

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et populaire

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

جامعة سعد دحلب البليدة
Université SAAD DAHLAB de BLIDA

كلية التكنولوجيا
Faculté de Technologie

قسم الإلكترونيك
Département d'Électronique



Mémoire de Projet de Fin d'Études

présenté par :

HAOULIA Salima

Pour l'obtention du diplôme de Master en Électronique option :

« TRAITEMENT DE L'INFORMATION ET SYSTEMES ELECTRONIQUES »

Thème

Détection morphologique et texturale de pathologies mammographiques pour l'aide à la décision

Proposé par : Reguieg F.Zohra & Benblidia Nadja

Année Universitaire 2012-2013

Dédicaces

À Mes parents pour leur contribution, leur soutien et leur **PATIENCE**,
jamais je ne pourrais les remercier assez.

À mes frères et sœurs : **Rabia, Sabiha, Sarah et Chakib** ;

À mes beaux-frères : Michael alias Adel et Sofiane ;

À mes amis de la classe **B2.3** et **C1.1** de l'institut français d'Alger (**I.F.A**) ;

À mes camarades et amis de L'École Militaire Polytechnique (les **GE**, les informaticiens, les chimistes et les mécaniciens), en particulier les **SPÉCIAUX** avec qui j'ai passé les pires et meilleurs moments de ma vie.

À l'incorrigible **Amel**, à **Hadia** l'autoritaire, à **Houda** la force bleue, (bien que ça sonne mieux en Arabe) à **Hind** : la **meilleure** binôme de tous les temps... la preuve après notre séparations je n'ai pas pu travailler en binôme ^^ je lui dédie ce travail en mémoire du bon vieux temps.

À Meriem, ma meilleure amie d'enfance.

À mon directeur Monsieur SEBTI Abdelouahab qui m'a permis de finir mon projet en m'accordant un congé d'une durée indéterminée ainsi qu'à SEBTI Madjdi pour le PC.

Une spéciale dédicace va à mes deux tantes chéries : **Khaltou** et **Souad**, qui m'ont toujours aidée et soutenue depuis mon enfance. Par leurs prières et leurs encouragements, j'ai pu surmonter tous les obstacles.

À ma grand-mère, **merci** pour tes prières qui m'ont accompagné tout au long de ma formation .Que **Dieu** te garde et te **protège**.

Pour finir j'adresse mes plus sincères remerciements à tous mes proches et amis, qui m'ont toujours soutenue et encouragée au cours de la réalisation de ce mémoire.

Merci à tous et à toutes :-)



Remerciements

En préambule à ce mémoire je remercie **Allāh** tout puissant et miséricordieux, qui m'a aidée et m'a donnée la patience et le courage durant ces longues années d'études et d'avoir répondu à mes prières.

Je tiens à remercier sincèrement mes promotrices Reguieg F.Zohra & Benblidia Nadja, pour le sujet, l'orientation, la confiance, la patience, les encouragements qui ont constitués un apport considérable sans lequel ce travail n'aurait pas pu être mené à bon port.

Un énorme merci à mes trois mousquetaires : **Google**, **You Tube** et **Yahoo! Questions/Réponses**, qui m'ont sauvée à maintes reprises.

Aux administrateurs de la page **Facebook** : « **Télécharger les articles scientifiques gratuitement** » Qui ont eu l'amabilité et la bonté de m'aider dans ma recherche bibliographique en me téléchargeant des articles scientifiques.

Un remerciement particulier va à Monsieur **Ben Laiid**, le rédacteur d'un tutoriel portant sur : **comment construire une interface graphique sous MATLAB**, grâce à dieu et à lui, et d'autre vidéos sur YouTube (38 milles vidéos à peu près et figurez-vous que certaines étaient en espagnol ou en italien...) j'ai pu apprendre les bases du guide sous **MATLAB**, entre autres.

J'exprime ma gratitude à tous les consultants et internautes rencontrés sur différents forums lors des recherches effectuées et qui ont accepté de répondre à mes interrogations.

Je suis reconnaissant envers tous ceux qui m'ont dit non, car grâce à eux,
je l'ai fait moi-même. - Albert Einstein

ملخص: تجزئة الصور الماموجرافية تلعب دورا هاما في عزل المناطق المحتمل أن تكون الأورام ويتم تحديد هذه المناطق في ثلاث خطوات: تعزيز النقيض ، عزل المناطق اللامعة وأخيرا تحليل الملمس وشكل للمناطق ذات الاهتمام لتصنيف كثافات. نستخدم ال region growing للكشف عن الكثافة. المناطق مجزأة تميز وفقا للصفات التكوينية و الشكلية وتصنف من قبل ناقلات آلات الدعم (SVM)، التي تميز فئتين ورم خبيث و ورم حميد ، لمساعدة التشخيص.

كلمات المفاتيح: صور الماموجرام؛ البنية التشكلية؛ region growing ؛ الشكل؛ النسيج؛ ناقلات آلات الدعم؛ المساعدة في التشخيص.

Résumé : La segmentation des images mammographiques joue un rôle majeur, dans l'isolation des zones susceptibles d'être des tumeurs. L'identification de ces zones est faite en trois étapes dans le cadre de ce mémoire : Un rehaussement de contraste, l'isolation des régions de surdensités et enfin l'analyse de la texture et de la forme des régions d'intérêt en vue d'une classification des masses. Nous utilisons l'algorithme de la croissance de région pour la détection de masses. Les régions d'intérêts segmentées, sont caractérisées suivant les descripteurs texturaux et topologiques et sont classifiées, par les machines à support de vecteurs SVM, qui les discriminent en deux classes : les masses malignes et la masses bénignes, pour une aide au diagnostic clinique.

Mots clés : image mammographique, morphologie mathématique, croissance de région, forme, texture, classification par les SVM, aide au diagnostic.

Abstract: The segmentation of mammograms plays a major role in isolating areas, which can be subject to tumors. The identification of these zones is done in three stages: contrast enhancement, overdense regions detection and finally texture and shape analysis of the regions of interest for a classification of masses. We use region growing algorithm for the masses detection. Classification of these structures is accomplished through Support Vector Machines, which separate them into two groups, Benign Tumors and Malignant Tumors, using shape and texture descriptors to aid clinical diagnosis.

Keywords: mammographic image; mathematical morphology; region growing; shape, texture; classification by SVM; computer-aided diagnosis

Table des matières

Introduction générale	1
-----------------------------	---

Chapitre 1

3

Introduction à la mammographie et cancer du sein

3

1.1 Anatomie mammaire	4
1.1.1 Cancers du sein non invasifs	5
1.1.2 Cancers du sein invasifs	6
1.2. Mammographie.....	7
1.2.1. Naissance de la radiologie	7
1.2.2 Définition de la mammographie	8
1.2.3 Circonstances de l'examen.....	8
1.2.4 Récepteurs d'image.....	9
1.2.4.1 Couple écran-film.....	10
1.2.4.2 Mammographie numérique	11
1.2.5 Les artefacts dans une mammographie	13
1.2.5.1 Étiquettes du film mammographique.....	14
1.2.5.2 Artefacts radiopaques	15
1.3 Anomalies mammaires.....	16
1.3.1 Masse.....	16
1.3.1.1 Les masses régulières	18
1.3.1.2 Les masses irrégulières	19
a. Distorsion architecturale	19
b. Masse avec des limites en spicules.....	19
1.3.2 Asymétrie de densité.....	19
1.3.3. Calcifications mammaires.....	20
1.3.3.1. Localisation anatomique.....	21
1.3.3.2. Analyse de la taille	22
1.3.3.3. Analyse du nombre.....	22
1.3.3.4. Analyse de la morphologie et de la texture.....	22
1.4 Conclusion	23

Chapitre 2 26

Techniques de segmentation en imagerie mammographique 26

2.1 Introduction	27
2.2 Système de détection automatique	27
2.2.1 Prétraitement des images	29
2.2.2 Généralités sur la segmentation	30
2.2.2.1. Définition	30
2.2.2.2 Travaux antérieurs sur la segmentation	32
2.3.1 Germes initiaux	34
2.3.2 Un processus itératif	34
2.3.3 Convergence, critère d'arrêt	35
2.4. Propriétés de la croissance de région.....	36
2.5. Paramètres de la croissance de régions	37
2.5.1. Les germes.....	37
2.5.2 Points candidats à l'agglomération	38
2.5.3 Croissance de régions sans concurrence	39
2.6. Morphologie mathématique.....	40
2.6.1 Chapeau haut de forme et chapeau bas de forme	41
2.6. Notre approche.....	42
2.6 Conclusion	44

Chapitre 3 47

Caractérisation et classification en imagerie mammaire 47

3.1 Extraction de caractéristiques.....	48
3.1.1 Les caractéristiques de texture	48
3.1.1.1 Notion de texture.....	49
3.1.1.2 Les matrices de cooccurrence	51
3.1.2 Les caractéristiques de formes.....	55
3.1.2.1 Compacité (ou facteur de circularité).....	55
3.1.2.2 Convexité	56
3.1.2.3 Excentricité	57
3.1.2.4 La rectangularité	57
3.2 Classification des images.....	59

3.3 Machines à vecteurs support.....	61
3.3.1 Principe.....	61
3.3.2 Classification multi-classes.....	63
3.3.2.1 Un contre tous.....	63
3.3.2.2 Un contre un.....	64
Chapitre 4.....	68
Application, résultats et discussion.....	68
4.1 Environnement de travail.....	69
4.1.1 Matériel utilisé.....	69
4.1.2 Qu'est-ce que MATLAB ?.....	69
4.2 Présentation du système développé.....	74
4.2.1 Images utilisées.....	77
4.2.2 Interface de l'application.....	78
4.3 Résultats et discussion.....	80
4.3.1 Rehaussement du contraste.....	81
4.3.2 Détection des masses.....	82
4.3.3. Extraction des caractéristiques.....	83
4.3.4 Classification.....	92
4.4 Conclusion.....	95
Conclusion générale.....	97

Liste des figures

FIGURE 1. 1 : ANATOMIE DU SEIN	4
FIGURE 1. 2 : CELLULES CANCÉREUSES.....	5
FIGURE 1. 3 : POSITIONNEMENT DE LA PATIENTE POUR LA PRISE D'UNE MAMMOGRAPHIE DE FACE.....	8
FIGURE 1. 4 : LES DIFFÉRENTS TYPES DE CAPTEURS ET TECHNOLOGIES EN MAMMOGRAPHIE NUMÉRIQUE	10
FIGURE 1. 5 :RELATION ENTRE LE LOGARITHME DE L'EXPOSITION ET LA DENSITÉ OPTIQUE.....	11
FIGURE 1. 6 : RÉSUMÉ DES TECHNOLOGIES DE DÉTECTION	13
FIGURE 1. 7 : UN EXEMPLE DE MAMMOGRAPHIES CONTENANT DES ARTEFACTS RADIOPAQUES.	14
FIGURE 1. 8 : EXEMPLES D'ÉTIQUETTES DU FILM MAMMOGRAPHIQUE	15
FIGURE 1. 9 : EXEMPLES DE MARQUEURS OPAQUES SUR UNE MAMMOGRAPHIE	15
FIGURE 1. 10 : DIFFÉRENTES FORMES DE MASSES	17
FIGURE 1. 11 : TYPES DE CONTOURS DES MASSES	17
FIGURE 1. 12 : DIFFÉRENTES DENSITÉS DE MASSES.....	18
FIGURE 1. 13 : MICROCALCIFICATIONS	21
FIGURE 1. 14 : MACROCALCIFICATIONS	22
FIGURE 2. 1 : SCHÉMA GÉNÉRAL D'UNE CHAÎNE DE DÉTECTION	28
FIGURE 2. 2: PRINCIPE DE LA CROISSANCE DE RÉGION SUR UNE IMAGE 2D	34
FIGURE 2. 3 : SEGMENTATION D'UNE IMAGE SYNTHÉTIQUE DE 256 NIVEAUX DE GRIS.	35
FIGURE 2. 4: SENSIBILITÉ DE LA CROISSANCE DE RÉGION AUX POINTS DE FUITES	37
FIGURE 2. 5 : SENSIBILITÉ DE LA CROISSANCE DE RÉGION VIS À VIS D'UN SEUIL.....	39
FIGURE 2. 6 : ALGORITHME DE SEGMENTATION DES MASSES.	42
FIGURE 3. 1 : INFORMATION TEXTURE ET INFORMATION COULEUR.....	46
FIGURE 3. 2 : PLUS PROCHES VOISINS DU PIXEL « X » SELON 4 DIRECTIONS.....	47
FIGURE 3. 3 : EXEMPLE DE MATRICES DE COOCCURRENCE	48
FIGURE 3. 4 : DIVERSES FORMES ET VALEURS CORRESPONDANT À LA COMPACTITÉ	52
FIGURE 3. 5 : ENVELOPPE CONVEXE.	52
FIGURE 3. 6 : EXCENTRICITÉ	53
FIGURE 3. 7 : ORGANIGRAMME D'EXTRACTION DES CARACTÉRISTIQUES	54
FIGURE 3. 8 : EXEMPLES DE SÉPARATION LINÉAIRE ET NON LINÉAIRE.	57
FIGURE 3. 9 : EXEMPLE DE SVM	57
FIGURE 3. 10 : EXEMPLE DE SVM.....	58
FIGURE 3. 11 : CONFIGURATION EXPLICITE DES SVM	59
FIGURE 3. 12 : ORGANIGRAMME DE PRISE DE DÉCISION.	61
FIGURE 4. 1 : ICÔNE DE MATLAB.....	69
FIGURE 4. 2: FENÊTRE DU GUI	71
FIGURE 4. 3: SYNOPTIQUE DE L'APPLICATION « MORPHOTEXTUROMAMM ».....	74
FIGURE 4. 4: IMAGES MAMMOGRAPHIQUES	77
FIGURE 4. 5: INTERFACE PRINCIPALE	78
FIGURE 4. 6 : APERÇU DE L'EXÉCUTION DE L'APPLICATION.....	79
FIGURE 4. 7 : CAS D'UN TRAITEMENT	80
FIGURE 4. 8: IMAGE ORIGINALE: MDB184 : MALIGNES.....	80
FIGURE 4. 9: ÉTAPES DU PRÉTRAITEMENT	81
FIGURE 4. 10: SENSIBILITÉ DE LA CROISSANCE DE RÉGION VIS À VIS D'UN SEUIL SUR QUATRE CAS DE MASSES	82
FIGURE 4. 11: SENSIBILITÉ DE LA CROISSANCE DE RÉGION VIS À VIS D'UN SEUIL POUR L'IMAGE :.....	83
FIGURE 4. 12: INFLUENCE DES CARACTÉRISTIQUES SUR DES MASSES BÉNIGNES.....	86
FIGURE 4. 13: INFLUENCE DES CARACTÉRISTIQUES SUR DES MASSES MALIGNES.	87
FIGURE 4. 14 : INFLUENCE DU CONTRASTE ET DE LA VARIANCE SUR DES MASSES MALIGNES.....	88
FIGURE 4. 15: INFLUENCE DU CONTRASTE ET DE LA VARIANCE SUR DES MASSES BÉNIGNES.	89
FIGURE 4. 16 : INFLUENCE DE L'ÉNERGIE SUR DES MASSES MALIGNES ET BÉNIGNES.	90
FIGURE 4. 17: CLASSIFICATION D'UNE MASSE BÉNIGNE.....	93
FIGURE 4. 18: CLASSIFICATION D'UNE MASSE MALIGNES	93
FIGURE 4. 19 : TAUX DE CLASSIFICATION.	94

Liste des tableaux

TABLEAU 1. 1 : DESCRIPTION D'UNE MASSE DANS LE CATALOGUE BI-RADS	16
TABLEAU 3. 1 : SÉLECTION DES DESCRIPTEURS TEXTURAUX.....	50
TABLEAU 3. 2 : SÉLECTION DES DESCRIPTEURS MORPHOLOGIQUES.....	51
TABLEAU 4. 1 : CARACTÉRISTIQUES DE L'IMAGE mdb135.....	83
TABLEAU 4. 2 : CARACTÉRISTIQUES DE QUELQUES MASSES BÉNIGNES.	85
TABLEAU 4. 3 : CARACTÉRISTIQUES DE QUELQUES MASSES MALIGNES.....	85



Une femme debout et une femme à genoux se regardent. Celle qui est debout a été à genoux et celle qui est à genoux sera debout. L'une est le futur de l'autre. Et leur regard dit cette certitude : oui, on se relève, oui on est tous un jour celui qui demande de l'aide, oui le cancer peut être dans le parcours d'une vie, la date d'une renaissance, une aventure qui vous emmène loin, c'est-à-dire plus près de vous-même. Ce fut mon aventure.

Hélène JOUSSE, sculpteur

Chapitre 4

Application, résultats et discussion

L’objectif de ce chapitre, est de présenter l’application développée, les traitements effectués sur des images mammographiques, dans le but de l’extraction et de la classification des masses, suivant les machines à support de vecteurs, en vue d’une aide à la décision.

4.1 Environnement de travail

Notre système, a été développé sous environnement Windows 7 Edition intégrale, à l'aide de l'outil de programmation **MATLAB** 7.11 (R2011a).

4.1.1 Matériel utilisé

L'implantation de nos codes, a été réalisée, grâce à un PC portable de type **Acer Aspire 7520 - AMD Turion 64 X2 1.9 GHz - 2 Go RAM - DD 120 Go - Ecran 17" LCD** .

4.1.2 Qu'est-ce que MATLAB ?

MATLAB (Matrix Laboratory) est un logiciel basé, sur le calcul matriciel. Son noyau est composé de bibliothèques écrites au début en Fortran, puis en **C++**.

Il permet:

- le calcul numérique.
- le tracé de graphiques.
- la programmation.
- l'interfaçage avec d'autres langages (**Fortran, C++**).
- le développement d'interface utilisateur (**GUI**).
- l'utilisation des boîtes à outils (Toolboxes) spécialisées.

L'icône du logiciel (cf. Figure 4.1) apparaît, dès qu'on l'ouvre.

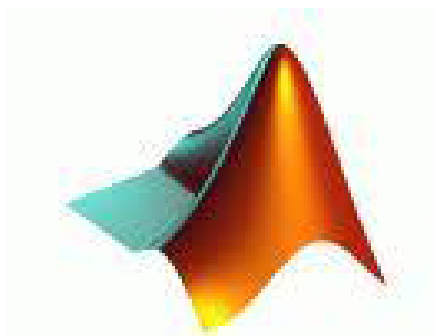


Figure 4. 1 : icône de MATLAB

MATLAB présente deux modes de fonctionnement:

1. le mode interactif : exécution des instructions entrées, par l'utilisateur dans la fenêtre de commande de **MATLAB** (le Workspace) ;
2. le mode exécutif : exécution de **programmes** ou **scripts**, qui sont des fichiers regroupant une suite d'instructions (les fichiers M).

● Première utilisation de Matlab

Une façon efficace, de découvrir Matlab est d'utiliser son aide en ligne:

- **help** : "help" tout seul donne la liste des aides générales possibles;
- **helpwin** : ouvre une fenêtre et donne accès à une aide détaillée ;
- **help + nom d'une commande**

Exemple : >> help plot , indique la syntaxe des graphes en 2D;

- **demo**: lance une démo générale de Matlab;
- **help demos** : donne une liste des démos existantes.

Quelques commandes importantes :

- **lookfor + nom de commande** : donne une liste de toutes les commandes, qui ont un rapport avec le nom de commande écrite ;
- **who** et **whos** : donne les noms de variables présentes dans l'espace de travail;
- **clear** : supprime les variables de l'espace du travail;
- **clc** : efface l'écran (sans toucher aux variables !);
- **exit** ou **quit** : permet de sortir de Matlab ;
- **CTRL + C** : (touches du clavier) permet d'interrompre un programme (ie : récupérer la main).

● Présentation du GUI

Le **GUI (Graphical User Interface)** permet de créer des interfaces où, le créateur choisit plusieurs types d'objets (*boutons, edit box, listbox...*) appelés *handles*. Ensuite, il doit réaliser la programmation pour obtenir l'interaction, qu'il souhaite obtenir entre ces

différents objets. La première étape, est la création d'un GUI en tapant le mot clé GUIDE, dans la fenêtre de commande **MATLAB** (cf. Figure 4.2).

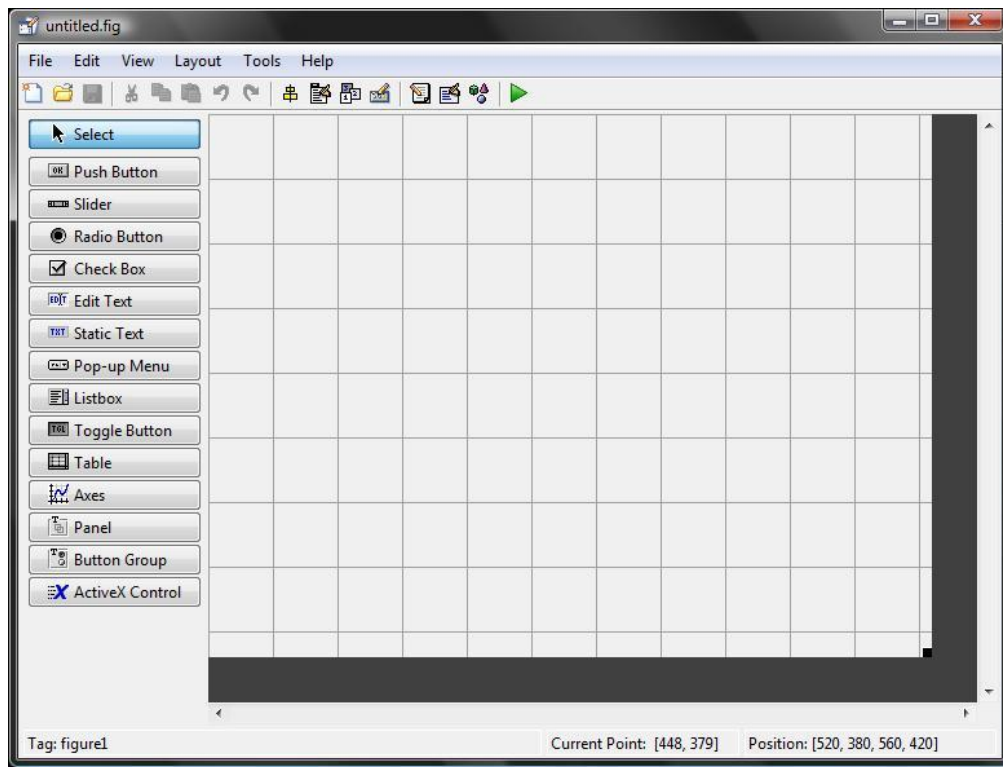


Figure 4. 2: fenêtre du GUI

● *Palette des composants visuels*

Une palette de divers composants, est offerte par matlab:

✓ « *Push buttons* » : (*boutons poussoirs*)

- Un seul état (stable).
- Identifié avec un court texte (paramètre 'String').
- Événement à traiter: clic ou relâchement du bouton gauche de la souris.

✓ « *Checkboxes* » : (*cases à cocher*)

- Deux états possibles.
- Boutons d'un même groupe sont indépendants (plusieurs options, peuvent être sélectionnées simultanément).

✓ « *Radio buttons* » : (*boutons radio*)

- Deux états (comme case à cocher).

- Boutons d'un même groupe sont mutuellement exclusifs (une seule option, peut être sélectionnée à la fois).
 - ✓ « Frames » : (cadres)
- Bordure rectangulaire délimitant un groupe de contrôles.

« Statictext » : (champ de texte fixe)

- Affichage de texte.
 - ✓ « Edittext » : (champ de texte éditable)
 - Saisie de texte.
 - Pas pour affichage.
 - ✓ « Sliders » : (barres de défilement)
 - Choix d'une valeur numérique à l'intérieur d'un intervalle.
 - Peuvent être orientées horizontalement ou verticalement.
 - Événement à traiter : déplacement de la barre.
 - ✓ « Listboxes » : (listes)
 - Choix d'un item parmi une liste.
 - Un groupe d'items, dont l'item sélectionné, est affiché.
 - Événement à traiter: item choisi par clic de la souris.
 - ✓ Menus : (à partir du MenuEditor)
 - Choix d'un item parmi, une liste d'options permises ou non.
 - Possibilité de clé de raccourcis.
 - Possibilité de sous-menus.
 - Événement à traiter: item de menus, choisi par clic de souris.
 - ✓ Axes :
 - Permet d'afficher un graphique tracé par **MATLAB**
 - Permet d'afficher des fichiers image
 - Événement à traiter: l'utilisateur clique sur la souris lorsque, le curseur se trouve à l'intérieur des bornes du composant.
- **Accès aux données associées aux composants visuels**
- À tout objet d'interface, Matlab associe un pointeur, qui permet d'accéder aux propriétés, le handler.

Certains handlers sont déjà définis par défaut:

- **gcf**: handler de la figure courante (getcurrent figure).
- **gca**: axes courants de tracé (getcurrent axes).
- **gcbf**: la figure cliquée (get callback figure).
- **gcbo**: l'objet qui appelle (celui sur lequel on a cliqué).
- **0** : la fenêtre interpréteur Matlab, ou racine.

Certaines instructions nécessitent un handler:

- **getet set** : respectivement pour lire et pour changer les valeurs des propriétés.
- **findobj**: pour retrouver le handler d'un objet.
- **propedit**: pour éditer les propriétés d'un objet.
- **delete**: pour effacer un objet.

4.2 Présentation du système développé

Le système (cf. Figure 4.3) que nous avons développé a pour objectif final, la classification des masses mammographiques moyennant les **SVM**. L'identification de ces zones, s'est faite par les étapes suivantes:

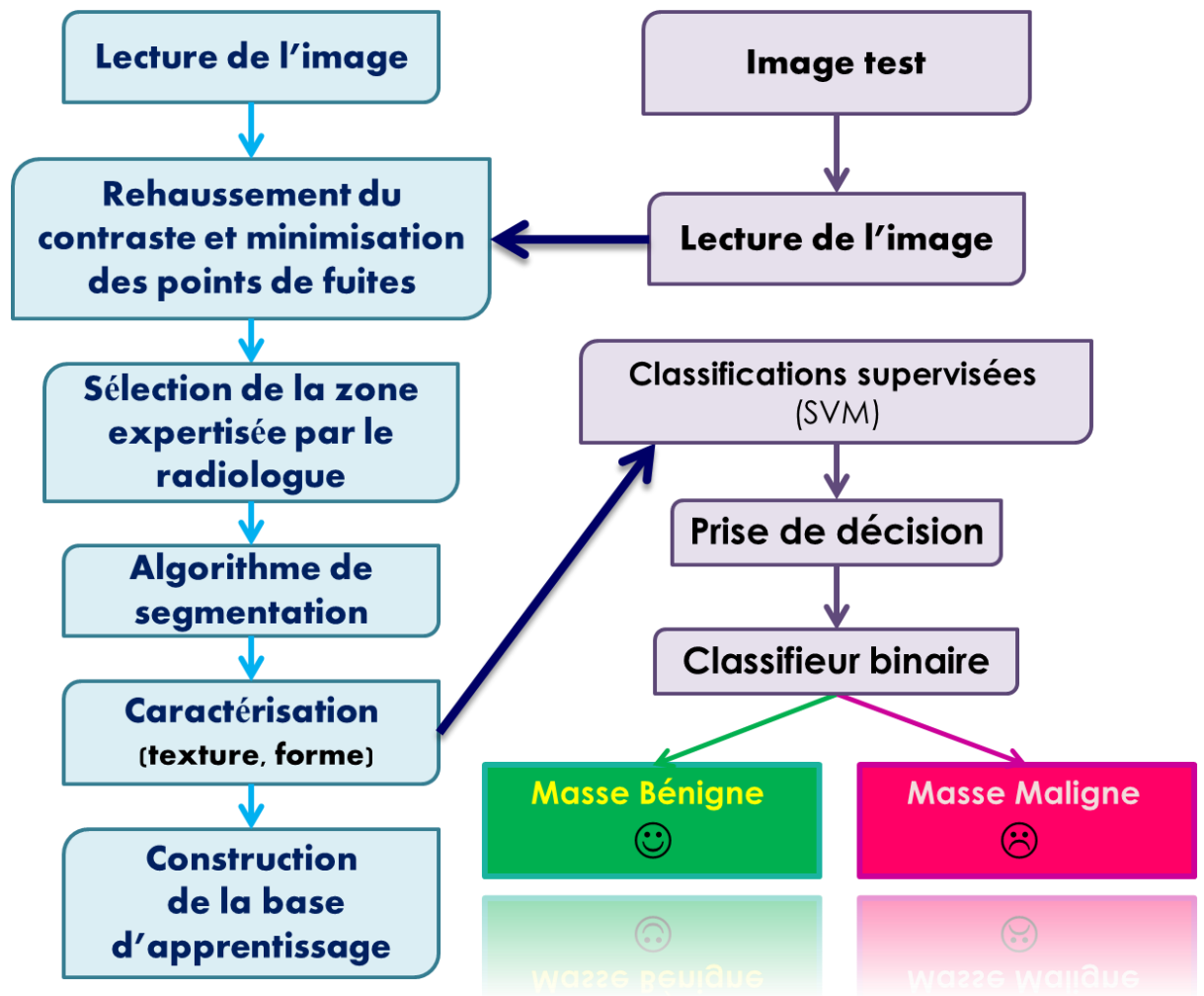


Figure 4. 3: synoptique de l'application « Morphotexturomamm »

Première étape

Nous avons utilisé les fonctions, de la morphologie mathématique pour faire un rehaussement du contraste, afin de faciliter la détection des masses sur la mammographie.

● Chapeau haut de forme (Black Top Hat)

Pour extraire les vallées (correspondant aux zones sombres) d'épaisseur inférieures à n , on a effectué la différence, entre la fermeture de l'image par un élément structurant B de taille n et l'image initiale cf. Figure 4.9 (a)).

● Chapeau bas de forme (White Top Hat)

En effectuant la différence entre l'image initiale et, son ouverture par un élément structurant B de taille n , on extrait les pics (correspondant aux zones claires) de l'image dont l'épaisseur, est inférieure à n (cf. Figure 4.9 (b)).

Nous avons pris, pour les deux cas, un élément structurant B : en forme de diamant et un autre, en forme de disque, afin de considérer les formes des pathologies mammographiques.

Nous avons fait varier sa taille¹ de $n= 100$ pixels, 110 pixels et 120 pixels

NB : Une taille (de l'élément structurant) inférieure à 90 pixels risquerait, d'éliminer les régions d'intérêts (masses) rendant ainsi la détection impossible.

Nous avons obtenu de meilleurs résultats avec l'élément structurant, en forme de diamant.

● Soustractions

La soustraction entre le résultat du Chapeau haut de forme et, celui du Chapeau bas de forme (cf. Figure 4.9 (c)) élimine, toute structure inférieure à l'élément structurant; afin de minimiser les points de fuite. Un point de fuite, se définit par la connexion de deux régions distinctes, aux propriétés statistiques semblables.

¹ Où n indique la distance de l'origine de l'élément structurant à son extrémité.

Pour un meilleur contraste, nous avons superposé le résultat de la soustraction, à lui-même. (cf. Figure 4.9 (d)).

Deuxième étape

Cette étape, permet la segmentation par l'algorithme de croissance de régions (pour plus de détails, voir plus loin), afin de détecter les régions d'intérêt.

Troisième étape

En isolant la région issue de l'algorithme de la croissance de régions, on obtient ainsi la région segmentée.

Quatrième étape

Cette étape, permet l'extraction des caractéristiques texturaux et topologiques, pour la région segmentée.

Cinquième étape

● Classification

La classification effectuée par le classifieur des SVM, permet l'identification du tissu normal et anormal, en utilisant des descripteurs de texture et de forme. L'apprentissage réalisé est en mode supervisé, car les images utilisées sont déjà diagnostiquées par des experts. Les résultats obtenus sont très encourageants, dans la mesure où une amélioration dans l'apprentissage peut réduire, le nombre de faux positifs.

Nous avons comparé chaque région segmentée avec la base d'apprentissage, suivant une mesure de la similarité (considérée par la distance euclidienne). Celle-ci a été créée, en ne considérant que les régions pathologiques expertisées pas des radiologues.

4.2.1 Images utilisées

● Images de la base MIAS² [1]

Dans notre étude, nous avons utilisé 92 images mammographiques provenant de la base d'images **MIAS** (Mammographic Image Analysis Society). Ces images de type MLO (incidence latérale oblique) sont numérisées avec une résolution de **50 μm** par pixel sur **8 bits**. La résolution spatiale de chaque image est de **1024 x 1024 pixels**. Les **92 images** sont réparties en **53 images** présentant des masses bénignes et **39 images** présentant des masses malignes. Ces pathologies ont été identifiées par les radiologues experts (cf. Figure 4.4)



Figure 4. 4: Images mammographiques

(a) mdb150 : distorsions architecturales, (b) mdb186: lésions spiculées

²MIAS:Mammographic Image Analysis Society

4.2.2 Interface de l'application

La Figure 4.5 présente l'interface principale, que nous avons baptisée « **Morphotexturomamm** ». Elle contient divers menus :

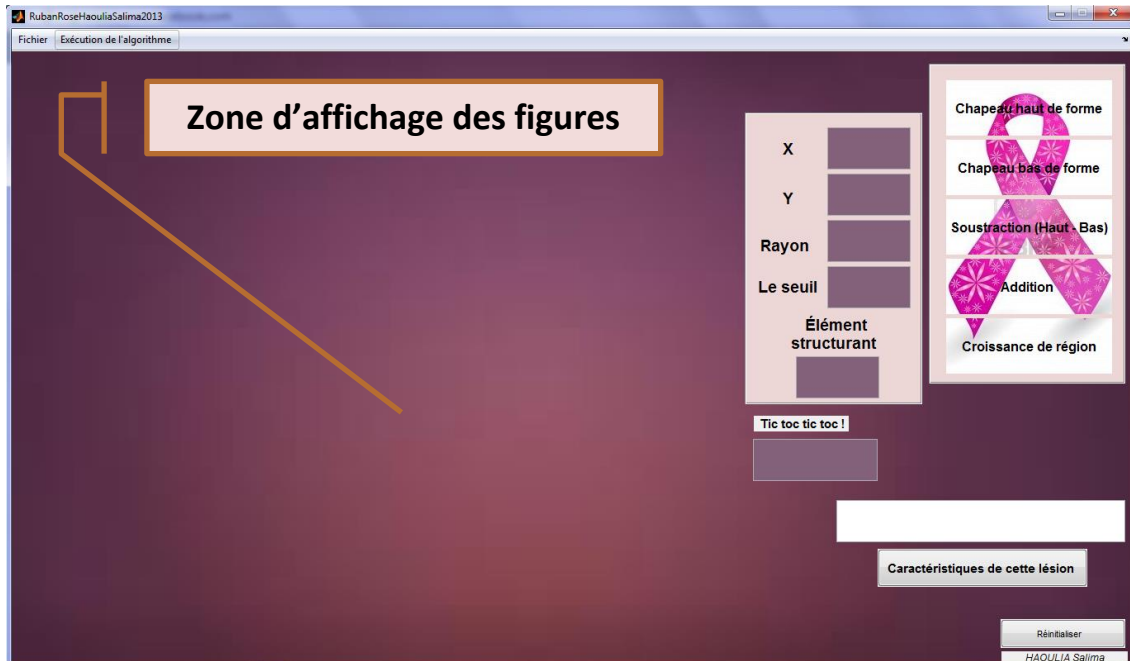


Figure 4. 5: Interface principale

● Fichier

Ce module réalise les opérations suivantes:

Ouvrir : Pour lire les images et les afficher;

Quitter : pour quitter l'application.

● Exécution de l'algorithme

Ce module, effectue toutes les opérations de l'algorithme que nous avons développées:

- le prétraitement;
- la segmentation en croissance de région;
- la caractérisation des régions d'intérêt, suivant des paramètres texturaux et topologiques;
- la classification selon la nature de la pathologie (maligne, bénigne).

Le résultat de chaque étape, nous est possible, en cliquant sur les boutons relatifs à chaque étape du programme; comme l'illustre la figure 4.6. On peut également, introduire les Coordonnées (X, Y) du centre de l'anomalie ainsi que le rayon approximatif (en pixels) d'un cercle entourant l'anomalie et, enfin le seuil.

- Après que l'exécution du programme soit finie, le type de la pathologie s'affiche sur la fenêtre principale pour nous dire, si la masse est maligne ou bien bénigne. Le temps d'exécution en secondes, est également donné. On peut également, introduire la taille de l'élément structurant, selon le type de masse (figure 4.6).
- le bouton « caractéristique de cette lésion » ouvre une autre fenêtre. Cette dernière, nous permet d'afficher les caractéristique texturales et topologiques de la région segmentée (figure 4.7).

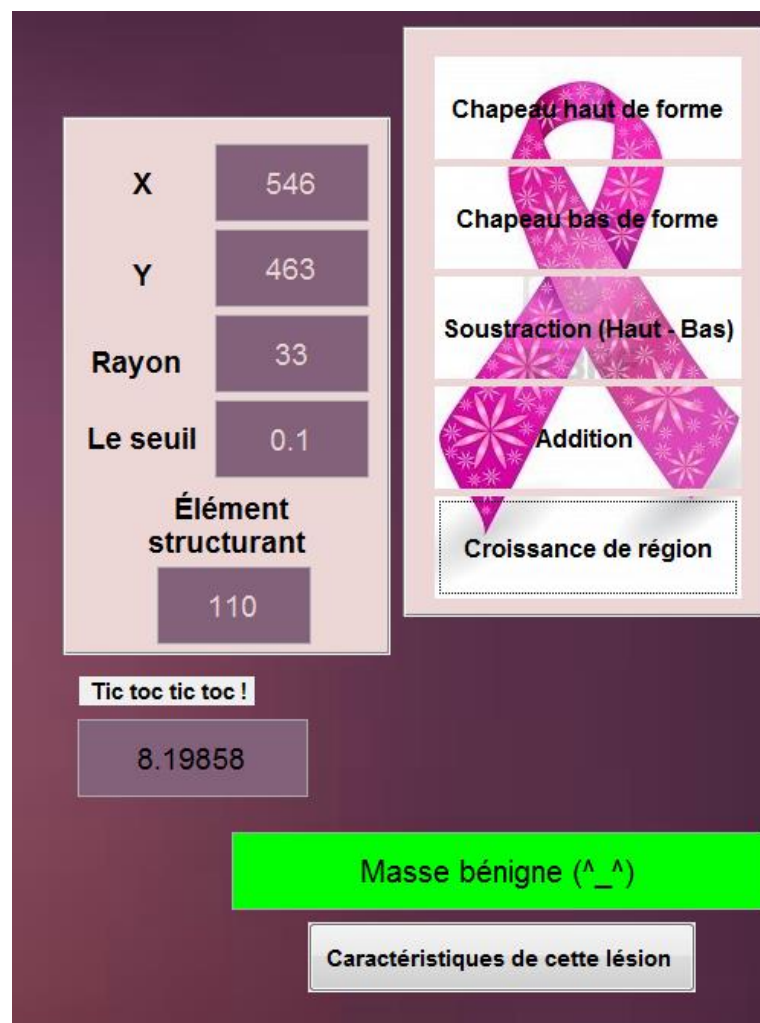


Figure 4. 6 : aperçu de l'exécution de l'application

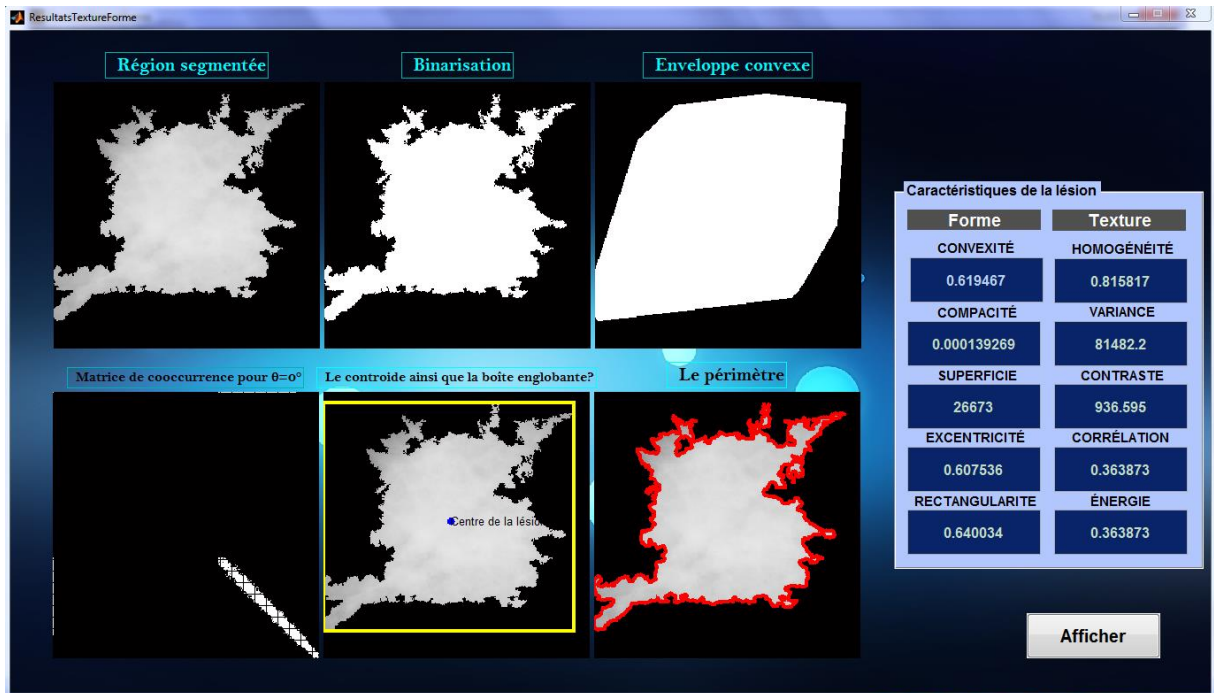


Figure 4. 7 : cas d'un traitement

4.3 Résultats et discussion

L'application « texturomorphomamm » parcourt l'analyse texturale et topologique des masses (cf. Figure 4.8), en vue d'une classification de ces données pour une aide à la décision.

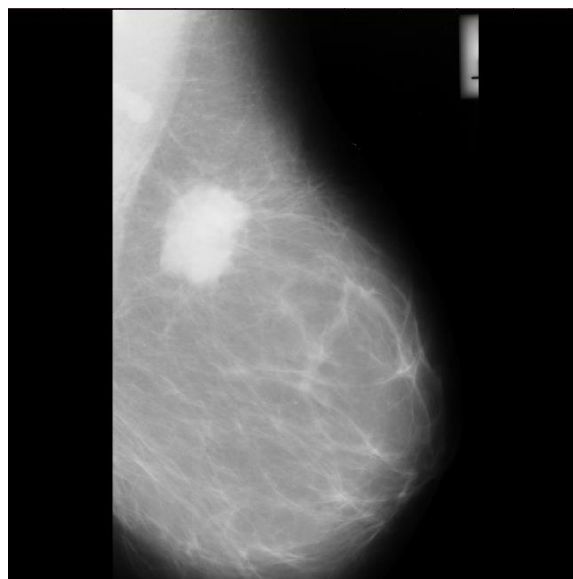


Figure 4. 8: image originale: mdb184 : Maligne

4.3.1 Rehaussement du contraste

Nous présentons dans la figure 4.9, les différentes étapes de l'amélioration de la visualisation, sur le cas de l'image de la figure 4.7 ; qui représente un cas malin.

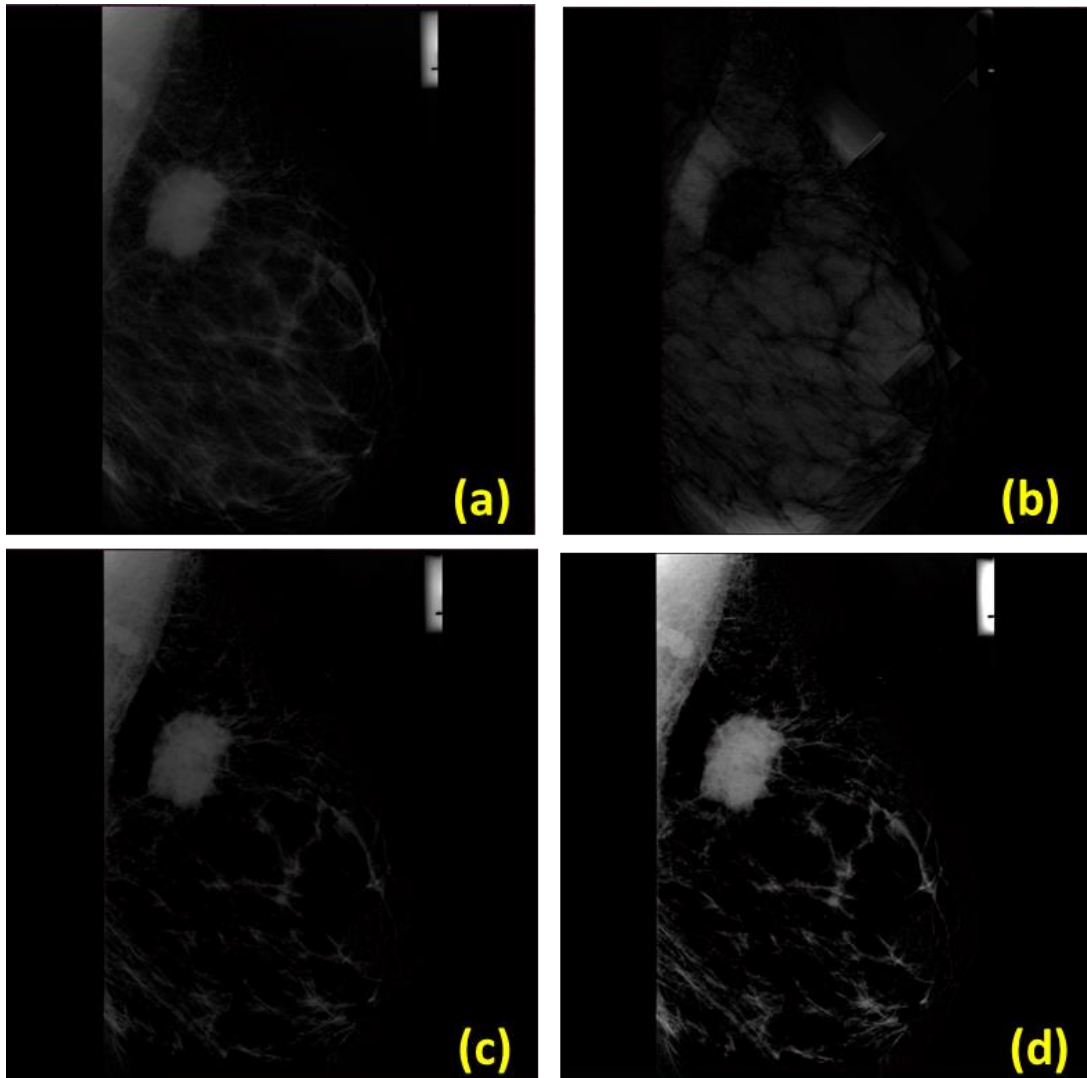


Figure 4. 9: étapes du prétraitement

(a)Chapeau haut de forme, (b) Chapeau bas de forme
(c) soustraction (a)-(b), (d) addition (c)+(c)

4.3.2 Détection des masses

La détection des masses, s'est faite par une technique de segmentation par croissance de régions. L'approche proposée pour segmenter l'image, consiste à faire croître chaque région autour d'un pixel de départ. Nous avons pris comme germe, le centre de la tumeur désignée par le médecin radiologue;

La différence entre la valeur de l'intensité d'un pixel et la moyenne de la région, est utilisée en tant que mesure de similitude. Le processus s'arrête, lorsque la différence d'intensité entre la région moyenne et, le nouveau pixel (relation 4.1) devient plus grande qu'un certain seuil « t ». Celui-ci a été désigné de manière empirique.

$$\boxed{|I(x, y) - M| \leq t} \quad 4.1$$

La valeur du seuil qui a fonctionné pour pratiquement toutes les images, vaut $t = 0.13$

Pour les autres cas, ce seuil a varié entre $0.09 \leq t \leq 0.2$.

Les résultats sont illustrés par les figures suivantes:

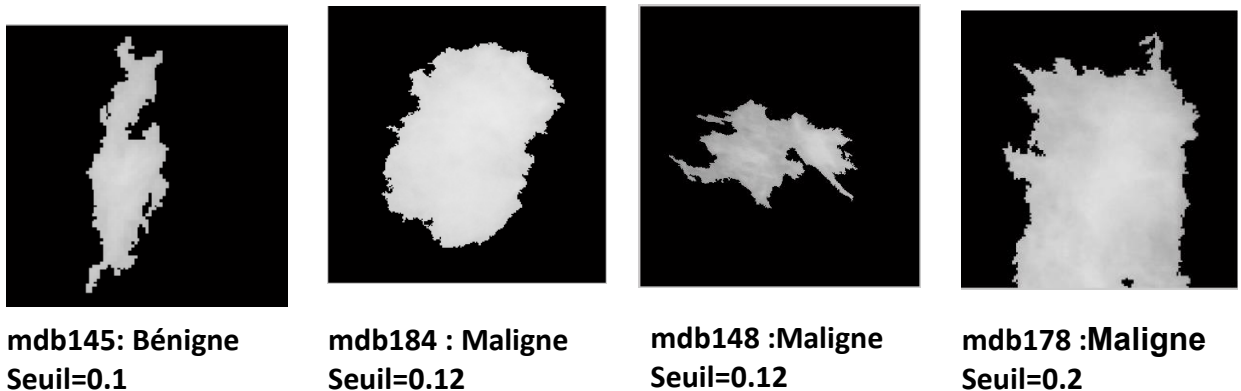


Figure 4. 10: sensibilité de la croissance de région vis à vis d'un seuil sur quatre cas de masses

Les figures suivantes montrent le résultat de la segmentation pour trois seuils différents et le même germe. Les structures segmentées diffèrent suivant le seuil.

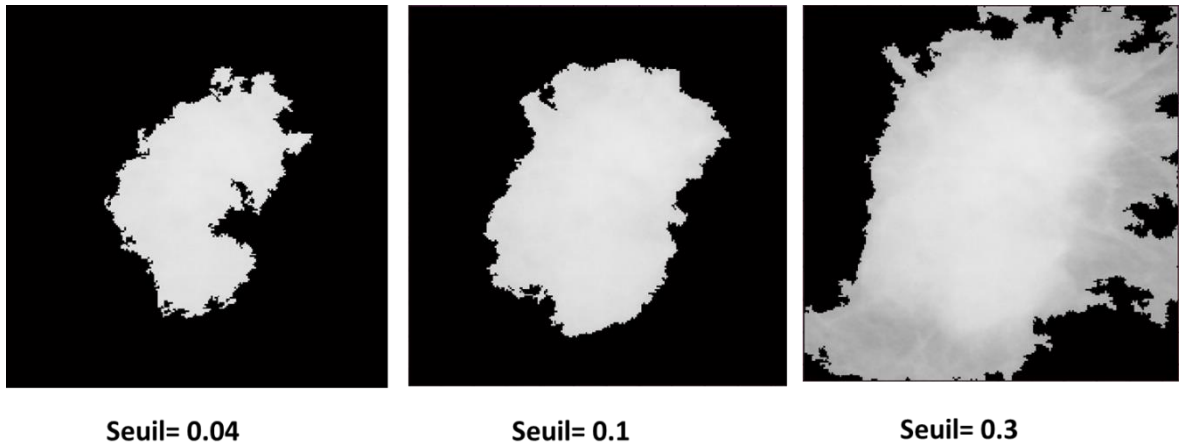


Figure 4. 11: sensibilité de la croissance de région vis à vis d'un seuil pour l'image : **mdb184 : Maligne**

4.3.3. Extraction des caractéristiques

Nos données sont issues de la base mammographique MIAS, réparties sur deux classes; des masses malignes et des masses bénignes. Nous avons analysé un ensemble de ces images (les deux tiers $\frac{2}{3}$) pour extraire les vecteurs caractéristiques, qui sont constitués par cinq paramètres texturaux et cinq paramètres topologiques (cf. Chapitre III). Nous avons ainsi, effectué la phase d'apprentissage, afin de concevoir la base d'apprentissage, pour la classification de nouvelles données. Des tests sont ensuite effectués pour la prise de décision. Le tableau ci-contre, représente les paramètres

Caractéristiques de la lésion	
Forme	Texture
CONVEXITÉ	HOMOGENÉITÉ
0.826211	0.692222
COMPACTITÉ	VARIANCE
0.000118363	156041
SUPERFICIE	CONTRASTE
24921	948.398
EXCENTRICITÉ	CORRÉLATION
0.683846	0.0798464
RECTANGULARITE	ÉNERGIE
0.72398	0.0798464

Tableau 4. 1:caractéristiques de l'image mdb135.

spécifiques (vecteurs caractéristiques), topologique et tissulaire de l'image mammographique« mdb315 » tel qu'ils apparaissent sur l'interface graphique.

Ces caractéristiques ont été calculées en ne prenant en compte que les parties expertisées par les radiologues experts, pour ne laisser que la partie suspecte afin que notre base d'apprentissage soit la plus précise possible. Ce qui, augmente les chances, et le taux de détection des cancers.

L'extraction de paramètres topologiques et texturaux des régions étiquetées, nous permet de constituer des vecteurs caractéristiques de chacune des formes étudiées, en vue d'une classification par les SVM, pour une prise de décision.

La mesure de ces descripteurs, nous permet de réaliser des graphes statistiques, afin de voir leur influence sur les différentes masses étudiées.

À cet effet, nous avons illustré dans les figures 4.12 et 4.13 la variation de l'homogénéité, de l'excentricité, de la convexité et de la rectangularité, suivant des masses bénignes et malignes, choisies parmi les tableaux 4.1 et 4.2. Dans les figures 4.14 et 4.15, nous avons illustré la variation du contraste, de la variance et de l'énergie, séparément, suivant quelques cas de masses choisis (dans les tableaux 4.1 et 4.2).

Caractéristiques de quelques masses bénignes

Convexité	Compacité	Superficie	Excentricité	rectangularité	homogénéité	Variance	contraste	corrélation	Énergie	Image
0,70021	0,00011649	30735	0,76492	0,67503	0,90096	42329	394,1	0,97697	0,63734	mdb001
0,64688	0,0002632	10068	0,83551	0,63408	0,73703	1,12E+05	1577,5	0,94203	0,21261	mdb002
0,93949	0,00050363	2484	0,66403	0,77097	0,59364	1,24E+05	1109,3	0,94681	0,098461	mdb005
0,86508	0,00062047	2071	0,42494	0,71715	0,55505	1,61E+05	1612,6	0,92381	0,062189	mdb005
0,87433	0,00053687	2442	0,7274	0,762	0,67596	1,18E+05	1759,8	0,93493	0,18736	mdb010
0,88028	0,00047435	2978	0,7094	0,757	0,72277	99849	1461,6	0,94703	0,28006	mdb012

Tableau 4. 2 : caractéristiques de quelques masses bénignes.

Caractéristiques de quelques masses malignes

Convexité	Compacité	Superficie	Excentricité	rectangularité	homogénéité	Variance	contraste	corrélation	Énergie	Image
0,81266	0,00064023	2156	0,33288	0,73087	0,61381	1,38E+05	2952,5	0,88763	0,11864	mdb023
0,90913	0,00025344	6513	0,52854	0,76899	0,74845	1,16E+05	1092	0,96175	0,22772	mdb028
0,85032	0,0006249	2011	0,62766	0,71251	0,59858	1,45E+05	1765,1	0,92716	0,097279	mdb058
0,87576	0,00078629	1586	0,83225	0,72901	0,71836	1,11E+05	2461,1	0,91414	0,23498	mdb072
0,78683	0,0014185	705	0,8415	0,71353	0,76113	69905	2531,4	0,89641	0,42179	mdb075
0,72298	0,00036698	5705	0,71953	0,66076	0,66812	1,25E+05	2076,8	0,92157	0,15619	mdb090

Tableau 4. 3: caractéristiques de quelques masses malignes.

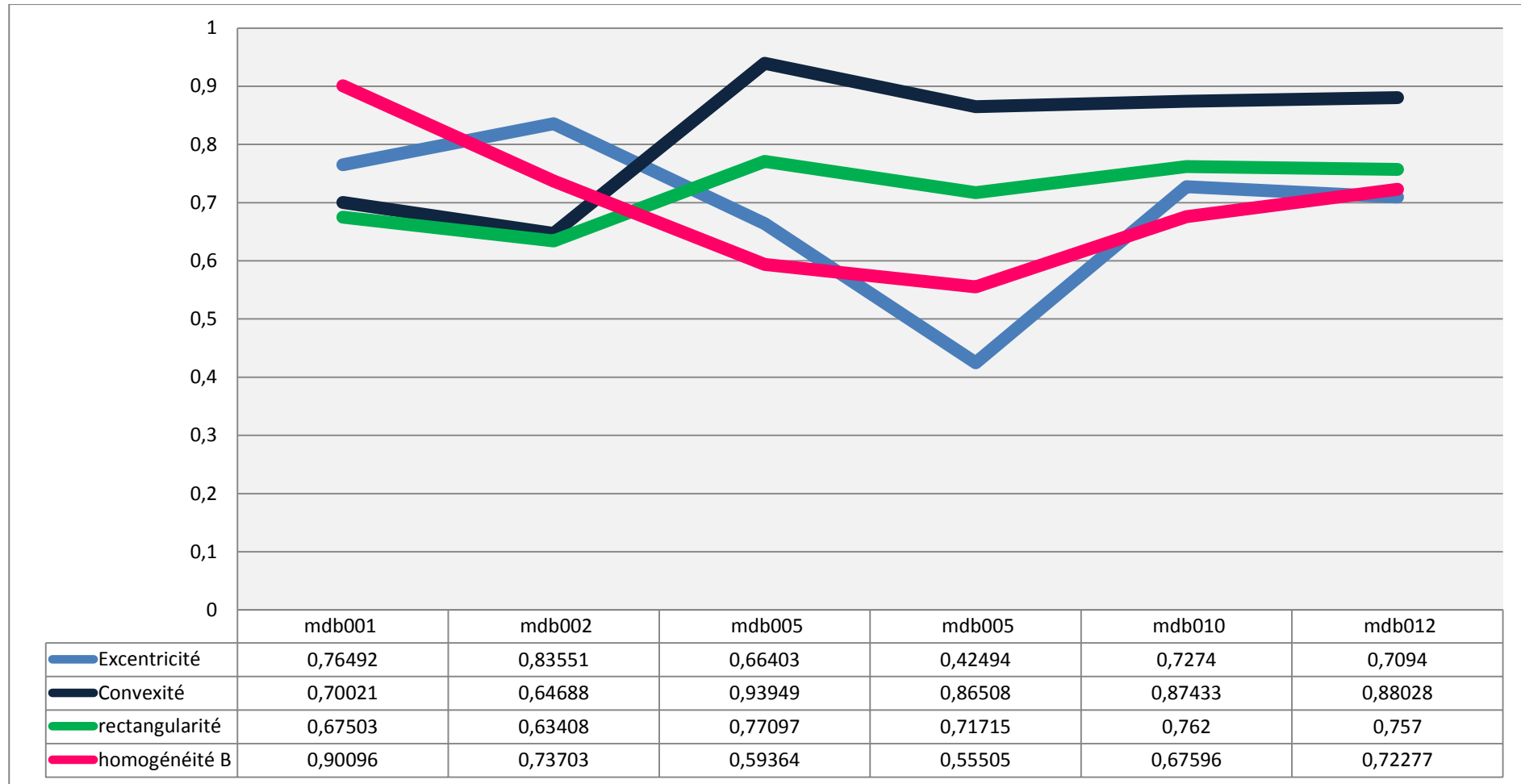


Figure 4. 12: influence des caractéristiques sur des masses bénignes.

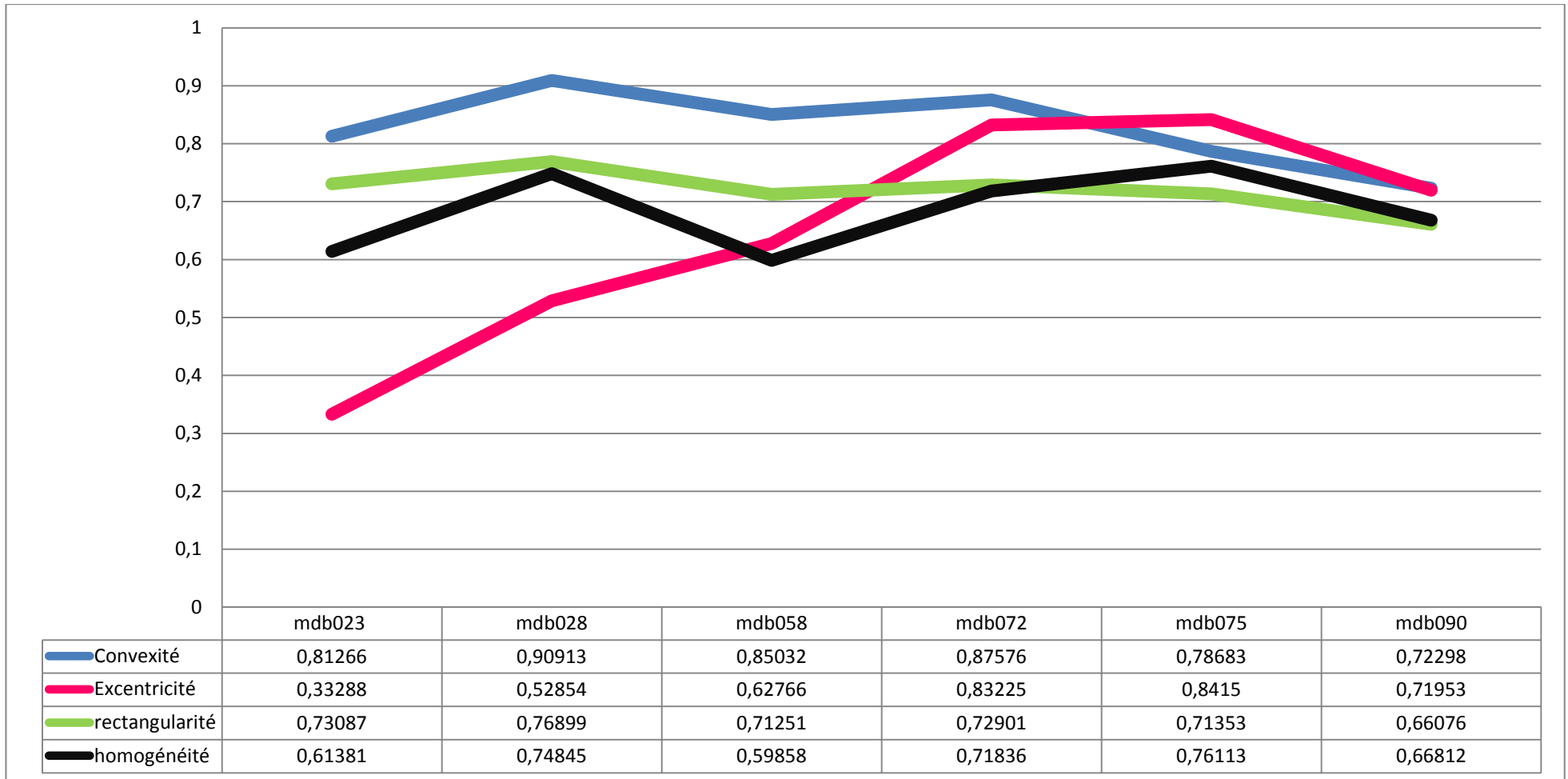


Figure 4. 13: influence des caractéristiques sur des masses malignes.

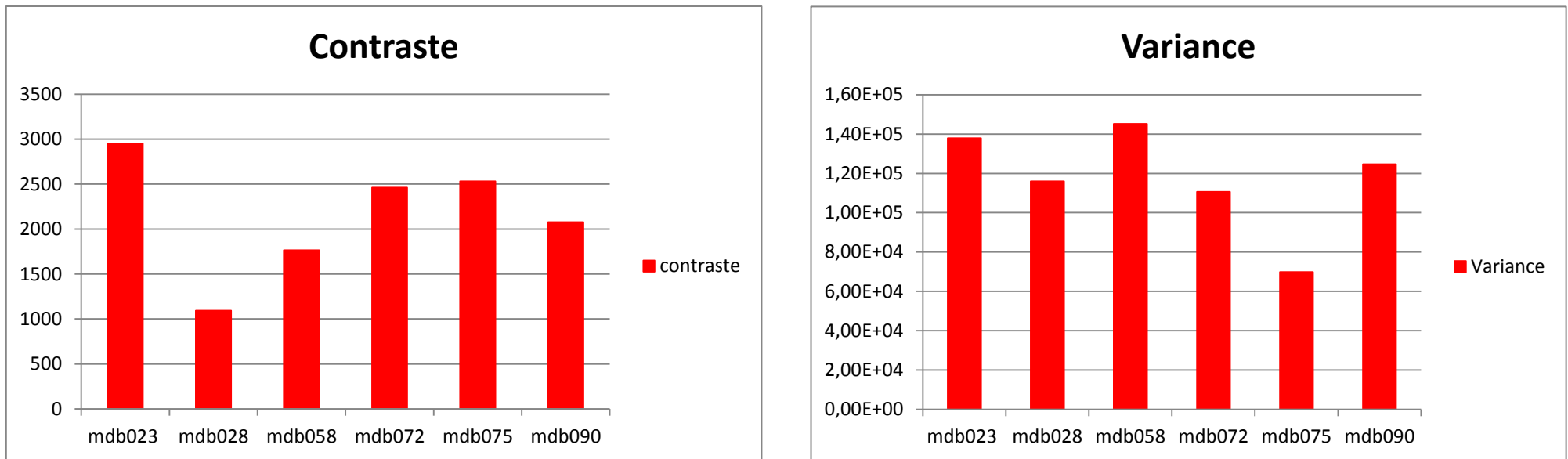


Figure 4. 14 : influence du contraste et de la variance sur des masses malignes.

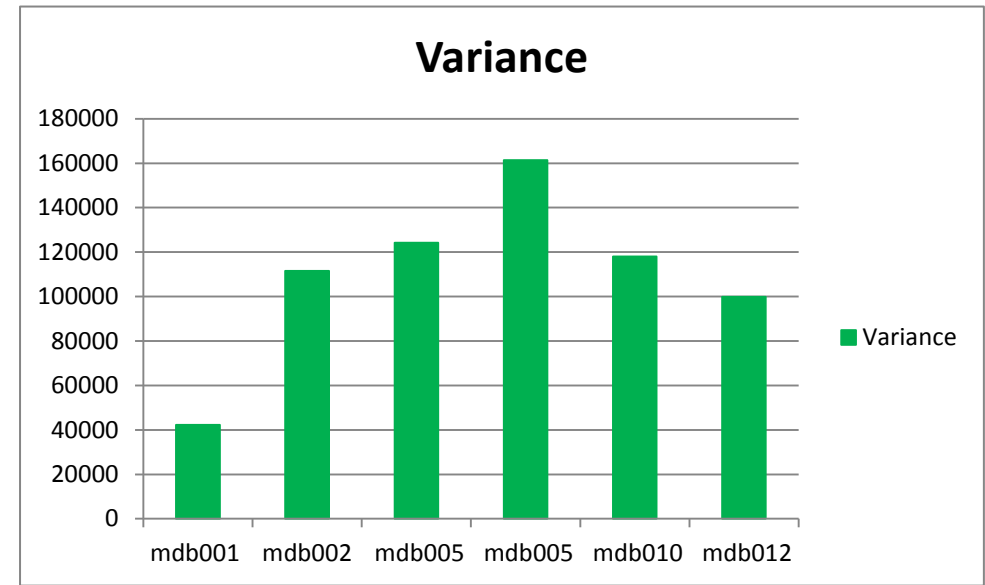
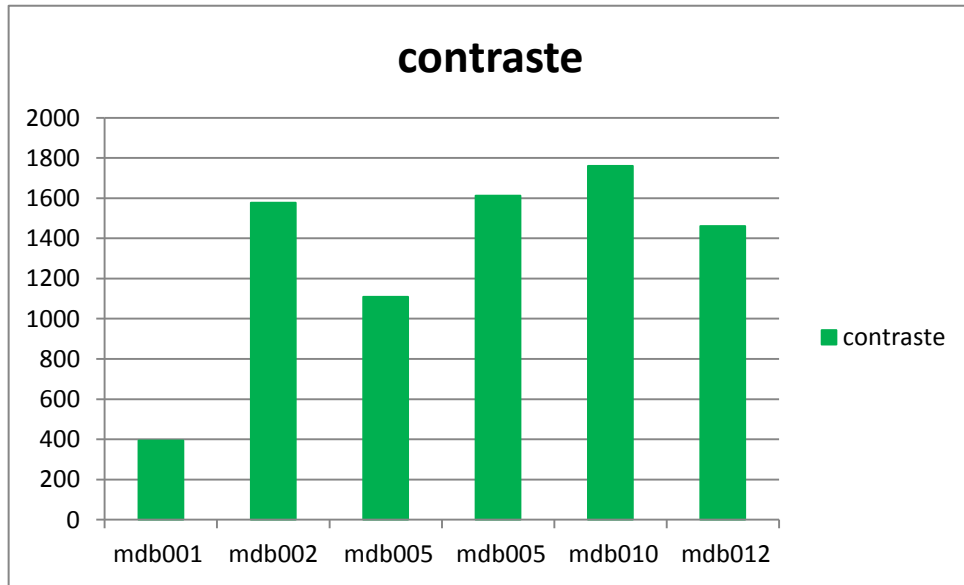
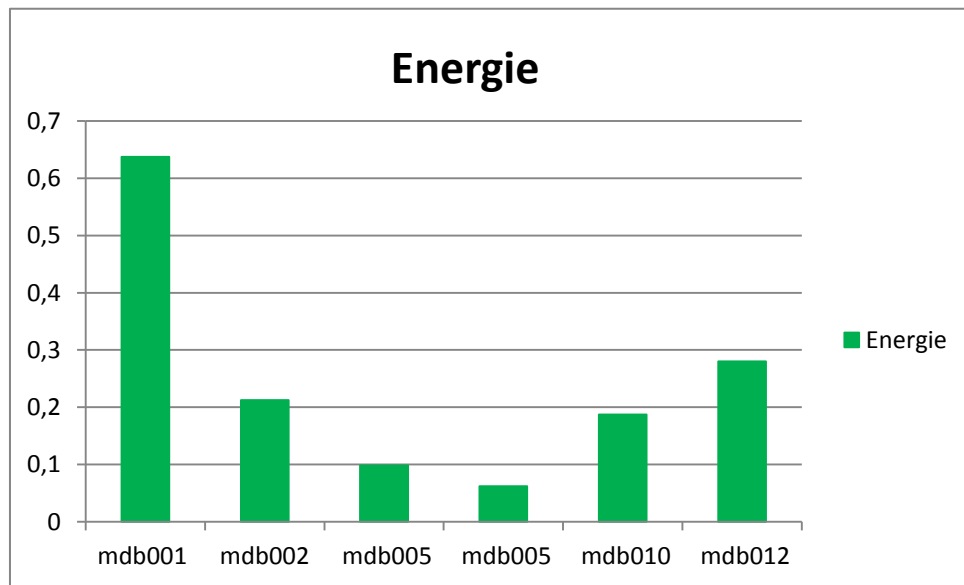
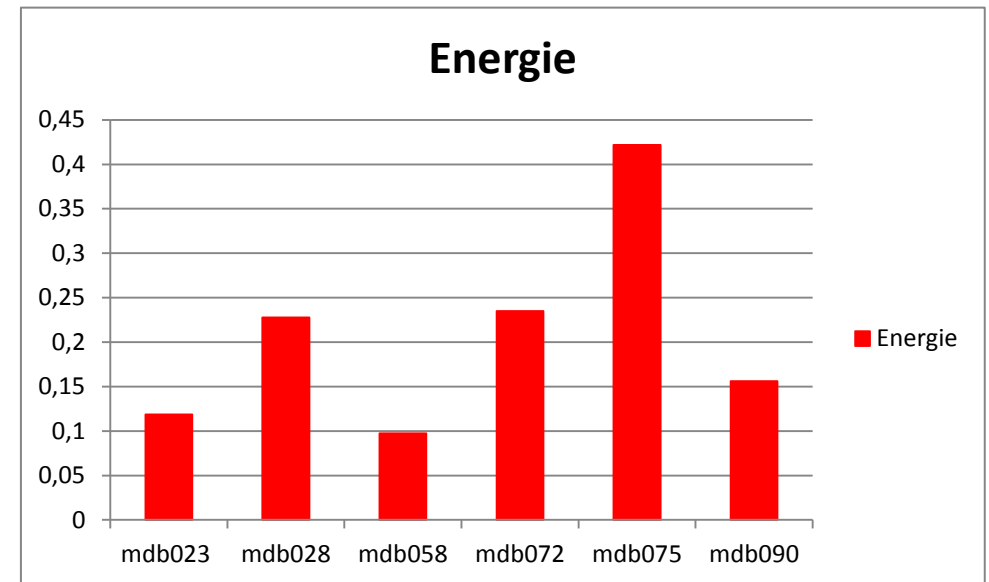


Figure 4. 15: influence du contraste et de la variance sur des masses bénignes.



(a)



(b)

Figure 4. 16 : influence de l'énergie sur des masses malignes et bénignes.

(a) : cas de masses bénignes, (b) : cas de masses malignes

Dans les graphes 4.12 et 4.13, la rectangularité influe sur le type de masse. Dans le cas bénin, elle est presque linéaire alors que, dans le cas malin, nous constatons des différences entre les différentes masses étudiées.

On remarque sur ces graphes (Figures 4.16 (b) et 4.16 (b)), la variation de la variance et du contraste, suivant le type de masse. Dans le cas, par exemple de l'image mdb 075, on voit une variance plus faible avec, une surface de la lésion moins élevée. Dans le même exemple, l'énergie est alors plus importante, pour la masse dont la compacité, est plus élevée.

4.3.4 Classification

Par l'étape de classification nous donnerons une réponse finale sur ce qui a été détecté et considéré comme un signe de lésion. L'idée, est de combiner les informations extraites précédemment, pour obtenir une décision.

Nous avons considéré un classifieur SVM. L'apprentissage a donc été accompli en considérant les machines à support de vecteurs binaires, pour constituer une base de connaissances qui considère deux types de masses (malignes et bénignes).

En prenant une image test sur laquelle, on appliquera l'intégralité des algorithmes (prétraitement, détection, segmentation, caractérisation), le classifieur est testé en lui donnant seulement les valeurs d'entrée mais pas les sorties désirées, et en regardant si le résultat obtenu, est proche du résultat désiré. La comparaison entre l'image test et, les images de la base de référence, est effectuée par une mesure de la similarité.

Le classifieur consacré aux masses (**spiculées, distorsions architecturales**), se chargera de les classer en deux groupes :

1. Masses malignes
2. Masses bénignes

La figure ci-dessous, illustre un exemple de classification d'une masse maligne et d'une masse bénigne.

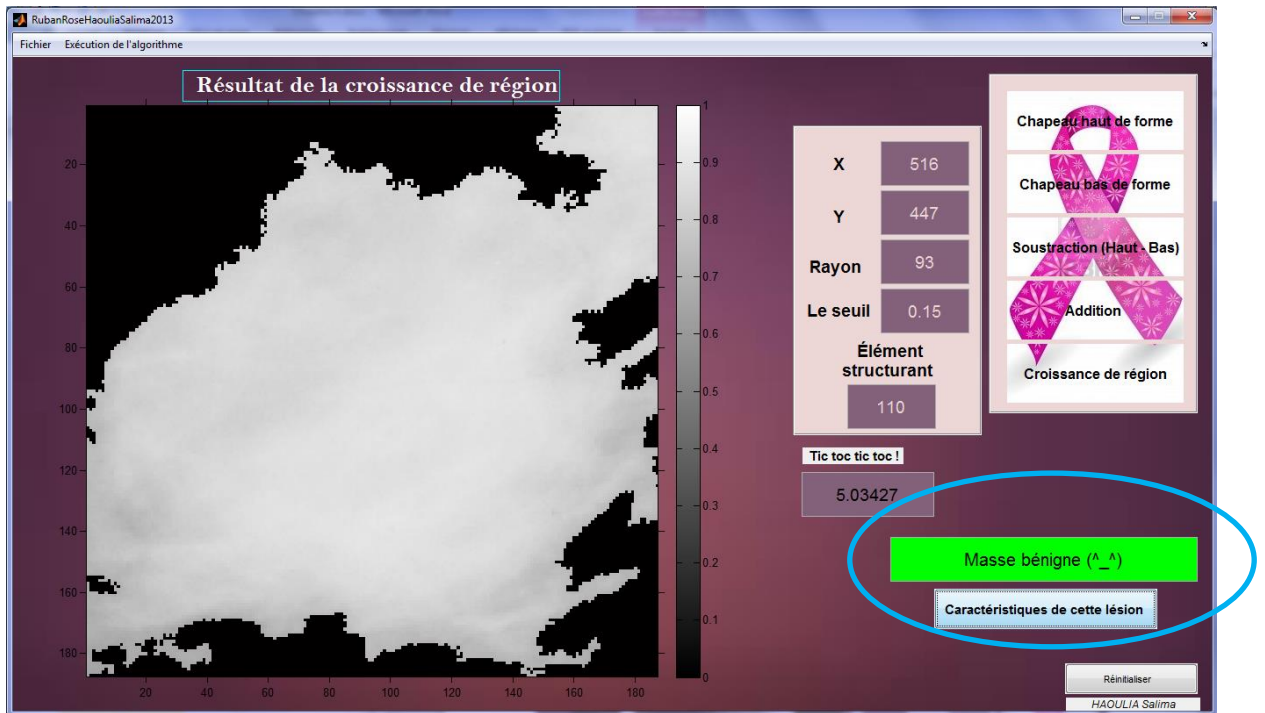


Figure 4. 17: Classification d'une masse bénigne

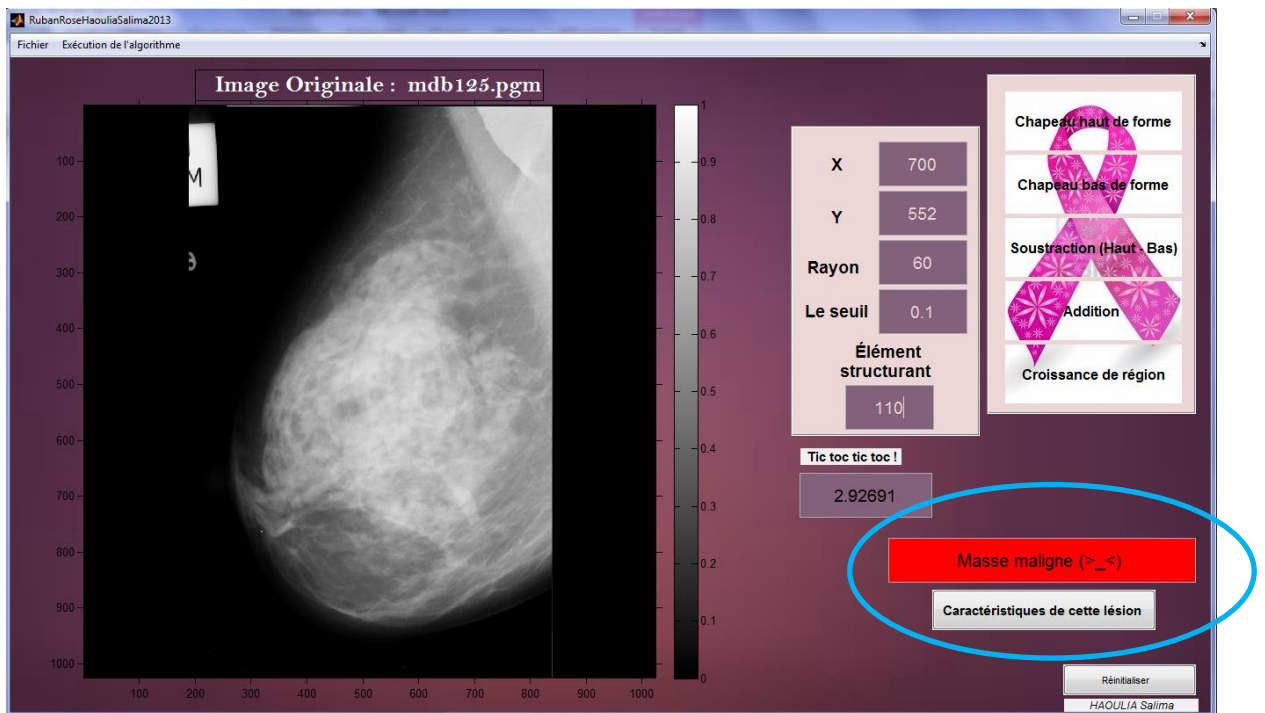


Figure 4. 18: classification d'une masse maligne

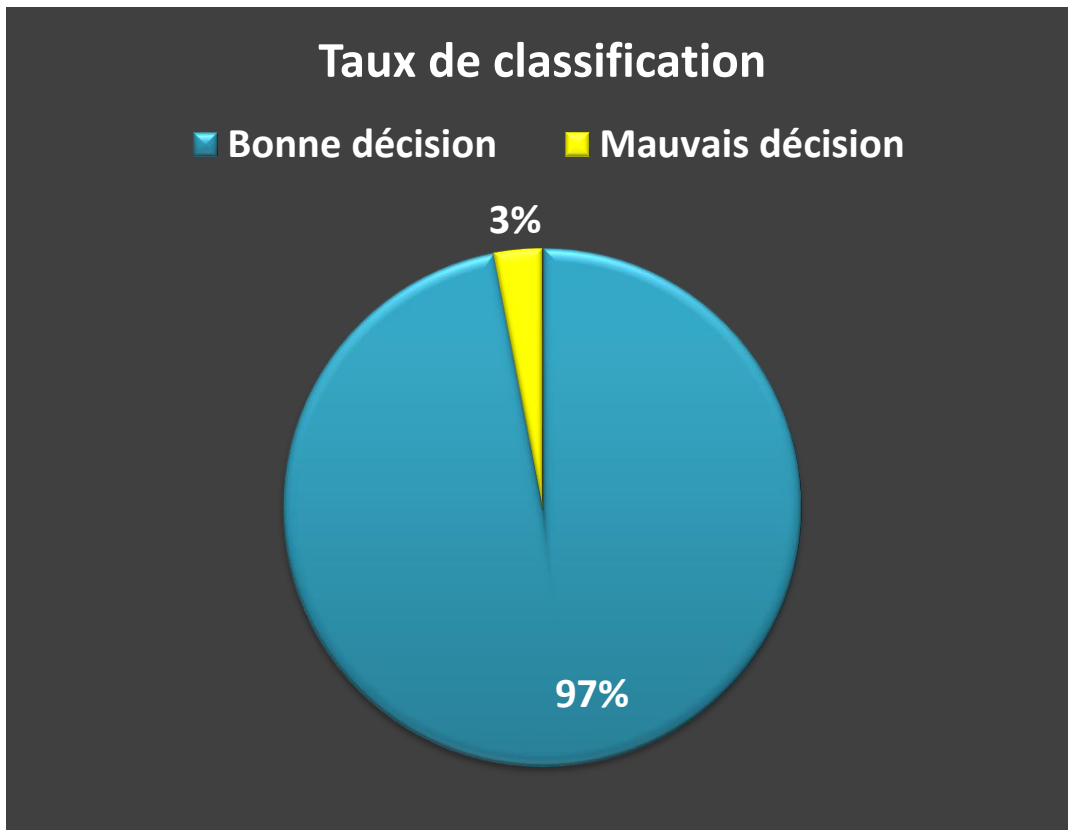


Figure 4. 19 : taux de classification.

- Nous avons pris 58 images pour la bases d'apprentissage dont 29 masses malignes et 29 masses bénignes (soit les $\frac{2}{3}$) et nous avons réservé 29 images pour le test (soit le $\frac{1}{3}$).
- Le système développé, a engendré un taux de classification de 97% comme l'illustre la figure 4.19.

Une base d'apprentissage plus grande donnerait un taux plus conséquent.

4.4 Conclusion

En dépit des efforts faits par les chercheurs, l'automatisation de détection des pathologies mammaires demeure toujours difficile. Ceci est dû principalement aux défauts d'acquisitions, la variété des types de cancers (microcalcifications, masses) et la variabilité des tissus d'une patiente à une autre. Tous ces paramètres et, bien d'autre, forment un obstacle devant les chercheurs, qui veulent atteindre un taux de réussite élevé. Notre méthodologie est fondée, sur une classification supervisée et l'apprentissage à partir d'exemples annotés par le radiologue. Par conséquent, le taux de détection des cancers est donc davantage concis, grâce à l'aide clinique.

Cette étude a confirmé la performance des **SVM**, dans le domaine de la reconnaissance des pathologies mammaires, suivant un choix judicieux des attributs de caractérisation.

Bibliographie

- [1] Medical Image/Video Engineering Group, MIAS: Mammographic Image Analysis Society, [En ligne]. Available: <http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html>.

Chapitre 3

Caractérisation et classification en imagerie mammaire

La prise de décision se fait généralement, après une étape de caractérisation. On fait des mesures, à partir des résultats de segmentation, dans le but de prendre dans un second temps, une décision à l'aide de méthodes de classification.

L'extraction de caractéristiques, focalise deux indices de bas niveaux à savoir la texture et la forme. Dans ce chapitre, nous étudions les principales caractéristiques usées dans le cadre de notre projet. À cet effet, cinq paramètres caractérisant les textures, sont extraits à partir de la matrice de cooccurrence: l'homogénéité, la variance, le contraste, la corrélation et l'énergie.

La description de la forme se fait suivant cinq attributs pertinents données par; la convexité, la circularité, l'excentricité, superficie et enfin la rectangularité.

Le chapitre 3, introduit également les concepts fondamentaux du Machine Learning supervisé dans le cas, de la classification. Nous nous intéressons aux machines à vecteurs de support (SVM), car elles limitent le risque de sur apprentissage du fait, de leur capacité de régularisation (ce risque étant particulièrement important lorsque le nombre de caractéristiques, c'est-à-dire la dimension, est grand face au nombre de données).

3.1 Extraction de caractéristiques

L'analyse de l'image à travers ses attributs (**niveaux de gris ou texture,...**) consiste souvent, à extraire un certain nombre de propriétés caractéristiques et, à les exprimer sous forme paramétrique. Ces derniers, permettent de décrire, de caractériser, d'analyser et de segmenter les images en question [32].

3.1.1 Les caractéristiques de texture

L'analyse des textures joue un rôle important, dans l'interprétation automatique des scènes. Un tel problème, est souvent rencontré dans divers domaines tels que l'imagerie médicale, le contrôle de qualité, la télédétection, etc. En effet, l'utilisation seule des niveaux de gris, ne semble pas être en mesure, de conduire à une bonne analyse de certaines images. Beaucoup d'images possèdent, des zones texturées non homogènes au sens de niveaux de gris, d'où la nécessité d'introduire l'information texturale. L'étude des propriétés d'une image, se fait alors par le traitement numérique de celle-ci via l'information texturale. Cette dernière montre que la distribution spatiale et l'organisation des niveaux de gris, d'une image, comportent beaucoup d'informations, qui méritent d'être formaliser par des paramètres quantitatifs, servant à l'identification de la texture [32]. Plusieurs méthodes de caractérisation de la texture des images, ont été ainsi proposées dans la littérature et, qu'on peut regrouper essentiellement en trois catégories:

1) Méthodes statistiques

- Matrice de cooccurrence ;
- Longueurs de plage.

2) Méthodes à base de modèle

- Décomposition de Wold ;
- Modèles Fractals ;
- Modèles AR (Autoregressive Models).

3) Méthodes fréquentielles (traitement du signal)

- Fourier ;
- Gabor ;
- Ondelettes.

Nous cherchons à avoir une représentation statistique des différentes textures des pathologies mammographiques, en utilisant une méthode statistique du deuxième ordre, appelée **Matrice de cooccurrence**. Les matrices de cooccurrence, seront complètement détaillées car elles serviront de référence pour la deuxième partie de ce mémoire, concernant la construction de paramètres destinés, à la discrimination de texture des différentes régions d'intérêts.

3.1.1.1 Notion de texture

La notion de texture, a toujours été rattachée à la notion de perception. Elle renvoie à l'apparence et à la consistance d'un objet. Plus précisément, une texture décrit la structure, l'agencement géométrique de celui-ci. Elle se manifeste par une information visuelle, et se décrit par des termes linguistiques qualitatifs comme la finesse, la granularité, la régularité, le contraste... Les chercheurs se sont efforcés, de formaliser les différents aspects visuels, par des paramètres mathématiques, pour obtenir des descripteurs quantitatifs [2] [32].

A la lumière de toutes les études menées, il semble clair qu'il n'existe pas de définition universellement admise, qui permette de caractériser la notion de texture au-delà du fait qu'elle restitue l'état de surface. En effet, toutes les définitions proposées sont, soit trop générales et imprécises, soit trop restrictives pour s'adapter à la diversité des cas rencontrés. Cependant nous retrouvons toujours un critère important, en commun. Il représente la notion d'arrangement spatial, des pixels dans une image. La texture est ainsi définie et, étudiée de manières différentes, en fonction de la façon dont on la caractérise. Ainsi, on trouve dans la littérature des définitions telles que:

- une texture est une région d'une image, pour laquelle il existe une fenêtre de dimension réduite, telle qu'une observation au travers de celle-ci, se traduise par une impression visuelle identique pour toutes les positions envisageables par translation, à l'intérieur de la région considérée [3].
- La texture peut être décrite, comme une structure hiérarchique à deux niveaux. Le premier concerne, la description d'éléments de base ou primitives, à partir

desquels est construite la texture. Le second niveau, est relatif à la description de l'organisation spatiale de ces primitives, à l'agencement des motifs texturaux entre eux, qui peut satisfaire aussi bien à des lois stochastiques que déterministe [4]. D'après **Haralick** [4][2] pour des détails fins (à petite échelle) on observe, souvent un objet élémentaire qui constitue la base de la texture; qui peut être régulier ou variable et, qui a tendance naturellement à disparaître, lorsqu'on observe le champ global de la texture (**une haute précision en espace et en fréquence ne peut être atteinte simultanément**). Pour une analyse grossière (à **plus grande échelle**), la texture apparaît comme une juxtaposition plus au moins régulière de motifs. Cette juxtaposition possède, ses propres lois statistiques (**corrélation, densité de puissance**) qui se combinent à celles, du motif de base. Une bonne analyse de texture, donnera des informations sur ces deux composantes [2] [32].

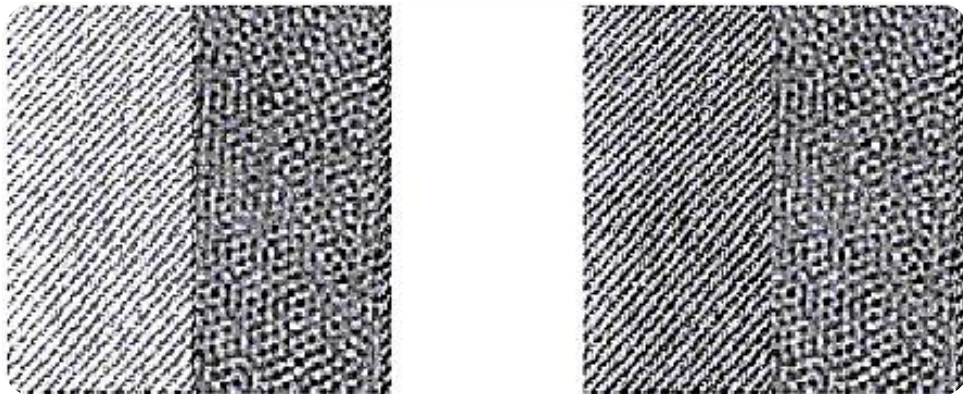


Figure 3. 1 : information texture et information couleur

La forme et la couleur, sont deux autres caractéristiques très importantes. Pour faire la distinction entre une orange et une cerise, les informations forme et couleur sont certainement plus significatives, que l'information texture. Mais l'information texture, devient primordiale lorsqu'on veut faire, la distinction entre deux zones d'une image de même couleur (ou bien de même niveau de gris). La figure 3.1 illustre, ce phénomène. Dans la première image, il est tout à fait possible de faire la différence entre les deux zones de l'image, simplement avec l'information couleur fournie par les niveaux de gris. En effet, la zone de gauche est globalement d'intensité lumineuse beaucoup plus forte que, celle de droite. Par contre, dans la deuxième image, il n'est pas possible de distinguer les deux zones à l'aide de leurs valeurs de gris. Dans ce cas,

c'est l'information texture qui prédomine. La zone de gauche, est caractérisée par une texture orientée, contrairement à celle de droite qui ressemble plus à une mosaïque [2].

3.1.1.2 Les matrices de cooccurrence

Du fait de leur richesse en information de texture, les matrices de cooccurrence sont devenues les plus connues et les plus utilisées, pour extraire ces caractéristiques de textures [2] [4].

Elles estiment des propriétés des images, relatives à des statistiques de second ordre. Une matrice de cooccurrence, mesure la probabilité d'apparition des paires de valeurs de pixels situés à une certaine distance dans l'image. Elle est basée, sur le calcul de la probabilité $P(i, j, \delta, \theta)$ qui représente le nombre de fois où un pixel de niveau de couleur i apparaît à une distance relative δ d'un pixel de niveau de couleur j et, selon une orientation θ donnée [4] [2].

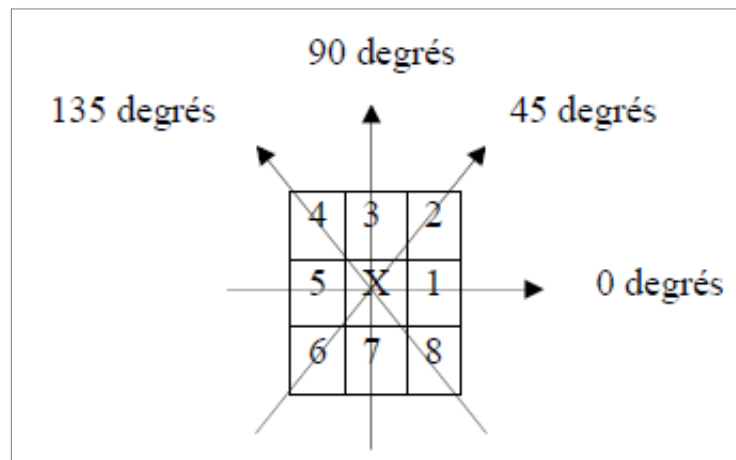


Figure 3. 2 : plus proches voisins du pixel « X » selon 4 directions[2].

Les directions angulaires θ classiquement utilisées sont **0, 45, 90** et **135** degrés. Les relations de voisinage entre pixels, nécessaires au calcul des matrices, sont illustrées dans la **figure 3.3**; par exemple, les plus proches voisins de 'X' selon la direction $\theta = 135$ degrés, sont les pixels **4** et **8**[2].

Les caractéristiques extraites à partir de ces matrices, contiennent des informations notamment sur l'homogénéité, les dépendances linéaires entre les niveaux de gris, le contraste et la complexité de cette image [2].

La **figure 3.3**, montre un exemple de calcul des $P(i, j)$ à partir d'une petite image 4×4 composée de quatre niveaux de gris (**0, 1, 2, 3**). Cet exemple se limite au cas $\delta = 1$ et $\theta = 0$. L'élément (**2, 3**) de la matrice $P(1, 0)$ est égal à **4**, cela signifie qu'il existe **4** configurations dans l'image où un pixel de niveau de gris 2, est séparé horizontalement d'un autre pixel de niveau de gris 3 par une distance 1. Ces configurations, sont représentées en trait gris sur l'image [2].

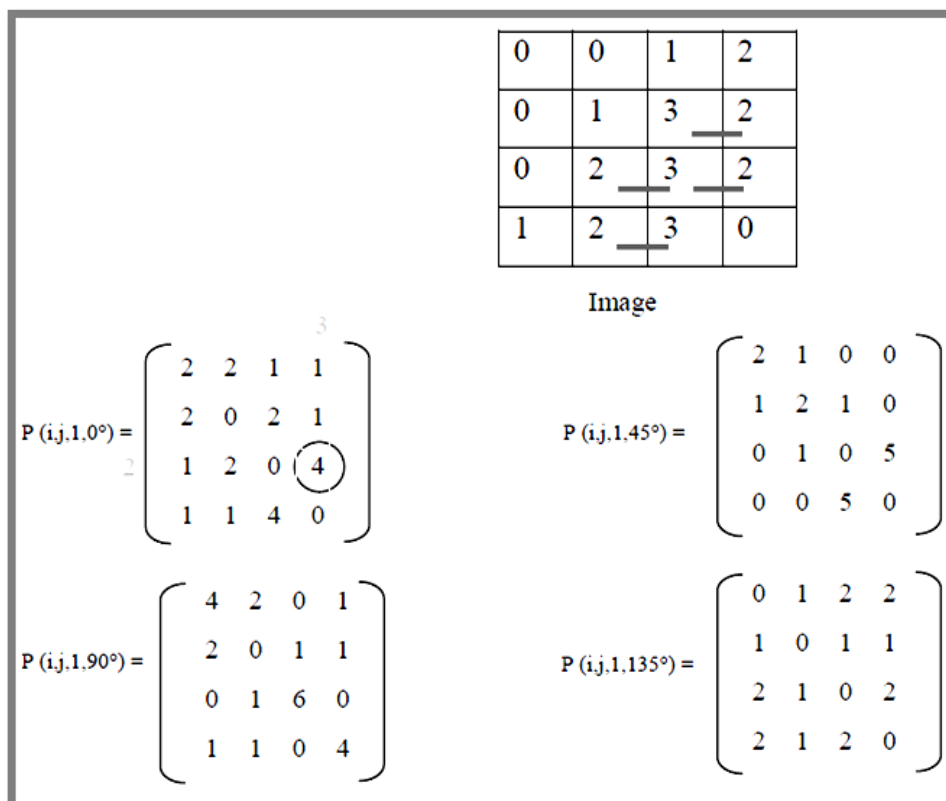


Figure 3. 3: exemple de matrices de cooccurrence construites à partir d'une image 4×4 composée de 4 niveaux de gris[2].

La plupart des images sont codées sur 256 niveaux de gris, par conséquent, la taille des matrices de cooccurrence est de 256×256 . On s'aperçoit ainsi que ces matrices comptabilisent une très grosse quantité d'informations, difficile à exploiter directement. C'est pourquoi, un certain nombre d'auteurs [4] [2] ont essayé d'extraire de l'information de ces matrices, afin de mettre en évidence la structure des textures.

Mais c'est Haralick et al. [4] qui ont proposé les premiers **14** paramètres, caractérisant les textures, issus de ces matrices. Nous présentons **6** paramètres considérés, comme étant les plus utilisés et les plus pertinents [2] :

■ L'énergie

Ce paramètre (**équation 3.1**) mesure l'uniformité de la texture. Il atteint, de fortes valeurs lorsque la distribution des niveaux de gris, est constante ou de forme périodique. Dans ce dernier cas, les valeurs élevées d'énergie sont obtenues pour les matrices $P(\delta, \theta)$ lorsque (δ, θ) correspond à la période [2].

$$ENE = \sum_i \sum_j (P_{ij}(\delta, \theta))^2 \quad 3.1$$

■ Le contraste

La valeur en est d'autant plus élevée, que la texture présente un fort contraste. Ce paramètre (**équation 3.2**) est fortement non corrélé, à l'énergie [2].

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

■ L'entropie

Ce paramètre (**équation 3.3**) mesure le désordre dans l'image. Contrairement à l'énergie, l'entropie atteint de fortes valeurs lorsque la texture, est complètement aléatoire (sans structure apparente). Elle est fortement corrélée (par l'inverse) à l'énergie [2].

$$ENT = - \sum_i \sum_j (\log P_{ij}(\delta, \theta) P_{ij}(\delta, \theta)) \quad 3.3$$

■ La variance

La variance (**équation 3.4**), mesure l'hétérogénéité de la texture. Elle augmente lorsque les niveaux de gris, différent de leur moyenne. La variance est indépendante du contraste [2].

$$VAR = \sum_i \sum_j ((i - \mu)^2 P_{ij}(\delta, \theta)) \quad 3.4$$

■ La corrélation

COR (équation 3.5) mesure la dépendance linéaire (**relativement à** (δ, θ)) des niveaux de gris de l'image. La corrélation n'est corrélée ni à l'énergie, ni à l'entropie [2].

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

■ Le moment inverse

Le moment inverse IDM (Inverse Difference Moment) mesure l'homogénéité de l'image. Ce paramètre (**équation 3.6**) est corrélé à une combinaison linéaire des variables ENE et CST [2].

$$IDM = \sum_i \sum_j \frac{P_{ij}(\delta, \theta)}{1+(i-j)^2} \quad 3.6$$

Pour l'analyse de texture nous avons opté pour les descripteurs présentés dans le tableau ci-dessous :

Descripteurs de texture	
1	HOMOGENÉITÉ
2	VARIANCE
3	CONTRASTE
4	CORRÉLATION
5	ÉNERGIE

Tableau 3. 1 : sélection des descripteurs texturaux.

3.1.2 Les caractéristiques de formes

L'analyse de forme est utilisée en imagerie médicale, pour la caractérisation tissulaire. Elle implique l'extraction de caractéristiques de forme, dans différentes **ROIs** saines ou pathologiques, afin de caractériser les tissus et, de les comparer entre eux. De telles méthodes, doivent pouvoir apporter des renseignements aux radiologues en termes de diagnostic, pronostic, suivi de pathologies, etc. [5]. Dans notre approche on va utiliser trois caractéristiques basées sur la forme:

Descripteurs de forme	
1	COMPACITÉ
2	EXCENTRICITÉ
3	CONVEXITÉ
4	RECTANGULAIRE
5	SUPERFICIE

Tableau 3. 2: sélection des descripteurs morphologiques.

3.1.2.1 Compacité (ou facteur de circularité)

La compacité (équation 3.7) est un facteur de circularité qui vaut 1, lorsque la région est un cercle et, diminue à mesure que le contour est très découpé ou que, la région est allongée [6].

Elle est calculée par la formule :

$$\text{Compacité} = \frac{4\pi S(R)}{P^2(R)} \quad 3.7$$

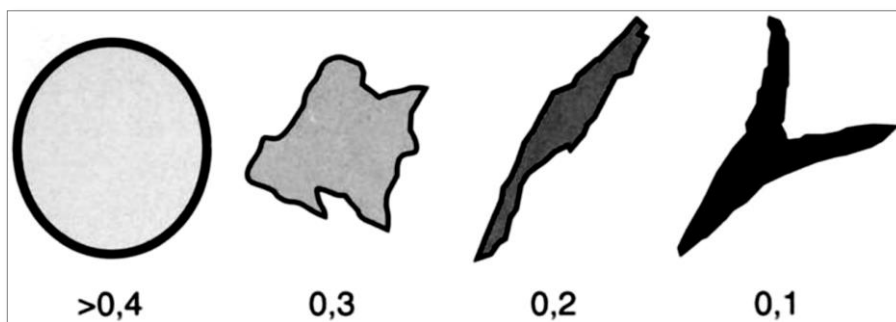


Figure 3. 4 : diverses formes et valeurs correspondant à la compacité [6].

3.1.2.2 Convexité

Elle représente, le rapport entre la surface de la région et, celle de son enveloppe convexe (Figure 3.5). Une région fortement convexe (i.e., qui épouse parfaitement son enveloppe convexe) a une valeur de convexité = 1. Une région faiblement convexe, a une valeur de convexité $\ll 1$ [7].

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

NB : Nécessite le calcul de l'enveloppe convexe et de son aire [7].

■ Enveloppe convexe

L'enveloppe convexe d'un ensemble de points, est le plus petit polygone convexe qui englobe tous les points. Elle est aussi le plus court chemin, entourant tous les points [7].

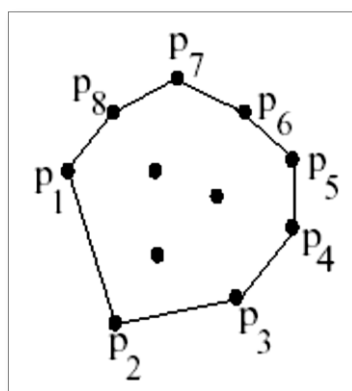


Figure 3. 5: enveloppe convexe [7].

3.1.2.3 Excentricité

L'excentricité correspond à la valeur moyenne du rapport des modules des deux rayons vecteurs pris, sur une même droite, entre le barycentre et les intersections avec le contour, le plus grand étant placé au numérateur (**figure 3.6**). Plus la forme est symétrique, plus ce paramètre est petit ; il est de **1** pour une forme présentant un centre de symétrie (**1 pour un carré**) [7] [6].

L'excentricité est calculée à partir des moments d'inertie :

$$excentricité = \frac{M_{2,0} + M_{0,2} - \sqrt{(M_{2,0} - M_{0,2})^2 + 4M_{11}^2}}{M_{2,0} + M_{0,2} + \sqrt{(M_{2,0} - M_{0,2})^2 + 4M_{11}^2}} \quad 3.9$$

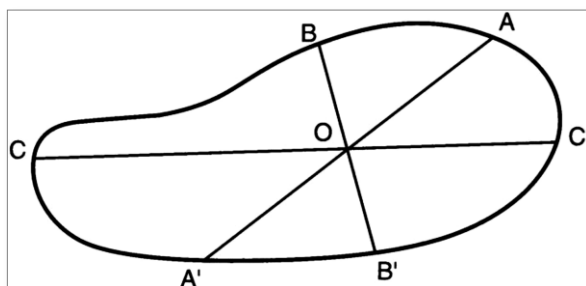
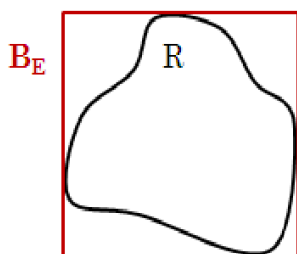


Figure 3. 6 : excentricité E définie par la valeur moyenne du rapport des deux segments délimités, sur une même droite, par le barycentre et le contour de la forme : $E = OA/OA' + OB/OB' + \dots N$ où N représente le nombre de droites prises en compte [6] [5].

3.1.2.4 La rectangularité

Un autre descripteur de forme intitulé rectangularité ou boîte englobante, est souvent



utilisé, pour décrire le taux de rectangularité et, d'allongement d'une région. On désigne par R la région étudiée et B_E la boîte englobante qui est définie, par le plus petit rectangle contenant la lésion étudiée. La mesure de rectangularité notée Rect peut alors, se définir comme suit :

$$Rect = \frac{aire(R)}{aire(B_E)} \quad 3.10$$

Selon cette formule, plus l'objet à décrire est rectangulaire, plus la valeur de rectangularité est proche de 1. Toutefois, cette formule de rectangularité ne tient pas compte de l'orientation de l'objet. En effet, la valeur de rectangularité d'un rectangle parfait et droit est égale à 1, alors que la rectangularité de ce même rectangle mais considéré incliné, est différente de 1. Ainsi, ce descripteur est sensible à la rotation [8].

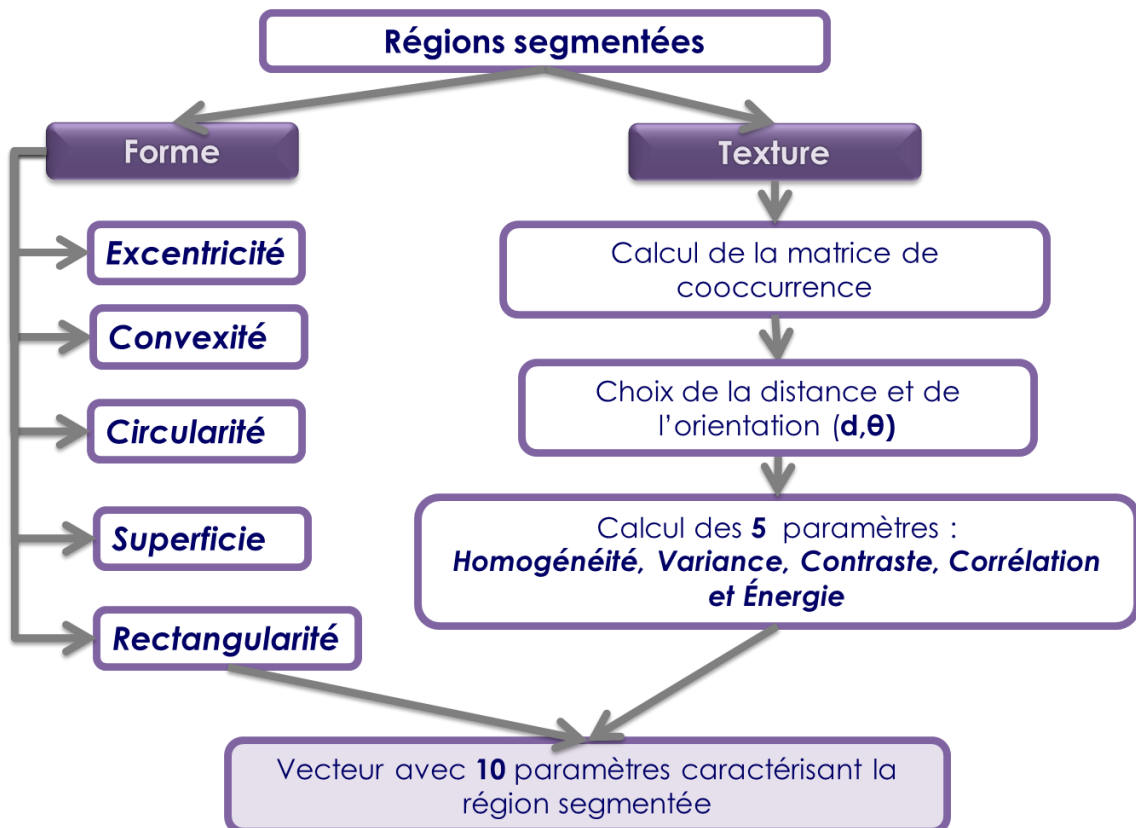


Figure 3. 7: organigramme d'extraction des caractéristiques

3.2 Classification des images

La classification, d'une manière générale, permet de partitionner un ensemble d'observations en différentes classes, en regroupant dans une même classe les observations présentant les mêmes caractéristiques. On parle de classification des images, lorsque les observations correspondent à des images et, où chaque image peut être caractérisée par un ensemble d'attributs tels que, les attributs de texture ou couleur.

La classification des images intervient comme une étape très importante dans des applications de reconnaissance des formes telles que l'identification des images de bois, la reconnaissance des visages, de la parole, des empreintes, etc..... Les méthodes de classification peuvent être divisées en deux catégories selon l'information disponible a priori, sur les observations.

Lorsque les différentes classes des observations sont connues au préalable, la classification est dite supervisée. Cette catégorie de méthodes est préférable pour la classification des images [2]. En revanche, dans le cas où l'on n'a aucune connaissance *a priori* sur l'appartenance des observations aux classes, on parle de classification non-supervisée ou **clustering**, celle-ci est plus adaptée pour la segmentation des images. Dans un contexte supervisé, on dispose d'échantillons (ensemble d'observations) où, chaque observation est caractérisée par un ensemble d'attributs.

L'appartenance de chaque observation de cet échantillon à une classe, est connue a priori. Dans ce cas ces observations sont appelées prototypes de la classe. La classification a pour but de définir à partir des prototypes, des modèles mathématiques pour chaque classe. Ces modèles peuvent être un simple centre de gravité, une fonction de densité de probabilité ou, une surface de séparation entre des classes. Les étapes de caractérisation et de classification des observations prototypes des classes, constituent ce qu'on appelle la phase d'apprentissage. Celle-ci se distingue de la phase de décision, qui consiste à affecter une nouvelle observation à l'une des classes définies lors de la phase d'apprentissage. Il existe une multitude de méthodes

de classification supervisée. Les plus classiques sont la classification optimale de Bayes, la règle des **K** plus proches voisins (**Kppv**) , les réseaux de neurones et les séparateurs à vaste marge (**SVM**). Dans notre projet, nous étudions la classification par les SVM.



Notre objectif : dans notre étude, l'étape de classification a pour but de donner la réponse finale sur ce qui est détecté et considéré, comme un signe de lésion. L'idée, est de combiner les informations extraites précédemment pour obtenir une décision. On peut voir cela de manière similaire à ce que fait le radiologue, pour prendre sa décision. Le problème de classification n'est pas propre à la détection de lésions dans le sein. En effet c'est un champ de recherche à part entière, qui peut être utilisé dans des applications très différentes.

3.3 Machines à vecteurs support

Les machines à vecteurs support ont été introduites en **1995** par **Cortes** et **Vapnik**[8]. Elles sont utilisées dans de nombreux problèmes d'apprentissage: reconnaissance de forme, catégorisation de texte ou encore diagnostic médical. Les **SVM** reposent sur deux notions: celle de marge maximale et, celle de fonction noyau.

Elles permettent de résoudre des problèmes, de discrimination non linéaire. La marge est la distance, entre la frontière de séparation et, les échantillons les plus proches appelés vecteurs support. Dans un problème linéairement séparable, les **SVM** trouvent une séparatrice qui maximise, cette marge. Dans le cas d'un problème non linéaire, nous utilisons une fonction noyau pour projeter les données dans un espace, de plus grande dimension où, elles seront linéairement séparables (Voir **Figure 3.8**) [9].

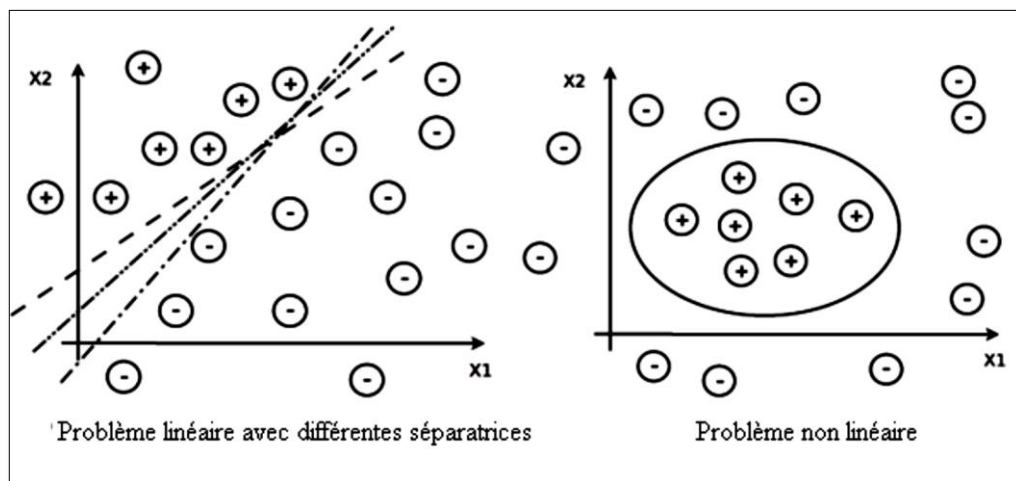


Figure 3. 8 : exemples de séparation linéaire et non linéaire [9].

3.3.1 Principe

La technique des *Séparateurs à Vaste Marge* (**SVM - Support Vector Machines**) (également appelées Machines à Vecteur de Support) repose sur le concept de plans de décision

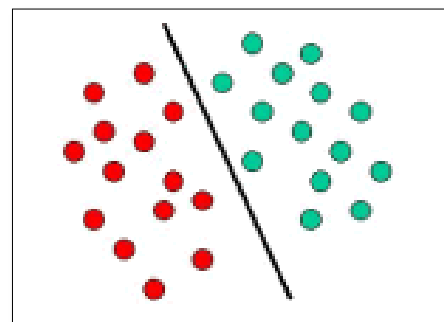


Figure 3. 9[8] : exemple de SVM

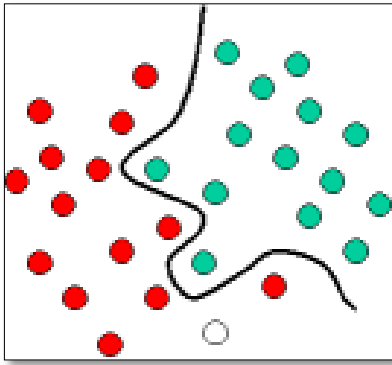


Figure 3. 10 : exemple de
SVM

ROUGE. La droite de séparation définit une frontière, à droite de laquelle, tous les objets sont **VERTS** et à gauche de laquelle tous les objets, sont **ROUGES**. Tout nouvel objet (matérialisé par un cercle blanc) apparaissant à gauche sera affecté à la classe **ROUGE** (ou au contraire affecté à la classe **VERTE** s'il se trouve à droite de la ligne de séparation) [8].

définissant, des frontières de décision. Un plan de décision, est un plan qui permet de séparer un ensemble d'objets, appartenant à des classes différentes.

L'illustration de la **figure 3.9** constitue un exemple représentatif. Dans cet exemple, les objets appartiennent soit à la classe **VERTE**, soit à la classe

L'exemple classique illustré ci-dessus, est un modèle de classification linéaire, c'est-à-dire un modèle de classification qui va répartir un ensemble d'objets dans leurs groupes respectifs (**VERT** et **ROUGE** dans le cas présent) à l'aide d'une droite. La plupart des tâches de classification, n'est malheureusement pas aussi simple et, il faut souvent recourir à des structures plus complexes pour trouver une séparation optimale, c'est-à-dire classer correctement de nouveaux objets (les observations de test) sur la base des exemples disponibles (observations d'apprentissage). C'est le cas dans l'illustration de la **figure 3.10**. Par rapport au schéma précédent, il est évident que nous avons besoin d'une courbe (qui est plus complexe qu'une droite), pour séparer totalement les objets **VERTS** des objets **ROUGES**. Les tâches de classification qui utilisent des lignes de séparation, pour faire la distinction entre des objets appartenant à des classes distinctes sont connues sous le nom de modèles de classification (ou classifieurs, ou encore classificateurs) par hyperplans de séparation. Les **SVM** sont particulièrement bien adaptés pour traiter ce type de tâches [8]. Le schéma de la **Figure 3.11** illustre le principe fondamental des **SVM**. Dans ce cas, nous pouvons voir que les objets originaux (situés à gauche du schéma) sont transformés, c'est-à-dire réorganisés, à l'aide d'un ensemble de fonctions mathématiques appelées noyaux. Le processus de réorganisation des objets, est appelé projection (transformation) [8].

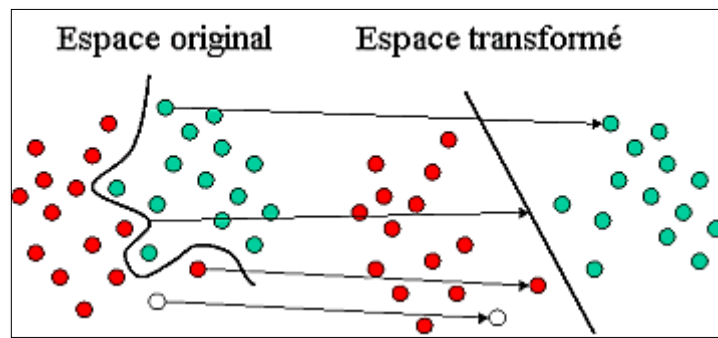



Figure 3. 11[8] : configuration explicite des SVM


 Dans cette nouvelle configuration, nous pouvons séparer les objets projetés (situés à droite du schéma) par une ligne droite et, par conséquent, au lieu de construire une courbe complexe (comme à gauche du schéma), nous pouvons nous contenter de trouver une droite, qui va séparer de façon optimale les objets **VERTS** des objets **ROUGES** [8].

3.3.2 Classification multi-classes

Les **SVM** vu précédemment effectuent une classification binaire. Dans la plupart des contextes nous cherchons à résoudre des problèmes multi-classes. Plusieurs méthodes ont été développées, elles se regroupent en deux approches : les approches parallèles et hiérarchiques. Ces dernières étant moins performantes, nous aborderons seulement les approches parallèles [9]. Les deux plus connues sont appelées, **one versus all**¹ et **one versus one** [13]².

3.3.2.1 Un contre tous

La méthode **one-versus-all** (appelée parfois **one-versus-the-rest**), consiste à construire M classifieurs binaires en attribuant le label **1**, aux échantillons de l'une des classes et le label **-1** à toutes les autres. En phase de test, le classifieur donnant la valeur de confiance (la marge) la plus élevée, remporte le vote [13].

¹One-versus-all : un contre tous.

²One-versus-one : un contre un.

3.3.2.2 Un contre un

La méthode **one-versus-one** consiste, à construire $\frac{M(M-1)}{2}$ classifieur binaires en confrontant chacune des M classes. En phase de test, l'échantillon à classer, est analysé par chaque classifieur et, un vote majoritaire permet de déterminer sa classe [13].

3.4 Prise de décision

Après la segmentation, vient l'étape de prise de décision. Celle-ci est très conditionnée, par les résultats de la première étape (segmentation). La **figure 3.15** illustre l'algorithme final de prise de décision, qui se compose d'une étape d'extraction de caractéristiques suivie d'une étape de classification supervisée: les machines à vaste marge (**SVM**).

Nous avons considéré un classifieur binaire en, utilisant la fonction prédéfinie de MATLAB « SVMTrain ».

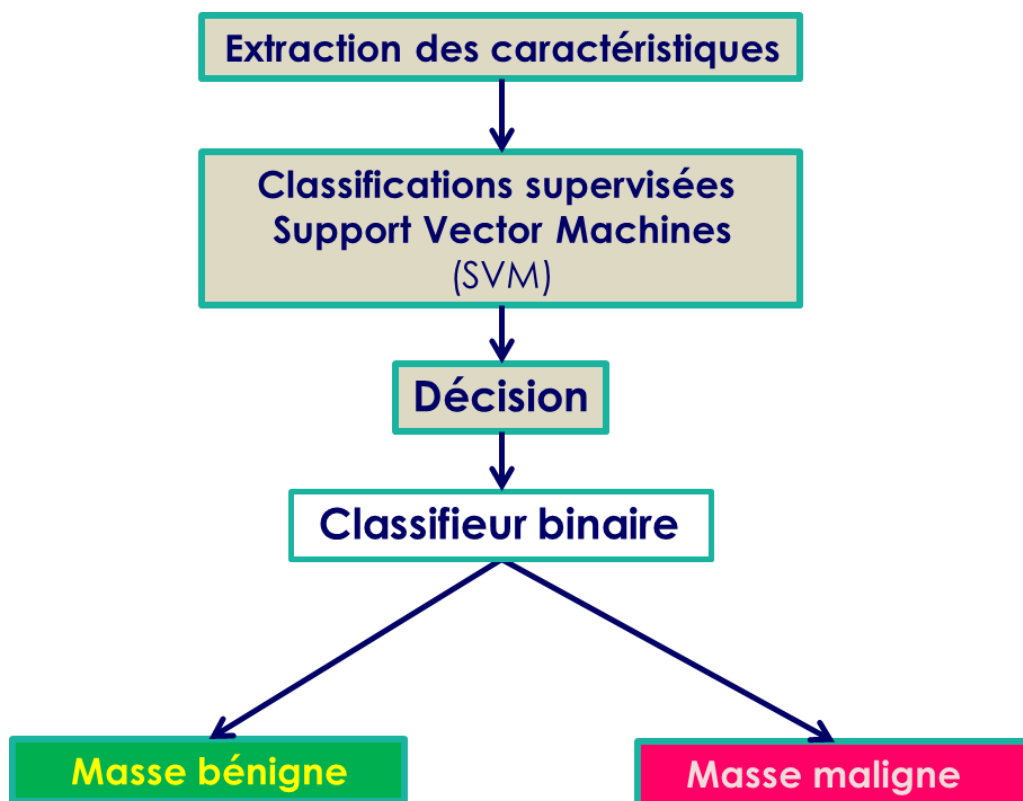


Figure 3. 12: organigramme de prise de décision.

3.5 Conclusion

L'analyse d'images englobe une multitude de domaines. Nous avons abordé celui de l'analyse de textures et de forme, dans ce chapitre. L'analyse de textures, est un problème difficile. Cette difficulté est due essentiellement, au fait qu'il n'existe pas de définition précise et rigoureuse, permettant de caractériser complètement la notion de texture.

Ce chapitre avait pour but de commenter les différents paramètres caractérisant les images mammaires, qui servent par la suite à leur interprétation et, à leur classification.

Nous avons également évoqué les algorithmes permettant l'extraction de ces paramètres, en l'occurrence les matrices de cooccurrence et le calcul de **la convexité, la circularité, l'excentricité, superficie et enfin la rectangularité.**

Dans notre travail, la discrimination de textures et de formes est réalisée, à partir des vecteurs d'attributs caractéristiques par une méthode rigoureuse et précise qui se base, sur les séparateurs à vaste marge.

Bibliographie

- [1] L. Malha, *Segmentation d'images texturées à partir des attributs fractals*, Tizi-Ouzou.
- [2] H. Majdoulayne, *Extraction de caractéristiques de texture pour la classification d'images satellites*, Toulouse, 2009.
- [3] M. Unser, *DESCRIPTION STATISTIQUE DE TEXTURES : APPLICATION À L' INSPECTION AUTOMATIQUE*, Lausanne, 1984.
- [4] K. S. I. D. R. M. Haralick, *Textural Features of Image Classification*, *IEEE Transactions on Systems*, 1973.
- [5] C. A. Cherif, *Elaboration d'un Descripteur de forme en vue de classification des clichés mammographiques (Les Masses)*.
- [6] A. Desgrez, C. Frouge, F. Archambaud, C. Iffenecker, B. O. Helal, L. André et Y. Bouchareb, *Numérisation et traitement des images médicales (III) Exemples de traitement d'image*.
- [7] C. BOUVRY, *RECONNAISSANCE DE CARACTÈRES ET DE SYMBOLES SUR VERRE GRAVÉ*.
- [8] I. CHEIKHROUHOU, *Description et classification des masses mammaires pour le diagnostic du cancer du sein*, Université d'Evry-Val d'Essonne, 2012.
- [9] CORTES et VAPNIK, *Support vector networks*, 1995, p. 273–297.
- [10] L. Mercier, *Les machines à vecteurs support pour la classification en imagerie hyperspectrale*, Grenoble, 2010.
- [11] StatSoft, «<http://www.statsoft.fr/index.php>,» StatSoft France, [En ligne]. Available: <http://www.statsoft.fr/concepts-statistiques/machine-learning/svm-reseaux-bayesiens-et-plus-proches-voisins.htm>. [Accès le 10 Septembre 2012].
- [12] F. Melgani et L. Bruzzone, *Classification of hyperspectral remote sensing images*, 2004.
- [13] 2008 Cambridge University Press, «stanford,» [En ligne]. Available: <http://nlp.stanford.edu/IR-book/html/htmledition/multiclass-svms-1.html>. [Accès le 10 Juin 2013].

Chapitre 2

Techniques de segmentation en imagerie mammographique

Ce chapitre est la clef de voûte de ce mémoire, nous y décrivons d'abord les méthodes permettant d'obtenir une segmentation automatique des régions des seins, ensuite nous rappelons brièvement le principe de la segmentation par croissance de région basé sur un processus itératif d'agrégation de pixels. Nous illustrons son principe sur quelques exemples et enfin, nous introduisons l'approche de segmentation que nous avons développée.

Dans le cadre de notre recherche sur les techniques employées pour la segmentation des mammographies, nous avons passé en revue un bon nombre d'approches mais avant de les présenter, il est nécessaire et important de définir la notion de « segmentation ».

2.1 Introduction [1]

Une image, est une projection d'une réalité tridimensionnelle sur un plan. Le traitement d'image est né, du désir de vouloir déléguer l'observation humaine aux machines. Le traitement d'image, se résume à l'analyse d'image. Celle-ci consiste à extraire d'une image numérique, des informations pertinentes, pour l'application chargée de les traiter et de les interpréter [2].

La segmentation d'images est une étape primordiale en traitement d'images, puisqu'elle conditionne la qualité de l'interprétation. Dans le domaine médical, son but est de délimiter les structures anatomiques et, pathologiques chez un patient d'une façon automatique. La segmentation automatique des cellules tumorales, dans une grande base d'images, est d'un grand intérêt: elle permet d'aider le médecin dans son diagnostic, en accélérant le processus d'analyse; en détectant les cellules tumorales, éventuellement oubliées.

La segmentation automatique des cellules, est aussi très utile pour l'extraction des caractéristiques de ces dernières, telle que leur dimension permettant ainsi aux praticiens, de déterminer le degré de gravité de la maladie [3]. Une bonne segmentation, aide le médecin par exemple à prendre une décision finale, avant son geste chirurgical [1].

La segmentation automatique d'images, est cependant une tâche difficile, du fait de la variabilité des images. Beaucoup de méthodes existantes dans la littérature, donnent de bons résultats mais elles nécessitent, des connaissances *a priori* qui sont généralement difficilement accessibles. En outre, les algorithmes de segmentation proposés dans la littérature, nécessitent un choix délicat d'opérateurs à utiliser et, un calcul compliqué de paramètres [3].

2.2 Système de détection automatique

De manière globale, un système de détection automatique de cancers en mammographie, se compose de deux branches: une dédiée à la détection des microcalcifications et, l'autre à la détection des opacités. Chacun de ces modules, peut se décomposer comme une étape de marquage, suivie d'une prise de décision [4].

Le marquage peut selon les cas, être composé d'une détection rapide suivie, d'une segmentation. La prise de décision se compose quant à elle, d'une étape d'extraction de caractéristiques suivie d'une étape de classification [4].

Certains de ces éléments, peuvent apparaître de manière plus ou moins implicite. La **figure 2.1** illustre la décomposition de haut niveau, des processus de détection automatique. Dans certains cas, un prétraitement des données permettant de mettre en évidence les signes recherchés, peut être utilisé. Néanmoins, cette étape montre généralement rapidement ses limitations dans la mesure où, pour vraiment mettre en évidence un motif, il faut être capable de le détecter, c'est le but principal de cette phase de prétraitement [4].

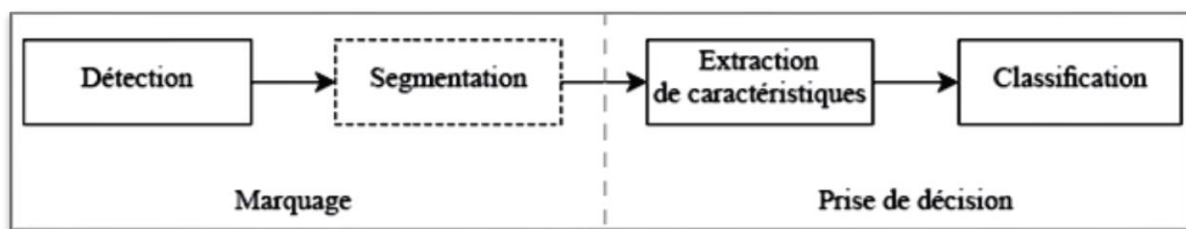


Figure 2.1 : schéma général d'une chaîne de détection de microcalcifications ou d'opacités en mammographie [4].

La chaîne de traitement est composée de trois étapes essentielles [3] :

1. l'étape de prétraitement qui consiste à améliorer la qualité de l'image acquise. Il peut s'agir d'éliminer le bruit ou de rehausser le contraste.
2. l'étape de segmentation qui sera détaillée, dans la suite de ce travail. Cette étape peut être suivie par une étape de post traitement (étape facultative), qui consiste à raffiner la segmentation afin d'améliorer les résultats obtenus.
3. l'étape d'interprétation, qui consiste à extraire l'information sémantique de l'image en identifiant les différents objets qui la constituent.

Il faut noter, que les résultats de la deuxième étape (**la segmentation**) conditionnent très fortement le résultat de la dernière étape (**l'interprétation**). Ceci, peut engendrer parfois des erreurs. Par conséquent, des méthodes d'évaluation de la segmentation sont généralement utilisées, afin de quantifier les résultats obtenus pour une meilleure interprétation de l'image [3].

2.2.1 Prétraitement des images

Une première étape, souvent utilisée, est donc la préparation des images avant la détection. En effet, les structures que l'on recherche n'étant pas toujours facilement discernables, une étape de prétraitement destinée à les mettre en évidence, peut faciliter leur détection [4].

Une approche couramment utilisée en traitement d'images, consiste à travailler sur l'histogramme; dans le but de définir une fonction de transfert sur les niveaux de gris permettant, de mettre en valeur les détails présents dans l'image. Néanmoins, le problème d'une telle approche, réside en sa limitation au niveau du traitement des textures, ce qui est gênant, puisque ces dernières portent une information importante pour la détection de signes radiologiques. D'autres approches plus locales, existent et se comportent mieux vis-à-vis des textures. Ainsi, **Braccialarghe et Kaufmann (1996)** [4] proposent une transformation des niveaux de gris en fonction d'éléments locaux comme, les contours ou les statistiques locales.

D'autres méthodes ont aussi été proposées comme l'utilisation de filtres d'amélioration de la netteté, ou encore la suppression, du fond de l'image. Cette dernière approche, consiste en la soustraction d'une version filtrée (par un opérateur passe bas) de l'image originale, à cette dernière. Certains auteurs proposent quant à eux, d'utiliser l'information d'orientation tandis que d'autres, utilisent une décomposition en ondelettes hexagonales [4].

D'autres formalismes sont aussi utilisés, pour accomplir la tâche d'amélioration du contraste. Ainsi, **Hassanien et Ali** [4], proposent d'utiliser les ensembles flous et, l'analyse de l'histogramme de l'image, pour atteindre ce but. L'idée d'un prétraitement, bien qu'intuitive, peut poser quelques problèmes. En effet, dans l'idéal on voudrait mettre en évidence, seulement les zones potentiellement suspectes dans le but de faciliter leur détection, ultérieurement [4]. Or, pour pouvoir accomplir cette tâche, il faudrait savoir quelles zones de l'image on doit améliorer, c'est-à-dire connaître les zones suspectes, ce qui est difficile puisque l'on cherche à améliorer l'image pour justement détecter ces structures. D'autre part, l'étape de prétraitement, peut aussi modifier certaines structures et, les rendre faussement suspectes; ce qui peut être problématique, pour l'étape de détection. De manière similaire, un prétraitement peut modifier substantiellement les propriétés de l'image, rendant la modélisation de l'étape d'extraction délicate [4].

2.2.2 Généralités sur la segmentation

2.2.2.1. Définition

La segmentation est un traitement de bas niveau, qui consiste à créer une partition d'une image A en sous-ensembles R_i , appelés régions, tels que [5]:

$$\left. \begin{array}{l} \forall i R_i \neq \emptyset \\ \forall i, j; i \neq j R_i \cap R_j = \emptyset \\ A = \bigcup_i R_i \end{array} \right\} \dots\dots\dots [5]$$

Une région est un ensemble connexe de pixels, ayant des propriétés communes (**intensité, texture, ...**) qui les différencient des pixels des régions voisines. Lorsque le nombre de régions est limité à deux, la segmentation prend le nom de **binarisation** et, se résume alors à séparer un objet, du fond. Dans certains cas, le nombre de régions est connu d'avance ainsi que, certaines caractéristiques de ces zones. Pour ces situations, la segmentation prend le nom de classification et, consiste à associer, chaque pixel de l'image à une de ces régions [5].

Il n'y a pas de méthode unique de segmentation. Selon **Cocquerez** [5], le choix d'une technique est lié à la nature de l'image, aux opérations situées en aval de la segmentation, aux primitives à extraire, aux contraintes d'exploitation. Du fait de cette diversité, il est difficile de définir une méthode universelle de segmentation qui convienne à toutes les situations. Cependant, en s'inspirant des systèmes de la vision humaine, la segmentation fait référence, aux notions de différence et de similarité. Il en découle deux approches courantes: l'approche « **frontière** » et, l'approche « **région** ». L'approche région, s'attache à faire apparaître des régions homogènes selon un critère (niveaux de gris, couleur ou texture). C'est une approche, qui fait référence à des groupements de points, ayant des caractéristiques similaires et, qui fait appel à la notion d'homogénéité; alors que, l'approche frontière tente de trouver des contours ou frontières de régions, présentant une variation rapide du même critère, i.e. une variation brutale de l'intensité lumineuse ou, une discontinuité entre les propriétés de deux ensembles, de points connexes. Ces deux approches, sont duales en ce sens, qu'une région définit une ligne par son contour et, qu'un contour fermé, définit une région. Elles amènent, cependant à des algorithmes différents et,

ne fournissent pas, en général, les mêmes résultats. Certaines méthodes, récentes tentent de faire coopérer plusieurs méthodes de segmentation, afin de combiner les avantages de chaque méthode ou d'exploiter leurs complémentarités [6].

Un algorithme de segmentation s'appuie donc sur :

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

Dans la plupart des cas, des prétraitements sont utilisés pour faciliter le travail de segmentation. Ces prétraitements, parfois intimement liés à la méthode de segmentation elle-même, sont appliqués à l'image selon sa nature, pour effectuer le plus souvent des transformations ponctuelles de luminance, une réduction du bruit (**lissage**) ou un rehaussement de contraste [2].

2.2.2.2 Travaux antérieurs sur la segmentation

Les approches les plus simples, reposent sur des seuillages¹ globaux de l'image. Une autre approche souvent usitée, fait appel aux méthodes de croissance de régions². L'idée est de faire croître, un masque représentant la lésion en agrégeant des pixels voisins de la région, qui ont des propriétés semblables, à cette dernière. La difficulté avec ce type d'approche, est la définition d'un critère d'arrêt pertinent, pour obtenir, des régions segmentées de manière optimale [4].

De manière plus générale, des méthodes de segmentation générique utilisant la morphologie mathématique [7] ou des outils de détection de contours, peuvent être appliquées en mammographie. Cependant ces outils assez simples, ne permettent bien souvent d'obtenir des résultats satisfaisants, qu'en combinaison avec des approches plus évoluées. De manière similaire, les méthodes de segmentation par contour actifs ou ensembles de niveaux ont aussi été utilisées, pour la segmentation de lésions. L'idée est de formuler, une énergie sur une représentation d'un contour (ensemble de points, forme paramétrique ou fonction implicite d'ordre supérieur) de telle manière qu'elle prenne en compte des informations *a priori* sur ce que l'on cherche à segmenter (la régularité du contour par exemple) ainsi que, des informations d'attache aux données (la segmentation doit correspondre aux zones de transition présentes dans l'image par exemple). Certains travaux, présentent une utilisation intensive des modèles déformables. Ainsi, ils proposent une première segmentation du noyau de la lésion, puis une segmentation de sa périphérie et le cas échéant, une troisième segmentation des spicules en utilisant des ensembles de niveaux. Dans d'autres méthodes, les contours actifs sont utilisés pour affiner une première segmentation assez grossière. La qualité des segmentations obtenues avec ce type d'approche, dépend beaucoup, de l'énergie utilisée pour juger de la qualité des contours. Cette dernière, comporte souvent beaucoup de paramètres et, sa minimisation est bien souvent locale. D'autres approches, reposant sur l'expression du problème encore sous forme d'énergie, mais cette fois avec un formalisme de chemins optimaux ont aussi été

¹Le **seuillage d'image** est la méthode la plus simple de segmentation d'image. À partir d'une image en niveau de gris, le seuillage d'image peut être utilisé pour créer une image comportant uniquement deux valeurs, noir ou blanc (monochrome) [19].

² Croissance de régions (**growing-region**) : cette technique consiste à faire progressivement grossir les régions autour de leur point de départ. L'algorithme se compose de deux étapes :

1. Trouver les points de départ des régions ;
2. Faire grossir les régions par agglomérations des pixels voisins.

présentées. L'idée, est d'utiliser des méthodes de programmation dynamique pour trouver le contour optimal de la lésion [4]. L'avantage de ce type d'approches, est d'offrir des outils efficaces pour optimiser l'énergie, cette fois ci de manière globale. Cependant, l'espace de recherche des contours est quant à lui limité, interdisant ainsi de considérer certains contours que, les modèles déformables peuvent trouver. Comme bien souvent dans d'autres champs applicatifs, la piste des méthodes de segmentation par régions a aussi été investiguée. Ainsi les champs de Markov aléatoires peuvent aussi être utilisés pour obtenir une segmentation des lésions. Dans ces derniers travaux, les champs de Markov sont combinés à une segmentation multirésolution. Des méthodes plus spécifiques du problème de segmentation de lésions dans le sein ont aussi été proposées. Ainsi **Kupinski et Giger** (1998) [4] proposent de considérer des partitions de la zone à segmenter et, de considérer que la meilleure segmentation, correspond à celle, qui maximise un indice de gradient radial. Les mêmes auteurs, proposent aussi une modélisation statistique pour trouver la partition optimale. La difficulté dans l'utilisation de ces approches, est d'obtenir un ensemble de partitions de qualité indépendamment de la structure à segmenter [4]. Ici aussi, l'utilisation de la logique floue pour le partitionnement d'images mammographiques a été investiguée. De même, des travaux relatant l'utilisation des méthodes par croissance de régions, dans le cadre du flou ont été proposés. On remarquera que l'utilisation des extensions floues, des différentes méthodes de segmentation courantes, se fondent généralement sur ce formalisme dans leur fonctionnement et, non dans leur décision finale. Ainsi si ambiguïté il y a, cette dernière risque d'être perdue une fois l'image segmentée. Des travaux un peu plus marginaux, proposent de traiter des lésions qui ne sont visibles qu'en partie dans l'image à cause, par exemple, d'une localisation trop proche du bord de cette dernière. Ce genre d'approche peut constituer une amélioration d'un système de détection, en traitant des cas identifiés comme difficiles. De manière plus générale, une revue assez complète des outils utilisés pour la segmentation ainsi que, pour les autres parties d'un système de détection a été proposée par **Cheng et al.** [4].

2.3 Présentation de la croissance de régions

Les méthodes par croissance de régions, de par leur relative simplicité de mise en œuvre, ont été l'une des premières familles d'outils utilisés pour la segmentation d'image. Dans [8], Zucker introduit les premiers travaux relatifs, à la croissance de région (figure 2.1).

À partir d'un ensemble de points initiaux, cette méthode agrège les pixels suivant un double critère: l'homogénéité et l'adjacence. Cette agrégation de pixels, est contrôlée par un prédicat (expression booléenne). Le processus s'applique sur une ou plusieurs régions. Pour chaque région, il comprend une phase d'initialisation et une phase itérative. La phase itérative, modifie la région R représentant l'objet segmenté. La région \bar{R} correspond à l'ensemble, des éléments non segmentés. Le processus de croissance de régions, étant itératif, le contenu d'une région à l'itération n est noté $R^{[n]}$ [9].

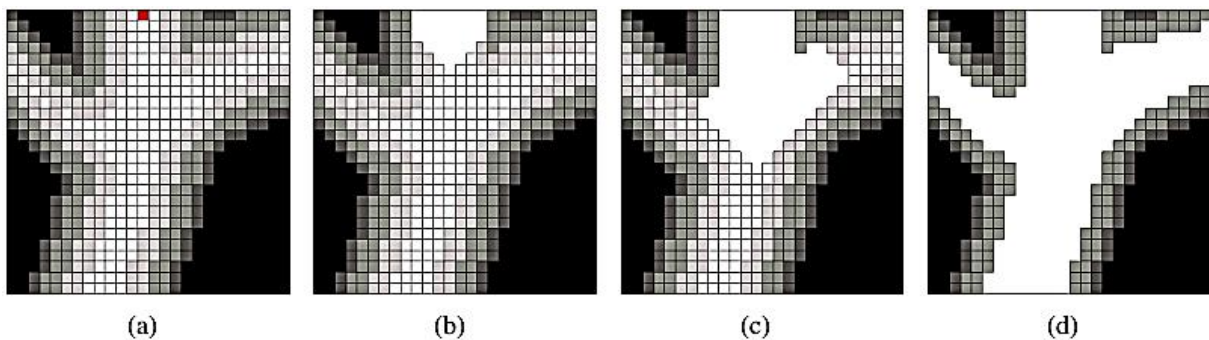


Figure 2. 2: principe de la croissance de région sur une image 2D : (a) image initiale à cinq niveaux de gris (point germe en rouge, en haut de l'image), (b) et (c) sont les étapes successives de la croissance de région jusqu'à obtention du résultat (d) [9].

2.3.1 Germes initiaux

L'initialisation de la croissance de régions, nécessite le positionnement de germes désignant la ou les premières régions (figure 2.1(a)). Bien entendu, le point ou la zone germe, doit faire partie de la région recherchée, au risque d'obtenir une segmentation partiellement, voire totalement erronée. Ces germes constituent donc, la région initiale $R^{[0]}$ [9].

2.3.2 Un processus itératif [9]

Suite à la phase d'initialisation, un processus itératif de déformation, ajoute progressivement des points situés à la périphérie de la région entrain de croître, s'ils respectent le prédicat. Ces nouveaux points, sont regroupés dans l'ensemble que l'on note $C^{[n]}$. Considérons une croissance de régions, ne faisant croître qu'une seule région. Étant donné une région

initiale, $R^{[0]}$, la région suivante $R^{[n+1]}$ est obtenue à partir de la région courante $R^{[n]}$ et, de la couronne $C^{[n]}$. Par la réunion suivante :

$$R^{[n+1]} = R^{[n]} \cup C^{[n]} \quad 2.1$$

Dans la croissance de régions, la condition d'agglomération implique la définition d'un terme de similarité, entre un point candidat et la région segmentée. Ce terme de ressemblance, appelé critère, est utilisé par le prédicat pour décider de l'ajout ou non, d'un pixel. En pratique, le critère implique presque toujours une mesure d'homogénéité sur les intensités des points de la région segmentée. La figure 2.2 présente un résultat de segmentation, par croissance de régions sur une image synthétique. Le processus a été initialisé, à partir de deux germes. Le résultat est donné à différentes itérations de l'évolution.

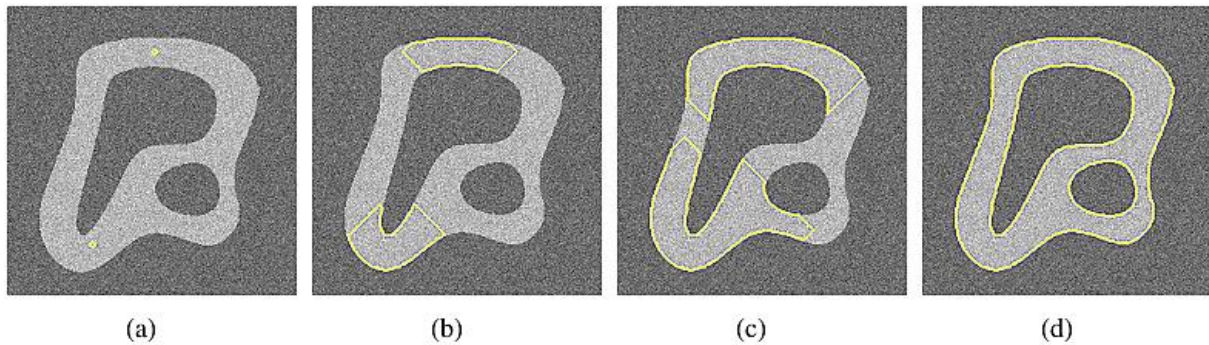


Figure 2. 3 : segmentation d'une image synthétique de 256 niveaux de gris. (b)-(d) évolution de la segmentation à partir de deux germes (a) [9].

2.3.3 Convergence, critère d'arrêt

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

- toutes les régions satisfont le prédicat et la segmentation comprend n régions;
- il existe $n - 1$ régions qui vérifient le prédicat, la $n^{\text{ième}}$ région comprend, les points ne vérifiant pas le prédicat.

2.4. Propriétés de la croissance de région

La croissance de régions, a été employée à de multiples reprises pour la segmentation des images médicales [9]. Cela s'explique par le fait que, c'est une méthode simple à initialiser et rapide. Dans cette section, les propriétés intrinsèques de cette méthode sont mises en évidence et illustrées sur des données médicales.

Quatre propriétés ont été retenues:

1. la croissance de régions, permet de segmenter un objet d'intérêt par la simple désignation de cet objet. Cette particularité la rend attrayante, pour les applications médicales.
2. La croissance de régions, n'intègre pas par défaut, de terme de régularisation. La méthode, ne nécessite pas de lisser le contour de la région segmentée, contrairement à d'autres méthodes. Dans le cadre de la quantification pour l'imagerie médicale, ce terme de régularisation peut provoquer une erreur de mesure.
3. Une propriété consécutive au point précédent et, à l'origine de la principale faiblesse de la méthode : sa sensibilité aux points de fuite. Un point de fuite, se définit par la connexion de deux régions distinctes, aux propriétés statistiques semblables. La figure 2.4(a), montre l'image d'origine sur laquelle est appliquée la croissance de régions. Un artéfact est visible en haut à droite dû à la modalité d'acquisition de l'image. La figure 2.4(b), montre, le contour du résultat de la segmentation. La croissance de région se propage dans l'artéfact et, empêche une segmentation correcte de l'objet recherché.
4. Cette méthode tolère les changements topologiques, de la carte de segmentation (en régions) au cours de son évolution. La figure 2.3, montre l'évolution de la croissance de régions sur une image contenant une forme avec deux anses. Cet exemple illustre la capacité, de cette méthode à détecter des objets à topologie complexe.

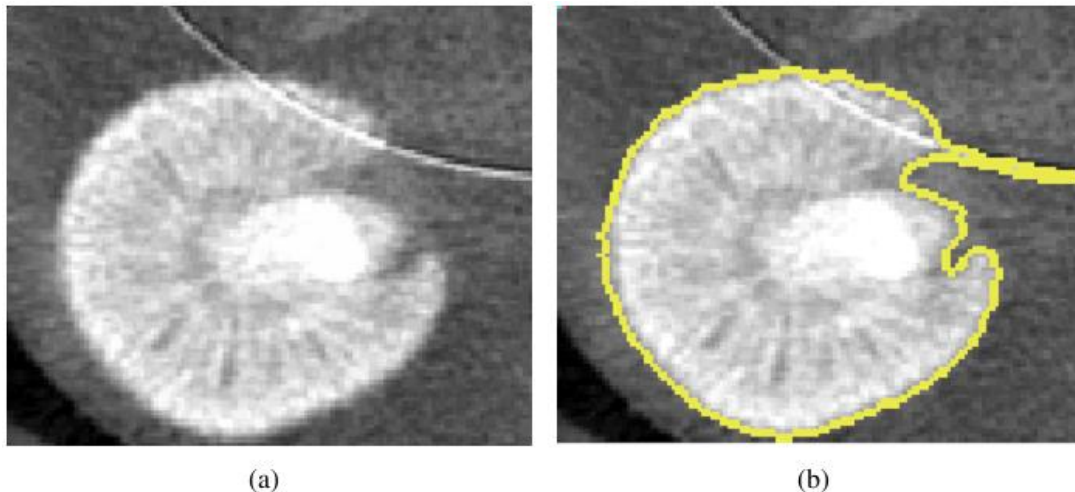


Figure 2. 4: sensibilité de la croissance de région aux points de fuites : (a) image initiale avec un point de fuite (artéfact) en haut à droite de l'image et (b) résultat de la segmentation [9]

2.5. Paramètres de la croissance de régions [9]

Dans le contexte de l'imagerie médicale, nous avons montré que la croissance de régions, est un choix pertinent pour répondre aux besoins de la segmentation. De nombreuses variantes, ont été proposées pour améliorer le processus. Elles peuvent être classées, en deux groupes suivant si l'effort porte sur :

- le type d'agrégation : sélection de points candidats à l'agglomération ;
- le choix de la mesure ou du critère : global, adaptatif.

Dans le premier groupe, les variantes portent sur l'évolution de la carte de segmentation et, agissent sur des paramètres tels que les germes ou, les points candidats à l'agglomération. Dans le second groupe, elles se focalisent sur l'élaboration de nouveaux critères globaux, adaptatifs ou concurrentiels. Dans ce chapitre, nous dressons une liste des différents paramètres de la croissance de régions.

2.5.1. Les germes

L'initialisation la plus intuitive des germes, est la sélection manuelle. Même si le positionnement d'un germe [10] (ou de plusieurs) reste l'approche la plus commune, cette intervention humaine n'est pas envisageable pour les applications automatisées. Plusieurs possibilités, existent pour procéder automatiquement à cette initialisation. Dans les croissances de régions appelées "unseeded region growing" Lin et al. [11], prennent un germe au hasard dans l'image et, construisent une première région à partir de ce germe. Lorsque la région converge, la méthode choisit un autre germe parmi les points non segmentés. Ce type d'approche nécessite, fréquemment une étape supplémentaire de

fusion de régions pour diminuer, le nombre de régions obtenues. La fusion, supprime les régions trop petites et, regroupe les régions ayant les mêmes caractéristiques [12]. La sélection des germes peut se faire également par des seuillages très sélectifs de l'image (seuil bas, seuil haut). Ainsi, à chaque seuillage, les germes d'une région particulière sont isolés. Garza [13] ajoute à cette technique une contrainte de concentration spatiale des germes pour supprimer les points aberrants. Un point de l'image ayant son intensité comprise dans l'intervalle de seuillage est considéré comme un germe si et seulement si un certain nombre de ses voisins ont leur intensité comprise dans le même intervalle. Une autre approche consiste à utiliser des connaissances a priori pour localiser grossièrement l'objet à segmenter. Dans le contexte de l'imagerie médicale par exemple, il est judicieux d'exploiter l'information anatomique et fonctionnelle. Lorsqu'on dispose d'un modèle d'organe, il peut être ajusté par recalage et servir de point de départ pour la segmentation.

2.5.2 Points candidats à l'agglomération

Les points candidats à l'agglomération [9], sont des pixels périphériques à la région entrain de croître et, sont généralement définis par le voisinage N_E . Ils appartiennent à un ensemble C tel que :

$$C = \{x \in \bar{R} | N_E(x) \cap R \neq \emptyset\} \quad 2.2$$

Notons que certaines approches, permettent une remise en cause des pixels agrégés. Il faut alors parler de points candidats au rejet. Dans ce cas, la région segmentée n'est plus nécessairement connexe. Différents critères basés sur des statistiques, comme nous l'avons vu plus haut, ont pour but, d'évaluer la qualité de la carte de segmentation à un instant donné. Dans cette section, nous détaillons quelques critères, se basant sur les statistiques de l'image. De nombreuses propositions, ont été publiées pour améliorer ces critères. Un critère peut être introduit, dans la croissance de régions de deux façons. La première présentée au paragraphe 2.5.3, consiste à comparer un terme de similarité à une valeur seuil. Tandis que la seconde, utilisée dans le cadre de croissance de régions concurrentielle, cherche à minimiser une erreur, en ajoutant le pixel le plus ressemblant à une région à chaque itération.

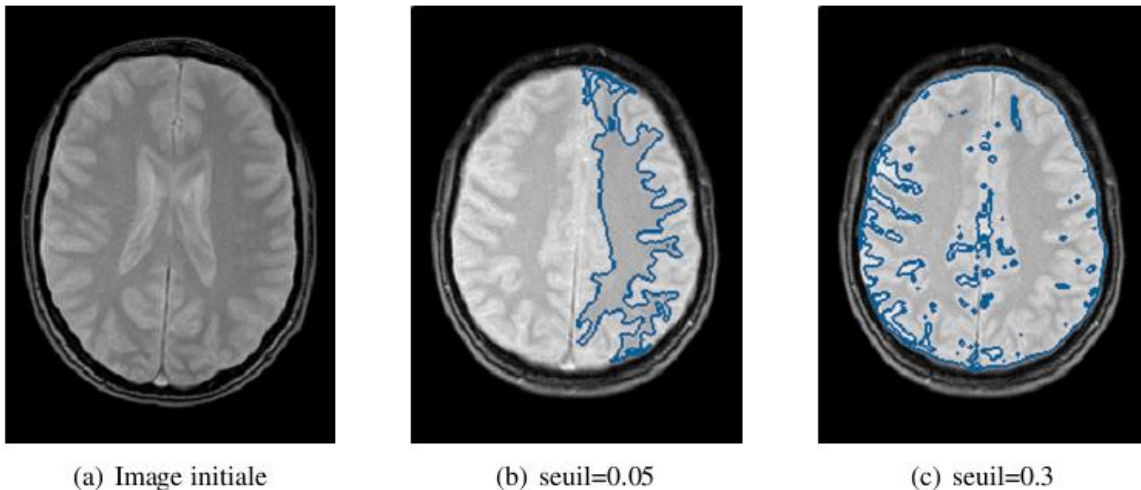


Figure 2.5 : sensibilité de la croissance de région vis à vis d'un seuil [9].

2.5.3 Croissance de régions sans concurrence

La croissance de régions [9], la plus simple consiste à agréger deux pixels x et y si leurs intensités $I(x)$ et $I(y)$ sont similaires. Ce critère de similarité $f(\cdot)$ est ensuite utilisé par le prédicat qui la compare à un seuil. Plus précisément, $f(I(x); I(y)) \leq \delta$, où δ est un seuil fixé. Cette inégalité définit le contrôle qui autorise ou non la déformation de la carte de segmentation. Cette méthode est attractive de par sa simplicité, mais la détermination de l'hyper-paramètre est un problème. L'ajout d'une condition sur l'agglomération des pixels voisins a contribué à l'amélioration de la méthode. Dans [14], un pixel x est aggloméré à un pixel adjacent y d'une région homogène, si la différence entre leurs niveaux de gris $I(x)$ et $I(y)$ est faible et si $I(x)$ est proche de l'intensité du germe s .

$$\begin{aligned} |I(\mathbf{x}) - I(\mathbf{y})| &\leq \delta \\ |I(\mathbf{x}) - I(\mathbf{s})| &\leq \beta, \end{aligned} \quad 2.3$$

Cette méthode, ne résout pas le problème du choix de δ et β qui a une forte influence sur le résultat de la segmentation. Pour un mauvais réglage de ces hyper-paramètres, la croissance de régions peut se propager à l'extérieur de la région à segmenter. La figure 2.4 montre le résultat, de la segmentation pour deux seuils différents et, le même germe. Les structures segmentées, diffèrent suivant les deux seuils. Tandis que la quasi-totalité du cerveau est segmentée pour un seuil, seulement la matière blanche de la partie droite du cerveau, est segmentée pour le second seuil.

2.5.4 Croissance de régions concurrentielle

La croissance de régions [9], dite concurrentielle, consiste à mettre en compétition deux ou plusieurs régions. Cette approche consiste, à faire évoluer des régions parallèlement. Le pixel qui minimise au mieux un critère, est ajouté à la région correspondante. La plupart de ces modèles de segmentation, sont basées sur une méthode appelée croissance de régions avec germes (SRG³). On n'en dira pas plus sur cette méthode, car elle n'a pas fait l'objet de notre étude.

2.6. Morphologie mathématique

La morphologie mathématique, est un ensemble de méthodes d'analyse d'images. Elle offre un grand nombre d'outils, très puissants de traitement et d'analyse d'images. Les outils proposés ont été développés au départ, pour traiter des images binaires : on fait alors de la morphologie mathématique **ensembliste**. Leur utilisation a été ensuite, étendue aux images en niveaux de gris : on parle, à ce moment-là, de morphologie mathématique **fonctionnelle**. Certains opérateurs fonctionnent désormais sur des images couleur [15].

Les opérations de base sont l'érosion⁴ et la dilatation⁵ [15] [16]. Ceux-ci sont à la base des autres opérateurs plus complexes tels que l'ouverture⁶, la fermeture⁷, le chapeau haut de forme [16] [17] (voir ci-après).

La morphologie mathématique, **ensembliste** ou **fonctionnelle**, s'appuie sur un élément structurant, dont on choisit la forme et la taille en fonction de ce que l'on souhaite faire [15]. Les transformations d'images, en morphologie mathématique fonctionnelle se pratiquent comme pour la morphologie mathématique ensembliste : l'élément structurant B est déplacé de façon à ce que son origine x, passe par toutes les positions de l'image. Pour chaque position, on comparera les valeurs prises par les pixels inclus dans le domaine de l'élément structurant. Le pixel central, prendra soit la valeur minimale (lors d'une érosion) soit la valeur maximale (lors d'une dilatation) [17].

³SRG : de l'anglais Seeded RegionGrowing

⁴**Érosion** : enlever un pixel sur les pourtours de la forme (agrandit les trous).

⁵**Dilatation** : ajouter un pixel sur les pourtours de la forme (boucher les trous).

⁶**Ouverture** : l'ouverture consiste en une érosion suivie d'une dilatation (supprime les pics mais préserve les vallées).

⁷**Fermeture** : la fermeture consiste en une dilatation suivie d'une [15]érosion (La fermeture comble les vallées, elle homogénéise et éclaircit l'image)

2.6.1 Chapeau haut de forme et chapeau bas de forme

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

Les propriétés de la fermeture et l'ouverture en teintes de gris, permettent de définir des filtres pouvant extraire les pics de la fonction $f(x)$ correspondant, aux petites zones claires de l'image ou les vallées, correspondant aux petites zones sombres. Ces filtres, sont appelés **Chapeaux Haut de Forme** et, consistent à effectuer une simple différence entre l'image d'origine et l'ouverture (ou la fermeture). L'appellation Chapeau se justifie par le fait que ces zones, sont représentées sur le graphe de la fonction $f(x)$ par des pics (resp. des vallées étroites) [16].

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

On appelle cette transformation **Chapeau Haut de Forme Blanc (White Top Hat)** [17] [15] [18].

$$\boxed{WTH(x) = f(x) - f_{Bn}(x)} \quad [16] \quad 2.4$$

Pour extraire les vallées d'épaisseur inférieures à n , on effectue la différence entre la fermeture de l'image par un élément structurant B de taille n et l'image initiale. Cette transformation est appelée **Chapeau Haut de Forme Noir (Black Top Hat)** [15] [19] [18].

$$\boxed{BTH(x) = f_{Bn}(x) - f(x)} \quad [16] \quad 2.5$$

Les Chapeaux Haut de forme blanc et noir, sont des filtres dits morphologiques et sont adaptés pour l'extraction d'objets linéaires fins, d'une certaine épaisseur, qu'il soient sombres ou clairs [16] localisant; ainsi des structures dans l'image qui sont plus petites que l'élément structurant [18]. Ils sont en général appliqués, seulement aux images à niveaux de gris [16].

2.6. Notre approche

L'organigramme de la **Figure 2.6** illustre les étapes, que nous avons élaborées afin de segmenter les différentes opacités à savoir : les masses⁸, les distorsions architecturales et enfin les masses spiculées.

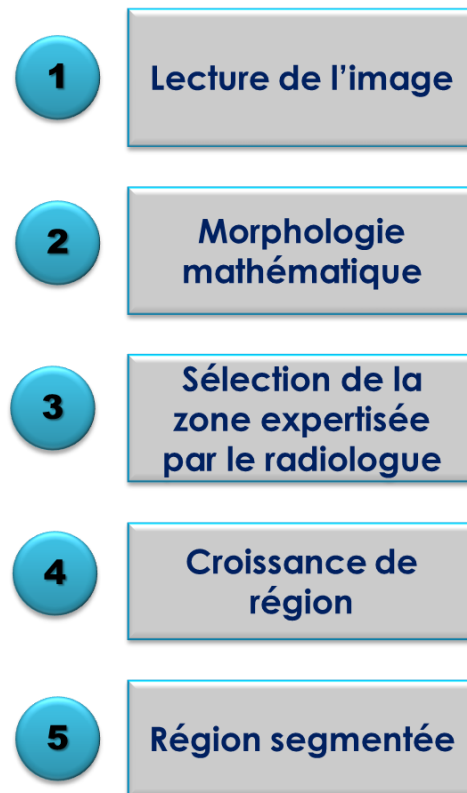


Figure 2.6 : algorithme de segmentation des masses.

1^{ère} étape :

Afin de lire sous **MATLAB** des images « classiques » il faut utiliser l'ordre **MATLAB** « imread ». Celui-ci reconnaît un certain nombre de formats standards (l'image « processing toolbox » en reconnaît la quasi-totalité)

2^{ème} étape :

Pour mettre en évidence les masses, nous avons utilisé les fonctions de la morphologie mathématique :

1. Chapeau haut de forme (**Black Top Hat**).
2. Chapeau bas de forme (**White Top Hat**).
3. Soustraction entre le résultat du Chapeau haut de forme et celui, du Chapeau bas de forme.

⁸ Voir chapitre 1: pathologies mammaires

4. pour un meilleur contraste nous avons superposé (additionner) le résultat de l'étape 3 à lui-même.

Dans notre étude, nous avons utilisé des images mammographiques identifiés, par les radiologues experts. Il est joint aux images un fichier comportant des informations sur chaque image et parmi celle-ci, nous trouvons les coordonnées (X, Y) du centre de l'anomalie ainsi que le rayon approximatif (en pixels) d'un cercle entourant l'anomalie. Nous avons donc exploité cette information pour une meilleure détection des masses, en choisissant le germe comme étant le centre de la tumeur, suivant les travaux de [20].

4ème étape :

- **Germe** = centre de la tumeur désignée par le médecin radiologue ;
- La région croit en comparant tous les pixels voisins non affectés à la région ;
- La différence entre la valeur de l'intensité d'un pixel et la moyenne de la région, est utilisée en tant que mesure de similitude.

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

2.6

- Le pixel avec la plus petite différence mesuré de cette façon, est affecté à la zone respective.

Ce processus s'arrête lorsque la différence d'intensité entre la région moyenne et, le nouveau pixel devient plus grand qu'un certain seuil « t »

2.6 Conclusion

La segmentation d'images, est une étape fondamentale dans le traitement des images médicales. Elle représente, toujours le point de départ des processus d'analyse de formes, de détection des lésions, de visualisation et d'estimation. La segmentation conditionne très fortement, le résultat de l'interprétation si bien, qu'une bonne segmentation assure une bonne caractérisation des régions segmentées. Dans le chapitre 3, nous passerons à l'étape de caractérisation, en utilisant des méthodes basées sur l'analyse de la forme et de la texture, pour étudier la classification des tumeurs en utilisant les machines à supports de vecteurs (SVM).

Bibliographie

- [1] Lotfi BEN ROMDHANE, Béchir AYEB Hechmi SHILI, Une Nouvelle Approche pour la Segmentation des Images Mammographiques, Mars 21-24, 2012.
- [2] MANGOUA SOFACK William, "Utilisation de la logique floue pour traitement," Hanoï, Juillet 2008.
- [3] Sarah GHANDOUR, Segmentation d'images couleurs par morphologie mathématique: application aux images microscopiques, juillet 12, 2010.
- [4] Giovanni Palma, Détection automatique des opacités en tomosynthèse numérique du sein, Février 23, 2010.
- [5] Sylvie Philipp Jean-Pierre Cocquerez, *Analyse d'image. Filtrage et segmentation*, Masson, Ed. Paris, 1995.
- [6] LEHAMEL Malha, "Segmentation d'images texturées à partir des attributs fractals," Tizi-Ouzou,.
- [7] C. VACHIER & S. MULLER & S.BOTHOREL, "ANALYSE MORPHOLOGIQUE DES CLICHES MAMMOGRAPHIQUE. EN VUE DE LA DETECTION DES OPACITES DU SEIN," 1997.
- [8] Steven W. Zucker, Region growing : Childhood and adolescence, Computer Graphics and Image Processing, 1976.
- [9] Jean-Loïc ROSE, Croissance de région variationnelle et contraintes géométriques tridimensionnelles pour la segmentation d'image, Décembre 5, 2008.
- [10] Christophe Lohou, Laurent Perroton & Gilles Bertrand Petr Dokladal, *Liver Blood Vessels Extraction by a 3-D Topological Approach. Medical Image Computing and Computer-Assisted Intervention.*, 1999.
- [11] Jesse Jin, Hugues Talbot Zheng Lin, *Unseeded region growing for 3D. ACM International Conference Proceeding Series.*, 2000.
- [12] Xiaobo Li Yian-Leng Chang, Adaptive image region-growing, Novembre 1994.
- [13] Meer P, Medina V. Garza-Jinich M, Robust retrieval of three-dimensional structures from image stacks., 1999.
- [14] Koichi Sano, Tetsuo Yokoyama Hiroyuki Sekiguchi1, Interactive 3-dimensional segmentation method based on region growing method, 1994.

- [15] <http://riadh.webuda.com>. [Online].
<http://riadh.webuda.com/SUPPORT%20DE%20COURS%20IMAGE/chapitre4.html>
- [16] Annick LEGELEY PADOVANI, "Contribution de la Morphologie Mathématique à la Cartographie de l'Occupation du SOI à partir d'Images SPOT," Bondy, 1999.
- [17] <http://urfist.enc.sorbonne.fr>. [Online].
http://urfist.enc.sorbonne.fr/anciensite/image_numerique/chapitre3_7.htm#intro
- [18] Annick LEGELEY-PADOVANI and Catherine MERING, "APPORT DES OPÉRATEURS DE MORPHOLOGIE MATHÉMATIQUE A L'EXTRACTION DES FAILLES À PARTIR D'IMAGES SPOT PANCHROMATIQUE," p. 212, 1997.
- [19] wikipedia. [Online]. http://fr.wikipedia.org/wiki/Seuillage_d'image
- [20] G. Rabottino, A. Mencattini, M.Salmeri, F. Caselli, and R.Lojacono, Performance evaluation of a region growing procedure for mammographic breast lesion identification, 2011.

Chapitre 1

Introduction à la mammographie et cancer du sein

Chaque année, environ 7500 cas du cancer du sein, sont enregistrés dans notre pays, tout en sachant que ce type de néoplasie¹ en Algérie, vient en tête des tumeurs malignes chez la femme et, constitue la première cause de mortalité chez la gent féminine, avec environ 3500 décès enregistrés chaque année [1].

Dans le but de réduire la mortalité par cancer du sein, une détection précoce est nécessaire motivant ainsi, des campagnes de dépistage chez les femmes à partir d'un certain âge variant entre 40 et 50 ans selon les pays [2]. Le type d'imagerie actuellement utilisée pour cette tâche, est la mammographie. Dans ce contexte, la quantité de clichés que le radiologue doit interpréter en un temps limité, est importante. De plus, la complexité de la tâche de détection est particulièrement élevée, en mammographie. Ces raisons motivent l'utilisation d'outils d'aide à la détection par ordinateur (CAD)², afin d'améliorer les performances du lecteur, en minimisant en particulier le nombre de signes radiologiques associés à des lésions malignes qui, pourraient ne pas être détectées [2].

Dans ce chapitre, nous définirons la notion du cancer du sein ainsi que, les différentes pathologies mammaires. Nous introduirons brièvement la mammographie conventionnelle et numérique dans le cadre de la détection du cancer du sein et, nous citerons l'un des problèmes majeurs qui est la présence, d'artefacts radiopaques sur le fond de ce type d'images médicales.

¹**Néoplasie:** une néoplasie est une tumeur ou une croissance anormale de nouveaux tissus.

²**CAD:** Computer Aided Medical Diagnosis.

1.1 Anatomie mammaire

Les seins, jouent un rôle important dans la féminité et dans l'image que la femme, a de son corps. La fonction biologique du sein, est de produire du lait afin de nourrir un nouveau-né. La structure du sein est complexe. Chaque sein (appelé aussi **glande mammaire**), est composé de quinze à vingt compartiments séparés par du tissu graisseux qui donne au sein la forme qu'on lui connaît. Chacun de ces compartiments, est constitué de lobules et de canaux. Le rôle des

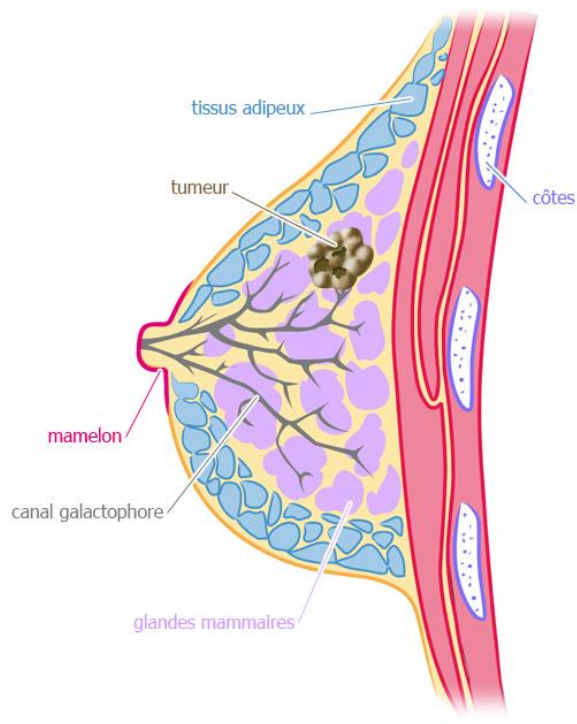


Figure 1. 1: anatomie du sein [4]

lobules est de produire le lait en période d'allaitement ; les canaux transportent ensuite le lait vers le mamelon [3]. Un sein (cf. **Figure 1.1**) contient divers types de tissus [4] :

- des glandes, formées de petits « sacs » (**les lobules**) où se trouvent, les cellules productrices de lait pendant l'allaitement ;
- des canaux de lactation, qui font communiquer les lobules avec le mamelon pour permettre le passage du lait ;
- de la graisse, pour soutenir lobules et canaux ;
- de la peau qui enveloppe le tout.

Un cancer est une maladie de la **cellule**: l'unité de base de la vie. Il en existe dans le corps plus de deux cents types différents. Toutes ont un rôle précis: **cellules musculaires, nerveuses, osseuses**, etc. Une cellule cancéreuse, est une cellule qui s'est modifiée. Habituellement, ces modifications sont réparées par l'organisme. Cependant, lorsque la cellule devient cancéreuse, elle perd ses capacités de

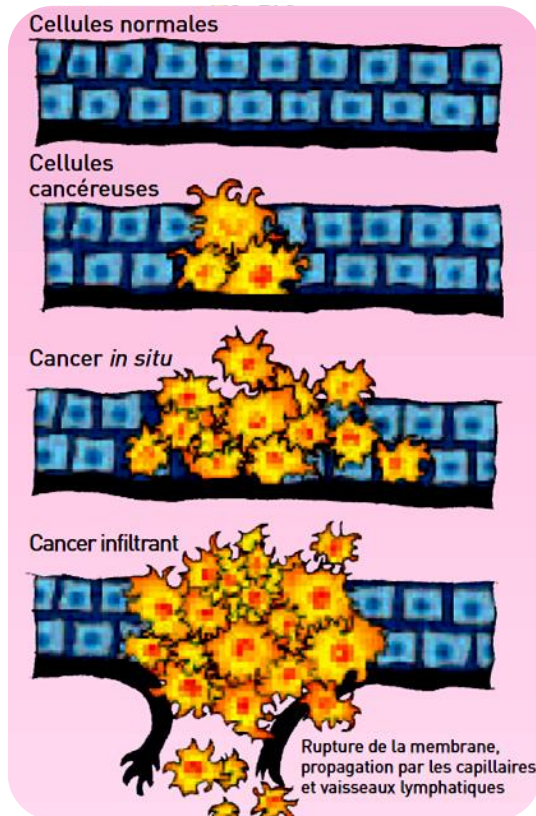


Figure 1. 2 : cellules cancéreuses [3]

cellules à l'origine du cancer et, selon

l'aspect de la tumeur, on distingue différents cancers du sein. Par exemple, un cancer du sein, qui touche les cellules qui bordent les canaux sera dit « **canalaire** » et, un cancer qui affecte les cellules des lobules sera dit : « **Lobulaire** ».

De plus, certains cancers du sein, sont provoqués par des cellules cancéreuses, dont la multiplication est favorisée par les hormones féminines, les estrogènes. Ce sont les cancers du sein dits « **hormonodépendants** » ou « **hormonosensibles** ». En cas de sensibilité des cellules cancéreuses aux estrogènes, le médecin peut bloquer la croissance de la tumeur, en supprimant ces hormones par le biais de médicaments adaptés [4].

1.1.1 Cancers du sein non invasifs

On parle de cancer du sein « **non invasif** » ou « **in situ** », lorsque la tumeur reste dans le tissu d'origine et n'envahit pas les tissus voisins (**cf. Figure 1.2**). Ces cancers sont

³**Métastase**: tumeur formée par des cellules cancéreuses, provenant d'une première tumeur. On parle également de maladie métastatique, ou de localisation secondaire du cancer [3].

plus facilement traités que les cancers, qui envahissent les tissus voisins (« cancers invasifs », voir ci-dessous). Dans **90 %** des cas, les cancers du sein non invasifs, sont de type canalaire (ils se forment à l'intérieur des canaux de lactation), les **10 %** restants étant de type lobulaire [4].

1.1.2 Cancers du sein invasifs

Les cancers du sein sont dits « **invasifs** » ou « **infiltrants** » (cf. **Figure1.2**) lorsque les cellules cancéreuses ne restent pas confinées à leur lieu d'origine et, envahissent les tissus avoisinants, les ganglions locaux (dits «ganglions axillaires » et situés sous l'aisselle) et, parfois, le reste du corps [4]. Elles s'accumulent alors dans les **ganglions**⁴lymphatiques voisins. On dit qu'il existe un envahissement ganglionnaire.

Les cellules cancéreuses, ont tendance à migrer dans d'autres organes ou parties du corps, et à y développer de nouvelles tumeurs qu'on appelle métastases. On dit dans ce cas, que le cancer est métastatique. Tous les cancers, ne se comportent pas de la même façon. C'est la raison pour laquelle, il est nécessaire d'envisager un traitement adapté à chaque cancer. Toutes les thérapies, ont pour but de supprimer les cellules cancéreuses. En l'absence de traitement adapté, le cancer risque d'évoluer plus vite [3].

⁴**Ganglion** : petit renflement réparti le long des vaisseaux lymphatiques. Disposés dans certaines parties du corps, les ganglions sont soit superficiels (dans le cou, l'aisselle, l'aîne), soit profonds (dans l'abdomen, le thorax). Les ganglions jouent un rôle essentiel, dans la protection du corps contre les infections ou les cellules cancéreuses. Ils mesurent normalement moins d'un centimètre de diamètre. Si leur taille est anormale, on parle d'adénopathie [3].

1.2. Mammographie

1.2.1. Naissance de la radiologie

C'est le 8 novembre 1895 que le physicien allemand **Wilhelm Konrad Röntgen (1845-1923)** [5] découvre, dans son laboratoire de l'Université **Julius-Maximilian** de **Würzburg**, un nouveau type de rayonnement capable de traverser la matière. Ne sachant comment désigner ces rayons pénétrants et invisibles, il les baptise "**rayons X**", **X** comme l'inconnue en mathématique. **Wilhelm Röntgen** a mis en évidence ce phénomène grâce à la fluorescence



écran au platino-cyanure de baryum excité par le rayonnement issu d'un tube de Crooks⁵ mis sous tension et, totalement enfermé dans une boîte en carton. À la suite de cette découverte, il s'aperçoit rapidement que l'atténuation des **rayons X** varie en fonction de la densité des matériaux ou des tissus traversés. L'utilisation d'un détecteur permettant la visualisation de l'absorption différentielle des **rayons X** par les tissus, complète la découverte; la radiologie est née. Ceci représente une véritable révolution dans le monde médical, car il est possible désormais, de visualiser "**in vivo**"⁶ et sans dommage apparent, la structure interne de l'organisme. Depuis plus de cent ans, l'évolution de l'imagerie médicale a été considérable, mais le principe découvert par **Röntgen** en 1895, reste toujours à la base de la radiologie moderne. Cependant, quelques semaines après la découverte des **rayons X**, on signalait déjà les premiers effets sanitaires dont, ils étaient responsables, comme l'apparition de brûlures sur la peau lors d'expositions importantes; l'irradiation du patient devient alors un sujet de préoccupation et tout est mis en œuvre, pour diminuer la dose administrée à celui-ci [5].

⁵ Tube portant le nom de son inventeur William Crooks (1832-1919). Crooks, physicien et chimiste britannique, découvrit le thallium (1861), inventa un tube électronique (1872) et montra que les rayons cathodiques sont des particules possédant une charge (1878) [5].

⁶ À l'intérieur de l'organisme vivant.

1.2.2 Définition de la mammographie

Une mammographie (ou **mastographie**) est un examen radiologique du (ou des) sein (s). Il permet d'obtenir, des images de l'intérieur du sein à l'aide de **rayons X** et, de détecter ainsi certaines anomalies [6]. Lors d'une mammographie, chaque sein est radiographié de face et de profil, pour visualiser l'intégralité de la glande mammaire. Chaque sein, est comprimé entre le détecteur du mammographe et, une plaque transparente (**palette de compression**), pour réduire ainsi l'épaisseur que les **rayons X** traverseront (cf. **Figure 1.3**). Cette compression, étalera le tissu du sein en évitant ainsi la superposition des structures et, elle limitera aussi le mouvement qui pourrait brouiller l'image radiographique. Le but de la mammographie est de produire de façon systématique, des images de haute résolution (pour une appréciation plus détaillée des microcalcifications) et hautement contrastées (pour une bonne détection des masses) tout en délivrant la plus petite dose de rayonnement possible [2].

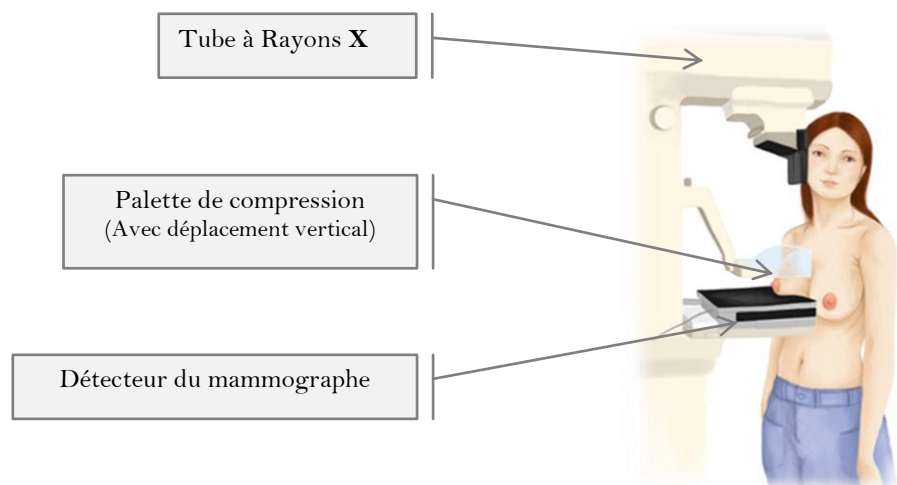


Figure 1. 3:positionnement de la patiente pour la prise d'une mammographie de face [2], [7].

1.2.3 Circonstances de l'examen [8]

Un examen mammographique peut se pratiquer dans 2 grands types de circonstances :

- soit la femme, a remarqué une anomalie d'un ou des **2** seins (autopalpation d'une masse, écoulement anormal, aspect inflammatoire de la peau), soit c'est son médecin traitant ou gynécologue qui, en l'examinant à titre systématique, a découvert une anomalie clinique: la mammographie est alors dite **diagnostique**, car

elle va chercher à objectiver, une anomalie radiologique qui est à l'origine du problème. Les douleurs de seins (mastodynies) sont un symptôme fréquent, de consultation mais, la très grande majorité des cancers du sein, ne sont pas douloureux et, les douleurs n'ont habituellement pas d'équivalence radiologique quel que soit le type d'imagerie utilisée;

- soit la patiente ne présente aucune symptomatologie fonctionnelle, aucune plainte, et l'examen clinique des seins réalisé par le médecin traitant ou le gynécologue, par ailleurs n'a révélé aucune anomalie visible ou palpable. La mammographie est alors réalisée dans un cadre de **dépistage** de cancer du sein afin de détecter un cancer dit « infra-clinique ».

*La quantité de **rayonsX**, délivrée lors d'une mammographie est très faible. Il n'y a donc pas de risque de développer un cancer du sein, à cause des **rayonsX** reçus, même si des mammographies de dépistage, sont répétées [6].*

1.2.4 Récepteurs d'image

Les récepteurs basés sur le **couple écran-film**, ont été les détecteurs de référence utilisés en mammographie conventionnelle. Mais, les nouveaux développements technologiques, en matière de détecteurs et d'ordinateurs ont altéré le parcours de la mammographie, en introduisant **les détecteurs numériques**. Actuellement les détecteurs utilisés en mammographie, sont classés principalement en deux catégories: **couples écran- film** et **détecteurs numériques (Figure 1.4)**. De ces derniers, ils existent des détecteurs qui effectuent, une conversion directe (détection de quanta et numérisation en une seule étape) et, des détecteurs qui réalisent une conversion indirecte (détection de quanta et numérisation réalisées en deux étapes). La **Figure 1.4**, montre une classification, des différents détecteurs et technologies utilisés en mammographie [2]. Bien que la partie détection, ne nous intéresse pas, une brève description de chaque technologie est donnée ci-après.

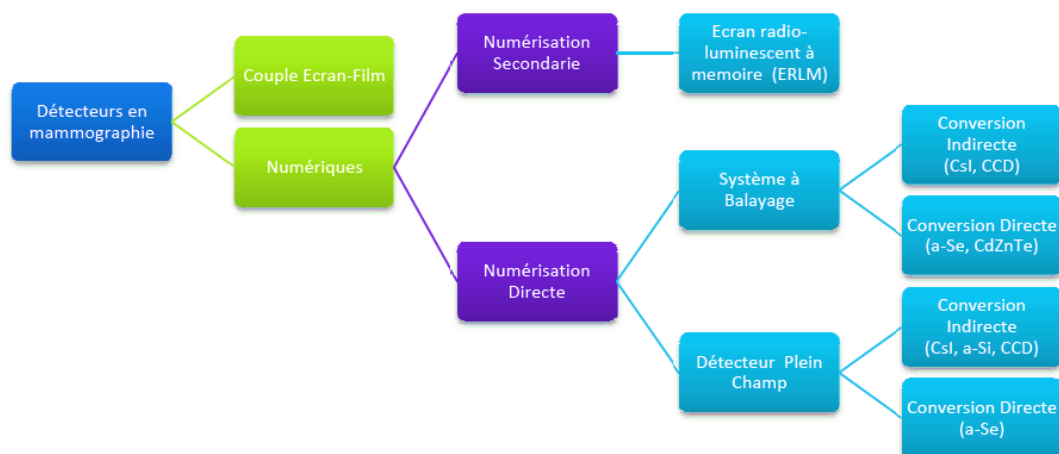


Figure 1. 4:schéma précisant les différents types de capteurs et technologies en mammographie numérique [2]

1.2.4.1 Couple écran-film

La mammographie analogique avec **couple écran-film**, est considérée depuis des années comme le « **gold standard** » pour le dépistage, des cancers du sein. Elle permet, comme l'ont démontré de grandes études randomisées⁷, de réduire la mortalité par cancer du sein. Cependant, environ **10 %** des cancers palpables restent occultes à la mammographie [9]. Les couples **écran-film** sont constitués, d'un film argentique monocouche et, d'un écran photo-luminescent(**gadolinium oxysulfure, Gd₂O₂S**) ayant, la propriété d'émettre de la lumière lorsqu'il est irradié par des **rayons X**. Cette lumière, est la responsable de l'exposition du film argentique. Quand les **photons X** sont absorbés, la lumière générée par l'écran **photo-luminescent**, illumine le film avec un nuage de photons lumineux qui expose le film. Donc, plus l'écran est épais, plus les **rayons X**, sont capturés pour former l'image ce qui se traduit par une meilleure efficacité d'utilisation de dose. Mais, un écran plus épais présente l'inconvénient de générer, plus de diffusion des photons lumineux. Ceci dégrade donc, la résolution spatiale du système, même si la sensibilité augmente. Pour cette raison, la relation entre la dose de radiation et la qualité de l'image, doit être optimisée, pour une application clinique spécifique. Les limites du couple écran-film sont [2]:

⁷ En statistiques, valider un résultat en fonction d'un échantillon choisi au hasard.

- une faible latitude d'exposition (voir **Figure 1.5**);
- la granularité du film, rend le film plus ou moins sensible, en variant la résolution spatiale. Plus le grain est fin, moins l'émulsion est sensible. Inversement, plus le grain est grand, plus l'émulsion est sensible mais ceci s'accompagne, d'une perte de résolution spatiale;
- le temps de développement, les difficultés de stockage et de transmission constituent également une limitation.

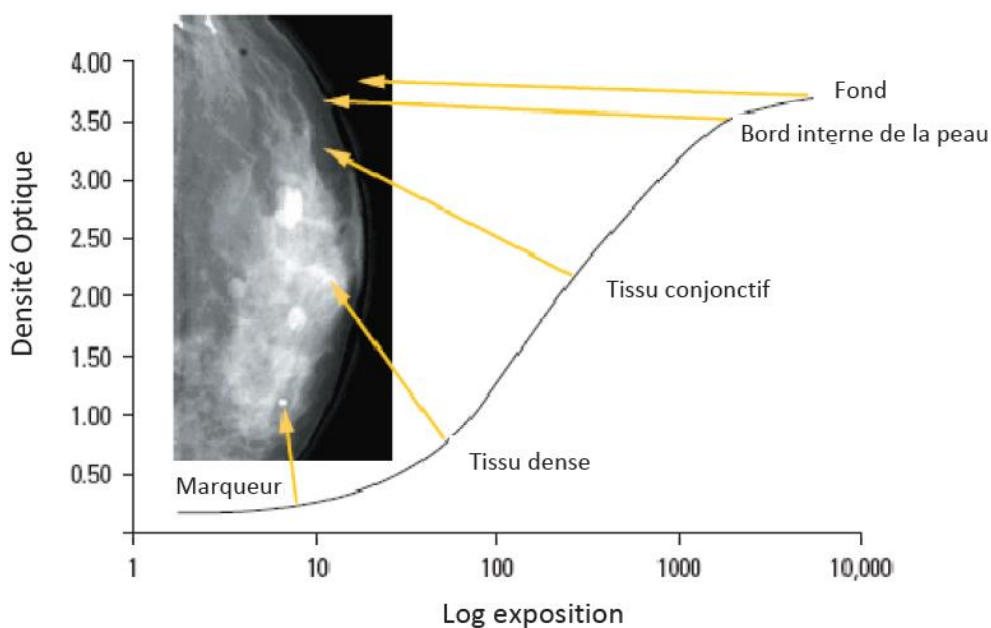


Figure 1. 5:relation entre le logarithme de l'exposition et la densité optique, d'un couple écran-film utilisé en mammographie [2].

1.2.4.2 Mammographie numérique

La mammographie numérique (**MN**) est une technologie, relativement récente qui a remplacé le film radiographique par des détecteurs solides, qui transforment les **rayons X** en signaux électriques. Les signaux électriques, produisent des images du sein qu'on peut afficher, à l'écran d'un ordinateur ou, imprimer sur un film spécial. Il est ainsi possible de faire un « **ZOOM** » ou d'améliorer les images, qui peuvent être archivées et transmises sur support électronique. Cette technologie devrait améliorer la qualité des images du sein, en particulier des tissus très denses sur le plan radiologiqueet, nécessiter une dose de rayonnement moins importante que la

mammographie classique sur film. Elle pourrait aussi réduire, le nombre d'examens d'appoint. La manipulation d'image, que permet cette technologie appuie l'établissement du diagnostic [10].

Au niveau de la détection, les systèmes numériques présentent une gamme dynamique beaucoup plus grande que le film, c'est-à-dire qu'ils fonctionnent sur une large plage d'exposition. En outre, ils offrent la possibilité d'effectuer des traitements de l'image, de modifier les paramètres d'affichage [5] et de transmettre, les données par réseau informatique. On peut encore ajouter à ceci, le fait qu'avec la radiologie numérique, il est possible de stocker et d'archiver les images, sous un format électronique. Nonobstant ces avantages, la radiologie numérique présente une "chaîne" plus complexe que celle des systèmes conventionnels, utilisant comme support de l'image, le film. Ainsi, le seul fait de caractériser les performances "physiques" du détecteur, n'est plus suffisant pour apprécier la qualité globale du système diagnostic; les paramètres liés au traitement informatique des images et, à leur affichage sur le support de visualisation, doivent être pris en compte [5]. On peut citer d'autres inconvénients, notamment [2]:

- une résolution spatiale plus petite que celle du couple écran-film;
- un coût élevé de la mise en œuvre et;
- une ergonomie des consoles de lecture encore insuffisante, pour permettre d'afficher 8 clichés sur un seul moniteur.

La mammographie numérique (cf. Figure 1.6), peut utiliser soit des détecteurs dédiés (numérisation directe), soit des détecteurs amovibles (**ERLM**, écran radio-luminescent à mémoire).

Même si l'on a annoncé que l'utilisation de la MN, pour le dépistage présenterait des avantages, il n'est pas clair que la nouvelle technologie, l'emportera sur la mammographie sur film avec écran (**MFE**) classique, dans la détection rapide du cancer du sein. La mammographie sur film, demeure l'étalon-or dans le dépistage du cancer du sein [10].

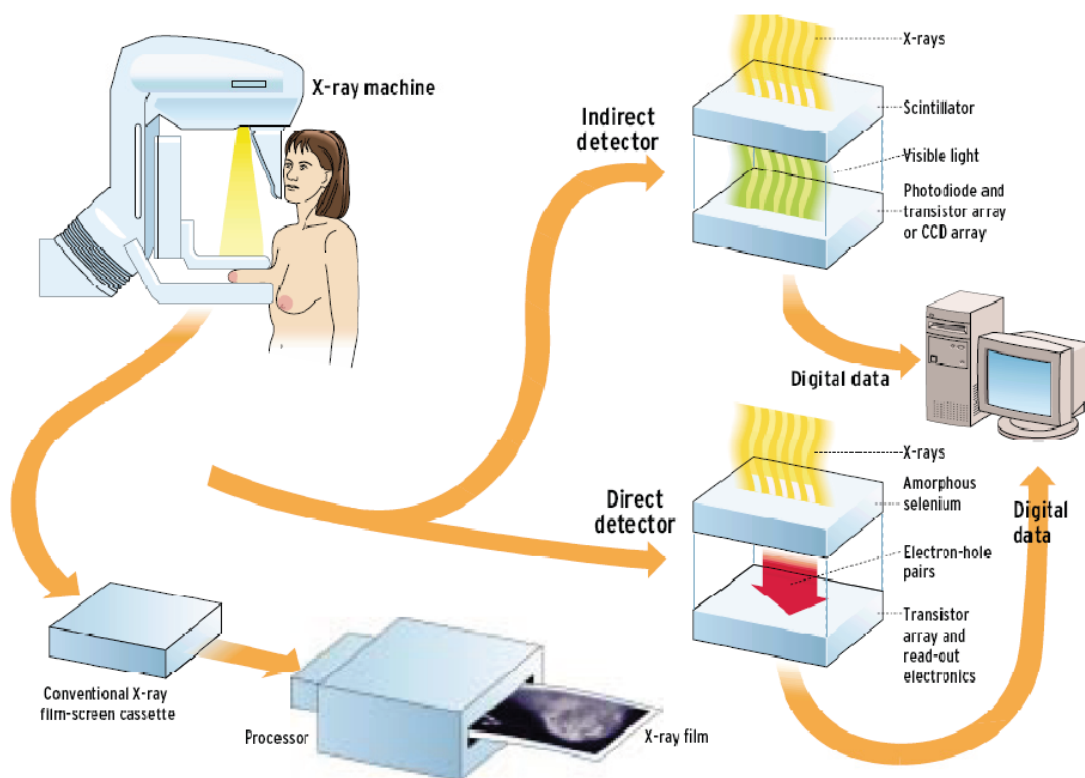


Figure 1. 6:résumé des technologies de détection actuellement employées dans la mammographie [11]

- La mammographie conventionnelle, utilise les cassettes film-écran.
- La mammographie numérique, utilise soit des détecteurs à conversion indirecte (photons X \Rightarrow lumière \Rightarrow Signal électrique), soit des détecteurs à conversion directe (photons X \Rightarrow Signal électrique) [2].

1.2.5 Les artefacts dans une mammographie [12]

Une mammographie numérisée, renferme généralement deux régions distinctives: la région exposée du sein et, la région non exposée (**fond**). L'interprétation visuelle de la mammographie, a souvent comme conséquences, l'identification d'artefacts radiopaques, qui peuvent être liés fortement au sujet, compliquant ainsi la segmentation des tissus mammaires et, la reconnaissance de structures anormales. Alors que le système visuel humain, peut facilement ignorer de tels objets pendant l'interprétation, un système automatisé en mammographie doit au préalable identifier et, classifier ces artefacts qui engendrent des erreurs d'interprétation lors de

l'analyse des images. Actuellement, plusieurs systèmes **CAD** fonctionnent, sur des mammographies numérisées. Sur de telles images, apparaissent souvent des artefacts radiopaques, sous plusieurs formes : étiquettes, éraflures, cales (voir *figure 1. 7*) [12]. Résumons ces artefacts, en deux catégories principales: étiquettes de film et artefacts opaques.

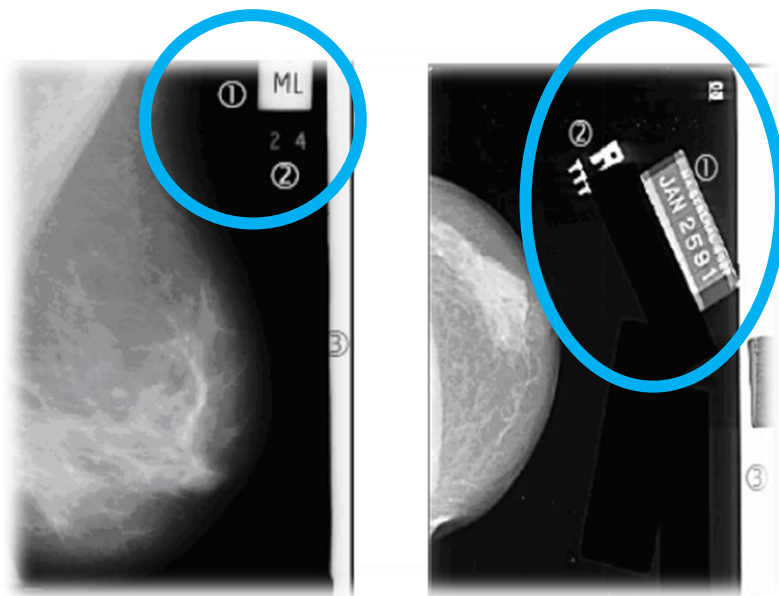


Figure 1. 7:un exemple de mammographies contenant des artefacts radiopaques.
1 : étiquettes de film, **2** : marqueurs radiopaques et **3** : coin [12].

1.2.5.1 Étiquettes du film mammographique [12]

Les mammographies, sont généralement marquées avec une certaine forme d'étiquette permanente, d'identification contenant l'information sur l'examen effectué. Ces labels, sont des indicateurs radiopaques, exhibant la latéralité de la mammographie (**R/L, Right/Left**) ainsi que, des indicateurs de projection **MLO /CC**⁸(voir *figure 1. 8*).

⁸Par exemple une vue mediolatérale oblique (MLO) du sein droit est marquée RMLO et, une vue craniocaudale (CC) du sein gauche, est marquée LCC.

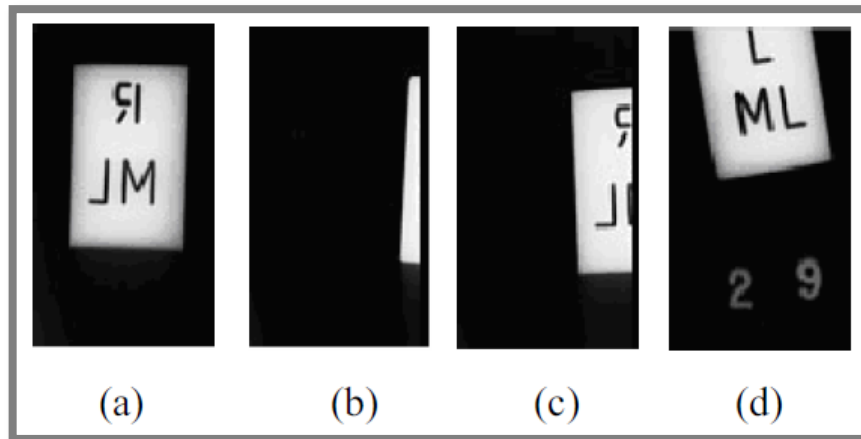


Figure 1. 8 :exemples d'étiquettes du film mammographique: (a) label entier et (b-c-d) labels partiels [12].

1.2.5.2 Artefacts radiopaques

On distingue deux types d'artefacts radiopaques: les bandes de haute intensité ou, les coins et, les marqueurs opaques. Ces marqueurs, sont des labels où, le texte est en haute intensité (le rectangle englobant le texte n'existe pas). Les coins, sont des bandes de haute intensité qui se situent, le long du bord de la mammographie (voir *figure 1.9*) [12].



Figure 1. 9 :exemples de marqueurs opaques sur une mammographie [12].

1.3 Anomalies mammaires

On peut classer les anomalies en trois catégories [13] :

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

1.3.1 Masse

Une masse est une opacité, visible sur les deux incidences. Elle se différencie du parenchyme⁹ mammaire normal par sa taille, son contour et sa densité [14]. La description d'une masse, doit contenir les informations quant à [14] :

- la taille ;
- le contour ;
- les signes associés ;
- la présence ou l'absence de calcification ;
- La localisation ;
- les modifications sous contrôle.

Dans le catalogue BI-RADS ¹⁰		
la forme est décrite soit :	Le contour peut être décrit comme :	La densité de l'image est décrite par rapport à la structure du parenchyme normal:
<ul style="list-style-type: none"> • Ovale ou ronde ; • polylobée ; • irrégulière ; • modification d'architecture. 	<ul style="list-style-type: none"> • bien limité ; • microlobulé ; • masqué (superposition partielle par des structures parenchymateuses voisines) ; • flou ; • avec des spicules (prolongement radiaire). 	<ul style="list-style-type: none"> • densité augmentée ; • isodense; • Hypodense ; • contenant des lobules adipeux ;

Tableau 1. 1 : description d'une masse dans le catalogue BI-RADS

⁹Ensemble des tissus assurant la fonction d'un organe

¹⁰Classification **BI-RADS** : il est souhaitable, de standardiser le compte rendu mammographique en employant les catégories telles qu'on les trouve dans le catalogue BI-RADS (**Breast Imaging Reporting and Data System**). Cela permet une standardisation du compte rendu mammographique et, du diagnostic [14].

La taille ne prédit pas le caractère malin, sauf sur des clichés successifs lorsque l'on voit la taille augmenter, régulièrement. On distingue traditionnellement **5** formes différentes de masses : **ronde**, **ovale**, **lobulaire**, **irrégulière** ou une **distorsion architecturale** (voir *Figure 1.10*) [13].

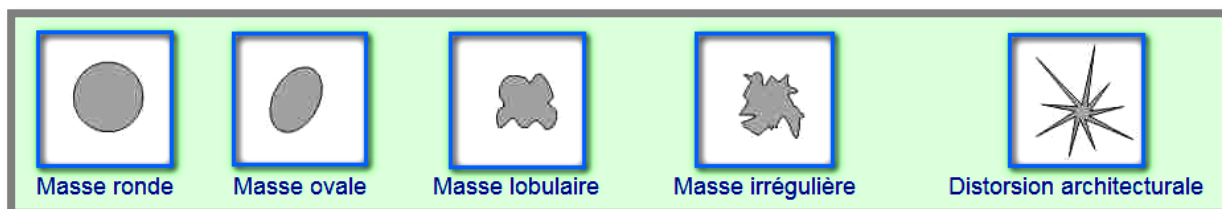


Figure 1. 10: différentes formes de masses [13]

La localisation de la masse, est bien analysée à partir de plusieurs clichés: celles qui sont petites, ne sont pas palpables. La localisation périphérique, est évocatrice de malignité. La taille ne prédit pas le caractère malin, sauf sur des clichés successifs, lorsque l'on voit la taille augmenter régulièrement. Les bords de la masse, sont très importants à étudier lors, de l'analyse sémiologique et, nécessitent parfois l'utilisation de la loupe. On décrit **5** types de bords : **circonscrits**, **effacés**, **micro-lobulés**, **mal définis** et avec des **spicules** (voir *Figure 1.11*) [13].

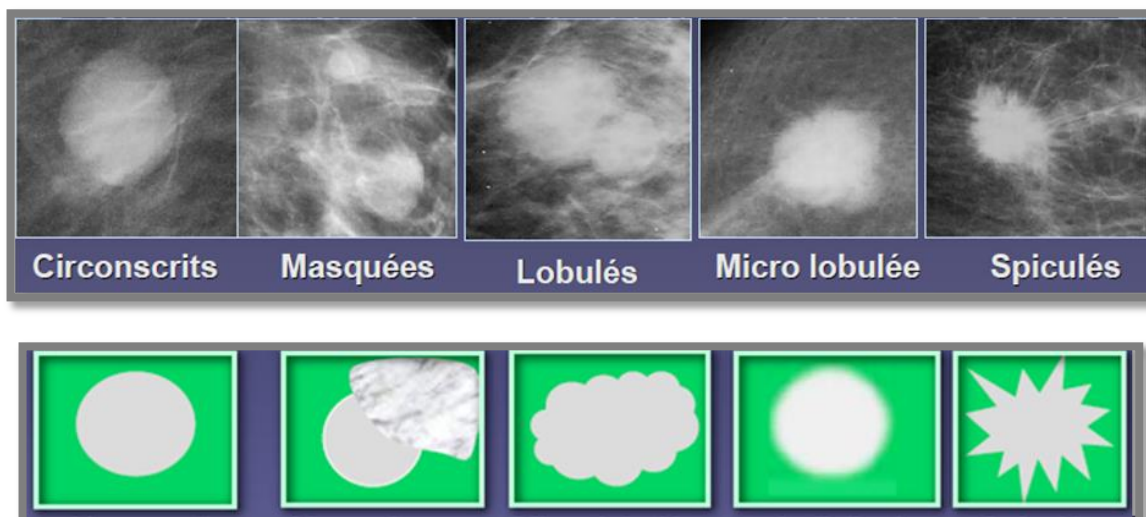


Figure 1. 11: types de contours des masses [15].

Les bords effacés, correspondent à une superposition du tissu adjacent. Les bords mal définis ou les spicules, correspondent à une invasion dans le tissu sain. Les masses sont en général denses (plus blanches) que, le tissu mammaire adjacent (voir *figure 1.11*)

Enfin, les masses peuvent s'accompagner de distorsions architecturales, de dilatation des canalicules, de modifications cutanées ou du mamelon [13].

1.3.1.1 Les masses régulières

Les masses rondes, ovales ou lobulées avec des limites bien définies, sont a priori des lésions bénignes. Cependant, cette règle n'est pas absolue et de petits cancers, peuvent être rencontrés. En principe, lorsqu'il y a plusieurs masses, l'aspect est plutôt évocateur de lésions bénignes [13].

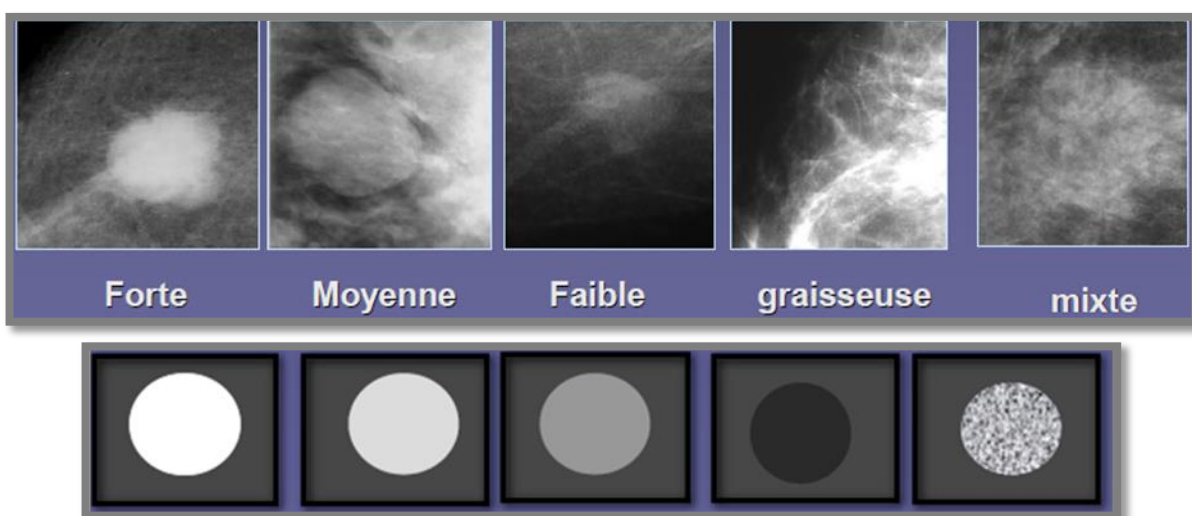


Figure 1. 12:différentes densités de masses [15].

a. Images avec des bords lobulés ou micro-lobulés

La plupart des masses, ne sont pas parfaitement rondes ou ovales mais, présentent un certain degré de lobulation. Plus le nombre de lobulations augmente, plus grand est le risque d'avoir une tumeur maligne [13].

b. Images à bords mal définis

Lorsque la masse, est bien visible et montre un bord mal défini, il faut redouter une lésion maligne. En échographie, ces lésions, si elles sont pleines, sont très suspectes de malignité [13].

1.3.1.2 Les masses irrégulières

a. Distorsion architecturale

Une distorsion architecturale, n'est pas une masse en soi, mais une réaction tissulaire traduisant, une interruption focale de la distribution normale du tissu mammaire. A la mammographie, elle donne l'impression d'une convergence du tissu normal vers un foyer central. On peut retrouver, une telle distorsion aussi bien près de la graisse sous-cutanée ou rétro-mammaire, qu'en plein parenchyme glandulaire.

Une telle distorsion, permet de localiser un cancer, mais on peut l'observer en cas de cicatrice chirurgicale, de modification fibro-kystique ou simplement parfois, en cas de superposition de tissus denses. Si aucune explication clinique n'est retrouvée, une telle lésion doit le plus souvent être vérifiée, par une biopsie [13].

b. Masse avec des limites en spicules

Une masse aux bords étoilés, est très suspecte d'un cancer: les spicules, représentent la modification tissulaire en rapport avec l'infiltration tumorale [13].

1.3.2 Asymétrie de densité [16]

Les asymétries de densité, sont le plus souvent bénignes. Elles correspondent soit à une asymétrie de répartition du tissu mammaire normal, soit à une asymétrie focale. Par définition, ce type d'anomalie ne peut être décrit par les autres qualificatifs de forme. Elles peuvent être détectées, sur une ou deux incidences mammographiques. En cas d'asymétrie focale, il ne faudra pas hésiter à compléter le bilan standard par des clichés localisés, voire des agrandissements, afin d'éliminer la présence de critères péjoratifs.

Les arguments en faveur de la bénignité sont :

- une forme variable dans l'espace selon les incidences ;
- l'absence de surdensité centrale, la présence de graisse au sein de la densité ;
- des bords concaves ;
- l'absence de microcalcifications ;

- l'absence de distorsion architecturale associée (bien regarder en lumière forte, l'alignement normal des crêtes de Duret sous-cutanées) ;
- la normalité de l'examen clinique, de grande valeur lorsque l'asymétrie décrite est de grande taille.

Les arguments en faveur d'une lésion suspecte, sont la présence d'un seul des critères décrits ci-dessus. L'échographie, est alors très utile pour rechercher une masse, une atténuation anormale des ultrasons dans la topographie de l'asymétrie mammographique. En cas de normalité de l'échographie et, en présence de signes cliniques péjoratifs, un prélèvement percutané sera discuté, afin de ne pas passer à côté d'un carcinome lobulaire infiltrant (la clinique prime sur l'imagerie) [16].

N.B

L'étude des asymétries, se fait en positionnant les mammographies en miroir: normalement, on doit observer une répartition grossièrement symétrique des densités dans les deux seins [13].

1.3.3. Calcifications mammaires

Les microcalcifications (voir **Figure 1.13**) sont des images calciques de petite taille (< **0,5 mm, en général comprises entre 200 et 500 μ**) ce qui les distingue des macrocalcifications (voir **Figure 1.14**) qui sont toujours bénignes. **10%** sont des cristaux d'oxalate de calcium, toujours bénignes et, **90%** correspondent à du phosphate de calcium et, peuvent être bénignes ou malignes. Elles sont facilement vues en mammographie du fait de leur forte densité: elles apparaissent comme des ponctuations blanches sur le gris des tissus [17]. Les calcifications, doivent être décrites selon leur localisation, leur nombre, leur topographie et, leur morphologie. Ces renseignements, ne peuvent être obtenus que par la comparaison des deux incidences [14].

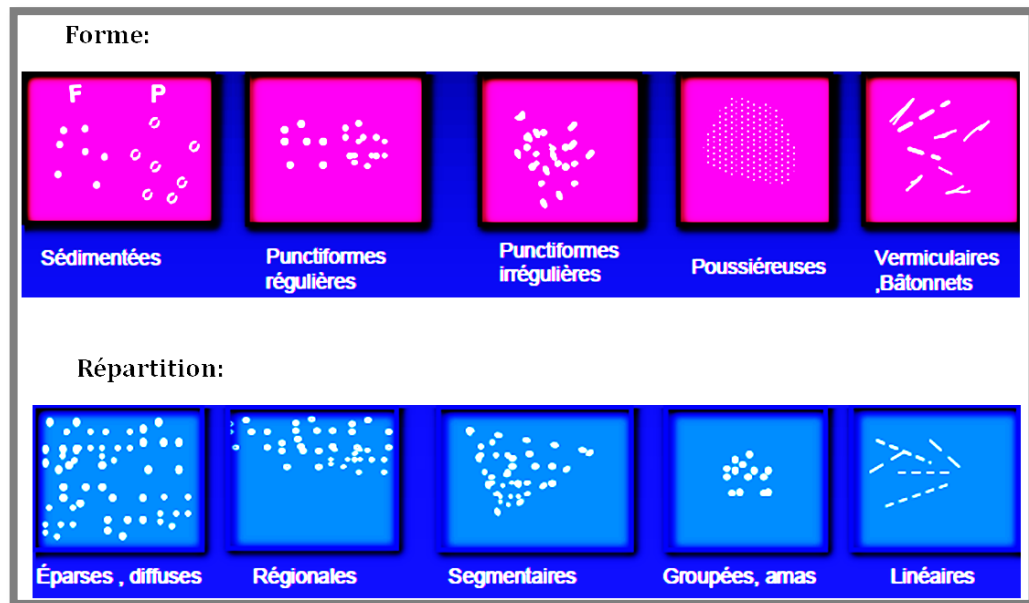


Figure 1. 13: microcalcifications [15].

1.3.3.1. Localisation anatomique [13]

Elles peuvent être:

- soit à l'intérieur ou au niveau des canalicules mammaires,
- soit au niveau des lobules,
- soit au niveau des structures vasculaires,
- soit au niveau du tissu conjonctif inter-lobulaire.

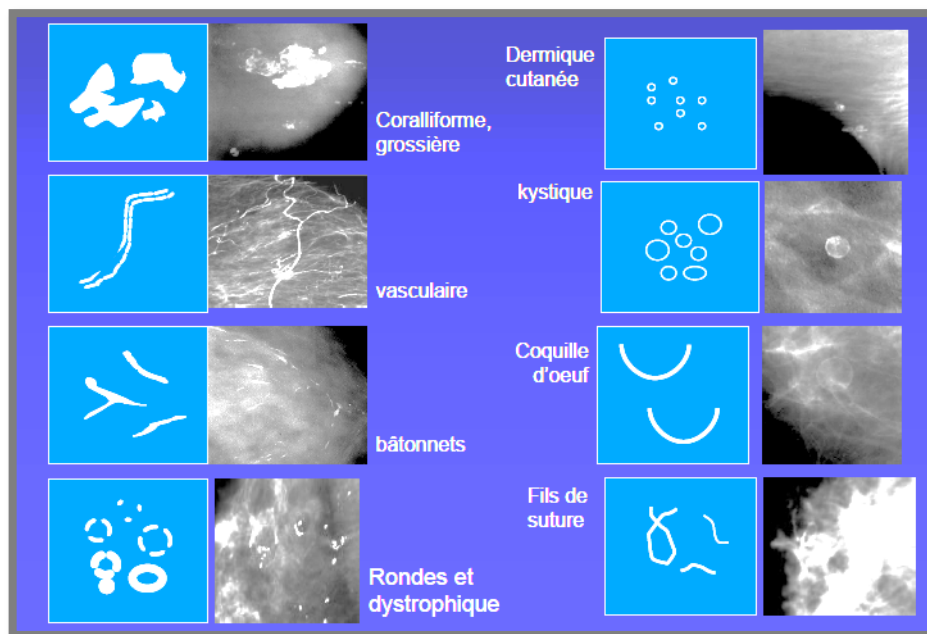


Figure 1. 14: macrocalcifications [15].

1.3.3.2. Analyse de la taille

En règle générale, les microcalcifications sont associées avec un processus malin et, les macrocalcifications avec un processus bénin. Les calcifications sont d'abord peu visibles puis augmentent de taille. On peut observer des calcifications, dont la taille est d'environ **0.2 à 0.3 mm**. Les calcifications de moins de **0.5 mm**, sont souvent associées à des cancers, celles de plus de **2 mm** sont assez typiques, de lésions bénignes [13].

1.3.3.3. Analyse du nombre

Les petites calcifications nombreuses, sont suspectes. La plupart des radiologues considèrent, qu'au-dessus de 4 à 6 (micro-) calcifications, il convient d'être prudent dans l'interprétation [13].

1.3.3.4. Analyse de la morphologie et de la texture

L'analyse de la morphologie, apporte beaucoup de renseignements et, permet le plus souvent de séparer microcalcifications bénignes et, malignes. Les calcifications arrondies ou ovales, uniformes dans leur taille et leur forme, sont probablement bénignes. A l'inverse, celles qui sont irrégulières dans leur taille, ressemblant par leur hétérogénéité à des '**débris de verre**', sont probablement malignes [13]. L'analyse de la texture permet d'évaluer la diffusion et la rugosité des lésions mammaires.

1.4 Conclusion

Beaucoup de lésions cancéreuses, ont un faible contraste et, souvent elles peuvent être confondues avec les tissus de la glande mammaire, par conséquent, l'interprétation des clichés mammographiques, devient une tâche très difficile qui demande une très grande habileté et, une forte concentration. De ce point, notre projet est né, afin d'apporter une aide automatique aux médecins radiologues dans leur diagnostic.

Le chapitre suivant, va parcourir les techniques employées, pour la segmentation des images mammographiques, afin de concevoir l'approche adoptée dans le cadre, de notre projet.

Bibliographie

- [1] algerie360. (2012, Octobre) algerie360. [Online]. <http://www.algerie360.com/algerie/cancer-du-sein-en-algerie-7500-nouveaux-cas-enregistres-par-an/>
- [2] Hector PEREZ-PONCE, "Corrélation entre les performances physiques mesurées des détecteurs et la qualité diagnostique de l'image en mammographie numérique, THESE," Rayonnement et Imagerie Médicale, Institut National Polytechnique de Lorraine, Nancy, Thèse de doctorat 2009.
- [3] Fédération nationale des centres de lutte contre le cancer (FNCLCC), "Comprendre le cancer du sein, Article," *SOR SAVOIR PATIENT*, 2007.
- [4] VIDAL. (2011, Novembre) EurekaSante. [Online]. <http://www.eurekasante.fr/maladies/cancers/cancer-sein.html>
- [5] Marc Pachoud, "Comparaison objective de systèmes d'imagerie conventionnelle et numérique en mammographie," Faculté des Sciences de l'Université de Lausanne, LAUSANNE, Thèse de doctorat 2003.
- [6] Fédération Nationale des Centres de Lutte Contre le Cancer (FNCLCC), "Comprendre. la mammographie," *SOR SAVOIR*, pp. 1-6, 2003.
- [7] ARCAGY - GINECO. (2011, octobre) InfoCancer. [Online]. <http://www.arcagy.org/infocancer/localisations/cancers-feminins/cancer-du-sein/detection-et-depistage/le-depistage/la-mammographie-de-depistage.html>
- [8] RIM Radiologie Imagerie Médicale. (2011, Avril) RIM Radiologie Imagerie Médicale. [Online]. <http://www.rim-radiologie.fr/mammographie-depistage.php>
- [9] A Noel, P Heid J Stines, "La mammographie numérique : paramètres utiles à l'évaluation des systèmes et offre industrielle en France," *Journal de Radiologie*, 2007.
- [10] Hailey D, Warburton R, MacGregor JH, Pisano ED, Joyce J Ho C, La mammographie numérique comparativement à la mammographie sur film avec écran : évaluation technique, clinique et économique, 2002.
- [11] S. K. Moore, *Better breast cancer detection.*: Ieee Spectrum, 2001, vol. 38.
- [12] L. BELKHODJA and N. BENAMRANE, Approche d'Extraction de la Région Globale d'Intérêt et Suppression des Artefacts Radiopaques dans une Image Mammographique, MEMOIRE.
- [13] Pr J.F. HERON. (2009, Juin) oncoprof.net. [Online]. http://www.oncoprof.net/Generale2000/g04_Diagnostic/Mammographie/g04

[mm04.html](#)

- [14] Ingrid Schreer Sylvia H. Heywang-Köbrunner, *Imagerie diagnostique du sein: mammographie, échographie, IRM, techniques interventionnelles*, 2nd ed., MASSON, Ed., 2007.
- [15] S MEHIRI, A SALEM, L KRIBI, and R HAMZA, CLASSIFICATION BI RADS de l'ACR, 2003.
- [16] Brigitte Séradour, *Le dépistage du cancer du sein : Un enjeu de santé publique*, 2nd ed., Springer Verlag France, Ed., 2003.
- [17] Dr. Marie-Hélène Dilhuydy. (2001, Mar.) europadonna. [Online]. www.europadonna.fr

Listes des acronymes et abréviations

AR: Autoregressive Models

BI-RADS: Breast Imaging Reporting and Data System

BTH: Black Top Hat

COR : correlation

CST : contraste

ENE: énergie

ENT : entropie

ERLM : écran radio-luminescent à mémoire

GUI : Graphical User Interface

IDM : Inverse Difference Moment

MFE : mammographie sur film avec écran

MN : mammographie numérique

ROI : Region of Interest

RX: rayon X

VAR: variance

WTH: White Top Hat

Introduction générale

La médecine, est le domaine le plus important dans la vie quotidienne de l'être humain. En effet, durant les dernières décennies, l'apparition des types de cancer est de plus en plus croissante dans tout le corps de l'être humain. L'imagerie médicale constitue aujourd'hui un outil extraordinaire d'aide au diagnostic, pour le médecin. Les modalités d'acquisition de l'information « **image** » sont nombreuses et la quantité d'images manipulées s'accroît. La radiographie par **rayons X** constitue, la part la plus importante du marché, l'imagerie ultrasonore, rapide, non invasive et peu onéreuse, gagne du terrain. Cependant, l'Imagerie par Résonance Magnétique (**IRM**), la tomographie de même que la radiologie **RX**, ont un avantage sur l'imagerie ultrasonore pour la qualité des images, en termes de résolution.

Dans ce contexte, l'imagerie médicale joue un rôle important aux niveaux diagnostic et thérapeutique pour tout type de cancer et, plus particulièrement pour le cancer du sein chez les femmes. En raison de la délicatesse du cancer de sein, un grand nombre de groupes de recherche travaille, sur le problème d'aide au diagnostic. Les premiers travaux datent de **1967**, mais la plupart des publications datent des deux dernières décennies. Depuis cette période, les nouveaux travaux ne cessent d'être publiés et, le besoin de la recherche dans ce domaine, ne cesse de s'accroître.

Notre travail s'inscrit dans le cadre de la détection et d'analyse des tumeurs, ayant comme objectif principal la détection des masses, qui sont des signes majeurs du cancer du sein. Certes, le problème de détection en traitement d'images est un problème classique, mais il est d'autant plus difficile et délicat, lorsque les données à traiter, sont de nature médicale.

Par rapport à la vision actuelle des traitements de détection des modèles, nous avons abouti à des systèmes d'aide à la détection "CAD" (Computer Aided Detection) et des systèmes d'aide au diagnostic "CADx" (Computer Aided Diagnosis) précis, en terme de détection et performant, en terme temps du calcul. Les perspectives offertes par un système de vision

artificielle permettant une analyse automatique des clichés mammographiques, sont prometteuses et peuvent être envisagées sous plusieurs angles: en tant qu'aide au diagnostic, en tant que lecteur potentiel d'images contribuant à valider ou à infirmer le diagnostic d'un expert radiologue, etc. L'analyse automatique des images mammographiques est, en plein essor; du fait du développement des techniques de mammographie numérique plein champ.

En effet, en nous basant sur les propriétés de malignité et de bénignité des tumeurs du sein, nous nous sommes intéressées à la complémentarité entre la segmentation et l'analyse, afin d'aboutir à des méthodes de détection et d'analyse des tumeurs du sein. Ces dernières touchent à deux indices de bas niveau à savoir la texture et la forme.

Dans ce cadre, l'objectif de notre mémoire, est de réaliser, un système d'aide à la décision, qui se base sur trois étapes principales: la première est la détection des masses par une segmentation en régions ; la seconde étape permet la caractérisation des régions d'intérêt suivant, la texture et la morphologie et la troisième étape, classe les tumeurs suivant les machines à support de vecteurs.

Plan du mémoire

Afin de mener à bien, notre prélude à la recherche, nous structurons notre mémoire selon quatre chapitres :

- ✂ **Le premier chapitre**, est consacré au contexte mammographique;
- ✂ **Le second chapitre**, parcourt les techniques de segmentation en imagerie mammographique;
- ✂ **Le troisième chapitre**, permet l'étude des descripteurs texturaux et morphologiques ainsi que la classification par les machines à support de vecteurs « **SVM** » ;
- ✂ **Le quatrième chapitre**, appréhende la méthodologie adoptée, son application sur des données mammographiques réelles ainsi, qu'une interprétation des résultats.

Conclusion générale

La mammographie est un domaine en évolution croissante, depuis plusieurs décennies. Ainsi, les systèmes analogiques ont bénéficié d'avancées technologies concernant les tubes à rayon X et les films, améliorant significativement la détection de lésions dans le sein. Plus récemment, les systèmes numériques ont été introduits permettant d'améliorer les performances cliniques, de réduire la durée de la procédure médicale et, ainsi d'améliorer la précision du geste et le confort de la patiente. Cependant, certaines limitations de ce type d'imagerie, sont toujours d'actualité, comme par exemple la superposition de tissus pouvant potentiellement cacher des lésions. En effet, certaines études montrent, que 76% des lésions qui sont manquées par les radiologues se trouvent dans des seins denses. Ainsi, la superposition des tissus serait la principale cause de la réduction de visibilité, des signes radiologiques. La détection automatique de cancers, permet de fournir au radiologue des outils très utiles, dans sa tâche de dépistage du cancer du sein, ou d'une analyse concise de tumeurs prononcées. Parmi les signes qu'il est intéressant de détecter, on peut citer les microcalcifications et les masses.

Les systèmes d'aide au diagnostic (Computer Aided diagnosis ou CAD) sont utilisées, pour identifier les régions suspectes d'une mammographie, dont les caractéristiques physiques ou perceptuelles sont similaires à celles d'un cancer.

Le travail abordé dans ce mémoire, appréhende l'image mammaire en se consacrant à la détection des masses maligne et bénignes. Dans un premier temps, les images mammographiques originales, ont été préparées en utilisant les fonctions de la morphologie mathématique pour faire un rehaussement de contraste afin, de faciliter la phase de segmentation, pour l'extraction des masses à partir d'images mammographiques.

La détection des masses s'est faite par une technique de segmentation par croissance de régions. L'approche proposée pour segmenter l'image consiste, à faire croître chaque région

autour d'un pixel de départ. Nous avons pris comme germe le centre de la tumeur désignée par le médecin radiologue.

Une caractérisation texturale des masses a été dans un second lieu abordé, suivant l'extraction de paramètres d'Haralick calculés à partir de la matrice de cooccurrence, ainsi que selon des descripteurs morphologiques. Cette description nous permet alors, de caractériser chaque type de pathologie. Les facteurs texturaux extraits, fournissent des mesures quantitatives de la granularité des masses et, les facteurs morphologiques, permettent une description de la géométrie. En effet, il a été remarqué, qu'une masse spinale présente en général, une valeur d'entropie plus élevée qu'une masse lisse.

Dans la dernière phase, les régions d'intérêts segmentées, sont classifiées en tissu malin et bénin via les machines à support de vecteurs.

Notre choix s'est porté sur les machines à vecteurs de support (SVM) car elles limitent le risque de sur apprentissage du fait de leur capacité de régularisation. Ce risque étant particulièrement important, lorsque le nombre de caractéristiques, c'est-à-dire la dimension, est grande face au nombre de données traitées.

Les résultats obtenus sont certes satisfaisants, mais des améliorations sont nécessaires pour l'enrichir. L'étude d'une méthode de segmentation coopérative, serait intéressante pour mieux délimiter la forme de la lésion dans les images mammographiques. Nous pouvons extraire d'autres types de caractéristiques texturales et morphologiques de l'image, en vue de caractériser davantage de pathologies. L'entreprise d'une étude de classifieurs connexionnistes ainsi que les multiSVM, serait envisageable pour une comparaison avec l'approche des SVM.

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