent, on obtient en fait les formules de ré-estimation classique de Baum-Welch [Baum et al., 1970]. Le second problème concerne les

paramètres du modèle TM $\boldsymbol{\theta}$ dont la nouvelle valeur $\overline{\boldsymbol{\theta}}$ est obtenue par la maximisation de $Q(\boldsymbol{\theta}, \boldsymbol{\Theta}^{'}) = \sum_{\mathbf{X}_{(.)}} ln\ P(\mathbf{X}_{(.)}; \boldsymbol{\theta}) P(\mathbf{X}_{(.)} | \mathbf{Z}_{(.)}; \boldsymbol{\Theta}^{'})$. En utilisant la définition du modèle TM, Q peut s'écrire comme :

$$Q(\boldsymbol{\theta}, \boldsymbol{\Theta}^{'}) \propto \sum_{k=1}^{n} (\alpha_k\ ln\ \mu_k + \beta_k\ ln\ (1-\lambda_k) + s_k\ ln\ \frac{\lambda_k + \mu_k - \lambda_k \mu_k}{\mu_k - \lambda_k \mu_k}) - ln \left(\sum_{k=1}^{n} \frac{\mu_k}{1-\lambda_k} \right) \ , \tag{3.12}$$

où la forme générale des quantités α_k, β_k et s_k est donnée par :

$$\alpha_k \triangleq \sum_{j=1}^{J} P(\mathbf{X}_{(.,j)} = k | \mathbf{Z}_{(.)}; \boldsymbol{\Theta}^{'}) \ , \tag{3.13}$$

$$\beta_k \triangleq \sum_{j=2}^{J} P(\mathbf{X}_{(.,j)} = k | \mathbf{Z}_{(.)}; \boldsymbol{\Theta}^{'}) \ , \tag{3.14}$$

$$s_k \triangleq \sum_{j=2}^{J} P(\mathbf{X}_{(.,j-1)} = k, \mathbf{X}_{(.,j)} = k | \mathbf{Z}_{(.)}; \boldsymbol{\Theta}^{'}) \ . \tag{3.15}$$

Ces quantités peuvent être évaluées avec le même algorithme *forward-backward* déjà utilisé pour la segmentation. La maximization de Q en fonction de $\boldsymbol{\lambda}$ et $\boldsymbol{\mu}$ sous les contraintes exprimées à l'équation (3.9) ne conduit pas à une formule traitable. Cependant, le dernier terme du membre de droite de l'équation (3.12) devient plus petit en fonction du premier terme au fur et à mesure que la taille des lignes et des colonnes augmente. En négligeant ce dernier terme et en maintenant l'expression de Q invariant à l'inversion des indices, on aboutit à l'approximation suivante :

$$Q(\boldsymbol{\theta}, \boldsymbol{\Theta}^{'}) \approx \hat{Q}(\boldsymbol{\theta}, \boldsymbol{\Theta}^{'}) = \sum_{k=1}^{n} Q_k \ , \tag{3.16}$$

avec

$$Q_k = \eta_k\ ln\ \mu_k (1-\lambda_k) + s_k\ ln\ \left(1 + \frac{\lambda_k}{\mu_k (1-\lambda_k)} \right) \ , \tag{3.17}$$

où $\eta_k \triangleq (\alpha_k + \beta_k)/2$. La maximisation de $\hat{Q}(\boldsymbol{\theta}, \boldsymbol{\Theta}^{'})$ sous les contraintes d'égalité $\sum_{k=1}^{n} \mu_k = 1$ fait intervenir plusieurs étapes. Premièrement, notons que pour une valeur donnée de $\boldsymbol{\mu}$, la valeur optimale de $\boldsymbol{\lambda}$ à travers une maximisation indépendante de chaque Q_k. Ceci conduit

$$\overline{\lambda}_k = \frac{s_k/\eta_k - \mu_k}{1 - \mu_k} \ , \tag{3.18}$$

et $\overline{\lambda}_k$ satisfait nécessairement la troisième contrainte de l'équation (3.9). La substitution de (3.18) dans (3.17) nous permet d'exprimer Q_k comme une fonction de μ_k modulo un facteur constant :

$$Q_k = \gamma_k\ ln\ \frac{\mu_k}{1-\mu_k}, \quad \text{avec } \gamma_k \triangleq \eta_k - s_k \geq 0 \ . \tag{3.19}$$

Ensuite, la technique du multiplicateur de Lagrange est utilisée pour maximiser Q en fonction de $\boldsymbol{\mu}$ sous les contraintes de l'équation (3.9). En fixant le gradient du critère correspondant à zéro, on obtient :

$$\forall k \in [1, n],\ \nu\overline{\mu}_k^2 - \nu\overline{\mu}_k + \gamma_k = 0\ , \tag{3.20}$$

avec ν le multiplicateur de Lagrange. Quand $\nu > 4\gamma_k$, l'équation ci-dessus admet deux racines distinctes situées de part et d'autre de 1/2 et qui sont respectivement désignées par $\mu_k^+(\nu)$ et $\mu_k^-(\nu)$. Ainsi, 2^n combinaisons doivent être comparées. Cependant, dans l'objectif de satisfaire la condition décrite en (3.9), $\overline{\boldsymbol{\mu}}(\nu)$ doit seulement contenir $\mu_k^+(\nu)$. Ceci a pour conséquence de réduire le nombre de combinaison à $n + 1$. Par ailleurs, des études poussées sur les propriétés des combinaisons qui ne prennent en compte qu'un seul $\mu_k^+(\nu)$ montrent que $\hat{Q}$ est maximisée par une combinaison qui contient $\mu_{\overline{k}}^+(\nu)$, où le nombre d'états $\overline{k}$ est défini par :

$$\overline{k} \triangleq \arg\max_k \gamma_k\ . \tag{3.21}$$

Seules deux combinaisons possibles sont maintenues : celle composée de tous les $\mu_k^-(\nu)$ et celle contenant $\mu_{\overline{k}}^+(\nu)$. Une analyse plus approfondie de leur propriété révèle qu'une et une seule combinaison satisfait les contraintes de l'équation (3.9) et donc, elle peut être sélectionnée selon un critère d'inégalité très simple [Idier et Goussard, 1995]. Pour finir, la valeur du multiplicateur de Lagrange doit être déterminée. Elle ne peut pas s'exprimer sous une forme explicite, mais les bornes inférieure et supérieur très rapprochées peuvent s'obtenir très facilement. Les techniques classiques d'interpolation numérique peuvent dont être utilisées pour satisfaire la contrainte $\sum_{k=1}^n \overline{\mu}_k = 1$ avec une précision arbitraire.

En conclusion, la procédure de segmentation proposée comporte deux étapes : dans la première, les paramètres du modèle télégraphique sont déterminés par les formules de ré-estimation décrites plus haut. Par construction, le modèle est compatible avec les propriétés de Pickard sur la mesure τ. La seconde étape quant à elle, permet la segmentation de l'image $\mathbf{Z}$ en exploitant les paramètres estimés à l'étape précédente. La méthode de MAP utilisée est similaire à celle décrite en [Devijver et Dekesel, 1988]. Pour finir, notons que les deux étapes sont basées sur le même algorithme *forward-backward*, ce qui simplifie la mise en œuvre de la méthode en raison de la nature éparse de la paramétrisation TM.

3.5 Méthodologie de validation

3.5.1 Matériel

Nous avons testé la méthode de segmentation proposée sur un jeu de données issu de la INbreast [Moreira et al., 2012]. Notre choix s'est porté sur cette base en raison de la précision de la segmentation manuelle des masses qu'elle contient. Les annotations qui y sont faites sont dans un premier temps réalisées par un radiologue, puis validées par un second spécialiste. La segmentation manuelle fournit par les deux experts est d'une très grande précision, ce qui nous permet de nous comparer de façon objective à la vérité terrain. Dans notre cas, 48 masses malignes de cette base ont été retenues pour constituer notre jeu de données. Ces 48 masses appartiennent aux classes BI-RADS IV, V et VI, et présentent des formes irrégulières dont les contours sont spiculés et sont embarquées dans des densités très élevées rendant presque impossible leur distinction

du fond de l'image. Ces mammographies ont été retenues parce qu'elles correspondent aux cas cliniques les plus critiques et également parce que leur segmentation présente le plus grand degré de difficulté. Pour chaque mammographie, la vérité terrain fournie par les radiologues est utilisée pour extraire la région où la masse est localisée. De cette façon, le jeu de données sur lequel notre approche est validée se compose uniquement des ROIs.

3.5.2 Evaluation de la performance de la segmentation

En vue d'évaluer de façon objective la performance de notre méthode de segmentation en comparaison de celles dédiées à la mammographie, nous avons calculé cinq mesures de performance que sont : *Area Overlap Ratio* (AOR), *Probabilistic Rand Index* (PRI), *Variation of Information* (VoI), *Global Consistency Error* (GCE) et *Boundary Displacement Error* (BDE) ; qui ont été choisies en raison de leur utilisation fréquente en mammographie – c'est le cas de l'AOR – et parce qu'elles appartiennent aux outils qui sont les plus corrélés à la segmentation humaine en terme de perception visuelle.

L'AOR [Duda et al., 2012] encore connu sous le nom de la mesure de similarité de Jaccard, compare le résultat de la segmentation automatique à la vérité terrain fournie par les radiologues. Soit $\mathbf{A}_{log-PRF}$ la superficie automatiquement segmentée et $\mathbf{B}_{expert}$ la superficie de la segmentation qui nous sert de référence. L'AOR est définie comme le rapport du nombre de pixels issus de l'intersection sur ceux de l'union. La valeur de l'AOR est entre 0 – pas de recouvrement - et 1, superposition parfaite.

$$AOR = \frac{\mathbf{A}_{log-PRF} \cap \mathbf{B}_{expert}}{\mathbf{A}_{log-PRF} \cup \mathbf{B}_{expert}} \quad . \tag{3.22}$$

Le PRI [Pantofaru et Hebert, 2005] est une mesure très intéressante car elle prend en compte la variabilité de la segmentation entre différents observateurs humains, elle est surtout utile en mammographie où la segmentation manuelle des masses par deux radiologues peut être très différente en fonction de leur expérience en mammographie. Le PRI compte le nombre de paires de pixels dont les étiquettes sont compatibles entre la segmentation automatique et la vérité terrain, et en fait une moyenne sur diverses vérités terrain afin de tenir compte de la variation de l'échelle dans la perception humaine. Sa valeur varie entre $[0, 1]$, plus elle est grande mieux c'est. Il convient de signaler que dans notre cas, nous ne disposons que d'une seule vérité terrain. Pour sa part, le VoI [Meila, 2005] calcule la distance entre deux segmentations en optimisant un critère de la théorie de l'information à savoir la moyenne de l'entropie conditionnelle, dans le but de mesurer de façon grossière le degré de la présence de l'aléatoire dans une segmentation. Le GCE [Martin et al., 2001] évalue la proportion à laquelle une segmentation peut être vue comme un raffinement d'une autre segmentation ; de manière à ce que les segmentations qui sont reliées de cette façon sont vues comme compatibles. Le GCE varie entre $[0, 1]$, plus elle est faible, mieux c'est. Le BDE [Freixenet et al., 2002] calcule la moyenne de l'erreur de déplacement – des pixels situés sur le contour – entre deux images segmentées. Notamment, il définit l'erreur d'un pixel du contour comme la distance entre ce pixel et le pixel le plus proche situé sur le contour de l'autre image. Le VoI et le BDE ont des valeurs qui se situent dans l'intervalle $[0, \propto)$, plus elles sont faibles, mieux c'est.

3.5.3 Approches retenues pour comparaison

En raison du caractère privé de certains jeux de données utilisés dans les travaux de recherche et la sélection aléatoire des images contenues dans les bases d'images pour la validation des méthodes de segmentation proposées dans l'état-de-l'art, il est difficile de faire une comparaison objective. De plus, la validation de certaines contributions se fait de façon visuelle sans utiliser des mesures quantitatives pertinentes. Dans le but d'évaluer l'efficacité et la robustesse de notre approche, une comparaison avec les résultats publiés dans le passé et obtenus dans les conditions similaires (ou presque) aux nôtres a été faite : segmentation des masses malignes avec une forte densité et divers contrastes, et une évaluation quantitative des résultats impliquant tout au moins l'AOR.

Les deux méthodes proposées par [Kupinski et Giger, 1998] : l'indice de gradient radial et une simple approche probabiliste. Le jeu de données qu'ils ont utilisé pour la validation est privé et contient uniquement des masses malignes ; cependant, ils n'ont fourni aucune information à propos de la densité du sein et du contour des masses. [Sahiner et al., 2001b] ont d'abord utilisé l'algorithme des K-means afin de détecter la forme de la ROI, après quoi, ils ont utilisé des modèles de contour actif – *snakes* – comme l'étape finale de leur approche. Leur validation a été faite sur un jeu de données contenant une part égale de masses bénignes et malignes. [Tao et al., 2010] ont quant à eux proposé une classification en plusieurs étapes des pixels de l'image en se servant d'un ensemble de descripteurs d'intensité, de forme et de texture dans le but de créer une matrice des probabilités conditionnelles nommée PM. Ensuite, la structure anatomique susceptible d'être la masse est extraite de PM à l'aide d'une connaissance a priori sur la forme et la situation des masses. Pour finir, ils ont utilisé une méthode d'optimisation par coupe minimale pour générer le masque final de la segmentation. Leur méthode a été évaluée sur une base privée de 54 images dont 51 malignes et 3 bénignes. Les masses utilisées dans leur étude ont une forme irrégulière, et des contours spiculés ; de plus, le jeu de données comporte tous les types de densités. L'approche introduite par [Ribeiro et al., 2013] utilise une technique appelée *Enhanced Independent Component Analysis Mixture Model* (EICAMM) pour la segmentation automatique des masses. Un jeu de données de 174 ROIs collecté sur la DDSM et la LAPIMO, avec 152 lésions bénignes et 22 masses malignes a été utilisé pour valider leur approche. Ces images contiennent des masses ayant tous les types de densité, de forme et de contour. Pour leur part, [Pereira et al., 2014] ont utilisé une méthode de seuillage multiple, couplée aux ondelettes et à un algorithme génétique pour la segmentation des masses. L'évaluation de la pertinence de leur méthode a été faite sur des images tirées de bases publiquement disponibles, avec le même nombre de lésions bénignes et malignes.

Par ailleurs, nous avons décidé de comparer notre méthode (log-PRF) à trois approches non-supervisées, qui sont bien connues des compétitions de segmentation d'images, en raison de leur grande efficacité et qui sont publiquement disponibles sur internet. Les deux premières exploitent une optimisation par coupure de graphe, tandis que la dernière à recours à l'ICM.

Felzenszwalb *et al.* [Felzenszwalb et Huttenlocher, 2004] (FH) a proposé une méthode efficace de segmentation de régions, qui est le résultat d'un prédicat sur l'évaluation de la présence pertinente d'une frontière entre les deux régions, à l'aide d'une représentation sous forme de graphe de l'image à segmenter. Cela passe par la comparaison de deux quantités que sont les différences d'intensité des pixels présents le long d'un contour, et celle des pixels voisins contenus

dans chaque région de l'image. La segmentation en fonction de ces deux informations est alors faite en utilisant l'algorithme de Kruskal – qui est une méthode de recherche d'arbres recouvrant de poids minimum. Cette approche se distingue par sa capacité à ajuster de façon adaptative le critère de segmentation basé sur un degré de variabilité dans les régions voisines de l'image. Elle produit des résultats qui capturent ces propriétés globales et a également l'avantage de s'éxécuter presque en temps réel.

[Salah et al., 2011] (KM) ont conçu une méthode de segmentation des images où peuvent cohabiter des modèles Gaussiens, Gamma, ou de Whishart, comme c'est le cas des mammographies et pour lesquelles les approches non-supervisées classiques de coupure de graphe échouent à donner de bons résultats. Leur méthode de segmentation multi -régions non-supervisée est fondée sur une transformation de l'image qui se traduit par une augmentation de la dimension de l'espace des pixels de l'image afin de traiter les cas les plus complexes ; ils se servent ensuite de l'optimisation par coupure de graphe de la fonction d'énergie. Pour y parvenir, ils réalisent dans un premier temps une transformation implicite de l'image par le biais d'une fonction à noyau, en l'occurence le *Radial Basis Function* (RBF), dans le but d'exploiter le modèle constant par morceaux lié à la formulation des coupures de graphe. Ensuite, ils ont proposé un nouveau terme de régularisation permettant de préserver les discontinuités entre les différentes régions de l'image et un nouveau terme d'attache aux données pour évaluer l'écart entre l'image précédement transformée à l'intérieur de chaque région segmentée et le modèle constant par morceaux. Quant à l'optimisation de la fonction d'énergie qui en résulte, elle est réalisée par plusieurs itérations de deux étapes successives : (a) la première étant l'étiquetage des régions de l'image tout en tenant compte des paramètres statistiques de chaque région ; (b) pour sa part, la seconde identifie la partition optimale de l'image – qui tient compte des paramètres obtenus à l'étape précédente – via une minimisation par coupe minimale qui utilise l'algorithme *swap* de [Boykov et al., 2001]. La méthode proposée par [Salah et al., 2011] améliore celle de Boykov *et al.* en terme d'efficacité et de flexibilité, tout en conservant les avantages liés à l'optimisation par coupe minimale et en étant non-supervisée ; c'est ce qui justifie son choix parmi les méthodes auxquelles nous nous sommes comparés dans cette étude.

[Mignotte, 2010] (PRIF) a proposé une approche de segmentation utilisant un modèle de fusion à base de MRF, dont le but est de combiner plusieurs résultats de segmentation obtenus par divers modèles de *clustering* et ce dans l'idée d'aboutir à une segmentation plus efficace et fiable. Le modèle de fusion utilisé par l'auteur, exploite une mesure de comparaison très répandue en segmentation, à savoir le PRI. Cette mesure a par la suite permis à l'auteur de déduire un modèle de fusion Bayésien Markovien attrayant de champs d'étiquettes exprimé comme une distribution de Gibbs associée à une segmentation correcte, dans laquelle le PRI sert de maximum de vraisemblance. Afin de trouver la configuration qui minimise la fonction coût, Mignotte a préféré une stratégie de minimisation multi-résolue exploitant l'ICM comme procédure d'optimisation.

3.5.4 Implémentation numérique

Notre méthode a été mise en oeuvre, conformément aux indications données dans les sections précédentes : après s'être servi de la transformation logarithmique pour améliorer le contraste de l'image, la segmentation non-supervisée à l'aide des champs de Pickard a été réalisée. Un

seul hyper-paramètre, *c-à-d.* le nombre de niveaux discrets du PRF a été spécifié. Nous avons constaté que des résultats convaincants sont obtenus pour des valeurs qui varient entre deux et quatre, ce qui peut être attribué à la présence de motifs de répétition souvent rencontrés en mammographie (nodulaire, linéaire, homogène et radio-transparent), *c-à-d.* les dénommés *blocs de Tabar* [Tabár et al., 2005].

Pour nos expériences, les paramètres internes des méthodes auxquelles nous nous comparons, sont fixés aux valeurs optimales, comme suggérés par les auteurs. Ces paramètres sont, pour FH : $\sigma = 0.5$ est utilisé pour lisser l'image avant de la segmenter, – k et min, qui sont respectivement la valeur de la fonction de seuil et la taille de la plus petite région, sont fixés de façon empirique, jusqu'à obtenir un résultat de segmentation satisfaisant ; pour KM : le nombre maximum initial de régions N_{reg} varie entre deux et quatre (un peu comme notre méthode) et le poids du terme de régularisation est α=0.6 ; finalement, pour PRIF : $K_1 = 18$, $K_2 = 10$, $K_3 = 2$, avec $K_.$ le nombre de classes. Il est important de noter que nous nous sommes servis du code Matlab® fournit par Allen Y. Yang pour calculer le PRI, le VoI, le GCE et le BDE. Ce code est publiquement disponible en ligne à l'adresse `http://people.eecs.berkeley.edu/yang/software/lossysegmentation/`.

3.6 Segmentation des masses : résultats et discussion

Dans l'idée d'évaluer l'efficacité de notre approche sur les cas les plus compliqués en mammographie, des tests ont été réalisés sur 48 masses malignes extraites de la INbreast.

Une synthèse des résultats est présentée dans le tableau 3.1. La valeur moyenne et l'écart-type de l'AOR sont respectivement 0.75 et 0.12, et, avec un seuil de 0.5 sur l'AOR, notre méthode est capable de segmenter correctement 97.92% des masses malignes. Ces chiffres montrent clairement que notre approche offre une segmentation fortement précise des mammographies et ce même dans des conditions très difficiles. La qualité du résultat obtenu a été également évaluée à la lumière des approches de segmentation des masses en mammographie auxquelles nous nous comparons dans ce chapitre. Les résultats obtenus par ces auteurs, sont reportés dans le tableau 3.2 :

— Les deux méthodes proposées par [Kupinski et Giger, 1998] donnent de bons taux de segmentation, respectivement 92% et 96% avec un seuil de 0.3 pour les valeurs de AOR.

— Les valeurs de AOR obtenues par [Sahiner et al., 2001b] sont 0.74 ± 0.13 et 0.74 ± 0.11 pour un jeu de données annoté par deux radiologues différents. La segmentation des masses lobulées avec une faible densité aboutie à une grande valeur de AOR de 0.92, mais les valeurs de l'AOR pour cette méthode se retrouvent dans l'intervalle 0.5 à 0.73 pour des masses lobulées embarquées dans des densités très élevées.

— Les résultats obtenus dans [Tao et al., 2010] montrent une valeur de AOR de 0.689 ± 0.16 pour la qualité de la segmentation des superficies des masses et 0.540 ± 0.164 pour les contours.

— [Ribeiro et al., 2013] ont quant à eux, obtenu un taux de bonne segmentation de 74.07% pour les 152 masses bénignes de leur jeu de données avec un seuil de 0.5 pour la mesure AOR. Comme indiqué par les auteurs, cette approche donne une bonne performance surtout dans le cadre de la segmentation des masses ayant une forme ronde ou ovale et un contour circonscrit, en raison de leur facilité à être appréhendées. Cependant, le taux de

bonne segmentation chute fortement à 50% avec les 22 masses malignes, et ce avec le même seuil de 0.5.

— Dans [Pereira et al., 2014], les auteurs ont eu un AOR de 0.79 ± 0.08. Aucun seuil n'a été utilisé et ils n'ont pas fourni des détails sur les valeurs de AOR obtenus séparément pour les masses bénignes et malignes de leur jeu de données.

Par ailleurs, il faut noter que la plupart des méthodes mentionnées plus haut, nécessitent la spécification de plusieurs hyper-paramètres qui en plus doivent être déterminés de manière expérimentale, après plusieurs tentatives et échecs, avant d'obtenir des résultats acceptables ; ceci est explicitement mentionné dans [Tao et al., 2010] et [Pereira et al., 2014]. Contrairement à ces approches, notre méthode ne nécessite que la spécification d'un seul hyper-paramètre qui peut être fixé assez aisément en raison du faible intervalle dans lequel il peut être choisi.

Ensuite, nous avons évalué la performance du log-PRF en comparaison du FH, du KM et du PRIF en nous servant des cinq métriques que sont : le PRI, le VoI, le GCE, le BDE et l'AOR. Les valeurs obtenues sont dans le tableau 3.3. Le meilleur résultat par mesure est mis en gras dans le tableau 3.3. A la lumière de ces chiffres, nous sommes parvenus aux conclusions suivantes ; premièrement, notre méthode offre la meilleure performance pour chacune des cinq métriques et ce avec une différence élevée comparée aux trois méthodes phares de l'état-de-l'art en matière de segmentation.

Deuxièmement, pour ce qui est du PRI, on observe que notre méthode donne des résultats très compétitifs (une valeur de PRI = 0.80 signifie qu'en moyenne 80% des paires d'étiquettes de pixels sont correctement classifiées par notre méthode de segmentation) avec une faible variance sur l'ensemble des images de notre jeu de données. Enfin, la méthode du log-PRF présente des performances exceptionnelles pour les mesures GCE et BDE. Cela signifie que notre technique aboutit à une excellente segmentation avec une meilleure adhérence et un faible déplacement du contour des masses, en comparaison de la vérité terrain.

Des exemples de nos résultats sont présentés à la figure 3.2 et ceux des trois méthodes non-supervisées auxquelles nous nous comparons sont exposés à la figure 3.3. Ces images ont été choisies parce qu'elles représentent trois cas typiques de densité mammaire élevée. On peut observer qu'en dépit de la densité dans laquelle les masses sont embarquées, notre méthode est capable d'extraire les contours spiculés des masses malignes avec une bonne précision, ainsi, les résultats obtenus sont très proches de la segmentation manuelle. Par ailleurs, il convient de remarquer que malgré le fait que quelques pixels ont été perdus – ce qui réduit la superficie de la masse et donc empiète quelque peu sur sa forme – notre algorithme préserve la partie essentielle de l'information contenue dans le contour des masses, qui est l'élément le plus pertinent. En effet, selon l'ACR, la différence essentielle entre les masses bénignes et malignes réside dans l'aspect de leur contour. De plus, comme l'ont démontré [Liberman et al., 1998], les masses spiculées représentent 81% des masses malignes. Donc, une extraction précise du contour d'une masse avec une grande précision est essentielle pour sa description dans les étapes ultérieures de la conception de notre système de DAOx.

Pour ce qui est de la rapidité numérique des différentes approches (tableau 3.4), le temps moyen de calcul (CTe) de FH est de (0.08 ± 0.11) s, tandis que, pour les méthodes KM et PRIF, le CTe est respectivement de (2.93 ± 2.67) s et (570.19 ± 367.79) s. Les deux derniers chiffres sont d'un ordre de grandeur bien plus grand que la méthode FH. Il convient néanmoins de souligner

que, FH et KM utilisent un algorithme d'optimisation à base de coupure de graphe, tandis que PRIF exploite l'ICM. C'est la raison pour laquelle le temps de calcul du PRIF est largement plus grand que FH et KM. Ceci vient prouver une fois encore que les méthodes d'optimisation basées sur la coupure de graphe sont plus rapides que l'ICM.

La méthode log-PRF présentée dans ce chapitre est quant à elle plus rapide que FH, elle est respectivement 47× et 9193× plus rapide que KM et PRIF, avec une moyenne et un écart-type de (0.062 ± 0.076) s pour une image de 300×300 pixels. Ce résultat peut être comparé à ceux obtenus par [Li et al., 1995], qui ont noté un temps de calcul de 6 s et 5 min respectivement pour des images de 64×64 et 500×1000 pixels, avec une optimisation par ICM. Ces résultats sont une indication claire de l'efficacité numérique de notre méthode – proclamée plus haut durant la présentation de notre approche – comparée aux techniques d'optimisation à base de coupure de graphe et de l'ICM, et ce malgré sa nature non-supervisée.

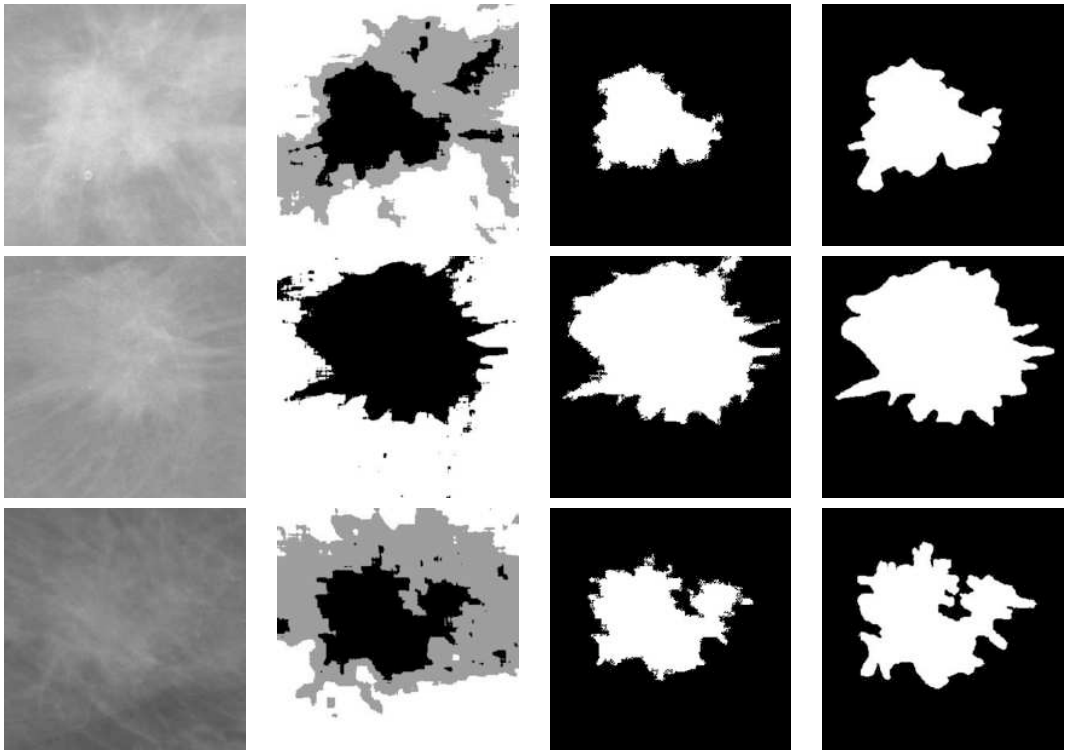
Afin de fournir une évaluation objective des performances de notre méthode de segmentation, nous avons besoin d'étudier les exemples pour lesquels notre algorithme échoue à produire de bons résultats. C'est le cas des images présentées à la figure 3.4. Nous pouvons observer que le log-PRF extrait non seulement la forme mais aussi le muscle pectoral et le fond de l'image. L'une des solutions pour améliorer la segmentation de ces types de masses est peut-être d'intégrér une information a priori sur la forme de certaines masses, en particulier celles qui se retrouvent dans le muscle pectoral. On peut également procéder à une élimination au préalable du muscle pectoral avant de segmenter la masse ; puis en se servant de l'information a priori reconstruire la partie manquante de la superficie de la masse. Le risque, c'est que ces solutions ne soient pas assez robuste pour gérer la variabilité des formes et des contours des masses mammaires.

TABLE 3.1 – Statistiques descriptives de la répartition des valeurs de AOR basées sur le log-PRF

Type de masses	min	max	moyenne ± écart-type	Taux de succès ($A0R > 0.5$)
Malignes (INbreast)	0.44	0.99	0.75 ± 0.12	97.92 %

TABLE 3.2 – Segmentation des masses mammaires : Evaluation de la performance des méthodes de l'état-de-l'art en mammographie

Methodes	Type des masses	moyenne ± écart-type	Taux de succès ($A0R > 0.5$)
[Kupinski et Giger, 1998]	Malignes	N/A	N/A
[Sahiner et al., 2001b]	Bénignes et malignes	0.74 ± 0.11	N/A
[Tao et al., 2010]	Malignes	0.689 ± 0.16	N/A
[Ribeiro et al., 2013]	Malignes	N/A	50 %
[Pereira et al., 2014]	Bénignes et malignes	0.79 ± 0.08	N/A



(a) (b) (c) (d)

FIGURE 3.2 – Segmentation des masses malignes en mammographie : (a) ROIs, (b) segmentation automatique par log-PRF, (c) image binaire de (b) et (d) vérité terrain.

TABLE 3.3 – Evaluation de la performance des meilleures méthodes de segmentation non-supervisées de l'état-de-l'art

Méthodes	PRI	GCE	VoI	BDE	AOR	Taux de succès
FH	0.72 ± 0.11	0.22 ± 0.08	1.67 ± 0.43	9.46 ± 5.15	0.60 ± 0.18	70.83 %
KM	0.76 ± 0.13	0.16 ± 0.09	1.29 ± 0.33	17.76 ± 18.68	0.64 ± 0.25	85.42 %
PRIF	0.73 ± 0.12	0.18 ± 0.08	1.46 ± 0.31	15.79 ± 17.59	0.60 ± 0.25	83.33 %
log-PRF	**0.83 ± 0.07**	**0.15 ± 0.07**	**1.20 ± 0.42**	**2.31 ± 2.79**	**0.75 ± 0.12**	**97.92** %

3.7 Extraction des spicules

En mammographie, la présence de structures convergentes est bien souvent la preuve de la présence d'une lésion maligne, *c-à-d.* distorsions architecturales ou masses spiculées. Un peu plus haut, nous avons présenté une méthode de segmentation robuste, efficace et rapide permettant

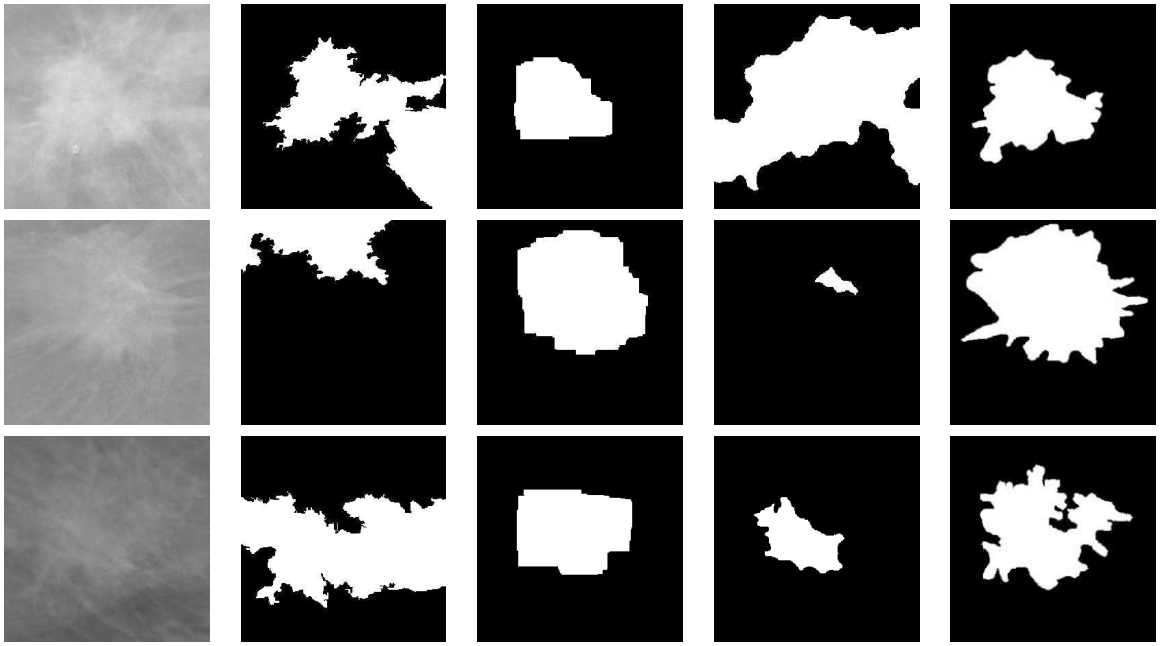


FIGURE 3.3 – Segmentation des masses malignes en mammographie : (a) ROIs, (b) FH, (c) KM, (d) PRI et (e) vérité terrain.

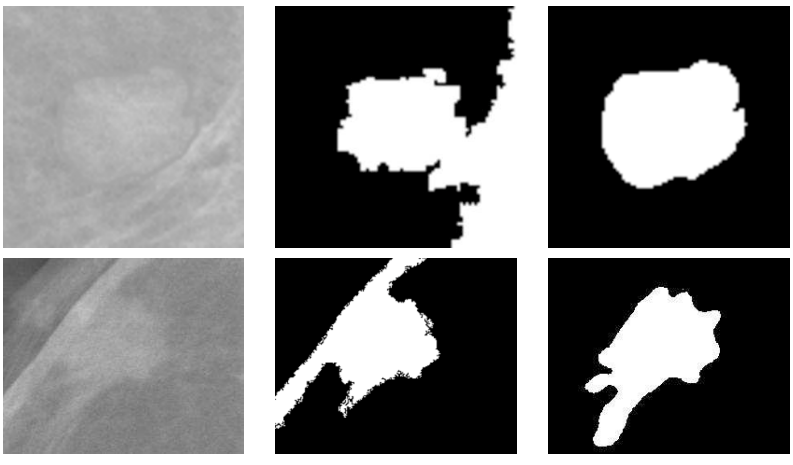


FIGURE 3.4 – Segmentation des masses malignes en mammographie : les cas d'échec. (a) ROIs, (b) segmentation automatique par log-PRF et (c) vérité terrain.

de segmenter les lésions à centre dense, qui sont un autre signe de suspicion. Toutefois, la sémiologie médicale nous apprend que ces deux symptômes ne sont pas toujours concomitants, et certaines lésions malignes peuvent uniquement se traduire par un motif de convergence. Des méthodes efficaces ont été proposées dans l'état de l'art pour tenter de résoudre ce problème (*p. ex.* [Karssemeijer et te Brake, 1996], [Zwiggelaar et al., 1999], [Palma et al., 2009a]) ; la plupart

TABLE 3.4 – Evaluation du temps de calcul

CTe (s)	FH	KM	PRIF	**log-PRF**
INbreast (48 images)	0.08 ± 0.11	2.93 ± 2.67	570.19 ± 367.79	**0.062 ± 0.076**

de ces techniques s'articulent en deux étapes : extraction de caractéristiques locales de l'image et classification des observations en affectant par exemple à chaque pixel une probabilité d'appartenir à la lésion. Toutefois, cette démarche présente une vraie limite qui est liée à l'acquisition des mammographies. En effet la superposition des diverses couches de tissus et des structures anatomiques qu'elles contiennent rend difficile une classification réaliste d'une éventuelle lésion exclusivivement sur la détermination d'un motif de convergence basé sur des descripteurs extraits sur chaque pixel de l'image. A cela s'ajoute le fait que les autres structures curvilignes anatomiques de l'image ont quasiment la même intensité de niveau de gris que celle des spicules sur lesquelles elles se superposent. Afin de contourner cette difficulté, un nombre limité de travaux ont proposé d'extraire les spicules de l'image (*p. ex.* [Kobatake et Yoshinaga, 1996], [Sampat et Bovik, 2003], [Muralidhar et al., 2010]) ; cependant, les performances obtenues par ces auteurs sont encore insuffisantes pour une utilisation en milieu clinique, surtout parce que les techniques proposées ne prennent pas en compte les différentes formes des structures anatomiques observées en mammographie.

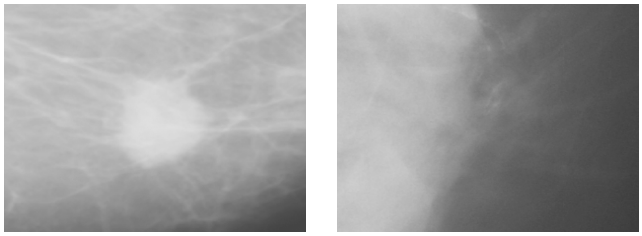
Dans cette partie, nous proposons une nouvelle méthode inspirée des méthodes utilisées en télédétection [Lacoste et al., 2005] et [Gao et Bischof, 2009] et des travaux sur la modélisation a contrario introduit par [Desolneux et al., 2000] et adaptée à la tomosynthèse par certains chercheurs dont [Palma et al., 2009a]. La méthode que nous avons proposé pour l'extraction des spicules – que nous considérons comme des structures de convergences – poursuit le même objectif que les travaux évoqués dans le dernier paragraphe, en essayant de capturer les caractéristiques saillantes des spicules : ainsi, après une discrétisation des diverses structures anatomiques de l'image en segments, les interactions locales entre les segments appartenant à ces structures curvilignes sont prises en compte grâce à un modèle de Markov que nous avons spécifiquement conçu pour cette application. Ensuite, la relation géométrique entre les spicules et la masse d'où elles émanent est exploitée à l'aide d'une méthode de détection a contrario. L'intérêt de l'approche que nous proposons est de s'affranchir de l'extraction et du calcul de descripteurs locaux de l'image qui seraient par la suite utilisés comme entrée d'un classifieur pour la classification, tâche dont la performance et le temps de calcul dépendent fortement des attributs extraits de l'image. Ceci a pour conséquence immédiate une réduction significative du temps de calcul et une amélioration des performances dont la robustesse est garantie par le modèle que nous avons développé.

Dans un premier temps nous rappellerons quelques généralités sur les lésions qui se traduisent par un motif de convergence dans les mammographies, puis nous verrons les approches existantes pour ce type de problème. Suite à cela, nous proposerons une nouvelle approche combinant à la fois l'extraction des structures anatomiques curvilignes de l'image par le biais des MRFs et la détection des spicules – ou structures convergentes – par une modélisation a contrario du problème. Nous finirons par une illustration de la méthode sur des mammographies ainsi qu'une

discussion sur les avantages et les limites de cette approche.

3.8 Les motifs de convergence en mammographie

En mammographie, il existe deux types de lésions qui se traduisent par un motif stellaire. Il s'agit des masses spiculées et des distorsions architecturales. Le premier est une lésion palpable caractérisée par un centre dense dû à la forte atténuation de rayons X et délimitée par des bords, de même qu'un voisinage contenant des structures anatomiques qui convergent vers son centre, qui sont appelées des *spicules* [Kim et Kim, 2005] et qui donnent à la lésion sa forme étoilée. Quant aux distorsions architecturales, elles se définissent comme étant une modification de l'architecture normale du parenchyme mammaire sans aucune masse visible, mais incluant des spicules qui émanent d'un point. La figure 3.5 présente un exemple de ces deux types de lésions.



(a) (b)

FIGURE 3.5 – Exemple d'une lésion spiculée (a) et d'une distorsion architecturale (b) en mammographie.

3.9 Travaux existants sur les lésions stellaires

Il existe bon nombre de techniques dédiées à l'extraction des descripteurs (centre dense et spicules) de l'image et toutes permettent de distinguer les régions suspectes des régions normales d'une mammographie. Parmi les méthodes existantes dans l'état-de-l'art, certaines utilisent l'une ou l'autre des attributs évoquées dans la section précédente, tandis que d'autres ont choisi de les combiner. Nous avons fait ci-dessous une synthèse des approches les plus pertinentes qui s'appuient sur ces attributs.

Au nombre des travaux combinant à la fois la détection des centres denses et des régions où les spicules sont organisées sous forme d'étoile, on peut citer les travaux de [Ng et Bischof, 1992] qui propose trois méthodes pour l'extraction des spicules, toutes fondées sur des théories concernant la forme idéale d'une lésion stellaire. Ici nous ne présenterons que celle qui donne les meilleurs résultats et qui est fondée sur l'hypothèse selon laquelle non seulement les spicules sont des structures visibles et ce de façon distincte, mais en plus elles ont une micro-structure dont l'extraction permettrait une recherche bien plus précise de leur centre. Aussi, après un étirement de l'histogramme, les auteurs ont filtré la ROI avec un filtre Gaussien afin de supprimer une certaine quantité de bruit tout en évitant d'endommager la fine colonne des structures

anatomiques linéaires de l'image. Ils ont ensuite introduit une régle dite de *cohérence du gradient* qui leur permet de noter l'absence ou la présence d'orientations dominantes dont ils se servent pour déterminer les pixels appartenant à l'axe principal de la structure radiale. Chacun des pixels de l'axe principal est ensuite décrit par des valeurs contenant des informations relatives à chaque spicule. La transformée de Hough est alors utilisée pour repérer le centre des structures radiales. Enfin, les spicules détectées dans l'image sont supprimées afin de faciliter la détection de la masse par une méthode de *template matching*. Quant à [Zwiggelaar et al., 1998] et [Zwiggelaar et al., 1999], ils ont d'abord utilisé des filtres *recursive median filter* (RMF) pour la détection des centres denses de l'image ; ensuite, ils ont modelisés les motifs linéaires entourant les masses en utilisant un détecteur de lignes multi-échelle. La combinaison de ces deux techniques leur a fourni une carte de confiance sur laquelle un seuillage a été appliqué pour détecter les lésions stellaires. Une approche différente a été proposée par [Liu et al., 2001] qui ont exploité l'information relative au gradient en y ajoutant des descripteurs de texture dans la modélisation des régions suspectes. Après avoir calculé l'orientation du gradient dans un voisinage, ils ont utilisé comme descripteur l'écart-type de l'histogramme de l'orientation du gradient, qui traduit l'idée selon laquelle toutes les orientations sont représentées sur des lésions spiculées et que par conséquent l'orientation locale doit avoir un histogramme plat. Cet attribut avait été introduit par [Kegelmeyer Jr, 1992] et [Kegelmeyer Jr et al., 1994] sous le nom de *Analysis of Local Oriented Edges* (ALOE) et se calcule en tout point de l'histogramme d'orientation locale dans un voisinage autour de ce pixel ; ce descripteur est l'écart-type de la hauteur des boîtes de cet histogramme. Ainsi, lorsque les valeurs de ALOE sont grandes, on se situe dans des zones de l'image où seul un petit nombre de directions sont privilégiées tandis que pour de petites valeurs, on est dans une région où l'orientation a un histogramme plat. ALOE est de ce fait un descripteur discriminant pour les lésions spiculées ; cependant, il est très sensible, ce qui implique un grand nombre de faux positifs. Pour limiter ce défaut, [Liu et al., 2001] ont ajouté de nouveaux descripteurs adaptés aux tissus sains et qui sont obtenus après la convolution de l'image avec quatre masques. Ces nouveaux descripteurs et ALOE sont ensuite utilisés dans un arbre de décision binaire afin de classifier chacun des pixels de la mammographie et leur attribuer un degré de suspicion. Enfin, pour la détection des centres denses, ils ont extrait deux descripteurs que sont l'intensité moyenne dans un voisinage ainsi que l'écart-type de l'intensité sur le voisinage de chaque pixel. Pour leur part, [Palma et al., 2010] et [Palma et al., 2014] ont proposé deux chaînes de traitement distinctes basées l'une sur la détection des masses et l'autre sur les distorsions architecturales. Pour détecter les centres denses, ils ont d'abord utilisé un filtre connexe flou [Palma et al., 2008] – qui possède l'avantage d'être adaptable au cas traité en ce sens qu'il ne necessite pas d'utiliser un seuillage dur – avec des éléments discriminants comme la taille, la compacité et les mesures de contraste des masses. Enfin, ils réalisent une extraction de descripteurs sur les masses obtenues afin de les classifier comme suspectes ou saines à l'aide d'un classifieur *Support Vector Machine* (SVM) [Vapnik, 2013]. En ce qui concerne les distorsions architecturales, ils ont proposé une modélisation a contrario du problème [Palma et al., 2009a] et [Palma et al., 2009b], dont l'idée est de définir une mesure de convergence et de sélectionner les réalisations qui ont très peu de chance d'apparaître dans des mammographies normales. La sortie de la modélisation a contrario est par la suite utilisée pour extraire des descripteurs qui sont classifiés par un SVM afin d'affecter une mesure de suspicion à chaque image. Les résultats de ces deux approches sont ensuite agrégés à l'aide d'une règle de disjonction définie par les

auteurs.

D'autres auteurs ce sont pour leur part intéressés à la détection des lésions spiculées sans prendre en compte les centres denses ; à titre d'exemple, [Karssemeijer, 1994], [Karssemeijer, 1999] et [Karssemeijer et te Brake, 1996] ont proposé une approche qui repose sur la matrice d'orientation obtenue en calculant l'orientation de chaque pixel de l'image initiale à l'aide des filtres Gaussien de second ordre et sur le fait que trois directions de filtrage seulement sont nécessaires. Après avoir réalisé une convolution de l'image avec ces filtres sur plusieurs échelles couvrant toutes les largeurs des spicules potentielles, ils ont retenu pour chaque pixel la réponse maximale sur l'ensemble des échelles parcourues ; ce qui leur a donc permis d'avoir une estimation spécifique de l'orientation des pixels appartenant soit aux structures anatomiques soit au fond de l'image. Sur la base de cette information, ils ont proposé deux descripteurs pour chaque pixel i dans un vosinage $\mathcal{N}_i$ formé des pixels j à distance r_{ij} de i, avec $r_{min} \leq r_{ij} \leq r_{max}$ et qui constituent l'entrée de leur classsifieur. Le premier descripteur est le nombre total n_i de pixels j orientés vers le centre i. Le second descripteur se sert du fait que les spicules sont disposées autour du centre de la lésion et sont donc dans plusieurs directions ; ainsi, les auteurs divisent le voisinage autour de i en K *bins* dans lequel ils comptent $n_{i,k}$ pixels parmi les $N_{i,k}$ pixels présents. Le second descripteur est le nombre n_+ parmi les k où $n_{i,k}$ est supérieur à la moyenne attendue. Ces deux attributs, sensibles aux motifs à symétrie radiale sont enfin utilisés dans un classifieur Bayésien dans le but de détecter les lésions stellaires. Dans des travaux plus récents, [te Brake et Karssemeijer, 1999] et [te Brake et Karssemeijer, 1998] ont amélioré cet algorithme en y incorporant de la multi-résolution. Une autre méthode utilisant l'information du contour a été proposée par [Jiang et al., 1998] qui après avoir amélioré la visibilité des spicules à l'aide des opérateurs de morphologie mathématique a ensuite calculé deux descripteurs – représentant la proportion de spicules – qui sont utilisées pour entrainer un classifieur basé sur une fonction discriminante.

Cependant, seul un nombre limité de travaux se proposent d'extraire les spicules de l'image. Ainsi, [Kobatake et Yoshinaga, 1996] ont proposé une approche qui consiste, premièrement à extraire l'axe principal des structures étroites et longues comme les spicules en se servant d'un filtre morphologique qui combine l'érosion et l'ouverture morphologique ; ils ont ensuite modélisé huit éléments structurants qui leur ont permis de supprimer un nombre assez important de faux positifs tout en conservant les structures anatomiques allongées qui ont une épaisseur supérieure à un seuil qu'ils ont préalablement fixé ; par la suite, ils ont proposé une version modifiée de la transformée de Hough afin de prendre en compte les structures allongées qui se retrouvent divisées en deux et situées de part et d'autre de la masse. Pour atteindre cet objectif, ils ont selectionné deux types de courbes, à savoir celles qui ont une amplitude de valeur négative et celles de valeur positive, puis, à l'aide des maxima locaux selectionnés dans l'accumulateur obtenu par la transformée de Hough, ils ont choisi les structures qui étaient susceptibles d'être des spicules. Cependant, les auteurs ont eux mêmes reconnus que leur approche n'est conçue que pour détecter les structures linéaires et échoue à extraire les spicules qui auraient une structure curviligne ; il en résulte donc un taux elévé de faux positifs. En 2003, [Sampat et Bovik, 2003] ont proposé une méthode qui utilise la transformée de Radon pour détecter les segments de l'image susceptibles d'être des spicules. En effet, la transformée de Radon associe à une image une fonction dans le domaine de la transformée et les droites contenues dans l'image de départ sont projetées sur des points dans le domaine de Radon. Ces points sont ensuite détectés par

un filtrage suivi d'un seuillage ; il suffit alors de réaliser la transformée de Radon inverse pour détecter les segments dans l'image de départ. Enfin, les convergences locales des spicules sont détectées en utilisant des filtres à symétrie radiale. Dans la même veine, [Muralidhar et al., 2010], ont utilisé la transformée de Radon pour atténuer les effets liés à la superposition des tissus durant l'acquisition de l'image ; puis en vue d'extraire les structures curvilignes de l'image, ils ont rehaussé le contraste des pics locaux du domaine de Radon qui correspondent aux structures linéaires dans le domaine spatial. Ensuite, ils utilisent des filtres Gaussien d'ordre quatre pour la détection des points d'initialisation où seront déployés des *snakes* ; pour l'évolution des *snakes* ils ont utilisé le *Vector Field Convolution* comme énergie externe, en raison de sa capacité à être robuste au bruit présent dans l'image.

3.10 Approche proposée pour l'annotation des spicules

La méthode que nous avons conçu est basée sur des hypothèses que nous avons formulé sur la structure des spicules telles qu'elles apparaissent dans des mammographies et ce en nous inspirant des descriptions faites dans l'état-de-l'art. Parmi celles que nous avons utilisé, quatre ce sont révélées être les plus importantes : (a) dans une mammographie 2D, les spicules qui ont un aspect quasi linéaire à une petite échelle en 3D, ont plutôt une forme curviligne en raison de la superposition des tissus et des occultations issues de la projection d'un objet 3D dans un plan 2D ; (b) quand on regarde toute la superficie d'une mammographie, les spicules ne sont pas les structures anatomiques les plus allongées du fait de la présence dans le parenchyme mammaire des vaisseaux sanguins, des canaux lymphatiques, des glandes mammaires, des fibres et de plusieurs autres structures ; (c) tous les segments d'une image n'appartiennent pas forcément à la même structure anatomique, des facteurs discriminants peuvent être obtenus en imposant des contraintes sur l'orientation qui existe entre les différents segments voisins et la distance qui séparent leurs extrémités, enfin, (d) les spicules sont des structures qui ont une intersection avec le contour de la masse et qui convergent vers le centre de celle-ci.

Dans l'idée d'exploiter les hypothèses formulées plus haut, la méthode proposée se décline comme suit : dans un premier temps, la mammographie est séparée en plusieurs patchs à l'intérieur desquels les structures curvilignes sont discrétisées en segments. Puis, à l'aide d'un modèle Markovien et de la prise en compte de l'information contextuelle, les positions des segments ont été affinées et ces derniers ont été associés en vue de recouvrer la forme initiale des vraies structures anatomiques de l'image. Enfin, les spicules sont détectées en nous basant sur les hypothèses (b) et (d).

Il faut souligner que la première partie de notre approche s'inspire des méthodes utilisées en télédétection [Lacoste et al., 2005] et [Gao et Bischof, 2009]. Dans le but de détecter les spicules parmi toutes les structures curvilignes, nous nous sommes servis de la modélisation a contrario proposée par [Desolneux et al., 2000] et utilisée pour la première fois en mammographie par [Palma et al., 2009a] pour les zones convergentes, que nous avons adapté pour la détection des spicules.

3.10.1 Extraction des structures allongées

Ici, l'idée générale est de séparer l'image en plusieurs sous-régions qui s'intersectent légèrement, de façon à ne retenir qu'un seul segment par sous-région. Après une étape d'initialisation, la position du segment dans chaque patch est estimée en se servant de la mammographie observée et du modèle Markovien dont le but est d'exploiter les propriétés géométriques des structures anatomiques. Les segments sont enfin regroupés en structures curvilignes qui sont par la suite traitées pour en déduire les spicules.

3.10.1.1 Discrétisation

Comme indiqué plus haut, l'image initiale est découpée en plusieurs patchs qui peuvent s'intersecter dans le but de faciliter le recouvrement des segments qui appartiennent à la même structure. La position initiale du segment contenu dans chaque patch est grossièrement déterminée à l'aide d'une transformée de Radon. Le rôle de ce dernier est de projeter plusieurs lignes droites du domaine spatial en un point dans le domaine de la transformée. A l'opposé des travaux ci-dessus cités [Muralidhar et al., 2010] où la transformée globale de Radon a été employée, nous utilisons une version locale, qui nous permet de prendre en compte la courbure des structures anatomiques. Cette technique est similaire à celle proposée par [Copeland et al., 1995], à la seule différence que dans notre cas, l'échelle correspond à la taille du patch. En 2D, la transformée de Radon est définie comme suit :

$$\mathfrak{R}(\rho,\theta) = \int\int_s f(x,y) \times \delta(\rho - x\cos(\theta) - y\sin(\theta))\, dx\, dy \ . \tag{3.23}$$

où $f(x,y)$ est l'intensité en niveau de gris du pixel situé à la position (x,y), δ la fonction Dirac, s le plan entier du patch et (ρ,θ) le couple des paramètres du domaine de Radon. Donc, chaque segment initial correspond au point d'intensité maximale dans le domaine de Radon. Cette configuration initiale est par la suite améliorée à l'aide de l'information contextuelle comme indiqué au paragraphe suivant.

3.10.1.2 Combinaison des segments

Le modèle probabiliste utilisé pour affiner la position de chaque segment est conçu comme suit : chaque patch s_n ; $1 \leq \ n \leq \ N$ est considéré comme un nœud d'un réseau cartésien S. Notre MRF a priori sur le modèle hiérarchique est essentiellement un modèle de Potts. En raison de l'équivalence entre les MRFs et les champs de Gibbs, la probabilité de notre réseau cartésien est donnée par la relation ci-dessous :

$$p(S) = 1/K \ \exp\left\{-\beta \sum_n \mathcal{U}(s_n)\right\} \ . \tag{3.24}$$

$$\mathcal{U}(s_n) = \mathcal{V}(s_n, s_n^-) + \mathcal{V}(s_n, s_n^+) \ . \tag{3.25}$$

où K, β et $\mathcal{U}(\cdot)$ sont respectivement le facteur de normalisation, le poids et l'énergie de chaque configuration du MRF. s_n^- et s_n^+ sont deux nœuds voisins définis dans un 8-voisinage et qui satisfont au mieux les critères de continuité définis entre les segments et traduits par $\mathcal{V}(\cdot)$. En

nous servant du maximum a priori qui est cohérent avec les champs de Markov, l'estimation des positions est réalisée grâce à la minimisation de l'énergie :

$$\mathcal{J} = \underbrace{\sum_{n} \mathcal{D}(s_n)}_{\text{terme d'attache aux données}} + \underbrace{\beta}_{\text{poids}} \underbrace{\sum_{n} \mathcal{U}(s_n)}_{\text{terme de régularisation}} \quad . \tag{3.26}$$

Le rôle du terme de l'attache aux données est de traduire la vraisemblance qu'un segment existe dans un patch donné. Ceci est fait de façon indépendante pour chaque segment obtenu à l'étape précédente. Pour des raisons de flexibilité, ce terme peut prendre diverses formes en fonction de l'application. En ce qui nous concerne, le terme d'attache aux données utilisé a été choisi en raison des résultats qu'il offre dans des images qui contiennent une grande proportion de bruit et peu de structures géométriques. Dans ce cas de figure, le contraste seul n'est plus suffisant pour faire la distinction entre le fond de l'image et les structures anatomiques. C'est donc la raison pour laquelle une mesure de vraisemblance découlant du *template matching* a été définie :

$$d_R(s_c) = \frac{\sum_{\rho \in I_\rho} \sum_{\theta \in I_\theta} \mathfrak{R}(\rho, \theta)}{\sum_{tous\ les\ \rho} \sum_{tous\ les\ \theta}} \quad . \tag{3.27}$$

où (ρ_c, θ_c) sont les coordonnées du segment s_c dans le domaine de Radon, $I_\rho = [\rho_c - \Delta_\rho, \rho_c + \Delta_\rho]$ $I_\theta = [\theta_c - \Delta_\theta, \theta_c + \Delta_\theta]$ et la somme au niveau du dénominateur est faite sur toutes les valeurs possibles de ρ et θ à l'intérieur de la région traitée. Ainsi, $d_R(s_c)$ estime la probabilité de l'existence d'un segment centré en (ρ_c, θ_c) contre tous les autres segments possibles dans le patch courant. En raison du terme de normalisation au dénominateur, l'affectation de la probabilité d'apparition de ce segment est localement invariant au changement de contraste et robuste à l'étirement d'histogramme. Enfin, le terme d'attache aux données de Radon est formulé comme suit :

$$\mathcal{D}_R(s_c) = 1 - 2 d_R(s_c) \quad . \tag{3.28}$$

et prend des valeurs comprises entre $(-1, 1)$; elle est faible pour les segments les plus saillants et élevée pour ceux qui sont moins remarquables.

Quant au terme de régularisation, il est le plus important car c'est lui qui comporte les informations a priori nécessaires au recouvrement des structures curvilignes. Les contraintes que nous imposons dans cette application, sont : une faible courbure, la continuité et le caractère plus ou moins allongé des structures que nous tentons de recouvrer. Mais avant de définir explicitement ces propriétés, il est important de rappeler que les mammographies sont des images qui sont corrompues par un bruit assez important, un flou non-négligeable et des artéfacts rendant ainsi presque impossible l'annotation des spicules. Toutefois, afin d'extraire les structures les plus réalistes possibles, nous avons imposé quelques contraintes qui se sont avérées utiles pour les mammographies. Premièrement, nous avons supposé que les structures anatomiques présentes dans une mammographie sont allongées, ce qui implique que les segments qui composent chacune de ces structures possèdent une orientation similaire. Ensuite, partant du constat que les occultations ont pu masquer les éventuelles relations entre les segments de l'image, ajouter aux informations perdues du fait de la discrétisation, on a supposé que des segments appartenant à la même structure géométrique ne devraient pas être distants d'une certaine distance. Prenant

en considération ces éléments, nous avons tenu à favoriser les configurations qui permettent d'assembler les segments de faible courbure, c'est-à-dire les segments voisins qui ont des orientations similaires, tout en assurant une continuté, ce qui implique que les segments de lignes doivent être connectés. Cela nous conduit donc à définir deux contraintes, une sur l'orientation et l'autre sur la distance.

Le fait d'utiliser une transformée de Radon locale nous permet de pouvoir réaliser une approximation des courbes de faible orientation par des segments ; cependant, il peut arriver qu'on se retrouve avec plusieurs segments de lignes ayant différentes orientations ; on a donc jugé utile de contraindre les orientations des segments de lignes adjacents. Cela a donc nécessité de pénaliser la non-similarité angulaire par le biais de la formule ci-dessous :

$$\mathcal{V}_o(s_n, s_n^+) = (\min(|\phi_n - \phi_n^+|, 180 - |\phi_n - \phi_n^+|)/90)^2 \ . \tag{3.29}$$

avec ϕ_n et ϕ_n^+ les orientations des deux segments s_n et s_n^+.

Les discontinuités entre les segments ont été également pénalisées de façon proportionnelle aux distances entre les segments qui appartienent à la même structure géométrique :

$$\mathcal{V}_d(s_n, s_n^+) = \min_{x \in s_n, y \in s_n^+} distance(x, y) \ . \tag{3.30}$$

En raison des fortes non-linéarités présentes dans l'expression $\mathcal{J}$ nous avons préféré utilisé le recuit simulé pour l'optimisation de ce terme d'énergie ; de plus étant donné la capacité du recuit simulé à converger vers un minimum global moyennant un temps de calcul plus important, nous avons jugé plus pertinent de le choisir à la place de l'ICM.

3.10.2 Détection des segments convergents

Dans le but de sélectionner les structures qui sont des spicules, nous avons dans un premier temps segmenté la masse à l'aide de l'approche que nous avons illustré dans la partie sur la segmentation des masses et qui possède l'avantage de surpasser les méthodes de segmentation actuelles dédiées à la mammographie. La tâche de la segmentation est très importante car elle nous permet d'extraire avec une très bonne précision les contours de la masse et de déterminer avec exactitude le centre de la masse vers lequel les spicules doivent converger. Pour y arriver, nous nous sommes servis de la modélisation a contrario proposée par [Palma et al., 2009a], dont nous proposons une brève description ci-dessous :

[Palma et al., 2009a] ont été les premiers à appliquer les méthodes a contrario à la détection des masses spiculées et des distorsions architecturales. Les auteurs ont alors modélisé les zones de convergence par des anneaux formés de deux cercles concentriques de rayon r et αr, avec $\alpha \in]0, 1[$, dans lesquels certains points sont orientés vers l'anneau interne. Ensuite, pour chaque pixel de l'image, l'orientation est calculée à l'aide du gradient – lequel gradient est plutôt calculé sur un bloc de 2×2 pixels qui ne s'intersectent pas, afin de s'assurer de l'indépendance entre les gradients des différents pixels. A la suite de cette étape, les auteurs obtiennent une carte des orientations, dont chaque valeur est différente des autres. Le modèle utilisé est basé sur l'indépendance entre les pixels et l'uniformité de l'orientation. Par la suite, pour déterminer les zones de convergence, ils comptent le nombre de pixels q dont l'orientation θ vérifie la condition $\tan(\theta)\|\vec{cq}\| \leq \alpha r$ et dont la localisation satisfait la relation $\alpha r \leq \|\vec{cq}\| \leq r$ où c est le centre

de l'anneau et q un point quelconque de cet anneau. Avec le modèle a contrario, la probabilité qu'un pixel q situé dans l'anneau soit orienté vers le centre de celui-ci est :

$$\mathbb{P}(\tan(\theta)\|\vec{cq}\| \leq \alpha r \wedge \alpha r \leq \|\vec{cq}\| \leq r) = \frac{2}{\pi} \times \tan^{-1}(\frac{\alpha r}{\|\vec{cq}\|}) \ . \tag{3.31}$$

Cette probabilité dépend de la distance entre le pixel q et le centre c. On note $Z_{c,r}$ la variable qui représente le nombre de pixels de l'anneau qui sont orientés vers c et qui désigne la somme sur tous les pixels de l'image des fonctions indicatrices $\mathbb{I}(\tan(\theta)\|\vec{cq}\| \leq \alpha r \wedge \alpha r \leq \|\vec{cq}\| \leq r) = \frac{2}{\pi} \times \tan^{-1}(\frac{\alpha r}{\|\vec{cq}\|})$. La loi de $Z_{c,r}$ est obtenue par sa fonction génératrice $\mathbb{G}_{Z_{c,r}}$, qui est calculée comme le produit des fonctions génératrices des indicatrices mentionnées plus haut. Lesdites fonctions sont des variables indépendantes sur les pixels q situés à l'intérieur de l'anneau. Enfin, les auteurs ont déterminé la valeur λ_r de $Z_{c,r}$ à partir de laquelle l'évènement $Z_{c,r} \geq \lambda_r$ peut être jugé suffisamment improbable pour suspecter une lésion. Ce seuil est donné par [Desolneux et al., 2000], qui définit les évènements ε-significatifs comme les évènements dont le nombre d'occurence est inférieur à ε :

$$\lambda_r = min\{\lambda_r \in \mathbb{N} | \mathbb{P}(Z_{c,r} \geq \lambda) \leq \frac{\varepsilon}{M}\} \tag{3.32}$$

avec M le nombre de tests.

En ce qui nous concerne, la procédure est légèrement différente de celle décrite plus haut, car à l'exception de [Palma et al., 2009a] nous ne considérons que les pixels qui sont les extrémités des structures curvilignes détectées ; ceci nous permet à la fois de simplifier le traitement des structures et de réduire le temps de calcul qui peut exploser en fonction de la taille de l'image. En plus de ce modèle, nous avons pris en compte l'intersection du dernier segment de la structure géométrique avec le contour de la masse afin de nous assurer de l'existence d'un lien entre les deux.

Il est important de remarquer que la segmentation de la masse et l'annotation des spicules ont été réalisées séparément dans l'objectif d'éviter l'influence significative de la suppression du bruit et du rehaussement de contraste sur le modèle a contrario.

3.11 Extraction des spicules : résultats et discussion

L'approche proposée a ensuite été testée sur 100 images de la DDSM dont 5 malignes. Nous avons travaillé plus spécifiquement sur les ROIs de dimension 300×300 pixels, puis sur chaque image, des patchs de taille s×s sont extraits, avec s=20 et δ=$\frac{1}{4}$ a été retenue comme l'intersection maximale qu'il peut y avoir entre deux patchs. Dans le but de produire la sortie de chaque image, les paramètres suivants ont été utilisés : $\beta_o = 3.0$ et $\beta_d = 0.2$, respectivement les poids sur l'orientation et la distance, obtenus de façon empirique, tandis que le nombre d'itérations du recuit simulé a été fixé à 50. Enfin, pour la transformée de Radon, nous avons utilisé un intervalle d'angle de $[0, 180)$ et autorisé qu'un nombre minimum $k = 4$ segments puissent former une structure curviligne.

Soit TP le nombre de vrais positifs, (*c-à-d.* le nombre de structures curvilignes qui sont vues comme des spicules par les radiologues), FN le nombre de faux négatifs, (*c-à-d.* le nombre de spicules évidentes qui sont manquées par notre approche) et FP le nombre de faux positifs, (*c-à-d.*

TABLE 3.5 – Comparaison de quelques travaux de recherche sur la base des métriques utilisées

Méthodes	$\mathcal{P}$	$\mathcal{R}$	FPI
Notre approche	0.83	0.9	0.2
Muralidhar *et al*	0.76	0.84	N/A
Zwiggelaar *et al*	N/A	N/A	0.22
Karssemeijer *et al*	N/A	N/A	1

le nombre de structures curvilignes qui sont obtenues par notre méthode et qui ne correspondent pas à des spicules). Nous avons évalué la pertinence de notre méthode par trois métriques que sont : la précision ($\mathcal{P} = \frac{TP}{TP+FP}$), le rappel ($\mathcal{R} = \frac{TP}{TP+FN}$) et le taux de faux positifs par image (FPI) ($\frac{FP}{TP}$). Les valeurs obtenues sur notre jeu de données et ceux obtenus par d'autres auteurs sont repertoriées dans le tableau 3.5.

Malgré les différences de composition dans les jeux de données et les différentes métriques utilisées, les résultats présentés dans le tableau 3.5 indiquent que notre méthode est meilleure que celles proposées dans l'état-de-l'art, même si, à ce stade, nos résultats ne peuvent être considérées que comme préliminaires. Ainsi, avec notre jeu de données, nous avons obtenu une précision d'une valeur de 0.83 et 0.9 pour le rappel contre respectivement 0.76 et 0.84 pour [Muralidhar et al., 2010] qui ont testé leur méthode sur un jeu de données de 20 ROIs.

Par la suite, nos résultats ont été favorablement comparés à ceux fournis dans la littérature par [Zwiggelaar et al., 1999] et [Karssemeijer et te Brake, 1996] sur différents jeux de données, et qui ont obtenus respectivement un FPI de 0.22 pour une région de 12mm et 1, tandis que nous avons obtenu un FPI de 0.20. Il est important de souligner que contrairement à [Karssemeijer et te Brake, 1996], notre méthode traite toutes les formes de structures anatomiques, *c-à-d.* à la fois les lignes droites et les structures curvilignes. Néanmoins, il faut noter que contrairement à [Karssemeijer et te Brake, 1996], nous n'essayons pas de détecter les régions suspectes (*c-à-d.* masses malignes avec des motifs étoilés) dans une mammographie entière, mais seulement de détecter les spicules reliées à une masse maligne, extraite à une étape précédente ; ce qui peut également expliquer nos résultats.

Par ailleurs, nous avons évalué le temps de calcul de notre approche ; en effet, l'une des forces de notre méthode est sa capacité à extraire des spicules tout en maintenant un temps de calcul très faible. Ainsi, pour une image de 300×300 pixels, son traitement requiert 4 secondes et environ 1 minute pour une image de 3000×4000 pixels. Afin de nous positionner par rapport à l'état-de-l'art, nous avons comparé notre temps de calcul à celui obtenu dans un papier récent de [Berks et al., 2011] dans lequel les auteurs ont utilisé une méthode d'apprentissage supervisée basée sur l'extraction de descripteurs par des ondelettes (*dual tree complex wavelet transform*, DT-CWT) couplée avec un classifieur *random forests* pour classifier les pixels comme appartenant à des spicules ou à des tissus sains. Dans un premier temps, ils ont crée un jeu de données synthétiques contenant des structures curvilignes proches de celles rencontrées en mammographie. Ensuite, ils ont extrait un ensemble de descripteurs sur les images synthétiques à l'aide desquels ils ont pu classifier les pixels comme appartenant soit à une spicule soit à une structure anatomique quelconque. Le même protocole a été utilisé sur des mammographies réelles. Les auteurs ont alors évalué le temps de calcul qu'ils ont estimé à 1 heure 30 minutes pour la classi-

fication des structures anatomiques présentes sur une mammographie de 2000×3000 pixels. Ceci nous permet d'affirmer que notre méthode est significativement plus rapide que celle décrite par [Berks et al., 2011].

Un exemple du résultat obtenu après exécution de notre méthode est illustré à la figure 3.6. Visuellement, les structures appropriées qui semblent satisfaire les contraintes que nous avons imposées ont été extraites. Aussi, pouvons-nous constater que nous avons mis en place un outil rapide et puissant capable d'annoter les spicules sans ajouter des artefacts à l'image. Néanmoins,

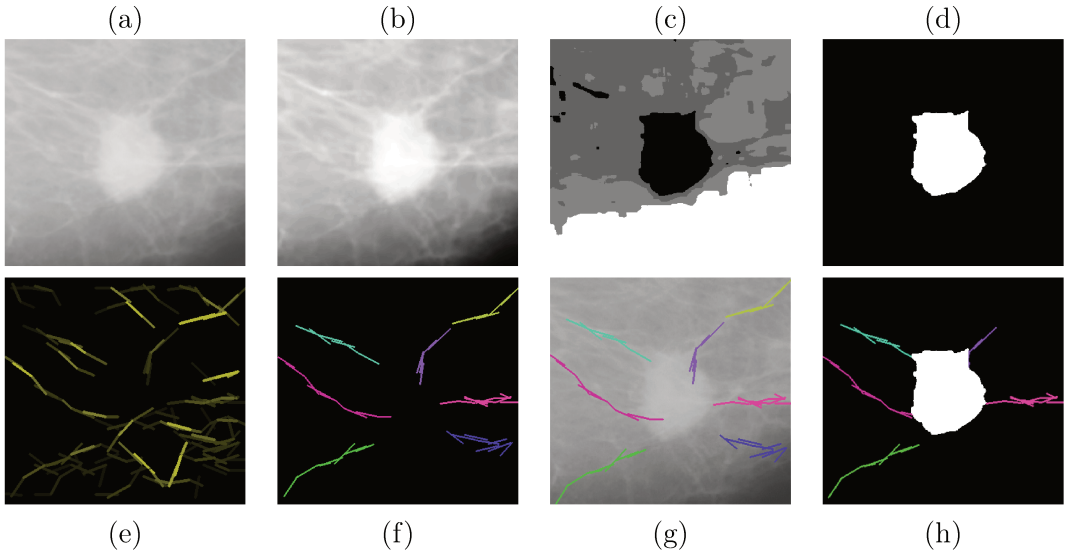


FIGURE 3.6 – Extraction des spicules : (a) Région d'intérêt (ROI), (b) suppression du bruit et amélioration de contraste, (c) segmentation de la masse, (d) version binaire de l'image en (c), (e) discrétisation des structures curvilignes, (f) spicules annotées, (g) superposition des images (a, f) et (h) résultat après réduction des faux positifs.

il convient de reconnaître que malgré son efficacité, notre méthode souffre de sa nature supervisée. En effet, les résultats obtenus sont fortement liés aux informations a priori à propos de la distance et de l'orientation entre les différents segments extraits de l'image, sans oublier le nombre de segments nécessaires pour reconstituer la structure curviligne qui a été discrétisée à l'aide de la transformée de Radon locale. Ces informations sont fournies par le radiologue et sont le fruit de leur expérience en mammographie ; ces éléments peuvent varier fortement d'un radiologue à un autre en fonction de leur expérience en imagerie du sein. En réalité, cet inconvénient cache une autre difficulté qui est le choix de la résolution à laquelle l'image doit être analysée, *c-à-d.* la taille du patch sur lequel la transformée de Radon doit être exécutée. Ainsi, quand la résolution est faible, on détecte des détails très fins, tandis qu'une très grande résolution nous fait perdre des spicules. La solution que nous avons trouvé consiste à extraire des ROIs et de considérer des blocs de 20×20 pixels pour la suite de la procédure ; cela permet de détecter la quasi totalité des spicules de l'image.

3.12 Conclusion

Dans cette partie, nous avons conçu et évalué une méthode automatique de segmentation des masses et des spicules en mammographie. L'approche de segmentation que nous proposons se résume en deux étapes que sont : (i) un rehaussement du contraste de l'image afin d'améliorer la visibilité des objets suspects et (ii) une segmentation s'inspirant des possibilités offertes par une modélisation d'images à l'aide des champs aléatoires markoviens de Pickard. Notre technique présente l'avantage d'être à la fois non-supervisée, efficace, robuste et numériquement rapide. Une comparaison avec les algorithmes de segmentation et de détection des masses les plus populaires et les plus efficaces a montré que non seulement notre méthode est celle qui se rapproche le plus de la segmentation manuelle, mais qu'elle est également très rapide, ce malgré sa composante markovienne. De plus, le fait que son utilisation ne nécessite que la spécification du nombre de niveaux observés dans l'image – tous les autres hyper- paramètres étant spécifiés de façon automatique –, en fait une technique totalement adaptée à tout usage et précisément dans un cadre hospitalier, où l'on désire avoir un outil à la fois temps réel et impliquant très peu ou pas d'interactions avec l'utilisateur. Cette méthode s'est ensuite avérée très utile pour la suite de nos travaux, notamment pour l'extraction des spicules.

Pour finir, une méthode automatique pour l'extraction des spicules sur des mammographies 2D a également été proposée et évaluée. Dans ce cadre, nous avons étendu l'idée de la transformée de Radon locale tout en introduisant les interactions locales entre les segments de chaque structure curviligne par le biais des champs de Markov. Nous avons pu démontrer que la méthode de régularisation proposée, basée sur les champs de Markov peut améliorer la détection des structures géométriques de façon générale. La méthode proposée est une approche supervisée qui s'avère prometteuse et se fonde sur des informations a priori aussi bien sur l'orientation que sur la distance entre les segments dans le but de sélectionner la configuration des structures curvilignes la plus vraisemblable dans une mammographie. De plus, l'usage de ces mêmes contraintes sur les structures curvilignes recomposées, nous a permis d'identifier celles qui sont des spicules. Enfin, la consistance et la précision de notre méthode ont été confirmées par des résultats expérimentaux qui utilisent des métriques comme la précision, le rappel et le FPI.

Chapitre 4

Description des masses et aide à la décision

4.1 Introduction

Après avoir réalisé le pré-traitement des images mammographiques suivie ensuite de la segmentation des masses et de l'extraction des spicules dans les chapitres précédents, nous abordons dans cette partie la dernière phase de la conception de notre système de DAOx. Il s'agit de l'extraction des descripteurs qui doivent servir à la caractérisation des masses et à une prise de décision de notre système à propos de la nature de la masse.

L'extraction des descripteurs est une étape fondamentale dans tout système de traitement d'images mammographiques, car ce sont eux qui vont traduire l'aspect pathologique et la spécificité des lésions par des quantités facilement utilisables par les algorithmes de classification. Aussi, plus les descripteurs seront en mesure d'appréhender la nature des lésions, plus le résultat de la classification sera précis. Afin de décrire au mieux les lésions mammaires, les chercheurs en vision par ordinateur se sont servis de la description du BI-RADS qui met l'accent sur la forme, le contour et la densité des masses. Ils ont alors investis beaucoup d'énergie dans la conception et l'utilisation des descripteurs de forme, de contour, d'intensité et de texture afin de classifier les masses comme bénigne ou maligne ou de classifier les différentes densités présentes dans chaque image. Les résultats obtenus se sont révélés prometteurs et ont suscité un grand engouement dans la communauté scientifique. Toutefois, les systèmes conçus sur de telles approches souffrent de plusieurs limitations qui justifient leur absence en milieu médical. Parmi ces limitations, nous pouvons évoquer le fait que les performances obtenues varient en fonction des descripteurs utilisés et du jeu de données ayant servi pour la validation, de plus, la combinaison de certains descripteurs peut dégrader les performances de ces systèmes, ce qui oblige à sélectionner de façon minutieuse les descripteurs les plus pertinents.

Face à ces difficultés, certains chercheurs ont préféré concevoir des systèmes de DAOx basés sur des techniques de *Content based image retrieval* (CBIR) qui auraient pour avantage de proposer aux radiologues des images disponibles dans une base d'images médicales, dont le contenu est connu et qui seraient similaires à une image requête pour laquelle le radiologue aurait des doutes. Toutefois, cette approche soulève également des problèmes de temps de recherche et de mesure de similarité adéquate entre l'image requête et celles contenues dans la base, sans oublier l'épineuse question du fossé sémantique.

De plus, à notre connaissance, à ce jour, aucune des approches proposées dans la littérature n'aborde le problème de la classification des masses malignes de classes IV et V – *c-à-d.* masses spiculées et distorsions architecturales – entre elles, et ce en dépit du fait que les distorsions architecturales soient les cas les plus manqués par les radiologues. Pour pallier les problèmes

évoqués ci-dessus, nous nous sommes intéressés à une approche d'apprentissage de représentation basée sur une extraction non-supervisée des descripteurs couplée à un classifieur efficace du type SVM. Cette approche est fondée sur une prise de décision sur la forme, sur la texture et les spicules afin de classifier les masses malignes entre elles.

Dans un premier temps, nous allons décrire brièvement les descripteurs de forme et de texture, puis, nous allons faire une synthèse des stratégies de classification des masses et des densités d'une image mammographique. Ensuite, nous nous intéresserons aux systèmes de DAOx basés CBIR. Puis, nous mettrons en évidence les difficultés liées à l'utilisation des approches existantes. Suite à cela, nous détaillerons la solution que nous avons proposé pour discriminer les masses malignes entre elles. Nous finirons par une évaluation de notre méthode en comparaison avec les techniques proposées dans l'état-de-l'art.

4.2 Description de la forme

Le degré de malignité d'une masse augmente avec la complexité de son contour. La figure 4.1 ci-dessous donne un léger aperçu de la diversité des contours qu'on peut retrouver en mammographie ; on y remarque clairement que plus la forme et le contour de la masse sont irréguliers, plus le risque qu'on soit en présence d'un cancer du sein est grand.

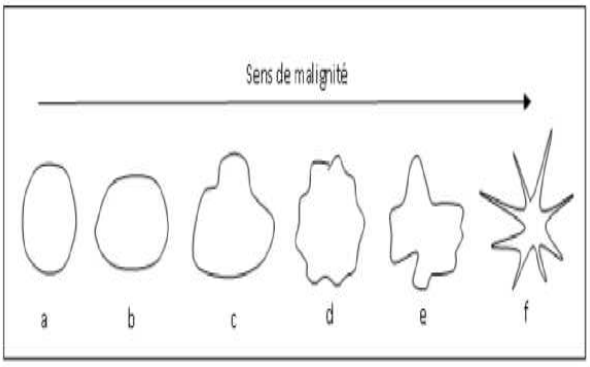


FIGURE 4.1 – Relation connue entre la complexité du contour et la malignité. (a) masse circonscrite circulaire, (b) masse circonscrite ovale, (c) masse macro lobulée, (d) masse micro lobulée, (e) masse irrégulière, (f) masse spiculée.

Afin de quantifier de façon effective la nature du contour des masses, plusieurs travaux se sont focalisés sur la conception de descripteurs de forme permettant de mieux appréhender l'information véhiculée par le contour et la forme des masses. De façon générale, l'aspect des opacités est modélisé soit à travers le contour ou l'intérieur de la masse. Que ce soit l'une ou l'autre de ces modélisations, les descripteurs de morphologie se basent sur des caractéristiques standards ou conçues spécifiquement pour la mammographie. Il existe dans la littérature un nombre important de descripteurs de morphologie qui ont été proposés et que nous mentionnerons dans la suite.

4.2.1 Descripteurs standards

Il y a un nombre relativement important de descripteurs standards utilisés dans divers types d'applications de la vie courante qui sont fréquemment exploités en imagerie du sein. Au nombre de ceux là, nous pouvons citer : l'*aire*, *c-à-d.* le nombre total de pixels contenus dans la région de la masse segmentée, le *périmètre*, qui est le nombre de pixels situés sur le contour de la masse. Il y aussi la *circularité*, qui correspond au degré de ressemblance entre la forme de la masse et un cercle ; de valeur comprise entre 0 et 1 – plus la région de la masse se rapproche d'un cercle, plus on tend vers 1 – ; elle est souvent utilisée en mammographie car une masse bénigne a plus de chance d'avoir une forme circulaire et régulière, ce qui permet de classifier les cas facilement séparables. La *rectangularité* ou boîte englobante, sert quant à elle à décrire le degré d'allongement d'une région. On peut également citer les *descripteurs de Fourier*, qui recherchent les fréquences des harmoniques qui composent le contour d'un objet ; ils sont cependant sensibles au bruit et aux transformations géométriques. La *compacité* qui évalue la complexité du contour par rapport à l'aire et l'*excentricité* – à savoir le rapport entre la longueur du grand axe et celle du petit axe de l'objet – sont eux aussi souvent employés en mammographie.

Pour finir, il y a les moments de Hu [Hu, 1962], ceux de Zernike [Teague, 1980] et de Krawtchouk [Yap et al., 2003] ; que nous avons décidé de détailler en annexe (A), en raison de leur utilisation récente en mammographie.

4.2.2 Descripteurs spécifiques

Malgré l'efficacité des descripteurs standards exposés plus haut, dans diverses applications de reconnaissance de forme ; ils offrent pour la plupart des performances médiocres dans la classification des masses mammaires. Afin d'améliorer la discrimination entre les masses mammaires bénignes et malignes, les chercheurs se sont intéressés à la conception de descripteurs plus précis et subsidiaires du point de vue de la forme. Au nombre de ces descripteurs, nous pouvons citer les descripteurs du *Normalized Radial Length* (NRL), ceux du *Normalized Chord Length* (NCL) – dont la description est donnée en annexe (A) et bien d'autres encore.

4.3 Description de la texture

En vision par ordinateur, il n'existe pas une définition unanimement acceptée de la texture. Les définitions mathématiques fondées sur des propriétés statistiques sont soit trop générales et donc imprécises, soit trop restrictives pour s'adapter aux différents types de texture qui existent. Nous pouvons néanmoins citer celles qui suivent : «*une texture est un champ de l'image qui apparaît comme un domaine cohérent et homogène,* c-à-d. *formant un tout pour un observateur*» ou «*la texture est l'ensemble des propriétés de voisinage locales de niveaux de gris*» [Livens et al., 1997]. C'est la propriété de cohérence de la texture placée dans un contexte d'un tout homogène par l'œil humain qui sera recherchée le plus souvent en traitement d'images, dans le but d'isoler les textures, soit pour segmenter l'image, soit pour reconnaître les régions.

Il existe plusieurs façon de représenter la texture, parmi lesquelles on peut citer l'utilisation des filtres de Gabor, des ondelettes, des matrices de co-occurences de type *grey-level co-occurence matrix* (GLCM) [Haralick et al., 1973], *run length statistics* (RLS) [Galloway, 1975] ou *gray level*

difference statistics (GLDS) [Weszka et al., 1976] et les descripteurs associés qui permettent une représentation plus condensée de l'information contenue dans ces matrices. Une description détaillée des filtres de Gabor, des ondelettes est proposée en annexe (B).

4.4 Evaluation de la performance de la classification

Une fois les différents descripteurs extraits et la classification réalisée, il existe diverses techniques permettant de vérifier la fiabilité des descripteurs utilisés et la pertinence de l'approche de classification. Parmi les techniques utilisées, les plus connues sont : la matrice de confusion, la courbe ROC et l'aire sous la courbe ROC.

4.4.1 La matrice de confusion

La matrice de confusion ou tableau de contingence sert à évaluer la qualité d'une classification. C'est un tableau 2D de taille K×K – où K est le nombre total de classes –, utilisé pour disposer les résultats d'une expérience de classification. Ainsi, la valeur à la ligne i et à la colonne j indique le nombre de fois qu'un objet dont la vraie classe est i, est classifié comme appartenant à la classe j. La diagonale principale de cette matrice indique le nombre de cas pour lesquels le classifieur a bien réussi sa tâche ; un classifieur parfait devrait afficher zéro pour tous les éléments situés en dehors de la diagonale principale.

4.4.2 La courbe ROC

De son nom complet *Receiver Operating Characteristic* (ROC), la courbe ROC est une technique de représentation graphique des performances d'un classifieur à deux classes [Metz, 1978] ; qui est une méthode d'évaluation par excellence utilisée dans la quasi totalité des travaux portant sur la classification. Elle représente la sensibilité en ordonnée en fonction de la quantité (1-spécificité) en abscisse. La sensibilité qui n'est rien d'autre que le taux de vrais positifs (TVP), représente la capacité d'un examen de diagnostic à fournir un résultat positif en présence de la maladie. Quant à la spécificité, elle équivaut à la capacité d'un examen à produire un résultat négatif en abscence de la maladie.

$$\text{TVP} = \text{Sensibilité} = \frac{VP}{VP + FN} \ . \tag{4.1}$$

$$\text{TFP} = 1 - \text{Spécificité} = \frac{FP}{FP + VN} \ . \tag{4.2}$$

avec VP (vrai positif) le nombre de masses malignes qui sont classées malignes, FP (faux positif) le nombre de masses bénignes qui sont classées malignes, VN (vrai négatif) le nombre de masses bénignes classées bénignes et FN (faux négatif) le nombre de masses malignes qui sont classées comme bénignes.

4.4.3 La mesure de la performance

Dans l'objectif d'avoir une description quantitative de la performance de la classification, à l'aide de la représentation ROC, on se sert de l'aire sous cette courbe, qu'on note généralement

A_z. L'évaluation de l'aire sous la courbe ROC peut être obtenue à partir du calcul de l'aire des rectangles juxtaposés occupant l'aire de la courbe. Une technique plus précise pour le calcul de A_z est la méthode trapézoïdale. Cette méthode consiste à calculer l'aire de chaque trapèze formé par deux points successifs de la courbe et leurs abscisses. Vu que l'aire sous la courbe est une portion de l'aire d'un carré unitaire, sa valeur varie toujours entre 0 et 1. Plus la valeur de A_z est proche de 1, plus le résultat de la classification est meilleur. Dans le cas parfait, l'aire sous la courbe vaut 1, tandis que dans la pire des situations, la valeur de l'aire équivaut à 0.5. Ainsi, la courbe ROC d'un bon classifieur monte rapidement vers le coin supérieur gauche. De façon explicite, l'interprétation que l'on peut faire des différentes valeurs de l'aire sous la courbe est la suivante :

— Si $A_z = 0.5$: le résultat de la méthode dont la performance est évaluée n'est d'aucune utilité,

— Si $0.5 < A_z < 0.7$: le résultat est peu informatif,

— Si $0.7 \leq A_z < 0.9$: le résultat est informatif,

— Si $0.9 \leq A_z < 1$: le résultat est très informatif,

— Si $A_z = 1$: le résultat est parfait.

4.5 Classification des masses mammaires

Afin de détecter et de classifier les tumeurs bénignes et malignes, [Brzakovic et al., 1990] ont d'abord segmenté les masses contenues dans les images à l'aide d'une approche de muli-résolution fondée sur une pyramide floue et dont une description plus détaillée se trouve au chapitre sur la segmentation. Une fois les masses obtenues, ils ont extraits plusieurs descripteurs à savoir la superficie, la taille et les variations d'intensité des pixels situés sur les bords des masses. Ces descripteurs sont ensuite utilisés dans un modèle de classification hiérarchique à quatre niveaux pour décider de la nature de la masse. Le premier niveau porte sur la superficie de la masse, qui est utilisée comme entrée d'un classifieur Bayésien afin de décider de la suite de la procédure de classification. Ils ont abouti à un taux de 85% de bonne classification sur un jeu de données de 25 images.

Après avoir extrait l'intensité moyenne et des descripteurs liés au contraste à la texture et à la spiculation du contour, [Huo et al., 1998] ont utilisé une approche de classification hybride exploitant successivement des règles de décision et un classifieur ANN dans le but de réduire les limitations liées à l'utilisation séparée de chacun de ces classifieurs. Les performances de leur classifieur est évaluée par la courbe ROC, avec une valeur de A_z=0.94. Ils ont également noté que les descripteurs liés à la spiculation du contour s'avèrent être très puissants pour discriminer entre masse bénigne et maligne, mais quand il s'agit de masses bénignes et malignes non spiculées, ce sont les descripteurs de contraste, de texture et l'intensité moyenne qui s'en sortent le mieux.

Dans le même ordre d'idée, [Rangayyan et al., 1997] ont proposé des descripteurs de forme basés sur les moments et un descripteur appelé *acutance* dont le rôle est de caractériser la transition entre la ROI et la masse en se basant sur la densité de chaque pixel du contour de cette dernière. Afin de prouver la supériorité de leurs descripteurs, ils ont extrait d'autres descripteurs de forme comme les descripteurs de Fourier, les moments de Hu, la compacité et les

descripteurs du NCL. Puis à l'aide de la mesure de Fisher utilisant la distance de Mahalanobis comme critère de sélection, ils ont évalué les performances de chacun des descripteurs. Les résultats obtenus pour une classification bénigne/maligne démontrent que l'*acutance* possède la meilleure performance soit 94.9% et que ce score ne change pas, même après une combinaison de l'*acutance* soit avec la compacité, soit avec les descripteurs de Fourier ou les moments. Par contre, quand il s'agit de comparer les masses dont le contour est circonscrit avec ceux dont le contour est spiculé, ce sont les descripteurs de Fourier qui ont le meilleur score de 92.2%. Toutefois, ces résultats sont à nuancer du fait de la taille du jeu de données utilisé, à savoir 39 ROIs de la MIAS et 15 d'une base privée.

Dans le même ordre d'idée, [Sahiner et al., 1998b] ont proposé une approche de caractérisation de la texture des masses afin de les classifier comme bénigne ou maligne. Après avoir extrait des descripteurs de texture de type GLCM et *run length statistics* (RLS) [Galloway, 1975]. Ils ont procédé à une sélection des descripteurs de texture les plus pertinents en se servant de l'algorithme génétique, couplée avec une mesure dite de surface partielle de la courbe ROC comme critère de sélection, dans le but d'améliorer la qualité de l'algorithme génétique. Par la suite, ils ont procédé à la classification des descripteurs obtenus à l'aide du critère de Fisher. Finalement, une comparaison avec le classifieur *linear discriminant analysis* (LDA) sur des descripteurs préalablement retenus – en se servant du lambda de Wilks comme critère de sélection – a été faite. Les auteurs sont parvenus à la conclusion selon laquelle l'utilisation d'un algorithme génétique couplée avec la surface partielle de la courbe ROC pour une classification avec la LDA offraient de bien meilleurs résultats en raison de l'élimination efficace des descripteurs dont le pouvoir discriminant réduisait les performances des descripteurs les plus pertinents.

Quant à [Oliver et al., 2010c], après avoir segmenté les masses à l'aide d'une méthode de *level-set*, ils ont extrait sur chacune d'elles les treize premiers moments de Zernike. Puis, en se servant du classifieur Gentleboost [Friedman et al., 2000] – qui repose sur le fait que l'utilisation de plusieurs classifieurs de faibles performances peut produire un classifieur de bonne qualité –, ils ont classifié les masses de la DDSM et de la MIAS comme bénigne ou maligne en fonction de leur morphologie. En passant par une validation croisée, ils ont obtenu une valeur de A_z de 0.89 pour la MIAS et 0.72 pour la DDSM.

De leur côté, [Guliato et al., 2008] ont développé une méthode de description de la masse basée sur des modèles polygonaux du contour de cette dernière et qui préservent les spicules de même que les détails importants pour le diagnostic. Une fois la modélisation faite, ils procèdent à l'extraction de plusieurs descripteurs de spiculation dont l'*acutance*. L'évaluation de la pertinence de leur technique sur un jeu de données composé de masses maligne et bénigne, donne un score de A_z de 0.94. Pour leur part, [Narváez et Romero, 2012] ont préféré utilisé les moments de Krawtchouk extraits sur les contours de masses bénigne et maligne, qu'ils ont ensuite utilisé comme entrée d'un classifieur kNN. Puis, une étude comparative a été menée avec les moments de Zernike afin d'évaluer la performance des moments de Krawtchouk. Ils ont alors remarqué que les moments de Krawtchouk étaient meilleurs que ceux de Zernike avec un A_z de 0.93 contre 0.85. [Kachouri et al., 2012] ont proposé de leur côté, trois descripteurs de morphologie à savoir les *points terminaux d'un squelette* (SEP), la *sélection des protubérances* (PS), et le *descripteur des masses spiculées* (SMD). Après extraction sur un jeu de données d'images bénigne et maligne issues de la DDSM, ils ont obtenu des scores de A_z de 0.92, 0.93 et 0.97 respectivement pour la SEP, la PS et la SMD. Ils ont également noté que ces performances étaient supérieures à celles

affichées par les descripteurs standards comme la compacité.

Dans un régistre légèrement différent, [Sahiner et al., 1996a] ont conçu une approche dans le but de classifier les ROIs des images mammographiques en région saine ou région maligne. Pour y parvenir, ils ont crée pour chaque ROI deux matrices l'une contenant les descripteurs de texture de type GLCM et l'autre des descripteurs de texture de type *gray level difference statistics* (GLDS) [Weszka et al., 1976]. Puis, ils ont évalué l'efficacité de plusieurs classifieurs ANN prenant en entrée divers types d'images. Les expériences ont été menées sur des jeux de données composées de ROIs saines et malignes, tandis que les performances de leur méthode ont été évaluées par la courbe ROC. Ils ont remarqué que le classifieur ANN qui donnait le meilleur score, soit un A_z de 0.87 est celui qui prend en entrée l'image de départ et les deux matrices de texture.

4.6 Classification de la densité mammaire

Depuis la création de correspondances entre la densité mammaire et le risque de développer un cancer, plusieurs travaux ont été menés dans le but de procéder à une classification automatique de la densité. Ces études se sont focalisées sur l'utilisation de l'histogramme de niveau de gris pour la classification des motifs présents dans une mammographie. C'est ainsi que [Karssemeijer, 1998] a proposé une méthode de détermination automatique des différentes densités du sein. Pour y arriver, il a d'abord réalisé un pré-traitement de l'image – pour supprimer le muscle pectoral –, à la suite duquel, il a extrait l'écart-type et le *skewness* comme descripteurs de densité ; ces descripteurs seront ensuite classifiés par le kNN [McLachlan, 2004]. Les résultats obtenus varient entre 67 et 80% de bonne classification pour l'ensemble des quatre classes de densité en fonction de la base d'image utilisée.

Cependant, des études ont prouvées que les informations relatives à l'histogramme de l'image ne sont pas suffisantes pour réaliser une classification de la densité qui puisse respecter le BI-RADS [Oliver et al., 2005]. En effet, bien qu'une image mammographique possède une densité répartie en quatre classes, les histogrammes de ces classes sont similiaires en terme de valeur de niveau de gris. Dans l'objectif de contourner ce problème, certains chercheurs se sont intéressés à l'utilisation de la texture. C'est ainsi que [Miller et Astley, 1992] ont décidé d'évaluer l'impact de la texture dans la discrimination entre les densités graisseuses et denses. Ils ont utilisé cinq masques de Law [Laws, 1980] qu'ils ont convolué avec l'image de départ, et sur chaque fenêtre de la taille du masque, la variance est calculée. Les différentes valeurs obtenues sont ensuite utilisées pour segmenter la densité de l'image. Enfin, pour chaque image test, les mêmes descripteurs sont extraits et classifiés à l'aide d'un classifieur Bayésien afin de fournir la probabilité qu'un pixel puisse avoir une densité graisseuse ou dense.

Quant à [Bovis et Singh, 2002] ils ont d'abord extrait différents descripteurs de texture sur les matrices de GLCM, suivis de l'énergie spectrale totale après la transformée de Fourier, puis l'énergie totale après la convolution de l'image avec les masques de Law et des descripteurs standards sur le résultat de la transformée en ondelettes. Ensuite, ils ont procédé à une réduction de dimension de leur vecteur de descripteurs à l'aide de l'ACP. Dans le but d'améliorer les performances et la robustesse de leur approche, ils ont exploité un paradigme de combinaison qui consiste à associer à travers des règles de somme et de produit – proposées

par [Kittler et al., 1998] – les sorties de différents classifieurs. Ce faisant et en utilisant onze classifieurs de type ANN sur différents jeux de données de la base d'apprentissage et de test, ils obtiennent des taux de bonne classification de 96.7% et 71.4% respectivement pour deux et quatre classes du BI-RADS, contre 77.3% et 58.3% quand ils ne procèdent pas à une combinaison des sorties des classifieurs.

[Petroudi et al., 2003], ont développé – dans le but de classifier la densité mammaire dans les quatre classes ACR I-IV – une approche qui utilise les *textons* comme descripteurs. Après avoir supprimé le muscle pectoral et pré-traité l'image, ils ont convolué les bancs de filtres *Maximum Response 8* (MR8) de [Varma et Zisserman, 2002] – composés de filtres Gaussien et Laplacien et qui possèdent l'avantage d'être invariants à la rotation – avec les mammographies. Les descripteurs obtenus après convolution sont regroupés en quatre classes de densité, et pour chacune des classes, dix centres de gravité sont déterminés à l'aide des K-means ; lesquels centres sont appelés les *textons*. Par la suite, un dictionnaire contenant les quarante *textons* est construit [Leung et Malik, 2001]. Pour ce qui est de la classification, ils procèdent comme suit : pour chaque image de la base de test, ils attribuent à chacun des pixels appartenant à la région du sein l'étiquette d'un *texton*, puis ils en déduisent un histograme de textons. Cet histograme est enfin comparé au modèle appris et la mammographie est alors affectée à la catégorie de BI-RADS correspondant au plus proche voisin au sens de la distribution χ^2. Ils ont obtenu un taux global de 76% de bonne classification ; mais avec des scores assez faibles pour les classes de densité qui nous intéresse, soit : 70% et 78% respectivement pour les classes III à IV.

Pour leur part, [Bosch et al., 2006] ont testé les *Scale Invariant Feature Transform* (SIFT) et les textons afin d'évaluer leur performance pour la classification de la densité mammaire. Ils ont ensuite construit deux dictionnaires visuels (BoW) différents – un par descripteur – après avoir procédé à une quantification des descripteurs par K-means. Une fois les dictionnaires obtenus, ils ont exploité le *probabilistic latent semantic analysis* (pSLA) [Hofmann, 2001] afin de retrouver la distribution de chaque tissue dans l'image. Les distributions des tissues ont enfin été utilisées pour la classification ; dans ce cadre, ils ont testé le kNN et le SVM et se sont rendus compte que le SVM donne de biens meilleurs résultats que le kNN à la fois sur la MIAS et le DDSM.

Dans le but d'optimiser les performances de classification de la densité en tenant compte du BI-RADS, [Oliver et al., 2008] ont décidé dans un premier temps, de procéder à la segmentation de l'image en deux régions à savoir une zone dense et une zone graisseuse à l'aide du FCM [Oliver et al., 2005]. Puis, ils ont extrait des descripteurs de texture sur les matrices GLCM d'une part sur la région dense et d'autre part sur la région graisseuse. Ils ont alors réalisé une étude comparative entre d'une part l'utilisation du classifieur kNN sur les descripteurs sélectionnés par SFS (kNN+SFS) et d'autre part la combinaison par une approche d'estimation Bayésienne du classifieur C4.5 et du kNN+SFS. Ils se sont alors rendus compte que la combinaison des deux classifieurs donnaient les meilleurs résultats à savoir 89% et 73% respectivement pour le BI-RADS III et le BI-RADS IV.

4.7 Les systèmes de DAOx basés CBIR

La précision et l'efficacité des systèmes de CBIR dépendent fortement des descripteurs visuels choisis. Aussi, un bon descripteur devrait faire un compromis entre la robustesse de la variance

intra-classe et le caractère discriminant de la différence inter-classes, sans oublier le temps de calcul nécessaire à son extraction. Comme nous l'avons vu plus haut, les descripteurs visuels peuvent décrire diverses propriétés sur toute l'image ou sur une région de celle-ci. Parmi les nombreux descripteurs locaux qui existent, les SIFT sont les plus utilisés en raison de leur excellente robustesse et de leur capacité de discrimination. Du fait de leur grande taille, les SIFT sont souvent quantifiés sous forme de *mot visuel* – par analogie au *mot* dans une recherche textuelle – et ce, de telle sorte qu'une image est caractérisée par un *bag of words* (BOW) *c-à-d.* les mots visuels les plus fréquents ; pour une recherche plus rapide.

Au cours de la dernière décénnie, plusieurs systèmes de DAOx basés CBIR pour la mammographie ont été proposés. Par exemple, [Tourassi et al., 2003] ont proposé un système de recherche par le contenu pour la détection des masses qui se sert de la connaissance experte présente dans les différentes mammographies qui composent la base d'images. Pour y arriver, ils utilisent dans un premier temps le *template matching* afin de retrouver parmi les images contenues dans la base celles qui sont similaires à la ROI requête présentée par l'utilisateur du système. Le choix des images qui ressemblent le plus à la ROI se fait sur la base d'une mesure de similarité appelée *information mutuelle* – qui est issue de la théorie de l'information –, dont le but est d'évaluer dans quelle mesure un objet peut expliquer l'autre auquel il est comparé. Afin de déterminer si la ROI requête contient une masse (peu importe sa nature) ou uniquement des tissus sains, ils ont proposés une mesure de décision qui combine efficacement plusieurs mesures de similarité sur les meilleures correspondances. Plus tard, [Tourassi et al., 2007] ont étudié les performances de huit mesures de similarité basées sur l'entropie, afin d'identifier celles qui possèdent la meilleure précision en ce qui concerne la qualité de la recherche dans la base, et celles qui sont les plus efficaces en matière de détection. Ils sont parvenus à la conclusion selon laquelle, les mesures de Kullback-Leibler donnaient les meilleurs résultats pour la qualité de la recherche, tandis que l'information mutuelle offrait la meilleure performance pour la détection. Pour leur part, [Alto et al., 2005] ont préféré extraire des descripteurs de texture, de forme et de netteté du contour ; ceux liés à l'intensité des pixels, la forme et à la texture ont été fusionnés par [Tao et al., 2007], afin de retrouver les masses similaires à celle contenue dans la ROI requête et de la classifier comme bénigne ou maligne. Pour une meilleure similitude visuelle, [Zheng et al., 2006] ont proposé un système offrant une interaction plus poussée avec l'utilisateur ; système dans lequel, ce dernier est invité à évaluer la nature de la masse spiculée de l'image requête afin que le système de CBIR ne recherche que des correspondances avec des degrés de spiculation similaires. Ce travail a par la suite été amélioré en supprimant de la base de recherche les ROIs qui donnaient les pires scores de similarité [Park et al., 2007].

Par ailleurs, plusieurs travaux ont tenté de retrouver dans les bases d'images, des masses mammaires ayant des caractéristiques similaires d'un point de vue du BI-RADS, telles que la forme, le contour et la pathologie. A titre d'exemple, les descripteurs de forme, d'intensité et de texture ont été combinés à l'aide d'un système de pondération adapté, exploitant également l'interaction de l'utilisateur afin d'optimiser la qualité des correspondances proposées [Wei et al., 2011]. [Narvaez et al., 2011] ont pour leur part proposé une méthode qui commence par fusionner les descripteurs de forme et de texture extraits sur les deux incidences (CC et MLO) du sein pour retrouver les meilleures correspondances, lesquelles images sont ensuite utilisées pour annoter la ROI requête. [Kinoshita et al., 2007] ont quant à eux utilisé des descripteurs d'intensité, de texture, de forme et les mesures de granulométrie afin de rechercher les

mammographies qui ont une composition de tissu similaire. [Liu et al., 2014] ont de leur côté introduit une recherche d'images basée sur une fonction de hachage afin de produire un diagnostic pour les masses contenues dans les ROIs requêtes. De façon plus spécifique, une fonction de hachage s'inspirant de la théorie des graphes et nommée *anchor* [Liu et al., 2011] a été employée pour compresser deux descripteurs, à savoir l'histogramme du SIFT et le GIST en des codes binaires ; finalement la recherche de similarité a été faite dans l'espace de Hamming.

4.8 Le fossé sémantique

La différence fondamentale entre les systèmes basés sur le contenu et ceux basés sur du texte, c'est que l'interaction humaine est une condition indispensable au bon fonctionnement de ce dernier. Ainsi, nous les humains avons tendance à utiliser des descripteurs de haut niveau tels que les mots-clés, les champs lexicaux afin d'interpréter et de mesurer la similarité des images que nous visualisons ; tandis que les descripteurs extraits automatiquement à l'aide de méthodes relatives à la vision par ordinateur, sont pour la plupart des concepts de bas niveau. En général, il n'existe aucun lien entre les concepts de haut niveau et ceux de bas niveau. Bien que bon nombre de méthodes aient été conçues pour décrire la couleur, la forme, le contour et la texture, ces algorithmes sont incapables de modéliser correctement la sémantique des images et font face à de nombreuses limites quand il s'agit de traiter de grandes bases d'images où il y a une forte variabilité à la fois sur la forme, le contour et la texture des objets. Des expériences poussées sur le fonctionnement des systèmes de CBIR ont montrées que les descripteurs de bas niveau échouent souvent à traduire les concepts sémantiques de haut niveau présents dans l'esprit de l'utilisateur de ces systèmes ; ce qui fait que les performances des systèmes de CBIR sont encore loin de satisfaire les attentes des utilisateurs.

Comme on peut le constater à travers la synthèse des méthodes d'aide à la décision en mammographie, les performances de classification des approches d'apprentissage automatique sont fortement liées au choix de la représentation que l'on fait du jeu de données dont on dispose. C'est pour cette raison que beaucoup d'efforts ont été consentis dans la conception des descripteurs de toute sorte, permettant une transformation des données dans le but de retenir l'information la plus pertinente. Cependant, l'extraction des descripteurs exige non seulement une réflexion importante en vue de proposer les descripteurs les plus utiles possibles, mais présente également une faiblesse majeure à savoir son incapacité à extraire et à organiser des descripteurs discriminants de haut niveau sur le jeu de donnée d'apprentissage ; c'est le problème évoqué plus haut à savoir le fossé sémantique. Les travaux réalisés dans la littérature [Krizhevsky et al., 2012], portant sur l'apprentissage profond ou le *deep learning* par le biais des réseaux de neurones à convolution entre autre, ont ouvert la voie vers une approche d'apprentissage automatique de représentation fondée sur une extraction non-supervisée des descripteurs (de bas, moyen et haut niveau), *c-à-d.* indépendante de toute intervention humaine qui pourrait affecter ces performances. La première perspective que nous proposons est donc d'envisager l'usage de ces architectures dans la discrimination entre les masses spiculées et les distorsions architecturales.

Dans un premier temps, nous allons décrire brièvement les ANNs puis nous présenterons de façon détaillée les réseaux de neurones à convolution, à savoir leur fonctionnement, la procédure utilisée pour leur apprentissage et nous finirons en décrivant mettant en œuvre notre idée sur la

combinaison de notre système de DAOe avec les réseaux de neurones à convolution pour une aide au diagnostic appropriée. Une dernière partie sera consacrée aux défis qu'il faut encore relever pour que le *deep learning* puisse permettre de s'affranchir entièrement de l'extraction supervisée des descripteurs et régler le problème du fossé sémantique.

4.9 Les réseaux de neurones artificiels

A l'instar des neurones biologiques, un ANN se compose de plusieurs couches disposées en série ou en parallèle. Chaque couche possède un ensemble de paramètres positifs ou négatifs noté θ permettant de transformer les données reçues en signal de sortie pour la couche suivante. Ainsi, cette architecture est similaire au neurone biologique qui reçoit des signaux par l'intermédiaire de ses dendrites, puis réalise une combinaison au niveau de son noyau afin de produire un signal d'inhibition ou d'excitation qui est envoyé au neurone adéquat par son axone. Par ailleurs, une transformation non-linéaire à l'image du seuil de déclenchement du neurone biologique peut s'ajouter à la transformation linéaire des entrées par les paramètres.

Plus précisément, un ANN est composé de plusieurs couches L ; chaque couche est représentée par $l \in 0, \cdots, L-1$ où la couche 0 est le vecteur d'entrée $x = y^0$ et $L-1$ la sortie à prédire $y = y^{L-1}$. Les couches se situant entre celle d'entrée et celle de sortie sont appellées des couches *cachées* ou latentes, car ce sont les couches pour lesquelles l'ANN apprend les paramètres. La procédure d'apprentissage d'un ANN consiste à estimer les poids W^l et les biais b^l de chaque couche cachée l. Le passage du vecteur d'entrée x en la cible de sortie y se résume en une succession d'opérations :

$$y^{(l)} = s(W^{(l)} y^{(l-1)} + b^l) \ . \tag{4.3}$$

avec $s(\cdot)$ la fonction d'activation. Les non-linéarités les plus courantes utilisées pour la fonction d'activation sont décrites à la section 4.9.2.1.

Parmi les réseaux de neurones artificiels les plus connus, nous avons les réseaux de neurones à convolution qui seront largement abordés dans la suite et les auto-encodeurs.

4.9.1 Les réseaux de neurones à convolution

Les réseaux de neurones à convolution ou *Convolutional Neural Networks* (CNN) sont l'une des approches les plus remarquables de l'apprentissage profond, dans laquelle plusieurs couches composées de neurones sont formées de manière robuste. Ils ont montré qu'ils étaient capables de faire preuve d'une capacité de généralisation impressionnante sur des jeux de données de grande taille comprenant des millions d'images [Krizhevsky et al., 2012], [Simonyan et Zisserman, 2014], [Russakovsky et al., 2015]. Ces résultats proviennent majoritairement de l'architecture particulière des CNNs qui tient compte de la topologie spécifique des tâches relatives au domaine de la vision par ordinateur qui exploitent des images en deux dimensions. D'autres dimensions peuvent également être prises en compte quand il s'agit d'images en couleurs avec plusieurs canaux.

Les CNNs se servent de la très forte corrélation qu'il y a au sein d'une structure locale en deux dimensions en restreignant le champ récepteur des unités cachées à se focaliser sur des variations locales [LeCun et al., 1998]. De part ce fait, un schéma de connectivité locale est appris sur les premières couches cachées afin de décrire des structures simples comme les

bords ou les coins. Le fait d'empiler des couches de convolution, *c-à-d.* augmenter la profondeur, contraint le réseau à apprendre des représentations plus abstraites, plus discriminantes et de haut niveaux, en combinant au sein des couches plus profondes les descripteurs locaux appris au cours des premières couches [LeCun et al., 1998].

Nous proposons dans ce paragraphe de prendre un exemple concret pour illustrer les représentations apprises par un CNN. Dans notre exemple, la tâche d'apprentissage est la reconnaissance de visages qui est illustrée à la figure 4.2. Le CNN doit apprendre à distinguer le visage des autres objets présents dans l'image. Au sein de la première couche – nommée *layer 2* sur la figure, les noyaux de convolution détectent des formes simples comme les bordures, et ceci selon plusieurs rotations. Ces noyaux de convolution sont très similaires au filtre de Gabor dont l'utilisation est courante en vision par ordinateur. Plus la couche cachée est profonde, plus les descripteurs appris ont tendance à être complexes et abstraits. A la seconde couche cachée, les noyaux de convolution résultent d'une convolution des noyaux précédents afin de former des motifs bien plus complexes, qui correspondent à des parties précises du visage, comme les yeux, le nez ou la bouche. Pour finir, la dernière couche cachée montre plusieurs représentations abstraites du visage. Cet exemple démontre la capacité du CNN à apprendre des descripteurs de haut niveau de façon non-supervisée et avec des performances bien meilleures que celles obtenues jusque là dans l'état-de-l'art. Ceci est la preuve que l'utilisation des CNNs peut nous permettre de nous affranchir du fossé sémantique évoqué plus haut.

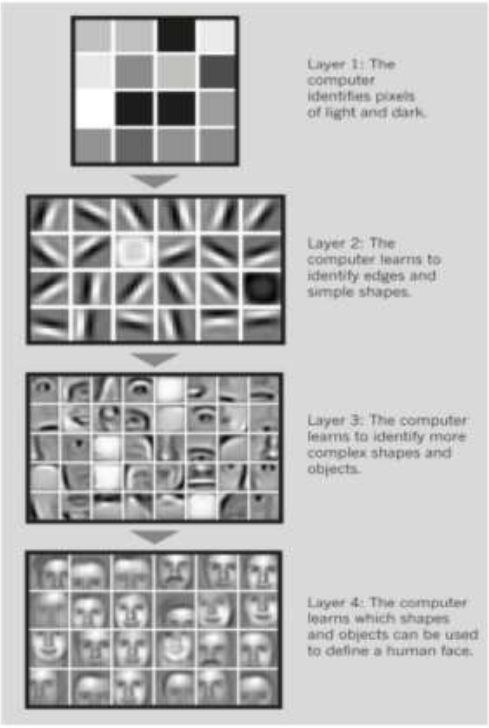


FIGURE 4.2 – Reconnaissance de visages avec un CNN. Image tirée de [Jones et al., 2014].

A tout ceci, s'ajoute le fait que les CNNs sont invariants à la translation en raison des

contraintes spécifiques à leur architecture. Les unités d'une couche de convolution sont disposées dans une topologie en 2D appelée noyau de convolution. A ce titre, une couche de convolution peut avoir un ou plusieurs noyaux de convolution ou filtres. Les paramètres d'un même noyau sont partagés, ce qui signifie que la même procédure est appliquée sur l'image en entrée par une technique de fenêtre glissante. Ce processus de partage des poids pousse le réseau à devenir invariant à la translation puisque les descripteurs appris dans le noyau seront appliqués à toute l'image et non à une région particulière de celle-ci.

A la couche de convolution s'ajoute une couche de sous-échantillonnage qui renforce davantage cette propriété. La couche de sous-échantillonnage réalise une opération de décimation locale où seule la valeur maximale de l'image obtenue après la convolution est conservée. Les trois principes évoqués plus tôt, à savoir la connectivité locale, le partage des paramètres et le sous-échantillonnage, constituent le cœur d'une couche de convolution dans un CNN. Les différentes opérations évoquées dans cette section sont illustrées à la figure 4.3. De plus en plus, une couche de convolution contient typiquement une ou plusieurs opérations de convolution suivie d'un sous-échantillonnage. Le fait d'empiler plusieurs couches de convolution engendre un CNN.

4.9.2 Apprentissage d'un réseau artificiel de neurones

4.9.2.1 Les fonctions d'activation

La fonction d'activation la plus simple est la fonction linéaire $s(x) = x$, qui renvoie tout simplement les valeurs en sortie de la couche sans aucune modification. Cette fonction est souvent utilisée dans des tâches de régression, où la couche de sortie doit fournir une valeur réelle, mais aussi dans la couche située au milieu d'un auto-encodeur empilé. Il existe également d'autres types d'activation pour que les ANNs puissent apprendre les fonctions non-linéaires. Les fonctions *sigmoïde* – comme la fonction logistique, $s(x) = \frac{1}{(1+exp(-x))}$ ou la tangente hyperbolique, $s(x) = \tanh(x)$ – sont parmi les plus connues. La tangente hyperbolique est souvent préférée à la fonction logistique en raison de sa symétrie par rapport à l'origine. Grâce à cette propriété, la tangente hyperbolique a tendance à produire des valeurs d'activation dont la moyenne vaut zéro. Ceci évite à la procédure d'apprentissage d'être ralentie à cause d'une mise à jour des paramètres qui se ferait dans la mauvaise direction [LeCun et al., 2012]. Toutefois, à ce jour, la fonction *rectificatrice linéaire* ou *unité rectificatrice linéaire* (ReLU) est devenue le standard pour l'apprentissage des ANNs. En effet, sa forme particulière amène l'ANN à apprendre des représentations parcimonieuses qui offrent bon nombre d'avantages [Glorot et al., 2011]. La représentation parcimonieuse incite l'ANN à apprendre à séparer les principales causes de variations contrairement aux représentations denses, à modéliser la vraie dimension des données, à mieux distinguer les données ou à propager l'information dans les neurones artificiels. De façon pratique, l'utilisation de la ReLU a montré des résultats meilleurs aux techniques utilisant un pré-apprentissage ; ceci se justifie par le fait que le réseau n'est plus sensible aux gradients qui s'effondrent, effet observable avec les fonctions sigmoïdes [Glorot et al., 2011].

4.9.2.2 Descente de gradient stochastique

La descente de gradient stochastique est une technique d'optimisation utilisée pour l'apprentissage des ANNs. C'est une procédure itérative durant laquelle la méthode d'apprentissage

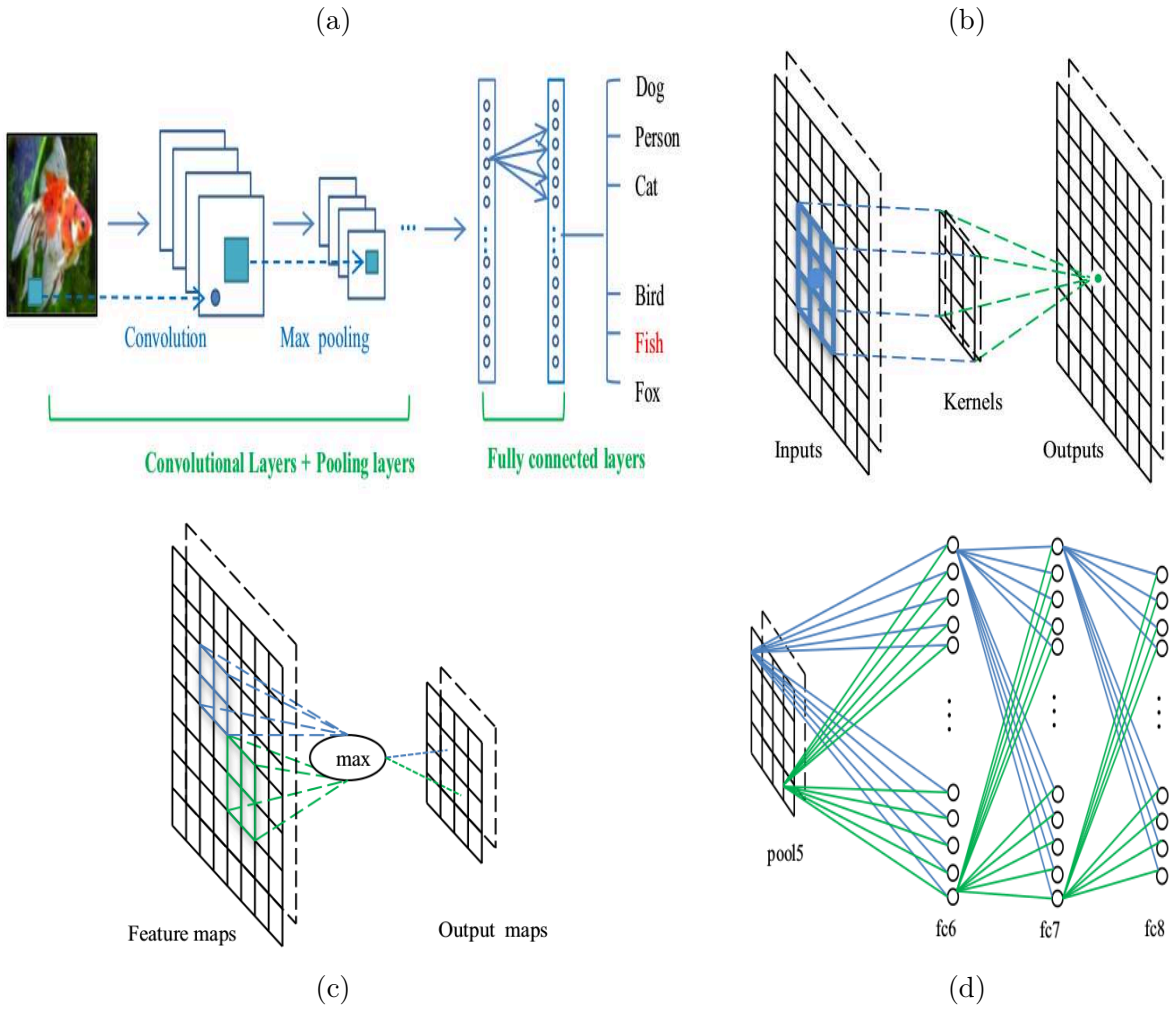


FIGURE 4.3 – Illustration des différents composants d'un CNN : (a) architecture générale d'un CNN, (b) fonctionnement d'une couche de convolution, (c) couche de sous- échantillonnage et (d) les couches *fully-connected* (FC).

ajuste ses paramètres pendant la phase d'apprentissage. A chaque mise à jour, des exemples de la base d'apprentissage sont présentés au réseau qui se charge de faire une prédiction et une estimation de l'erreur associée. Par la suite, le gradient moyen des erreurs est calculé pour indiquer la direction à choisir pour diminuer le risque associé. En ce qui concerne les méthodes de représentation faisant intervenir plusieurs couches cachées, la mise à jour des paramètres se fait grâce à une rétro-propagation du gradient de la couche de sortie vers celle d'entrée par le principe des dérivées en chaîne :

$$\theta = \theta - \varepsilon \nabla f(\theta) \ . \tag{4.4}$$

avec $f(\theta)$ la fonction apprise pour minimiser le risque empirique et $\varepsilon > 0$ le taux d'apprentissage. Il est à noter qu'une itération représente un parcours de toutes les données de la base d'apprentissage, ce qui implique donc un grand nombre de mise à jour des paramètres.

4.9.2.3 Initialisation des paramètres et pré-apprentissage

La descente de gradient stochastique à tendance à rendre les réseaux sensibles aux valeurs initiales des paramètres, de telle sorte que, si les valeurs s'écartent d'une solution convenable, l'apprentissage des ANNs devient très difficile [Hinton et Salakhutdinov, 2006]. De plus, une grande saturation des fonctions d'activation peut engendrer une mauvaise propagation des gradients [Glorot et Bengio, 2010]. A cela vient s'ajouter le fait que les paramètres appris peuvent être une fonction identité, ce qui rend l'apprentissage inutile [Glorot et Bengio, 2010]. La méthode ayant permis l'engouement récent pour l'apprentissage profond est basée sur un modèle qui sert de pré-apprentissage. Le modèle en question a été conçu pour reconstruire les données en entrée à partir d'une version qui est corrompue. Des auto-encodeurs débruitants ou des machines de Boltzmann restrictives peuvent être exploitées pour atteindre le même objectif. Dès que le pré-apprentissage de chaque couche cachée a été réalisé, le réseau de neurones peut procéder au raffinement des paramètres en tenant compte de l'objectif d'apprentissage définit. Le pré-apprentissage permet d'accélerer le processus d'apprentissage de manière à améliorer sa capacité de généralisation. Malgré leur popularité, ces méthodes de pré-apprentissage non-supervisées sont de plus en plus délaissées, en raison du fait qu'elles peuvent en réalité apprendre de mauvais paramètres, car il est difficile de contrôler leur capacité [LeCun et al., 2015]. Ceci justifie le recours aux méthodes d'apprentissage supervisée où il semble plus aisé de contrôler ce que l'algorithme apprend. Cette possibilité est offerte par des mécanismes comme les unités rectificatrices linéaires et la régularisation par *dropout*, couplé à la grande disponibilité de données labélisées – pour des applications mettant en scène des objets du quotidien – qui permettent d'apprendre des réseaux à plusieurs couches cachées sans passer par des pré-apprentissages.

D'autres approches plus simples qui n'ont pas recours au pré-apprentissage ont été testées avec succès ; c'est le cas de l'initialisation par une distribution Gaussienne [Krizhevsky et al., 2012] et [Simonyan et Zisserman, 2014]. Cependant, l'apprentissage de réseaux très profonds peut se montrer fastidieux lorsque le nombre de couches est supérieur à dix [Simonyan et Zisserman, 2014]. Une autre solution consiste à utiliser une distribution uniforme, c'est-à-dire une moyenne de zéro et un écart-type de un [LeCun et al., 2012]. [Glorot et Bengio, 2010] ont toutefois démontré que l'initialisation des poids de chaque couche du réseau par cette dernière approche peut aboutir à un gradient dont les valeurs s'écroulent pour un régime linéaire. Pour y rémédier, la solution est de normaliser l'initialisation des poids.

4.9.2.4 La norme des paramètres

On peut restreindre la capacité des ANNs en ajoutant la norme des paramètres à la fonction coût. Cela consiste soit en une régularisation de la norme *l1* ou la norme *l2*. La régularisation *l1* favorise la parcimonie au sein du réseau en mettant à zéro les paramètres non-discriminants [Bengio, 2012] et peut donc permettre une sélection des attributs.

4.9.2.5 L'arrêt prématuré

Le nombre de mise à jour des paramètres du réseau joue un rôle prépondérant sur les performances du modèle. En effet, un modèle qui n'est pas suffisament entraîné conduit à un sous-apprentissage, tandis qu'un modèle entraîné sur une trop longue durée aboutit au sur-

apprentissage. Il faut donc trouver un compromis afin d'optimiser la performance de généralisation du système. Une des solutions est l'arrêt prématuré, qui sert à fixer les autres hyperparamètres et à mesurer l'erreur du système sur un jeu de validation au cours de l'entraînement et à l'arrêter quand cette erreur augmente. Pour eviter de s'arrêter trop tôt, une solution a été proposée et consiste à spécifier un nombre minimum – nommé *patience* – de mises à jour des paramètres à observer lorsqu'une certaine valeur est obtenue sur la base de validation. Si un nouveau minimum est trouvé durant la *patience* choisi, alors sa longueur est étendue ; dans le cas contraire, l'entraînement s'interrompt et les paramètres reviennent à l'état du dernier minimum atteint [Bengio, 2012].

4.9.2.6 Les ensembles

En raison de l'optimisation non-convexe des ANNs, il existe de nombreux minima locaux. D'après [Bengio et Courville, 2016], un minimum local est un point dans l'espace des paramètres qui produit une erreur plus faible que celle de ses voisins ; il est donc difficile d'en sortir en faisant des petits pas de gradient. Pour ce qui est de l'apprentissage des représentations, les minima locaux permettent cependant d'obtenir des résultats intéressants. Le challenge se situe au niveau des points de selle qui sont plus préjudiciables pour l'optimisation des ANNs [Bengio et Courville, 2016]. La présence de plusieurs minima locaux constituent une réelle source de difficultés, car cela signifie que selon les valeurs initiales des paramètres, les résultats des modèles appris vont varier [Hansen et Salamon, 1990] ; autrement dit, ces modèles feront des erreurs durant la classification mais pas forcément les mêmes erreurs. C'est là qu'intervient la notion des *ensembles* dont l'idée est de tirer partie de cette propriété afin de produire une prédiction collective, sous prétexte qu'un ensemble à tendance à produire une meilleure prédiction qu'un modèle pris individuellement [Hansen et Salamon, 1990]. Il existe plusieurs façons pour créer des ensembles ; chaque modèle peut être entrainé avec différentes initialisations ou encore des mini-lots présentés dans un ordre différent [Hansen et Salamon, 1990]. Pour ce qui est de la prédiction collective, elle peut provenir des votes majoritaires – décision ayant recueilli au moins la moitié des voix – ou des votes pluraux, *c-à-d.* décision ayant recueilli le plus de voix. La technique des ensembles est utilisée par bien de chercheurs pour atteindre de bonnes performances de généralisation ; c'est notamment le cas de [Russakovsky et al., 2015].

4.9.2.7 *Dropout*

Proposée par [Hinton et al., 2012], le *dropout* est une procédure de corruption qui consiste à réaliser une transformation stochastique afin d'imposer que certaines dimensions de la sortie d'une couche soient égales à zéro [Srivastava et al., 2014]. De ce point de vue, cette transformation est similaire au procédé de corruption d'un auto-encodeur débruitant ; la différence entre les deux outils se trouve dans l'utilisation de la corruption. Au cours de l'entraînement, la régularisation par *dropout* ne conserve qu'un certain nombre d'unités selon une certaine proportion, ce qui engendre un ANN de plus petite taille. Seules les petites unités sont mises à jour, *c-à-d.* que pour chaque mise à jour des paramètres, un réseau différent est échantillonné. Pendant la phase de test, ce grand ensemble d'ANNs différents de petite taille est approximé en faisant la multiplication entre les paramètres des unités cachées par leur proportion p. Un des avantages

du *dropout* est qu'elle empêche la co-adaptation entre les unités cachées dans le but de mettre en avant l'apprentissage de descripteurs discriminants [Srivastava et al., 2014]. En effet, il arrive qu'une unité corrige les erreurs effectuées par une autre unité de son voisinage, ce qui entraîne des effets de co-adaptation complexe entre les unités cachées. La co-adaptation produit de graves conséquences sur l'apprentissage d'un ANN car il entraîne un sur-apprentissage. L'autre atout du *dropout* est qu'il encourage l'apprentissage de représentations parcimonieuses qui possède plusieurs avantages permettant de faciliter l'apprentissage d'un ANN [Srivastava et al., 2014]. Le *dropout* a été étudié de près par [Warde-Farley et al., 2013] qui ont analysé sa potentialité et suggéré que c'est une méthode d'apprentissage d'ensemble extrêmement efficace.

Néanmoins, la régularisation par *dropout* nécessite un temps d'apprentissage deux à trois fois supérieur à une procédure d'apprentissage sans *dropout* [Srivastava et al., 2014]. Ceci s'explique par le fait que chaque mise à jour des paramètres est calculée à partir d'un nouveau réseau échantillonné aléatoirement. Une explication plus exhaustive sur le *dropout* est proposée par [Baldi et Sadowski, 2013]. Il existe une autre variante du *dropout* connue sous le nom de *DropConnect* et qui a été largement étudiée par [Wan et al., 2013].

4.9.2.8 Momentum

Le momentum est une méthode destinée à accélerer la descente de gradient et qui consiste à calculer une moyenne mobile des gradients à travers le temps. Dans le cas des ANNs, il sert à accumuler un vecteur de vitesse v dans les directions permettant de réduire l'erreur estimée par la fonction objective [Sutskever et al., 2013]. La règle de mise à jour des paramètres est alors la suivante :

$$v_{t+1} = \mu v_t - \varepsilon \nabla f(\theta_t) \ . \tag{4.5}$$

$$\theta_{t+1} = \theta_t + v_{t+1} \ . \tag{4.6}$$

avec $\mu \in [0, 1]$ le coefficient du momentum et ε le taux d'apprentissage. Il existe une variante du momentum classique qui se nomme le momentum de Nesterov et dont le but est de réaliser une mise à jour des paramètres comme suit :

$$v_{t+1} = \mu v_t - \varepsilon \nabla f(\theta_t + \mu v_t) \ . \tag{4.7}$$

$$\theta_{t+1} = \theta_t + v_{t+1} \ . \tag{4.8}$$

La différence entre les deux méthodes du momentum se situe dans le calcul du gradient. Le momentum de Nesterov réalise une première mise à jour avant le calcul du gradient, rendant ainsi le vecteur vitesse plus stable evitant les grandes oscillations souvent constatées avec le momentum classique [Sutskever et al., 2013].

4.9.2.9 La sélection des hyper-paramètres

En apprentissage automatique et particulièrement pour l'apprentissage profond, le choix des hyper-paramètres est une tâche à la fois chronophage et fastidieuse. Le nombre d'hyper-paramètres est très important et faire le choix des valeurs permettant d'obtenir des performances

élevées n'est pas chose aisée, car cela se fait sur un ensemble de données de validation qui peut être très important en fonction des problèmes. Bien vrai, il existe certaines valeurs d'hyper-paramètres obtenues de façon empirique ; toutefois, elles ne sont pas écrites à l'encre de chine car à chaque application et jeu de données correspond un ensemble d'hyper-paramètres différent [Bengio, 2012]. La façon de procéder la plus simple est de faire une recherche manuelle, ce qui fait que les méthodes d'apprentissage automatique sont fortement liées à l'interaction avec l'utilisateur qui doit donner la valeur de chacun des hyper-paramètres. Une autre heuristique consiste à fournir une grille dont la dimension est égale au nombre d'hyper-paramètres. Le but étant d'entraîner un modèle pour toutes les combinaisons possibles de la grille afin de choisir un ensemble d'hyper- paramètres produisant la plus faible erreur possible ; cette approche est néanmoins très vague. Une solution différente est d'échantillonner les valeurs des hyper-paramètres selon une loi uniforme [Bergstra et Bengio, 2012]. Les valeurs des hyper-paramètres sont alors obtenues de façon aléatoire. Cette solution peut être étendue pour tirer profit de toutes les décisions passées ; cela passe par l'optimisation Bayésienne [Snoek et al., 2012] qui va décider de la zone à explorer à l'itération suivante et ce en se fiant aux évaluations passées.

4.9.2.10 L'augmentation artificielle de données

L'augmentation artificielle de données est une heuristique souvent utilisée pour éviter le sur-apprentissage des données, qui possède également l'avantage de ne pas introduire un coût additionnel lié à l'éventuel ajout de données labélisées. De façon pratique, cela signifie que durant l'entraînement, des transformations aléatoires seront appliquées à chacun des exemples présentés au CNN. Toutefois, vue la nature aléatoire du processus, il n'est pas exclue qu'aucune transformation n'ait été faite, ou que l'exemple ne soit que peu ou fortement modifié. Les transformations affines à savoir : les translations, les mises l'échelle, les réflexions et les rotations sont les techniques les plus utilisées. D'autres transformations plus avancées peuvent être exploitées ; c'est le cas du AlexNet [Krizhevsky et al., 2012] qui a employé deux formes distinctes d'augmentation de données : la première consiste à générer des translations et des réflexions horizontales, tandis que la seconde a servi à altérer les intensités des canaux couleurs RGB en ajoutant du bruit aux composantes principales des images de la base d'entraînement. [Howard, 2013] s'est dans un premier temps servi du AlexNet comme modèle de base, auquel il a ajouté des transformations supplémentaires qui ont améliorées l'invariance à la translation et à la couleur par des transformations dans l'espace des couleurs. Cette méthode d'augmentation de données a été exploitée par d'autres chercheurs [Szegedy et al., 2015], [He et al., 2015]. Quant à [Dosovitskiy et al., 2013], il a proposé une approche d'apprentissage non-supervisée des descripteurs basée sur l'augmentation de données : premièrement il procède à une extraction aléatoire de plusieurs patchs de l'image qu'il déclare comme classe de substitution ; puis, il étend chacune de ces classes en réalisant des transformations affines de couleur et de contraste. Ensuite, il a appris un CNN afin de distinguer les diverses classes de substitution ; les descripteurs appris par ce réseau ont montré de bonnes performances sur une variété de tâches de classification. Plus tard, [Wu et al., 2015] a utilisé des techniques de distorsions, d'aberrations chromatiques et de vignettage pour produire plus d'exemple d'apprentissage.

4.10 Apprentissage de représentation profonde par *Deep learning*

En raison de la petite taille des jeux de données qui existent en mammographie, nous avons jugé plus judicieux – pour faire de l'apprentissage profond – de nous servir d'un CNN pré-entrainé qui permettra d'extraire de façon non-supervisée des descripteurs sur toute la base d'entraînement. Plusieurs modèles de CNN, parmi lesquels le AlexNet [Krizhevsky et al., 2012]. Ce choix se justifie par le fait que c'est le premier modèle moderne de CNN qui est à l'origine du regain d'intérêt de la communauté scientifique pour les CNNs ; de plus, il a gagné le concours ILSVRC2012 [Russakovsky et al., 2015] après un entraînement poussé de six jours sur la base ImageNet, pour un taux d'erreur de 15%.

4.10.1 L'apprentissage et composition du AlexNet

La composition et le fonctionnement de ce modèle se résument comme suit : avant tout, et dans le but de réduire le sur-apprentissage, les auteurs ont procédé à un pré-traitement de l'image qui consiste en une translation et une réflexion horizontale de l'image, par extraction aléatoire des patchs de taille 224×224 sur les images de départ ; lesquels patchs ont ensuite servi de base d'apprentissage. Puis, afin d'assurer l'invariance à l'illumination et à la couleur, ils ont altéré les intensités des canaux couleurs RGB en ajoutant du bruit aux composantes principales des images de la base d'apprentissage. En ce qui concerne les couches, les auteurs ont utilisé huit couches dont cinq de convolution et trois couches FC. La première et la seconde couche de convolution sont suivies des couches de normalisation et de *max-pooling*, tandis que les trois dernières couches de convolution sont reliées les unes aux autres sans aucune intervention de normalisation et de *pooling*. De plus, il y a des astuces que les auteurs ont employé et qui leur ont permis de concevoir un modèle de CNN plus efficace que les précédents. Premièrement la fonction d'activation utilisée est une ReLU, qui comme nous l'avons vu un peu plus tôt, contribue à réduire de façon significative le temps d'entraînement du CNN en comparaison de *tanh (x)*. Deuxièmement, ils ont adopté une normalisation locale, qui facilite la généralisation. Enfin, ils ont utilisé une technique de sous-échantillonnage chevauchante, qui améliore l'invariance des matrices de descripteurs aux diverses transformations. Traditionnellement, les neurones voisins sous-échantillonnés ne se chevauchaient pas ; en adoptant un sous-échantillonnage chevauchant, ils ont pu réduire le taux d'erreur lié à l'apprentissage. A la suite des couches de convolution, il y a deux couches FC dénommées FC1 et FC2 avec 4096 neurones. La dernière couche de sortie, qui est reliée à la couche FC2, contient 1000 neurones qui est une prédiction pour chacune des 1000 classes de la base ImageNet. En tout, il y a un peu plus de 60 millions d'hyper-paramètres dont dépend les performances de l'AlexNet.

4.11 La solution proposée

Pour l'application qui nous concerne, à savoir la discrimination d'une part entre les masses bénignes et malignes et d'autre part entre les masses du BI-RADS IV et celles du BI-RADS V, nous proposons une approche à trois niveaux de décision permettant aux radiologues de confirmer ou d'infirmer son diagnostic à propos de la nature de la masse.

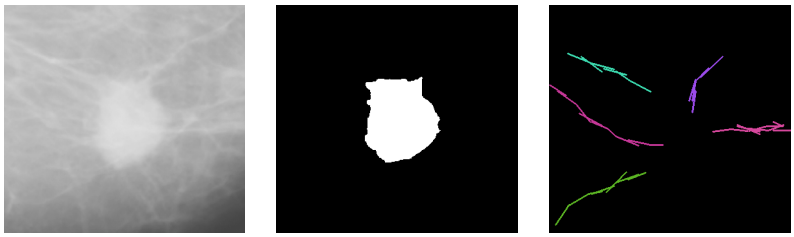
Dans un premier temps le AlexNet est appliqué sur la base des ROIs, afin d'extraire de

façon non-supervisée les descripteurs de texture et de densité de haut niveau pouvant permettre de prendre une décision sur la densité du centre de la masse à savoir si elle est graisseuse ou dense. Bien que la densité du centre de la masse compte très peu dans la distinction entre une masse bénigne et une masse maligne, il faut souligner que si une masse est embarquée dans un parenchyme mammaire d'une densité importante l'analyse de sa morphologie sera fortement compromise. Par contre, c'est un aspect qui sera déterminant dans la discrimination entre les masses spiculées et les distorsions architecturales.

Ensuite, on exécute le AlexNet sur les images obtenues après segmentation par le log-PRF. Ici, le rôle du AlexNet est d'extraire différents descripteurs de forme et de contour relatifs aux différentes masses contenues dans notre base d'apprentissage.

Enfin, pour ce qui est des spicules, nous proposons de laisser les radiologues juger de la pertinence de leur prise en compte dans leur décision. En effet, le caractère aléatoire des formes, de la densité des masses et des types (longueur et nombre) des spicules dans le cas du cancer du sein, rend irréaliste la proposition de toute règle qui serait gravée dans le marbre. Néanmoins, nous avons la certitude à travers le BI-RADS que les spicules permettent de classifier presque parfaitement les masses bénignes de celles malignes ou des distorsions architecturales, car aucune masse bénigne ne pourrait produire des spicules. De plus, la quantité et la longueur des spicules pourrait permettre de discriminer entre les masses spiculées et les distorsions architecturales. Cependant, seule l'expertise et l'expérience des radiologues pourraient aider à aboutir à ce degré de subtilité. L'aide que nous proposons à travers la segmentation des spicules étant de tirer leur attention sur les vraies structures d'intérêt en éliminant celles qui pourraient prêter à confusion.

La figure 4.4 ci-dessous rappelle par souci de clarté les trois types d'information prise en compte dans l'aide à la prise de décision.



(a) (b) (c)

FIGURE 4.4 – Illustrations des images intervenant dans le processus d'aide à la décision : (a) ROI, (b) image binaire obtenue après segmentation et (c) les spicules.

Une fois chacune des catégories de descripteurs extraits, on les utilise séparément dans un classifieur de type SVM afin d'apprendre deux modèles différents qui aideront à prendre une décision soit sur leur forme, soit sur leur densité. Ces deux modèles sont stockés une fois pour toutes et cette première étape constitue la partie *off-line* de notre système d'indexation et de reconnaissance du cancer du sein. La figure 4.5 ci-dessous donne un aperçu des étapes d'apprentissage de nos deux modèles.

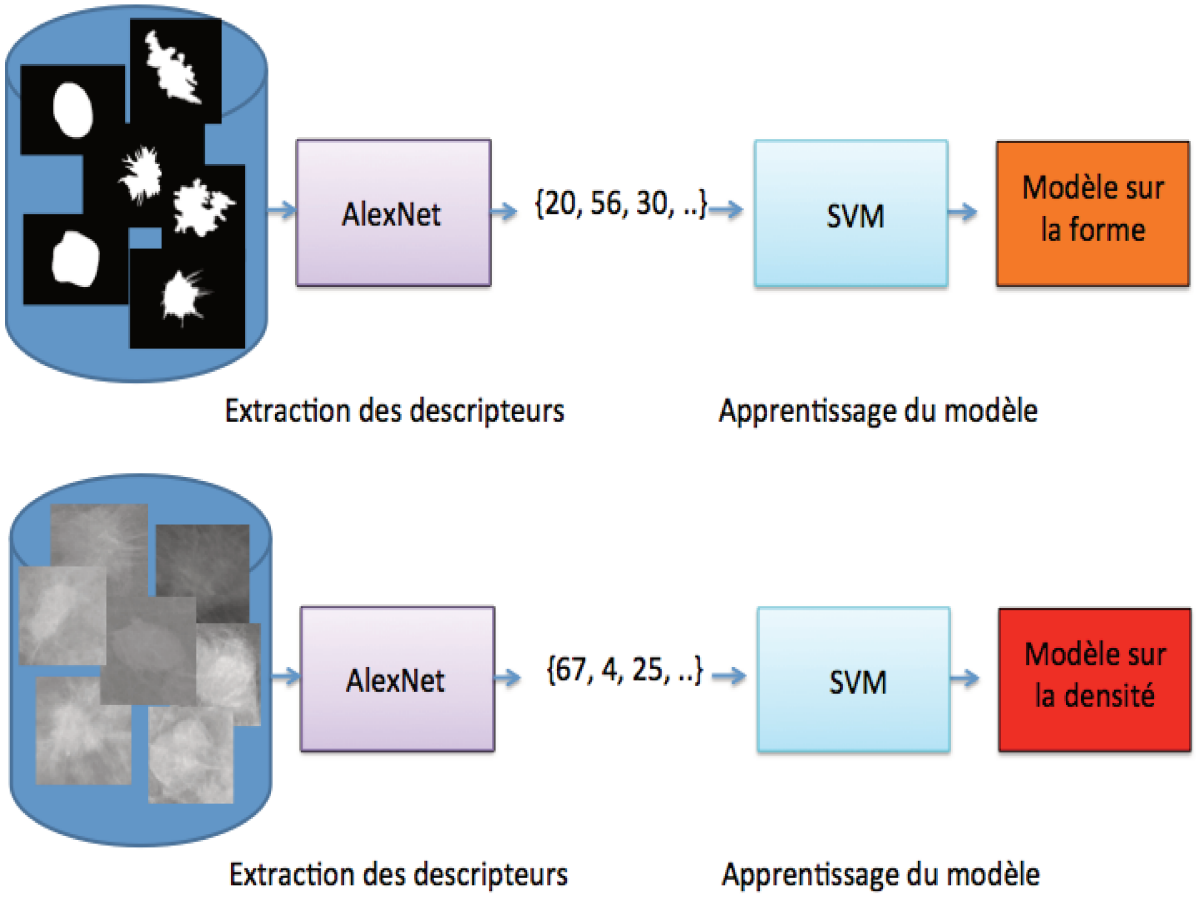


FIGURE 4.5 – Illustration du déroulement de la phase *off-line*.

4.11.1 Implémentation numérique

Afin de construire notre classifieur de morphologie des masses mammaires, nous avons utilisé un jeu de données – le même que celui utilisé par [Kachouri et al., 2012] dans le but de réaliser une comparaison objective – que nous avons divisé en trois partitions différentes qui sont : $\mathcal{D}_{train}$, $\mathcal{D}_{valid}$, $\mathcal{D}_{test}$. $\mathcal{D}_{train}$ représente la base d'images sur laquelle le modèle est entrainé afin d'ajuster ses paramètres. $\mathcal{D}_{valid}$ est le jeu de données sur lequel se fait la sélection des hyper-paramètres et l'évaluation de la performance du modèle. Quant à $\mathcal{D}_{valid}$, il représente le jeu de données servant à évaluer la performance de généralisation globale. Pour cela, des images issues de la DDSM et ne contenant que les masses sont sélectionnées afin d'être utilisées dans la partie expérimentale. Le jeu de données choisi est composé de 248 images dont 114 masses malignes et 128 masses bénignes. Cette base est répartie comme suit : $\mathcal{D}_{train}$ (57 masses malignes / 57 masses bénignes), $\mathcal{D}_{valid}$ (58 masses malignes / 56 masses bénignes) et enfin $\mathcal{D}_{test}$ (6 masses malignes / 14 masses bénignes). En ce qui concerne le jeu de données pour la classification de la densité, il provient de la MIAS en raison de la qualité visuelle des images. Il est composé de 234 images réparties comme suit : $\mathcal{D}_{train}$ (54 images denses / 54 images graisseuses), $\mathcal{D}_{valid}$ (55 images denses / 51 images

graisseuses) et $\mathcal{D}_{test}$ (10 images denses / 10 images graisseuses). Il est important d'ajouter que nous aurions pu procéder à une classification de la densité en quatre groupes telles que proposées par l'ACR (ACR I à ACR IV), mais nous avons préféré une classification dense/graisseuse pour tenir compte de la description du BI-RADS à propos du centre des masses spiculées (qui sont censées être dense *c-à-d.* BI-RADS V) et de celle des distorsions architecturales (qui sont censées être graisseuses, *c-à-d.* BI-RADS IV).

Etant donné l'architecture de notre CNN (5 couches de convolution et 3 couches FC) et compte tenu du nombre de paramètres (plus de 60 millions) à apprendre, nous avons utilisé une carte graphique de NVIDIA GEFORCE GTX 1070 ayant les caractéristiques suivantes : 1920 *cores*, une fréquence de 1506 MHz, une mémoire de 8GB et une vitesse de mémoire de 8 Gbps. Les librairies MatConvNet et/ou Caffe ont été utilisées pour les calculs. Pour finir, nous parcourons 100 fois chacune des images de notre base par itération aussi bien pour extraire les descripteurs de $\mathcal{D}_{train}$ que ceux de $\mathcal{D}_{valid}$.

4.11.2 Résultats

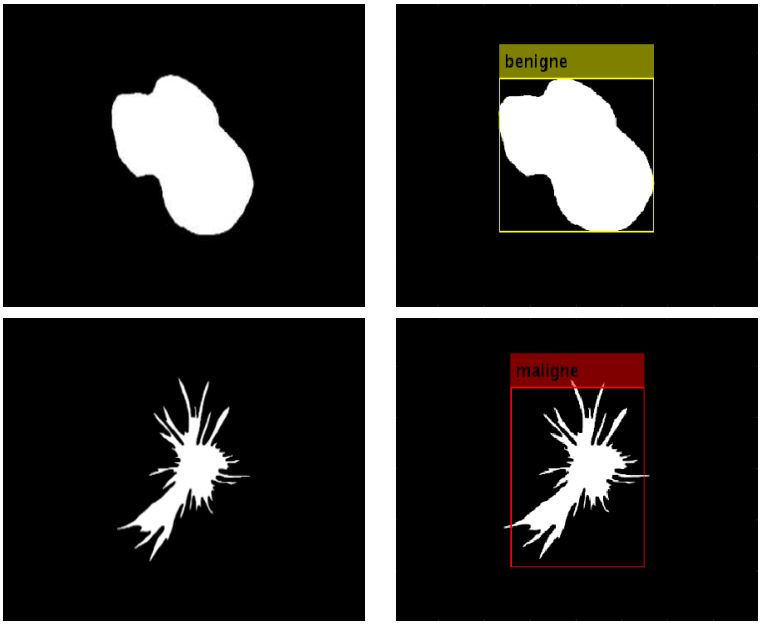
Comme évoqué durant la présentation du AlexNet, l'architecture d'un CNN se compose de plusieurs couches dont les FC qui contiennent plusieurs neurones destinés à fournir les descripteurs extraits. Dans notre cas, nous avons utilisé la seconde couche FC qu'on va appeler FC7, composée de 4096 neurones comme couche de descripteurs. Par la suite, ces descripteurs ont été utilisés dans un classifieur de type SVM pour apprendre un modèle d'aide à la décision. Ce schéma a été utilisé pour apprendre un premier modèle permettant de classifier les masses comme bénigne ou maligne en se focalisant sur leur morphologie et un second modèle qui produit une classification dense ou graisseuse du centre de la masse.

Une illustration du résultat obtenu pour la généralisation après extraction et classification des descripteurs de morphologie est présentée à la figure 4.6 ci-dessous. Quant à la figure 4.7, elle montre les résultats obtenus pour la classification de la densité du centre de la masse.

4.11.3 Discussion

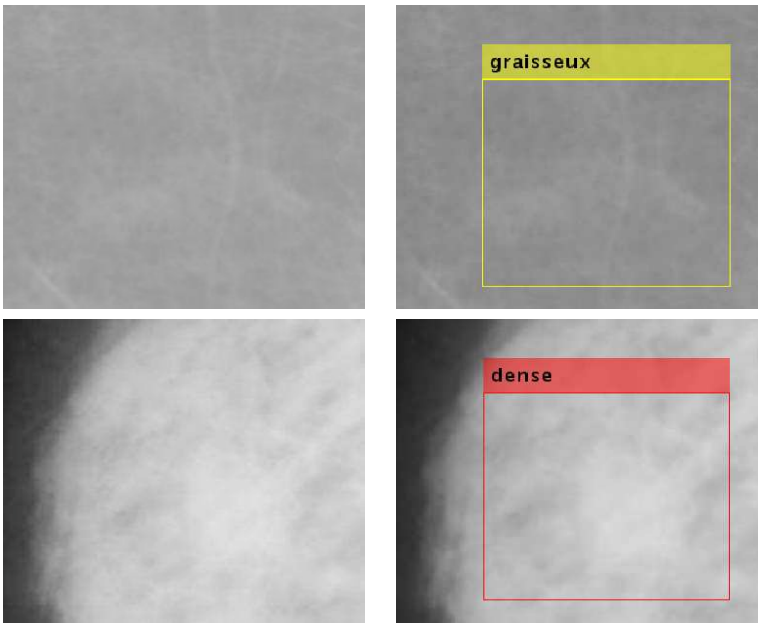
Afin de vérifier la pertinence du modèle appris sur notre jeu de données à la fois pour la qualité de la classification des masses selon la forme (bénigne ou maligne) ou la densité (dense ou graisseuse) nous avons décidé de faire une étude comparative avec les méthodes les plus efficaces de l'état-de-l'art. Pour la forme, les attributs utilisés pour la comparaison sont ceux proposés par [Kachouri et al., 2012] à savoir la SEP, la PS et la SMD, d'une part en raison du fait que l'auteur a prouvé que les performances de ces descripteurs de forme et de contour surpassent à la fois les descripteurs standards et ceux spécifiques et d'autre part parce que nous disposons exactement du même jeu de données – *c-à-d.* les mêmes régions et les mêmes images – que celui utilisé par [Kachouri et al., 2012] pour évaluer les peformances de ses descripteurs. Dans l'objectif de comparer rigoureusement les performances des descripteurs extraits par le AlexNet avec la SEP, la PS et la SMD, l'évaluation sera faite dans la suite à l'aide des métriques que sont la matrice de confusion et l'AUC, étant donné que ce dernier a également été utilisé par [Kachouri et al., 2012].

Pour ce qui est de la densité/texture, nous avons fait la comparaison avec les travaux



(a) (b)

FIGURE 4.6 – Classification des masses en mammographie selon leur morphologie : (a) ROIs bénigne et maligne, (b) diagnostic fourni par notre système.



(a) (b)

FIGURE 4.7 – Classification des masses en mammographie selon la densité de leur centre : (a) ROIs bénigne et maligne, (b) diagnostic fourni par notre système.

de [Bovis et Singh, 2002], [Petroudi et al., 2003] et [Bosch et al., 2006]. Notre choix s'est porté sur ces travaux car parmi les approches offrant les meilleures performances pour la classification des masses en fonction de leur densité, ce sont ceux qui ont pris en compte une classification des masses dans un contexte graisseux et dense. Afin de nous comparer de façon objective à ces différentes contributions, nous avons utilisé la matrice de confusion étant donné que c'est la métrique utilisée dans les travaux ci-dessus cités.

Aussi bien l'évaluation de la performance moyenne – en terme de classification – de notre système de reconnaissance à travers $\mathcal{D}_{valid}$, que celle du processus de généralisation ($\mathcal{D}_{test}$) se font à l'aide du protocole de test suivant dont une illustration est fournie à la figure 4.8 ci-dessous. Pour chacune des images contenues soit dans $\mathcal{D}_{valid}$ ou $\mathcal{D}_{test}$, AlexNet extrait un ensemble de 4096 descripteurs à l'aide de la FC7. Une fois les descripteurs obtenus, on utilise l'un des deux modèles appris plus haut en *off-line* afin de prédire la nature de la masse soit bénigne ou maligne ou la densité du centre de cette dernière soit graisseuse ou dense. La précision moyenne du système est enfin calculée à l'aide des métriques évoquées plus haut et détaillées dans la suite.

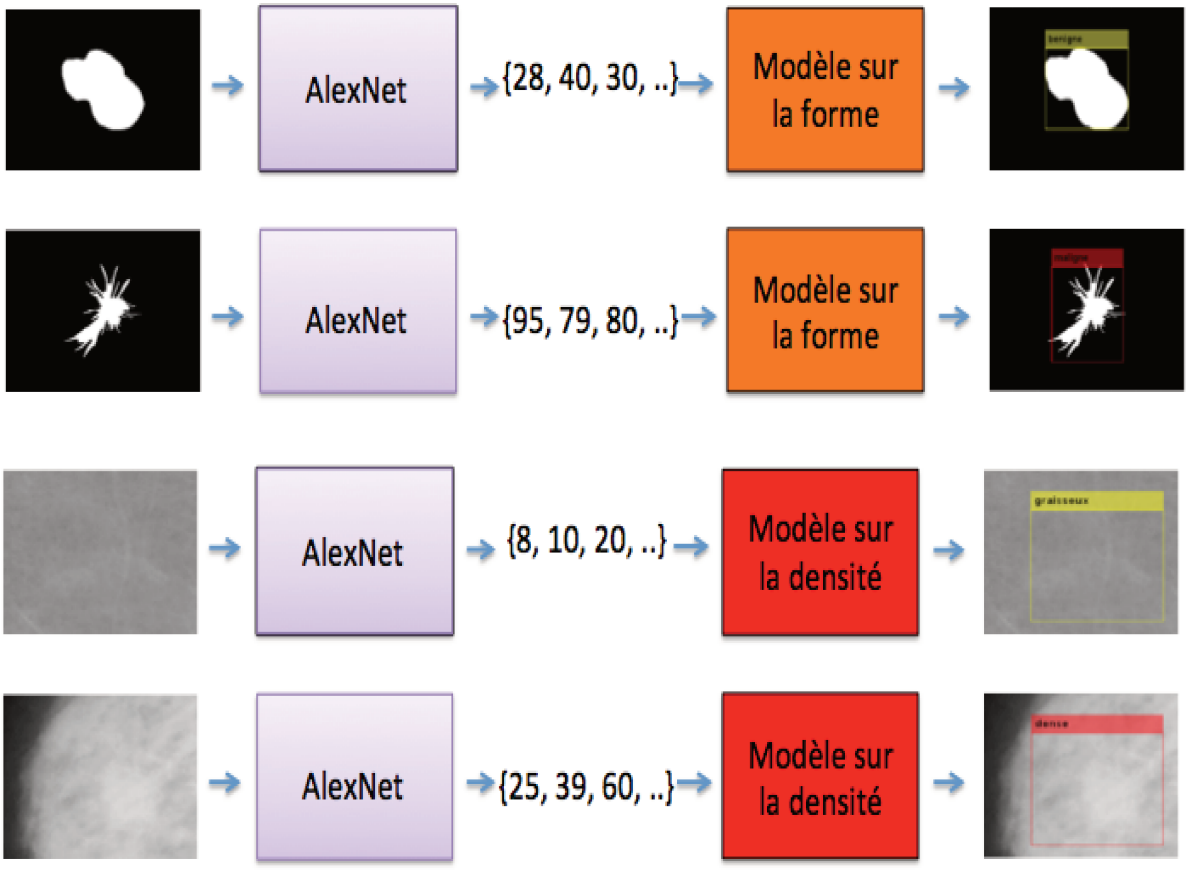


FIGURE 4.8 – Illustration du déroulement de la phase *on-line*.

De façon générale, notre système de DAOx (la partie *on-line*) fonctionne comme suit : *(1)*

les descripteurs sont extraits et classifiés pour une ROI requête – fournie par l'utilisateur *p. ex.* le radiologue – et une décision est prise à propos de la nature dense ou graisseuse de l'image ou du centre de la masse (comme décrit dans le protocole de test plus haut) ; puis, *(2)* pour la même image, notre système procède dans un premier temps au pré-traitement de l'image à l'aide du K-SVD-ACP afin de réduire le bruit. Ensuite, le contraste de l'image est amélioré et la segmentation de la masse contenue dans l'image est réalisée par l'intermédiaire du log-PRF. Ce n'est qu'à la suite de la segmentation que les différents descripteurs sont extraits par le CNN. Finalement, la prédiction est faite sur la nature maligne ou bénigne de la masse en tenant compte de l'aspect que présentent sa forme et son contour. *(3)* pour les distorsions architecturales, les spicules sont également segmentées. Ces trois éléments (forme/contour, graisseux/dense et spicules) permettront ensuite au radiologue de prendre la décision adéquate concernant la masse.

4.11.3.1 Morphologie

Une fois le modèle appris, la base de validation a été utilisée pour évaluer la qualité du modèle construit. Conformément à ce qui a été annoncé plus haut, nous avons calculé la matrice de confusion, la courbe ROC, illustrées respectivement par les figures 4.9 et 4.10, enfin nous avons déduit la valeur de l'AUC dont les détails figurent dans le tableau 4.1.

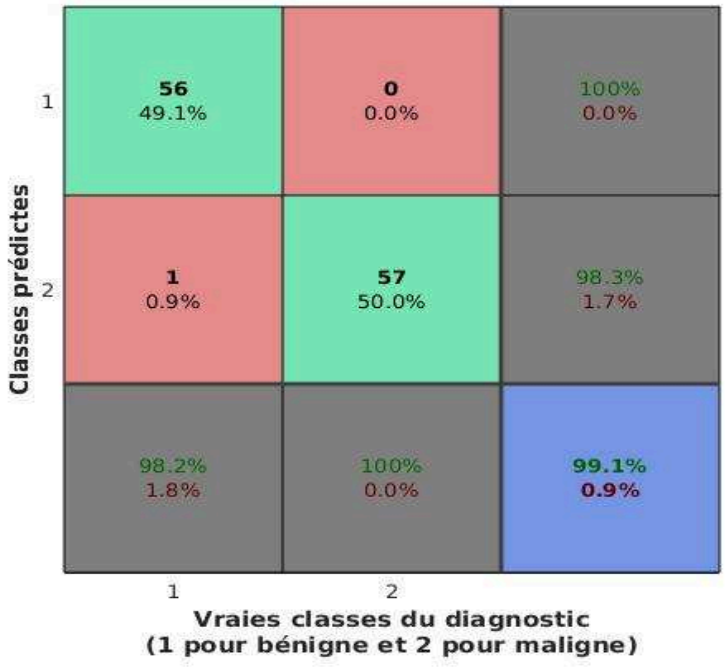


FIGURE 4.9 – Matrice de confusion de la classification des masses dans les classes bénigne et maligne.

Dans la matrice de confusion, on observe que les 56 masses bénignes, ont été parfaitement bien classifiées ce qui équivaut à une sensibilité de 100% ; par contre sur les 58 masses malignes,

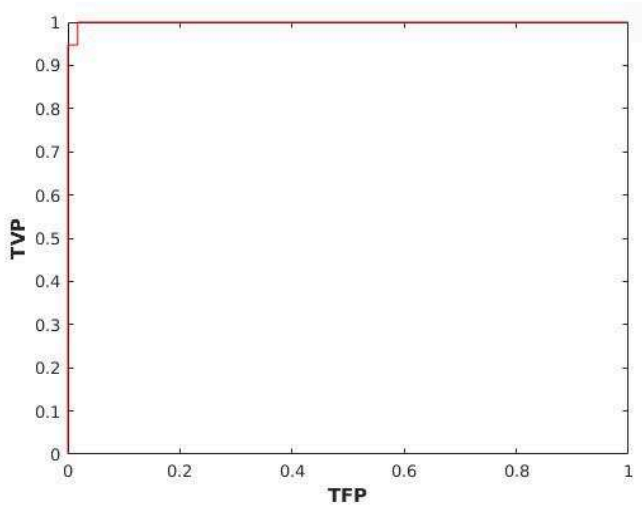


FIGURE 4.10 – Courbe ROC des descripteurs extraits par AlexNet.

TABLE 4.1 – Performance des différents descripteurs de forme en terme de l'aire sous la courbe ROC.

Descripteurs	Valeur de A_z
Descripteurs CNN	0.99
SMD	0.97
PS	0.93
SEP	0.92
NRL	0.92
NCL	0.87

57 ont été correctement classifiées tandis qu'une image a été rangée comme bénigne, ce qui donne une sensibilité de 98.3%. La courbe ROC qui peut être vue comme une représentation graphique de la matrice de confusion pour un descripteur donné, vient confirmer la répartition fournie par cette matrice. En effet, cette courbe avoisine très vite le coin supérieur gauche du cadre unitaire et prend une allure parfaite ; cette tendance est confirmée par la valeur de l'AUC à savoir A_z=0.99. Ce qui signifie que notre résultat est très informatif voire parfait. Pour ce qui est des masses utilisées en généralisation, elles ont toutes été bien classifiées.

Il convient donc de remarquer à la lumière de ces résultats que les performances obtenues pour la classification des masses dans les catégories bénigne/maligne sur la base de leur morphologie sont bien meilleures avec l'approche que nous avons proposé. En effet, dans le tableau 4.1, on observe les différentes valeurs de l'AUC prises par divers descripteurs de morphologie extraits sur le même jeu de données, parmi lesquels les descripteurs proposés par [Kachouri et al., 2012] et dont les valeurs de A_z obtenues sont : 0.92, 0.93 et 0.97 respectivement pour la SEP, la PS et la SMD. Dans le même tableau, on note les valeurs de A_z obtenues à l'aide des descripteurs spécifiques de forme comme ceux liés à la NRL [Kilday et al., 1993] et à la NCL [Kilday et al., 1993],

tous présentés dans l'annexe (A). Les performances de ces descripteurs sont inférieurs à celles de [Kachouri et al., 2012] et aux nôtres. Ceci démontre qu'une bonne qualité de segmentation combinée avec le CNN permet d'aboutir à des taux assez impressionants de bonne classification, ce qui nous permet de dire que notre modèle est parfaitement fiable pour une aide à la décision sur la nature bénigne ou maligne de la masse en se fondant sur sa morphologie.

4.11.3.2 Densité

Pour la densité, les résultats obtenus après la classification des images contenues dans notre base de validation par le modèle construit pour l'aide à la décision sur la densité du centre de la masse, ont permis de générer la matrice de confusion qui se trouve à la figure 4.11.

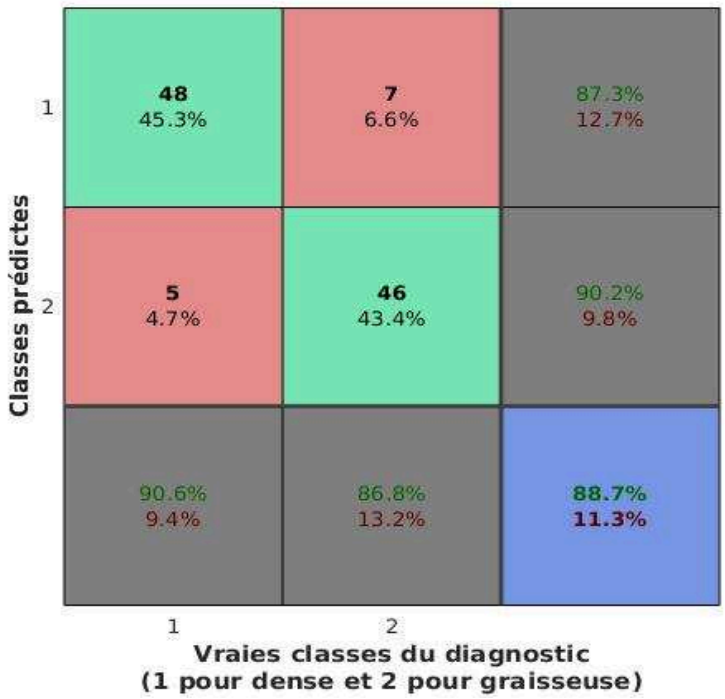


FIGURE 4.11 – Matrice de confusion de la classification des masses dans les classes graisseuse et dense.

On observe ainsi que les descripteurs de texture et de densité extraits de manière non-supervisée par le AlexNet et classifiés par le SVM nous permettent d'aboutir à une sensibilité de 90.2% et 87.3% respectivement pour la classe graisseuse et pour la classe dense et ce pour un taux global de bonne classification de 88.7%. Ce qui signifie que sur les images 51 images de densité graisseuse, 46 ont été bien classifiées, de même que 48 des 55 images de forte densité ayant servi à la validation.

Quant à [Bovis et Singh, 2002], rappelons qu'après avoir extrait différents descripteurs de texture sur les matrices de : GLCM, la transformée de Fourier, les masques de Law et la transformée en ondelettes, ils ont procédé à une réduction de la taille de leur vecteur de descripteurs

à l'aide de l'ACP. Après utilisation de plusieurs classifieurs ANN, ils ont obtenu une sensibilité globale de 77.3% pour les classes graisseuse et dense avec le classifieur offrant le meilleur taux de bonne classification. Ce score est bien inférieur au 88.7% que nous obtenons en utilisant un seul classifieur SVM. Toutefois, quand ils procèdent à une combinaison des classifieurs offrant les meilleures performances par l'intermédiaire de la règle de la somme de [Kittler et al., 1998], ils aboutissent à un taux de 96.7%. Ce qui est bien meilleur que le score que nous avons obtenu. Rappelons que [Petroudi et al., 2003] ont pour leur part développé une stratégie qui utilise les *textons* comme descripteurs et la distribution χ^2. Ils ont obtenu un taux de 91% et 94% respectivement pour les classes graisseuse et dense. En ce qui concerne [Bosch et al., 2006], ils ont d'abord construit un BoW en se servant des *textons* ; ensuite, ils ont exploité le pSLA afin de retrouver la distribution de chaque tissue dans l'image. Ces distributions ont enfin été utilisées pour la classification par SVM. Les valeurs de sensibilité obtenues sont de 99.5% et 98.24% respectivement pour les catégories graisseuse et dense.

Ces résultats montrent qu'en matière de discrimination par la densité/texture, l'utilisation d'une stratégie bien pensée fondée sur les descripteurs extraits de manière supervisée en l'occurence les *textons* – les stratégies de [Petroudi et al., 2003] et [Bosch et al., 2006] – permet d'obtenir des résultats bien meilleurs aux nôtres. Ce qui signifie qu'il serait plus indiqué de réfléchir à une approche qui combine de façon intelligente les descripteurs de textons à ceux extraits par AlexNet afin d'améliorer les scores de ces deux classes. Néanmoins, les résultats affichés sont intéressants et permettent d'être optimistes quant au potentiel des méthodes basées sur le *Deep learning*.

A toute fin utile, précisons que contrairement à ces auteurs – qui se servent des descripteurs extraits sur toute l'image pour réaliser la classification –, nous utilisons des ROIs centrés sur les masses afin d'évaluer la nature dense ou graisseuse du centre de la masse et non celle du parenchyme mammaire. Cette situation est bien évidemment plus compliquée que celles traitées dans l'état de l'art ; ce qui explique sans doute que les peformances obtenues par ces auteurs soient supérieures aux nôtres.

4.12 Les défis et les tendances du *Deep Learning*

Malgré les performances prometteuses obtenues à l'aide de l'apprentissage profond, une analyse pointue de l'état-de-l'art permet d'identifier plusieurs défis importants auxquels l'apprentissage profond est confronté, de même que les futures tendances en la matière. Une description détaillée de tout ceci est faite dans la suite.

4.12.1 Compréhension des principes théoriques

En dépit des résultats très intéressants obtenus dans l'exécution de certaines tâches du domaine de la vision par ordinateur grâce aux approches issues de l'apprentissage profond, les aspects théoriques sous-jacents ne sont pas encore bien compris et on n'a toujours pas une compréhension très claire des architectures ou réseaux qui offrent de meilleures performances que les autres. Il est donc difficile de dire quelle structure, combien de couches voire le nombre de neurones dont on doit disposer sur chaque couche afin d'obtenir de bons résultats pour une tâche bien précise. A cela s'ajoute, le manque de connaissance sur le choix des valeurs critiques

à savoir le taux d' apprentissage et la fiabilité de l'optimisation, etc ··· De ce fait, la conception de l'architecture a toujours été faite de façon *ad hoc*.

[Chu et Krzyżak, 2014] ont certes proposé une approche théorique pour déterminer le nombre optimal de matrices de descripteurs, mais cette technique n'est valable que pour un nombre très faible de neurones. Pour mieux comprendre le comportement des architectures des CNNs, [Zeiler et Fergus, 2014] ont développé une technique de visualisation qui donne un aperçu des couches de descripteurs intermédiaires. En révélant les descripteurs présents dans des modèles interprétables, ils ont apporté de nouvelles possibilités pour une meilleure conception des architectures de CNN. [Yu et al., 2014] ont également proposé une technique de visualisation similaire à celle de [Zeiler et Fergus, 2014]. Mise à part la visualisation des descripteurs, le réseau proposé par [Girshick et al., 2014] a tenté de découvrir le modèle d'apprentissage du CNN. Pour y arriver, il a testé la performance de chaque couche du modèle durant la phase d'apprentissage, puis il est parvenu à la conclusion selon laquelle, les couches de convolution peuvent apprendre plus de descripteurs d'ordre général et contiennent la grande partie de la capacité de représentation du CNN, tandis que les couches supérieures entièrement connectées sont spécifiques au domaine d'application. Hormis l'analyse des descripteurs du CNN, [Agrawal et al., 2014] ont étudié l'impact de certaines stratégies très répandues – telles que le pré-apprentissage et le réglage des paramètres – sur les performances du CNN et ont fourni des preuves soutenues permettant d'utiliser de façon moins empirique les modèles de CNNs sur des problèmes de la vision par ordinateur.

Malgré les progrès réalisés sur les principes théoriques de l'apprentissage profond, il reste encore beaucoup de chemin à faire pour une meilleure compréhension du fonctionnement et de l'optimisation des architectures du CNN, en vue d'améliorer les propriétés souhaitables telle que l'invariance à toutes les transformations géométriques et la discrimination entre les classes qui peuvent exister dans diverses applications de la classification.

4.12.2 La vision à l'échelle humain

La vision humaine possède une compétence remarquable pour les tâches spécifiques au domaine de la vision par ordinateur, aussi bien pour les simples représentations visuelles ou des transformations géométriques, des variations du fond de l'image et des occultations. En ce sens, on peut s'y référer pour trouver une solution efficace permettant de combler le fossé sémantique en terme de précision, ou s'en servir pour apporter de nouvelles perspectives d'études sur l'intégration des mécanismes du fonctionnement du cerveau humain dans des architectures d'apprentissage profond.

Comparé aux descripteurs bas niveau standards utilisés en vision par ordinateur, le CNN imite la structure du cerveau humain en construisant plusieurs couches d'activation pour des descripteurs de moyen et haut niveau. Le travail de [Wan et al., 2014] visait à évaluer le degré d'amélioration auquel on peut aboutir dans des tâches de recherche par le contenu en se servant des techniques d'apprentissage profond ; de plus ils ont tenté de déterminer si l'apprentissage profond était la clé pour combler à long terme le fossé sémantique. Sans pour autant répondre à cette question de façon nette, on peut constater qu'au fil des années, le taux d'erreur d'apprentissage sur la base ImageNet, s'est considérablement réduit, allant de 15% en 2012 [Krizhevsky et al., 2012] à 4.82% en 2015 [Ioffe et Szegedy, 2015]. Cette amélioration

prometteuse, vient confirmer l'efficacité des CNNs. En particulier, les résultats obtenus par [Ioffe et Szegedy, 2015] ont dépassé la précision des êtres humains. Cependant, on n'est pas encore en mesure de conclure que les performances de représentation des CNNs rivalisent avec celles du cerveau humain [Cadieu et al., 2014]. Par exemple, il est facile de produire des images qui sont complètement méconnaissables par l'être humain, tandis qu'un CNN affirmera (à tort) à 99.99% qu'elles contiennent des objets reconnaissables [Nguyen et al., 2015]. Ce résultat met en lumière la différence qui existe entre la vision humaine et les modèles actuels de CNNs. [Cadieu et al., 2014] a trouvé que, à l'image du cortex temporal inférieur, les récents modèles de CNNs pourraient générer des espaces de descripteurs similaires pour les mêmes catégories et distincts pour des images de catégories différentes. Ce résultat indique également que les CNNs peuvent donner un aperçu du fonctionnement du processus lié à la vision chez les primates.

Certes, il reste encore à démontrer que les modèles de CNNs – qui reposent sur des mécanismes de calcul – sont similaires au système visuel des primates, mais les modèles dont nous disposons aujourd'hui ont le potentiel de s'améliorer encore à condition d'imiter et d'exploiter les mécanismes de fonctionnement du système visuel des primates.

4.12.3 Apprentissage avec des données limitées

Les modèles de plus en plus larges font preuve d'un grand potentiel et ont le vent en poupe ; cependant, le manque de jeu de données pour l'apprentissage, peut réduire la capacité d'apprentissage de ces modèles, surtout quand il est coûteux d'obtenir des données étiquetées, comme dans notre cas. Comment pallier le besoin d'un important jeu de données et comment apprendre de façon efficace des réseaux dont la taille ne cesse de croître ? Ce sont autant de questions qui restent en suspens.

A ce jour, il existe deux solutions couramment utilisées pour obtenir beaucoup plus de données d'apprentissage. La première solution est de générer davantage de données d'apprentissage à partir des jeux de données existants, et ce à l'aide de techniques d'augmentation artificielle de données telles que la mise à l'échelle, les translations, les rotations et le *cropping*. La seconde solution est de collecter un plus grand nombre de données d'apprentissage en se servant de faibles classifieurs. Récement, il y a eu d'importants travaux de recherche sur l'apprentissage de concepts visuels sur les moteurs de recherche d'images [Chen et al., 2013], [Divvala et al., 2014]. Afin de développer des systèmes de reconnaissance en vision par ordinateur, [Zhou et al., 2015] a proposé le *ConceptLearner* qui pourrait apprendre des milliers de détecteurs de concepts visuels sur des jeux de données faiblement annotés. En outre, pour réduire le coût énorme de l'annotation par boîte englobante dans le cadre de la détection des objets, de nombreuses approches faiblement supervisées ont vu le jour [Song et al., 2014]. Toutefois, nous pensons qu'on serait plus avisés de développer des techniques pour générer ou collecter des jeux de données plus complets pour un meilleur apprentissage.

4.12.4 La complexité en temps de calcul

Les premières architectures de CNNs étaient vues commes des réseaux nécessitant beaucoup de ressources matérielles pour le calcul et n'étaient pas du tout adaptées à des applications temps réel. Aussi, l'une des tendances actuelles dans le domaine de l'apprentissage automatique

est la conception de nouvelles architectures de CNNs permettant une exécution temps réel. [He et al., 2015] ont mené une série d'expériences avec comme contrainte un coût de calcul limité et les modèles qui ont été proposés se sont révélés rapides pour des applications de la vie courante tout en étant aussi compétitifs que les modèles de CNNs existants. En effet, l'amélioration de la complexité en temps de calcul, permet également de comprendre les effets de plusieurs facteurs à savoir la profondeur du réseau, le nombre et la taille des noyaux de convolution. Par ailleurs, [Li et al., 2014] ont pour leur part, éliminé tous les calculs redondants liés à la rétropropagation du gradient, ce qui conduit à une accélération d'un facteur de 1500×. L'approche proposée jouit d'une flexibilité robuste pour différents modèles de CNNs – avec des conceptions et des structures différentes – et offre une performance élevée en raison de son implémentation basée sur un GPU assez puissant.

4.12.5 Des modèles plus puissants

Etant donné que les algorithmes d'apprentissage profond ont amélioré de façon significative des résultats de l'état-de-l'art pour les tâches relatives à la vision par ordinateur, il est de plus en plus difficile de faire mieux que les performances obtenues à ce jour. Toutefois, on peut tenter de dégager des pistes pour la mise en place de modèles plus puissants :

La première solution serait d'accroître la capacité de généralisation des CNNs en augmentant leur taille [Szegedy et al., 2015] et [Simonyan et Zisserman, 2014]. Des réseaux plus grand devraient normalement améliorer la qualité des performances ; mais il faudra porter une attention particulière au revers de la médaille, à savoir le sur-apprentissage et des ressources de calcul très importants.

Une autre piste serait de combiner les informations provenant de diverses sources. La fusion des descripteurs a longtemps été une approche très répandue en raison de son caractère attrayant et de ses bonnes performances. Dans le cas des CNNs, cette fusion peut se décliner en deux types :
(a) la combinaison des descripteurs de chaque couche du réseau, en raison du fait que différentes couches peuvent apprendre divers descripteurs [Girshick et al., 2014]. Cette approche peut donner des résultats prometteurs à condition que nous puissions développer une technique permettant de rendre complémentaire les descripteurs appris sur chaque couche. A titre d'exemple, [Liu et al., 2015] ont proposé d'intégrer plusieurs descripteurs de CNNs en incorporant soit différentes couches d'un modèle ou plusieurs couches de modèles distincts.
(b) la combinaison des descripteurs de divers types. On peut ainsi obtenir des modèles plus interprétables en incorporant d'autres types de descripteurs comme ceux de Gabor, le HOG ou les SIFT. [Zheng et al., 2014] ont par exemple combiné les SIFT aux descripteurs de CNNs extraits sur plusieurs patchs afin d'améliorer la puissance du *matching*.

Enfin, la dernière perspective vers des modèles plus puissants est de convevoir des architectures profondes spécifiques. En pratique, la plupart des approches basées CNNs utilisent un réseau déjà pré-appris sur ImageNet et qui de ce fait n'est pas assez spécialisé. Une meilleure façon de procéder serait d'apprendre des réseaux plus spécifiques, *c-à-d.* des réseaux conçus pour l'objet de notre intérêt, à savoir les masses dans notre cas. [Ouyang et al., 2014] a démontré qu'une annotation basée sur un objet d'intérêt dans l'image était plus utile qu'une annotation basée sur l'image entière dans le cadre d'une tâche de détection d'objet. Ceci peut donc être vu

comme une sorte de réseau spécialisé qui se concentre uniquement sur l'objet à rechercher au lieu de l'image. Une autre solution possible, serait d'apprendre différents réseaux sur diverses catégories d'images, comme l'ont fait [Yan et al., 2015] avec d'excellents résultats sur la base CIFAR100.

4.13 Conclusion

Dans cette dernière partie, nous avons abordé le problème de l'extraction et de la classification des descripteurs en vue de la reconnaissance du cancer du sein à travers d'une part la nature de la morphologie de la masse, et d'autre part le caractère graisseux ou dense du centre de cette dernière. Notre choix s'est porté sur l'utilisation d'un CNN (AlexNet) pour l'extraction des descripteurs et du SVM pour l'apprentissage d'un modèle et la classification. Ensuite, nous avons réalisé une comparaison avec les méthodes les plus efficaces de l'état-de-l'art. Les résultats prouvent que l'apprentissage non-supervisée de descripteurs par le CNN est très efficace et est en mesure d'appréhender – sans aucune intervention humaine – les différents descripteurs de haut niveau pouvant permettre d'apprendre les représentations les plus profondes liées aux données traitées. De plus, les descripteurs appris n'ont pas fait l'objet d'un pré-traitement explicite tel que la sélection des attributs dont la combinaison offrirait le meilleur taux de bonne classification avant la phase d'apprentissage. Ceci vient donc nous conforter dans le choix que nous avons fait de nous tourner vers une approche qui s'éloigne de la pratique qui a toujours eu cours en apprentissage automatique à savoir la conception de descripteurs permettant de décrire un comportement bien précis, avant l'extraction et la classification de ces derniers, avec les risques que l'on connaît. Pour finir, il est important de souligner que pour tirer le meilleur des CNNs, il faut nécessairement disposer d'une grande base de données ; dans notre cas, les bases utilisées sont de très petites tailles comparées aux applications évoquées dans la littérature, aussi les performances obtenues auraient peut-être été bien meilleures – notamment pour la densité/texture – si on avait fait une évaluation sur une base bien plus plus grande.

Conclusion générale et perspectives

Conclusion générale

Ce travail de thèse s'est intéressé de façon générale à l'étude des systèmes de DAOx en vue de la détection et de la discrimination des masses mammaires et plus précisément des masses malignes spiculées et des distorsions architecturales. Le choix de ce type de cancer se justifie par le fait que les statistiques de l'OMS et divers organismes de santé publique français ont montré que le cancer du sein représente une menace considérable pour les femmes à travers le monde et plus précisément en France où il est la première cause de mortalité par cancer chez les femmes. Ces études ont également démontré qu'une détection précoce et une prise en charge rapide de ce type de cancer durant les premières phases de son développement contribuent grandement à améliorer les chances de survie des femmes concernées tout en évitant d'avoir recours à des solutions invasives telles que la radiothérapie ou la mastectomie. La détection précoce passant obligatoirement par des campagnes de dépistage, les radiologues ont vu la quantité d'images mammographiques à interpréter exploser, ce qui a pour conséquence des taux d'erreur de plus en plus elevés variant entre 10 et 30% pour des masses manquées durant le dépistage. Cet état de choses s'explique par plusieurs facteurs dont la faible qualité des mammographies rendant difficile la détection des masses surtout quand elles sont embarquées dans des densités importantes, l'expérience des radiologues qui joue un rôle prépondérant dans leur diagnostic, la fatigue visuelle engendrée par une quantité importante d'images à interpréter. Cela a justifié le recours aux outils numériques permettant d'automatiser la procédure de détection ou de diagnostic. L'idée étant non pas de se substituer aux radiologues mais de les aider dans la prise de décision. Pour concevoir de tels systèmes, il est crucial de prendre connaissance des informations relatives au cancer du sein. Nous avons donc étudié dans un premier temps l'anatomie du sein, puis nous nous sommes intéressés au lien existant entre les images mammographiques et les structures anatomiques du sein. Finalement, l'accent a été mis sur la caractérisation des masses mammaires en fonction de leur degré de suspicion au travers de l'étude de la classification BI-RADS de l'ACR.

La conception de systèmes de DAOx fait intervenir une chaîne de traitement composée de quatre étapes que sont : le pré-traitement, la segmentation, l'extraction de divers descripteurs et l'apprentissage/classification. Dans ce travail, nous nous sommes intéressés à toutes ces étapes. Afin de concevoir notre système de DAOx dédié à la détection des masses spiculées et des distorsions architecturales – qui rappelons le constituent le signe de suspicion de cancer du sein le plus manqué par les radiologues, soit 35%, nous avons proposé une approche originale à toutes les étapes de la chaîne de traitement.

Premièrement, nous avons conçu une approche de pré-traitement, qui allie à la fois la robustesse, la rapidité numérique et supprime toute interaction avec l'utilisateur et ce dans le but d'améliorer la qualité de l'image pour faciliter l'éxécution des autres étapes. Notre méthode repose sur l'apprentissage d'un dictionnaire parcimonieux sur notre base d'images à l'aide du K-SVD. Cependant, l'apprentissage d'un dictionnaire soulève le problème de la taille de ce dernier, *c-à-d.* qu'elle est la taille offrant les performances optimales ? Dans la pratique, plus la taille est grande et meilleures sont les performances, mais cela implique un temps de calcul important.

Nous avons donc proposé de combiner le K-SVD avec l'ACP en partant de l'hypothèse selon laquelle la présence des motifs répétitifs présents en mammographie peuvent aider à réduire de façon importante la taille du dictionnaire appris. Cette combinaison a permis de trouver une taille adéquate pour un dictionnaire offrant des performances optimales. Une comparaison avec les méthodes les plus efficaces de l'état-de-l'art, *p. ex.* les ondelettes de Daubechies et les *curvelet* nous a permis de démontrer la pertinence de notre approche. Cet algorithme a surpassé la suppression du bruit basée sur les ondelettes de Daubechies, mais requiert un temps de calcul plus grand.

A la suite du pré-traitement, nous avons travaillé sur la segmentation des masses et plus précisément sur la segmentation des masses malignes embarquées dans des densités très élevées rendant difficile leur détection. L'étude préliminaire des méthodes de segmentation les plus utilisées en imagerie du sein a permis d'observer que les techniques les plus efficaces sont celles qui utilisent des formulations variationnelles. Nous avons par la suite mentionné la difficulté rencontrée par les approches basées sur des formulations variationnelles dans l'optimisation des différentes fonctions d'énergie qui convergent vers un minimum local très éloigné du minimum global. Dans ce contexte, les approches d'optimisation par coupure de graphes ont été abordées du fait de leur capacité à fournir un minimum local proche de l'optimum global. Nous avons par la suite évoqué les méthodes de segmentation par MRF en raison de leur aptitude à prendre en compte les interactions locales entre les pixels. Une étude des techniques de segmentation basées MRF permet de constater que la quasi totalité de ces méthodes est supervisée, ce qui engendre un temps de calcul exponentiel et la spécification de plusieurs hyper-paramètres. Ceci nous a conduit à la conception d'une méthode de segmentation non-supervisée basée sur les MRFs qui est capable de donner de bonnes performances en un laps de temps, tout en ne nécessitant la spécification que d'un seul paramètre. L'approche proposée (log-PRF) est fondée sur l'utilisation des champs de Pickard et exploite une méthode d'optimisation basée sur l'utilisation des chaînes de Markov télégraphique et d'une estimation des paramètres par le biais de l'algorithme EM. Les performances obtenues démontrent la supériorité de notre approche comparée à toutes les méthodes de segmentation proposées dans l'état-de-l'art et ce sur tous les aspects : temps de calcul, robustesse, nombre d'hyper-paramètres à fournir. Pour ce qui est de la rapidité numérique, le log-PRF est respectivement 47× et 9193× plus rapide que les approches basées sur les coupures minimales de graphes et l'ICM.

Après avoir constaté qu'en dehors de quelques travaux, la majorité des techniques existantes n'ont jamais tenté de segmenter les structures anatomiques afin de s'en servir pour justifier la nature maligne ou bénigne de la masse et étant donné l'importance que revêt la segmentation des spicules aussi bien pour la distinction entre les masses malignes et bénignes que pour la caractérisation des distorsions architecturales, nous avons proposé une méthode originale s'inspirant des techniques utilisées en télédétection, pour extraire convenablement les spicules. L'approche conçue s'affranchit de l'utilisation de descripteurs extraits manuellement et classifiés et repose sur la discrétisation des structures curvilignes de l'image à l'aide d'une transformée locale de Radon ; ensuite en se servant d'un champ de Markov prenant en compte des contraintes géométriques sur la nature des structures anatomiques, nous avons recomposé les structures curvilignes de l'image de départ. Pour finir, l'utilisation d'un modèle a contrario nous a permis de retenir uniquement les structures curvilignes qui sont des spicules. Ce qui a permis à la fois de caractériser les masses spiculées et les distorsions architecturales. Une étude comparative avec les

méthodes utilisées dans l'état-de-l'art permet de prouver la qualité de notre approche. Ainsi, les résultats obtenus après évaluation indiquent que notre méthode est meilleure que celles proposées dans l'état-de-l'art, même si, à ce stade, nos résultats ne peuvent être considérés que comme préliminaires. Par ailleurs, nous avons évalué le temps de calcul de notre approche ; en effet, l'une des forces de notre méthode est sa capacité à extraire des spicules tout en maintenant un temps de calcul très faible. Ainsi, le traitement d'une image de 300×300 pixels requiert 4 secondes et environ 1 minute pour une image de 3000×4000 pixels, contre 90 minutes pour la classification des structures anatomiques présentes sur une mammographie de 2000×3000 pixels par la meilleure des approches proposées dans l'état-de-l'art.

Pour finir, nous avons proposé un modèle pour l'aide au diagnostic basé sur une combinaison d'une méthode de *Deep learning* avec un classifieur de type SVM. Dans un premier temps nous utilisons un CNN nommé AlexNet pour extraire de façon non-supervisée les divers descripteurs de bas, moyen et haut niveau sur nos images, puis les descripteurs extraits sont appris par un SVM afin de construire un modèle pouvant permettre la discrimination des masses d'une part sur leur morphologie et d'autre part sur la densité de leur centre. L'approche proposée a été comparée aux performances des meilleurs descripteurs de forme spécifiques existants dans la littérature afin d'évaluer son efficacité dans la classification des masses selon leur forme. Les résultats montrent que cette approche a dépassé les performances obtenues dans la littérature. De plus, on s'est comparé aux travaux réalisés dans l'état-de-l'art sur la classification de la densité des images mammographiques ; les résultats obtenus sont très intéressants, même s'ils sont encore loin des performances des meilleures méthodes proposées dans la littérature. Finalement, nous avons proposé d'intégrer les spicules extraits dans le processus d'interprétation des images afin de discriminer définitivement les différents cas de pathologie en imagerie du sein. A notre sens, ceci, ne relève que de la seule expertise du radiologue.

Perspectives

Les perspectives que nous envisageons dans le prolongement de ce travail de thèse sont d'une part, les défis liés à l'utilisation du *Deep learning* et d'autre part, les avantages et inconvénients liés à l'acquisition des images du sein par le biais de la tomosynthèse du sein dans le but d'améliorer la détection des distorsions architecturales et des masses spiculées. Par ailleurs, dans son état actuel, notre système ne donne que des indications au radiologue afin de l'orienter dans son diagnostic. Il serait peut-être convenable d'approfondir les expériences en vue de concevoir un système capable de fournir un avis tranché. Cela comporte bien évidemment des risques importants pour la santé des patientes en cas de mauvais diagnostic.

Annexe A

Les descripteurs de forme

A.1 Les descripteurs standards

A.1.1 Les moments de Hu

Un moment est une somme pondérée de tous les pixels en fonction de leur localisation dans l'image ; ils sont pour cela des descripteurs de forme très puissants. Plusieurs moments peuvent être calculés à partir de la silhouette d'un objet ou de son contour ; la silhouette est moins sensible au bruit et représente un indicateur de la forme brute de l'objet, tandis que le contour est plus sensible aux éléments de hautes fréquences présents dans l'image. Le calcul des moments est assez simple, mais présente trois inconvénients : le premier est la difficulté relative à l'association des moments avec la perception humaine, le second est le besoin de calculer des moments d'ordre élevé afin d'être le plus discriminant possible dans une tâche de classification et enfin le dernier est la sensibilité des moments d'ordre élevé à l'erreur d'échantillonnage issu de la numérisation des contours de l'objet [Duda et al., 2012].

Le moment m_{pq} d'ordre $(p+q)^{eme}$ et le centre de gravité $(\overline{x}, \overline{y})$ d'un objet sont définis comme :

$$m_{pq} = \sum_i \sum_j i^p j^q I(i,j), \quad p,q = 0,1,2,\cdots \ . \tag{A.1}$$

$$\overline{x} = \frac{m_{10}}{m_{00}}, \quad \overline{y} = \frac{m_{01}}{m_{00}}, \ . \tag{A.2}$$

où la somme se fait sur tous les pixels de la silhouette ou le contour du ROI, et $I(i,j)$ est la valeur du niveau de gris du pixel situé à l'emplacement (i,j). [Hu, 1962] a proposé un ensemble de sept moments centraux et invariants aux transformations géométriques de faible ordre.

$$\begin{aligned}
M_1 &= \mu_{20} + \mu_{02} \ , \\
M_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \ , \\
M_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \ , \\
M_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} - \mu_{03})^2 \ , \\
M_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \times [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})]^2 + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \ , \\
M_6 &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \ , \\
M_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12}) \times [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - (\mu_{03} - 3\mu_{12})(\mu_{21} + \mu_{03}) \ .
\end{aligned} \tag{A.3}$$

avec $\mu_{00} = m_{00} = \mu$, $\mu_{10} = \mu_{01} = 0$, $\mu_{20} = m_{20} - \mu\overline{x}^2$, $\mu_{11} = m_{11} - \mu\overline{x}\overline{y}$, $\mu_{02} = m_{02} - \mu\overline{y}^2$, $\mu_{30} = m_{30} - 3m_{20}\overline{x} + 2\mu\overline{x}^3$, $\mu_{21} = m_{21} - m_{20}\overline{y} - 2m_{11}\overline{x} + 2\mu\overline{x}^2\overline{y}$, $\mu_{12} = m_{12} - m_{02}\overline{x} - 2m_{11}\overline{y} + 2\mu\overline{x}\overline{y}^2$ et $\mu_{03} = m_{03} - 3m_{02}\overline{y} + 2\mu\overline{y}^3$.

A.1.2 Les moments de Zernike

Les travaux menés par [Teague, 1980] ont suggérés que l'utilisation des bases orthogonales en termes de polynômes de Zernike sont à la fois de bons descripteurs de forme, mais sont également invariants à la rotation, robuste au bruit, tout en étant une représentation multi-échelle de la forme [Kim et Kim, 2000]. Les polynômes de Zernike basés sur la représentation se révèlent être plus robustes au bruit, permettant une reconstruction minimisant la perte d'information [Wee et Paramesran, 2007]. Les polynômes de Zernike sont par définition, un ensemble de polynômes complexes qui forment un ensemble orthogonal complet $V_{pq}(x, y)$ à l'intérieur d'un cercle unitaire et sont définis comme suit :

$$V_{pq}(x, y) = R_{pq}(r)e^{jq\theta}, \quad \text{avec } r[-1, 1] \ . \tag{A.4}$$

où $r = \sqrt{x^2 + y^2}$ est l'amplitude et $\theta = \tan^{-1}(\frac{y}{x})$ l'angle. En général, les moments de Zernike sont définis comme :

$$Z_{pq} = \frac{p+1}{\pi} \int_{-\pi}^{\pi} \int_{0}^{1} [V_{pq}(r, \theta)]^* f(r, \theta) r dr d\theta \ . \tag{A.5}$$

où $f(r, \theta)$ est l'image dans le repère polaire. Avec une approximation numérique, les moments complexes de Zernike sont déduits de la valeur réelle du polynôme radial donné par :

$$R_{pq}(r) = \sum_{x=0}^{(p-|q|)/2} (-1)^s \frac{(p-s)!}{s!(\frac{p+|q|}{2} - s)!(\frac{p-|q|}{2} - s)!} r^{p-2s} \ . \tag{A.6}$$

où p et q sont choisies de telle sorte que $p - |q|$ est paire, avec $0 \leq |q| \leq p$ et $p \geq 0$. Donc les moments complexes de Zernike d'ordre p avec q répétitions de la fonction d'intensité de l'image $f(x, y)$ sont donnés par :

$$Z_{pq} = \frac{p+1}{\pi} \sum_{x} \sum_{y} V_{pq} * (x, y) f(x, y), r dr \ . \tag{A.7}$$

où le symbole $*$ a été mis pour signifier la conjuguée complexe de $V_{pq}(x, y)$.

A.1.3 Les moments de Krawtchouk

Récemment, un nouvel ensemble des fonctions de moments orthogonaux discrets basés sur les polynômes discrets de Krawtchouk a été introduit par [Yap et al., 2003], qui ont démontré que les moments orthogonaux discrets donnent de meilleurs résultats par rapport aux moments orthogonaux continus en terme de représentation de l'image. Des expériences menées à la suite de ces travaux ont prouvées que les moments de Krawtchouk sont meilleurs que ceux de Zernike pour l'analyse de la forme. Par ailleurs, contrairement aux moments orthogonaux qui ne permettent qu'une description globale de l'image, les moments de Krawtchouk peuvent être utilisés pour extraire des descripteurs locaux dans une image.

Les fonctions de base des moments de Krawtchouk sont des polynômes orthogonaux discrets de Krawtchouk satisfaisant la relation suivante :

$$\sum_{x=0}^{N} j(x) k_n(x) k_n(x) = \rho(N, n, p) \delta_{pq}, \quad \text{tel que : } 0 \leq m, n \leq N \ . \tag{A.8}$$

où $\rho(N,n,p) = \begin{pmatrix} N \\ n \end{pmatrix} p^n(1-p)^n$, $q = (1-p)$ et le poids $j(x) = \begin{pmatrix} N \\ x \end{pmatrix} p^x q^{N-x}$. La représentation hyper-géométrique (F) explicite du polynôme de Krawtchouk est donnée par [Yap et al., 2003] comme étant :

$$k_n(x) = q^n \begin{pmatrix} x \\ n \end{pmatrix} F(-n, x-N; x-n, \frac{-p}{q}) \ . \tag{A.9}$$

Les moments de Krawtchouk, contrairement aux moments de Zernike et de Legendre appartiennent à la classe des moments orthogonaux discrets, ce qui fait que leur implémentation n'implique aucune approximation numérique. De plus, les polynômes de Krawtchouk ne requièrent aucune transformation des coordonnées spatiales. Les moments de Krawtchouk d'ordre $m+n$ sont définis comme suit :

$$k_{nm}(x) = [\rho(N,n,p)\rho(N,m,p)]^{-1} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} j(x)k_n(x)j(y)k_m(y)f(x,y) \ . \tag{A.10}$$

avec $m, n = 0, 1, 2...N$ où les valeurs de $k_n(x)$ et $k_m(y)$ données par A.9 seront utilisées comme l'ensemble de base. La transformée inverse du moment calculée comme ci-dessous est enfin utilisée pour reconstruire l'image.

$$f(x,y) = \sum_{m=0}^{N} \sum_{n=0}^{N} k_{nm}(x)k_n(x)k_m(y) \ . \tag{A.11}$$

A.2 Les descripteurs spécifiques

A.2.0.1 La longueur radiale normalisée

Dans le but de décrire convenablement le contour des masses, [Kilday et al., 1993] ont proposé six descripteurs basés sur la longueur radiale normalisée, traduction de l'anglicisme *Normalized Radial Length* (NRL). Premièrement, la longueur radiale est calculée comme étant la distance euclidienne entre le centre de gravité de la tumeur et chaque pixel situé sur le contour de la masse ; puis elle est normalisée par le maximum de la distance radiale trouvée. De façon plus formelle, chaque longueur radiale est calculée comme ci-dessous :

$$r_k = \frac{\sqrt{(x(k)-x_c)^2 + (y(k)-y_c)^2}}{max(r(k))}, \quad \text{avec } k = \{1, 2, \cdots N\} \ . \tag{A.12}$$

où $(x(k), y(k))$ et (x_c, y_c) sont respectivement les coordonnées du k^{ieme} pixel et du centre de gravité de la masse et N le périmètre. Par la suite, les six descripteurs suivants ont été extraits :

1. **la moyenne de la NRL** : l'objectif de ce descripteur (m) est de mesurer la variation à l'échelle macrocospique du contour, d'une façon similaire à la circularité.

$$m = \frac{1}{N} \sum_{k=1}^{N} r(k) \ . \tag{A.13}$$

2. **l'écart-type de la NRL** : le but de l'écart-type (σ) est d'évaluer le caractère irrégulier du contour de la masse ; plus sa valeur est grande et plus le contour de la masse est irrégulier.

$$\sigma = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(r(k) - m(k))^2} \ . \tag{A.14}$$

3. **l'entropie de la NRL** : calculée à l'aide de l'histogramme de la NRL, c'est une métrique probabiliste qui détermine la façon dont la longueur radiale d'une tumeur pourrait être estimée. L'entropie (E) se calcule comme suit :

$$E = \sum_{k=1}^{100} p_k log(p_k) \ . \tag{A.15}$$

où p_k est la probabilité que la longueur radiale varie entre r et $r + \Delta$, où Δ est un pas régulier d'une valeur de $1/100$. L'entropie ainsi calculée, intègre à la fois la notion de circularité et d'irrégularité.

4. **le ratio de surface de la NRL** : ce descripteur (A_r) évalue le pourcentage de la partie de l'objet sans la région circulaire, définie par la moyenne de la NRL.

$$A_r = \frac{1}{m \times N}\sum_{k=1}^{N}(r(k) - m) \ . \tag{A.16}$$

où

$$A_r = 0 \quad \forall r(k) \leq m \ . \tag{A.17}$$

5. **la rugosité de la NRL** : elle a pour but d'isoler la forme macrocospique de la masse à partir de la structure des bords. Elle permet d'obtenir la moyenne des pixels voisins. Ainsi, plus les contours sont irréguliers, plus leur valeur de rugosité (R) est élevée.

$$R = \frac{1}{N}\sum_{k=1}^{N}(r(k) - r(k+1)) \ . \tag{A.18}$$

6. **le taux de croissement en zéro de la NRL** : cette mesure sert à extraire les informations concernant les petites variations du contour. De façon pratique, elle calcule le nombre de fois que la droite définie par la moyenne NRL intercepte le contour de la masse. Elle sert donc à évaluer le degré de spiculation du contour.

A.2.0.2 La longueur de la corde

[You et Jain, 1984] ont proposé quatre descripteurs fondés sur la longueur de la corde ou *Normalized Chord Length* (NCL) afin de discriminer des objets 2D ayant un contour fermé. Une corde L_i est définie comme un segment d'une ligne qui relie une paire de points situés sur le contour et normalisée par la longueur de la plus grande corde. L'ensemble complet des cordes pour un objet donné, est l'ensemble de toutes les cordes possibles, dessinées en partant d'un pixel situé sur le bord de la masse vers un autre pixel situé sur le même bord. Les auteurs

ont considéré les $K = N(N-1)/2$ longueurs de la corde des N pixels qui forment le contour d'un objet comme un échantillon d'un ensemble de distribution et ont calculé les métriques de Kolmogorov-Smirnov comme des descripteurs de dissimilarité. Les auteurs ont démontré que les descripteurs extraits en se servant de la méthode de la corde sont invariants à l'échelle, à la translation et à la rotation ; de plus, ces derniers sont robustes au bruit. Les descripteurs de forme basés sur la longueur de la corde sont les suivants :

1. **la moyenne de la NCL** :

$$M_{c1} = \frac{1}{K} \sum_{i=1}^{K} L_i \ , \tag{A.19}$$

2. **l'écart-type de la NCL** :

$$M_{c2} = \sqrt{\frac{1}{K} \sum_{i=1}^{K} (L_i - M_{c1})^2} \ , \tag{A.20}$$

3. **l'asymétrie ou *skewness* de la NRL** :

$$M_{c3} = \frac{1}{M_{c2}^3} \frac{1}{K} \sum_{i=1}^{K} (L_i - M_{c1})^3 \ , \tag{A.21}$$

4. **l'aplatissement ou *skewness* de la NRL** :

$$M_{c4} = \frac{1}{M_{c2}^4} \frac{1}{K} \sum_{i=1}^{K} (L_i - M_{c1})^4 \ . \tag{A.22}$$

Annexe B

Les descripteurs standards de texture

B.1 Les filtres dyadiques de Gabor

Ce sont des fonctions de noyau Gaussien modulé avec une sinusoide, souvent employés pour la détection de contour, ils essaient de copier les fonction de la vision humaine. En analyse de texture, la méthode consiste à convoluer l'image avec des filtres de Gabor. Les textures sont alors représentées par leurs réponses à ces filtres extrayant une information fréquentielle localisée et orientée. Un filtre de Gabor horizontal en $(0, 0)$ de paramètres $(\sigma_1, \sigma_2, f_0)$ – où f_0 est la fréquence radiale et (σ_1, σ_2) les écart-type – a pour réponse impulsionnelle [Jain et Farrokhnia, 1991].

$$\psi(x, y) = e^{-\frac{1}{2}(\frac{x^2}{\sigma_1} + \frac{y^2}{\sigma_2})} \cos(2\pi f_0 x) \ . \tag{B.1}$$

Les autres filtres s'obtiennent par des rotations de θ et des translations de (x_0, x_0). Certains auteurs ont conseillé l'utilisation de cinq fréquences radiales différentes et de quatre orientations, c'est le cas de [Jain et Farrokhnia, 1991]. Pour finir, la texture est caractérisée par différentes statistiques sur les réponses aux filtres.

B.2 Les ondelettes

La transformation en ondelettes discrètes permet de choisir une famille complète de filtres de décomposition en sous-bandes. L'image est alors décomposée en ses composantes fréquentielles ; cette transformation offre une localisation spatiale et une flexibilité des fonctions de bases qui sont connues sous le nom d'ondelettes. Le résultat de la transformée en ondelettes dépend de la fonction de base, dont les plus célèbres sont celles de Haar, Daubechies, Coifman et Mallat. Toutefois les ondelettes de Daubechies [Daubechies, 1992] demeurent l'une des bases d'ondelettes les plus utilisées pour l'analyse de texture en raison de leur bonne efficacité de calcul, de leurs performances pour séparer les fréquences ainsi que de la variété des fonctions qu'elles autorisent.

De façon générale, le résultat obtenu à chaque décomposition est une image de résolution inférieure à celle de départ – appelée coefficient d'approximation – et trois images contenant les informations de haute résolution en position horizontale, verticale et diagonale, appelées coefficients de détails. Par ailleurs, le fait que l'image soit décomposée dans des gammes de fréquences plus basses permet une analyse en multi-résolution [Porter et Canagarajah, 1997].

Annexe C

Approches de sélection des descripteurs

C.1 La sélection par mérite individuel

La sélection par filtrage selon le mérite individuel utilise un critère d'évaluation qui se calcule sur chaque descripteur et ce de façon indépendante les uns des autres. Ce critère évalue la pertinence d'un descripteur selon les propriétés mesurées sur le jeu de données utilisé pour l'entrainement. L'objectif est alors de calculer un score et de retenir uniquement les descripteurs ayant obtenus la meilleure valeur associée à ce score. Il existe plusieurs critères utilisés dans la littérature, comme celui de Fischer [Duda et al., 2012], celui de la corrélation et l'information mutuelle [Guyon et Elisseeff, 2003]. Nous n'évoquerons ici que le score de Fischer en raison de son utilisation répétée en mammographie.

Le critère de Fischer évalue la divergence D entre deux classes i et j en se basant sur la moyenne μ et l'écart-type σ estimés du descripteur k pour ces deux classes :

$$D(k) = \frac{(\mu_i - \mu_j)^2}{\sigma_i^2 + \sigma_j^2} \quad . \tag{C.1}$$

Ce critère va être optimisé dans le but de trouver la séparation la plus grande possible entre les deux classes. On souhaite donc éviter le moins de d'intersection possible entre les deux classes une fois que les échantillons sont projetés sur l'axe du descripteur pour lequel le critère de Fischer est calculé. Dans un contexte de multi-classe, le critère devient la divergence moyenne globale des critères de Fischer pour des classes prises deux à deux. Une autre approche consiste à évaluer la divergence de la moyenne de la classe par rapport à la moyenne des échantillons [Gu et al., 2012].

C.2 La sélection séquentielle

Les méthodes de sélection séquentielle sont des techniques itératives dont l'objectif est la recherche d'un ensemble sous-optimal de descripteurs. Elles consistent à évaluer les performances d'un modèle pour un ensemble de descripteurs dont la taille augmente ou diminue. Les méthodes les plus simples sont la *sequential forward selection* (SFS) et la *sequential backward selection* (SBS). En dépit de leur efficacité, le SFS et le SBS souffrent d'un problème majeur à savoir qu'elles ne réalisent pas le parcours total de l'espace de recherche des descripteurs. Afin de pallier à ce problème, des techniques bien plus élaborées ont été développées, au nombre desquelles nous avons la *sequential floating forward selection* (SFFS) et la *sequential floating forward selection* (SFFS). Ici, nous nous contenterons d'exposer la méthode dont l'utilisation est la plus récurrente en mammographie, à savoir la SFS.

La procédure de la SFS débute avec un modèle vide de sous-ensemble des descripteurs sélectionnés. A chaque itération, le meilleur descripteur selon un certain critère de performance

est retenu parmi l'ensemble des descripteurs à sélectionner, puis il est supprimé de cet ensemble et ajouté au modèle. Ce processus est réitéré jusqu'au critère d'arrêt qui correspond d'habitude soit à une performance qui stagne ou à l'atteinte du nombre maximum d'itérations. Comme on peut l'observer, cette technique est simple et rapide à mettre en œuvre ; toutefois, elle ne prend en compte que l'interaction d'un descripteur avec le sous-ensemble déjà sélectionné et non un ensemble plus grand [Marill et Green, 1963].

C.3 La sélection par algorithmes génétiques

Conçus dans les années 70, les algorithmes génétiques constituent une méthode d'optimisation efficace. Dans un algorithme génétique, on a une population de base qui est d'habitude composée d'une chaîne de caractères représentant chacune un chromosome. Le but de l'algorithme génétique est de réaliser des copies de chaînes et des échanges de morceaux de chaîne. Globalement, la procédure d'un algorithme génétique se décline en trois étapes que sont la sélection, le croisement et la mutation. La sélection équivaut à la copie de chaque chaîne en fonction des valeurs d'une fonction à optimiser. Une fois la sélection faite, on réalise le croisement, qui correspond successivement à un appariement et à une permutation aléatoires des éléments obtenus après sélection. Enfin, intervient la mutation dont le rôle est de modifier de façon aléatoire un élément de la chaîne [Kudo et Sklansky, 2000]. Ces trois étapes sont itérées jusqu'à ce qu'on trouve des résultats qui n'évoluent plus.

Liste des publications

Revues internationales avec comité de lecture

Sègbédji R. T. J. Goubalan, Y. Goussard, H. Maaref. log-PRF : Unsupervised malignant mammographic breast mass segmentation algorithm based on Pickard Markov random field. En cours de soumission à *IEEE Transactions on Biomedical Engineering*.

Sègbédji R. T. J. Goubalan, Y. Goussard, H. Maaref. Segmentation of spicules of malignant masses and architectural distortion in mammography. En cours de soumission à *Pattern Recognition*.

Conférences internationales avec actes et comité de lecture

Sègbédji R. T. J. Goubalan, Y. Goussard, H. Maaref. Segmentation of spiculated breast mass. *42th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 5-9 Mars 2017, New-Orleans, USA* (accepté).

Sègbédji R. T. J. Goubalan, Y. Goussard, H. Maaref. Unsupervised malignant mammographic breast mass segmentation algorithm based on Pickard Markov random field. *23th IEEE International Conference on Image Processing (ICIP), 25-28 September 2016, p.2653–2657, Phoenix, USA*.

Sègbédji R. T. J. Goubalan, Y. Goussard, H. Maaref. Markovian Approach to Automatic Annotation of Breast Mass Spicules Using an A Contrario Model. *13th International Workshop on Digital Mammography (IWDM), 19-22 June 2016, p.461–468, Malmö, Sweden* (*Best paper 1000$ US*).

Sègbédji R. T. J. Goubalan, K. Djemal, H. Maaref. Optimization of the Dictionary Size Selection : An Efficient Combination of K-SVD and PCA to Denoise and Enhance Digital Mammography Contrast. *International Symposium on Intelligent Computing Systems, 16-18 March 2016, p.1–15, Merida, Mexico*.

Exposé

Sègbédji R. T. J. Goubalan, Y. Goussard, H. Maaref. Analyse d'images médicales pour la reconnaissance du cancer du sein. *Rencontre des jeunes chercheurs africains en France (APSA), 1-2 Décembre 2016, Institut Henri Poincaré, Paris, France*.

Sègbédji R. T. J. Goubalan, K. Djemal, H. Maaref. Masses indexation for breast cancer recognition. *MAP Breast, 13-17 Juillet 2015, Porto, Portugal*.

Poster

Sègbédji R. T. J. Goubalan, Y. Goussard, H. Maaref. Vers un DAOx dédié au cancer du sein. *Evry Sciences Innovation, Juin 2016, Evry, France*.

Bibliographie

[Abdel-Dayem et El-Sakka, 2005] Abdel-Dayem, A. R. and El-Sakka, M. (2005). Fuzzy entropy based detection of suspicious masses in digital mammogram images. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 4017–4022. IEEE. (Cité en page 53.)

[ACR et al., 2003] ACR, of Radiology, A. C., et al. (2003). Breast imaging reporting and data system atlas (bi-rads atlas). *Reston, VA : American College of Radiology*, 98. (Cité en pages 19 et 21.)

[Agrawal et al., 2014] Agrawal, P., Girshick, R., and Malik, J. (2014). Analyzing the performance of multilayer neural networks for object recognition. In *European Conference on Computer Vision*, pages 329–344. Springer. (Cité en page 114.)

[Aharon et al., 2006] Aharon, M., Elad, M., and Bruckstein, A. (2006). K-svd : An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11) :4311–4322. (Cité en pages 34 et 35.)

[Akaike, 1974] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6) :716–723. (Cité en page 57.)

[Allen et al., 2003] Allen, B. H., Oxley, M. E., and Collins, M. J. (2003). A universal segmentation platform for computer-aided detection. In *Digital Mammography*, pages 164–168. Springer. (Cité en page 51.)

[Alto et al., 2005] Alto, H., Rangayyan, R. M., and Desautels, J. L. (2005). Content-based retrieval and analysis of mammographic masses. *Journal of Electronic Imaging*, 14(2) :023016–023016. (Cité en page 94.)

[Bailar, 1976] Bailar, J. C. (1976). Mammography : a contrary view. *Annals of Internal Medicine*, 84(1) :77–84. (Cité en page 20.)

[Baker et al., 2003] Baker, J. A., Rosen, E. L., Lo, J. Y., Gimenez, E. I., Walsh, R., and Soo, M. S. (2003). Computer-aided detection (cad) in screening mammography : sensitivity of commercial cad systems for detecting architectural distortion. *American Journal of Roentgenology*, 181(4) :1083–1088. (Cité en pages 24 et 29.)

[Baldi et Sadowski, 2013] Baldi, P. and Sadowski, P. J. (2013). Understanding dropout. In *Advances in Neural Information Processing Systems*, pages 2814–2822. (Cité en page 102.)

[Baum et al., 1970] Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains. *The Annals of Mathematical Statistics*, 41(1) :164–171. (Cité en page 62.)

[Bengio et Courville, 2016] Bengio, I. G. Y. and Courville, A. (2016). Deep learning. Book in preparation for MIT Press. (Cité en page 101.)

[Bengio, 2012] Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. In *Neural Networks : Tricks of the Trade*, pages 437–478. Springer. (Cité en pages 100, 101 et 103.)

[Bergstra et Bengio, 2012] Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(Feb) :281–305. (Cité en page 103.)

[Berks et al., 2011] Berks, M., Chen, Z., Astley, S., and Taylor, C. (2011). Detecting and classifying linear structures in mammograms using random forests. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 510–524. Springer. (Cité en pages 82 et 83.)

[Besag, 1986] Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 48(3) :259–302. (Cité en page 57.)

[Beucher et Lantuéjoul, 1979] Beucher, S. and Lantuéjoul, C. (1979). Use of watersheds in contour detection. (Cité en page 49.)

[Bezdek, 2013] Bezdek, J. C. (2013). *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media. (Cité en page 52.)

[Binder-Foucard et al., 2013] Binder-Foucard, F., Belot, A., Delafosse, P., Remontet, L., Woronoff, A., Bossard, N., et al. (2013). Estimation nationale de lâincidence et de la mortalité par cancer en france entre 1980 et 2012. *Etude à partir des registres des cancers du réseau Francim. Partie*, 1 :122. (Cité en pages vii, 9 et 10.)

[Bishop, 2006] Bishop, C. M. (2006). Pattern recognition. *Machine Learning*, 128. (Cité en page 52.)

[Blake et al., 2004] Blake, A., Rother, C., Brown, M., Perez, P., and Torr, P. (2004). Interactive image segmentation using an adaptive gmmrf model. In *European Conference on Computer Vision*, pages 428–441. Springer. (Cité en page 58.)

[Blessing et al., 2006] Blessing, M., Godfrey, D., Lohr, F., and Yin, F. (2006). Th-c-330a-05 : Analysis of the point spread function of isocentric digital tomosynthesis (dts). *Medical Physics*, 33(6) :2266–2266. (Cité en page 18.)

[Bosch et al., 2006] Bosch, A., Munoz, X., Oliver, A., and Marti, J. (2006). Modeling and classifying breast tissue density in mammograms. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 1552–1558. IEEE. (Cité en pages 93, 109 et 113.)

[Bovis et Singh, 2002] Bovis, K. and Singh, S. (2002). Classification of mammographic breast density using a combined classifier paradigm. In *Medical image understanding and analysis (MIUA) conference, Portsmouth*. Citeseer. (Cité en pages 92, 109 et 112.)

[Boyd et al., 1995] Boyd, N., Byng, J., Jong, R., Fishell, E., Little, L., Miller, A., Lockwood, G., Tritchler, D., and Yaffe, M. J. (1995). Quantitative classification of mammographic densities and breast cancer risk : results from the canadian national breast screening study. *Journal of the National Cancer Institute*, 87(9) :670–675. (Cité en page 20.)

[Boykov et Jolly, 2001] Boykov, Y. and Jolly, M.-P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 105–112. IEEE. (Cité en page 58.)

[Boykov et al., 2001] Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11) :1222–1239. (Cité en pages 58 et 67.)

[Brzakovic et al., 1990] Brzakovic, D., Luo, X., and Brzakovic, P. (1990). An approach to automated detection of tumors in mammograms. *IEEE Transactions on Medical Imaging*, 9(3) :233–241. (Cité en pages 52 et 90.)

[Cadieu et al., 2014] Cadieu, C. F., Hong, H., Yamins, D. L., Pinto, N., Ardila, D., Solomon, E. A., Majaj, N. J., and DiCarlo, J. J. (2014). Deep neural networks rival the representation of primate it cortex for core visual object recognition. *PLoS Comput Biol*, 10(12) :e1003963. (Cité en page 115.)

[Candes et al., 2006] Candes, E., Demanet, L., Donoho, D., and Ying, L. (2006). Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, 5(3) :861–899. (Cité en pages 33 et 40.)

[Candes, 2012] Candes, E. J. (2012). Curvelab-2.1.3. online. (Cité en page 39.)

[Canny, 1986] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6) :679–698. (Cité en page 51.)

[Cao et al., 2004a] Cao, A., Song, Q., Yang, X., and Liu, S. (2004a). Breast mass segmentation on digital mammograms by a combined deterministic annealing method. In *IEEE International Symposium on Biomedical Imaging : Nano to Macro, 2004*, pages 1303–1306. IEEE. (Cité en page 52.)

[Cao et al., 2004b] Cao, A., Song, Q., Yang, X., and Wang, L. (2004b). Breast mass segmentation based on information theory. In *ICPR 2004. Proceedings of the 17th International Conference on Pattern Recognition, 2004*, volume 3, pages 758–761. IEEE. (Cité en page 52.)

[Carneiro et al., 2015] Carneiro, G., Nascimento, J., and Bradley, A. P. (2015). Unregistered multiview mammogram analysis with pre-trained deep learning models. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 652–660. Springer. (Cité en page 55.)

[Cascio et al., 2006] Cascio, D., Fauci, F., Magro, R., Raso, G., Bellotti, R., De Carlo, F., Tangaro, S., De Nunzio, G., Quarta, M., Forni, G., et al. (2006). Mammogram segmentation by contour searching and mass lesions classification with neural network. *IEEE Transactions on Nuclear Science*, 53(5) :2827–2833. (Cité en page 51.)

[Catarious Jr et al., 2004] Catarious Jr, D. M., Baydush, A. H., and Floyd Jr, C. E. (2004). Incorporation of an iterative, linear segmentation routine into a mammographic mass cad system. *Medical Physics*, 31(6) :1512–1520. (Cité en page 53.)

[Catarious Jr et al., 2006] Catarious Jr, D. M., Baydush, A. H., and Floyd Jr, C. E. (2006). Characterization of difference of gaussian filters in the detection of mammographic regions. *Medical Physics*, 33(11) :4104–4114. (Cité en page 53.)

[Champagnat et al., 1998] Champagnat, F., Idier, J., and Goussard, Y. (1998). Stationary markov random fields on a finite rectangular lattice. *IEEE Transactions on Information Theory*, 44(7) :2901–2916. (Cité en page 60.)

[Chan et al., 1987] Chan, H.-P., Vyborny, C. J., MacMAHON, H., Metz, C. E., Doi, K., and Sickles, E. A. (1987). Digital mammography : Roc studies of the effects of pixel size and unsharp-mask filtering on the detection of subtle microcalcifications. *Investigative Radiology*, 22(7) :581–589. (Cité en pages 32 et 33.)

[Chang et al., 1996a] Chang, Y.-H., Zheng, B., and Gur, D. (1996a). Computerized identification of suspicious regions for masses in digitized mammograms. *Investigative Radiology*, 31(3) :146–153. (Cité en page 54.)

[Chang et al., 1996b] Chang, Y.-H., Zheng, B., and Gur, D. (1996b). Robustness of computerized identification of masses in digitized mammograms : A preliminary assessment. *Investigative Radiology*, 31(9) :563–568. (Cité en page 54.)

[Chen et Lee, 1997] Chen, C. H. and Lee, G. G. (1997). On digital mammogram segmentation and microcalcification detection using multiresolution wavelet analysis. *Graphical Models and Image Processing*, 59(5) :349–364. (Cité en page 52.)

[Chen et Pavlidis, 1979] Chen, P. C. and Pavlidis, T. (1979). Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm. *Computer graphics and image processing*, 10(2) :172–182. (Cité en page 49.)

[Chen et al., 1989] Chen, S., Billings, S. A., and Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of control*, 50(5) :1873–1896. (Cité en page 35.)

[Chen et al., 2013] Chen, X., Shrivastava, A., and Gupta, A. (2013). Neil : Extracting visual knowledge from web data. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1409–1416. (Cité en page 115.)

[Chu et Krzyżak, 2014] Chu, J. L. and Krzyżak, A. (2014). Analysis of feature maps selection in supervised learning using convolutional neural networks. In *Canadian Conference on Artificial Intelligence*, pages 59–70. Springer. (Cité en page 114.)

[Ciatto et al., 2003] Ciatto, S., Del Turco, M. R., Risso, G., Catarzi, S., Bonardi, R., Viterbo, V., Gnutti, P., Guglielmoni, B., Pinelli, L., Pandiscia, A., et al. (2003). Comparison of standard reading and computer aided detection (cad) on a national proficiency test of screening mammography. *European Journal of Radiology*, 45(2) :135–138. (Cité en page 23.)

[Comer et al., 1996] Comer, M. L., Liu, S., and Delp, E. J. (1996). Statistical segmentation of mammograms. In *Proceedings of the 3rd International Workshop on Digital Mammography*, pages 475–478. (Cité en pages 57 et 58.)

[Constantinidis et al., 2001] Constantinidis, A., Fairhurst, M. C., and Rahman, A. F. R. (2001). A new multi-expert decision combination algorithm and its application to the detection of circumscribed masses in digital mammograms. *Pattern Recognition*, 34(8) :1527–1537. (Cité en page 54.)

[Copeland et al., 1995] Copeland, A. C., Ravichandran, G., and Trivedi, M. M. (1995). Localized radon transform-based detection of ship wakes in sar images. *IEEE Transactions on Geoscience and Remote Sensing*, 33(1) :35–45. (Cité en page 78.)

[Daubechies, 1992] Daubechies, I. (1992). *Ten lectures on wavelets*, volume 61. SIAM. (Cité en pages 40 et 129.)

[Dempster et al., 1977] Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 39(1) :1–38. (Cité en page 57.)

[Deriche, 1987] Deriche, R. (1987). Using canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1(2) :167–187. (Cité en page 51.)

[Desolneux et al., 2000] Desolneux, A., Moisan, L., and Morel, J.-M. (2000). Meaningful alignments. *International Journal of Computer Vision*, 40(1) :7–23. (Cité en pages 73, 77 et 81.)

[Devijver et Dekesel, 1988] Devijver, P. A. and Dekesel, M. (1988). Champs aléatoires de pickard et modélisation d'images digitales. *Traitement du Signal*, 5(5) :131–150. (Cité en pages 60, 61, 62 et 64.)

[Dhungel et al., 2015a] Dhungel, N., Carneiro, G., and Bradley, A. P. (2015a). Automated mass detection in mammograms using cascaded deep learning and random forests. In *Digital Image Computing : Techniques and Applications (DICTA), 2015 International Conference on*, pages 1–8. IEEE. (Cité en page 55.)

[Dhungel et al., 2015b] Dhungel, N., Carneiro, G., and Bradley, A. P. (2015b). Deep learning and structured prediction for the segmentation of mass in mammograms. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 605–612. Springer. (Cité en page 54.)

[Dhungel et al., 2015c] Dhungel, N., Carneiro, G., and Bradley, A. P. (2015c). Deep structured learning for mass segmentation from mammograms. In *Image Processing (ICIP), 2015 IEEE International Conference on*, pages 2950–2954. IEEE. (Cité en page 54.)

[Dhungel et al., 2015d] Dhungel, N., Carneiro, G., and Bradley, A. P. (2015d). Tree re-weighted belief propagation using deep learning potentials for mass segmentation from mammograms. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 760–763. IEEE. (Cité en page 55.)

[Diahi et al., 1996] Diahi, J., Frouge, C., Giron, A., and Fertil, B. (1996). Artificial neural networks for detection of breast cancer in mammography. In *Proceedings of the 3rd International Workshop on Digital Mammography, Chicago, USA*, pages 9–12. (Cité en page 54.)

[Divvala et al., 2014] Divvala, S. K., Farhadi, A., and Guestrin, C. (2014). Learning everything about anything : Webly-supervised visual concept learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3270–3277. (Cité en page 115.)

[Dobbins III et Godfrey, 2003] Dobbins III, J. T. and Godfrey, D. J. (2003). Digital x-ray tomosynthesis : current state of the art and clinical potential. *Physics in Medicine and Biology*, 48(19) :R65. (Cité en page 17.)

[Domınguez et Nandi, 2007] Domınguez, A. R. and Nandi, A. (2007). Detection of masses in mammograms using enhanced multilevel-thresholding segmentation and region selection based on rank. In *IASTED International Conference on Biomedical Engineering. Austria : Innsbruck*, volume 1, pages 370–375. (Cité en page 26.)

[Dosovitskiy et al., 2013] Dosovitskiy, A., Springenberg, J. T., and Brox, T. (2013). Unsupervised feature learning by augmenting single images. *arXiv preprint arXiv :1312.5242*. (Cité en page 103.)

[Dubes et al., 1990] Dubes, R., Jain, A., Nadabar, S., and Chen, C. (1990). Mrf model-based algorithms for image segmentation. In *Proceedings of 10th International Conference on Pattern Recognition, 1990*, volume 1, pages 808–814. (Cité en page 57.)

[Duda et al., 2012] Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons. (Cité en pages 65, 123 et 131.)

[Eccles et al., 2013] Eccles, S. A., Aboagye, E. O., Ali, S., Anderson, A. S., Armes, J., Berditchevski, F., Blaydes, J. P., Brennan, K., Brown, N. J., Bryant, H. E., et al. (2013). Critical research gaps and translational priorities for the successful prevention and treatment of breast cancer. *Breast Cancer Research*, 15(5) :1. (Cité en page 20.)

[Elter et Horsch, 2009] Elter, M. and Horsch, A. (2009). Cadx of mammographic masses and clustered microcalcifications : a review. *Medical Physics*, 36(6) :2052–2068. (Cité en page 23.)

[Fauci et al., 2005] Fauci, F., Raso, G., Magro, R., Forni, G., Lauria, A., Bagnasco, S., Cerello, P., Cheran, S. C., Torres, E. L., Bellotti, R., et al. (2005). A massive lesion detection algorithm in mammography. *Physica Medica*, 21(1) :23–30. (Cité en page 51.)

[Felzenszwalb et Huttenlocher, 2004] Felzenszwalb, P. F. and Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2) :167–181. (Cité en page 66.)

[Freixenet et al., 2002] Freixenet, J., Muñoz, X., Raba, D., Martí, J., and Cufí, X. (2002). Yet another survey on image segmentation : Region and boundary information integration. In *European Conference on Computer Vision*, pages 408–422. Springer. (Cité en page 65.)

[Freixenet et al., 2008] Freixenet, J., Oliver, A., Martí, R., Lladó, X., Pont, J., Pérez, E., Denton, E. R., and Zwiggelaar, R. (2008). Eigendetection of masses considering false positive reduction and breast density information. *Medical Physics*, 35(5) :1840–1853. (Cité en page 54.)

[Friedman et al., 2000] Friedman, J., Hastie, T., Tibshirani, R., et al. (2000). Additive logistic regression : a statistical view of boosting (with discussion and a rejoinder by the authors). *The Annals of Statistics*, 28(2) :337–407. (Cité en page 91.)

[Fu et Mui, 1981] Fu, K.-S. and Mui, J. (1981). A survey on image segmentation. *Pattern recognition*, 13(1) :3–16. (Cité en page 52.)

[Gal et al., 1984] Gal, M. L., Chavanne, G., and Pellier, D. (1984). Valeur diagnostique des microcalcifications groupées découvertes par mammographie. *Bull cancer*, 71 :57–64. (Cité en page 21.)

[Galloway, 1975] Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4(2) :172–179. (Cité en pages 88 et 91.)

[Gao et Bischof, 2009] Gao, R. and Bischof, W. F. (2009). Detection of linear structures in remote-sensed images. In *International Conference Image Analysis and Recognition*, pages 896–905. Springer. (Cité en pages 73 et 77.)

[Geman et al., 1987] Geman, D., Geman, S., and Graffigne, C. (1987). Locating texture and object boundaries. In *Pattern Recognition Theory and Applications*, pages 165–177. Springer. (Cité en page 59.)

[Geman et Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6) :721–741. (Cité en page 57.)

[Gennaro et al., 2008] Gennaro, G., Baldan, E., Bezzon, E., La Grassa, M., Pescarini, L., and di Maggio, C. (2008). Clinical performance of digital breast tomosynthesis versus full-field digital mammography : Preliminary results. In *Digital Mammography*, pages 477–482. Springer. (Cité en page 17.)

[Girshick et al., 2014] Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587. (Cité en pages 114 et 116.)

[Glorot et Bengio, 2010] Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Aistats*, volume 9, pages 249–256. (Cité en page 100.)

[Glorot et al., 2011] Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Aistats*, volume 15, page 275. (Cité en page 98.)

[Godfrey et al., 1980] Godfrey, R., Muir, F., and Rocca, F. (1980). Modeling seismic impedance with markov chains. *Geophysics*, 45(9) :1351–1372. (Cité en page 62.)

[Gomez et al., 2010] Gomez, W., Leija, L., Alvarenga, A., Infantosi, A., and Pereira, W. (2010). Computerized lesion segmentation of breast ultrasound based on marker-controlled watershed transformation. *Medical physics*, 37(1) :82–95. (Cité en page 50.)

[Gordon et Rangayyan, 1984] Gordon, R. and Rangayyan, R. M. (1984). Feature enhancement of film mammograms using fixed and adaptive neighborhoods. *Applied optics*, 23(4) :560–564. (Cité en page 33.)

[Goto et al., 1998] Goto, M., Morikawa, A., Fujita, H., Hara, T., and Endo, T. (1998). Detection of spicules on mammograms based on a multistage pendulum filter. In *Digital Mammography*, pages 135–138. Springer. (Cité en page 53.)

[Goudail et Réfrégier, 2004] Goudail, F. and Réfrégier, P. (2004). Contrast definition for optical coherent polarimetric images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(7) :947–951. (Cité en page 58.)

[Grady, 2006] Grady, L. (2006). Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11) :1768–1783. (Cité en page 58.)

[Greig et al., 1989] Greig, D. M., Porteous, B. T., and Seheult, A. H. (1989). Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 271–279. (Cité en page 57.)

[Gu et al., 2012] Gu, Q., Li, Z., and Han, J. (2012). Generalized fisher score for feature selection. *arXiv preprint arXiv :1202.3725*. (Cité en page 131.)

[Guliato et al., 1998] Guliato, D., Rangayyan, R. M., Carnielli, W. A., Zuffo, J. A., and Desautels, J. L. (1998). Segmentation of breast tumors in mammograms by fuzzy region growing. In *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, volume 2, pages 1002–1005. IEEE. (Cité en page 49.)

[Guliato et al., 2008] Guliato, D., Rangayyan, R. M., Carvalho, J. D., and Santiago, S. A. (2008). Polygonal modeling of contours of breast tumors with the preservation of spicules. *IEEE Transactions on Biomedical Engineering*, 55(1) :14–20. (Cité en page 91.)

[Gulsrud et al., 2006] Gulsrud, T. O., Engan, K., and Hanstveit, T. (2006). Watershed segmentation of detected masses in digital mammograms. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 3304–3307. IEEE. (Cité en page 50.)

[Guyon et Elisseeff, 2003] Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3(Mar) :1157–1182. (Cité en page 131.)

[Hansen et Salamon, 1990] Hansen, L. K. and Salamon, P. (1990). Neural network ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12 :993–1001. (Cité en page 101.)

[Haralick et al., 1973] Haralick, R. M., Shanmugam, K., et al. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, (6) :610–621. (Cité en page 88.)

[Hatanaka et al., 2001] Hatanaka, Y., Hara, T., Fujita, H., Kasai, S., Endo, T., and Iwase, T. (2001). Development of an automated method for detecting mammographic masses with a partial loss of region. *IEEE Transactions on Medical Imaging*, 20(12) :1209–1214. (Cité en page 54.)

[He et al., 2015] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9) :1904–1916. (Cité en pages 103 et 116.)

[Heath et al., 1998] Heath, M., Bowyer, K., Kopans, D., Kegelmeyer Jr, P., Moore, R., Chang, K., and Munishkumaran, S. (1998). Current status of the digital database for screening mammography. In *Digital mammography*, pages 457–460. Springer. (Cité en page 26.)

[Heath et Bowyer, 2000] Heath, M. D. and Bowyer, K. W. (2000). Mass detection by relative image intensity. In *Proceedings of the 5th International Workshop on Digital Mammography (IWDM-2000)*, pages 219–225. (Cité en page 53.)

[Hejazi et Ho, 2005] Hejazi, M. R. and Ho, Y.-S. (2005). Automated detection of tumors in mammograms using two segments for classification. In *Advances in Multimedia Information Processing-PCM 2005*, pages 910–921. Springer. (Cité en page 49.)

[Heron, 2009] Heron, J. (2009). La mammographie. En Ligne. page consultée le 16 septembre 2016. (Cité en pages vii et 15.)

[Hinton et Salakhutdinov, 2006] Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786) :504–507. (Cité en pages 55 et 100.)

[Hinton et al., 2012] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv :1207.0580*. (Cité en page 101.)

[Hofmann, 2001] Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, 42(1-2) :177–196. (Cité en page 93.)

[Hong et Brady, 2003] Hong, B.-W. and Brady, M. (2003). A topographic representation for mammogram segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 730–737. Springer. (Cité en page 51.)

[Howard, 2013] Howard, A. G. (2013). Some improvements on deep convolutional neural network based image classification. *arXiv preprint arXiv :1312.5402*. (Cité en page 103.)

[Hu, 1962] Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8(2) :179–187. (Cité en pages 88 et 123.)

[Huang et Chen, 2004] Huang, Y.-L. and Chen, D.-R. (2004). Watershed segmentation for breast tumor in 2-d sonography. *Ultrasound in medicine & biology*, 30(5) :625–632. (Cité en page 50.)

[Hummel et Zucker, 1983] Hummel, R. A. and Zucker, S. W. (1983). On the foundations of relaxation labeling processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (3) :267–287. (Cité en page 57.)

[Huo et al., 1995] Huo, Z., Giger, M. L., Vyborny, C. J., Bick, U., Lu, P., Wolverton, D. E., and Schmidt, R. A. (1995). Analysis of spiculation in the computerized classification of mammographic masses. *Medical Physics*, 22(10) :1569–1579. (Cité en page 48.)

[Huo et al., 1998] Huo, Z., Giger, M. L., Vyborny, C. J., Wolverton, D. E., Schmidt, R. A., and Doi, K. (1998). Automated computerized classification of malignant and benign masses on digitized mammograms. *Academic Radiology*, 5(3) :155–168. (Cité en page 90.)

[IARC, 2013] IARC (2013). Globocan 2012 :estimatied cancer incidence, mortality and prevalence worldwide in 2012. En Ligne. page consultée le 12 février 2015. (Cité en page 9.)

[Idier et Goussard, 1995] Idier, J. and Goussard, Y. (1995). Formules de réestimation pour un modèle de chaîne de markov cachée stationnaire réversible. In *15 Colloque sur le traitement du signal et des images, FRA, 1995*. GRETSI, Groupe dâEtudes du Traitement du Signal et des Images. (Cité en page 64.)

[Ioffe et Szegedy, 2015] Ioffe, S. and Szegedy, C. (2015). Batch normalization : Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv :1502.03167*. (Cité en pages 114 et 115.)

[Jain et Farrokhnia, 1991] Jain, A. K. and Farrokhnia, F. (1991). Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24(12) :1167–1186. (Cité en page 129.)

[Jain et al., 1999] Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering : a review. *ACM computing surveys (CSUR)*, 31(3) :264–323. (Cité en page 51.)

[Jiang et al., 1998] Jiang, H., Tiu, W., Yamamoto, S., and Iisaku, S. (1998). A method for automatic detection of spicules in mammograms. *J. Comput. Aided Diagn. Med. Images*, 2(4) :23–31. (Cité en page 76.)

[John et Ewen, 1989] John, V. and Ewen, K. (1989). Ct scanning of the breast in problem cases. *Strahlentherapie und Onkologie : Organ der Deutschen Rontgengesellschaft...[et al]*, 165(9) :657–662. (Cité en page 16.)

[Jones et al., 2014] Jones, N. et al. (2014). The learning machines. *Nature*, 505(7482) :146–148. (Cité en pages viii et 97.)

[Kachouri et al., 2012] Kachouri, I. C., Djemal, K., and Maaref, H. (2012). Characterisation of mammographic masses using a new spiculated mass descriptor in computer aided diagnosis systems. *International Journal of Signal and Imaging Systems Engineering*, 5(2) :132–142. (Cité en pages 91, 106, 107, 111 et 112.)

[Karssemeijer, 1993] Karssemeijer, N. (1993). Adaptive noise equalization and recognition of microcalcification clusters in mammograms. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(06) :1357–1376. (Cité en page 33.)

[Karssemeijer, 1994] Karssemeijer, N. (1994). *Recognition of stellate lesions in digital mammograms*. Elsevier Science BV. (Cité en page 76.)

[Karssemeijer, 1998] Karssemeijer, N. (1998). Automated classification of parenchymal patterns in mammograms. *Physics in medicine and biology*, 43(2) :365. (Cité en page 92.)

[Karssemeijer, 1999] Karssemeijer, N. (1999). Local orientation distribution as a function of spatial scale for detection of masses in mammograms. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 280–293. Springer. (Cité en page 76.)

[Karssemeijer et te Brake, 1996] Karssemeijer, N. and te Brake, G. M. (1996). Detection of stellate distortions in mammograms. *IEEE Transactions on Medical Imaging*, 15(5) :611–619. (Cité en pages 72, 76 et 82.)

[Kegelmeyer Jr, 1992] Kegelmeyer Jr, W. P. (1992). Computer detection of stellate lesions in mammograms. In *SPIE/IS&T 1992 Symposium on Electronic Imaging : Science and Technology*, pages 446–454. International Society for Optics and Photonics. (Cité en page 75.)

[Kegelmeyer Jr et al., 1994] Kegelmeyer Jr, W. P., Pruneda, J. M., Bourland, P. D., Hillis, A., Riggs, M. W., and Nipper, M. L. (1994). Computer-aided mammographic screening for spiculated lesions. *Radiology*, 191(2) :331–337. (Cité en page 75.)

[Kilday et al., 1993] Kilday, J., Palmieri, F., and Fox, M. D. (1993). Classifying mammographic lesions using computerized image analysis. *IEEE Transactions on Medical Imaging*, 12(4) :664–669. (Cité en pages 111 et 125.)

[Kim et Kim, 2005] Kim, H. and Kim, W. (2005). Automatic detection of spiculated masses using fractal analysis in digital mammography. In *International Conference on Computer Analysis of Images and Patterns*, pages 256–263. Springer. (Cité en page 74.)

[Kim et Kim, 2000] Kim, H.-K. and Kim, J.-D. (2000). Region-based shape descriptor invariant to rotation, scale and translation. *Signal Processing : Image Communication*, 16(1) :87–93. (Cité en page 124.)

[Kinoshita et al., 2007] Kinoshita, S. K., de Azevedo-Marques, P. M., Pereira Jr, R. R., Rodrigues, J. A. H., and Rangayyan, R. M. (2007). Content-based retrieval of mammograms using visual features related to breast density patterns. *Journal of Digital Imaging*, 20(2) :172–190. (Cité en page 94.)

[Kirby et Sirovich, 1990] Kirby, M. and Sirovich, L. (1990). Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1) :103–108. (Cité en pages 34 et 37.)

[Kittler et al., 1998] Kittler, J., Hatef, M., Duin, R. P., and Matas, J. (1998). On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3) :226–239. (Cité en pages 93 et 113.)

[Kobatake et Murakami, 1996] Kobatake, H. and Murakami, M. (1996). Adaptive filter to detect rounded convex regions : Iris filter. In *Proceedings of the 13th International Conference on Pattern Recognition, 1996*, volume 2, pages 340–344. IEEE. (Cité en page 53.)

[Kobatake et al., 1999] Kobatake, H., Murakami, M., Takeo, H., and Nawano, S. (1999). Computerized detection of malignant tumors on digital mammograms. *IEEE Transactions on Medical Imaging*, 18(5) :369–378. (Cité en page 51.)

[Kobatake et Yoshinaga, 1996] Kobatake, H. and Yoshinaga, Y. (1996). Detection of spicules on mammogram based on skeleton analysis. *IEEE Transactions on Medical Imaging*, 15(3) :235–245. (Cité en pages 73 et 76.)

[Kom et al., 2007] Kom, G., Tiedeu, A., and Kom, M. (2007). Automated detection of masses in mammograms by local adaptive thresholding. *Computers in Biology and Medicine*, 37(1) :37–48. (Cité en page 53.)

[Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105. (Cité en pages 95, 96, 100, 103, 104 et 114.)

[Krupinski et Nishikawa, 1997] Krupinski, E. A. and Nishikawa, R. M. (1997). Comparison of eye position versus computer identified microcalcification clusters on mammograms. *Medical Physics*, 24(1) :17–23. (Cité en page 24.)

[Kudo et Sklansky, 2000] Kudo, M. and Sklansky, J. (2000). Comparison of algorithms that select features for pattern classifiers. *Pattern Recognition*, 33(1) :25–41. (Cité en page 132.)

[Kupinski et Giger, 1998] Kupinski, M. and Giger, M. (1998). Automated seeded lesion segmentation on digital mammograms. *IEEE Transactions on Medical Imaging*, 17(4) :510–517. (Cité en pages 48, 51, 66, 68 et 70.)

[Lacoste et al., 2005] Lacoste, C., Descombes, X., and Zerubia, J. (2005). Point processes for unsupervised line network extraction in remote sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10) :1568–1579. (Cité en pages 73 et 77.)

[Lai et al., 1989] Lai, S.-M., Li, X., and Biscof, W. (1989). On techniques for detecting circumscribed masses in mammograms. *IEEE Transactions on Medical Imaging*, 8(4) :377–386. (Cité en page 54.)

[Laine et al., 1994] Laine, A. F., Schuler, S., Fan, J., and Huda, W. (1994). Mammographic feature enhancement by multiscale analysis. *IEEE Transactions on Medical Imaging*, 13(4) :725–740. (Cité en page 33.)

[Lanyi, 1985] Lanyi, M. (1985). Morphologic analysis of microcalcifications. In *Early Breast Cancer*, pages 113–135. Springer. (Cité en page 21.)

[Lau et al., 2008] Lau, B. A., Reiser, I. S., and Nishikawa, R. M. (2008). Microcalcification detectability in tomosynthesis. In *Medical Imaging*, pages 69134L–69134L. International Society for Optics and Photonics. (Cité en page 18.)

[Laws, 1980] Laws, K. I. (1980). Textured image segmentation. Technical report, DTIC Document. (Cité en page 92.)

[LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553) :436–444. (Cité en page 100.)

[LeCun et al., 1998] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324. (Cité en pages 96 et 97.)

[LeCun et al., 2012] LeCun, Y. A., Bottou, L., Orr, G. B., and Müller, K.-R. (2012). Efficient backprop. In *Neural networks : Tricks of the trade*, pages 9–48. Springer. (Cité en pages 98 et 100.)

[Leung et Malik, 2001] Leung, T. and Malik, J. (2001). Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, 43(1) :29–44. (Cité en page 93.)

[Li et al., 1995] Li, H., Kallergi, M., Clarke, L., Jain, V., and Clark, R. (1995). Markov random field for tumor detection in digital mammography. *IEEE Transactions on Medical Imaging*, 14(3) :565–576. (Cité en pages 57 et 70.)

[Li et al., 1996] Li, H., Liu, K., Wang, Y., and Lo, S. (1996). Morphological filtering and stochastic modeling-based segmentation of masses on mammographic images. In *Nuclear Science Symposium, 1996. Conference Record., 1996 IEEE*, volume 3, pages 1792–1796. IEEE. (Cité en page 57.)

[Li et al., 1997a] Li, H., Liu, K., Wang, Y., and Lo, S.-C. (1997a). Nonlinear filtering enhancement and histogram modeling segmentation of masses for digital mammograms. In *Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE*, volume 3, pages 1045–1046. IEEE. (Cité en page 57.)

[Li et al., 2001] Li, H., Wang, Y., Liu, K., Lo, S.-C. B., and Freedman, M. T. (2001). Computerized radiographic mass detection. i. lesion site selection by morphological enhancement and contextual segmentation. *IEEE Transactions on Medical Imaging*, 20(4) :289–301. (Cité en page 57.)

[Li et al., 2014] Li, H., Zhao, R., and Wang, X. (2014). Highly efficient forward and backward propagation of convolutional neural networks for pixelwise classification. *arXiv preprint arXiv :1412.4526*. (Cité en page 116.)

[Li et al., 2002] Li, L., Clark, R. A., and Thomas, J. A. (2002). Computer-aided diagnosis of masses with full-field digital mammography. *Academic Radiology*, 9(1) :4–12. (Cité en page 52.)

[Li et al., 1997b] Li, L., Qian, W., and Clarke, L. P. (1997b). Digital mammography : computer-assisted diagnosis method for mass detection with multiorientation and multiresolution wavelet transforms. *Academic radiology*, 4(11) :724–731. (Cité en page 52.)

[Li et al., 1999] Li, L., Qian, W., Clarke, L. P., Clark, R. A., and Thomas, J. A. (1999). Improving mass detection by adaptive and multiscale processing in digitized mammograms. In *Medical Imaging'99*, pages 490–498. International Society for Optics and Photonics. (Cité en page 52.)

[Liberman et al., 1998] Liberman, L., Abramson, A., Squires, F., Glassman, J., Morris, E., and Dershaw, D. (1998). The breast imaging reporting and data system : positive predictive value of mammographic features and final assessment categories. *American Journal of Roentgenology*, 171(1) :35–40. (Cité en page 69.)

[Liu et al., 2014] Liu, J., Zhang, S., Liu, W., Zhang, X., and Metaxas, D. N. (2014). Scalable mammogram retrieval using anchor graph hashing. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 898–901. IEEE. (Cité en page 95.)

[Liu et al., 2001] Liu, S., Babbs, C. F., and Delp, E. J. (2001). Multiresolution detection of spiculated lesions in digital mammograms. *IEEE Transactions on Image Processing*, 10(6) :874–884. (Cité en page 75.)

[Liu et al., 2011] Liu, W., Wang, J., Kumar, S., and Chang, S.-F. (2011). Hashing with graphs. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 1–8. (Cité en page 95.)

[Liu et al., 2015] Liu, Y., Guo, Y., Wu, S., and Lew, M. S. (2015). Deepindex for accurate and efficient image retrieval. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, pages 43–50. ACM. (Cité en page 116.)

[Livens et al., 1997] Livens, S., Scheunders, P., Van de Wouwer, G., and Van Dyck, D. (1997). Wavelets for texture analysis, an overview. In *Image Processing and Its Applications, 1997., Sixth International Conference on*, volume 2, pages 581–585. IET. (Cité en page 88.)

[MacQueen et al., 1967] MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA. (Cité en page 52.)

[Marill et Green, 1963] Marill, T. and Green, D. (1963). On the effectiveness of receptors in recognition systems. *IEEE Transactions on Information Theory*, 9(1) :11–17. (Cité en page 132.)

[Marr, 1976] Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London B : Biological Sciences*, 275(942) :483–519. (Cité en page 51.)

[Marr et Hildreth, 1980] Marr, D. and Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B : Biological Sciences*, 207(1167) :187–217. (Cité en page 51.)

[Marroquin et al., 1987] Marroquin, J., Mitter, S., and Poggio, T. (1987). Probabilistic solution of ill-posed problems in computational vision. *Journal of the American Statistical Association*, 82(397) :76–89. (Cité en page 58.)

[Martí et al., 2003] Martí, J., Freixenet, J., Munoz, X., and Oliver, A. (2003). Active region segmentation of mammographic masses based on texture, contour and shape features. In *Pattern Recognition and Image Analysis*, pages 478–485. Springer. (Cité en page 49.)

[Martin et al., 2001] Martin, D., Fowlkes, C., Tal, D., and Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 416–423. IEEE. (Cité en page 65.)

[Matheus et Schiabel, 2011] Matheus, B. R. N. and Schiabel, H. (2011). Online mammographic images database for development and comparison of cad schemes. *Journal of digital imaging*, 24(3) :500–506. (Cité en page 26.)

[Matsubara et al., 1998] Matsubara, T., Fujita, H., Hara, T., Kasai, S., Otsuka, O., Hatanaka, Y., and Endo, T. (1998). Development of a new algorithm for detection of mammographic masses. In *Digital Mammography*, pages 139–142. Springer. (Cité en page 53.)

[Matsubara et al., 1997] Matsubara, T., Fujita, H., Kasai, S., Goto, M., Tani, Y., Hara, T., and Endo, T. (1997). Development of new schemes for detection and analysis of mammographic masses. In *Intelligent Information Systems, 1997. IIS'97. Proceedings*, pages 63–66. IEEE. (Cité en page 53.)

[McLachlan, 2004] McLachlan, G. (2004). *Discriminant analysis and statistical pattern recognition*, volume 544. John Wiley & Sons. (Cité en page 92.)

[McLoughlin et al., 2004] McLoughlin, K. J., Bones, P. J., and Karssemeijer, N. (2004). Noise equalization for detection of microcalcification clusters in direct digital mammogram images. *IEEE transactions on medical imaging*, 23(3) :313–320. (Cité en page 33.)

[Meila, 2005] Meila, M. (2005). Comparing clusterings : an axiomatic view. In *Proceedings of the 22nd International Conference on Machine Learning*, pages 577–584. ACM. (Cité en page 65.)

[Mencattini et al., 2008] Mencattini, A., Salmeri, M., Lojacono, R., Frigerio, M., and Caselli, F. (2008). Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing. *IEEE Transactions on Instrumentation and Measurement*, 57(7) :1422–1430. (Cité en page 33.)

[Metz, 1978] Metz, C. E. (1978). Basic principles of roc analysis. In *Seminars in nuclear medicine*, volume 8, pages 283–298. Elsevier. (Cité en page 89.)

[Meyer et al., 1990] Meyer, J. E., Eberlein, T. J., Stomper, P. C., and Sonnenfeld, M. R. (1990). Biopsy of occult breast lesions : analysis of 1261 abnormalities. *Jama*, 263(17) :2341–2343. (Cité en page 24.)

[Mignotte, 2010] Mignotte, M. (2010). A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation. *IEEE Transactions on Image Processing*, 19(6) :1610–1624. (Cité en page 67.)

[Mihcak et al., 1999] Mihcak, M. K., Kozintsev, I., Ramchandran, K., and Moulin, P. (1999). Low-complexity image denoising based on statistical modeling of wavelet coefficients. *IEEE Signal Processing Letters*, 6(12) :300–303. (Cité en page 33.)

[Miller et Astley, 1992] Miller, P. and Astley, S. (1992). Classification of breast tissue by texture analysis. *Image and Vision Computing*, 10(5) :277–282. (Cité en page 92.)

[Monnin et al., 2007] Monnin, P., Gutierrez, D., Bulling, S., Guntern, D., and Verdun, F. (2007). A comparison of the performance of digital mammography systems. *Medical physics*, 34(3) :906–914. (Cité en page 39.)

[Moreira et al., 2012] Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., and Cardoso, J. S. (2012). Inbreast : toward a full-field digital mammographic database. *Academic radiology*, 19(2) :236–248. (Cité en pages 27 et 64.)

[Muralidhar et al., 2010] Muralidhar, G. S., Bovik, A. C., Giese, J. D., Sampat, M. P., Whitman, G. J., Haygood, T. M., Stephens, T. W., and Markey, M. K. (2010). Snakules : A model-based active contour algorithm for the annotation of spicules on mammography. *IEEE Transactions on Medical Imaging*, 29(10) :1768–1780. (Cité en pages 73, 77, 78 et 82.)

[Nakagawa et al., 2004] Nakagawa, T., Hara, T., Fujita, H., Iwase, T., Endo, T., and Horita, K. (2004). Automated contour extraction of mammographic mass shadow using an improved active contour model. In *International Congress Series*, volume 1268, pages 882–885. Elsevier. (Cité en page 51.)

[Narvaez et al., 2011] Narvaez, F., Diaz, G., and Romero, E. (2011). Multi-view information fusion for automatic bi-rads description of mammographic masses. In *SPIE Medical Imaging*, pages 79630A–79630A. International Society for Optics and Photonics. (Cité en page 94.)

[Narváez et Romero, 2012] Narváez, F. and Romero, E. (2012). Breast mass classification using orthogonal moments. In *International Workshop on Digital Mammography*, pages 64–71. Springer. (Cité en page 91.)

[Ng et Bischof, 1992] Ng, S. L. and Bischof, W. F. (1992). Automated detection and classification of breast tumors. *Computers and Biomedical Research*, 25(3) :218–237. (Cité en page 74.)

[Nguyen et al., 2015] Nguyen, A., Yosinski, J., and Clune, J. (2015). Deep neural networks are easily fooled : High confidence predictions for unrecognizable images. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 427–436. IEEE. (Cité en page 115.)

[Nishikawa, 1998] Nishikawa, R. M. (1998). Mammographic databases. *Breast Disease*, 10(3, 4) :137–150. (Cité en page 25.)

[Oliver et al., 2010a] Oliver, A., Freixenet, J., Marti, J., Pérez, E., Pont, J., Denton, E. R., and Zwiggelaar, R. (2010a). A review of automatic mass detection and segmentation in mammographic images. *Medical Image Analysis*, 14(2) :87–110. (Cité en page 26.)

[Oliver et al., 2008] Oliver, A., Freixenet, J., Marti, R., Pont, J., Pérez, E., Denton, E. R., and Zwiggelaar, R. (2008). A novel breast tissue density classification methodology. *IEEE Transactions on Information Technology in Biomedicine*, 12(1) :55–65. (Cité en page 93.)

[Oliver et al., 2005] Oliver, A., Freixenet, J., and Zwiggelaar, R. (2005). Automatic classification of breast density. In *IEEE International Conference on Image Processing 2005*, volume 2, pages II–1258. IEEE. (Cité en pages 92 et 93.)

[Oliver et al., 2010b] Oliver, A., Lladó, X., Pérez, E., Pont, J., Denton, E. R., Freixenet, J., and Martí, J. (2010b). A statistical approach for breast density segmentation. *Journal of Digital Imaging*, 23(5) :527–537. (Cité en page 26.)

[Oliver et al., 2010c] Oliver, A., Torrent, A., Llado, X., and Marti, J. (2010c). Automatic diagnosis of masses by using level set segmentation and shape description. In *2010 20th International Conference on Pattern Recognition (ICPR)*, pages 2528–2531. IEEE. (Cité en page 91.)

[Otsu, 1975] Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285-296) :23–27. (Cité en page 53.)

[Ouyang et al., 2014] Ouyang, W., Luo, P., Zeng, X., Qiu, S., Tian, Y., Li, H., Yang, S., Wang, Z., Xiong, Y., Qian, C., et al. (2014). Deepid-net : multi-stage and deformable deep convolutional neural networks for object detection. *arXiv preprint arXiv :1409.3505*. (Cité en page 116.)

[Özekes et al., 2005] Özekes, S., Osman, O., and Camurcu, A. Y. (2005). Mammographic mass detection using a mass template. *Korean journal of radiology*, 6(4) :221–228. (Cité en page 53.)

[Palma et al., 2008] Palma, G., Bloch, I., and Muller, S. (2008). Fuzzy connected filters for fuzzy gray scale images. In *IPMU*, volume 8, pages 667–674. (Cité en page 75.)

[Palma et al., 2010] Palma, G., Bloch, I., and Muller, S. (2010). Spiculated lesions and architectural distortions detection in digital breast tomosynthesis datasets. In *International Workshop on Digital Mammography*, pages 712–719. Springer. (Cité en page 75.)

[Palma et al., 2014] Palma, G., Bloch, I., and Muller, S. (2014). Detection of masses and architectural distortions in digital breast tomosynthesis images using fuzzy and a contrario approaches. *Pattern Recognition*, 47(7) :2467–2480. (Cité en page 75.)

[Palma et al., 2009a] Palma, G., Muller, S., Bloch, I., and Iordache, R. (2009a). Convergence areas detection in digital breast tomosynthesis volumes using a contrario modeling. In *SPIE Symposium on Medical Imaging : Computer-Aided Diagnosis, Lake Buena Vista, FL, USA*. (Cité en pages 72, 73, 75, 77, 80 et 81.)

[Palma et al., 2009b] Palma, G., Muller, S., Bloch, I., and Iordache, R. (2009b). Fast detection of convergence areas in digital breast tomosynthesis. In *2009 IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, pages 847–850. IEEE. (Cité en page 75.)

[Pantofaru et Hebert, 2005] Pantofaru, C. and Hebert, M. (2005). A comparison of image segmentation algorithms. *Robotics Institute*, page 336. (Cité en page 65.)

[Pappas, 1992] Pappas, T. N. (1992). An adaptive clustering algorithm for image segmentation. *IEEE Transactions on Signal Processing*, 40(4) :901–914. (Cité en page 52.)

[Park et al., 2007] Park, S. C., Sukthankar, R., Mummert, L., Satyanarayanan, M., and Zheng, B. (2007). Optimization of reference library used in content-based medical image retrieval scheme. *Medical Physics*, 34(11) :4331–4339. (Cité en page 94.)

[Parr et al., 1994] Parr, T., Astley, S., Boggis, C., et al. (1994). The detection of stellate lesions in digital mammograms. In *International Workshop on Digital Mammography*, pages 231–239. (Cité en page 51.)

[Pereira et al., 2014] Pereira, D. C., Ramos, R. P., and do Nascimento, M. Z. (2014). Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. *Computer Methods and Programs in Biomedicine*, 114(1) :88–101. (Cité en pages 66, 69 et 70.)

[Petrick et al., 1996] Petrick, N., Chan, H.-P., Sahiner, B., and Wei, D. (1996). An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. *IEEE Transactions on Medical Imaging*, 15(1) :59–67. (Cité en page 51.)

[Petrick et al., 1995] Petrick, N., Chan, H.-P., Sahiner, B., Wei, D., Helvie, M. A., Goodsitt, M. M., and Adler, D. D. (1995). Automated detection of breast masses on digital mammograms using adaptive density-weighted contrast-enhancement filtering. In *Medical Imaging 1995*, pages 590–597. International Society for Optics and Photonics. (Cité en page 51.)

[Petrick et al., 2002] Petrick, N., Sahiner, B., Chan, H.-P., Helvie, M. A., Paquerault, S., and Hadjiiski, L. M. (2002). Breast cancer detection : Evaluation of a mass-detection algorithm for computer-aided diagnosisâexperience in 263 patients 1. *Radiology*, 224(1) :217–224. (Cité en page 49.)

[Petroudi et al., 2003] Petroudi, S., Kadir, T., and Brady, M. (2003). Automatic classification of mammographic parenchymal patterns : A statistical approach. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 1, pages 798–801. IEEE. (Cité en pages 93, 109 et 113.)

[Pickard, 1977] Pickard, D. (1977). A curious binary lattice process. *Journal of Applied Probability*, pages 717–731. (Cité en pages 59 et 60.)

[Pickard, 1980] Pickard, D. K. (1980). Unilateral markov fields. *Advances in Applied Probability*, pages 655–671. (Cité en pages 59 et 60.)

[Pizer et al., 1987] Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. B., and Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3) :355–368. (Cité en page 50.)

[Polakowski et al., 1997] Polakowski, W. E., Cournoyer, D. A., Rogers, S. K., DeSimio, M. P., Ruck, D. W., Hoffmeister, J. W., and Raines, R. A. (1997). Computer-aided breast cancer detection and diagnosis of masses using difference of gaussians and derivative-based feature saliency. *IEEE Transactions on Medical Imaging*, 16(6) :811–819. (Cité en page 53.)

[Porter et Canagarajah, 1997] Porter, R. and Canagarajah, N. (1997). Robust rotation-invariant texture classification : wavelet, gabor filter and gmrf based schemes. *IEE Proceedings-Vision, Image and Signal Processing*, 144(3) :180–188. (Cité en page 129.)

[Poulos et al., 2003] Poulos, A., McLean, D., Rickard, M., and Heard, R. (2003). Breast compression in mammography : how much is enough ? *Australasian radiology*, 47(2) :121–126. (Cité en page 39.)

[Prewitt, 1970] Prewitt, J. M. (1970). Object enhancement and extraction. *Picture Processing and Psychopictorics*, 10(1) :15–19. (Cité en page 50.)

[Qi et Snyder, 1998] Qi, H. and Snyder, W. E. (1998). Lesion detection and characterization in digital mammography by bézier histograms. In *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, volume 2, pages 1021–1024. IEEE. (Cité en page 49.)

[Qian et al., 1998a] Qian, W., Li, L., Clarke, L., Mao, F., Clark, R. A., and Thomas, C. J. (1998a). A computer assisted diagnostic system for mass detection. In *Digital Mammography*, pages 79–86. Springer. (Cité en page 52.)

[Qian et al., 1999] Qian, W., Li, L., and Clarke, L. P. (1999). Image feature extraction for mass detection in digital mammography : Influence of wavelet analysis. *Medical Physics*, 26(3) :402–408. (Cité en page 52.)

[Qian et al., 1998b] Qian, W., Li, L., Clarke, L. P., Mao, F., and Clark, R. A. (1998b). Adaptive cad modules for mass detection in digital mammography. In *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, volume 2, pages 1013–1016. IEEE. (Cité en page 52.)

[Qian et al., 2001] Qian, W., Sun, X., Song, D., and Clark, R. A. (2001). Digital mammography : wavelet transform and kalman-filtering neural network in mass segmentation and detection. *Academic Radiology*, 8(11) :1074–1082. (Cité en page 52.)

[Ram, 1982] Ram, G. (1982). Optimization of ionizing radiation usage in medical imaging by means of image enhancement techniques. *Medical Physics*, 9(5) :733–737. (Cité en page 32.)

[Rangayyan et al., 1997] Rangayyan, R. M., El-Faramawy, N. M., Desautels, J. L., and Alim, O. A. (1997). Measures of acutance and shape for classification of breast tumors. *IEEE Transactions on Medical Imaging*, 16(6) :799–810. (Cité en pages 49 et 90.)

[Raso et al., 2004] Raso, G., Magro, R., and Fauci, F. (2004). Mammogram segmentation by contour searching and massive lesion classification with neural network. In *IEEE nuclear science symposium conference record*. (Cité en page 51.)

[Reddi et al., 1984] Reddi, S., Rudin, S., and Keshavan, H. (1984). An optimal multiple threshold scheme for image segmentation. *IEEE Transactions on Systems, Man, and Cybernetics*, (4) :661–665. (Cité en page 53.)

[Ribeiro et al., 2013] Ribeiro, P. B., Romero, R. A. F., Oliveira, P. R., Schiabel, H., and Verçosa, L. B. (2013). Automatic segmentation of breast masses using enhanced ica mixture model. *Neurocomputing*, 120 :61–71. (Cité en pages 66, 68 et 70.)

[Rissanen, 1978] Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5) :465–471. (Cité en page 57.)

[Roberts, 1963] Roberts, L. G. (1963). *Machine perception of three-dimensional solids*. PhD thesis, Ph. D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA. (Cité en page 50.)

[Rogova et al., 1999] Rogova, G. L., Ke, C.-C., Acharya, R. S., and Stomper, P. C. (1999). Feature choice for detection of cancerous masses by constrained optimization. In *Medical Imaging'99*, pages 1440–1447. International Society for Optics and Photonics. (Cité en page 52.)

[Rose, 1998] Rose, K. (1998). Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE*, 86(11) :2210–2239. (Cité en page 52.)

[Rother et al., 2004] Rother, C., Kolmogorov, V., and Blake, A. (2004). Grabcut : Interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 309–314. ACM. (Cité en page 58.)

[Rubinstein et al., 2010] Rubinstein, R., Zibulevsky, M., and Elad, M. (2010). Double sparsity : Learning sparse dictionaries for sparse signal approximation. *IEEE Transactions on Signal Processing*, 58(3) :1553–1564. (Cité en page 35.)

[Russakovsky et al., 2015] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3) :211–252. (Cité en pages 96, 101 et 104.)

[Sahiner et al., 1998a] Sahiner, B., Chan, H.-P., Petrick, N., Helvie, M. A., and Goodsitt, M. M. (1998a). Computerized characterization of masses on mammograms : the rubber band straightening transform and texture analysis. *Medical Physics*, 25(4) :516–526. (Cité en page 52.)

[Sahiner et al., 1998b] Sahiner, B., Chan, H.-P., Petrick, N., Helvie, M. A., and Goodsitt, M. M. (1998b). Design of a high-sensitivity classifier based on a genetic algorithm : application to computer-aided diagnosis. *Physics in medicine and biology*, 43(10) :2853. (Cité en page 91.)

[Sahiner et al., 2001a] Sahiner, B., Chan, H.-P., Petrick, N., Helvie, M. A., and Hadjiiski, L. M. (2001a). Improvement of mammographic mass characterization using spiculation measures and morphological features. *Medical Physics*, 28(7) :1455–1465. (Cité en page 51.)

[Sahiner et al., 1996a] Sahiner, B., Chan, H.-P., Petrick, N., Wei, D., Helvie, M. A., Adler, D. D., and Goodsitt, M. M. (1996a). Classification of mass and normal breast tissue : a convolution

neural network classifier with spatial domain and texture images. *IEEE Transactions on Medical Imaging*, 15(5) :598–610. (Cité en page 92.)

[Sahiner et al., 1996b] Sahiner, B., Chan, H.-P., Wei, D., Petrick, N., Helvie, M. A., Adler, D. D., and Goodsitt, M. M. (1996b). Image feature selection by a genetic algorithm : Application to classification of mass and normal breast tissue. *Medical Physics*, 23(10) :1671–1684. (Cité en page 52.)

[Sahiner et al., 2001b] Sahiner, B., Petrick, N., Chan, H.-P., Hadjiiski, L. M., Paramagul, C., Helvie, M. A., and Gurcan, M. N. (2001b). Computer-aided characterization of mammographic masses : accuracy of mass segmentation and its effects on characterization. *IEEE Transactions on Medical Imaging*, 20(12) :1275–1284. (Cité en pages 51, 66, 68 et 70.)

[Salah et al., 2011] Salah, M. B., Mitiche, A., and Ayed, I. B. (2011). Multiregion image segmentation by parametric kernel graph cuts. *IEEE Transactions on Image Processing*, 20(2) :545–557. (Cité en page 67.)

[Sameti et Ward, 1996] Sameti, M. and Ward, R. (1996). A fuzzy segmentation algorithm for mammogram partitioning. *Digital Mammography*, pages 471–474. (Cité en page 53.)

[Sameti et al., 1997] Sameti, M., Ward, R. K., Morgan-Parkes, J., and Palcic, B. (1997). A method for detection of malignant masses in digitized mammograms using a fuzzy segmentation algorithm. In *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 2, pages 513–516. IEEE. (Cité en page 53.)

[Sampat et Bovik, 2003] Sampat, M. P. and Bovik, A. C. (2003). Detection of spiculated lesions in mammograms. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, volume 1, pages 810–813. IEEE. (Cité en pages 73 et 76.)

[Saslow et al., 2007] Saslow, D., Boetes, C., Burke, W., Harms, S., Leach, M. O., Lehman, C. D., Morris, E., Pisano, E., Schnall, M., Sener, S., et al. (2007). American cancer society guidelines for breast screening with mri as an adjunct to mammography. *CA : a Cancer Journal for Clinicians*, 57(2) :75–89. (Cité en page 16.)

[Scharcanski et Jung, 2006] Scharcanski, J. and Jung, C. R. (2006). Denoising and enhancing digital mammographic images for visual screening. *Computerized Medical Imaging and Graphics*, 30(4) :243–254. (Cité en pages 32 et 33.)

[Shi et al., 2008] Shi, J., Sahiner, B., Chan, H.-P., Ge, J., Hadjiiski, L., Helvie, M. A., Nees, A., Wu, Y.-T., Wei, J., Zhou, C., et al. (2008). Characterization of mammographic masses based on level set segmentation with new image features and patient information. *Medical physics*, 35(1) :280–290. (Cité en page 51.)

[Sibala et al., 1981] Sibala, J. L., Chang, C. J., Lin, F., and Jewell, W. R. (1981). Computed tomographic mammography : Diagnosis of mammographically and clinically occult carcinoma of the breast. *Archives of Surgery*, 116(1) :114–117. (Cité en page 16.)

[Sickles, 2007] Sickles, E. A. (2007). Wolfe mammographic parenchymal patterns and breast cancer risk. *American Journal of Roentgenology*, 188(2) :301–303. (Cité en page 20.)

[Simoncelli et Adelson, 1996] Simoncelli, E. P. and Adelson, E. H. (1996). Noise removal via bayesian wavelet coring. In *Proceedings., International Conference on Image Processing*, volume 1, pages 379–382. IEEE. (Cité en page 33.)

[Simonyan et Zisserman, 2014] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv :1409.1556*. (Cité en pages 96, 100 et 116.)

[Snoek et al., 2012] Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951–2959. (Cité en page 103.)

[Sobel, 1970] Sobel, I. (1970). Camera models and machine perception. Technical report, DTIC Document. (Cité en page 50.)

[Soille, 2013] Soille, P. (2013). *Morphological image analysis : principles and applications*. Springer Science & Business Media. (Cité en page 49.)

[Song et al., 2010] Song, E., Xu, S., Xu, X., Zeng, J., Lan, Y., Zhang, S., and Hung, C.-C. (2010). Hybrid segmentation of mass in mammograms using template matching and dynamic programming. *Academic Radiology*, 17(11) :1414–1424. (Cité en page 26.)

[Song et al., 2014] Song, H. O., Lee, Y. J., Jegelka, S., and Darrell, T. (2014). Weakly-supervised discovery of visual pattern configurations. In *Advances in Neural Information Processing Systems*, pages 1637–1645. (Cité en page 115.)

[Srivastava et al., 2014] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout : a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1) :1929–1958. (Cité en pages 101 et 102.)

[Stathaki et Constantinides, 1994] Stathaki, T. and Constantinides, A. (1994). Neural networks and higher order spectra for breast cancer detection. In *Proceedings of the 1994 IEEE Workshop Neural Networks for Signal Processing*, pages 473–481. IEEE. (Cité en page 54.)

[Suckling et al., 1994] Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S., et al. (1994). The mammographic image analysis society digital mammogram database. In *Exerpta Medica. International Congress Series*, volume 1069, pages 375–378. (Cité en page 26.)

[Suliga et al., 2008] Suliga, M., Deklerck, R., and Nyssen, E. (2008). Markov random field-based clustering applied to the segmentation of masses in digital mammograms. *Computerized Medical Imaging and Graphics*, 32(6) :502–512. (Cité en pages 52 et 57.)

[Sutskever et al., 2013] Sutskever, I., Martens, J., Dahl, G. E., and Hinton, G. E. (2013). On the importance of initialization and momentum in deep learning. *ICML (3)*, 28 :1139–1147. (Cité en page 102.)

[Szegedy et al., 2015] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9. (Cité en pages 103 et 116.)

[Szekely et al., 2006] Szekely, N., Toth, N., and Pataki, B. (2006). A hybrid system for detecting masses in mammographic images. *IEEE Transactions on Instrumentation and Measurement*, 55(3) :944–952. (Cité en page 57.)

[Szummer et al., 2008] Szummer, M., Kohli, P., and Hoiem, D. (2008). Learning crfs using graph cuts. In *European Conference on Computer Vision*, pages 582–595. Springer. (Cité en page 54.)

[Tabár et Dean, 1982] Tabár, L. and Dean, P. B. (1982). Mammographic parenchymal patterns : risk indicator for breast cancer ? *Jama*, 247(2) :185–189. (Cité en page 20.)

[Tabár et al., 2005] Tabár, L., Tot, T., and Dean, P. (2005). *Breast cancer : the art and science of early detection with mammography : perception, interpretation, histopathologic correlation*. Thieme. (Cité en pages 20, 37 et 68.)

[Tao et al., 2010] Tao, Y., Lo, S.-C. B., Freedman, M. T., Makariou, E., and Xuan, J. (2010). Multilevel learning-based segmentation of ill-defined and spiculated masses in mammograms. *Medical Physics*, 37(11) :5993–6002. (Cité en pages 66, 68, 69 et 70.)

[Tao et al., 2007] Tao, Y., Lo, S.-C. B., Freedman, M. T., and Xuan, J. (2007). A preliminary study of content-based mammographic masses retrieval. In *Medical Imaging*, pages 65141Z–65141Z. International Society for Optics and Photonics. (Cité en page 94.)

[te Brake et Karssemeijer, 1998] te Brake, G. M. and Karssemeijer, N. (1998). Comparison of three mass detection methods. In *Digital Mammography*, pages 119–126. Springer. (Cité en page 76.)

[te Brake et Karssemeijer, 1999] te Brake, G. M. and Karssemeijer, N. (1999). Single and multiscale detection of masses in digital mammograms. *IEEE Transactions on Medical Imaging*, 18(7) :628–639. (Cité en pages 51 et 76.)

[Teague, 1980] Teague, M. R. (1980). Image analysis via the general theory of moments*. *JOSA*, 70(8) :920–930. (Cité en pages 88 et 124.)

[Timp et Karssemeijer, 2004] Timp, S. and Karssemeijer, N. (2004). A new 2d segmentation method based on dynamic programming applied to computer aided detection in mammography. *Medical physics*, 31(5) :958–971. (Cité en page 51.)

[Tourassi et al., 2007] Tourassi, G. D., Harrawood, B., Singh, S., Lo, J. Y., and Floyd, C. E. (2007). Evaluation of information-theoretic similarity measures for content-based retrieval and detection of masses in mammograms. *Medical Physics*, 34(1) :140–150. (Cité en page 94.)

[Tourassi et al., 2003] Tourassi, G. D., Vargas-Voracek, R., Catarious Jr, D. M., and Floyd Jr, C. E. (2003). Computer-assisted detection of mammographic masses : A template matching scheme based on mutual information. *Medical Physics*, 30(8) :2123–2130. (Cité en pages 54 et 94.)

[Tropp, 2004] Tropp, J. A. (2004). Greed is good : Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10) :2231–2242. (Cité en page 35.)

[Tsochantaridis et al., 2005] Tsochantaridis, I., Joachims, T., Hofmann, T., and Altun, Y. (2005). Large margin methods for structured and interdependent output variables. *Journal of Machine Learning Research*, 6(Sep) :1453–1484. (Cité en page 54.)

[Vachier et Meyer, 2005] Vachier, C. and Meyer, F. (2005). The viscous watershed transform. *Journal of Mathematical Imaging and Vision*, 22(2-3) :251–267. (Cité en page 50.)

[Vapnik, 2013] Vapnik, V. (2013). *The nature of statistical learning theory*. Springer Science & Business Media. (Cité en page 75.)

[Varela et al., 2007] Varela, C., Tahoces, P. G., Méndez, A. J., Souto, M., and Vidal, J. J. (2007). Computerized detection of breast masses in digitized mammograms. *Computers in Biology and Medicine*, 37(2) :214–226. (Cité en page 53.)

[Varma et Zisserman, 2002] Varma, M. and Zisserman, A. (2002). Classifying images of materials : Achieving viewpoint and illumination independence. In *European Conference on Computer Vision*, pages 255–271. Springer. (Cité en page 93.)

[Velthuizen, 2000] Velthuizen, R. P. (2000). Computer diagnosis of mammographic masses. In *Applied Imagery Pattern Recognition Workshop, 2000. Proceedings. 29th*, pages 166–172. IEEE. (Cité en page 52.)

[Vincent, 1992] Vincent, L. (1992). Morphological algorithms. *Optical Engineering-New York Marcel Dekker Incorporated*, 34 :255–255. (Cité en page 49.)

[Vincent et Soille, 1991] Vincent, L. and Soille, P. (1991). Watersheds in digital spaces : an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6) :583–598. (Cité en page 49.)

[Wan et al., 2014] Wan, J., Wang, D., Hoi, S. C. H., Wu, P., Zhu, J., Zhang, Y., and Li, J. (2014). Deep learning for content-based image retrieval : A comprehensive study. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 157–166. ACM. (Cité en page 114.)

[Wan et al., 2013] Wan, L., Zeiler, M., Zhang, S., Cun, Y. L., and Fergus, R. (2013). Regularization of neural networks using dropconnect. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pages 1058–1066. (Cité en page 102.)

[Warde-Farley et al., 2013] Warde-Farley, D., Goodfellow, I. J., Courville, A., and Bengio, Y. (2013). An empirical analysis of dropout in piecewise linear networks. *arXiv preprint arXiv :1312.6197*. (Cité en page 102.)

[Wee et Paramesran, 2007] Wee, C.-Y. and Paramesran, R. (2007). On the computational aspects of zernike moments. *Image and Vision Computing*, 25(6) :967–980. (Cité en page 124.)

[Wei et al., 2011] Wei, C.-H., Li, Y., and Huang, P. J. (2011). Mammogram retrieval through machine learning within bi-rads standards. *Journal of Biomedical Informatics*, 44(4) :607–614. (Cité en page 94.)

[Wei et al., 1995] Wei, D., Sahiner, B., Chan, H.-P., and Petrick, N. (1995). Detection of masses on mammograms using a convolution neural network. In *ICASSP-95., 1995 International Conference on Acoustics, Speech, and Signal Processing, 1995*, volume 5, pages 3483–3486. IEEE. (Cité en page 54.)

[Weszka et al., 1976] Weszka, J. S., Dyer, C. R., and Rosenfeld, A. (1976). A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 4(SMC-6) :269–285. (Cité en pages 89 et 92.)

[Whatmough, 1991] Whatmough, R. (1991). Automatic threshold selection from a histogram using the âexponential hullâ. *CVGIP : Graphical Models and Image Processing*, 53(6) :592–600. (Cité en page 53.)

[Wolfe, 1976] Wolfe, J. N. (1976). Breast patterns as an index of risk for developing breast cancer. *American Journal of Roentgenology*, 126(6) :1130–1137. (Cité en pages 20 et 21.)

[Wu et al., 2015] Wu, R., Yan, S., Shan, Y., Dang, Q., and Sun, G. (2015). Deep image : Scaling up image recognition. *arXiv preprint arXiv :1501.02876*, 22 :388. (Cité en page 103.)

[Yan et al., 2015] Yan, Z., Zhang, H., Piramuthu, R., Jagadeesh, V., DeCoste, D., Di, W., and Yu, Y. (2015). Hd-cnn : hierarchical deep convolutional neural networks for large scale visual recognition. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2740–2748. (Cité en page 117.)

[Yap et al., 2003] Yap, P.-T., Paramesran, R., and Ong, S.-H. (2003). Image analysis by krawtchouk moments. *IEEE Transactions on Image Processing*, 12(11) :1367–1377. (Cité en pages 88, 124 et 125.)

[Yin et al., 2003] Yin, L., Deshpande, S., and Chang, J. K. (2003). Automatic lesion/tumor detection using intelligent mesh-based active contour. In *Tools with Artificial Intelligence, 2003. Proceedings. 15th IEEE International Conference on*, pages 390–397. IEEE. (Cité en page 51.)

[You et Jain, 1984] You, Z. and Jain, A. K. (1984). Performance evaluation of shape matching via chord length distribution. *Computer Vision, Graphics, and Image Processing*, 28(2) :185–198. (Cité en page 126.)

[Yu et al., 2014] Yu, W., Yang, K., Bai, Y., Yao, H., and Rui, Y. (2014). Visualizing and comparing convolutional neural networks. *arXiv preprint arXiv :1412.6631*. (Cité en page 114.)

[Yuan et al., 2007] Yuan, Y., Giger, M. L., Li, H., Suzuki, K., and Sennett, C. (2007). A dual-stage method for lesion segmentation on digital mammograms. *Medical physics*, 34(11) :4180–4193. (Cité en page 51.)

[Zeiler et Fergus, 2014] Zeiler, M. D. and Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European Conference on Computer Vision*, pages 818–833. Springer. (Cité en page 114.)

[Zhang et al., 2004] Zhang, H., Foo, S. W., Krishnan, S. M., and Thng, C. H. (2004). Automated breast masses segmentation in digitized mammograms. In *2004 IEEE International Workshop on Biomedical Circuits and Systems*, pages S2–2. IEEE. (Cité en page 49.)

[Zhang et Say-Wei, 2006] Zhang, H. and Say-Wei, F. (2006). Computer aided detection of breast masses from digitized mammograms. *IEICE transactions on information and systems*, 89(6) :1955–1961. (Cité en page 49.)

[Zhang et al., 2006] Zhang, Y., Chan, H.-P., Sahiner, B., Wei, J., Goodsitt, M. M., Hadjiiski, L. M., Ge, J., and Zhou, C. (2006). A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis. *Medical physics*, 33(10) :3781–3795. (Cité en page 18.)

[Zhen et Chan, 2001] Zhen, L. and Chan, A. K. (2001). An artificial intelligent algorithm for tumor detection in screening mammogram. *IEEE Transactions on Medical Imaging*, 20(7) :559–567. (Cité en pages 57 et 58.)

[Zheng et al., 1995] Zheng, B., Chang, Y.-H., and Gur, D. (1995). Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. *Academic radiology*, 2(11) :959–966. (Cité en pages 49 et 54.)

[Zheng et al., 2006] Zheng, B., Lu, A., Hardesty, L. A., Sumkin, J. H., Hakim, C. M., Ganott, M. A., and Gur, D. (2006). A method to improve visual similarity of breast masses for an interactive computer-aided diagnosis environment. *Medical Physics*, 33(1) :111–117. (Cité en page 94.)

[Zheng et al., 2014] Zheng, L., Wang, S., He, F., and Tian, Q. (2014). Seeing the big picture : Deep embedding with contextual evidences. *arXiv preprint arXiv :1406.0132*. (Cité en page 116.)

[Zhou et al., 2015] Zhou, B., Jagadeesh, V., and Piramuthu, R. (2015). Conceptlearner : Discovering visual concepts from weakly labeled image collections. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1492–1500. (Cité en page 115.)

[Zwiggelaar et al., 1998] Zwiggelaar, R., Astley, S., and Taylor, C. (1998). Detecting the central mass of a spiculated lesion using scale-orientation signatures. In *Digital Mammography*, pages 63–70. Springer. (Cité en page 75.)

[Zwiggelaar et al., 1999] Zwiggelaar, R., Parr, T. C., Schumm, J. E., Hutt, I. W., Taylor, C. J., Astley, S. M., and Boggis, C. R. (1999). Model-based detection of spiculated lesions in mammograms. *Medical Image Analysis*, 3(1) :39–62. (Cité en pages 72, 75 et 82.)

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre : Contributions à l'analyse d'images médicales pour la reconnaissance du cancer du sein

Keywords : Cancer du sein, BI-RADS IV, BI-RADS V, Diagnostic assistée par ordinateur, Champs de Markov, *Deep learning*

Résumé : Le diagnostic assisté par ordinateur du cancer du sein suscite de plus en plus un réel engouement en raison de la quantité sans cesse croissante d'images mammographiques issues des campagnes de dépistage. L'accent est mis sur les opacités mammaires en raison du risque élevé de cancer qui leur est associé. En effet, la variabilité des formes rencontrées et la difficulté à discerner les masses surtout quand ces dernières sont embarquées dans des densités importantes exigent une nouvelle stratégie plutôt adaptée aux cas les plus complexes à savoir les masses appartenant aux classes BI-RADS IV et V, *c-à-d.* respectivement les masses malignes spiculées et les distorsions architecturales. Dans ce travail, un système de diagnostic assisté par ordinateur entièrement automatique et conçu pour la segmentation et la classification des opacités dans les catégories bénigne/maligne ou graisseuse/dense est abordé. Dans un premier temps, nous avons développé une approche de pré-traitement des images fondée sur l'apprentissage d'un dictionnaire parcimonieux sur les bases d'images, combinée à une réduction de dimension afin de supprimer de façon efficace et rapide le bruit de numérisation des images mammographiques présentes dans les bases utilisées pour concevoir notre système de diagnostic en comparaison des approches déjà existantes. Une fois les images pré-traitées, nous avons mis en place une procédure de segmentation non-supervisée des masses basée sur les champs de Markov et qui a l'avantage d'être à la fois plus rapide, plus efficace et plus robuste que les meilleures techniques de segmentation disponibles dans l'état-de-l'art. De plus, la méthode proposée s'affranchit de la variabilité des masses et ce quelque soit la densité de l'image. Dans l'idée de décrire convenablement les lésions malignes spiculées, nous avons conçu une méthode de segmentation des spicules qui présente la particularité de ne pas recourir à l'utilisation de descripteurs extraits manuellement dont les performances peuvent varier en fonction de leur qualité. L'approche proposée repose sur des hypothèses que nous avons formulées concernant l'aspect des spicules. Celles-ci nous ont conduits à développer un modèle Markovien combiné à une transformée de Radon locale pour extraire les structures curvilignes de l'image. Ensuite, nous servant d'un modèle a contrario, nous avons pu extraire les spicules de l'ensemble des structures détectées. Cette phase, vient clore la première partie de la conception de notre système, qui est en mesure d'extraire soit des masses spiculées, soit des distorsions architecturales. Afin de finaliser sa conception, nous avons procédé à la création d'un modèle d'aide à la décision qui, à l'inverse de ce qui s'est toujours fait dans l'état-de-l'art pour la discrimination des masses, procède à une extraction non-supervisée des descripteurs à l'aide d'une méthode issue du *Deep learning*, à savoir les réseaux de neurones à convolution. Les descripteurs extraits, sont ensuite utilisés dans un classifieur SVM pour apprendre un modèle. Ce modèle servira par la suite à la reconnaissance du cancer du sein. Les résultats obtenus pour chacune des étapes du système de diagnostic sont très intéressants et viennent combler un vide important dans la classification des masses en général et dans la distinction des masses malignes entre elles en particulier en se fondant sur trois niveaux de décision que sont la forme, la densité et les spicules.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Title : Contributions to medical images analysis for breast cancer recognition

Keywords : Breast cancer, BI-RADS IV, BI-RADS V, Computer-aided diagnosis, Markov random field, Deep learning

Abstract : Computer-aided diagnosis of breast cancer is raising increasingly a genuine enthusiasm because of the ever-increasing quantity of mammographic images from breast cancer screening campaigns. The focus is on breast masses due to the high risk of cancer associated with them. Indeed, the variability of shape encountered and the difficulty to discern the masses especially when they are embedded in a high density require a new approach especially suited for the most complex cases namely the masses which belong to classes BI-RADS IV and V, *i.e.*, spiculated breast mass and architectural distortion. In this work, a fully automatic computer-aided diagnosis system is designed for the segmentation and classification of breast mass. Initially, we developped a pre-processing method combined with the reduction of the dictionary size in order to remove effectively and quickly the digitization noise of the mammographic images that make up the database used to design our computer-aided diagnosis system in comparison with the existing approaches. After the image pre-processing, we have proposed an unsupervised segmentation method based on a Markov random field which has the advantage of being faster, more efficient and more robust than the state-of-art segmentation methods. Furthermore, the proposed method overcomes the variability of the breast masses whatever the image density. In purpose to describe correctly the spiculated malignant lesions, we proposed an approach which avoid the computation and extraction of local features, and to rely on general-purpose classification procedures whose performance and computational efficiency can greatly vary depending on design and image characteristics. The proposed method is based on several assumptions on the structure of spicules as they appear in mammograms which have been reported in the literature. In order to make use of the above assumptions, the proposed method proceeds the following steps: first the mammogram is separated into patches onto which the curvilinear structures are discretized into segments due to Radon transform. Then, Markov modeling and contextual information are used to refine the segment positions and associate segments into curvilinear structures. Finally, spicules are detected based on a contrario model. This stage conclude the first part of the design of our computer-aided diagnosis system, that is able to extract both spiculated masses and architectural distortion. In order to complete the design of the diagnosis system, we carried out the creation of a decision support model which, contrary to what has always been done in the state-of-art for discrimination of the masses, conducts an unsupervised extraction of features through Deep learning approach – namely convolutional artificial neural networks –, combined with an SVM-type classifier. The obtained model is then stored and used as a classifier for breast cancer recognition tasks during the generalization phase. The results obtained for each step of the design of our system are very interesting and come to fill an important gap in the distinction of different type of malignant masses.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France