

DEDICACE

Je dédie ce travail de Mastère à ma chère mère et mon père pour leurs sacrifices illimités pour ma réussite.

Que Dieu leur préserve une bonne santé et une longue vie.

A mes frères et sœurs

Et spécialement à Taoufik

A mes amis.

En tout modeste témoignage de ma très grande affection et ma haute considération, je vous dédie ce travail en témoignage de ma gratitude et mon dévouement.

REMERCIEMENTS

C'est avec un grand plaisir que je réserve ces lignes comme un signe de gratitude et de reconnaissance à tous ceux qui ont contribué, de près ou de loin, au bon déroulement de ce Mastère. Qu'ils trouvent ici l'expression de mes sincères remerciements.

Je tiens de remercier vivement mon cher encadreur M. Adel MAHFoudhi, ces conseils bienveillantes et surtout son soutien moral et ses sentiments parentaux pendant les moments les plus critiques.

Je tien de remercier vivement aussi M. Hedi TMAR pour son aide, sa disponibilité et ces directives durant toute la période de réalisation de ce projet de Mastère.

J'adresse mes sincères remerciements à M. CHOKRI BEN AMAR qui m'a fait l'honneur d'avoir accepté de présider le jury de ma soutenance pour ce Mastère.

J'exprime mes remerciements à M. Moncef BOUSSALMI pour l'honneur qu'il m'a accordé en acceptant de juger mon travail.

Je tiens également à remercier tous mes collègues, et mes amis à l'ENIS. Ainsi que les membres du laboratoire de recherche CES (Computing, Electronic & Smart engineering system design) à l'ENIS et en particulier mes amis AHMED KHARRAT et Atef BOUJELBEN. Je leurs souhaite une bonne continuation.

Je tiens enfin à exprimer toute ma gratitude à tous ceux qui, par leur aide morale ou matérielle, ont contribué à l'accomplissement de ce projet de Mastère.

Table des matières

INTRODUCTION GENERALE	1
CHAPITRE -1- GENERALITE SUR LA MAMMOGRAPHIE	3
1. Anatomie du sein	3
2. Cancer de sein	4
2.1. Pathologies mammaires	4
2.2. Diagnostic du cancer du sein	4
3. Mammographie et dépistage	5
3.1. Dépistage du cancer du sein	5
3.1.1. Bénéfices	6
3.1.2. Limites	6
3.2. Mammographie	7
4. Anomalies radiologiques	8
4.1. Masse	8
4.1.1. Description par région	9
4.1.2. Description par contour	9
4.1.3. Description par densité	10
4.2. Calcifications	11
4.3. Distorsion architecturale	13
5. Systèmes CAD/CADx	13
5.1. Technologie R2	14
5.2. Intelligent Systems Software	15
5.3. CADx Medical Systems	15
6. Conclusion	15
CHAPITRE -2- ETAT DE L'ART	17
1. Analyse mammographique	17
2. Détection assistée par ordinateur	21
2.1. Objectif de CAD	21
2.2. Détection de masses	22
2.2.1. Identification des régions	22
2.2.2. Discrimination entre les régions	24
3. Diagnostique assistée par ordinateur	25
3.1. Objectif de CADx	25
3.2. Diagnostic des masses	26
3.2.1. Segmentation	26
3.2.1.1. Méthodes basées sur la région	27

Table des matières

3.2.1.2. Méthodes basées sur le contour	27
3.2.1.3. Méthodes basées sur le regroupement de pixels (clustering)	28
3.2.2. Extraction des caractéristiques	28
3.2.2.1. Caractéristiques de forme	28
3.2.2.2. Caractéristiques combinées (texture et forme)	37
3.2.3. Classification	38
4. Flôt de diagnostic mammographique proposé	39
5. Conclusion	41
CHAPITRE -3- EXTRACTION DES CARACTERISTIQUES	42
1. Forme	42
2. Description de la forme	43
3. Méthodes de description de forme	44
3.1. Caractéristiques de forme simples	45
3.1.1. Aire et périmètre	45
3.1.2. Circularité	45
3.1.3. Excentricité	46
3.2. Enveloppe convexe	47
3.2.1. Convexité (CVX)	48
3.2.2. Valeur moyenne résiduelle (NRV)	50
3.3. Mesure de la Distance Radiale (RDM)	51
3.3.1. Principe	51
3.3.2. Caractéristiques	53
3.3.2.1. Distance moyenne normalisée (d_{moy})	53
3.3.2.2. Déviation standard (σ)	54
3.3.2.3. Zero Crossing (ZC)	55
3.3.2.4. Rugosité (R)	56
3.3.2.5. Rapport d'aire (AR)	57
4. Descripteur de forme proposé	58
4.1. Caractéristiques basées sur le contour	59
4.1.1. RDM étendue	59
4.1.2. Indice d'Angle (IA)	61
4.1.3. Convexité (CVX)	63
4.2. Les caractéristiques basées sur la région	64
4.2.1. Circularité (C)	64
4.2.2. Cercle Interne Externe (CIE)	64
4.2.3. Valeur moyenne résiduelle (NRV)	66
5. Conclusion	66
CHAPITRE -4- EVALUATION ET VALIDATION DU FLOT PROPOSE	68
1. Présentation du flôt d'analyse	68
1.1. Identification de la masse (méthodes de segmentation)	69
1.1.1. Prétraitement (élimination de bruit)	71
1.1.2. Extraction de la région (seuillage)	71
1.1.3. Extraction du contour (le filtrage de Sobel)	71
1.1.4. Extraction des points du contour	72
1.2. Extraction des caractéristiques	73

1.3. Classification	76
2. Evaluation des caractéristiques	77
2.1. Base d'images utilisée	78
2.2. Evaluation des caractéristiques basées sur le contour et sur la région	79
2.3. Evaluation des caractéristiques RDM	81
2.4. Evaluation de la caractéristique d'angle (IA)	82
2.5. Evaluation de la caractéristique de convexité (CVX)	83
2.6. Comparaison avec d'autres approches	84
3. Conclusion	85
CONCLUSION GENERALE	86

Liste des Figures

Fig. 1.1. L'anatomie du sein.....	3
Fig. 1.2. La masse : (a) nodulaire et (b) stellaire.....	8
Fig. 1.3. Caractéristiques selon la région d'une masse.....	9
Fig. 1.4. Caractéristiques selon le contour d'une masse.....	10
Fig. 1.5. Caractéristiques selon la densité d'une masse.....	10
Fig. 1.6. (a) Mammographie montrant un foyer de microcalcifications et (b) agrandissement de la région contenant l'amas de microcalcifications	11
Fig. 1.7. Distorsion architecturale	13
Fig. 1.8. Détection automatique des régions suspectes (les masses sont représentées par une étoile et les microcalcifications par un triangle)	14
Fig. 2.1. Flôt général d'analyse : (a) détection et (b) diagnostic	18
Fig. 2.2. Flôt de détection/ diagnostic	19
Fig. 2.3. Flôt de détection de masses stellaires	19
Fig. 2.4. Flôt de détection/ diagnostic	20
Fig. 2.5. (a) Les directions de rayons d'une masse stellaire sont différentes des directions de marquages linéaires normaux et (b) l'écart type de l'histogramme de l'orientation de gradient différencie la région stellaire de celle normale	23
Fig. 2.6. (a) Une tumeur bénigne et (b) sa signature.....	31
Fig. 2.7. (a) Une tumeur maligne et (b) sa signature	31
Fig. 2.8. Signature par l'angle ϕ de la tangente au point P_i et (b) la signature du contour	32
Fig. 2.9. Calcul de la caractéristique FD pour 6 cas de masses dont 3 sont bénignes (a, b et c) et 3 sont maligne (d, e et f)	34
Fig. 2.10. Flôt de diagnosticmammographique proposé	40
Fig. 3.1. Forme binnaire des différents objets	42
Fig. 3.2. Notion de similarité : (a)régions similaires et (b) contours différents	43
Fig. 3.3. Vecteur des caractéristiques.....	43
Fig. 3.4. Transformations appliquées à l'image : (a)changement de dimension et (b) rotation	44
Fig. 3.5. Une (a) forme circulaire et (b) une forme irrégulière (non circulaire).	46
Fig. 3.6. Calcul d'excentricité pour une (a) tumeur bénigne et (b) une tumeur maligne	47
Fig. 3.7. Caractérisation non fiable pour deux tumeurs malignes : (a) Cas 1 et (b) Cas 2	47
Fig. 3.8. L'enveloppe convexe E d'une forme F	48
Fig. 3.9. (a)forme convexes et (b) forme non convexe	49
Fig. 3.10. (a) Région régulière (tumeur bénigne) et (b) région irrégulière (tumeur maligne)	49
Fig. 3.11. Calcul de NRV: (a) forme irrégulière et son enveloppe convexe, (respectivement (b) forme régulière et son enveloppe convexe)	51
Fig. 3.12. Principe de la méthode RDM.....	52
Fig. 3.13. Signature de contour dans le cas (a) d'une tumeur bénigne et (b) d'une tumeur maligne.....	52
Fig. 3.14. Principe de SDEV pour (a) un contour stellaire et (b) un contour lobulé.....	54
Fig. 3.15. Signature de Zéro Crossing pour (a) une tumeur maligne et (b) une tumeur bénigne.....	55

Fig. 3.16. Principe de la rugosité (a) contour stellaire, (b) contour microlobulé et (c) contour lisse.....	56
Fig. 3.17. Principe d'AR pour (a) un contour stellaire et (b) un contour lobulé.....	57
Fig. 3.18. Descripteur de forme : (a) caractéristiques de région et (b) caractéristiques de contour	58
Fig. 3.19. Ensemble des points concaves et convexes (a) cas d'une tumeur maligne et (b) d'une tumeur bénigne.....	60
Fig. 3.20. Cas traité par la caractéristique IA : (a) contour stellaire et (b) contour micro-lobulé.....	61
Fig. 3.21. Calcul des angles θ_i et φ_i	62
Fig. 3.22. Calcul de convexités : $CVX_1 = (1-h)^2$ et $CVX_2 = 1/2(1-h)$	63
Fig. 3.23. Mauvaise description par l'excentricité : (a) forme ronde bénigne et (b) d'une forme allongée maligne.....	65
Fig. 3.24. Caractéristique CIE pour une (a) forme ronde et (b) une forme allongée.....	66
Fig. 4.1. Flot d'analyse proposé.....	68
Fig. 4.2. Flot d'identification d'une masse	70
Fig. 4.3. Filtrage Sobel : (a) Image initiale, (b) application de deux filtres 3×3 et (c) contour détecté.....	72
Fig. 4.4. Chaînage de Freeman dans le cas de direction à (a) 4 connexité et à (b) 8 connexité.....	72
Fig. 4.5. Caractéristiques (a) de région et (b) de contour	73
Fig. 4.6. Les régions à décrire : (a) rondes, (b) ovales, (c) lobulées et (d) irrégulières	74
Fig. 4.7. Contours à décrire (a) stellaires et (b) micro-lobulés	74
Fig. 4.8. Contours (a) indistincts et (b) masqués	74
Fig. 4.9. Méthodes de classification	76

Liste des Tables

<i>Tab. 1.1 Classification des microcalcifications.....</i>	<u>12</u>
<i>Tab. 2.1 Méthode d'extraction des caractéristiques basées sur la forme.....</i>	<u>29</u>
<i>Tab. 2.2 Performances obtenues avec des caractéristiques de forme, de texture et combinées</i>	<u>38</u>
<i>Tab. 2.3 Résultat de test de quelques travaux utilisant des classifieurs supervisés.....</i>	<u>39</u>
<i>Tab. 3.1 Evaluation de la caractéristique NRV.....</i>	<u>51</u>
<i>Tab. 3.2 Comparaison des valeurs retenues par la méthode RDM et notre RDM étendue.....</i>	<u>61</u>
<i>Tab. 4.1 Liste des caractéristiques utilisées</i>	<u>75</u>
<i>Tab. 4.2 Résumé de l'ensemble de paramètres utilisés</i>	<u>78</u>
<i>Tab. 4.3 Résumé de l'ensemble d'image DDSM utilisées dans ce travail</i>	<u>79</u>
<i>Tab. 4.4 Résultat de test des caractéristiques basées sur le contour</i>	<u>79</u>
<i>Tab. 4.5 Résultat de test des caractéristiques basées sur la région.....</i>	<u>80</u>
<i>Tab. 4.6 Résultat de test de notre vecteur caractéristique basé sur le contour et sur la région</i>	<u>80</u>
<i>Tab. 4.7 Résultats de test des caractéristiques RDM vs RDM étendue.....</i>	<u>81</u>
<i>Tab. 4.8 Résultats de test avec intégration de la technique RDM</i>	<u>82</u>
<i>Tab. 4.9 Résultats de test de la caractéristique d'angle (IA)</i>	<u>83</u>
<i>Tab. 4.10 Résultats de test de la caractéristique de convexité (CVX).....</i>	<u>83</u>
<i>Tab. 4.11 Comparaison avec d'autres approches</i>	<u>84</u>

Introduction Générale

Le cancer du sein est la première cause de mortalité féminine par cancer en plusieurs pays.

Des études ont montré qu'en moyenne 12% des femmes risquent de développer un cancer du sein durant leur vie [1]. L'agence de recherche sur le cancer de l'organisation mondiale de la santé basée à Lyon estime que plus de 150 000 femmes dans le monde meurent par le cancer du sein par an [2]. Ces chiffres montrent l'importance d'une détection précoce de cette maladie. La prévention de cette maladie est très difficile. Pour cela, le seul moyen permettant de réduire le taux de la mortalité est le dépistage précoce. Ce dernier, permet de faire le tri entre les patientes ne présentant aucun signe de cancer et celles ayant probablement un cancer. À l'heure actuelle, la chaîne conventionnelle de diagnostic repose sur l'examen clinique, la mammographie X et l'échographie en mode B. En cas de doute sur la nature d'une structure comprise à l'intérieur du sein, on procède à un examen cytologique. L'imagerie par résonance magnétique étant plus coûteuse, elle est pratiquée essentiellement pour des études de recherches. Finalement, la mammographie reste la seule méthode de dépistage du cancer du sein à un stade précoce.

La mammographie basée sur l'examen radiologique du sein, permet principalement de repérer trois signes indicateurs de pathologies : les anomalies de la distorsion architecturale, de l'opacité (masse) et de microcalcifications.

En général, la recherche en traitement/analyse d'images et plus particulièrement dans le domaine médical progresse tous les jours. La mise à disposition de tels outils technologiques au médecin lui permet d'améliorer son diagnostic et certainement d'optimiser son environnement de travail. D'un point de vue « imagerie », l'analyse automatique des clichés mammographiques est un défi qui fait l'objet de nombreuses recherches depuis plusieurs décennies ; c'est une succession de plusieurs phases qui ont pour objectif la détection ou le diagnostic de la tumeur de sein. La description de forme représente une phase parmi les phases importantes d'analyse mammographique vu son rôle discriminatif. En ce qui concerne la détection, elle permet la distinction des régions suspectes de celles normales. Alors que pour la phase de diagnostic, elle est utilisée pour la distinction entre les tumeurs malignes et celles bénignes.

L'objectif de notre travail est la proposition d'une méthode de description des masses basée sur la forme. Cette dernière doit satisfaire un compromis qualité de description / temps de calcul. Pour cela, on commence, dans ce mémoire de mastère, par une étude théorique et bibliographique. On présente, ensuite, notre approche de description de forme. On procède, enfin, par une évaluation de notre approche d'analyse mammographique avec une interprétation des résultats obtenus.

Ce mémoire de Mastère comprend quatre chapitres :

- Le chapitre 1 présente les principales pathologies du sein et leur classification. Il finit par une présentation du domaine de recherche en exposant quelques applications développées pour l'analyse des masses.
- Le chapitre 2 présente un état de l'art sur les méthodes d'analyse mammographique.
- Le chapitre 3 explicite la méthode adoptée pour la description des masses mammographiques en indiquant les critères du choix des caractéristiques.
- Le dernier chapitre est consacré à la présentation des méthodes utilisées pour l'identification et de classification des masses pour évaluer les caractéristiques de forme proposées, ainsi que l'interprétation des résultats trouvés.

Enfin, on trace des conclusions et on donne quelques perspectives de ce travail.

Chapitre 1

Généralité sur la mammographie

Le cancer du sein est le cancer le plus répandu chez la femme à l'échelle mondiale. Il est aussi le plus fréquent dans les pays industrialisés [1], [2]. L'incidence est en hausse surtout chez les femmes post ménopausées. Par contre, chez les jeunes femmes les taux d'incidence se stabilisent ou sont même en régression. Etant donnée que la prévention du cancer de sein reste encore difficile, le dépistage reste le seul moyen permettant de diminuer le taux de mortalité ; car plus la maladie est détectée à un stade précoce plus les chances de guérison sont grandes. Dans ce contexte, d'une part, la mammographie est actuellement l'examen de référence pour le dépistage du cancer du sein, d'autre part, les progrès des technologies informatiques ont comme objectif d'aide soit à la détection soit au diagnostic afin de réduire encore le taux de mortalité.

Dans ce chapitre, on présente quatre parties : la première donne les principales pathologies du sein. Quant à la deuxième, elle présente un aperçu sur la mammographie et son importance dans la phase de dépistage et dans le diagnostic du cancer de sein. Dans la troisième partie, on focalise les propriétés de malignités et de bénignités à prendre en considération par les radiologues lors de l'analyse mammographique. Enfin, on donne un aperçu sur les outils d'analyse automatisés permettant d'aider à la réduction de taux de mortalité chez les femmes.

1. Anatomie du sein

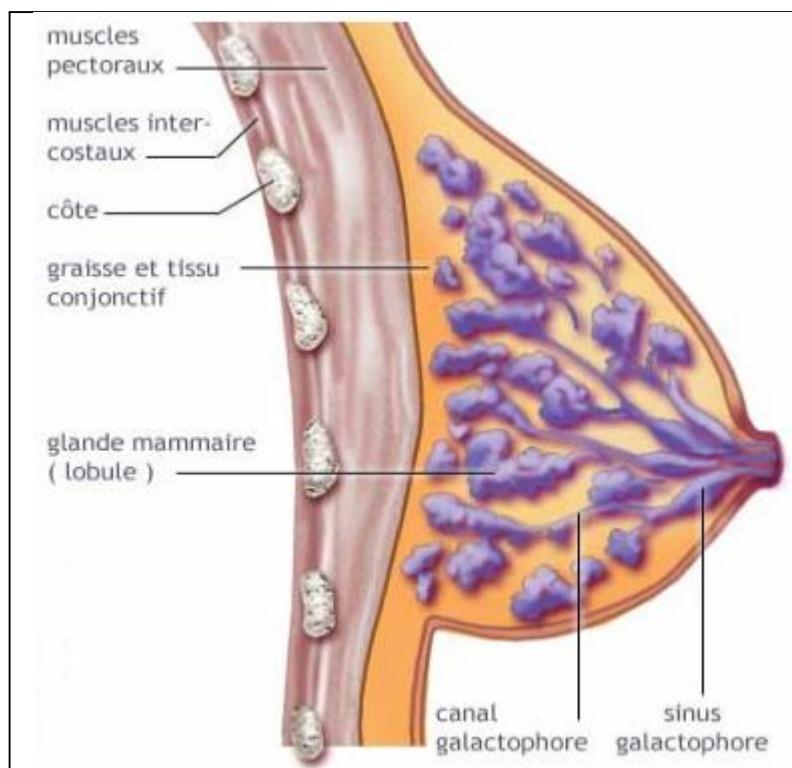


Fig. 1.1. *L'anatomie du sein.*

Le sein est un organe constitué d'un tissu graisseux qui repose sur les muscles pectoraux. Chez la femme, sa principale fonction est la lactation. Il comporte 15 à 25 lobules faits de glandes mammaires responsables de la sécrétion du lait. Celui-ci est conduit par les canaux galactophores vers le mamelon. Les seins sont soutenus par la peau et un tissu fibreux qui pénètrent à l'intérieur de la glande mammaire, participant ainsi à leur maintien.

2. Cancer de sein

Le cancer de sein est une maladie où les cellules anormales se multiplient hors de tout contrôle. La croissance de ces millions de cellules dans un seul endroit comme le sein constitue une tumeur maligne ou un cancer de sein. À partir d'une phase du cancer de sein, les cellules cancéreuses se voient propager vers d'autre partie du corps où elles s'installent et se multiplient pour y former une autre tumeur. La prévention de la maladie est très difficile car, mis à part l'âge, on ne connaît pas de facteurs de risque déterminants. Il est donc difficile de réduire l'incidence du cancer de sein par la prévention. En effet, le seul moyen de réduire la mortalité est le dépistage précoce.

On présente par la suite les différentes pathologies ainsi que les diagnostics correspondants.

2.1. *Pathologies mammaires*

Les pathologies mammaires sont classées en quatre groupes qui sont les pathologies des lobules, les pathologies des lobes, les pathologies des canaux collecteurs et les pathologies des enveloppes. Les principales pathologies bénignes et malignes du sein rencontrées sont indiquées ci-dessous :

- Les pathologies bénignes sont : L'adénofibrome, le lipome, le kyste, le nodule de mastose.
- Les pathologies malignes sont divisées en deux catégories :
 - Cancer in situ : les cellules cancéreuses se développent strictement à l'intérieur des canaux galactophores. Les cellules ne peuvent pas se déplacer dans d'autres parties du corps. Le cancer in situ est l'étape qui précède l'apparition d'un cancer infiltrant.
 - Cancer infiltrant : les cellules cancéreuses, localisées dans la paroi des canaux, détruisent et franchissent cette paroi.

2.2. *Diagnostic du cancer du sein*

L'examen clinique est la première modalité de diagnostic du cancer du sein qui doit se pratiquer, si possible, en début du cycle menstruel. Le praticien pratique une inspection comparative des deux seins et enfin effectue une palpation de la glande mammaire. Lorsque le praticien remarque

un écoulement mammelonnaire unilatéral, une infection cutanée, et une masse ou un ganglion palpable, il demande des examens complémentaires.

Malgré l'utilité de cet examen clinique, il ne peut pas détecter des cancers de moins de 0.5 cm de diamètre, et encore moins les cancers non palpables. C'est pour cela, un examen mammographique est nécessaire et est considéré comme le principal examen d'imagerie réalisé à l'aide du mammographe (ou scénographe). Ce qui montre aussi l'utilité et l'importance de cet examen est que chez les patientes âgées de 40 à 50 ans, 20 % des cancers sont ainsi découverts en plus de ceux qui le sont par examen clinique et ce taux passe à 40 % après l'âge de 50, car à cet âge le taux de graisse dans les seins est plus important et de ce fait les radiographies sont plus faciles à lire [4].

En plus de la mammographie, d'autres examens complémentaires sont pratiqués comme l'échographie, la scintigraphie, la tomodensitométrie, la résonance magnétique. Mais seule la mammographie est utile comme test de détection du cancer du sein à un stade précoce.

3. Mammographie et dépistage

La prévention de la maladie est très difficile car les facteurs de risques sont soit mal connus (facteurs nutritionnels, comportementaux ou environnementaux) soit peu influençables (risque génétique, facteurs hormonaux tels que l'âge des premières règles, de la première grossesse ou de la ménopause,...). Des études scientifiques [3] ont permis de mieux comprendre le développement des cancers, mais il n'est pas encore possible de savoir pourquoi telle personne développe tel cancer. Il est à noter que seulement 5 à 10 % des cancers du sein ont une origine héréditaire liée à la transmission de gènes délétères dont les plus fréquemment incriminés sont BRCA1 et BRCA2 (Breast Cancer Acronymes pour 1/2) associés à une prédisposition à la maladie.

Il est donc difficile de réduire l'incidence de cancer du sein par la prévention. En effet, le seul moyen de réduire la mortalité est le dépistage précoce qui est basé sur la mammographie. Dans ce qui suit, on donne le principe de la phase de dépistage et le rôle de la mammographie dans cette phase.

3.1. *Dépistage du cancer du sein*

Le dépistage, aussi appelé prévention secondaire, est défini comme une mesure pour réduire la gravité de la maladie. Il consiste, dans le cas du cancer du sein, à détecter la maladie au stade le plus précoce possible ; c'est à dire à un stade où le volume de la tumeur est le plus petit possible. Plus simplement, le dépistage revient à rechercher les signes d'une maladie avant qu'elle ne se déclare.

On note la présence de deux types de dépistage : soit individuel, soit organisé (par exemple un examen chaque deux ans). Remarquons aussi, qu'il convient de distinguer entre dépistage et diagnostic. Concernant, le dépistage il s'agit de faire le tri entre les femmes ne présentant pas de signes de cancer et celles ayant peut être un cancer. Alors que dans l'activité de diagnostic, il s'agit de caractériser la pathologie et de décider le traitement à suivre.

Donc malgré l'utilité du dépistage dans le monde médical et particulièrement en mammographie, il comporte, comme toute procédure, des risques et des bénéfices qu'il faut bien les connaître pour les maîtriser.

3.1.1. *Bénéfices*

Le dépistage par la mammographie s'est jugé obligatoire chez les femmes de plus de 50 ans et préférable chez les femmes de 40 à 49 ans. L'efficacité a été démontrée dans les années 80 aux Etats-Unis [5] et en Suède [9]. Le dépistage diminue la mortalité par cancer du sein et permet de renforcer l'effet du traitement chez les femmes de plus de 50 ans, à condition d'être périodiquement effectué chez un nombre suffisant de femmes et de se dérouler dans des conditions optimales de qualité.

Depuis, la mise en place de campagnes de dépistage de ce cancer est un thème primordial de santé publique dans des nombreux pays industrialisés. Par exemple en France, en 1989, le Fond National de Prévention, d'Evaluation et d'Informations Sanitaires (FNPEIS) a initié un vaste programme de dépistage de masse : il concernait seulement 6 départements en 1990, puis 21 départements en 1994 et aujourd'hui il porte sur toute la France ; les femmes de 50 à 74 ans sont invitées tous les deux ans par leur régime d'assurance maladie à passer une mammographie gratuite. Chaque mammographie est interprétée par le premier radiologue puis par un second lecteur plus spécialisé. Cette deuxième lecture permet de « ratrapper » 15 % des cancers détectés [11].

3.1.2. *Limites*

Les femmes sont exposées à des inconvénients du dépistage [13], aussi bien dans le cadre des programmes organisés que dans le cadre du dépistage individuel. Il faut bien connaître les inconvénients du dépistage afin de les contrôler et de les limiter. Parmi les inconvénients, on trouve :

- Les faux négatifs : le résultat de l'examen mammographique est négatif (cas bénigne), alors que réellement il s'agit d'un cas maligne.
- Les faux positifs : le résultat de l'examen mammographique est positif (cas maligne), alors que réellement il s'agit d'un cas bénigne.

Malgré les limites, le dépistage semble être le seul moyen efficace de lutter contre un cancer.

3.2. Mammographie

La mammographie (ou mastographie) est le principal moyen de dépistage et de diagnostic. C'est un examen radiographique (par rayons X) des deux seins, généralement sous deux incidences (un cliché de face et un autre en oblique) qui permet de détecter certaines anomalies. On considère qu'en moyenne le cancer est visible sur la mammographie de 1 à 3 ans avant le premier signe clinique. D'autres modalités d'imagerie (échographie, ultrasons, IRM) peuvent être employées mais à titre d'examen complémentaire en cas de diagnostic difficile.

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

- L'examen de dépistage : il comporte une mammographie de base. Après l'âge de 45 ans, cet examen peut être réalisé dans le cadre du programme de dépistage national généralisé, à raison d'une fois tous les deux ans gratuitement. Ce type d'examen très efficace permet de détecter les cancers à leur stade le plus précoce, permettant d'en guérir plus de 80% [12].
- L'examen de diagnostic : il est réalisé lorsqu'une anomalie a été détectée au cours d'un examen de dépistage ou lorsqu'une anomalie dans le sein (douleur, écoulement du mamelon, rougeur de la peau, palpation d'une "boule" dans le sein, etc.) a été découverte par la patiente ou par son médecin traitant. Cet examen aide à déterminer s'il s'agit d'une anomalie bénigne ou maligne. En plus des clichés de base, des images complémentaires peuvent être réalisées. Le médecin radiologue peut ensuite compléter la mammographie par une échographie et/ou un prélèvement (c'est une action d'extraire quelque chose du corps pour l'analyser au microscope).

On peut dire que la mammographie reste l'outil principal de dépistage pour plusieurs raisons [8] :

- Le coût moyen d'un examen mammographique est faible comparé à d'autres techniques d'imagerie comme l'IRM. De plus l'appareillage est mobile et peu encombrant ;
- L'examen n'est pas traumatisant. En effet les tissus du sein sont mous et perméables aux rayons X, rendant faibles les doses d'irradiation. D'autre part, la compression du sein, nécessaire pour obtenir une meilleure uniformité de la densité optique et une réduction du bruit de diffusé, ne doit pas en général être douloureuse ;
- La résolution spatiale d'une radiographie est de l'ordre de quelques dizaines de microns, ce qui permet de visualiser les microcalcifications, contrairement aux autres techniques d'imagerie ;
- Les mammographies numériques sont de plus en plus utilisées, ce qui permet de disposer du signal directement sous forme numérique, avec les avantages associés en terme d'archivage et

de consultation de dossier médical à distance. Pour autant, la mammographie numérique reste coûteuse pour des performances au niveau des diagnostics équivalents à celles des techniques conventionnelles.

Malgré le rôle important de la mammographie dans la phase de dépistage ou de diagnostic, ça n'empêche pas qu'elle possède des certaines limitations tels que l'irradiation par des rayons X, la réalisation d'un examen inconfortable, l'interprétation difficile et la détection manquée d'un cas maligne.

4. Anomalies radiologiques

Parmi les anomalies radiologiques les plus courantes observées sur une mammographie, on peut citer les masses (ou encore les opacités), les calcifications et les distorsions architecturales. Dans ce qui suit-on donne un aperçu sur chacune de ces trois anomalies.

4.1. *Masse*

La masse mammographique est une anomalie qui peut infecter le sein. Il existe deux types de masses (nodulaires et stellaires). On les distingue selon la forme. La masse nodulaire (figure 1.2 (a)) a une forme arrondie compacte. Elle est considérée comme tumeur bénigne si sa forme est régulière et arrondie avec un contour net et bien défini. Mais elle est considérée comme tumeur maligne si son contour est flou même sur une partie seulement.

Dans le cas de masse stellaire (figure 1.2 (b)), la forme est comparable à la forme d'une étoile. Elle est considérée comme tumeur maligne, et contient des spicules sur sa frontière.

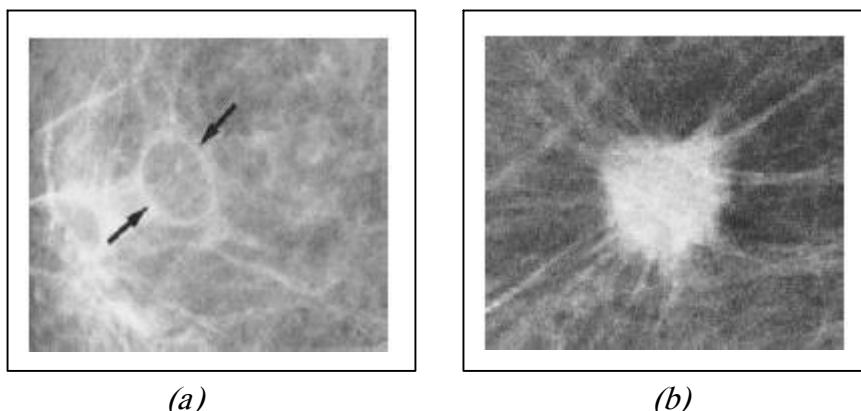


Fig. 1.2. *La masse : (a) nodulaire et (b) stellaire.*

Afin d'arriver à analyser la masse, il faut savoir les différentes caractéristiques possibles à utiliser pour distinguer entre les formes malignes de celles bénignes. Ces caractéristiques sont réparties en trois catégories : une catégorie qui décrit la région, une autre qui décrit le contour et une dernière qui caractérise la densité de la masse.

Dans ce qui suit, on présente les différentes propriétés de malignité et de bénignité qui décrivent les masses mammographiques.

4.1.1. Description par la région

La région de la forme est considérée comme un signe indiquant l'état pathologique qui peut avoir la masse mammographique. En se basant sur la région, les radiologues spécifient quatre caractéristiques de région (ronde, ovale, lobulée et irrégulière) permettant de classifier une masse en tant que maligne ou bénigne.

Les caractéristiques que peut avoir une région sont données par la figure 1.3 et décrites comme suit :

- **Ronde** : la région est sphérique ou circulaire.
- **Ovale** : la région est elliptique ou en forme d'œuf.
- **Lobulée** : la région contient de légère ondulation.
- **Irrégulière** : la région ne peut être décrite par les termes repris ci-dessus.

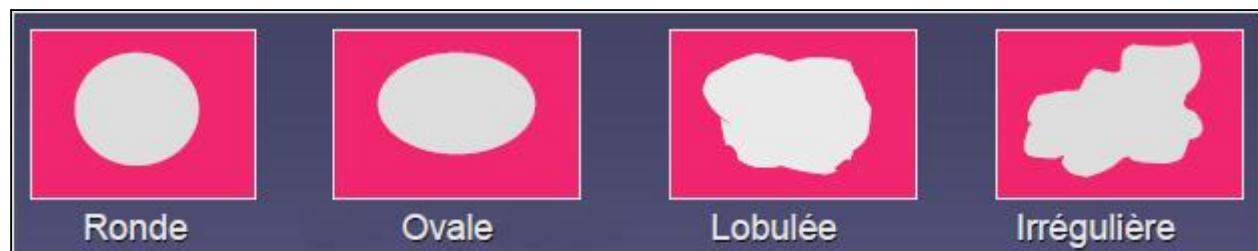


Fig. 1.3. Caractéristiques selon la région d'une masse.

Si la région est ronde, ovale ou lobulée on parle d'une masse bénigne. Alors que lorsqu'il s'agit d'une région irrégulière la masse est dite maligne.

4.1.2. Description par contour

Le contour d'une masse est pris en considération par les radiologues pour classifier les masses mammographiques. Il tient donc une place importante dans la description pathologique d'une masse. Les radiologues spécifient cinq propriétés (circonscrit, micro-lobulé, masqué, mal défini ou stellaire) que peut avoir le contour de la masse.



Fig. 1.4. Caractéristiques selon le contour d'une masse.

Les caractéristiques que peut avoir le contour d'une masse mammographique sont données par la figure 1.4 et décrites comme suit :

- **Circonscrit** : le contour est bien défini.
- **Micro-lobulé** : il y à présence de plusieurs ondulation sur le contour.
- **Masqué** : le contour est masqué par une superposition par d'autres investigations.
- **Indistinct** : le contour est mal défini pouvant correspondre à une infiltration.
- **Spiculé**: il y à présence des lignes radiaires sur le contour de la masse.

Les radiologues identifient la masse bénigne par la présence d'un contour circonscrit. Alors qu'ils identifient la masse maligne par la présence d'un contour micro-lobulé, masqué, mal-défini ou stellaire.

4.1.2. Description par densité (ou l'atténuation)

L'intensité ou l'atténuation de rayon X de la région de masse de tissu est décrite comme densité. La densité ici est la densité relative, c'est à dire plus haut, inférieur ou semblable au tissu environnant.

La plupart des cancers du sein présentent une densité supérieure ou égale au tissu environnant. Il est rare (mais non impossible) qu'un cancer du sein soit de densité inférieure.

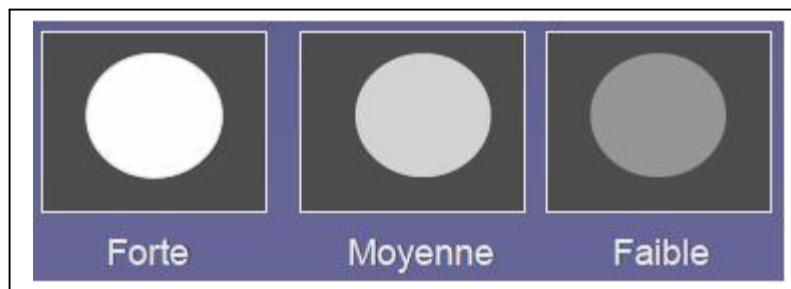


Fig. 1.5. Caractéristiques selon la densité d'une masse.

La densité est décrite en utilisant les caractéristiques données par la figure 1.5 et décrites ci-dessous :

- **Forte** (hyperdense) : densité forte par rapport au tissu environnant.
- **Moyenne** (isodense) : densité moyenne par rapport au tissu environnant.

- **Faible** (hypodense) : densité faible par rapport au tissu environnant.

4.2. *Calcifications*

Les calcifications sont des dépôts de sels de calcium qui apparaissent sous forme de points « blancs et brillants » à la mammographie. En fonction de leur taille, on parle de micro ou macrocalcifications. En effet, les macrocalcifications sont souvent bénignes alors que les microcalcifications demandent plus d'attention.

Le diagnostic dépend de leur position dans le sein, de leur arrangement géométrique (linéaire, en forme d'agrégats, ...) et du nombre de microcalcifications dans un amas, mais aussi de la comparaison avec une mammographie antérieure. La figure 1.6 (a) présente une mammographie contenant un amas de microcalcifications, et 1.6 (b) un agrandissement de la région contenant l'amas.

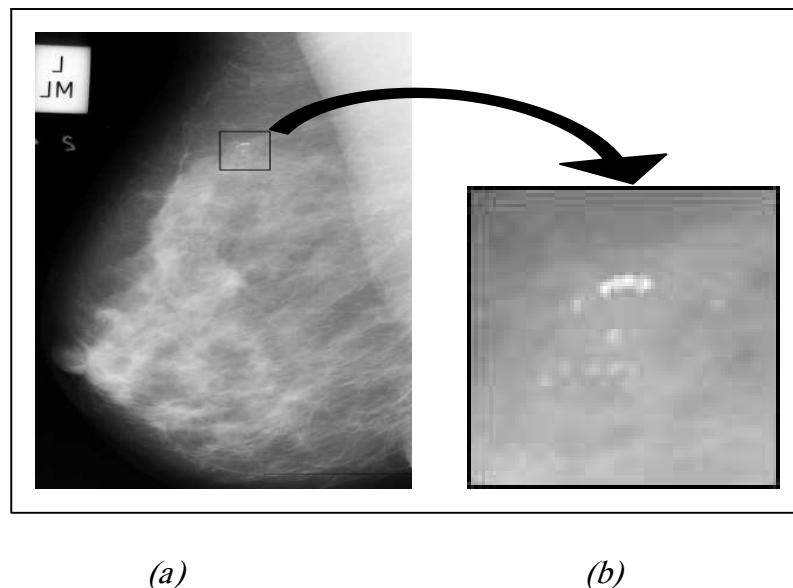


Fig. 1.6. (a) *Mammographie montrant un foyer de microcalcifications et (b) agrandissement de la région contenant l'amas de microcalcifications.*

Les radiologues caractérisent les microcalcifications à travers plusieurs propriétés [4]. Parmi les propriétés importantes pour un radiologue, on trouve :

- La taille des microcalcifications : plus la taille est faible plus le risque de malignité augmente.
- La forme : plus la forme est arrondie, moins serait le risque de malignité.
- L'orientation : les amas malignes ont tendance à avoir une forme triangulaire ou losangique orientée vers le mamelon.

- Le nombre : plus le nombre de microcalcifications est élevé, plus le risque de malignité augmente.
- Le polymorphisme : si les microcalcifications dans un amas ont des formes variées, l'amas a plus de risque d'être malinige.

La présence des calcifications dans les cancers de seins est de 30 à 50% de la totalité des catégories existants [4]. De nombreuses classifications des calcifications existent. Parmi les plus utilisées, celle de Le Gal et al [6]. Elle présente cinq types de calcifications (voir le tableau 1.1) donnés comme suit :

- Type 1: il s'agit des calcifications annulaires, arciformes ou polyédriques. Le risque de cancer du sein est quasi nul.
- Type 2: ce sont des calcifications rondes et de tailles variables avec un risque de carcinome : 22%.
- Type 3: ce sont des calcifications poussiéreuses, pulvérulentes avec un risque de cancer : 36%.
- Type 4: ce sont des calcifications irrégulières associées à un risque de cancer : 56%.
- Type 5: ce sont des calcifications vermiculaires ou branchées avec un risque de carcinome : 90%.

Type 1	Type 2	Type 3	Type 4	Type 5
				
Annulaires, arrondies au centre clair	Formes rondes, densités homogènes	poussiéreuses	Formes irrégulières	Vermiculaires Branchées
Malignité				
0%	19%	36%	52%	90%

Tab. 1.1. Classification des microcalcifications [6].

4.3. Distorsion architecturale

Il s'agit d'une distorsion de l'architecture mammaire (figure 1.7). Ces images peuvent être caractérisées par de fins rayons radiaux, ainsi que des rétractions ou des distorsions du bord du parenchyme.



Fig. 1.7. Distorsion architecturale.

Il s'agit d'images difficiles à voir et à caractériser. Leur identification requiert une bonne habitude de la mammographie pour les différencier de l'architecture normale et de ses superpositions. Elles sont particulièrement difficiles à reconnaître dans les glandes denses et hétérogènes. Ces images stellaires posent, en effet, des problèmes de diagnostic difficiles.

Le taux de mortalité a diminué pendant ces dernières années, dues en partie à l'utilisation de la mammographie et du dépistage dans la plupart des pays [14]. Du fait de ce dépistage, le nombre de mammographies à analyser est en augmentation, ce qui pose le problème de la charge de travail des spécialistes, qui varient dans leur interprétation des mammographies. Dans cette optique plusieurs systèmes automatisés d'analyse mammographique [15], [16], [17], ont été développés. On trouve le système CAD (Computer Aided Detection) et celui de CADx (Computer Aided Diagnostic) : le CAD sert à la détection d'une masse suspecte à partir d'une image mammographique, alors que les systèmes CADx aident les radiologues à la prise de décision concernant le degré de malignité des masses.

5. Systèmes CAD/CADx

Les systèmes de détection ou de diagnostic ont pour rôle l'aide à identifier d'une manière efficace les tumeurs de sein. Plusieurs études ont été faites pour évaluer les performances des

différents systèmes CAD/CADx. Dans leur étude, C. Marx et al [14], ont démontré les possibilités des systèmes CAD et CADx, en détectant la moitié de cancers du sein manqués. Plus récemment, dans une autre étude de C. Marx et al [18], il a été prouvé que l'exactitude dans la lecture des mammographies peut être augmentée d'environ de 2% en utilisant un système de CAD ou CADx. La proportion des cancers manqués peut être trouvée grâce à la double lecture avec la méthode du système CAD, avec un taux de fausse détection FP (taux noté faux positif qui représente le nombre des cas mal classifiés; c'est à dire que le système donne une classification au contraire de la réalité : par exemple pour une région réellement normale le système indique qu'elle est cancéreuse) inférieur à 1 par image. Il est connu qu'un taux élevé de FP peut constituer un problème pendant l'utilisation réelle d'un système CAD.

À notre connaissance, trois systèmes CAD/CADx [15], [16], [17] sont disponibles sur le marché. La qualité de tels systèmes dépend du taux de la détection de tumeur ainsi que de taux de FP :

- Le premier système est le système « ImageChecker » (R2 Technologie, Los Altos, CA, USA).
- Le second système est nommé « MammoReader » (Intelligent Systems Software Inc, Clearwater, FL, USA).
- Le troisième système est le « Second look » (CADx Medical Systems, Quebec, Canada).

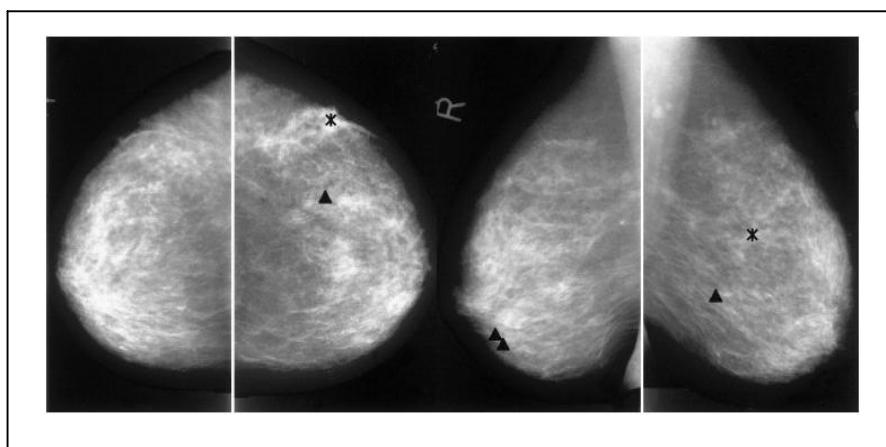


Fig. 1.8. Détection automatique des régions suspectes (les masses sont représentées par une étoile et les microcalcifications par un triangle).

5.1. Technologie R2

Image Checker® de la technologie R2 était le premier système mammographique commercial approuvé par la FDA. Ce système est conçu pour rechercher les signes qui peuvent être associés au

cancer du sein. Les masses sont identifiées par un astérisque tandis que les microcalcifications sont identifiées par un triangle comme il est indiqué à la figure 1.8. Avec ce système, l'exactitude de détection des microcalcifications a atteint 98.5% de sensibilité (la capacité d'un système à découvrir les positifs ; c'est-à-dire le taux de détection des régions cancéreuses suspectes ou malignes) avec un taux de FP de l'ordre 0.74 par image alors que l'exactitude de détection des masses a atteint 85.7% de sensibilité et un taux FP de l'environ 1.32 par image.

5.2. Intelligent Systems Software

Le FDA a approuvé le système MammoReader™ en 2002. Ce système a été conçu pour détecter des signes fondamentaux de cancer de sein dans les images mammographiques y compris les groupes des microcalcifications, les masses bien et mal définies, les masses stellaires et les distorsions architecturales.

La sensibilité totale rapportée était 89.3% (91.0% dans le cas où les microcalcifications étaient le seul signe de cancer et 87.4% dans les cas restants où les masses malignes étaient présentes).

5.3. CADx Medical Systems

CADx Medical Systems étaient la troisième système recevant l'approbation pour un système CAD mammographique appelée SecondLook™. Ce système a été conçu pour marquer les régions des images mammographiques qui donnent une indication sur des cancers. Il marque les masses avec des cercles et le groupe des microcalcifications avec des rectangles. La sensibilité du système est de l'ordre 85% (combinaison de masses et des microcalcifications) et le taux de FP est de l'ordre de 0.28 par image.

En conclusion, nous notons que les systèmes de CAD et CADx, même dans les versions actuelles, se présentent comme des outils additionnels utiles pour éviter les biopsies (technique de prélèvement d'un petit morceau de tissu afin de l'analyser au microscope) inutiles et pour augmenter l'exactitude de la lecture de la mammographie.

6. Conclusion

Le cancer du sein est un cancer qui conduit à un taux de mortalité élevé chez les femmes. Heureusement, la mammographie exécutée à des intervalles réguliers chez les femmes peut en affaiblir la gravité des conséquences.

L'automatisation a incité les informaticiens à toucher à l'élaboration des outils d'aide au diagnostique et détection. Pour atteindre cet objectif, plusieurs laboratoires de recherches se sont intéressés à la proposition de nouveaux algorithmes et méthodes touchants à la détection et au diagnostique. Dans le chapitre 2 on présente l'état de l'art lié à ces aspects.

Chapitre 2

Etat de l'art

Le dépistage ainsi que l'apparition de la mammographie dans le domaine médical ont augmenté le taux de détection des cancers de sein à un stade précoce ce qui permet de réduire le taux énorme de mortalité chez les femmes. Dans ce cadre, des études ont montré le dépistage permet de réduire le taux de mortalité de 40% chez les femmes âgées entre 50 et 60 ans [10]. L'automatisation et les progrès technologiques ont encouragé les chercheurs à élaborer des outils d'aide basés sur les systèmes CAD et CADx permettant de rendre encore plus efficace l'analyse mammographique.

Dans ce chapitre, on donne l'état de l'art des travaux menés dans le cadre de ces deux types de systèmes après avoir introduit quelques notions de base concernant l'analyse mammographique. On présente, à la fin, le flôt de travail proposé.

1. Analyse mammographique

L'analyse des clichés mammographiques est une succession de plusieurs étapes depuis l'acquisition jusqu'à la décision finale. Ces étapes sont étroitement liées ; la réussite d'une étape dépend de la réussite de l'étape précédente. Dans cette partie, on présente la démarche générale qu'on doit utiliser dans l'analyse des clichés mammographiques et on donne quelques flôts utilisés dans ce sens.

Le flôt général d'analyse mammographique, comme montre la figure 2.1, est composé de deux étapes : la détection et le diagnostic des anomalies mammographiques (particulièrement les masses). Le point commun entre ces deux étapes est qu'elles essaient d'identifier les cas cancéreux, tandis que la différence entre elles est que la détection commence par une région qui peut avoir une anomalie alors que l'étape de diagnostic doit commencer par une région contenant l'anomalie.

Il faut distinguer entre le système de détection (CAD) et celui de diagnostic (CADx) : le premier à pour objectif la discrimination régions suspectes/régions normales quant à la deuxième, il vise la différenciation entre régions suspectes malignes et régions suspectes bénignes.

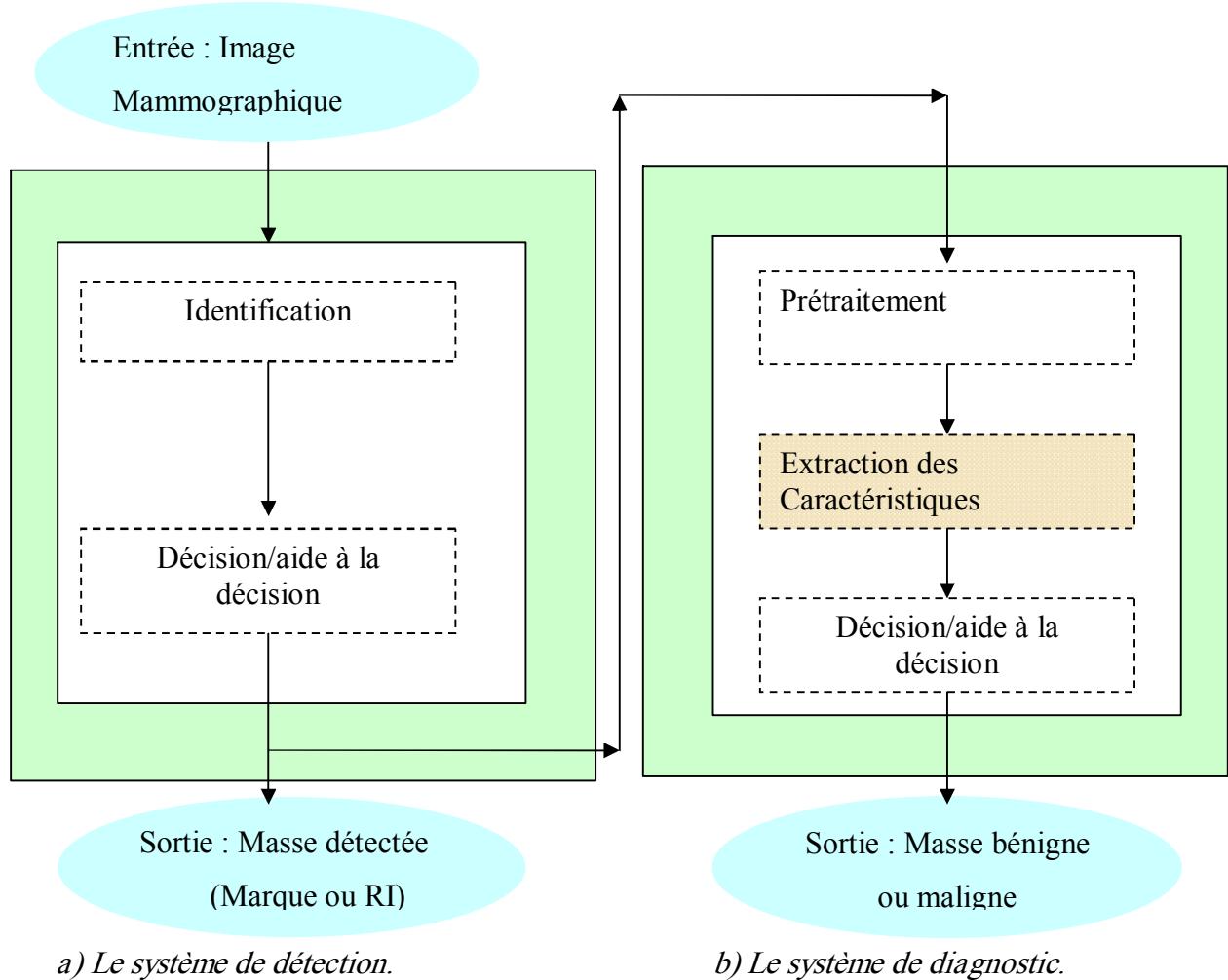
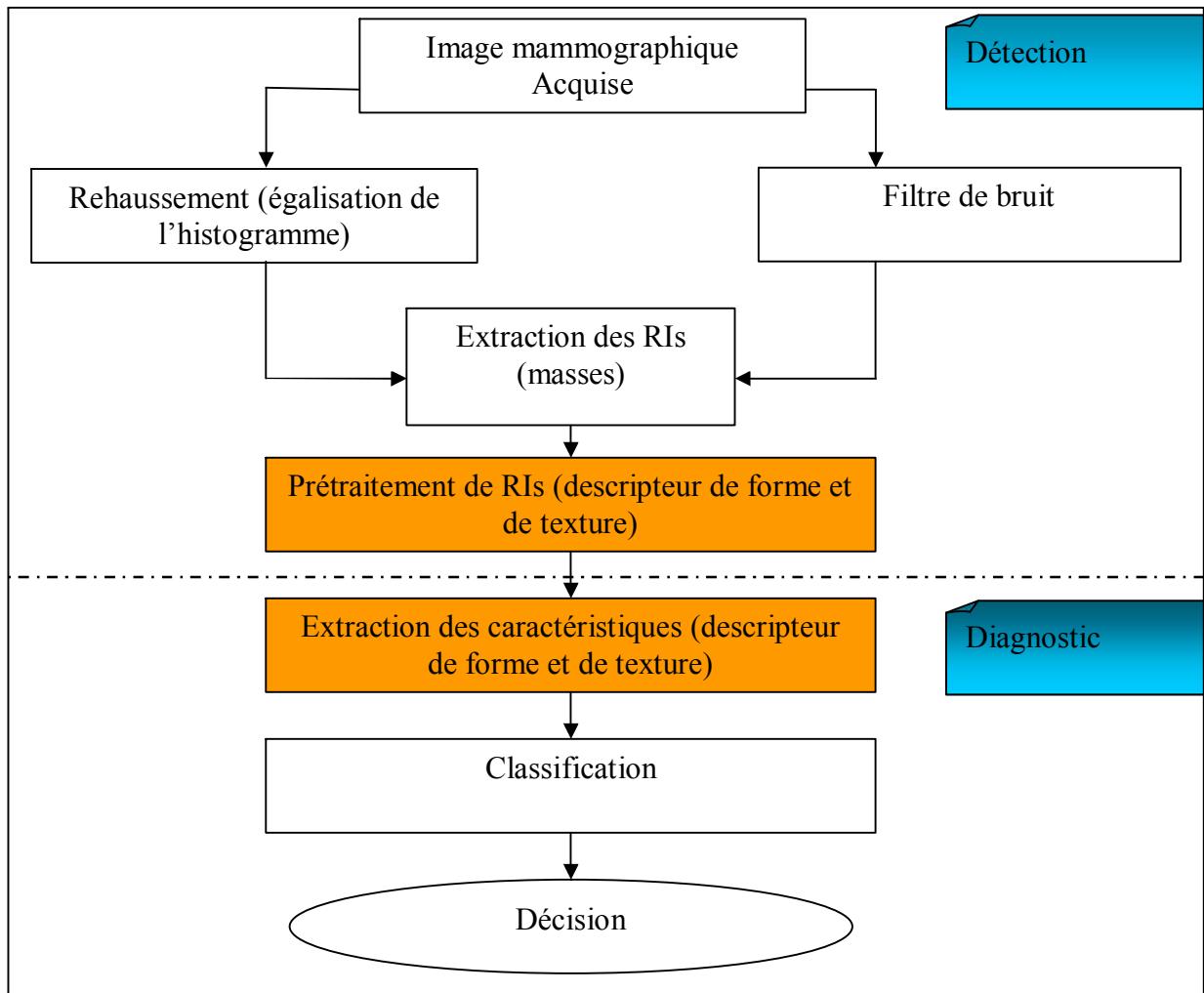


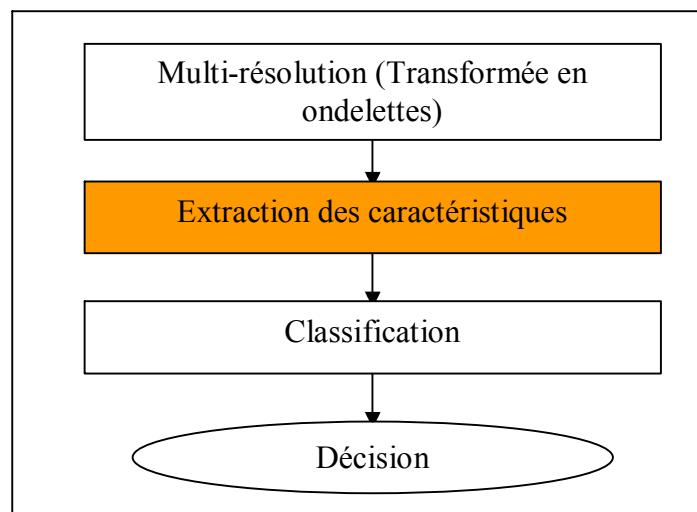
Fig. 2.1. Flot général d'analyse : (a) détection et (b) diagnostic.

Dans la littérature, on trouve plusieurs flots d'analyse mammographique [67], [68], [69] dédiés soit à la détection des masses soit au diagnostic. Par exemple, dans leur travail de d'analyse, S. Singh et al [67] ont utilisé un flot (figure 2.2) commençant par une étape de prétraitement afin d'améliorer la qualité d'image. La phase de détection de masses est basée sur la segmentation à travers la méthode de la croissance de régions. Pour éliminer les régions qui ne représentent pas des masses ils ont utilisé un ensemble des caractéristiques basé sur la forme et la texture.

Concernant la phase de diagnostic un vecteur caractéristique, basé sur la forme et la texture, a été utilisé pour permettre la prise de décision afin de donner finalement le type de la masse (bénigne ou maligne).

**Fig. 2.2.** Flôt de détection/diagnostic – [67].

Concernant les mammographies de haute résolution, la détection de la masse entraîne un temps de calcul important. Pour cette raison, F. Djidet et F. Boumghar [68] ont utilisé une méthode basée sur la multi-résolution dans leur analyse de masses (voir la figure 2.3).

**Fig. 2.3.** Flôt de détection des masses stellaires – [68].

Elles ont utilisé la transformée par ondelettes pour décomposer l'image en une pyramide d'images de résolution réduite de moitié. À partir de cette représentation, ils ont réalisé l'extraction des paramètres caractérisant à la fois le centre de la masse et les spicules. Ces paramètres sont utilisés comme une entrée à un classifieur neuronal (Perceptron multicouches MLP).

Dans [69], N. Székely et al ont proposé un système des masses mammographiques. Une méthode de segmentation globale est appliquée pour trouver les régions d'intérêts (RIs) dans l'image. Pour ces régions ils ont calculé un vecteur de caractéristiques basé sur la texture. Ce vecteur servira comme entrée à un classifieur basé sur l'arbre de décision. Dans la phase de diagnostic ils ont appliqué, dans un premier temps, une segmentation locale (combinaison de la méthode de binarisation et celle d'histogramme de Bézier). Ils ont extrait, dans un deuxième temps, des caractéristiques de forme et de texture. En se basant sur ces caractéristiques, la décision finale est réalisée pour indiquer s'il s'agit d'une masse maligne ou bénigne.

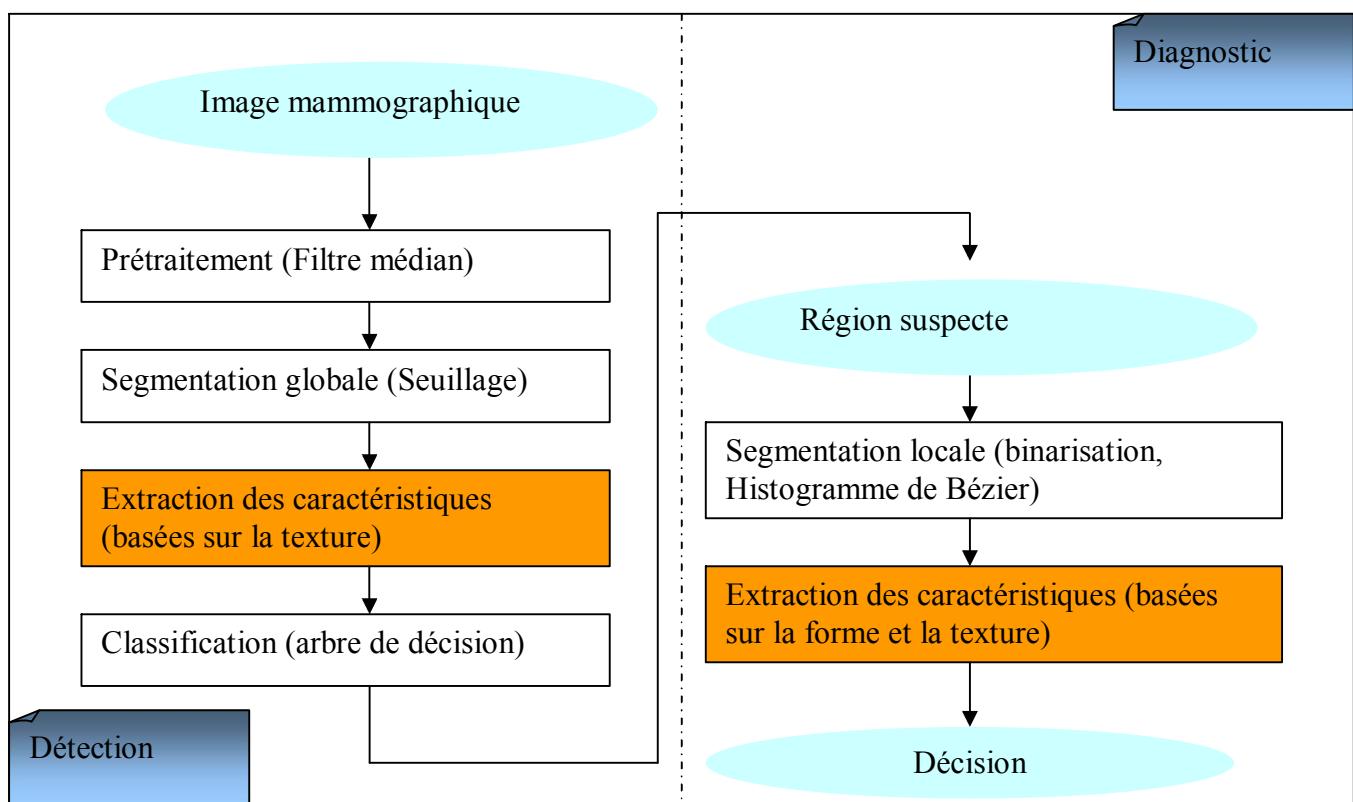


Fig. 2.4. Flot de détection/ diagnostic - [69].

Techniquement, les flots d'analyses se basent généralement sur des méthodes de segmentation, des méthodes de description et des méthodes de classification. La description des masses, à travers l'extraction des caractéristiques, joue deux rôles principaux dans l'analyse mammographique : elle

permet le raffinement de la détection en éliminant les régions qui paraient normales, et la discrimination régions malignes/régions bénignes.

À partir des travaux présentés ci-dessus, on voit l'utilité et l'importance de la phase d'extraction des caractéristiques dans la détection et le diagnostic des masses. C'est ce qui nous amène à se concentrer sur cette phase et à développer ainsi une méthode basée sur la forme.

2. Détection assistée par ordinateur

L'étape de détection de masses fait l'objectif de plusieurs travaux de recherche vu son importance dans l'analyse mammographique. Elle aide les radiologues à identifier les régions suspectes dans les clichés mammographiques.

Dans cette section, on présente l'objectif de cette étape de détection ainsi qu'un aperçu sur quelques méthodes utilisées.

2.1. Objectif de CAD

Les systèmes CAD sont utilisés pour assister le radiologue à identifier les régions suspectes d'une image mammographique. Le radiologue garde la responsabilité du diagnostic final. Les systèmes CAD permettent aussi de détecter davantage de cancers. En effet, l'interprétation des mammographies de dépistage est un véritable défi, compte tenu du grand nombre de clichés mammographiques à lire par le radiologue pour détecter un petit nombre de cancers (3 à 10 cancers pour 1000 patientes dépistées). Il en résulte quelques cancers parfois manqués.

Fort heureusement, les avancées technologiques récentes en mammographie digitale (amélioration des contrastes, de la netteté et de la qualité des images), ainsi que les progrès des systèmes CAD, devraient apporter une aide appréciable au radiologue et permettraient de détecter davantage les cancers tout en réduisant le nombre de patientes rappelées pour des examens complémentaires.

Les chercheurs dans le domaine médical font appel à des critères d'évaluation de performances, la sensibilité et la spécificité (voir chapitre 4, section 2), pour juger l'exactitude des résultats et la précision des systèmes CAD.

$$\text{Sensibilité} = \frac{\text{Nombre de vrai positives (VP)}}{\text{Nombre des masses suspectes}} \quad (\text{Eq. 2.1.})$$

Avec

VP : nombre de vraie classification des masses suspectes.

La sensibilité est la capacité de la technique à découvrir les positifs (cas suspects) alors que la spécificité est la capacité de la technique à découvrir les négatifs (cas normaux)

$$\text{Spécificité} = \frac{\text{Nombre de vrai négatives(VN)}}{\text{Nombre de régions normales}} \quad (\text{Eq. 2.2.})$$

Avec

VN : nombre de vraie classification des régions normales.

Donc on peut dire que le principal objectif du CAD est d'améliorer le taux de sensibilité en aidant des radiologues pour détecter la masse suspecte qui pourrait autrement avoir été manqué.

2.2. Détection de masses

Plusieurs travaux ont traité le problème de la détection. La plupart des méthodes de détection des masses se composent de deux étapes :

- 1) L'identification des régions
- 2) La discrimination entre régions suspectes et régions normales.

Dans l'étape d'identification, on essaie de détecter les régions d'intérêt. Après ceci, les régions sont classifiées en deux groupes : masses ou tissus normaux. On présentera ci-dessous dans le détail ces deux étapes en citant quelques travaux relatifs à chacune d'entre elles.

2.2.1. Identification des régions

Le but de cette étape de détection est de séparer les régions qui contiennent des tumeurs (bénignes et malignes) des régions normales. Elle est conçue pour avoir une sensibilité très élevée et un grand nombre de faux positifs (FP) est acceptable puisqu'on s'attend à ce qu'ils soient enlevés dans la deuxième étape.

Les masses mammographiques avec les marges stellaires ont une très haute probabilité de malignité (voir chapitre 1 section 3.2.1) et donc plusieurs méthodes ont été développées spécifiquement pour la détection de ce type de masses. Parmi les approches développées certaines approches sont basées sur la détection de la masse centrale de la région stellaire, d'autres sont basées sur la détection de spicules. Dans ce sens, W.P. Kegelmeyer et al [25] ont développé une méthode

utilisant un ensemble de cinq caractéristiques. Ils ont utilisé l'écart type d'un histogramme de l'orientation du contour **ALOE** (Analysis of Local Oriented Edges) et la sortie de quatre filtres spatiaux. La raison d'utilisation de la caractéristique ALOE est que le tissu normal a des orientations dans une direction particulière alors que, dans les régions suspectes qui contiennent des masses stellaires, les lignes radiales du contour sont orientées dans de nombreuses directions. Pour détecter cette différence, ils ont calculé des orientations de contour afin de produire un histogramme. Cette idée est représentée dans la figure 2.5. La caractéristique ALOE est définie comme l'écart-type du nombre d'éléments de chaque classe de l'histogramme et est décrite par l'équation ci-dessous :

$$ALOE(\partial_{ij}) = \sqrt{\frac{\sum_{n=0}^{255} (hist_{ij}(n) - \overline{hist(i,j)})^2}{255}} \quad (\text{Eq. 2.3.})$$

Avec

$hist_{ij}$: est l'histogramme d'orientations de contour dans une fenêtre autour de pixel localisé à (i, j).

$\overline{hist(i,j)}$: est la hauteur moyenne de l'histogramme $hist_{ij}$.

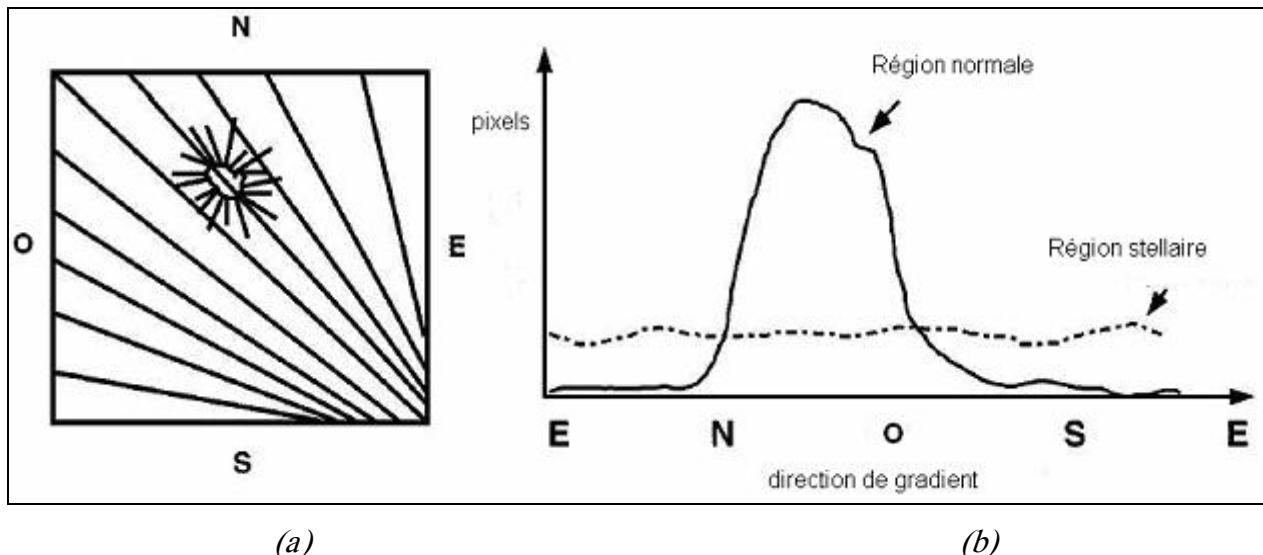


Fig. 2.5. (a) Les directions de rayons d'une masse stellaire sont différentes des directions de marquages linéaires normaux et (b) l'écart type de l'histogramme de l'orientation de gradient différencie la région stellaire de celle normale- [22].

La transformée de Hough a été utilisée dans plusieurs travaux mammographiques pour l'objectif de détection des masses stellaires ou entourées. Par exemple B.R. Groshong et W.P. Kegelmeyer [33], ont utilisé la transformée de Hough cercle pour la détection des masses entourées : Tout d'abord, ils ont

fait la détection de contour à l'aide de l'opérateur de Canny et ont sélectionné un sous-ensemble des contours en se basant sur la longueur et l'intensité. Ce sous-ensemble de contours était l'entrée à la transformée de Hough cercle. Deux caractéristiques ont été extraites du domaine de Hough afin de les classer en masse ou tissu normal.

Dans [35], W. Zhang et al ont montré que la présence de masses stellaires est liée aux changements dans la texture locale de la mammographie. Ils ont proposé qu'un tel changement pourrait être détecté dans le domaine de Hough. Ils ont découpé l'image en plusieurs régions et ont calculé les transformées de Hough pour chaque région. Ils ont utilisé un seuillage pour détecter les changements de textures locaux afin de déterminer la présence ou l'absence d'une masse stellaire.

L'inconvénient majeur de méthodes citées précédemment est que, si la résolution de l'image est grande, elles entraînent des temps de calcul prohibitifs. Pour cette raison, d'autres approches [30], [68] ont proposé des méthodes de détection des masses stellaires à partir d'une représentation en multi résolution. Dans ce contexte, les ondelettes ont été utilisées par Liu et al [30] pour décomposer l'image selon plusieurs résolutions, et pour chaque résolution extraire des informations directionnelles afin de détecter des spicules.

Vu que les masses ne sont pas toutes stellaires, la détection des autres types de masses est également importante. Dans ce cadre, H.D. Li et al [27] ont développé un système à deux phases : dans la première, un seuillage adaptatif de niveau de gris a été utilisé afin d'obtenir une segmentation initiale des régions suspectes ; dans la deuxième étape un arbre de décision flou a été développé pour classer les régions identifiées en deux groupes (masse ou tissus normal). La décision a été faite à l'aide d'un vecteur caractéristique basé sur la forme et sur la texture.

Le filtre de différence Gaussian (Difference of Gauss), qui est un filtre passe-bande, est parmi les techniques de segmentation les plus figurant dans la littérature [29] utilisé pour la détection des masses dans une image mammographique.

2.2.2. Discrimination entre les régions

La classification, qui est la deuxième étape du processus de détection, permet de séparer les régions représentatives des masses de celles normales. Elle permet également de réduire le nombre de fausses détections qui ont été produites à la fin de la première étape.

Dans le domaine médical, les radiologues utilisent plusieurs propriétés de l'image pour discriminer entre les masses et les tissus normaux. Les chercheurs ont essayé d'imiter ce processus, par l'extraction de plusieurs caractéristiques de texture ou de forme à partir des régions suspectes et

l'utilisation de ces caractéristiques par la suite dans la phase de classification. Dans ce cadre, B.Sahiner et al [39] ont exploité un ensemble des caractéristiques basé sur la texture dans leur système de classification. Dans un autre travail, D. Wei et al [42] ont développé un système utilisant les caractéristique de texture calculées à partir de la matrice de cooccurrence et l'analyse du discriminant linéaire (Linear Discriminant Analysis) pour la classification. Dans un travail plus récent, D. Wei et al. [19] a testé l'usage de l'analyse en multi-résolution globale et locale afin d'extraire des caractéristiques de texture.

La détection de masses, à travers les méthodes citées ci-dessus, permet d'aider le radiologue à identifier les masses à traiter. Ce qui reste par la suite est de voir s'il s'agit d'une tumeur bénigne ou maligne d'où notre objectif dans ce travail qui permet d'aider le radiologue lors de la phase de diagnostic.

3. Diagnostic assistée par ordinateur

Comme pour le cas de la détection, le diagnostic joue un rôle important dans le processus d'analyse mammographique. Elle est utile dans la phase de décision et permet de séparer les tumeurs malignes de celles bénignes afin d'éviter plusieurs traitements à réaliser par le radiologue qui, souvent, s'intéresse seulement aux tumeurs malignes.

Dans cette partie on donne les objectifs des systèmes CADx ainsi que les différentes techniques utilisées en littérature.

3.1. Objectif de CADx

Les systèmes CADx sont utilisés pour aider le radiologue à identifier les régions malignes et bénignes d'un cliché mammographique. La diversité les signes des cas anormaux (bénignes et malignes) dans les images mammographiques sont variés, rend le diagnostic difficile. Depuis des années, des campagnes de dépistage systématique ont été organisées dans les pays développés. Ces derniers engendrent un flux énorme de clichés mammographiques ce qui met les radiologues experts face à un problème d'imprécision d'analyse. Toutes ces limites, mènent à des traitements (biopsies chirurgicales) inutiles. Afin d'éviter ces derniers, un système de diagnostic séparant les cas malignes de celles bénignes est plus que nécessaire.

Pour évaluer la performance des systèmes CADx les chercheurs utilisent la sensibilité et la spécificité données par les équations suivantes:

$$\text{Sensibilité} = \frac{\text{Nombre de vrai positives}(VP)}{\text{Nombre des masses malignes}} \quad (\text{Eq. 2.4.})$$

Avec

VP: nombre de vraie classification des masses malignes.

$$\text{Spécificité} = \frac{\text{Nombre de vrai négatives}(VN)}{\text{Nombre de masses bénignes}} \quad (\text{Eq. 2.5.})$$

Avec

VN: nombre de vraie classification des masses bénignes.

L'objectif principal du CADx est d'améliorer le taux de spécificité (taux d'identification des masses bénignes), sans diminuer le taux de sensibilité (taux de détection des masses malignes).

3.2. Diagnostic des masses

Les systèmes de diagnostic ont comme entrée une région d'intérêt (RI) contenant les masses suspectes. Ils y agissent en trois reprises :

- 1) Segmentation de masse dans la RI.
- 2) Extraction des caractéristiques.
- 3) Classification.

Dans l'étape de segmentation, la masse est séparée du tissu normal. En se basant sur les propriétés de la forme, des caractéristiques discriminantes sont extraites. A la fin, les masses sont classifiées comme malignes ou bénignes. Un état de l'art de ces différentes étapes est présenté par la suite.

3.2.1. Segmentation

L'étape de la segmentation sert à l'identification des masses à partir des clichés mammographiques. Elle peut être manuelle, semi automatique, ou encore automatique. Elle est extrêmement importante puisque la réussite d'un système d'analyse mammographique dépend de cette étape. En imagerie médicale, et particulièrement en mammographie, il y'a trois types de méthodes de segmentation utilisées : les méthodes basées sur la région, les méthodes basées sur le contour et celles basées sur le regroupement. Dans ce qui suit, on donne les principaux travaux utilisant ces méthodes de segmentation.

3.2.1.1. Méthodes basées sur la région

La méthode de croissance de région est parmi les méthodes utilisées fréquemment dans la segmentation des masses mammographiques. Par exemple, Z. Huo et al [23] ont développé une approche semi-automatique de croissance de région dans laquelle le pas de croissance a été calculé automatiquement suite à un choix manuel du point de départ (seed point). Une approche différente a été proposée par D. Guliato et al [31], avec une implémentation d'une version de l'algorithme de croissance de région basée sur la logique floue.

L'inconvénient majeur de la croissance de région réside dans le choix du point de départ puisqu'un mauvais choix mène à un mauvais résultat. Pour surmonter ce problème, cet algorithme a été modifié de façon à identifier l'ensemble optimal des points de départ (seeds points) [38].

3.2.1.2. Méthodes basées sur le contour

Concernant la segmentation basée sur le contour, il y a seulement un nombre limité de travaux vu la difficulté qui réside dans la distinction entre la masse et le tissu normal.

Les algorithmes typiques pour trouver les contours sont basés sur le filtrage de l'image pour rehausser les contours pertinents avant l'étape de la détection. Par exemple, dans le but d'améliorer le contraste, N.Petrick et al [37] ont proposé une technique de segmentation basée sur la méthode de Gaussian-Laplacien. Une approche différente a été proposée par H. Kobatake et Y. Yoshinaga [26] qui part d'une sous image qui contient une masse. L'algorithme cherche les lignes radiales de la masse stellaire et se résume dans trois étapes : i) extraction des lignes radiales par skeletonisation, ii) application de la transformée de Hough ligne modifiée pour l'extraction des lignes radiales qui passent par le centre de la masse, iii) raffinement et sélection des masse dont le nombre de ligne dépasse un seuil fixé d'avance.

Dans les approches récentes, l'information de contour a été utilisée pour raffiner les résultats initiaux de la segmentation. Par exemple, B. Sahiner et al [40] ont utilisé le modèle de contour actif (Snake) comme une étape finale de leur approche.

En mammographie, l'utilisation des méthodes de détection de contour ne peut pas être appliquée directement sur une image mammographique mais sur des régions d'intérêts (RIs) prétraitées.

3.2.1.3. Méthodes basées sur le regroupement de pixels (clustering)

Les méthodes par regroupement (Clustering) servent à diviser une région d'intérêt en plusieurs classes. Plusieurs travaux, utilisent les méthodes par classification pour la segmentation des clichés mammographiques.

Un traditionnel algorithme est l'algorithme C-Means Flou (FCM), qui a été utilisé dans différents objectifs dans les travaux de Velthuizen [28], Chen et Lee [44]. Pendant que Velthuizen l'a utilisé pour grouper les pixels avec valeurs similaires du niveau de gris dans l'image originale, Chen et al l'ont utilisé sur l'ensemble de caractéristiques locales extraites de l'application.

La méthode de seuillage est une autre approche qui a été fréquemment utilisée en mammographie. Elle forme un cas spécial des méthodes de segmentation par classification où seulement deux groupes sont considérés. Dans le travail de Matsubara et al [32], en se basant sur l'analyse de l'histogramme, différentes valeurs de seuil ont été considérées. Plus récemment, Mudigonda et al [34] ont utilisé un algorithme de seuillage à plusieurs niveaux afin de détecter les contours fermés. Cet algorithme peut être considéré comme une technique de croissance de région où dans chaque itération, les voisins qui possèdent des niveaux de gris similaires sont groupés.

Dans d'autres cas, le seuillage n'est pas appliqué à l'image mammographique directement, mais plutôt à une version rehaussée de l'image originelle. Par exemple, C. Varela et al. [24] ont appliqué « un filtre Iris » afin de rehausser des masses arrondies. En utilisant un seuillage adaptatif, ils ont fait la détection des masses.

3.2.2. Extraction des caractéristiques

L'étape d'extraction des caractéristiques est sans doute l'étape la plus importante dans la phase de diagnostic vu son rôle discriminatif permettant la différenciation entre les masses bénignes et celles malignes. Elle est fortement liée à l'étape de détection et sa réussite fait réussir la phase de détection. En littérature, on trouve deux méthodes qui sont fréquemment utilisées : les méthodes basées sur la texture et celles basées sur la forme.

L'identification de la tumeur selon l'avis des radiologues se base sur l'analyse de la forme de la région suspecte. On se propose, dans le cadre de ce travail, de donner un panorama de techniques et de méthodes permettant l'analyse selon la forme.

3.2.2.1. Caractéristiques de forme

La forme des régions suspectes se caractérise par des propriétés de malignité dépendant du contenu (ou encore la région) et par des propriétés de malignité selon le contour (ou frontière). Pour

cette raison les caractéristiques qui forment le descripteur de forme (ou vecteur de caractéristiques) sont divisées en deux catégories : les caractéristiques qui décrivent la région et celles permettant la description du contour. Dans la littérature, on trouve plusieurs méthodes qui ont été exploitées pour l'objectif de la description. On en résume dans le tableau 2.1 les plus importantes.

Caractéristiques	Rôle
Les caractéristiques de forme simple.	Description de la régularité de la région.
Les caractéristiques de l'enveloppe convexe.	Description de la convexité de la région.
La méthode RDM (Radial Distance Mesure).	Mesure de la régularité du contour.
Les moments.	Description de la région/contour.
L'analyse de fractale.	Description de la rugosité du contour.
La signature des contours.	Description du contour.

Tab. 2.1. Méthodes d'extraction des caractéristiques basées sur la forme.

Ces caractéristiques seront détaillées plus tard.

a) Caractéristiques basées sur la région

Les propriétés de malignité et de bénignité peuvent être tirées à partir de la région. Dans ce cadre, plusieurs travaux ont traité les propriétés d'irrégularité de la région lors du choix des méthodes de description.

- *Caractéristiques simples :*

Pour décrire la régularité de la forme d'une masse mammographique, plusieurs travaux font appel à la caractéristique de **circularité (C)** (ou encore **la compacité**). Par exemple A. Retico et al [53], Pohlman et al [46], Kilday et al [45] et A.V. Alvarenga et al [47], ont utilisé pour leurs descripteurs de forme des caractéristiques géométriques simples telles que **la circularité (C)** afin de distinguer les masses mammographiques circulaires de celles irrégulières. Pour évaluer la performance de leurs caractéristiques, A.V. Alvarenga et al [47] ont montré que la circularité, à part sa simplicité, est la deuxième meilleure caractéristique parmi un ensemble de caractéristiques utilisé pour la description.

Pour décrire les formes allongées qui représentent des tumeurs malignes, la caractéristique d'**'excentricité (Exc)** a été utilisée dans le travail de U. Bottigli et al [48]. Elle donne un indice

d'elongation de la forme qui peut aider à la discrimination entre les masses malignes et celles bénignes.

- *L'enveloppe convexe :*

La notion de la convexité peut être employée pour détecter l'irrégularité des régions. Pour cette raison on trouve dans d'autres travaux [47], [53] plusieurs méthodes qui s'appuient sur *l'enveloppe convexe* pour générer des caractéristiques de convexité discriminantes telles que *la convexité (CVX) et la valeur résiduelle normalisée (NRV)*. Ces dernières ont donné des performances acceptables dans le travail de A.V Alvarenga et al [47].

Les caractéristiques basées sur la région se caractérisent par la simplicité de calcul et aussi l'efficacité de différenciation entre les masses bénignes et celles malignes. Pour cette raison, dans notre travail, on exploite ces caractéristiques.

b) Caractéristiques basés sur le contour

Ces caractéristiques résument l'information du contour pour décrire les masses. En mammographie, plusieurs techniques ont été employées pour générer des caractéristiques discriminantes à partir du contour.

- *La signature des contours :*

La signature des contours sert à représenter un contour comme une fonction 1D (réelle ou complexe) afin de réduire la dimensionnalité des contours 2D correspondantes. En mammographie, la représentation du contour par signature est exploitée par plusieurs approches pour la description de la rugosité du contour. En effet, dans la littérature, on trouve plusieurs types de signatures dont les plus utilisées sont : la signature par le centroïde [46] et la signature par l'angle de la tangente [49].

La signature par le centroïde est basée sur la fonction de distance $d(t)$ [65] qui est exprimée par la distance euclidienne d des points de frontière de coordonnées $(x(t), y(t))$ du centre C de coordonnées (x_c, y_c) . La fonction de distance $d(t)$ est donnée par l'équation suivante :

$$d(t) = \sqrt{(x(t) - x_c)^2 + (y(t) - y_c)^2} \quad (\text{Eq. 2.6.})$$

Les figures 2.6 et 2.7 présentent deux signatures (figure 2.6 (b) et (figure 2.7 (b))) différentes des tumeurs bénignes et malignes (figure 2.6 (a) et (figure 2.7 (a))). En regardant la signature de la tumeur maligne, on constate qu'il y'a une grande différence par rapport à la signature de la tumeur bénigne. Cette différence peut être exploitée pour l'extraction des caractéristiques discriminantes.

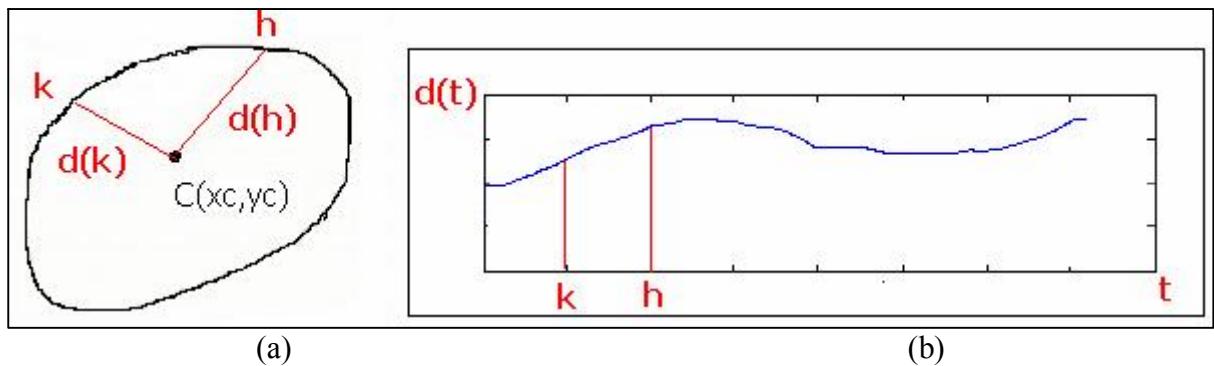


Fig. 2.6. (a) Une tumeur bénigne et (b) sa signature.

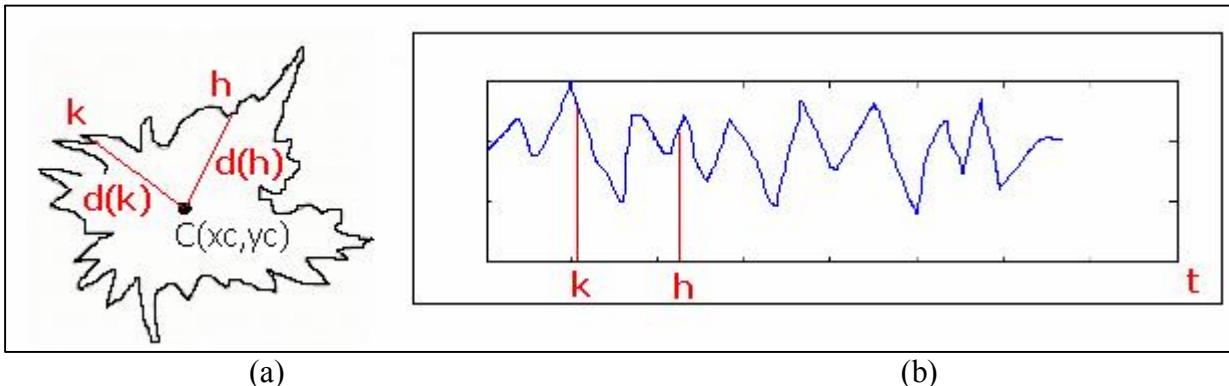


Fig. 2.7. (a) Une tumeur maligne et (b) sa signature.

En mammographie, la signature par le centroïde est utilisée dans plusieurs travaux de recherches tel que le travail de S. Pohlman et al [46] qui, à partir de la représentation du contour, font l'extraction d'un ensemble des caractéristiques et servent à la classification des masses en deux catégories (bénigne ou maligne).

L'avantage d'utilisation de la signature par le centroïde est qu'elle est simple à implémenter. Mais cette signature n'est pas invariante au changement de la dimension et à la rotation et par la suite il suffit de changer l'orientation ou la dimension du contour pour avoir une signature totalement différente.

D'autres travaux ont fait appel à une autre façon de représentation du contour, notée la signature par l'angle de tangente [49] (figure 2.8 (b)), pour extraire des caractéristiques qui peuvent distinguer les tumeurs bénignes de celles malignes.

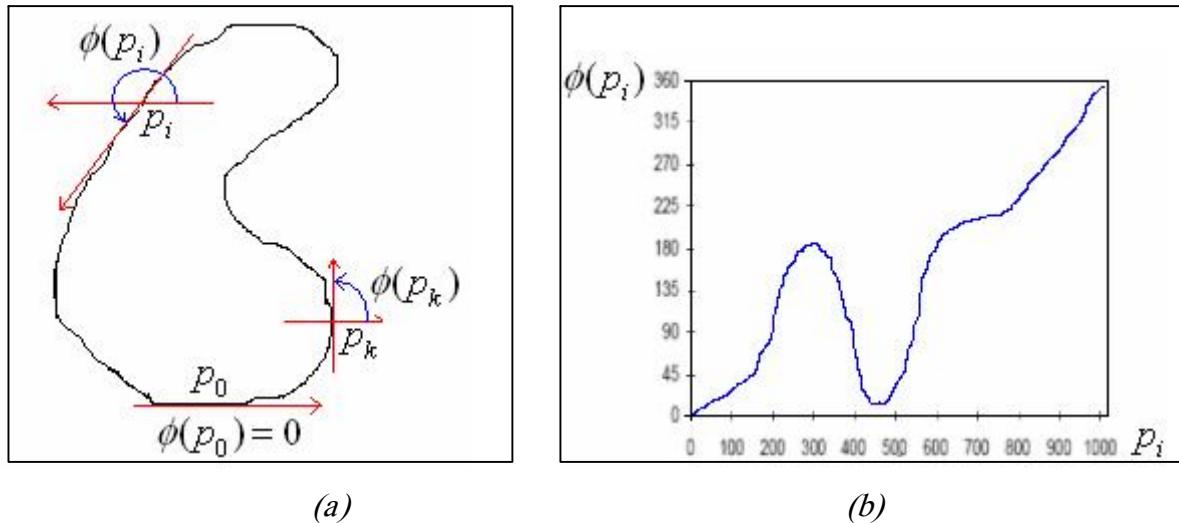


Fig. 2.8. Signature par l'angle ϕ de la tangente au point P_i et (b) la signature du contour-[49].

Le principe de cette signature est que pour chaque point P_i du contour on fait le calcul de l'angle ϕ à partir de la tangente à ce point (Figure 2.8 (a)). Après le calcul des angles sur toutes les tangentes on aura une représentation (figure 2.8 (b)) qui sera exploitée par la suite pour l'extraction des caractéristiques pouvant être utiles pour la description du contour de la masse. Dans ce cadre, Rangaraj et al [49], à partir de cette méthode, ont extrait deux caractéristiques : la première notée « *indice de spicule* » (SI_{TA}) permettant de mesurer le degré de rugosité du contour et la deuxième notée « *indice de convexité* » (IC_{TA}) représentant le degré de convexité du contour. Pour évaluer la performance de leur système, ils ont comparé les résultats fournis avec les résultats des travaux de Rangayyan et al. [50], [51] basé sur l'analyse fractale des contours. Ils ont montré qu'avec la signature par l'angle de la tangente, les caractéristiques proposées permettent une bonne discrimination entre les masses bénignes et celles malignes.

Le problème de ce type de signature réside dans la complexité temporelle car le calcul des angles, le passage à la représentation du contour et le calcul des caractéristiques demande beaucoup de traitement à faire. Elle dépend aussi de la rotation ; c'est-à-dire ; pour deux rotations du même contour on aura deux signatures différentes. Ces inconvénients nous découragent à utiliser la signature par l'angle de la tangente en tant qu'une méthode de description des masses mammographiques.

– *descripteur de Fourier :*

Une des techniques les plus prometteuses pour la description de masses mammographique est celle basée sur des *descripteurs de Fourier* [56], [57]. Cette méthode s'effectue en quatre étapes : i) une étape de représentation des contours en tant que nombre complexe, ii) une étape de calcul des

composants pour les N points du contours, iii) une phase de normalisation des composants et finalement iv) une phase de calcul d'une caractéristique pour la description des contours.

Partant de l'ensemble N des points du contour, nous pouvons regarder la région comme étant placée dans l'espace de représentation complexe, avec l'ordonnée étant l'axe imaginaire et l'abscisse étant l'axe réel. Puis, les coordonnées (x et y) de chaque point du contour à analyser peuvent être représentées en tant que nombres complexes ($x + jy$). L'ordre du contour peut alors être décrit comme ordre complexe Z_i :

$$Z_i = x_i + jy_i, \quad i=0,1,2,\dots,N-1 \quad (\text{Eq. 2.7.})$$

Les composants de Fourier (FDS) à utiliser dans la description du contour sont définis comme :

$$A(n) = \frac{1}{N} \sum_{i=0}^{N-1} Z_i e^{\left[\frac{-j2\pi ni}{N} \right]}, \quad n=0,1,\dots,N-1 \quad (\text{Eq. 2.8.})$$

Une normalisation de ces composants est calculée pour rendre ces quantités invariantes aux transformations affines. Les composants normalisés deviennent :

$$NFD(k) = \begin{cases} 0; & k=0 \\ A(k)/A(1); & k=1,2,\dots,N/2 \\ A(k+1)/A(1); & k=-1,-2,\dots,-N/2+1 \end{cases} \quad (\text{Eq. 2.9.})$$

Finalement la caractéristique notée **FF** (Fourier Fraction) est utilisée pour la discrimination entre les masses des contours réguliers et celles des contours irréguliers.

$$FF = \left[\sum_{k=-N/2+1}^{N/2} \|NFD(k)\| / \|k\| \right] / \sum_{k=-N/2}^{N/2} \|NFD(k)\| \quad (\text{Eq. 2.10.})$$

Dans les travaux [56], [57], on trouve des taux acceptables en exploitant la caractéristique FF. Dans le travail de L.Shen et al [56], l'utilisation de FF a donné un taux d'exactitude de l'ordre de 84.81%. Plus récemment, R. M. Rangayyan et al [57] ont trouvé un taux d'exactitude de l'ordre de 88.9% dans la distinction entre les contour stellaires et ceux circonscrits.

Malgré que les résultats trouvés soient acceptables, le problème de cette méthode réside dans la complexité temporelle à cause des traitements réalisés dans la représentation complexe du contour, le calcul des composants et le passage à la normalisation.

- **Dimension fractale (FD) :**

Le concept de la **dimension fractale (FD)** peut être exploité pour mesurer la complexité ou l'irrégularité de la frontière d'un objet. En mammographie, ce concept est utilisé pour caractériser la complexité des contours (2D) des masses, aussi bien que leurs signatures (1D). Les tumeurs cancéreuses montrent un certain aspect aléatoire lié à leur croissance, et sont en général irrégulières et complexes en terme de forme ; donc, l'analyse de fractale peut fournir une meilleure idée concernant leurs modèles en comparaison avec la géométrie euclidienne conventionnelle [55].

Dans leurs travail, Pohlman et al [46], ont obtenu un taux d'exactitude de classification de plus de 80% avec l'analyse de fractale des signatures des contours de masses. Dans un travail récent, Matsubara et al [52], ont obtenu un taux d'exactitude de 100% dans la classification de 13 masses. La méthode a exigé le calcul d'une série de valeurs de FD pour plusieurs contours d'une masse donnée obtenue par seuillage à plusieurs niveaux. La variation de FD a été employée pour classer les masses en deux catégories (bénignes ou malignes).

Deux méthodes ont été décrites dans la littérature pour estimer la dimension de fractale: i) La méthode de boîte (**box-counting**) et ii) la méthode de règle (**ruler**) [51], [55]. Dans un travail plus récent, pour l'objectif d'évaluer ces deux méthodes, Nguyen et Rangayyan [51] ont estimé le calcul de FD d'un ensemble de 111 contours (2D) des masses aussi bien que leurs signatures 1D. La meilleure performance de classification avec un taux d'exactitude de l'ordre de 89% a été obtenue avec la méthode de règle appliquée aux signatures 1D des contours.

Malgré qu'elle puisse être utile pour la description des masses, l'analyse fractale présente certaines limites qui nous découragent à l'utiliser dans notre travail. En regardant la figure 2.9, on constate que les valeurs retenues par le FD sont très proches et par la suite la séparation entre les cas malignes et ceux bénignes est difficile. Aussi les valeurs retournées ne sont pas normalisées c'est-à-dire elles ne sont pas limitées à un intervalle bien déterminé.

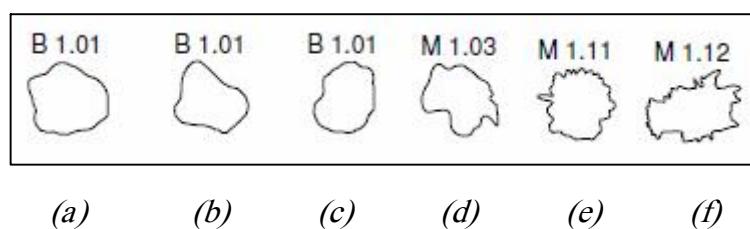


Fig. 2.9. Calcul de la caractéristique FD pour 6 cas de masses dont 3 sont bénignes (a, b et c) et 3 sont maligne (d, e et f)-[51].

- ***Les moments :***

Pour étudier l'irrégularité des masses mammographiques, plusieurs travaux ont fait appel aux concepts des ***moments du contour*** [56], [41]. Les moments p+q sont définis par :

$$m_{pq}(I) = \sum_0^{N-1} \sum_0^{M-1} x^p y^q I(x, y) \quad (\text{Eq. 2.11.})$$

Où I est l'image au sein de laquelle la forme à étudier a été isolée, (M, N) sont les paramètres de taille d'image et (x, y) les coordonnées d'un pixel dans l'image I. En utilisant l'équation 11 et les coordonnées du centre ($\bar{x} = \frac{m_{10}}{m_{00}}$ et $\bar{y} = \frac{m_{01}}{m_{00}}$ correspondent à la position du centre de gravité de la forme) les moments centrés peuvent être ainsi:

$$\mu_{pq}(I) = \sum_0^{N-1} \sum_0^{M-1} (x - \bar{x})^p (y - \bar{y})^q I(x, y) \quad (\text{Eq. 2.12.})$$

Les moments centrés sont basés sur le calcul des distances euclidiennes des pixels de la région et le point central de la forme. Les différentes valeurs de μ_{pq} permettent de caractériser une forme : par exemple la valeur μ_{01} mesure l'allongement d'une forme.

Les formes ou les contours de deux objets sont alors comparables grâce à leurs moments respectifs. Dans ce cadre, HU [58] a dérivé un ensemble de sept caractéristiques, notées les moments invariants, qui sont à base des moments centrés de second et de troisième ordre. Ces caractéristiques sont définies comme suit:

$$\phi_1 = \frac{1}{\mu_{00}^2} (\mu_{20} + \mu_{02}) \quad (\text{Eq. 2.13.})$$

$$\phi_2 = \frac{1}{\mu_{00}^4} ((\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2) \quad (\text{Eq. 2.14.})$$

$$\phi_3 = \frac{1}{\mu_{00}^5} ((\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2) \quad (\text{Eq. 2.15.})$$

$$\phi_4 = \frac{1}{\mu_{00}^5} ((\mu_{30} - \mu_{12})^2 + (\mu_{21} - \mu_{03})^2) \quad (\text{Eq. 2.16.})$$

$$\begin{aligned} \phi_5 = & \frac{1}{\mu_{00}^{10}} \{ (\mu_{30} - 3\mu_{12})(\mu_{30} - \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{30} + \mu_{21})^2] \\ & - (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[(\mu_{03} + \mu_{21})^2 - 3(\mu_{30} + \mu_{12})^2] \} \end{aligned} \quad (\text{Eq. 2.17.})$$

$$\phi_6 = \frac{1}{\mu_{00}^7} \{ (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2] - 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \} \quad (\text{Eq. 2.18.})$$

$$\phi_7 = \frac{1}{\mu_{00}^{10}} \{ 3(\mu_{21} - \mu_{03}).(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{03} + \mu_{21})^2] \\ - (\mu_{30} - \mu_{12}).(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2] \}$$
(Eq. 2.19.)

Les moments invariants de HU figurent dans plusieurs études pour différents objectifs ; elles ont été exploitées dans le travail de Rangayyan et al [57] afin de distinguer entre les contours stellaires et ceux circonscrits alors que dans un travail récent de N. Székely et al [61] elles ont été utilisées pour l'objectif de détection des masses.

Dans une autre étude, L. Shen et al [56] ont utilisé une représentation unidimensionnelle du contour afin d'estimer les moments. En se basant sur l'ensemble des coordonnées ($x(i)$ et $y(i)$) des points du contour et sur les distances euclidienne $Z(i)$ entre le point central et les points du contour, le moment d'ordre p a été défini ainsi :

$$m_p = \frac{1}{N} \sum_{i=1}^N [Z(i)]^p$$
(Eq. 2.20.)

Afin d'avoir des caractéristiques invariantes, une normalisation des moments est faite :

$$M_p = \frac{1}{N} \sum_{i=1}^N [Z(i) - m_1]^p$$
(Eq. 2.21.)

Lors de la description des masses stellaires, trois caractéristiques (F_1, F_2, F_3) ont été générées à partir du moment M_p normalisé :

$$F_1 = \frac{1}{m_1} \sqrt{\frac{1}{N} \sum_{i=1}^N (Z_i - m_1)^2}$$
(Eq. 2.22.)

$$F_2 = \frac{1}{m_1} \left[\sqrt{\frac{1}{N} \sum_{i=1}^N (Z_i - m_1)^3} \right]^{\frac{1}{3}}$$
(Eq. 2.23.)

$$F_3 = \frac{1}{m_1} \left[\sqrt{\frac{1}{N} \sum_{i=1}^N (Z_i - m_1)^4} \right]^{\frac{1}{4}}$$
(Eq. 2.24.)

En se basant sur les caractéristiques définies ci-dessus, une autre caractéristique a été définie comme la différence entre les caractéristiques F_1 et F_3 :

$$MF_{1-3} = F_1 - F_3$$
(Eq. 2.25.)

L. Shen et al [56], ont montré que cette dernière permet une meilleure description des contours stellaires par rapport aux autres caractéristiques.

- ***RDM (Radial Distance Measure)*** :

Une autre technique similaire à la méthode des moments est ***la méthode RDM (Radial Distance Measure)*** qui a été utilisée pour l'extraction des caractéristiques permettant la description des contours stellaires. Elle est basée sur la mesure de la distance euclidienne $d(i)$ entre un point i du contour et le point central de la masse (on revient en détail à cette méthode dans le prochain chapitre). Dans [45], [53], [97], les auteurs ont fait appel à cette méthode vu qu'elle est la moins complexe par rapport aux autres techniques. Dans le travail de A. Retico et al [53] plusieurs caractéristiques telles que la distance moyenne d_{moy} , la déviation standard (*SDEV*) et la rugosité (*R*) sont extraites à partir de la méthode ***RDM***.

La méthode ***RDM*** possède plusieurs avantages par rapport aux autres méthodes, vues précédemment, qui nous poussent à l'utiliser pour résoudre notre problème. Elle est simple à implémenter, et est aussi invariante aux transformations affines et ne demande pas un prétraitement à appliquer sur le contour ce qui nous permet de décrire, d'une manière efficace, les contours pour n'importe quel position, dimension...

Dans notre travail on essaye de réduire encore la complexité temporelle de cette méthode en procédant par la minimisation du nombre des points du contour qui entrent dans le calcul des caractéristiques.

3.2.2.2. Caractéristiques combinées (texture et forme)

Il est parfois nécessaire de calculer des caractéristiques de texture [74] lorsque le contour de ces objets n'est pas suffisamment discriminant. Ce problème est plus complexe que la description à base du contour car la description doit prendre en compte tant l'information sur le contour que sur la texture interne des objets. Dans ce cadre, On trouve pas mal de travaux [45], [48], [56], [59] qui combinent les caractéristiques de forme et celles de texture afin de mieux représenter les masses et améliorer le taux d'exactitude de classification. Par exemple, J. Panchal et al [59], ont montré que l'utilisation des caractéristiques combinées basées sur la technique RDM (3 caractéristiques) et la matrice de co-occurrences (5 caractéristiques) donne un taux d'exactitude qui dépasse 78%. Aussi, dans le travail de U.Bottigli et al [48] un ensemble de caractéristiques de forme (RDM, Forme géométrique simple) et de caractéristiques de texture (la moyenne, l'entropie...), a été utilisé pour prendre en considération les trois types de propriétés d'une tumeur (le contour, la région et la texture).

Dans un travail plus récent, A.V. Alvarenga et al [59], ont montré que l'utilisation des caractéristiques combinées basées sur RDM (4 caractéristiques) et la matrice de co-occurrences (6 caractéristiques) augmente encore la performance (sensibilité et spécificité) et rend la caractérisation des masses mammographiques plus précise (voir le tableau 2.2).

<i>Caractéristiques</i>	<i>Sensibilité (%)</i>	<i>Spécificité (%)</i>
<i>Forme</i>	88.0	90.4
<i>Texture</i>	89.0	86.5
<i>Combinées</i>	92.0	94.2

Tab. 2.2. Performances obtenues avec des caractéristiques de forme, de texture et combinées [73].

3.2.3. Classification

Plusieurs techniques ont été utilisées pour la classification des cancers de sein. Parmi ces techniques, on cite les méthodes supervisées (*réseau de neurones*, *C-means*,...) et d'analyse du discriminant linéaire (*LDA*).

Dans le cadre d'analyse mammographique, on trouve plusieurs utilisations du système de classification (LDA) [45], [47], [57] pour l'objectif de la prise de décision en ce qui concerne la malignité ou la bénignité d'une tumeur mammographique. Dans le travail de Kilday [45], l'utilisation du classifieur LDA a abouti à un taux de sensibilité de l'ordre de 69%. Ce taux a augmenté à 95% dans un travail récent de Rangayyan et al [57]. Dans une étude d'évaluation de la performance des caractéristiques de forme, A. Alvarenga et al [47] ont trouvé un taux de sensibilité (88%) et de spécificité (90.4%).

Une autre forme de classification basée sur le modèle de réseau de neurones est celle de multicouche (*MLP*) qui a été utilisé dans plusieurs travaux mammographiques [48], [53]. Par exemple, dans le travail de U. Bottigli et al [48] une comparaison d'un certain système de classification des masses a été présentée. Les résultats trouvés, montrent que le classifieur MLP donne la meilleure performance par rapport aux deux autres classificateurs *KNN* et *SVM* avec un taux de sensibilité de l'ordre de 88%.

Kim et al [60] ont conçu un nouveau type de classifieur qui combine un modèle non supervisé et un modèle supervisé pour la classification des masses. Le modèle non supervisé est basé sur la

théorie de la résonance adaptative *ART* qui a groupé les masses en plusieurs classes séparées. Les classes ont été divisées en deux types: un qui contient seulement les masses malignes et un autre contient un mélange des masses malignes et bénignes. Quelques masses malignes sont séparées et ont été classées par ART et les masses bénignes et malignes les moins distinguables ont été classées par LDA.

On présente dans le tableau 2.3, quelques travaux réalisés pour l'objectif de diagnostic avec les résultats trouvés. Ces derniers paraissent être comparables entre eux. Cependant, réellement, on ne peut pas les comparer vu que les critères d'évaluation, les bases d'images utilisées et les types des classificateurs sont différents.

Travail	Sensibilité	Spécificité	Ac	Classifieur	Base utilisée
Rangayyan et al [57].	95%	--	--	LDA	Base locale (39 cas).
L. M. Bruce et al [43].	80%	--	--	LDA	Base locale (60 cas).
A. Retico et al [53].	78,1%	79,1%	--	MLP	Base locale (226 cas avec 109 malignes et 117 bénignes).
R. Feng Chang et al [64].	88,89%	92,5%	--	SVM	Base locale (210 cas avec 90 cas malignes et 120 cas bénignes).
U. Bottigli et al [48].	88%	--	--	MLP	Base CALMA (320cas).
Kilday [45].	69%	--	--	LDA	Base locale (82 cas).
A. Alvarenga et al [47]	88,0%	90,4	88,8%	LDA	Base locale (152 cas).

Tab. 2.3. Résultat de test de quelques travaux utilisant des classificateurs supervisés.

4. Flôt de diagnostic mammographique proposé

L'analyse mammographique se base sur trois étapes fortement liées quelque soit le flôt à utiliser: l'étape de segmentation, l'étape d'extraction de caractéristiques et celle de classification.

Dans ce travail, on présente une démarche de diagnostic contenant trois parties (Fig. 2.10). La première consiste en l'identification des masses. Quant à la deuxième, elle aborde la description de

forme. Dans la troisième partie, on passe à la prise de décision concernant les masses suspectes (bénigne ou maligne).

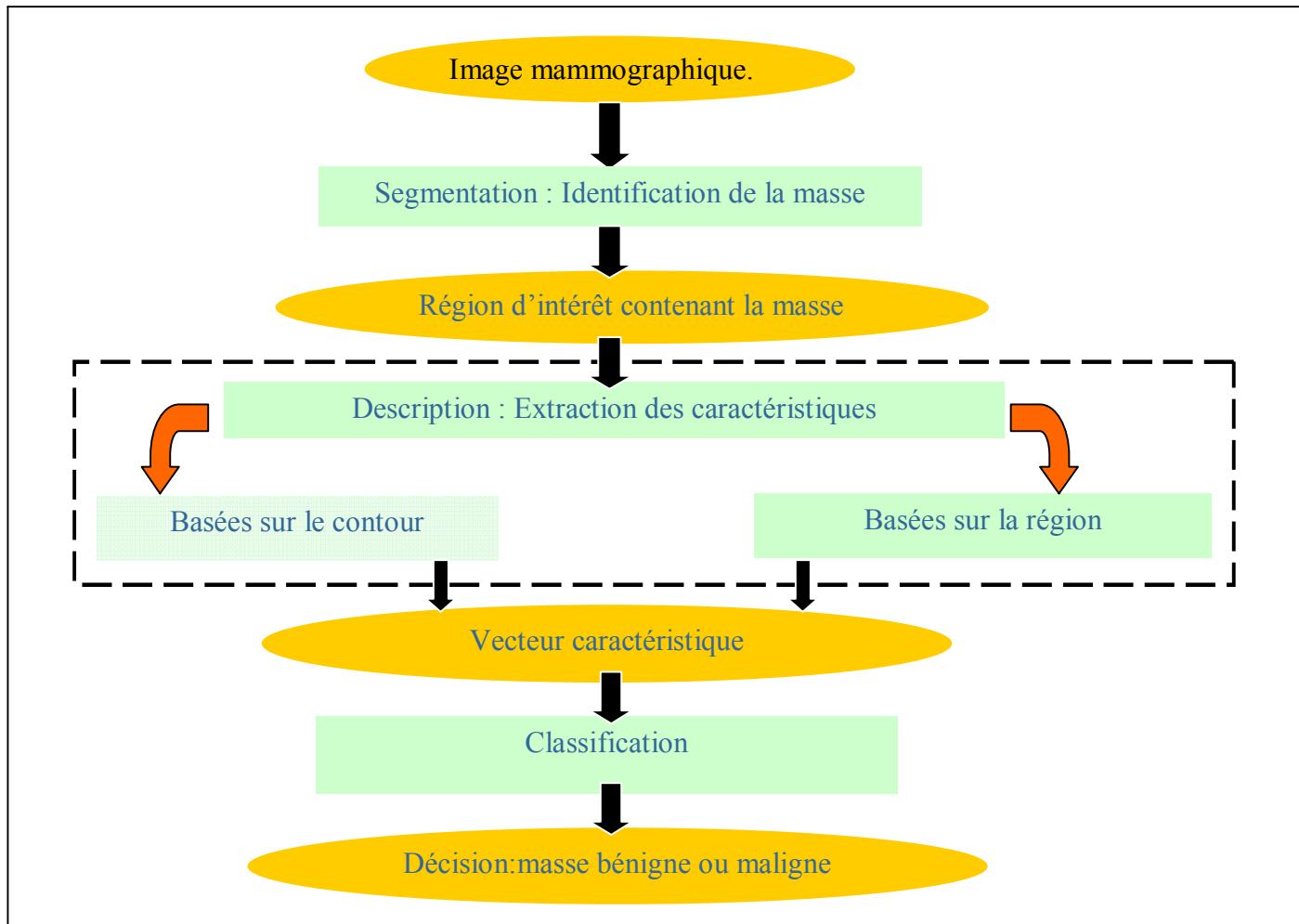


Fig. 2.10. Flot de diagnostic mammographique proposé.

Une fois la région suspecte contenant la masse est connue (on suppose que la région est détectée à travers un système CAD), la tâche à réaliser est l'identification de masses consistant à séparer le tissu de la masse du tissu environnant normal. Cette tâche peut être manuelle (le radiologue trace le contour autour la masse) ou automatique (à travers des méthodes de segmentation). Cette phase sera l'objectif du quatrième chapitre de notre travail.

Après la récupération de masse, on passe à la phase de l'extraction des caractéristiques à l'aide d'un descripteur de forme. Vu que les propriétés de malignité et de bénignité se basent sur la région et le contour des masses, on propose de procéder par une méthode de description qui combine des caractéristiques de région et du contour. On revient sur cette méthode en détail dans le chapitre suivant.

Finalement, pour évaluer la performance des caractéristiques générées dans la deuxième étape, on passe à l'étape de décision à travers des méthodes de classification.

Il est à noter qu'on s'intéresse à l'étape de description alors que les deux autres sont utilisés seulement pour évaluer notre travail.

5. Conclusion

Dans deuxième partie, on a étudié les deux parties constituant un système d'analyse des masses mammographiques : la partie de détection et la partie de diagnostic. Dans les deux parties, l'étape d'extraction des caractéristiques est très importante et doit exister pour augmenter la performance de la détection et/ou de la classification des tumeurs.

Dans notre travail, on s'intéresse à la deuxième partie d'analyse (partie diagnostic) et plus exactement à la partie de description basée sur la forme. Cette dernière, sera l'objectif du chapitre suivant.

Chapitre 3

Extraction des caractéristiques

L'analyse d'images est une discipline dont les fondements théoriques reposent sur les mathématiques et les fondements pratiques sur l'informatique. En effet, les techniques d'analyse d'images jouent un rôle très important dans plusieurs applications médicales, particulièrement en mammographie.

En général, les applications incluent l'extraction automatique de caractéristiques qui seront utilisées dans la prise de décision (par exemple distinguer les tissus normaux de ceux suspects). L'extraction des caractéristiques, qui constitue la deuxième phase d'analyse, est l'étape sur laquelle repose tout ce qui suit car un faut départ peut dévier la décision et engendrer des résultats erronés.

Les caractéristiques extraites à partir d'une image donnent des informations sur les propriétés de forme, les propriétés de couleur ou certaines propriétés de texture [66]. Dans notre travail on s'intéresse aux caractéristiques de forme et c'est dans ce sens que s'oriente ce chapitre qui commence par un aperçu sur les méthodes utilisées pour la description des formes mammographiques suivie par une description détaillée sur notre méthode proposée.

1. Forme

La forme de l'objet est une image binaire représentant l'ampleur de l'objet ; c'est-à-dire la forme peut être considérée comme une silhouette de l'objet. La figure 3.1 montre quelques exemples :

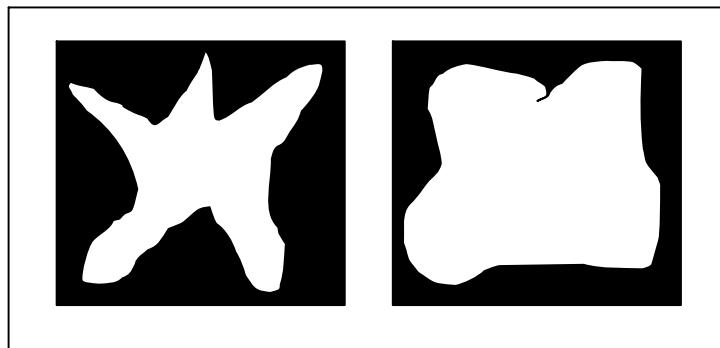


Fig. 3.1. Forme binaire des différents objets.

En (2D), on peut définir deux notions de similarité entre les formes : la similarité de la région et celle du contour. Ceci est illustré dans la figure 3.2 : les objets de la figure 3.2 (a) ont des similarités spatiales dans la distribution des pixels et sont similaires selon des critères régions, alors qu'ils sont clairement différents selon leur contour (Figure 3.2 (b)). On peut déduire donc que l'analyse des formes selon leurs régions ne suffit pas, et qu'il faut ajouter l'analyse basée sur le contour pour avoir

une meilleure précision dans la caractérisation des formes. Dans ce contexte, nous proposons une approche d'analyse des masses mammographique suivants leurs formes et leurs contours.

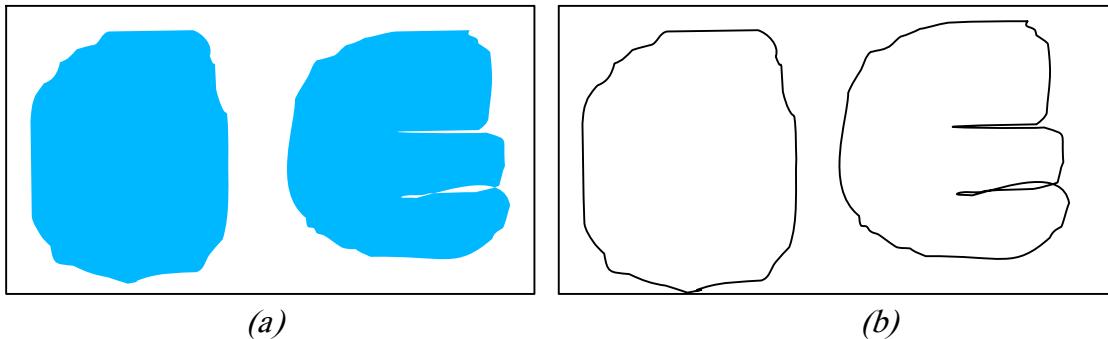


Fig. 3.2. Notion de similarité : (a)régions similaires et (b) contours différents.

2. Description de la forme

Le descripteur de forme se rapporte aux méthodes qui fournissent comme résultat une description numérique de la forme. Il génère un vecteur de caractéristiques d'une forme donnée ; à un objet (contour ou région) donné dans une image, on fait correspondre un vecteur caractéristique dont les composantes sont les différents paramètres calculés.

Les descripteurs, pour différentes formes, devraient être assez différents pour que les formes puissent être distinguées. Ceci implique que le problème rencontré dans l'analyse d'images est le choix d'un bon descripteur et par la suite on ne peut dire qu'un tel descripteur est meilleur qu'un autre que si la différence des valeurs retenues, pour des formes différentes, est grande et que si la différence des valeurs retenues, pour les formes semblables, est petite.

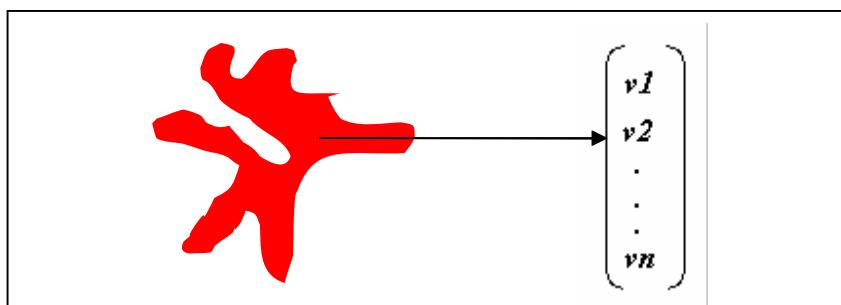


Fig. 3.3. Vecteur des caractéristiques.

Il est important de distinguer les types de descripteurs de forme utilisés pour la discrimination ou la reconnaissance des objets. En général, les descripteurs de formes peuvent être divisés en deux grandes classes : les uns s'appuient sur la forme entière et les autres sur le contour. Les descripteurs

basés sur le contour prennent en compte seulement les pixels de la frontière d'une forme alors que les descripteurs basés sur la région prennent en compte tous les pixels de la forme.

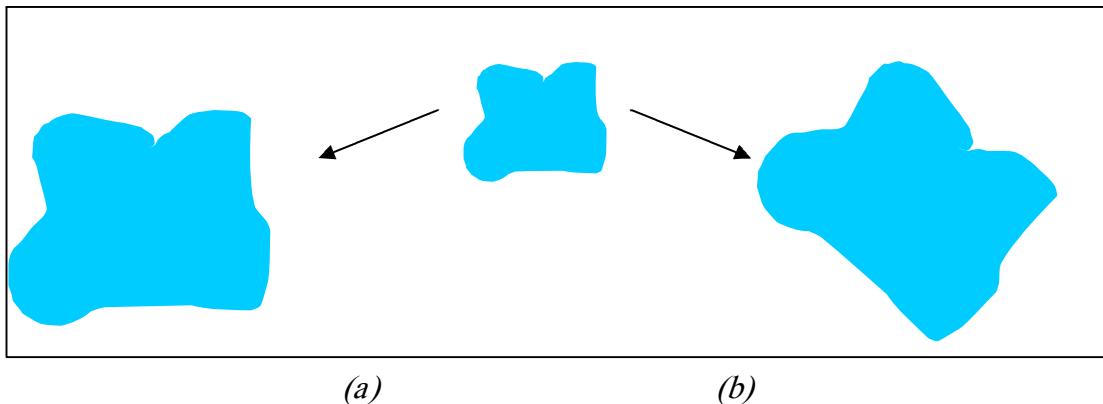


Fig. 3.4. Transformations appliquées à l'image : (a) changement de dimension et (b) rotation.

Malgré que l'objectif de la description soit de caractériser la forme à travers un vecteur caractéristique, les descripteurs de formes doivent satisfaire certains critères parmi lesquels :

- l'invariance à un ensemble de transformations géométriques comme la translation, le changement de dimension et la rotation (figure 3.4) afin d'aboutir à une description efficace d'une forme donnée.
- la rapidité de calcul.

3. Méthodes de description de forme

Une fois la masse est identifiée, nous procédons par une extraction des caractéristiques à partir du contour et de la région pour faire la distinction entre les masses bénignes et celles malignes. En littérature, une variété des méthodes a été utilisée pour l'objectif de description des masses.

Dans le cadre d'analyse des masses mammographiques, de nombreux descripteurs ont été proposés. Dans [45], [46] les auteurs utilisent un ensemble de *caractéristiques géométriques simples* telles que le périmètre, la circularité et l'elongation pour la caractérisation des masses. Dans [47], [53], les auteurs ont, pour leur part, exploité des caractéristiques de *l'enveloppe convexe* telle que la convexité. *La méthode RDM*, est utilisée fréquemment dans diverses approches [53], [59] pour générer des caractéristiques à utiliser dans la phase de description.

Dans cette partie on donne un aperçu sur les méthodes qu'on peut exploiter dans l'élaboration de notre descripteur de forme.

3.1. *Caractéristiques de forme simples*

Plusieurs caractéristiques géométriques simples sont utilisées pendant l'analyse mammographique. Ces caractéristiques peuvent être utiles dans la description de forme. En présente par la suite un panorama non exhaustif des caractéristiques les plus utilisées telles que l'aire, le périmètre, la circularité et l'excentricité.

3.1.1. *Aire et périmètre*

L'aire **A** et le périmètre **P** sont calculés pour caractériser la taille d'une forme donnée. L'aire est définie comme le nombre de pixels à l'intérieur d'une forme. Tandis que le périmètre est défini comme le nombre de pixels du contour d'une forme. Ces deux caractéristiques ne peuvent pas être considérées seules en mammographie, et en analyse des masses afin de distinguer entre les tumeurs bénignes et malignes. Ils ne sont pas suffisants car la taille de la tumeur de sein n'est pas un signe de maladie. Mais ils sont utiles dans le calcul des plusieurs autres caractéristiques qu'on va détailler par la suite.

3.1.2. *Circularité*

La circularité (**C**) est une caractéristique géométrique qui permet de décrire les régions qui peuvent être circulaires. Elle est donnée par l'équation suivante:

$$C = \frac{4\pi A}{P^2} \quad (\text{Eq. 3.1.})$$

Avec

$$\begin{cases} P: \text{Le périmètre de la région.} \\ A: \text{L'aire de la région.} \end{cases}$$

Vu qu'on traite, en mammographie, des masses qui peuvent être soit des formes rondes ou ovales (dans le cas d'une tumeur bénigne), soit des formes irrégulières (dans le cas d'une tumeur maligne), la circularité peut être utile dans ce sens et peut donner une indication sur la régularité d'une forme donnée.

Dans la figure 3.5, on présente deux formes : la première représente une forme régulière (figure 3.5 (a)) qui a une circularité $C=1$ et la deuxième représente une forme irrégulière (figure 3.5 (b)) avec

une circularité $C \ll 1$. Par la suite on peut constater que la différence entre les valeurs retenues dans les deux cas peut être exploitée pour regrouper les cas réguliers de ceux irréguliers.

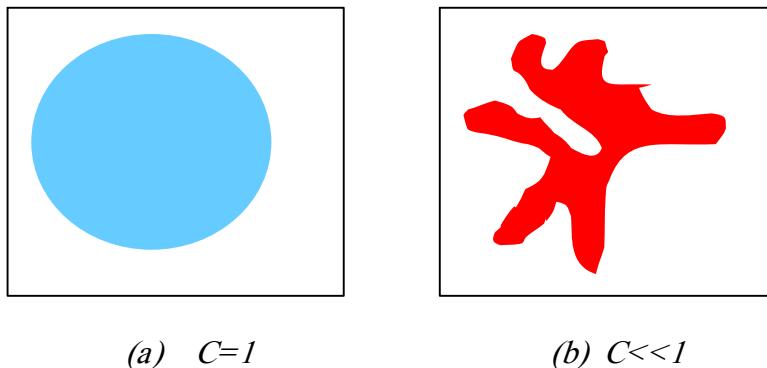


Fig 3.5. Une (a) forme circulaire et une (b) forme irrégulière (non circulaire).

Plus la circularité est haute ($C \approx 1$), plus la forme tend à être plus ronde et par la suite la probabilité d'une masse d'être bénigne est plus grande. Par contre lorsque C est proche de 0, la forme tend à être irrégulière et par conséquent il s'agit d'une tumeur maligne.

En mammographie, la circularité C est utilisée fréquemment dans des travaux destinés au diagnostic [53], [59] vu qu'elle est simple à implémenter et possède l'avantage qu'elle est invariante à toute transformation affine (translation, rotation, etc.).

3.1.3. Excentricité

L'excentricité est une autre caractéristique qui peut être utilisée dans la discrimination des masses. Elle donne une indication sur l'élongation d'une masse et est définie par le rapport entre l'axe mineur et l'axe majeur :

$$Ex = \frac{\text{Axe_Mineur}}{\text{Axe_Majeur}} \quad (\text{Eq. 3.2.})$$

L'axe majeur est la plus longue droite qui traverse le centre alors que l'axe mineur est la ligne qui est perpendiculaire à l'axe majeur et qui passe par le centre. Lorsque l'axe mineur est proche de l'axe majeur il s'agit d'une forme ronde alors que si l'axe mineur est loin de l'axe majeur on dit qu'il s'agit d'une forme allongée.

La figure 3.6, illustre un exemple de calcul de l'excentricité pour une masse bénigne et une autre maligne. Pour une forme bénigne, l'excentricité vaut 1 alors que pour une forme maligne

l'excentricité est loin de 1. Donc, on voit l'utilité de cette caractéristique dans la discrimination entre les cas malins et ceux bénins.

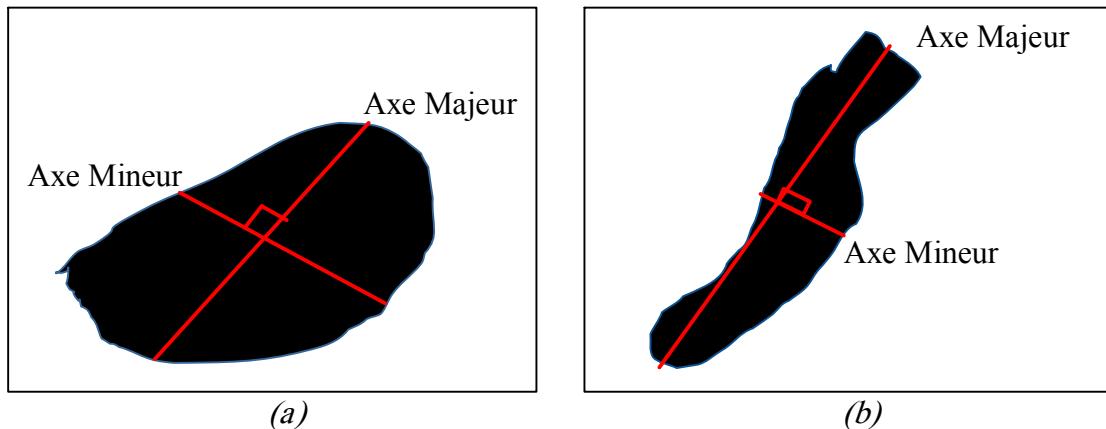


Fig. 3.6. Calcul d'excentricité : pour une (a) tumeur bénigne et (b) une tumeur maligne.

Malgré que cette caractéristique puisse être utile dans la phase de description et qu'elle soit invariante à toute transformation affine, elle présente des limites dans plusieurs cas. La figure 3.7 illustre deux cas dont les valeurs retournées ($Exc \approx 1$) par l'excentricité indiquent la présence des tumeurs bénignes alors qu'au contraire il s'agit des tumeurs malignes.

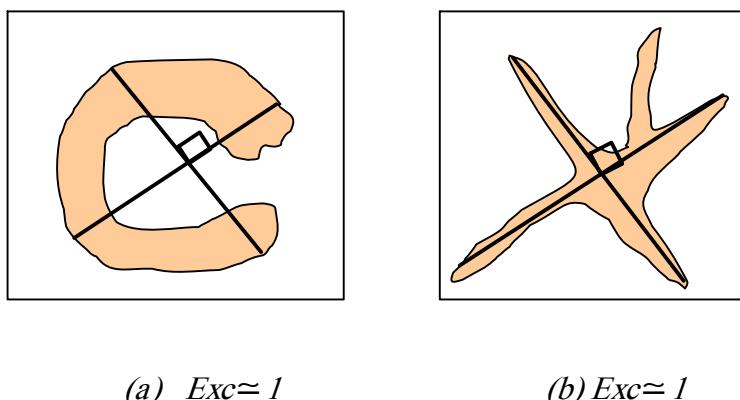


Fig. 3.7. Caractérisation non fiable pour deux tumeurs malignes : (a) Cas 1 et (b) Cas 2.

3.2. Enveloppe convexe

L'enveloppe convexe notée E est largement utilisée dans plusieurs domaines d'analyse tel que l'analyse d'images médicales et en particulier dans l'analyse des masses mammographiques [47], [53].

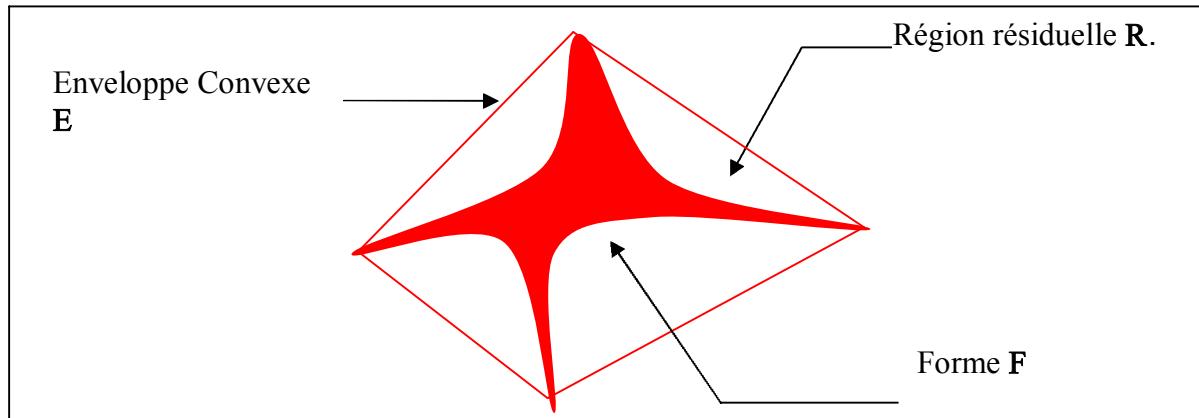


Fig. 3.8. L'enveloppe convexe E d'une forme F .

Elle est définie comme la plus petite surface convexe contenant une forme F (voir la figure 3.8). L'objectif de l'utilisation de l'enveloppe convexe est l'extraction des caractéristiques qui servent à décrire les formes et qui permettent de nous aider dans la distinction entre les formes régulières (rondes, ovales, lobulaires) et celles irrégulières.

On fait appel à plusieurs caractéristiques données comme suit :

- $A(F)$: L'aire (la surface) de la forme F .
- $A(R)$: L'aire (la surface) de la région résiduelle R .
- $A(E)$: L'aire (la surface) de l'enveloppe Convexe E .
- $P(E)$: Le périmètre de l'enveloppe Convexe E .
- $P(F)$: Le périmètre de la forme F .

Les principales caractéristiques qu'on peut utiliser dans notre étude mammographique sont la convexité (**CVX**) et la valeur résiduelle normalisée (**NRV**). Ces deux dernières seront l'objectif des sous-sections suivantes.

3.2.1. Convexité (**CVX**)

On dit qu'un ensemble $X \subseteq E$ est convexe si $rx + (1 - r)y \in X$ pour tout $x, y \in X$ avec $r \in [0, 1]$. Autrement dit, tout segment reliant deux points quelconques (x et y) de X doit être entièrement contenu dans X . À la figure 3.9, on a un objet convexe (figure 3.9 (a)) et un autre qui est non convexe (figure 3.9 (b)).

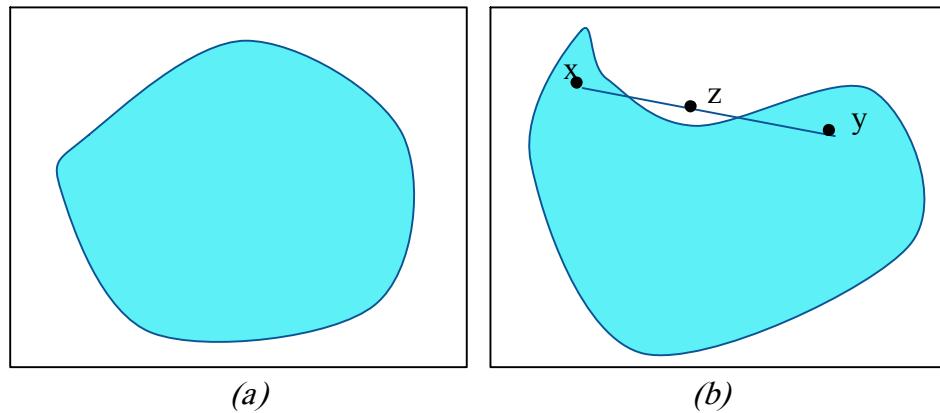


Fig. 3.9. (a) forme convexe et (b) forme non convexe.

Une région fortement convexe (c'est à dire qui épouse parfaitement son enveloppe convexe) a une valeur de convexité proche de 1 (figure 3.10 (a)). Alors qu'une région faiblement convexe possède une valeur de convexité loin de 1 (figure 3.10 (b)).

Les formes irrégulières des tumeurs malignes sont stellaires ou microlobulées ; ces formes sont faiblement convexes. Par contre les formes régulières des tumeurs bénignes (rondes, ovales et lobulées) sont fortement convexes. D'où, la nécessité de calculer une valeur de la convexité pour caractériser les formes des masses mammographique.

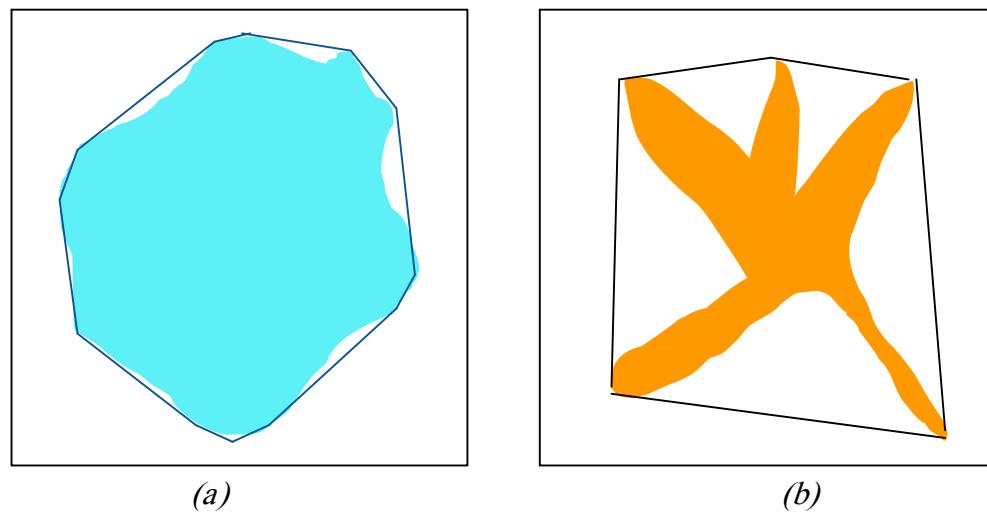


Fig. 3.10. (a) Région régulière (tumeur bénigne) et (b) région irrégulière (tumeur maligne).

En mammographie, pour calculer la convexité (CVX) d'une masse donnée, on utilise l'enveloppe convexe. Cette caractéristique est définie comme le rapport entre l'aire de la forme A (F) et l'aire de l'enveloppe convexe A (E) :

$$CVX = \frac{A(F)}{A(E)} \quad (\text{Eq. 3.3.})$$

La caractéristique de convexité a plusieurs avantages. En effet, elle est facile à calculer. Son calcul est normalisé (entre 0 et 1) ; elle est invariante à toute transformation affine. Mais, la convexité ne détecte pas des défauts énormes sur les frontières des formes qui ont un impact relativement petit sur les surfaces des formes.

3.2.2. Valeur résiduelle normalisée (NRV)

La caractéristique **NRV** (ou Normalised Residual Value) est une autre caractéristique qu'on peut l'extraire de l'enveloppe convexe. Elle est définie par le rapport entre l'aire **A(R)** de la région résiduelle R et le périmètre **P(E)** de l'enveloppe convexe E. La région résiduelle **R** (région en blanc dans la figure 3.16 (b)) est la région résultante de la différence entre l'enveloppe convexe E et la forme **F** :

$$R = E - F \quad (\text{Eq. 3.4.})$$

On dit qu'une forme F est convexe lorsque l'aire **A(R)** de la région résiduelle tend vers 0. La caractéristique **NRV** est donnée par l'équation suivante:

$$NRV = \frac{A(R)}{P(E)} \quad (\text{Eq. 3.5.})$$

Lorsque l'aire de la région résiduelle est plus grande que son périmètre, on aura une valeur de NRV grande et par la suite il s'agit d'une forme non convexe irrégulière (figure 3.11 (a) et (b)). Dans le cas contraire, où le périmètre est proche à l'aire de la région résiduelle, on aura une valeur $NRV \approx 1$, qui permet de conclure que la forme est régulière (figure 3.11 (c) et (d)).

L'avantage de NRV est qu'elle est simple à implémenter, elle est invariante à toute transformation affine et qu'elle donne des résultats encourageants dans plusieurs travaux [47], [72] d'analyse mammographiques. D'après une étude d'évaluation, comme indique le tableau 3.1, d'un ensemble des caractéristiques A. Alvarenga et al [47] ont montré que la caractéristique NRV possède une meilleure sensibilité (94.6%) par rapport à la CVX (81.5%) et aussi une meilleure spécificité (91.7%) par rapport à la CVX (86.7%). Mais l'inconvénient de cette caractéristique est que les valeurs calculées ne sont pas limitées dans un intervalle fini.

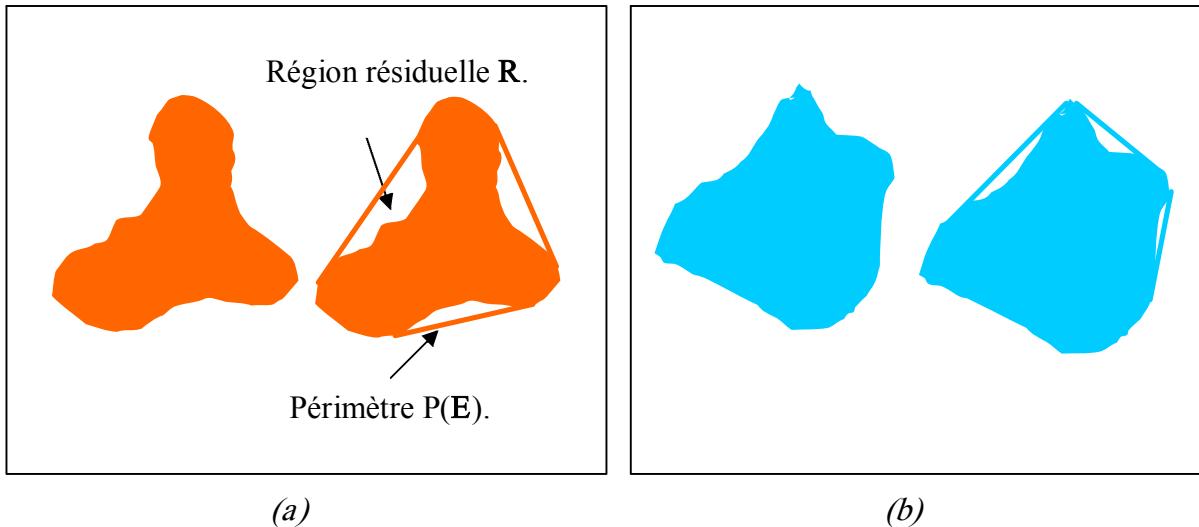


Fig. 3.11. *Calcul de NRV: (a) forme irrégulière et son enveloppe convexe, (respectivement (b) forme régulière et son enveloppe convexe).*

Caractéristiques	Sensibilité (%)	Spécificité (%)
NRV	94.6	91.7
C	88.0	88.3
CVX	81.5	86.7

Tab 3.1. *Evaluation de la caractéristique NRV[47].*

3.3. Mesure de la Distance Radiale (RDM)

La technique de RDM est utilisée fréquemment dans l'analyse mammographique [53], [59] vu qu'elle est efficace pour décrire des masses. À travers cette technique, on fait l'extraction de plusieurs caractéristiques qui peuvent être utilisées par la suite dans des vecteurs caractéristiques. Dans ce qui suit, on donne le principe de cette technique et quelques caractéristiques qu'on peut utiliser pour répondre à nos besoins.

3.3.1. Principe

La mesure de la distance radiale RDM est une méthode qui se base sur le calcul des distances euclidiennes d_i calculées du centre C de la tumeur aux points élémentaires P_i du contour comme il est indiqué à la figure 3.12.

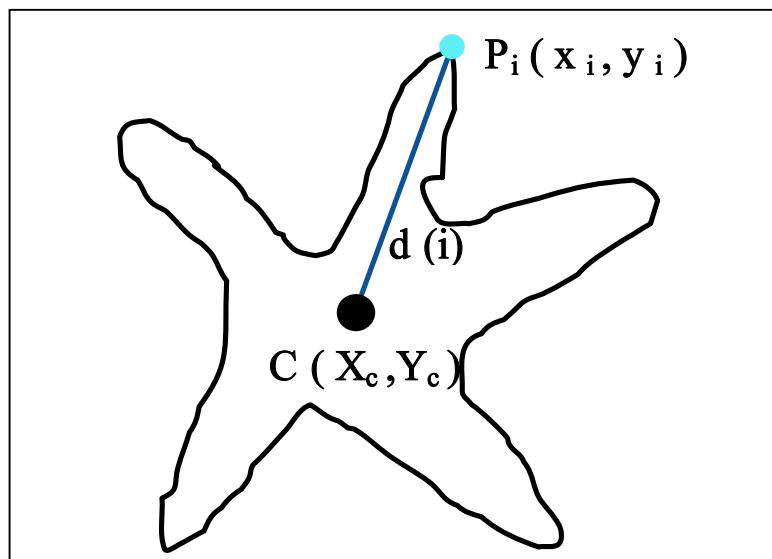
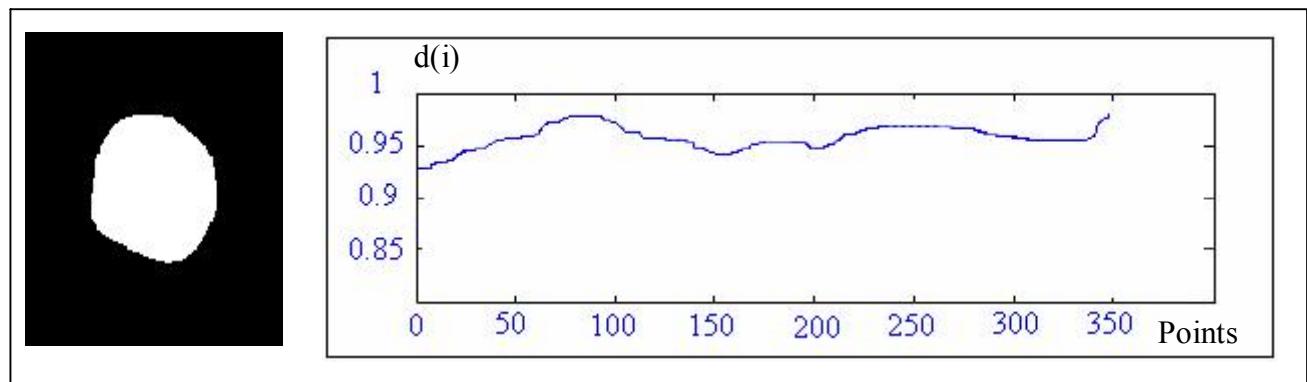
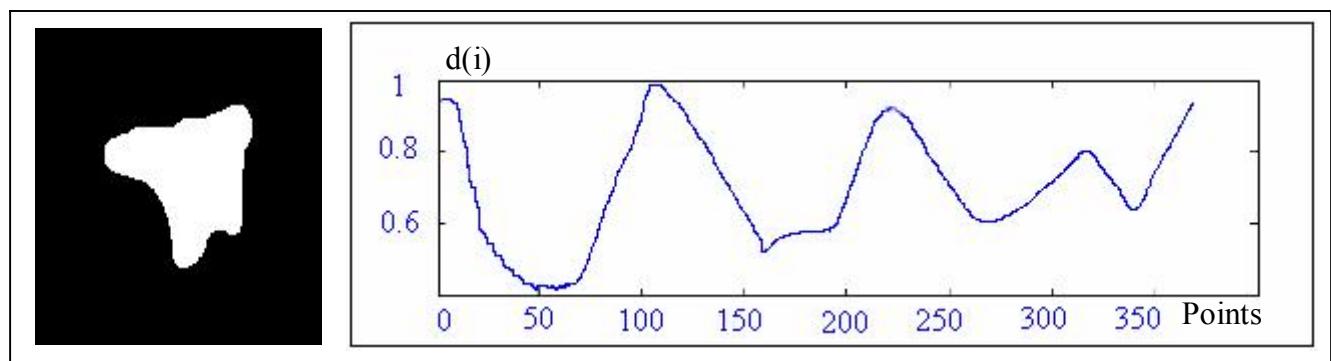


Fig 3.12. Principe de la méthode RDM.

Pour voir l'utilité de cette approche, on étudie la signature du contour afin de distinguer entre un contour régulier et un autre irrégulier. À la figure 3.13 on a deux représentations : l'une pour une tumeur bénigne (figure 3.13 (a)) et l'autre pour une tumeur maligne (figure 3.13 (b)).



(a)



(b)

Fig. 3.13. Signature de contour dans le cas (a) d'une tumeur bénigne et (b) d'une tumeur maligne.

En comparant les deux signatures, on remarque une grande variation entre les distances des rayons (les rayons varient entre 0.4 et 1) pour une tumeur maligne alors que pour une tumeur bénigne il y'a une petite variation entre les distances des rayons (les distances varient entre 0.93 et 1). La grande variation s'explique par la présence d'un contour irrégulier et par la suite on peut conclure que l'utilisation de la technique RDM peut nous mener à des bons résultats lors de la différenciation entre les contours des masses.

Dans la partie suivante on présente les caractéristiques utilisées qui sont extraites à partir de cette méthode pour décrire des masses mammographiques.

3.3.2. Caractéristiques

Plusieurs caractéristiques sont extraites à partir de la méthode RDM tels que la moyenne (d_{moy}), la déviation standard (SDEV), le zero-crossing (ZC), la rugosité (R) et le rapport d'aire (AR). Lors du calcul de ces caractéristiques, on utilise plusieurs paramètres:

- La distance $d(i) = \sqrt{(X_i - X_c)^2 + (Y_i - Y_c)^2}$: la distance euclidienne entre le point i (de coordonnées X_i et Y_i) du contour et le point central de la forme (de coordonnées X_c et Y_c).
- La distance $d_n(i) = \frac{d(i)}{\max[d(i)]}$: la distance euclidienne normalisée.
- Le nombre N : le nombre de points (pixels) du contour d'une forme (appelé aussi le périmètre de contour).
- Les coordonnées $X_c = \frac{\sum_N X}{N}$ et $Y_c = \frac{\sum_N Y}{N}$: sont les coordonnées de point central d'une forme donnée.

Toutes les distances radiales étaient normalisées en utilisant la valeur maximale des distances radiales. Cette normalisation a pour objectif la limitation de la valeur de distance $d_n(i)$ dans l'intervalle [0, 1], quelque soit la dimension de la tumeur et par conséquent rendre la technique RDM invariante à toute transformation affine.

3.3.2.1. Distance moyenne normalisée (d_{moy})

Cette caractéristique peut être utilisée directement pour la description, comme elle peut être utilisée pour calculer d'autres caractéristiques. Elle est importante pour la caractérisation des tumeurs de formes stellaires.

La caractéristique d_{moy} est donnée par l'équation suivante :

$$d_{moy} = \frac{1}{N} \sum_{i=1}^N d_n(i) \quad (\text{Eq. 3.6.})$$

Plus la valeur rentrée par cette caractéristique est proche de 1, plus la forme est régulière ; respectivement, plus la valeur de d_{moy} est loin de 1, plus la forme est irrégulière.

3.3.2.2. Déviation standard (σ)

La caractéristique de déviation standard (Standard Deviation ou σ) de la distance radiale normalisée est définie par la variance des distances autour du rayon (le rayon est la distance d_{moy} définie précédemment) d'un cercle (figure 3.14).

Cette caractéristique est utilisée pour la description des contours stellaires. Elle est donnée par l'équation suivante :

$$\sigma = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N (d_n(i) - d_{moy})^2 \right)} \quad (\text{Eq. 3.7.})$$

Dans la figure 3.14, on montre la caractéristique SDEV qui est représentée par une ligne pointillée qui est la différence (SDEV) entre la distance $d_n(i)$ et la distance moyenne d_{moy} .

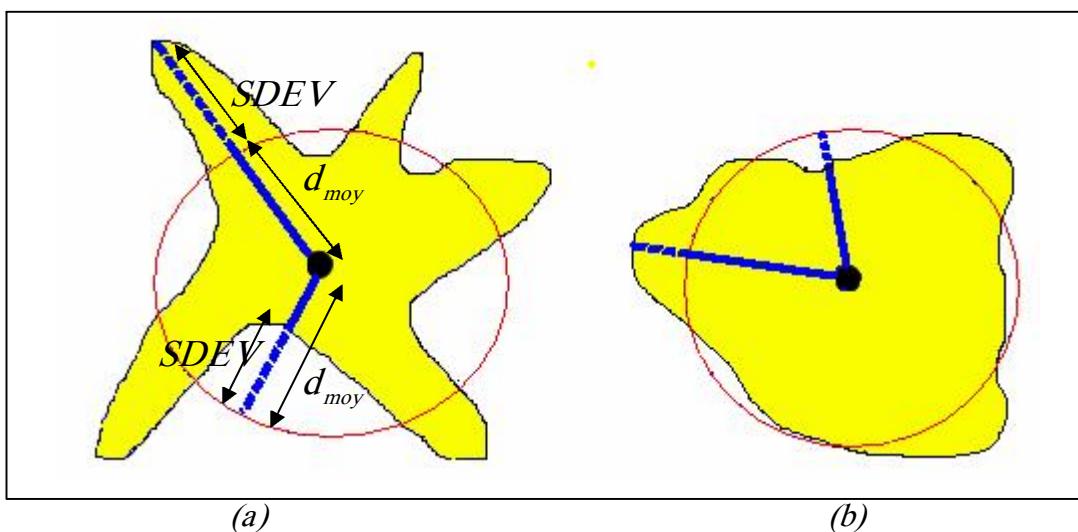


Fig. 3.14. Principe de SDEV pour (a) un contour stellaire et (b) un contour lobulé.

Plus cette différence est grande, plus le contour étudié est irrégulier (contour stellaire maligne à la figure 3.14 (a)). Au contraire, plus cette différence est petite plus la probabilité de régularité du contour est grande (contour lobulé bénigne à la figure 3.14 (b)). Par la suite, on constate que le SDEV peut donner une bonne indication sur l'irrégularité du contour. En effet, lorsqu'il s'agit d'un contour stellaire maligne la valeur de SDEV augmente, par contre dans le cas d'un contour bien entouré ou lobulé bénigne le SDEV diminue et tend vers 0.

3.3.2.3. Zero Crossing (ZC)

C'est le nombre de fois où la distance $d_n(i)$ est supérieure à la distance moyenne d_{moy} . Plus la valeur de Zero Crossing est grande plus le contour est irrégulier. D'après les signatures des contours, à la figure 3.15, on remarque que le nombre de fois où la distance $d_n(i)$ est supérieure à la distance moyenne d_{moy} est grand pour une tumeur maligne (figure 3.15 (a)) et il est petit pour une tumeur bénigne (figure 3.15 (b)).

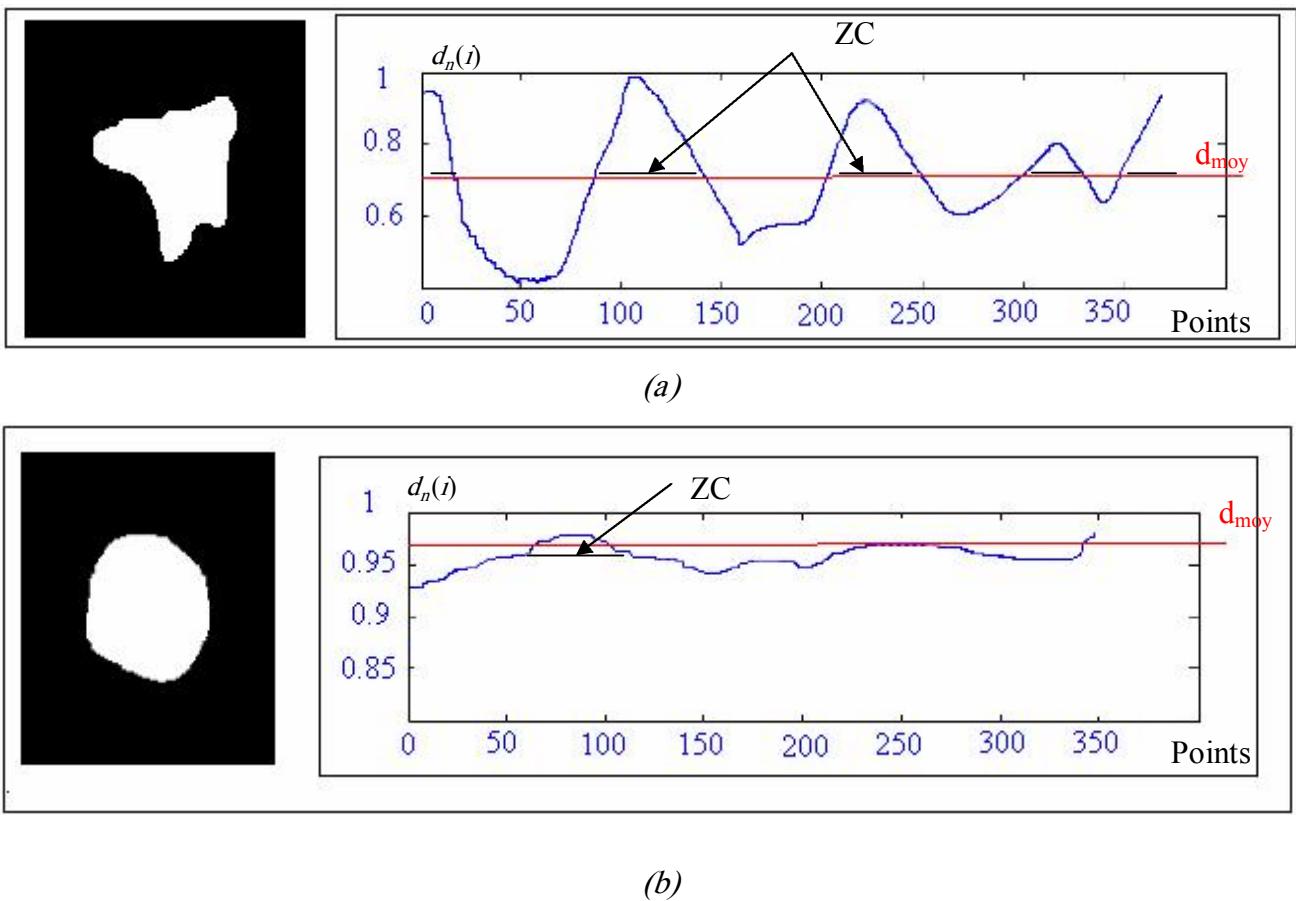


Fig. 3.15. Signature de Zéro Crossing pour (a) une tumeur maligne et (b) une tumeur bénigne.

Le problème est que les valeurs retournées par cette caractéristique, pour une forme donnée, ne sont pas normalisées et par la suite il faut définir un intervalle des valeurs pour n'importe quelle forme.

3.3.2.4. Rugosité (R)

Puisque, en mammographie, on traite des contours qui peuvent être stellaires ou microlobulés, la rugosité peut être utilisée pour nous aider à la description de ces types de contours anguleux (contours qui contiennent des segments concaves).

La rugosité est définie, par l'équation suivante :

$$R = \frac{1}{N} \sum_{i=1}^N |d_n(i) - d_n(i+1)| \quad (\text{Eq. 3.8.})$$

La figure 3.16 illustre le principe de la rugosité donnée par la différence R entre le rayon $d_n(i)$ et le rayon $d_n(i+1)$. On remarque que R augmente pour un contour stellaire (figure 3.16 (a)), diminue pour un contour microlobulé (figure 3.16 (b)) et proche de 0 pour un contour bien entouré (figure 3.16 (c)). Généralement, on peut dire que la rugosité augmente pour une tumeur maligne et diminue pour une tumeur bénigne.

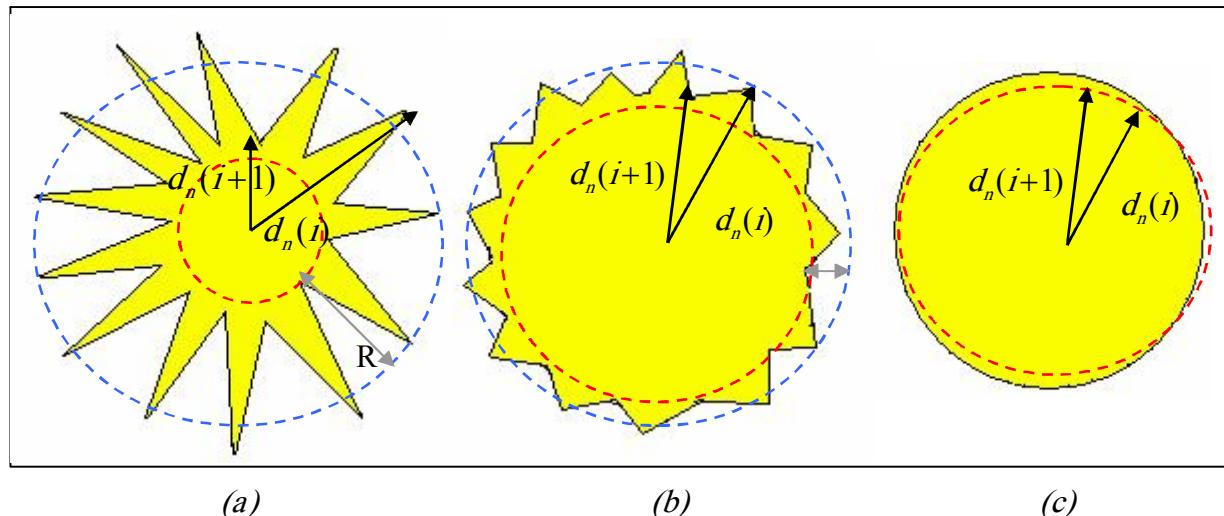


Fig. 3.16. Principe de la rugosité (a) contour stellaire, (b) contour microlobulé et (c) contour lisse.

3.3.2.5. Rapport d'aire (AR)

Le rapport d'aire A_R (ou Area Ration) est une autre forme de caractéristique qu'on peut utiliser pour décrire les contours des masses stellaires. Elle permet de donner le pourcentage que la masse soit en dehors de la région circulaire du rayon d_{moy} . L'équation de cette caractéristique est donné par :

$$A_R = \frac{1}{N^* d_{moy}} * \sum_{i=1}^N (d_n(i) - d_{moy}) \quad (\text{Eq. 3.9.})$$

Avec $A_R = 0 \quad \forall d_n(i) \leq d_{moy}$

À la figure 3.17, on présente la façon de calcul de la caractéristique A_R (représentée par la ligne pointillée en bleu dans la figure) donnée par la différence entre la distance $d_n(i)$ et la distance moyenne d_{moy} mais seulement dans le cas où la distance $d_n(i)$ est strictement supérieure à la distance d_{moy} .

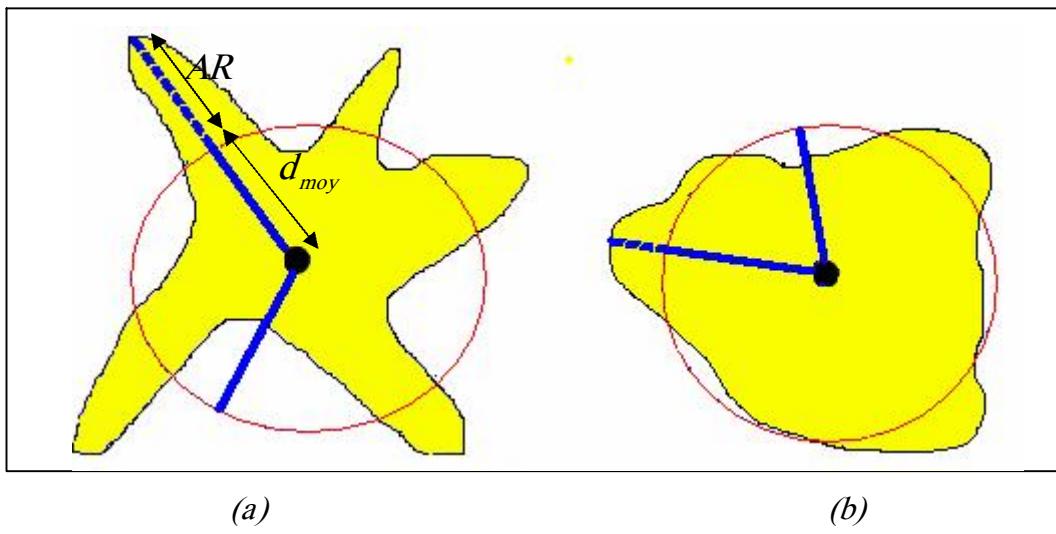


Fig. 3.17. Principe d' A_R pour (a) un contour stellaire et (b) un contour lobulé.

Comme pour le cas de la caractéristique **SDEV**, plus la ligne pointillée est grande, plus le contour est irrégulier (contour stellaire malin à la figure 3.17 (a)). Au contraire, plus la ligne pointillée est petite plus la probabilité de régularité du contour est grande (contour lobulé bénigne à la figure 3.17 (b)). Par conséquent, la caractéristique A_R peut être utile dans la discrimination des contours des masses mammographique. En effet, lorsqu'il s'agit d'un contour stellaire malin la valeur d' A_R augmente, par contre dans le cas d'un contour bien entouré ou lobulé bénigne l' A_R diminue et tend vers 0.

4. Descripteur de forme proposé

Notre méthode proposée est une combinaison des caractéristiques basées sur la région et des caractéristiques basées sur le contour, comme indique la figure 3.17, afin de décrire d'une manière plus efficace les masses mammographiques.

En utilisant une méthode de segmentation, qu'on va décrire dans le chapitre suivant, et à partir de la région de masse (figure 3.18 (a)) isolée, on fait l'extraction de trois caractéristiques : i) la circularité **C**, ii) la valeur résiduelle normalisée **NRV** et iii) la caractéristique d'elongation notée **CIE** (Cercle Interne Externe). Ces caractéristiques, ont pour rôle la description du contenu des formes mammographiques.

À partir du contour de la masse (figure 3.18 (b)), on extrait cinq caractéristiques : trois caractéristiques provenant de la méthode RDM modifiée (section 4.1.1), une caractéristique de convexité **CVX** et une caractéristique, propre à notre travail, calculée à partir des angles notée l'indice d'angle **IA**.

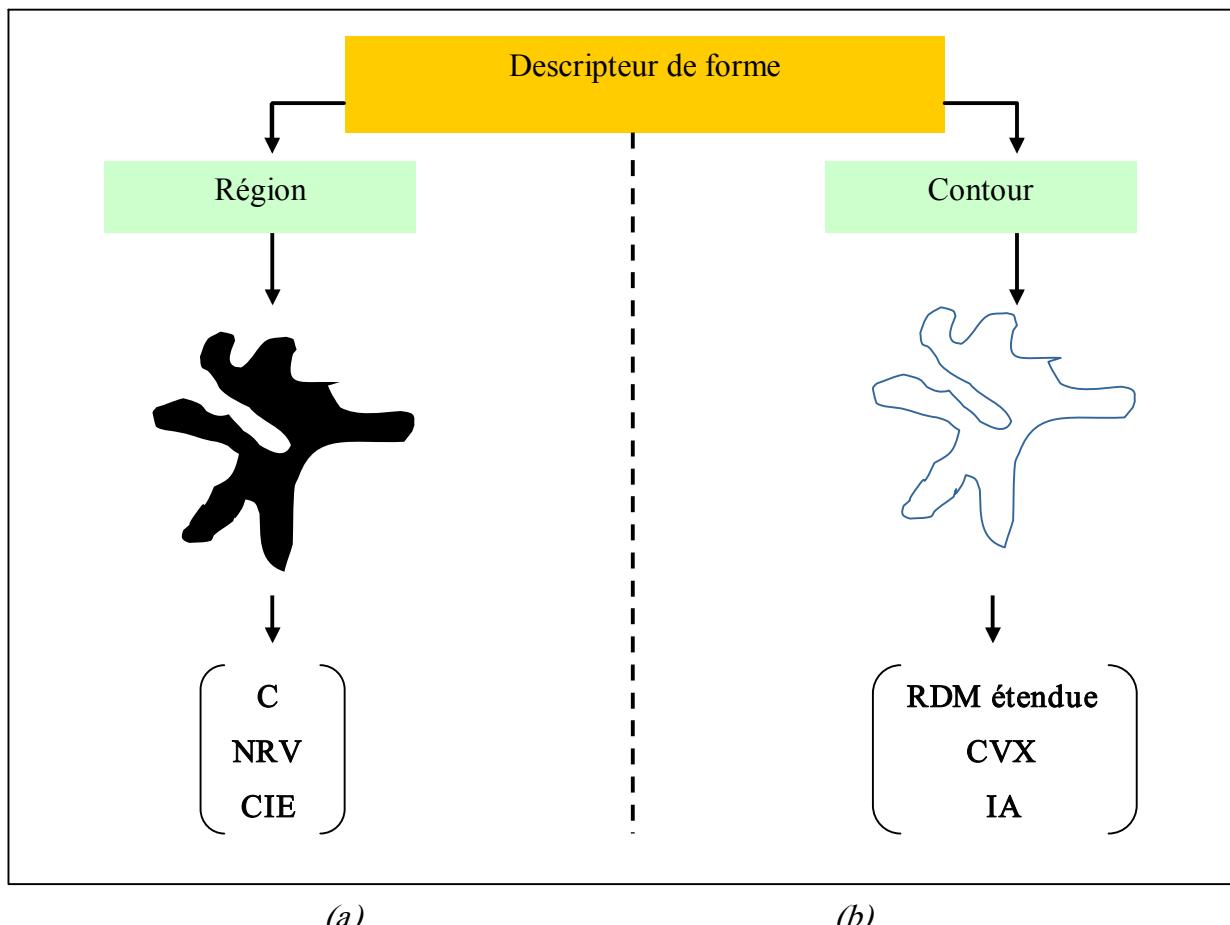


Fig. 3.18. Descripteur de forme : (a) caractéristiques de région et (b) caractéristiques de contour.

Finalement, on a un ensemble de huit caractéristiques qui sont choisies en respectant les critères de simplicité en terme d'implémentation d'un côté, et d'efficacité d'un autre côté.

Par la suite on présente notre vecteur de caractéristiques en indiquant à chaque fois l'intérêt de chaque caractéristique lors de la phase de description.

4.1. Caractéristiques basées sur le contour

Les caractéristiques basées sur le contour sont utilisées pour décrire la frontière d'une forme. En mammographie, on traite les contours des masses qui peuvent être stellaires, microlobulées, circonscrits, mal définis et indistincts (vues dans le chapitre 1).

Les caractéristiques à utiliser sont celles extraites à travers la méthode RDM : le **SDEV**, la rugosité **R** et le rapport **AR**. L'apport dans notre étude est l'amélioration de ces caractéristiques en optimisant les temps de calculs et en augmentant la pertinence des valeurs retenues. Aussi on utilise la convexité **CVX** à partir de l'enveloppe convexe.

La contribution ici est la réalisation d'une nouvelle caractéristique notée (**IA**) jouant un rôle dans la distinction entre les contours réguliers et les contours microlobulés.

4.1.1. RDM étendue

Les caractéristiques basées sur RDM sont celles dont la complexité temporelle liée au calcul est optimale relativement aux autres caractéristiques basées sur le calcul d'angle tel que « turning angle » [49].

Partant de la technique RDM décrite dans la section 3.4, qui est basée sur le calcul de tous les points constituant le contour (complexité temporelle élevée), on se propose de minimiser encore sa complexité en ne prenant en compte que quelques points de contours. Ces points représentent les extrema du contour. Ces extrema représentent soit les points concaves soit les points convexes décrites comme suit :

- Un point concave **P_{concave}** (*i*) de contour est un point qui possède une distance radiale *d* (*i*) inférieure à la distance radiale *d* (*i*-1) et inférieure à la distance radiale *d* (*i*+1).
- Un point convexe **P_{convexe}** (*i*) de contour est un point qui possède une distance radiale *d* (*i*) supérieure à la distance radiale *d* (*i*-1) et supérieure à la distance radiale *d* (*i*+1).

La totalité des points sur lesquels on travaille est donnée comme suit:

- **P_{concave}** (*i*) = { *i* | *d*(*i*) < *d*(*i*-1) et *d*(*i*) < *d*(*i*+1)}

- $P_{\text{convexe}}(i) = \{ i \mid d(i) > d(i-1) \text{ et } d(i) > d(i+1) \}$

En fixant les points convexes et concaves, on réduit le nombre de points du contour à utiliser dans le calcul et par conséquent on réduit la complexité temporelle de calcul des caractéristiques extraites à partir de la méthode RDM. Dans la figure 3.19 on a deux représentations : l'une pour une tumeur maligne (figure 3.19 (a)) et l'autre pour une tumeur bénigne (figure 3.19 (b)).

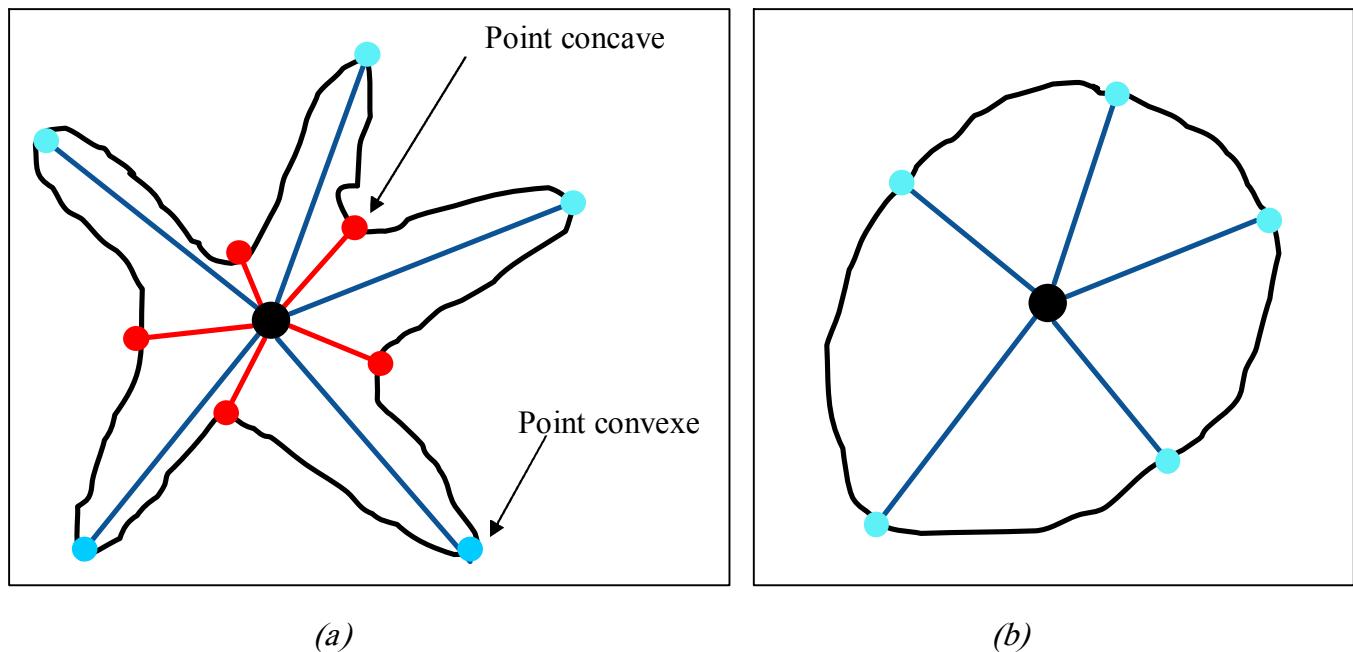


Fig. 3.19. Ensemble des points concaves et convexes (a) cas d'une tumeur maligne et (b) d'une tumeur bénigne.

D'après la table 3.2, on remarque que les valeurs retenues par notre RDM améliorée respectent la séparation entre les masses maligne et celles bénigne ainsi que la méthode RDM. En plus de ça, la séparation à travers notre méthode est plus claire ; car si on prend par exemple le cas de la caractéristique R on constate une distinction plus claire (les valeurs des cas malignes varient entre 0.13 et 0.21 alors qu'elles varient entre 0.013 et 0.018 pour les cas bénignes) par rapport à la caractéristique R de la méthode RDM (les valeurs des cas malignes sont entre 0.02 et 0.03 alors qu'elles sont entre 0.018 et 0.029 pour les cas bénignes).

Parmi les caractéristiques extraites de la méthode RDM, on peut utiliser dans notre travail les caractéristiques suivantes: l'écart type **SDEV**, la rugosité **R** et le rapport **AR** (voir la section 3.2.3) puisqu'elles caractérisent mieux les contours stellaires.

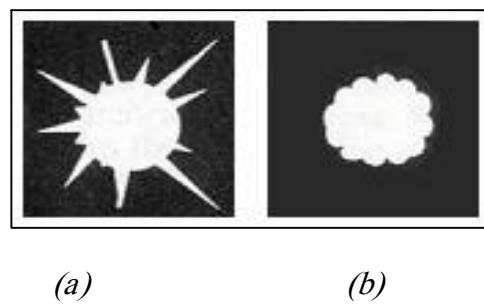
				
RDM étendue	d_{moy}	0.5723	0.6048	0.8602
	A_R	0.1260	0.1470	0.0352
	σ	0.1693	0.2108	0.0715
	R	0.1326	0.2102	0.0187
RDM	d_{moy}	0.6224	0.6053	0.8598
	A_R	0.1012	0.1140	0.0305
	σ	0.1518	0.1693	0.0629
	R	0.0238	0.0363	0.029

Tab 3.2. Comparaison des valeurs retenues par la méthode RDM et RDM étendue.

4.1.2. Indice d'Angle (IA)

L’importance de la technique RDM réside dans la meilleure description du contour stellaire. En contre partie, cette technique ne permet pas de distinguer, d’une manière efficace, entre les cas des contoursstellaires et ceux microlobulés. Ce problème peut être résolu si on prend en considération la notion d’angles.

Partant de la méthode RDM et en utilisant les points concaves et convexes vus dans la section précédente, on ajoute à notre descripteur de forme une caractéristique, propre à notre application, qu’on l’appelle Indice d’Angle (IA). Cette dernière permet la description de contoursstellaires (figure 3.20 (a)) et de contours microlobulés (figure 3.20 (b)).

**Fig. 3.20.** Cas traité par la caractéristique IA : (a) contour stellaire et (b) contour micro-lobulé.

Elle est basée sur le calcul d'angle et est définie par le rapport entre les angles φ_i (calculés à partir des points convexes et concaves) et les angles θ_i (calculés à partir des points convexes et ses voisins). La caractéristique d'indice d'angle est donnée par l'équation suivante :

$$IA = \frac{\sum_{i=1}^N \varphi_i}{\sum_{i=1}^N \theta_i} \quad (\text{Eq. 3.10.})$$

Avec $\varphi_i < \theta_i$

À la figure 3.21, on présente un contour d'une masse microlobulée et la manière de calcul d'angle θ_i et d'angle φ_i . Pour un point i du contour, l'angle φ_i est l'angle fait par le point convexe $P_{\text{convexe}}(i)$ et ses voisins concaves gauche ($i-1$) et droite ($i+1$). L'angle θ_i est l'angle fait par le point convexe $P_{\text{convexe}}(i)$ et point convexe gauche $P_{\text{convexe}}(i+1)$ et droite $P_{\text{convexe}}(i-1)$.

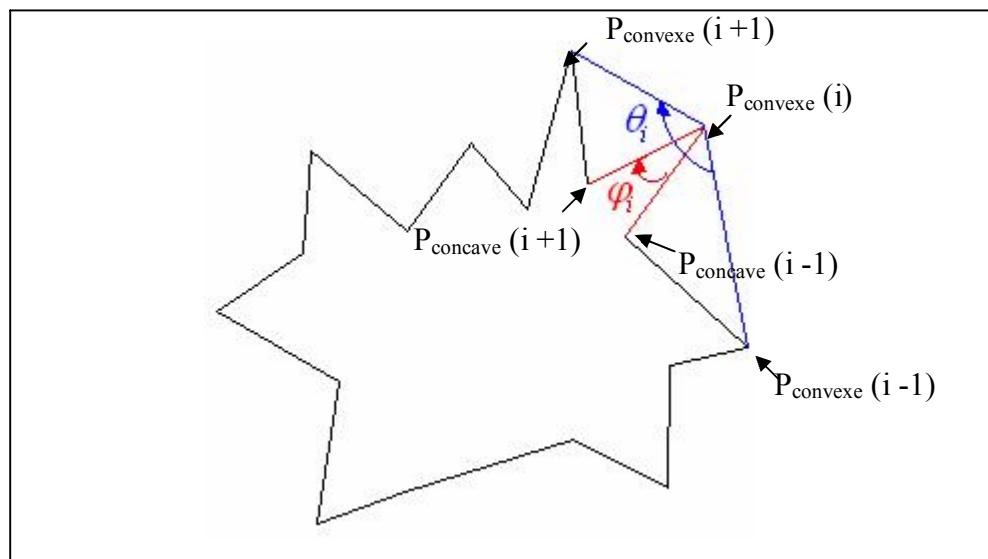


Fig. 3.21. Calcul des angles θ_i et φ_i .

On peut déduire que, plus les valeurs des angles φ_i soient proches des valeurs des angles θ_i , plus la forme de tumeur est bénigne. Contrairement, lorsque les valeurs des angles φ_i soient loin des valeurs des angles θ_i la forme de tumeur est maligne. Les valeurs retournées par l'indice d'angle (IA)

sont limitées dans l'intervalle [0 ; 1]. En effet, pour une forme convexe avec un contour régulier la valeur de IA ≈ 1 alors que pour une forme possédant un contour microlobulé la valeur IA $<<1$.

L'avantage de cette caractéristique, est qu'elle est normalisée et est invariante à toute transformation affine. Aussi, elle a une meilleure complexité par rapport aux autres caractéristiques basées sur le calcul d'angle [49] vu qu'on se limite à un nombre réduit de l'ensemble des points du contour alors que dans les autres méthodes le calcul d'angle englobe tous les points du contour.

4.1.3. Convexité (CVX)

La convexité est une caractéristique utile dans la discrimination entre une tumeur bénigne et autre maligne et elle est aussi simple à calculer.

En général, le calcul de la convexité CVX peut s'effectuer à l'aide de deux manières :

- Soit on utilise le rapport entre l'aire de la forme A (F) et l'aire de l'enveloppe convexe A (E):

$$CVX_1 = \frac{A(F)}{A(E)} \quad (\text{Eq. 3.11.})$$

- Soit on utilise le rapport entre le périmètre l'enveloppe convexe P(E) et le périmètre de la forme P (F):

$$CVX_2 = \frac{P(E)}{P(F)} \quad (\text{Eq. 3.12.})$$

Des travaux d'analyse mammographiques font appel à la caractéristique de convexité basée sur le calcul d'aires (équation 11). Cependant, dans une étude comparative sur l'efficacité de la convexité dans l'analyse des formes J. Zunic et al [71] ont montré que le calcul de cette caractéristique en se basant sur le périmètre au lieu de l'aire permet d'augmenter l'écart entre les valeurs de convexité pour les formes convexes avec celles non convexes. La figure 3.22, présente une forme irrégulière non convexe d'une tumeur maligne.

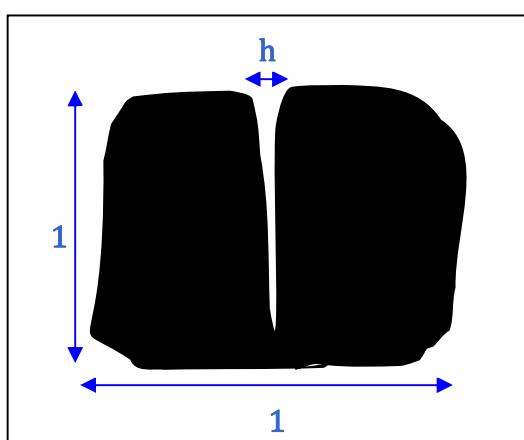


Fig. 3.22. Calcul de convexités : $CVX_1 = (1-h)$ et $CVX_2 = 2/(3-h)$.

En prenant une valeur de h proche de 0, on aura une valeur de CVX1 qui est proche de 1 ($\lim_{h \rightarrow 0} (1-h) = 1$) et une valeur de CVX2 qui tends vers 2/3 ($\lim_{h \rightarrow 0} \left(\frac{2}{(3-h)} \right) = \frac{2}{3}$), par la suite on constate que CVX2 décrit d'une manière plus exacte la forme au contraire de CVX1 qui va conclure sur le fait que la forme est convexe ce qui n'est pas le cas. Donc la convexité basée sur le calcul d'aires peut mener à des décisions erronées ce qui nous empêche à utiliser cette caractéristique malgré qu'elle est la plus utilisée dans la littérature.

Pour rendre notre description des contours plus fiable, nous proposons d'utiliser la convexité calculée à partir des périmètres de la forme et de l'enveloppe convexe.

4.2. Caractéristiques basées sur la région

Les caractéristiques basées sur la région sont utilisées pour décrire le contenu d'une forme. En mammographie, on traite des masses qui peuvent être rondes, ovales, lobulaires et irrégulières. Dans notre étude, on opte pour l'utilisation de la circularité **C** vu qu'elle est simple à implémenter et qu'elle peut être utile pour distinguer les masses régulières et les masses irrégulières. En plus de la circularité la caractéristique **NRV**, extraite de l'enveloppe convexe, peut nous aider à décrire les masses et peut nous donner de meilleurs résultats. Notre contribution est l'intégration d'une nouvelle caractéristique. Cette dernière, notée **Cercle Interne Externe (CIE)**, constitue une amélioration de la caractéristique d'excentricité (voir la section 3.1.3).

4.2.1. Circularité (C)

L'utilisation de cette caractéristique ne fait pas partie de nos contributions, mais son utilisation fréquente dans plusieurs approches d'analyse et sa simplicité d'implémentation nous encourage à l'utiliser. D'où l'idée de profiter de la circularité **C** (présentée dans la section 3.1.2) qui peut donner des résultats importants [47] et surtout lorsqu'on étudie des formes qui peuvent être circulaires.

4.2.2. Cercle Interne Externe (CIE)

Pour calculer le taux d'elongation d'une masse mammographique, on utilise la caractéristique d'excentricité. Mais l'excentricité, comme on le montre par la suite, ne peut pas être efficace dans plusieurs cas ce qui donne un mauvais effet lors de la caractérisation des formes mammographiques.

Sur la figure 3.23 on donne un exemple de calcul d'excentricité pour deux formes mammographiques. En calculant le rapport entre l'axe mineur et l'axe majeur de la forme maligne (figure 3.23 (b)), on remarque que la valeur retournée est la même pour le cas de la forme bénigne (figure 3.23 (a)). Par conséquent, l'excentricité, dans ce cas, conduira vers une fausse décision.

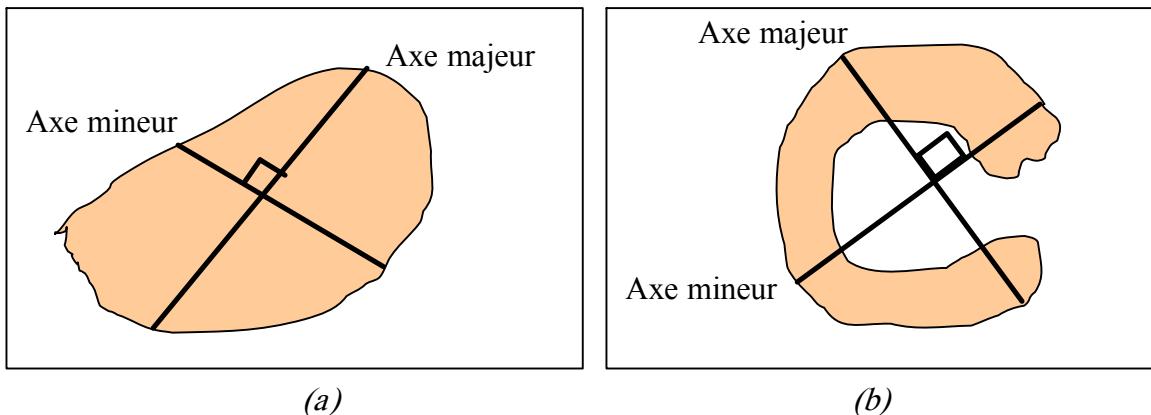


Fig. 3.23. Mauvaise description par l'excentricité : (a) forme ronde bénigne et (b) d'une forme allongée maligne.

Pour remédier à ce problème, nous nous proposons d'utiliser une autre caractéristique notée « Cercle Inscrit Circonscrit » (**CIE**). Cette dernière a été utilisée par C. Chettaoui et al [54] dans leur travail d'analyse des globules drépanocytaires. Ils ont montré que la caractéristique CIE peut palier les défauts d'utilisation d'excentricité dans la description des régions qui peuvent être allongées. Cette caractéristique est basée sur le calcul de rapport entre deux rayons : le rayon **Rgi** du plus grand cercle interne (le cercle en bleu à la figure 3.24) et le rayon **Rpe** du plus petit cercle externe (le cercle en rouge à la figure 3.24) de la forme étudiée. Elle est donnée par l'équation suivante :

$$CIE = \frac{R_{gi}}{R_{pe}} \quad (\text{Eq. 3.13.})$$

En analysant la figure 3.24, on remarque que pour une forme ronde (figure 3.24 (a)) la valeur de CIE est proche de 1 puisque la valeur de **Rgi** est très proche de la valeur de **Rpe**, alors que pour une forme allongée (figure 3.24 (b)) la valeur de CIE devient proche de 0 puisque la valeur de **Rgi** est loin de la valeur de **Rpe**.

L'avantage de cette caractéristique est qu'elle est invariante à toute transformation affine. Elle est adéquate à notre travail et elle représente une amélioration de la caractéristique l'excentricité. Mais

son calcul est moins rapide, étant donné que pour chaque forme on doit parcourir tous les points pour déterminer le cercle inscrit dans l'objet qui contient ce point.

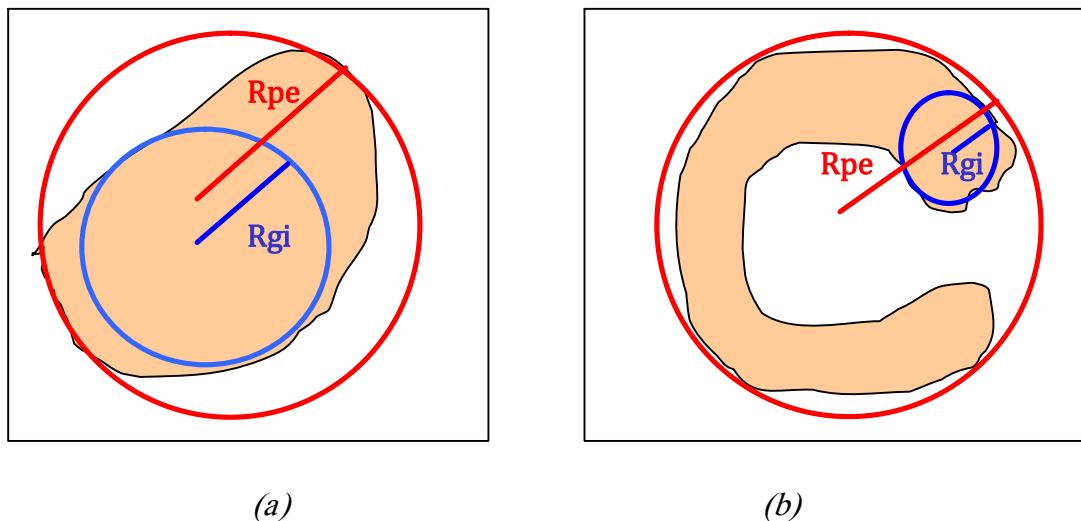


Fig. 3.24. Caractéristique CIE pour une (a) forme ronde et (b) une forme allongée.

4.2.3. Valeur moyenne résiduelle (NRV)

À partir de l'enveloppe convexe, on utilise la caractéristique NRV (décrise dans la partie 3.2.2) qui donne les meilleures performances par rapport aux caractéristiques qu'ont peut extraire, et qui peut être utile dans la distinction entre les régions régulières et irrégulières.

5. Conclusion

Dans ce chapitre, une étude sur les méthodes de description de forme, utilisées pour la discrimination entre les masses malignes et bénignes, a été faite. Le choix de la méthode a pris en compte deux catégories de caractéristiques : la première est basée sur le contour et elle sert à décrire la régularité de la frontière d'une masse. Quant à la deuxième, elle est basée sur la région et permet la description de contenu d'une masse.

Parmi les méthodes de description du contour, on a montré que la méthode de la mesure de la distance radiale (RDM) possède une complexité de calcul minimale par rapport aux autres méthodes telles que les moments et les méthodes basées sur le calcul d'angles. Nous avons proposé une méthode de calcul permettant l'optimisation de la technique RDM afin de réduire de plus en plus la complexité

temporelle. En utilisant la méthode RDM on a introduit une nouvelle caractéristique (IA) qui est basée sur le calcul d'angle et a pour rôle la description des contours microlobulés.

En plus des caractéristiques basées sur le contour, on a utilisé trois autres caractéristiques basées sur la région : la première est la circularité (C) qui décrit les régions rondes, la deuxième est le Cercle Interne Externe (CIE) qui décrit les régions allongées. Tandis que, la troisième caractéristique, extraite à partir de l'enveloppe convexe, notée la NRV, est destinée à décrire la convexité de la région. Ces caractéristiques seraient évaluées dans le chapitre suivant.

Chapitre 4

Evaluation & Validation du flôt proposé

Dans le chapitre précédent, on a défini un descripteur de forme à utiliser dans la discrimination entre les tumeurs malignes et celles bénignes. Ce descripteur contient deux catégories de caractéristiques : les caractéristiques basées sur la région et celles basées sur le contour. Dans ce chapitre, on fait recours à la phase de validation de notre approche en se basant sur un flot d'analyse. Pour cela, on commence par présenter le flot d'analyse proposée, ensuite on détaille chaque phase d'analyse toute seule en indiquant son objectif, et enfin on termine par présenter et discuter les résultats donnés par le descripteur de forme proposé.

1. Présentation du flôt d'analyse

Le flôt d'analyse proposé décrit les phases à suivre pour analyser les masses (voir la figure 4.1) mammographiques. Ces phases sont respectivement: la segmentation, l'extraction des caractéristiques et la classification.

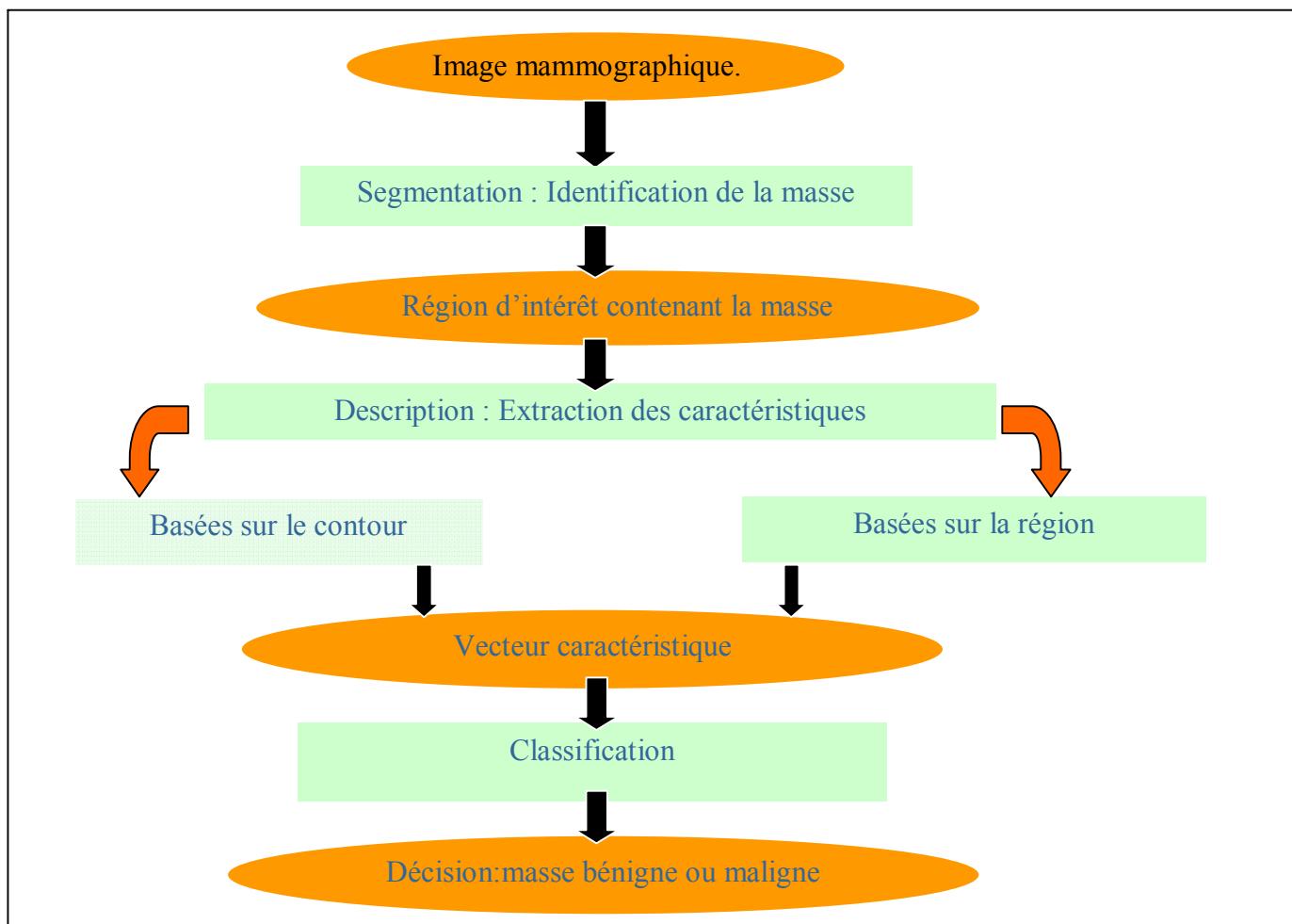


Fig. 4.1. Flôt d'analyse proposé.

Afin de d'extraire les caractéristiques, définies dans le chapitre précédent, on a besoin d'identifier la masse sur laquelle on fait tous les calculs. Pour cela, on se propose d'utiliser des techniques de segmentation classiques. Dans cette phase, on commence par une sélection manuelle de la région suspecte et à partir de laquelle on applique une méthode de segmentation pour identifier la forme. À partir de cette forme on applique une deuxième méthode de segmentation qui permet d'isoler le contour.

Dans la phase d'extraction des caractéristiques, à travers la région et le contour déjà identifiés, on réalise notre descripteur de forme en calculant un ensemble de 8 caractéristiques. Après avoir calculé l'ensemble des caractéristiques, et pour évaluer les résultats trouvés, on passe à une étape de classification. Cette dernière nous permet d'avoir une information sur la décision ainsi que sur le taux de performance de notre descripteur de forme.

Afin de voir l'impact de la classification sur les caractéristiques qu'on a défini on a utilisé trois classifieurs supervisés de différents types : un probabiliste, un autre basé sur l'apprentissage et un troisième exacte.

Dans cette section, on présente les détails de chacune des trois phases composant le flot d'analyse proposé.

1.1. Identification de la masse (méthodes de segmentation)

La phase de segmentation est sans doute la colonne vertébrale de toute étude sur le traitement/analyse d'images. De nombreuses techniques, méthodes, algorithmes peuvent être trouvés dans la littérature. Le niveau de connaissance dans le domaine est tel qu'il est nécessaire, à l'heure actuelle, de choisir une technique de segmentation adaptée à une application donnée. En mammographie, plusieurs méthodes de segmentation ont été utilisées pour l'extraction de région d'intérêt (RI) à partir d'une image. Ces méthodes sont classées en trois catégories : les méthodes basées sur la région [23], [38], les méthodes de regroupement de pixels (clustering) [24] et les méthodes basées sur le contour [37].

Vu que notre descripteur de forme proposé est une combinaison des caractéristiques qui décrivent la région et le contour d'une forme, nous proposons d'utiliser deux méthodes de segmentation : une méthode qui permet l'identification de la région et une autre pour la détection du contour.

Notre méthode proposée pour l'identification des masses est semi-automatique et commence depuis la sélection manuelle d'une zone suspecte, dans l'image, par le radiologue : La figure 4.2 présente le flôt d'extraction de masse proposé.

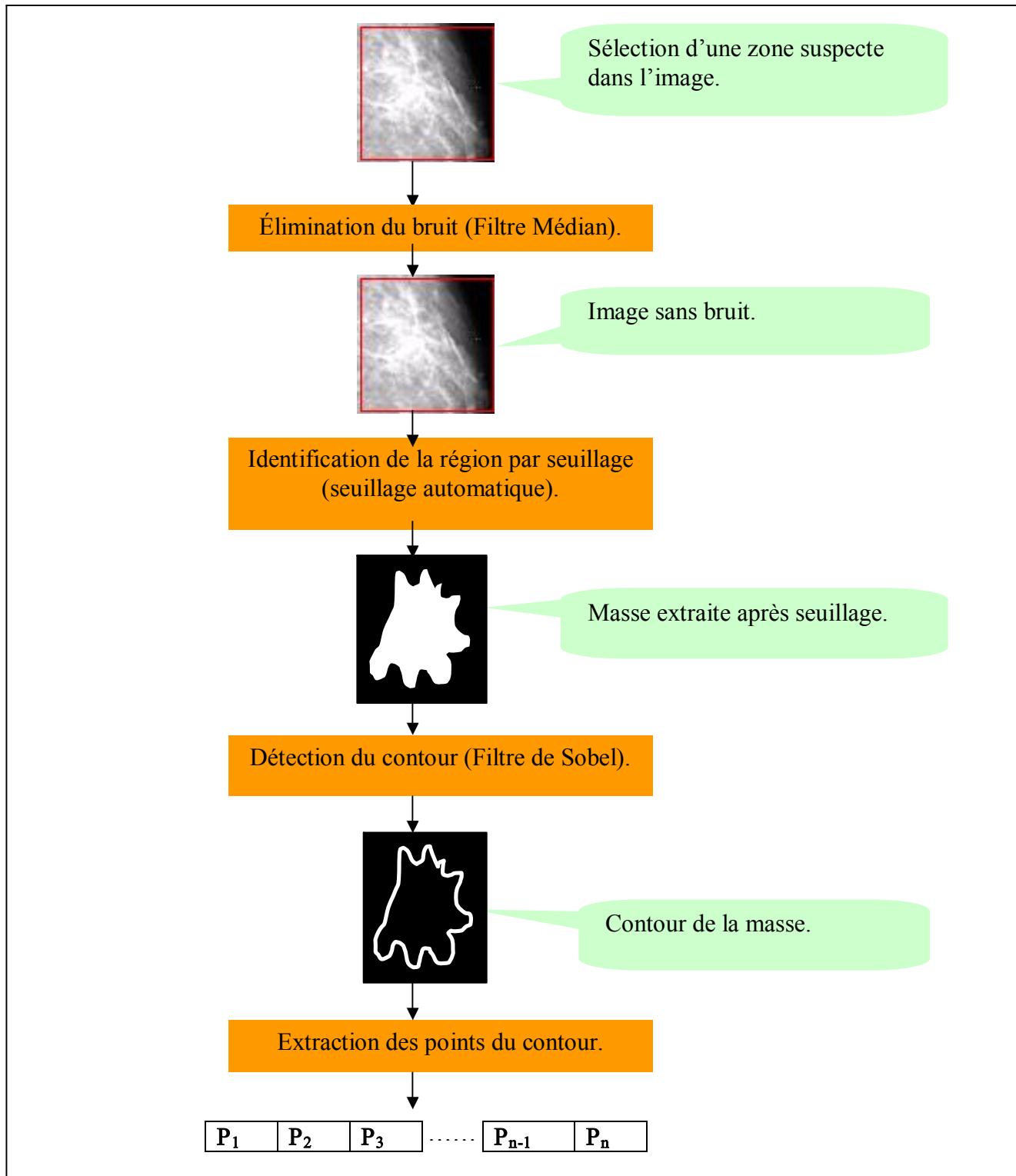


Fig. 4.2. Flôt d'identification d'une masse.

Le flot proposé se compose de quatre étapes : élimination du bruit d'une zone sélectionnée manuellement, extraction de la région, extraction du contour à partir de la région et extraction des points du contour.

1.1.1. Prétraitement (élimination de bruit)

Cette première étape de prétraitement a pour but d'améliorer la qualité de l'image de façon à ce que le traitement qui suivra soit optimal en qualité. Elle permet d'atténuer le bruit de l'image. Pour cela plusieurs filtres passe-bas peuvent être utilisés. Parmi ces filtres passe-bas qu'on peut utiliser est le filtre médian.

1.1.2. Identification de la région (seuillage)

Pour extraire une région d'une image I , on peut utiliser le seuillage binaire qui est un cas particulier des méthodes par regroupement. Les méthodes de binarisation consistent à transformer une image de niveau de gris en une image binaire à travers un seuil S bien déterminé ; c'est-à-dire pour f une fonction de binarisation et $i(x, y)$ un pixel de l'image, on a $f(i(x, y)) \in \{0, 1\}$.

L'implémentation du seuillage binaire d'une image est très simple mais la problématique est dans le choix de seuil. C'est pour cette raison, qu'on peut utiliser une méthode de seuillage automatique comme celle décrite dans [73].

1.1.3. Détection du contour (le filtrage de Sobel)

Après le seuillage, on passe à l'extraction du contour à l'aide d'une méthode de détection. Parmi les méthodes possibles on cite la méthode de Sobel, la méthode de Laplacien ou la méthode de Snake.

Dans cette étude, on opte pour la méthode de Sobel pour l'extraction de contour car c'est la méthode la plus simple à implémenter et qui permet d'aboutir à nos besoins. Cette méthode est basée sur le filtrage passe-haut et elle consiste à appliquer les deux filtres (G_x pour l'axe horizontal (X) et G_y pour l'axe vertical(Y)) sur l'image (voir figure 4.3).

$$\begin{array}{c} G_x \quad \quad \quad G_y \\ \left(\begin{array}{ccc} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{array} \right) \quad \quad \left(\begin{array}{ccc} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{array} \right) \end{array}$$

Le contour résultat est donnée par $G = \sqrt{(g_x^2 + g_y^2)}$ où g_x et g_y sont les résultats de l'application des filtres G_x et G_y sur l'image (c'est-à-dire $g_x = G_x * I$ et $g_y = G_y * I$).

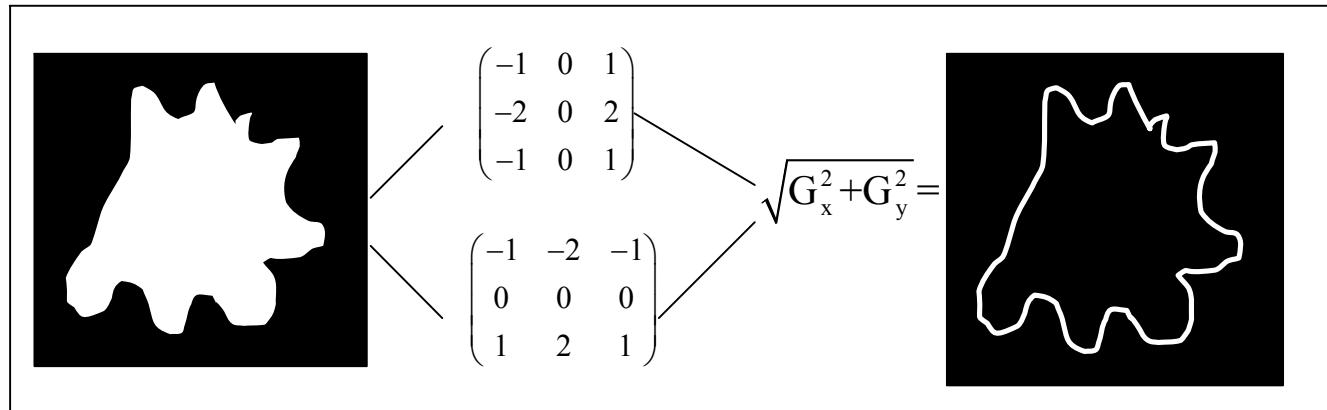


Fig. 4.3. Filtrage Sobel : (a) Image initiale, (b) application de deux filtres 3×3 et (c) contour détecté.

1.1.4. Extraction des points du contour

L'extraction des points du contour nous permet d'obtenir une image binaire ; c'est-à-dire des pixels blancs représentant les points de contour et des pixels noirs représentant les autres points de l'image. Il existe plusieurs méthodes qu'on peut utiliser comme celle basée sur l'algorithme de tortue [62], ou celle basée sur le chaînage de Freeman [63]. Dans notre étude, nous avons choisi la méthode basée sur le chaînage de Freeman pour le suivi du contour comme indique la figure 4.4.

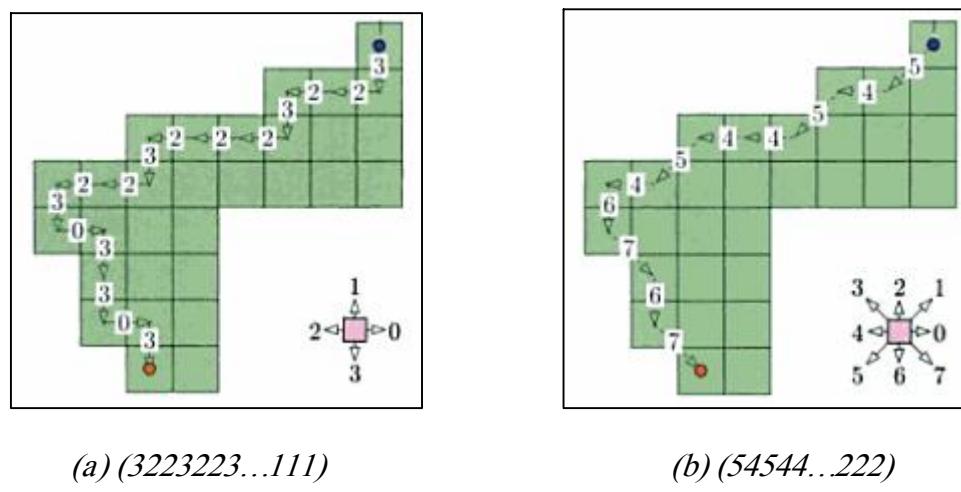


Fig. 4.4. Chaînage de Freeman dans le cas de direction à (a) 4 connexité et à (b) 8 connexité

Son principe est de réaliser le chaînage par suivi des pixels du contour suivant une direction choisie dans un ensemble fini. La direction est codée par une suite comprise entre 0 et 3 en 4 connexités (4 voisins pour chaque pixel) et une suite comprise entre 0 et 7 en 8 connexités (8 voisins pour chaque pixel).

1.2. Extraction des caractéristiques

Pendant l'étape d'identification de la masse, on a utilisé la méthode de seuillage et la méthode de Sobel. L'utilisation du seuillage est utile dans le calcul des caractéristiques qui décrivent la région alors que l'utilisation de la méthode de Sobel est utile lors du calcul des caractéristiques qui décrivent le contour. On a choisi ces méthodes vu leur simplicité d'implémentation. Par la suite, on passe à l'étape d'extraction de caractéristiques qui constitue la deuxième phase de notre flot d'analyse et qui représente l'objectif de notre travail.

Dans la figure 4.5 on présente deux images : la première (a) représente la région de la tumeur avec leur caractéristiques (**C, NRV, CIE**), et la deuxième représente le contour avec l'ensemble des caractéristiques que lui sont associé (**CVX, σ , R, AR, IA**).

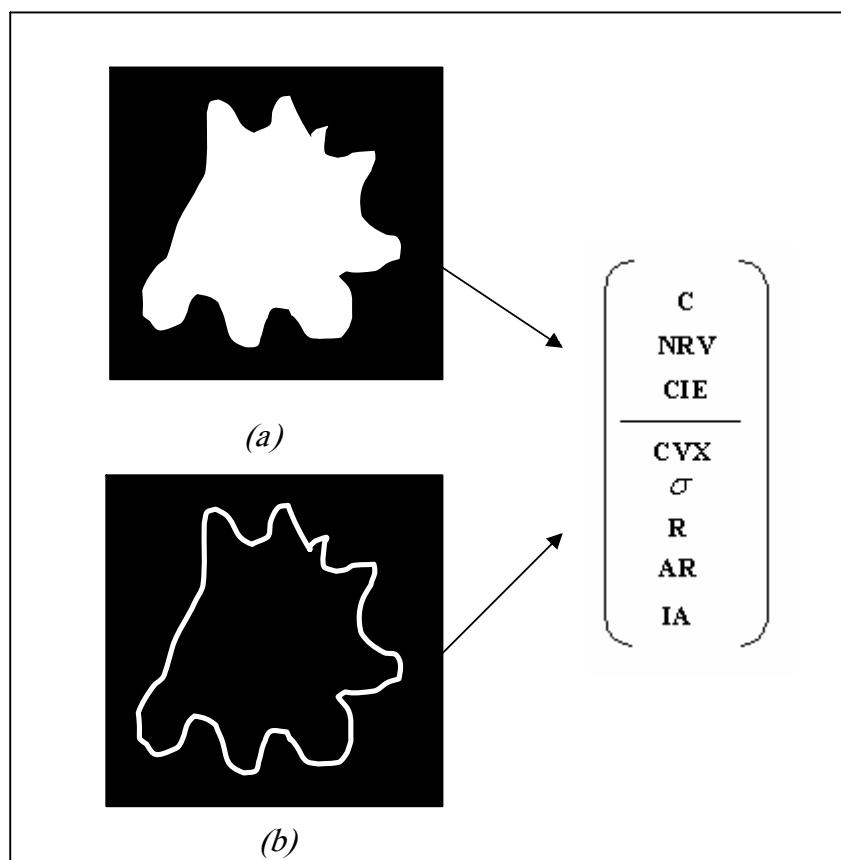


Fig. 4.5. Caractéristiques (a) de région et (b) de contour.

Les caractéristiques de **CIE** (Cercle Interne Externe), **C** (la circularité) et **NRV** (l'indice de convexité), qui décrivent la région, jouent le rôle de discriminants entre les masses régulières (rondes, ovales, lobulaires) et les masses irrégulières (voir la figure 4.6).

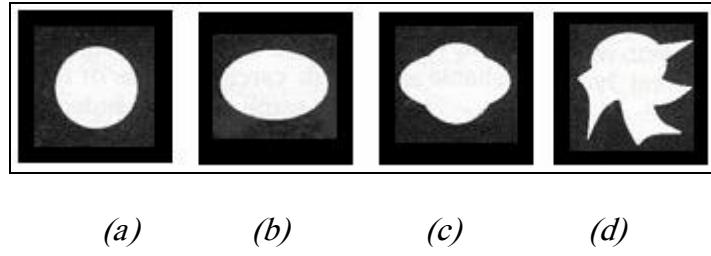


Fig. 4.6. Les régions à décrire : (a) rondes, (b) ovales, (c) lobulées et (d) irrégulières.

Les caractéristiques qui décrivent le contour ont pour objectif la distinction entre les contours stellaires, micro-lobulés et irréguliers (voir la figure 4.7). Pour cela on utilise la convexité (**CVX**), l'indice d'angle (**IA**) et trois autres caractéristiques à partir de la technique RDM qui sont la déviation standard (**σ**), la rugosité (**R**) et le rapport (**AR**).

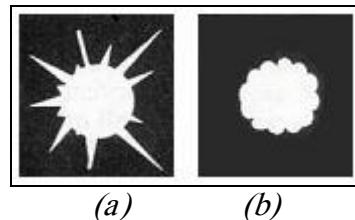


Fig. 4.7. Les contours à décrire : (a) stellaires et (b) microlobulés.

Concernant la description des contours indistincts et masqués (figure 4.8), plusieurs travaux utilisaient des caractéristiques de texture et d'intensité [45], [48] vu que leurs descriptions avec des caractéristiques de forme, sont difficile. Pour cette raison, Dans notre travail, on ne traite pas ce genre des contours.

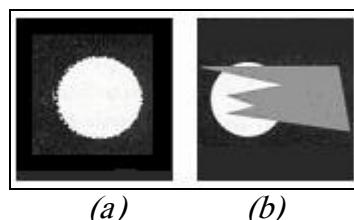


Fig. 4.8. Contours (a) indistincts et (b) masqués.

Finalement, on a huit caractéristiques dont chacune d'entre elles possède un rôle bien déterminé dans la phase de discrimination. Ces caractéristiques, ainsi que leurs rôles sont décrites dans la table ci-dessous :

	Caractéristique	Rôle
<i>Basées sur la région</i>		
(1)	$CIE = \frac{R_{gi}}{R_{pe}}$	Cette caractéristique donne une indication sur l'elongation de la forme. Elle est utile dans la distinction entre les formes rondes ou elliptiques et les formes irrégulières.
(2)	$C = \frac{4\pi A}{P^2}$	Cette caractéristique indique le degré de circularité de la forme.
(3)	$NRV = \frac{A(R)}{P(E)}$	Elle indique le degré de régularité de la région d'une masse. Elle permet la description des formes lobulaires, ovales et rondes.
<i>Basées sur le contour</i>		
(4)	$CVX = \frac{P(E)}{P(F)}$	Elle indique le degré de convexité du contour.
(5)	$\sigma = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N (d_n(i) - d_{moy})^2 \right)}$	Ces trois caractéristiques sont utiles dans la description des contours stellaires. Elles permettent la distinction entre les masses stellaires et celles qui ont des contours lisses.
(6)	$R = \frac{1}{N} \sum_{i=1}^N d_n(i) - d_n(i+1) $	
(7)	$A_R = \frac{1}{N * d_{moy}} * \sum_{i=1}^N (d_n(i) - d_{moy})$	
(8)	$IA = \frac{\sum_{i=1}^N \varphi_i}{\sum_{i=1}^N \theta_i}$	Elle donne une indication sur la régularité du contour et permet la description des contours microlobulés.

Tab 4.1. Liste des caractéristiques utilisées.

1.3. Classification

Une fois l'extraction des caractéristiques est effectuée par les méthodes étudiées dans la section précédente, on peut procéder par une classification des masses issues des images mammographiques en se basant sur le vecteur caractéristique. La classification des images mammographiques constitue la phase finale de notre travail qui sert à partitionner les masses en deux classes : les masses malignes et celles bénignes. Pour cela, il faut choisir le classifieur adéquat et le plus performant, c'est-à-dire, celui qui prédit correctement les classes.

Dans cette section, on donne les méthodes retenues qui serviront à une prise de décision afin d'aider le médecin dans son travail.

La procédure de classification sera réalisée automatiquement à partir d'une base d'exemples (un exemple consiste en une description d'un cas avec la classification correspondante). Un système d'apprentissage doit alors, à partir de cette base d'exemples, déterminer une procédure de classification. Le problème est donc un problème inductif ; il s'agit en effet d'extraire une règle générale à partir des données observées. La méthode devra classer les exemples d'échantillons mais surtout avoir un bon pouvoir prédictif pour classer correctement les nouvelles descriptions.

Il existe plusieurs méthodes de classification réparties en deux catégories : les méthodes supervisées et celles non supervisées.

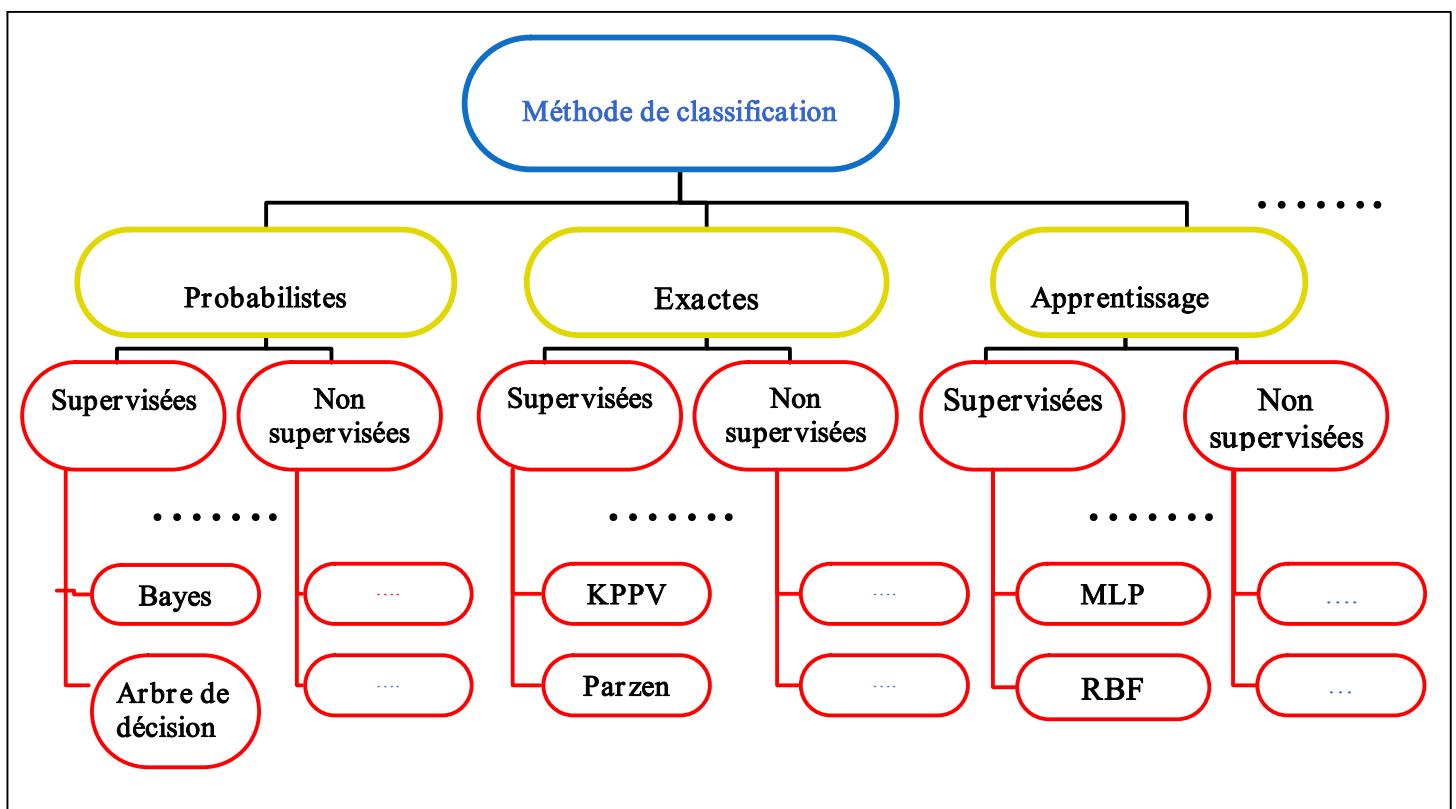


Fig. 4.9. Méthodes de classification.

La classification supervisée suppose qu'on connaît les classes possibles et qu'on dispose d'un ensemble d'instances déjà classées, servant de base d'apprentissage. Le problème est alors d'être capable d'associer à tout nouvel objet sa classe la plus adaptée en se servant des exemples déjà étiquetés. Dans ce sens, on peut citer quelques méthodes de classification automatique supervisée :

- Le classifieur KPPV (K-plus proches voisins ou K-NN) : c'est un classifieur exact.
- Le classifieur Bayes : c'est un classifieur probabiliste.
- Le classifieur MLP (Multi-Layer Perceptron) : qui est un classifieur basé sur l'apprentissage.

Dans la classification non supervisée, les classes possibles ne sont pas connues à l'avance et les exemples disponibles sont non étiquetés. Le but est donc de regrouper dans un même groupe (ou cluster) les objets considérés comme similaires, pour constituer les classes. L'apprentissage non supervisé ne peut fonctionner que si les données ont déjà une structure interne.

Dans notre travail on s'intéresse uniquement à la classification supervisée (voir la figure 4.9). Notre objectif n'est pas de proposer ou d'évaluer une nouvelle méthode de classification pour la prise de décision, mais d'utiliser une ou plusieurs méthodes qui permettent de répondre à nos besoins. En effet, pour notre étude, on a choisi des méthodes de différentes catégories : une de la classe des méthodes exactes (KPPV), une des méthodes probabilistes (Bayes) et une méthode de la classe basée sur l'apprentissage (MLP) pour pouvoir comparer et tirer des résultats intéressantes servant à affiner nos prédictions. Une comparaison entre les résultats de ces classificateurs sera détaillée dans ce qui suit afin de choisir le meilleur dans notre cas.

2. Evaluation des caractéristiques

En imagerie médicale, on utilise des critères afin d'évaluer la performance d'un classifieur. Les critères d'évaluation les plus utilisés en mammographie sont la sensibilité, la spécificité et le taux d'exactitude :

- La sensibilité est la capacité d'une telle technique à découvrir les positifs c'est à dire les cancers malins. L'équation de la sensibilité est donnée par :

$$\text{Sensibilité} = \frac{VP}{VP + FN} \quad (\text{Eq. 4.1.})$$

- La spécificité est la capacité d'une telle technique à découvrir les négatifs c'est à dire les cancers bénignes. L'équation de la spécificité est donné par:

$$\text{Spécificité} = \frac{VN}{VN + FP} \quad (\text{Eq. 4.2.})$$

- Le taux d'exactitude est la capacité d'une telle technique à découvrir les négatifs et les positifs à la fois. Elle est définie comme suit:

$$Exa = \frac{VN + VP}{VN + FP + VP + FN} \quad (\text{Eq. 4.3.})$$

Les paramètres utilisés par les critères d'évaluation sont résumés dans le tableau ci dessous :

Cas	Classifieur	En réalité
Faux Positif (FP)	maligne	bénigne
Faux négatif (FN)	bénigne	maligne
Vrai Positif (VP)	maligne	maligne
Vrai négatif (VN)	bénigne	bénigne

Tab 4.2. Résumé de l'ensemble de paramètres utilisés.

Afin de voir l'impact des classificateurs sur les résultats trouvés, on présente les différents résultats trouvés par chacune des trois classificateurs. Tout d'abord, on fait une comparaison entre les résultats trouvés par les caractéristiques basées sur la région et ceux trouvés par les caractéristiques basées sur le contour.

Pour évaluer le changement de calcul sur la technique RDM et sur la convexité CVX, on fait une comparaison entre les résultats trouvés par notre méthode et ceux trouvés par les autres travaux incluant, dans leurs approches, ces techniques. Après, on procède par une évaluation de la caractéristique (IA) basée sur le calcul d'angle.

Finalement, on donne quelques résultats générés de différentes approches d'analyse des masses mammographiques.

2.1. Base d'images utilisée

La Base d'images Numérisées de l'Université de Florida DDSM (The Digital Database for Screening Mammography) est une base de données d'images mammographiques digitalisée disponible en ligne [70]. Elle est divisée en 43 volumes:15 volumes malades, 16 bénignes et 12 normales. Au total, le nombre de cas est de 2620. À chaque cas est associé un ensemble d'informations concernant le patient (Age, densité, description pour les caractères anormaux...) et des informations concernant l'image (Scanner, résolution, ...). De plus, pour les images qui contiennent des régions suspectes, des informations sont associées sur l'emplacement et les types des régions.

Pour la phase de test, on a utilisé un ensemble d'images qui contiennent des masses sélectionnées de la base de données DDSM avec la composition suivante :

	<i>Cas maligne</i>	<i>Cas bénigne</i>
<i>Base</i>	120	120
<i>Test</i>	130	130

Tab 4.3. Résumé de l'ensemble d'image DDSM utilisées dans ce travail.

On a choisi d'utiliser la base de données DDSM vu qu'elle gratuite par opposition à d'autres bases telles que MIAS (The Mammographic Image Analysis Society).

2.2. Evaluation des caractéristiques basées sur le contour et sur la région

En regardant les résultats des caractéristiques basées sur le contour présentés par le tableau 4.4, on remarque un meilleur taux de détection des cas malignes (sensibilité), par rapport au taux de détection des cas bénignes (spécificité), pour les trois classificateurs utilisés.

Malgré que le classifieur Bayesien donne le meilleur taux de spécificité (94.80%) et le classifieur MLP donne le meilleur taux de sensibilité (97,90%) et le meilleur taux d'exactitude (95,98%), on ne peut pas dire qu'un classifieur est mieux qu'un autre vu que la différence entre les résultats donnés par les trois classificateurs n'est pas grande.

Caractéristiques basées sur le contour			
	Kppv	Bayes	MLP
Spécificité (%)	93,67	94,80	94,20
Sensibilité (%)	95,10	95,16	97,90
Taux d'Exa (%)	94,37	94,97	95,98

Tab 4.4. Résultat de test des caractéristiques basées sur le contour.

Concernant les résultats des caractéristiques basées sur la région, comme l'indique le tableau 4.5, on remarque un meilleur taux sensibilité donné par les trois classificateurs. Aussi, on remarque que le classifieur MLP donne le meilleur taux d'exactitude (95,58%) et le meilleur taux de spécificité (94,50%), alors que le classifieur Bayes mène à un meilleur taux de sensibilité (97,37%).

Caractéristiques basées sur la région			
	Kppv	Bayes	MLP
Spécificité (%)	92,96	90,33	94,50
Sensibilité (%)	95,45	97,37	96,70
Taux d'Exa (%)	94,17	93,57	95,58

Tab 4.5. Résultat de test des caractéristiques basées sur la région.

En comparant les résultats fournis par chaque classifieur dans les deux tables 4.4 et 4.5, on constate que les deux classifiants Kppv et Bayes donnent le meilleur taux de spécificité et d'exactitude pour le cas des caractéristiques basées sur le contour (respectivement, ils donnent le meilleur taux de sensibilité pour le cas des caractéristiques basées sur la région). Le classifieur MLP donne le meilleur taux de sensibilité et d'exactitude concernant les caractéristiques basées sur le contour (respectivement, pour le cas des caractéristiques basées sur la région, ils donnent le meilleur taux de spécificité).

Caractéristique de contour			
	Kppv	Bayes	MLP
Spécificité (%)	93,67	94,80	94,20
Sensibilité (%)	95,10	95,16	97,90
Taux d'Exa (%)	94,37	94,97	95,98
Caractéristiques de région			
	Kppv	Bayes	MLP
Spécificité (%)	92,96	90,33	94,50
Sensibilité (%)	95,45	97,37	96,70
Taux d'Exa (%)	94,17	93,57	95,58
Vecteur caractéristique (basé sur le contour et la région)			
	Kppv	Bayes	MLP
Spécificité (%)	94,53	96,04	95,63
Sensibilité (%)	97,10	97,55	96,74
Taux d'Exa (%)	95,78	96,78	96,18

Tab 4.6. Résultat de test de notre vecteur caractéristique basé sur le contour et sur la région.

L'apport réel de ce travail, réside dans la combinaison des caractéristiques basées sur la région et celles basées sur le contour. D'après la table 4.6, on remarque que les trois classificateurs utilisés nous donnent la meilleure performance. En effet, les taux des critères d'évaluations sont augmentés avec la combinaison des caractéristiques et par la suite leur fusion permet d'augmenter le taux de détection des tumeurs malignes et celles bénignes. Aussi, en comparant les résultats donnés par les trois classificateurs, on remarque que le classificateur Bayesien possède la meilleure sensibilité (97,55%), la meilleure spécificité (96,04%) et le meilleur taux d'exactitude (96,78%) par rapport aux classificateurs Kppv et MLP, malgré que la différence entre les performances de trois classificateurs n'est pas grande.

Finalement on peut déduire que l'utilisation d'une combinaison d'un descripteur basé sur la région et d'un autre basé sur le contour a amélioré la description des masses mammographiques.

2.3. Evaluation des caractéristiques RDM

Afin de montrer l'efficacité de RDM étendue, on a fait une comparaison (tableau 4.7) des résultats trouvés et ceux donnés par RDM [45], [47], [53]. On constate que les trois classificateurs utilisés donnent une meilleure sensibilité, une meilleure spécificité et un meilleur taux d'exactitude pour notre proposition. En effet, le taux de spécificité donné par RDM étendue est de 88% à 91% alors qu'il est de 83% à 89% par les caractéristiques RDM utilisées dans d'autres travaux. Le taux de sensibilité qu'on a trouvé est de 87% à 92% alors que taux est de 83% à 85% concernant les caractéristiques RDM d'autres travaux.

RDM			
	Kppv	Bayes	MLP
<i>Spécificité (%)</i>	89,74	83,93	86,88
<i>Sensibilité (%)</i>	85,22	83,93	85,43
<i>Taux d'Exa (%)</i>	87,34	83,93	86,14
RDM étendue.			
	Kppv	Bayes	MLP
<i>Spécificité (%)</i>	90,28	91,17	88,88
<i>Sensibilité (%)</i>	89,64	87,69	92,82
<i>Taux d'Exa (%)</i>	89,95	89,35	90,76

Tab 4.7. Résultats de test des caractéristiques RDM vs RDM étendue.

Ce qu'on peut déduire est que la modification appliquée sur la méthode RDM a augmentée l'efficacité de cette dernière dans la séparation des cas représentant des tumeurs malignes de ceux représentant des tumeurs bénignes.

Pour voir l'apport de l'utilisation des caractéristiques RDM dans notre vecteur caractéristique, on a fait une comparaison entre les résultats trouvés en utilisant la RDM étendue avec ceux trouvés en utilisant la RDM figurant dans la littérature. On constate d'après le tableau 4.8 que les deux classifiants Bayes et MLP donnent la meilleure performance dans le cas de la RDM étendue.

Notre vecteur caractéristique incluant la technique RDM			
	Kppv	Bayes	MLP
<i>Spécificité (%)</i>	96	94,16	94,86
<i>Sensibilité (%)</i>	96,37	97,09	96,32
<i>Taux d'Exa (%)</i>	96,18	95,58	95,58
Notre vecteur caractéristique incluant la technique RDM étendue			
	Kppv	Bayes	MLP
<i>Spécificité (%)</i>	94,53	96,04	95,63
<i>Sensibilité (%)</i>	97,10	97,55	96,74
<i>Taux d'Exa (%)</i>	95,78	96,78	96,18

Tab 4.8. Résultats de test avec intégration de la technique RDM.

Donc on peut dire que notre amélioration de calcul des caractéristiques RDM (voir le chapitre 3, section 4.1.1.) a amélioré davantage la description des masses mammographiques et la performance de classification. En effet, on a optimisé le temps de calcul et on a augmenté la performance des caractéristiques extraites de la méthode RDM.

2.4. Evaluation de la caractéristique d'angle (IA)

Le tableau 4.9 présente les résultats donnés concernant la caractéristique IA. Les résultats trouvés en utilisant les trois types de classification sont proches. En effet, les taux donnés par les trois critères d'évaluation permettent d'indiquer la réussite de cette caractéristique dans la description des contours microlobulés.

	Kppv	Bayes	MLP
<i>Spécificité (%)</i>	86,9	84,72	87,45
<i>Sensibilité (%)</i>	87,8	92,82	91,9
<i>Taux d'Exa (%)</i>	87,34	88,35	89,55

Tab 4.9. Résultats de test de la caractéristique d'angle (IA).

On ne peut vraiment pas comparer notre résultat trouvé avec les résultats trouvés pour d'autres caractéristiques d'angle, puisque les critères d'évaluation et la base d'images utilisés sont différents.

2.5. Evaluation de la caractéristique de convexité (CVX)

Comme dans le cas de la RDM, on a fait une comparaison des résultats trouvés lors de calcul de la convexité (CVX) comme indique le tableau 4.10. D'après cette dernière, on remarque que les deux classifiants Kppv et MLP donnent la meilleure performance concernant la caractéristique de convexité proposée. En effet, la différence entre les résultats trouvés concernant la caractéristique de convexité proposée et ceux concernant la convexité utilisée par d'autres travaux [45], [47], [53], est assez grande. Donc, l'amélioration apportée à la caractéristique de convexité a augmenté l'efficacité de la description d'une masse mammographique.

Le vecteur caractéristique incluant la caractéristique CVX			
	Kppv	Bayes	MLP
<i>Spécificité (%)</i>	88,67	86,64	87,54
<i>Sensibilité (%)</i>	90,9	95,92	95,55
<i>Taux d'Exa (%)</i>	89,75	90,76	91,16

Le vecteur caractéristique incluant la caractéristique CVX proposée.			
	Kppv	Bayes	MLP
<i>Spécificité (%)</i>	96	82,8	97,58
<i>Sensibilité (%)</i>	96,37	98,5	97,2
<i>Taux d'Exa (%)</i>	96,18	89,15	97,38

Tab 4.10. Résultats de test de la caractéristique de convexité (CVX).

2.6. Comparaison avec d'autres approches

Dans la littérature, on trouve plusieurs travaux qui étudiaient l'analyse de masses mammographiques [53], [57], [64]. On présente dans le tableau 4.11, ceux qui sont performants. Cependant, puisque les méthodes de description de forme, les bases de test et les systèmes de classification utilisés par les autres approches sont différents, on ne peut pas dire qu'on a le meilleur résultat malgré que les taux de sensibilité et de spécificité trouvés par notre travail soient les meilleurs.

D'après les résultats qu'on a trouvé et aussi les résultats trouvés dans d'autres travaux, on constate que les caractéristiques de forme peuvent caractériser les types des masses mammographiques. Vu que les signes de malignité de la tumeur de sein concernent la forme et la texture, les caractéristiques de forme restent insuffisantes toutes seules pour une description plus efficace. C'est pour cette raison qu'on trouve, en mammographie, pas mal d'approches [45], [48], [56], [59] qui prennent en considération les propriétés de malignité et de bénignité selon la texture des masses. Donc il vaut mieux d'ajouter des caractéristiques de texture à nos descripteurs pour augmenter encore le taux d'exactitude concernant la discrimination entre les masses bénignes et celles malignes.

	<i>Sensibilité (%)</i>	<i>Spécificité (%)</i>	<i>Exactitude (%)</i>	<i>Classifieur</i>	<i>Images utilisées</i>
Rangayyan et al. [57]	95	--		LDA	Base locale (39 cas)
A. Retico et al [53]	78,1	79,1		MLP	Base locale (226 cas avec 109 malignes et 117 bénignes)
R.Feng Chang et al [64]	88,89	92,5		SVM	Base locale (210 cas avec 90 cas malignes et 120 cas bénignes)
U. Bottigli et al [48]	88	--		MLP	Base CALMA (320cas)
Kilday [45]	69	--		LDA	Base locale (82 cas)
A. Alvarenga et al [47]	88,0	90,4	88,8	LDA	Base locale (152 cas)
Notre proposition	97,55	96,04	96,78	Bayes	Base DDSM (500 cas avec 250 malignes et 250 bénignes)

Tab 4.11. Comparaison avec d'autres approches.

3. Conclusion

Via ce chapitre on a donné un aperçu sur la méthode d'isolation des masses à partir d'une image mammographique. Cette méthode permet d'isoler deux composantes de l'image : une contenant la région de la masse et une autre contenant le contour. À partir de ces deux composantes, on a fait l'extraction des caractéristiques représentant notre descripteur de forme. Ce dernier sera l'entrée d'un classifieur permettant de prendre une décision concernant le type de la masse.

Dans notre travail, on a utilisé une base d'images DDSM et trois classificateurs de différent types afin d'évaluer notre descripteur de forme. Le classificateur Bayésien a donné la meilleure performance par rapport aux deux autres (MLP et KPPV). On a obtenu un taux de **96,04%** pour la spécificité, un taux de **97,55%** pour la sensibilité et un taux de **96,78%** concernant l'exactitude de classification. Aussi, la modification apportée à la méthode RDM et au calcul de la convexité a permis d'aboutir à une meilleure performance par rapport aux autres travaux qui utilisaient ces techniques.

Conclusions et perspectives

Dans ce mastère, on s'est intéressé à la réalisation d'un descripteur de forme pour l'aide à l'analyse des pathologies mammographiques. Afin de réaliser ce travail on a commencé par une étude bibliographique consacrée à la présentation de différentes pathologies de sein et leurs caractéristiques. On a trouvé que chaque pathologie possède ses caractéristiques de forme qui sont différentes aux autres. En raison de la difficulté de travailler sur toutes les caractéristiques de toutes les pathologies et la difficulté de les intégrer dans un seul descripteur de forme, on a choisi de travailler sur une seule pathologie qui est la masse. Cette dernière se caractérise par des propriétés de malignité et de bénignité selon le contour et la région.

Après l'étude des pathologies mammographiques on a donné un aperçu sur les méthodes les plus utilisées pour l'objectif de description des masses mammographiques.

Afin d'élaborer notre descripteur de forme, plusieurs approches ont été étudiées. Dans ce travail, on a procédé par une combinaison des caractéristiques basées sur la région et autres basées sur le contour. Les caractéristiques décrivant la région ont été choisies pour distinguer les formes rondes, ovales, lobulées et celles irrégulières. Parmi les caractéristiques utilisées dans la littérature on a choisi celles les plus performantes en terme de description et celles les plus simples à implémenter. Les caractéristiques basées sur le contour, qu'on a utilisé, ont été choisies pour la description de la frontière des masses et la distinction des contours réguliers de celles stellaires ou microlobulés. Dans ce cadre, on a ajouté une nouvelle caractéristique basée sur le calcul d'angle, on a optimisé le calcul des caractéristiques extraites à partir de la méthode RDM et on a modifié le calcul de la caractéristique de convexité.

Pour évaluer l'ensemble des caractéristiques proposées, on a utilisé trois types de classificateurs supervisés de différents types (Kppv, Bayes et MLP). Le classifieur Bayes a donné la meilleure performance avec un taux de l'ordre de **96,04%** pour la spécificité et un taux de l'ordre de **97,55%** concernant la sensibilité. Aussi, et grâce à notre optimisation de calcul, on a obtenu la meilleure

Conclusions et perspectives

performance (sensibilité et spécificité) pour les caractéristiques RDM étendue et la convexité modifiée, utilisées dans notre descripteur.

Comme perspectives de ce travail de recherche, il est possible de réaliser une méthode de détection automatique des masses mammographiques. Aussi la réalisation des caractéristiques de forme pour la description des autres pathologies de sein tel que les microcalcifications.

Glossaire

ALOE : Analysis of Local Oriented Edges.

AR : Area Ration.

ART : Adaptatif Resonance Therory.

BRCA : BReast Cancer Acronymes.

CAD : Computer Aided Detection.

CADx : Computer Aided Diagnosis.

CIE : Cercle Interne Externe.

CVX : ConVeXité.

DDSM : Digital Database for Screening Mammography.

DoG : Difference of Gauss.

FD : Fractal Dimension.

FDA : Food and Drug Administration.

FN : Faux Négatif ou False Negatif.

FP : Faux Positif ou False Positif.

IA : Indice d'Angle.

KNN : K-Nearest Neighbors.

KPPV : K Plus Proches Voisins.

LDA : Linear Discriminant Analysis.

MIAS : Mammographic Image Analysis Society.

MLP : Multi Layer Perceptron.

MRI : Magnetic Resonance Imaging.

NRV : Normalised Residual Value.

RBF : Radial Basis Function.

RDM : Radial Distance Measure.

SDEV : Standard DEViation.

SVM : Support Vector Machine.

VN : Vrai Négatif.

VP : Vrai Positif.

ZC : Zero Crossing.

Bibliographie

- [1] L. Zhang, R. Sankar et W. Qian, “*Advances in micro-calcification clusters detection in mammography*”, Computers in Biology and Medicine, vol. 32, no. 6, pp 515-528, 2002.
- [2] S. Yu et L. Guan, “*A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films*”, IEEE Transactions on Medical Imaging, vol. 19, n°. 2, pp. 115-126, 2000.
- [3] Lansac, “*Cancer du sein*”, Le monde, pp. 19, 9 octobre 1985.
- [4] M. Lanyi, “*Diagnosis and Differential Diagnosis of Breast calcifications*”, Springer-Verlag, Berlin, Heidelberg, 1988.
- [5] S. Shapiro, P. Venet, P. Strax et R. Roeser , “*Ten to fourteen years effect of breast cancer screening on mortality*”, J Natl Cancer Inst, vol. 69, pp. 349-355, 1982.
- [6] Le Gal, J.C. Durant, Laurent et M. Pellier, “*Conduite à tenir devant une mammographie révélatrice de microcalcifications groupées, sans tumeur palpable*”, la nouvelle presse Médicale, vol. 5, pp. 1623-1627, 1976.
- [7] KELSY, “*A review of the epidemiology of human breast cancer*”, Epidemiol, vol. 15, pp. 36-47, 1993.
- [8] H. Guillemet, H. BENALI, F. Preteux et R. DI Paola, “*Noisy Fractional Brownian Motion for Detection of Perturbations in Regular Textures*”, Proceedings SPIE Conference on Statistical and Stochastic Methods for Image Processing, vol. 2823, pp. 40-51, Denver, CO, 1996.
- [9] L.Tabár, C. Fagerberg, A. Gad, L. Balderup, L.H. Holmberg et O.Grøntoft, “*Reduction in mortality from breast cancer after mass screening with mammography: Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare*”, The Lancet, vol. 325, no. 8433, pp. 829-832, 1985.
- [10] B.K Rimer, N. Resch, E. King, E. Ross, C. Lerman, A. Boyce, H. Kessler et P.F. Engstrom, “*Multistrategy health education program to increase mammography use among women ages 65 and older*”, Public Health Rep, vol. 107, no. (4), pp. 369-380, 1992.
- [11] B.Séradour, M.H.Dilhuydy, “*Dépistage organisé des cancers du sein : un enjeu de santé publique*”, Springer éd, pp.158, 2003.

- [12] H.D. Cheng, C. Xiaopeng, C. Xiaowei, H. Liming et L. Xueling, “*Computer-aided detection and classification of microcalcifications in mammograms : a survey*”, Pattern Recognition, vol. 36, no. 12, pp. 2967-2991, 2003.
- [13] S.F. Hurley, J.M. Kaldor, “*The benefits and risks of mammographic screening for breast cancer*”, Epidemiologic Reviews, vol.14, pp.101-30, 1992.
- [14] C. Marx, B. Schutze, M. Fleck, K. O'Shaughnessy et W.A. Kaiser, “*Computer aided diagnosis in mammography*”, European Journal of Radiology, vol. 7, pp. 7-82, 1997.
- [15] Y-H. Chang, B. Zheng, David Gur, “*Computer-aided detection of clustered microcalcifications on digitized mammograms: a robustness experiment*”, Acad Radiol., vol 4, no. 6, pp. 415-418, 1997.
- [16] A. Malich, C. Marx, M. Facius, T. Boehm, M. Fleck et W.A. Kaiser, “*Tumour detection rate of a new commercially available computer-aided detection (CAD) system*”, European Journal of Radiology, vol. 12, pp. 14-20, 2001.
- [17] E. Thurfjell, M.G. Thurfjell, E. Egge et N. Bjurstam, “*Sensitivity and specificity of computer-assisted breast cancer detection in mammography screening*”, Acta Radiol., vol. 39, pp. 384-388, 1998.
- [18] C. Marx, A. Malich, M. Facius, U. Grebenstien, D. Sauner, S. Pleiderer et W.A. Kaiser, “*Are unnecessary follow-up procedures by computer-aided diagnosis (CAD) in mammography ? Comparaison of mammographic diagnosis with and without use of CAD*”, European Journal of Radiology., vol. 51, pp. 66-72, 2003.
- [19] D. Wei, H. P. Chan, N. Petrick, B. Sahiner, M.A. Helvie, D.D. Adler et M.M. Goodsitt, “*Falsepositive reduction technique for detection of masses on digital mammograms: global and local multiresolution texture analysis*”, Medical Physics, vol. 24, pp. 903-14, 1997.
- [20] G. M. Brake, N. Karssemeijer et J.H. Hendriks, “*An automatic method to discriminate malignant masses from normal tissue in digital mammograms*”, Physics in Medicine & Biology, vol. 45, pp. 2843-2857, 2000.
- [21] G.D. Tourassi, R. Vargas-Voracek, D.M. Catarious, et C.E. Floyd, “*Computer-assisted detection of mammographic masses: a template matching scheme based on mutual information*”, Medical Physics, vol. 30, pp. 2123-2130, 2003.
- [22] S.L. Liu, C.F. Babbs et E.J. Delp, “*MultiResolution Detection of spiculated Lesions in Digital Mammograms*”, IEEE Transactions on Image Processing, vol. 10, no. 6, pp. 874-884, 2001.
- [23] Z. Huo, M.L. Giger, C.J. Vyborny, U. Bick, P. Lu, D.E. Wolverton et R.A. Schmidt, “*Analysis of spiculation in the computerized classification of mammographic masses*”, Medical Physics, vol. 22, no. 10, pp.1569-1579, 1995.

- [24] C. Varela, P.G. Tahoces, A. J. Méndez, M. Souto et J.J. Vidal , “Computerized detection of breast masses in digitized mammograms ”, Computers in Biology and Medicine, vol. 37, no. 2, pp. 214-226, 2007.
- [25] W.P. Kegelmeyer, J.M. Pruneda, P.D. Bourland, A. Hillis, M.W. Riggs et M.L. Nipper, “*Computer-aided mammographic screening for spiculated lesions*”, Radiology, vol. 191, pp. 331-337, 1994.
- [26] H. Kobatake et Y. Yoshinaga, “*Detection of spicules on mammogram based on skeleton analysis*”, IEEE Transactions on Medical Imaging, vol. 15, no. 3, pp. 235-245, 1996.
- [27] H.D. Li, M.~Kallergi, L. P. Clarke, V.K. Jain, et R.A. Clark, “*Markov random field for tumor detection in digital mammography*”, IEEE Transactions on Medical Imaging, vol. 14, no. 3, pp. 565-576, 1995.
- [28] R.P. Velthuizen, “*Computer diagnosis of mammographic masses*”, IEEE Transactions In Proc. of the App. Imagery Pattern Recognition Workshop, pp. 166-172, 2000.
- [29] W.E. Polakowski, D.A. Cournoyer, S.K. Rogers, M.P. DeSimio, D.W. Ruck, J. W. Hoffmeister et R. A. Raines, “*Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency*”, IEEE Transactions on Medical Imaging, vol. 16, pp.811-819, 1997.
- [30] W. Qian, L. Li, L. Clarke, R. A. Clark, et J. Thomas, “*Comparison of adaptive and non adaptive cad methods for mass detection*”, Academic Radiology, vol. 6, pp. 471-480, 1999.
- [31] D. Guliato, R.M. Rangayyan, W.A. Carnielli, J.A Zuffo et J.E.L Desautels, “*Segmentation of breast tumors in mammograms by fuzzy region growing*” In Proc International Conference IEEE Engineering in Medecine and Biology Society”, vol. 20, pp. 1002-1005, 1998.
- [32] T. Matsubara, H. Fujita, T. Hara, S. Kasai, O. Otsuka, Y. Hatanaka, et T. Endo, “*Development of a new algorithm for detection of mammographic masses*”, In Proc. International Workshop on Digital Mammography, pp. 139-142, 1998.
- [33] B.R. Groshong et W. P. Kegelmeyer, “*Evaluation of a Hough Transform Method for Circumscribed Lesion Detection*”, In Digital Mammography, pp. 361-366, 1996.
- [34] N.R. Mudigonda, R.M. Rangayyan et J.E.L. Desautels, “*Detection of breast masses in mammograms by density slicing and texture flow-field analysis*”, IEEE Transactions on Medical Imaging, vol. 20, pp. 1215-1227, 2001.
- [35] W. Zhang, K. Doi, M.L. Giger, R.M. Nishikawa et R.A. Schmidt, “*An improved shift-invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms*”, Medical Physics., vol. 23, pp. 595-601, 1996.

- [36] R. Prokop. et A. Reeves, “*A survey of moment based techniques for unoccluded object representation and recognition*”, Graphical Models and Image Processing, vol. 54, no 5, pp. 438-460, 1992.
- [37] N. Petrick, H.P. Chan, B. Sahiner, D. Wei, D.D. Alder et M.A. Helvie, “*Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification*”, Medical Physics, vol. 23, pp.1685-1696, 1996.
- [38] H. Qi et W.E. Snyder, “*Lesion detection and characterization in digital mammography by Bézier histograms*”, In Proc. International Conference IEEE Engineering in Medicine and Biology Society, vol. 2, pp. 1021-1024, 1998.
- [39] B. Sahiner, H.-P. Chan, N. Petrick, D. Wei, M.A. Helvie, D.D. Adler, et M.M. Goodsitt, “*Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images*”, IEEE Transactions on Medical Imaging, vol. 15, pp. 598-610, 1996.
- [40] B. Sahiner, H.P. Chan, N. Petrick, M.A. Helvie, et L.M. Hadjiiski, “*Improvement of mammographic mass characterization using spiculation measures and morphological features*”, Medical Physics, vol. 28, pp. 1455-1465, 2001.
- [41] M.K. Hu, “*Visual pattern recognition by moment invariants*”, IRE Trans.Inform. Theory, vol. IT-8, no. 2, pp. 179–187, 1962.
- [42] D. Wei, H.P. Chan, M.A. Helvie, B. Sahiner, N. Petrick, D.D. Adler et M.M. Goodsitt, “*Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis*”, Medical Physics, vol. 22, pp. 1501-13, 1995.
- [43] L.M. Bruce et R.R. Adhami, “*Classifying mammographic mass shapes using the wavelet transform modulus-maxima method*”, IEEE Transactions on Medical Imaging, vol. 18, pp. 1170-7, 1999.
- [44] C.H. Chen et G.G. Lee, “*On digital mammogram segmentation and microcalcification detection using multiresolution wavelet analysis*”, Graphical Models and Image Processing, vol. 59, pp. 349-364, 1997.
- [45] J. Kilday, F. Palmieri et M.D. Fox, “*Classifying mammographic lesions using computerized image analysis*”, IEEE Transactions on Medical Imaging, vol. 12, pp. 664-9, 1993.
- [46] S. Pohlman, K.A. Powell et N.A. Obuchowski, “*Quantitative classification of breast tumors in digitized mammograms*”, Medical Physics, vol. 23, pp.1337-1145, 1996.
- [47] A. V. Alvarenga, A. F. Catelli Infantosi, W. C. de Albuquerque Pereira, C. M. de Azevedo, “*Assessing the performance of the normalised radial length and convex polygons in distinguishing breast tumours on ultrasound images*”, Pereira Revista Brasileira de Engenharia Biomédica, ISSN 1517-3151, vol. 22, no. 3, pp. 181-189, 2006.

- [48] U. Bottigli, D. Cascio, F. Fauci, B. Golosio, R. Magro, G.L. Masala, P. Oliva, G. Raso, et S. Stumbo “*Massive Lesions Classification using Features based on Morphological Lesion Differences*”, Proceedings of World Academy Of Science, vol. 12, ISSN 1307-6884, 2006.
- [49] R.M. Rangayyan, D. Guliato, J.D. Carvalho, et S.A. Santiago, “*Feature extraction from the turning angle function for the classification of breast tumors*”, In Int. Special Topics Conf. Inf. Technol.Biomed, Greece, 2006.
- [50] R.M. Rangayyan, N.R. Mudigonda et JEL. Desautels, “*Boundary modelling and shape analysis methods for classification of mammographic masses*”, Medical and Biological Engineering and Computing, vol. 38, no. 5, pp.487–496, ISSN 0140-0118, 2000.
- [51] T.M. Nguyen et R.M. Rangayyan, “*Shape analysis of breast masses in mammograms via the fractal dimension*”, In Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society , pp. 3210-3213, Shanghai, China, 2006.
- [52] T. Matsubara, H. Fujita, S. Kasai, M. Goto, Y. Tani, T. Hara et T. Endo, “*Development of new schemes for detection and analysis of mammographic masses*”, In Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS’97), pp. 63–66, Grand Bahama Island, Bahamas,1997.
- [53] A. Retico, P. Delogu, M.E. Fantacci et P. Kasaé, “*An Automatic System to Discriminate Malignant from Benign Massive Lesions on Mammograms*”, Medical Physics, vol. 14, pp. 596-600, 2007.
- [54] C. Chettaoui, K. Djamel, A. Djouak et H. Maaref, “*Etude de formes des globules drépanocytaires par traitement numérique des images*”, In Int. Conf. SETIT, Tunisia, 2005.
- [55] T.M. Nguyen et R.M. Rangayyan , “*Fractal analysis of contours of mammographic masses*”, In Third IASTED International Conference on Biomedical Engineering, pp. 186–191, Innsbruck, Austria, 2005.
- [56] L. Shen, R.M. Rangayyan, et J.E.L. Desautels, “*Application of shape analysis to mammographic calcifications*”, IEEE Transactions on Medical Imaging, vol.13, no. 2, pp. 263–274, 1994.
- [57] R.M. Rangayyan, N.M. El-Faramawy, J.E.L. Desautels, et O.A. Alim, “*Measures of acutance and shape for classification of breast tumors*”, IEEE Transactions on Medical Imaging, vol. 16, no. 6, pp. 799-810, 1997.
- [58] M.K. Hu, “*Visual pattern recognition by moment invariants*,” IRE Trans.Inform. Theory, vol. IT-8, no. 2, pp. 179–187, 1962.
- [59] A.V. ALVARENGA, A.F.C. INFANTOSI, W.C.A. PEREIRA, C.M. AZEVEDO, “*Morphometric and Texture Parameters in Distinguishing Breast Tumours on Ultrasound Images*”, In: WC2006,

World Congress on Medical Physics and Biomedical Engineering, IFMBE Proceedings. Korea : IFMBE, vol. 14, pp. 2156-2159, Seoul, 2006.

[60] J.K. Kim, J.M. Park, K.S. Song et H.W. Park, “*Adaptive mammographic image enhancement using first derivative and local statistics*”, IEEE Transactions on Medical Imaging, pp: 495–502, vol.16, no. 5, 1997.

[61] N. Székely, N. Tóth et B. Pataki, “*A Hybrid System for Detecting Masses in Mammographic Images*”, in *Proc. of the IEEE Instrumentation and Measurement Technology Conference, IMTC '2004*, pp. 2065-2070, Italy, 2004.

[62] A. Tréneau, Ch. Fernandez-Maloigne et P. Bonton, “*Images numériques Couleur*”, DUNOD, pp. 100-113, 2004.

[63] H. Freeman, “On the encoding of arbitrary geometric configurations”, IRE Trans. Electron. Comput. EC-10, pp. 260-268, 1961.

[64] R.F. Chang, W.Wu, W.K. Moon, et D.R.Chen, “*Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors*”, Springer, vol. 89, no. 2, pp.179–185, 2005.

[65] D. Zhang et G. Lu, “*A Comparative Study of Fourier Descriptors for Shape Representation and Retrieval*”, In Proc. of the Fifth Asian Conference on Computer Vision, pp. 646-651, Melbourne, Australia, 2002.

[66] P. S. P. Wang, C. H. Chen et L.F. Pau, “*The Handbook of Pattern Recognition and Computer Vision*”, (2nd Edition), pp. 207-248, World Scientific Publishing Co, 1998.

[67] S. Singh et R. Al-Mansoori, “*Identification of regions of interest in digital mammograms*”, Journal of Intelligent Systems, vol. 10,no. 2, pp. 183-217, 2000.

[68] F. Djidet et F. Boumghar, “*Détection de Lésions Stellaires par la Multirésolution et les Réseaux de Neurones*”, 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, Tunisia , 2005.

[69] N. Székely, N. Tóth, L. Lasztovicza et B. Pataki, “*Combining Methods For Mass Detection In Mammograms*”, In Proceedings of the 17th Biennial International EURASIP Conference BIOSIGNAL, pp. 281-283, Brno, Csehország, 2004.

[70] M. Heath, K. Bowyer, D. Kopans, R. Moore et P.J. Kegelmeyer, “*The Digital Database for Screening Mammography*”, In Proc. International Workshop on Digital Mammography, pp. 212-218, 2000.

[71] J. Zunic et P.L. Rosin. “*A New Convexity Measure for Polygons*”. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 7, pp. 923 – 934, 2004.

[72] A.V. ALVARENGA, J.L.R. Macrini, W.C.A. PEREIRA, C.E. PEDREIRA, A.F.C. INFANTOSI, “*Breast Ultrasound Images Classification Using Morphometric Parameters Ordered by Mutual*

Information”, In Proc of the 11th Mediterranean Conference on Medical and Biological Engineering and Computing - MEDICON2007, Ljubljana - Slovenia. IFMBE Proceedings, vol. 16, pp. 1025-1029, 2007.

[73] O. Otsu, “A threshold selection method from gray-level histogram,” *IEEE Trans. on System, Man and Cybernetics*, vol. 9, pp. 62–66, 1979.

[74] A. Boujelben, H. TMAR, J. MNIF et M. ABID, “*Influence des caractéristiques issues de la matrice de co-occurrence sur le diagnostic du cancer du sein*”, In 2th Workshop AMINA : Application Medicale en Informatique: Nouvelle Approche, pp.9-14, Monastir, Tunisia, 2006.

Elaboration d'un descripteur de forme en vue de classification des clichés mammographiques

Ali Chérif CHAÂBANI

الخلاصة : إن تفعيل آلية الأنظمة دفع الباحثين في مجال الإعلامية إلى التطرق إلى تطوير أدوات المساعدة في التشخيص والكشف عن سرطان الثدي. في هذا السياق، قمنا باقتراح طريقة لوصف شكل الكوبيرات المومغرافية (*masses*) من أجل المساعدة في تشخيص سرطان الثدي. إن واصف الشكل الذي أنجزناه يجمع بين خصائص تستمد من منطقة الشكل و أخرى تستمد من حدود الشكل من أجل ضمان أفضل وصف.

خلال هذا العمل، قمنا بإدخال تغيرات على طريقة *RDM* التي تعتمد قيس المسافات، بين مركز الشكل والنقط الموجدة عند حدوده، قصد تحسين التعقيد الزمني. كما اقترحنا خاصية مؤشر الزاوية (*IA*) التي تعتمد على حساب الزاوية لوصف الأشكال (*microlobulées*) التي لها أقواس مقرعة عند حدودها. خصائص أخرى مثل مؤشر دائرة الشكل (*C*) و مقياس تحدب الشكل (*CIE*) و مؤشر التندد (*NRV*)، وقع استخدامها لوصف منطقة الشكل.

لتقدير فعالية واصف الشكل، استخدمنا قاعدة بيانات الصور *DDSM* و ثلاثة أنواع مختلفة من المصنفين (*Kppv* ، *Bayes* ، *MLP*). المصنف « *Bayesian* » قدم أفضل أداء مع نسبة (96.04%) من حيث دقة كشف الأورام الحميدة ، وبنسبة (97.55%) من حيث دقة كشف الأورام الخبيثة.

Résumé : L'automatisation a incité les informaticiens à élaborer des outils d'aide au diagnostic et à la détection du cancer de sein. Dans ce contexte, on s'est intéressé à la proposition d'une méthode de description de forme pour aider au diagnostic du cancer de sein. Pour mieux caractériser la masse, notre descripteur de forme combine des caractéristiques basées sur la région et autres basées sur le contour. Pendant ce travail, on a modifié la méthode de la mesure de la distance radiale (*RDM*) afin de réduire encore l'ordre de complexité temporelle. Aussi, on a proposé une caractéristique, basée sur le calcul d'angle, notée Indice d'Angle (*IA*) afin de mieux décrire les masses microlobulées. Autres caractéristiques, comme la circularité (*C*), la valeur moyenne (*NRV*) et l'indice d'elongation (*CIE*), ont été utilisées pour représenter la région.

Afin d'évaluer l'efficacité de notre descripteur de forme, on a utilisé une base d'images *DDSM* et trois classificateurs supervisés de différent types (*Kppv*, *Bayes* et *MLP*). Le classifieur Bayesian a donné la meilleure performance avec un taux de 96,04% en terme de spécificité et un taux de 97,55% en terme de sensibilité.

Abstract : Automation has prompted scientists to develop tools to cope with breast cancer detection and diagnosis. In this context, we proposed a new method to describe the shape to assist in diagnosis of breast cancer. For better characterizing the mass, our shape descriptor combines shape and boundary features. During this work, we modified the Radial Distance Measure (*RDM*) method in order to reduce the temporal complexity. Also, we proposed a feature, based on the measure of angle, denoted (*IA*) to describe the microlobulated masses. Other features like the circularity (*C*), the average value (*NRV*) and the index of elongation (*CIE*) are used to represent the region.

To evaluate the effectiveness of our shape descriptor, we used a *DDSM* database and three supervised classifiers (*Kppv*, *Bayes* and *MLP*). The Bayesian classifier gave the best performance in terms of specificity (96.04%) and sensitivity (97.55%).

المفاتيح : المؤوبستي , تحليل , واصف الشكل , سرطان الثدي , مموجرافيا.

Mots clés : Masse; Analyse; Descripteur de forme; Cancer du sein; Mammographie.

Key-words : Mass; Analysis; Shape descriptor; Breast cancer; Mammography.