

Conclusion générale et perspectives

Malgré les évolutions des nouvelles technologies pour améliorer la qualité des images mammographiques numériques, l'interprétation et l'analyse de ces images restent une tâche difficile même pour un radiologue étant donné que ces images sont fortement texturées. D'où le besoin de développer des systèmes d'analyse et d'interprétation automatique des images mammographiques numériques, comme outil d'aide à la décision, augmentant la fiabilité de l'interprétation par un expert.

En analyse d'images médicales, la segmentation est sans aucun doute la tâche qui mobilise le plus d'efforts. Elle nécessite des connaissances a priori sur le type et la qualité d'images à traiter et, la structure à segmenter.

Dans ce mémoire nous avons étudié et développé un système de détection et de classification neuro-génétique de structures pathologiques mammographiques établies par les masses et les calcifications, en vue d'une aide à la décision.

Dans un premier temps, les images mammographiques originales ont été prétraitées, afin de rehausser le contraste. Cette opération fondée sur un ensemble d'analyse d'histogramme et d'opérations morphologiques, contribue en particulier à l'amélioration des résultats ultérieurs de la segmentation des lésions mammaires. Par la suite, Nous avons prouvé la faisabilité de la segmentation par le modèle de Chan et Vese, des masses et des calcifications mammaires. La troisième partie de ce mémoire, a montré les possibilités offertes par les caractéristiques de la texture et de la forme, pour la caractérisation des lésions. La dernière partie de notre système, est la phase de classification des anomalies mammaires, suivant le modèle du perceptron multicouche, en cascade avec les algorithmes génétiques. Notre classifieur NG, nous a permis d'obtenir une base d'entraînement comportant des exemples de chaque type avec des performances intéressantes. Les résultats de la classification sont encourageants. Chacun des deux réseaux réalisés, a pu classifier les deux types de classes (masses et calcifications) avec un taux de reconnaissance de 100%.

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

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

1.1 INTRODUCTION

Le cancer du sein est un problème de santé publique, il touche environ une femme sur 9 à travers le monde [1]. La mortalité par le cancer du sein, reste élevée, ce qui fait de cette maladie, l'une des principales causes de décès chez la femme dans le monde.

L'imagerie médicale est certainement l'un des domaines de la médecine, qui a connu une véritable révolution pendant ces vingt dernières années. Ces récentes découvertes permettent non seulement un meilleur diagnostic, mais offrent aussi de nouveaux espoirs de traitement pour de nombreuses maladies (comme le cancer du sein). La mammographie est la plus performante pour la détection des cancers du sein, en matière de surveillance et de dépistage précoce du cancer du sein. Elle permet de mettre en évidence les anomalies mammaires, telles que les masses et, les calcifications qui peuvent traduire des lésions malignes. Cependant, tous les radiologues reconnaissent la difficulté de l'interprétation des mammographies, qui s'accroît selon, le type du tissu examiné.

Dans ce chapitre, nous présentons un aperçu général sur l'imagerie mammographique, ses modalités quelques notions sur l'anatomie du sein, pour aborder enfin les pathologies mammaires et, situer ainsi, le cadre de notre application.

1.2 Anatomie du sein

Le sein est une glande exocrine qui se développe, au cours de la vie de la femme. C'est un organe pair et symétrique de forme hémisphérique, situé en avant du thorax, entre la troisième et la cinquième côte, au-dessus du muscle pectoral (figure 1.1).

La plus grande partie du sein est constituée de tissu adipeux qui comble l'espace situé, entre les différentes structures du sein [2].

La glande mammaire est constituée de 2 compartiments cellulaires: le compartiment mésenchymateux (les vaisseaux sanguins et les nerfs) et le compartiment épithélial qui s'articule autour d'un réseau de canaux galactophores et, de lobules renfermant les alvéoles. Ces deux compartiments sont séparés par une membrane basale de collagène, de laminine et de glycosaminoglycanes. Cette architecture se construit tout au long de la vie, du stade foetal à la ménopause, sous l'influence d'hormones (œstrogènes et progestérone) et d'un certain nombre de facteurs de croissance ainsi que, d'une coopération permanente qui s'établit entre les deux compartiments cellulaires [2] [3].

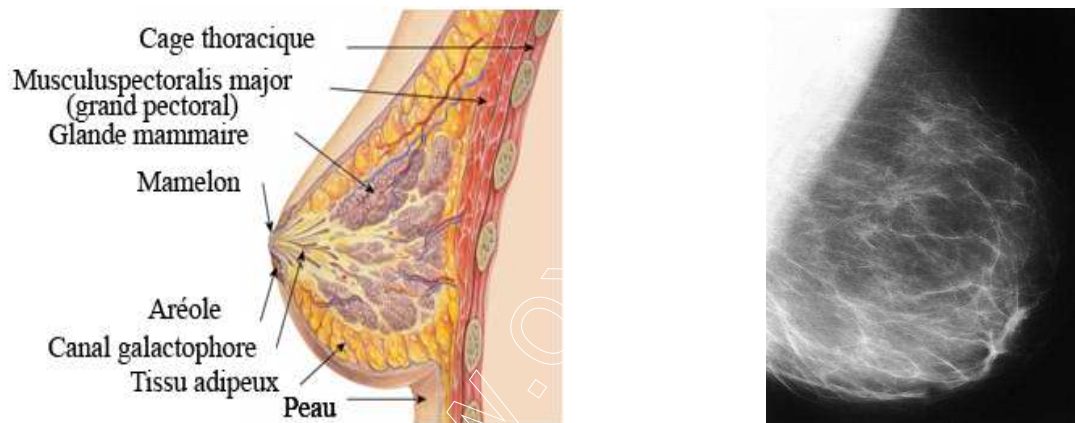


Figure 1.1 : anatomie de la glande mammaire [2][3]

1.3 Les tumeurs [4][5]

Au cours de la vie, la plupart des femmes connaissent plusieurs changements au niveau de leurs seins. Certaines modifications sont normales, d'autres peuvent faire peur. La pathologie, est détectée par l'examen du tissu au microscope qui permet de savoir s'il est cancéreux ou non et, de déterminer le type de la tumeur qu'il forme. Une masse au sein peut être d'origine bénigne (kyste, fibroadénome, maladie fibrokystique, abcès, galactocèle) ou d'origine maligne.

La probabilité de malignité augmente avec l'âge, plus particulièrement à partir de 50 ans.

La découverte d'une masse dans un sein n'est jamais une partie de plaisir, mais près de 80 % des tumeurs du sein sont bénignes, et certains symptômes, même s'ils peuvent sembler affolants, ne riment pas toujours avec cancer.

1.3.1 Les tumeurs bénignes

Les tumeurs bénignes ont des contours bien délimités. Elles ont une croissance lente et restent localisées, dans le tissu ou l'organe dans lequel elles sont apparues. Elles n'entraînent pas de métastases dans d'autres endroits du corps. Les tumeurs bénignes sont composées de cellules, qui ressemblent aux cellules normales du tissu concerné [4].

L'exploration devra permettre de classer les tumeurs bénignes en trois groupes (hormis les tumeurs néoplasiques du sein) :

- lésions bénignes sans risque carcinologique.
- lésions bénignes avec risque carcinologique minime.
- lésions épithéliales avec risque carcinologique.

a. Les lésions bénignes sans risque carcinologique

- Le fibroadénome : c'est la forme la plus fréquemment rencontrée. Cliniquement, il s'agit d'une masse ferme, mobile, caoutchouteuse, bien limitée. En imagerie, cette masse est ovalaire, horizontale, régulière, homogène. La biopsie, lorsqu'elle est réalisée, confirme aisément le diagnostic. Il n'y a pas d'indication opératoire systématique sauf si le nodule fait plus de 4 cm, augmente de taille, déforme le sein, ou s'il y a une demande de la patiente ou si la patiente a plus de 40 ans. Le traitement de référence est la tumorectomie.
- L'hamartome : cliniquement, il s'agit d'une masse assez molle. L'imagerie est pathognomonique, en particulier en échographie qui met en évidence une lésion limitée par une capsule. La biopsie n'est pas indiquée dans ce cas, car peu contributive. Il n'y a pas d'indication opératoire systématique.

- Les états fibrokystiques : il s'agit de formations nodulaires associant en proportion variable fibrose, kystes, adénose, hyperplasie épithéliale simple. Il y a fréquemment des calcifications rétionnelles (sédiment). Le nodule s'accompagne de symptômes à type de douleurs prémenstruelles en période d'activité génitale. La palpation peut retrouver soit un nodule à limites nettes et arrondies, mobile, parfois rénitent, soit des placards douloureux confluents éventuellement en masse pseudo-tumorale. Il n'y a pas a priori d'indication opératoire.
- Le kyste mammaire : le diagnostic de kyste simple est effectué le plus souvent facilement par l'examen clinique et par l'échographie (nodule strictement liquidien). Il ne nécessite pas d'autre exploration, s'il s'agit d'une découverte incidente. En cas de symptomatologie douloureuse, une ponction est réalisée. L'existence d'un kyste complexe ou d'une masse solide résiduelle après ponction nécessite une triple exploration (clinique, échographique et biopsique) qui précède une intervention d'exérèse.
- Le lipome : cliniquement, il s'agit d'une masse molle. L'imagerie retrouve une opacité lobulée, radio transparente. Il n'y a pas d'indication opératoire.
- Le nodule de cytotéatonecrose : il est le plus souvent secondaire à un traumatisme (60 % des cas). L'échographie est souvent caractéristique et, en cas de ponction, le contenu est huileux et il n'y a pas, d'indication opératoire systématique.

b. Les lésions bénignes avec risque cancérologique minime

Ce groupe est essentiellement composé par des pathologies, de type adénose microglandulaire, adénose sclérosante complexe et cicatrice radiaire. Ces lésions souvent nodulaires, ont une imagerie différente des précédentes, en particulier en mammographie; où l'on remarque des lésions stellaires avec centre clair (scléreux hyalin). Lorsque la taille de ces lésions fait plus de 1 cm ou qu'il y a plusieurs foyers, on parle d'adénose complexe. Il s'agit d'une image radiologique difficile à surveiller avec un risque de cancer non-invasif ou invasif d'environ 17 %. Une histologie préopératoire est toujours nécessaire et, ce diagnostic doit déboucher sur une exérèse chirurgicale simple.

c. Les lésions bénignes avec risque carcinologique

Ces lésions, sont diagnostiquées par le prélèvement biopsique qui permet d'adapter la conduite à tenir, le plus souvent chirurgicale. Le prélèvement histologique des nodules retrouve des atypies cellulaires avec des risques de transformations malignes majorées. Ces lésions sont classées en 4 groupes: l'hyperplasie canalaire atypique, la métaplasie cellulaire atypique et l'hyperplasie lobulaire atypique (HLA) ainsi que la néoplasie intralobulaire (LIN).

1.3.2 Les tumeurs malignes

Une tumeur maligne est une maladie caractérisée par une prolifération cellulaire (tumeur) anormalement importante au sein d'un tissu normal de l'organisme, de telle manière que la survie de ce dernier est menacée. Ces cellules dérivent toutes d'un même clone, cellule initiatrice du cancer qui a acquis certaines caractéristiques lui permettant, de se diviser indéfiniment. Au cours de l'évolution de la maladie, certaines cellules peuvent migrer de leur lieu de production et former des métastases. Pour ces deux raisons, le dépistage du cancer doit être le plus précoce possible.

a. Types du cancer du sein

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

- **Cancer du sein non-invasif**: c'est le type le plus fréquent du cancer du sein non-invasif chez la femme. C'est une forme de cancer très précoce. Comme son nom l'indique, cette pathologie, se développe à l'intérieur des canaux de lactation du sein. Dans la majorité des cas, il est non-palpable. Parfois, le cancer canalaire in-situ se situe dans une minime portion du sein. Ce type de cancer ne se dissémine pas aux ganglions ou à d'autres organes. Il n'a pas le potentiel de faire des métastases. Il est avantageux de le diagnostiquer à ce stade, car ce cancer est guérissable dans plus de 98 % des cas. Par conséquent, on peut être très optimiste et il est primordial d'effectuer un bon traitement. S'il n'est pas traité, le cancer « in

situ » poursuit son développement et peut devenir un cancer « infiltrant » qui peut se propager à l'extérieur des canaux et former une masse.

- Cancer du sein invasif : cette catégorie de cancer est aussi appelée cancer infiltrant. Ce terme signifie que les cellules cancéreuses ont traversé, la paroi du canal de lactation. Le cancer est souvent palpable et a la possibilité, de se propager, aux ganglions de l'aisselle. Dans la plupart des cas, le cancer invasif est encore bien guérissable. Cependant, si on néglige de le traiter il pourrait se propager dans diverses parties du corps (métastases).

Il existe plusieurs types du cancer du sein invasif, à savoir le carcinome lobulaire infiltrant et le carcinome canalaire infiltrant. Ce dernier type peut être sans particularité aucune ou encore, démontrer des caractéristiques particulières qui sont souvent associées à un meilleur pronostic: le carcinome mucineux (ou colloïde), tubulaire, médullaire et, papillaire.

Une forme distincte du cancer du sein invasif, est le cancer du sein inflammatoire qui se caractérise principalement par un sein qui peut devenir rouge, enflé et chaud. La peau du sein peut aussi prendre l'aspect d'une orange. Il passe souvent pour une banale mastite ou une infection du sein. Cependant, cette « infection » ne se guérit pas avec des antibiotiques. Une autre forme particulière du cancer du sein, se nomme la maladie de Paget qui se caractérise par une petite plaie qui, ne guérit pas au niveau du mamelon.

b. Signes Cliniques et Symptômes

Le cancer du sein ne donne pas de signe, au début de la maladie. Quand la tumeur est suffisamment grosse, elle est palpable au niveau du sein ou de l'aisselle. C'est à ce moment-là que la patiente consulte son médecin, après avoir découvert un nodule mammaire. À l'examen, le nodule est dur, irrégulier, indolent, peu mobile sur les plans voisins. Parfois, en regard de la tumeur, la peau change d'aspect: peau d'orange, rétractée, ridée ou d'allure inflammatoire (rouge et chaude). Le mamelon peut aussi être déformé, être le siège d'un eczéma ou d'un écoulement de sang, avoir un aspect croûteux. Il est aussi possible de découvrir des ganglions axillaires et sus-claviculaires. En clinique, on examine le côté opposé (13% des patientes ont des foyers controlatéraux).

1.4 Facteurs de risques [4][6]

De nombreuses maladies ont une cause unique et identifiable. Au contraire, les cancers, sont des maladies « multifactorielles », les causes étant multiples et inconnues, mais les travaux scientifiques ont pu identifier, certains facteurs de risques pour développer un cancer du sein.

1.4.1 Les facteurs de risque non modifiables

- Le sexe et l'âge: le cancer du sein est plus courant chez les femmes que les hommes car les cellules des canaux mammaires des femmes, sont plus développées que les cellules des hommes. En effet, les cellules des seins de femmes sont continuellement exposées à l'hormone femelle œstrogène, hormone essentielle à la bonne croissance, au bon développement et à la vie reproductive normale de la femme, mais qui constitue également un facteur de risque du cancer du sein. Le risque de cancer du sein augmente significativement avec l'âge, mais les femmes de tout âge, peuvent présenter un cancer du sein.
- Les antécédents personnels du cancer: le fait d'avoir un antécédent personnel du cancer, est considéré comme un facteur de risque de récurrence. Le cancer du sein peut réapparaître tôt après que le premier cancer ait été diagnostiqué et traité ou; peut survenir après de nombreuses années. Le cancer peut réapparaître dans le même sein du cancer d'origine, dans l'autre sein ou dans une autre partie du corps.
- Les antécédents familiaux du cancer et les facteurs génétiques: de nombreuses personnes croient que si le cancer est fréquent dans leur famille, ceci constitue la cause la plus courante du cancer du sein. Les recherches indiquent que seulement 5 à 10 pour cent des cancers du sein sont, de nature héréditaire.
- La menstruation précoce et la ménopause tardive: l'âge à laquelle une femme commence et termine ses menstruations, peut influencer sur son risque de présenter un cancer du sein. Si les premières menstruations ont commencé tôt « avant 12 ans » ou si la ménopause s'est manifestée, à un âge tardif « après 55 ans ».
- La densité des seins: le sein est constitué de différents types de tissus: le tissu glandulaire, le graisseux et le conjonctif. La densité des seins décrit la quantité

relative, des différents tissus dans le sein. Les seins denses présentent moins de tissu graisseux et plus de tissu conjonctif et, glandulaire. Le tissu mammaire devient généralement moins dense avec l'âge, bien que certaines femmes aient toujours des seins denses sans égard à leur âge. La recherche indique que les femmes qui ont des seins denses présentent un risque accru du cancer du sein.

- Les affections des seins : il existe un certain nombre d'affections bénignes (non cancéreuses) du sein. Certaines de ces affections sont associées, à un risque accru du cancer du sein. Une affection proliférative (les cellules de seins grossissent et se reproduisent plus rapidement, que la normale) est toujours bénigne, mais elle indique qu'il y a une augmentation modérée de risque, d'avoir un cancer du sein.

1.4.2 Les facteurs de risque modifiables

- Le poids corporel : la recherche montre que les femmes qui ont un surplus du poids, peuvent présenter un risque plus élevé du cancer du sein. Ce risque augmente si la prise du poids, se fait après la ménopause.
- La consommation d'alcool et le tabagisme: la consommation d'alcool et Le tabagisme sont des facteurs de risque majeurs de maladies chroniques, incluant le cancer.
- L'exposition aux hormones synthétiques: les hormones synthétiques «œstrogènes et progestérone », peuvent également accroître le risque pour une femme, de présenter un cancer du sein.
- L'exposition aux rayonnements dans le cadre médical: les recherches ont prouvé que l'exposition à des niveaux élevés de rayonnement, est liée à la formation de différents types du cancer, incluant le cancer du sein.

1.5 Outils de dépistage du cancer du sein

L'imagerie médicale a révolutionné la médecine et devient incontournable dans de nombreuses situations. Elle permet non seulement d'établir un meilleur diagnostic, mais aussi d'évaluer la sévérité d'une pathologie et tester, l'efficacité d'un traitement. Il existe, à

l'heure actuelle, un certain nombre de techniques d'imagerie du corps humain couramment employées dans le domaine médical. Chacune d'elles est sensible à un type de contraste particulier et trouve ses applications pour des organes différents. Plusieurs techniques peuvent, également, apporter des informations complémentaires sur un même organe. Les outils de l'imagerie médicale utilisés pour le dépistage ainsi que le diagnostic du cancer du sein sont: l'échographie (imagerie par ultrasons), l'IRM (Imagerie par Résonance Magnétique) et la mammographie (imagerie par rayons X) [7] [8].

1.5.1. La palpation

La palpation du sein est un examen pratiqué chaque année par le médecin, ou le gynécologue. Il s'adresse à toutes les femmes, à partir de 35 ans. En palpant les seins, le médecin peut identifier une grosseur inhabituelle [4].

1.5.2. L'échographie

L'échographie, introduite en médecine dans les années 1950, utilise les ultrasons non irradiants, ondes sonores imperceptibles à l'oreille humaine. Quand une sonde émet des ultrasons en direction d'un objet solide, ceux-ci rebondissent sur l'objet et reviennent, au point de départ. L'échographie mesure le temps nécessaire pour ce trajet (écho) et restitue une image, en temps réel permettant de distinguer les différentes structures [7].

La figure 1.2 présente deux exemples d'images échographiques du sein: une lésion maligne du sein droit (a) et une lésion bénigne du sein gauche, de la même patiente (b).

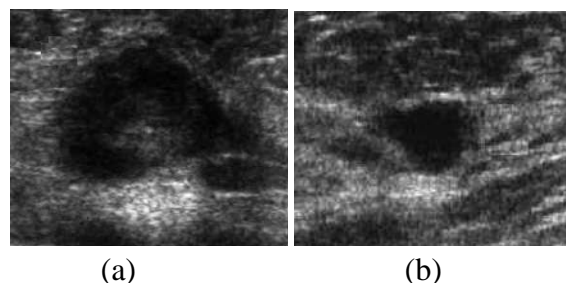
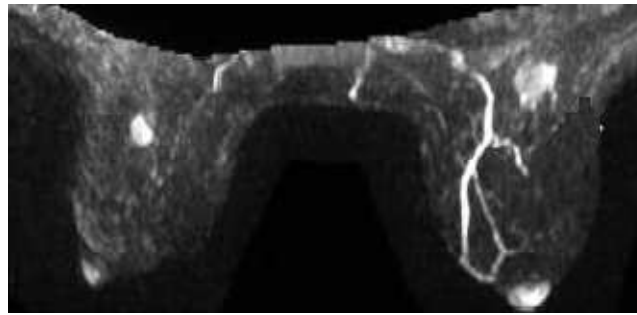


Figure 1.2 : exemples d'images échographiques du sein [7]

L'échographie est un examen complémentaire à celui de la mammographie, non utilisée en première intention pour le dépistage du cancer du sein mais seulement, en cas d'anomalie mammographique. L'échographie ne remplace pas la mammographie car il ne permet pas de détecter les microcalcifications et, même les petites masses, très importantes en dépistage. De plus, les images ultrasonores sont généralement altérées par un bruit spécifique appelé la granularité (reconnu aussi sous le nom de speckle) qui est causé essentiellement par, les interférences entre les ondes. En conséquence, cette méthode d'imagerie médicale n'est pas généralisée aux campagnes de dépistage. Elle est souvent exploitée comme moyen de repérage lors d'une ponction, ou d'une biopsie et aussi comme moyen de repérage préopératoire, pour marquer l'emplacement de la lésion.

1.5.3. L'imagerie par résonance magnétique

L'imagerie par résonance magnétique « I.R.M » est apparue au début des années 1980, du phénomène de la résonance magnétique nucléaire « R.M.N » découvert en 1946 par les physiciens américains Edward Mills Purcell et Felix Bloch [7]. L'IRM mammaire, qui n'est pas un examen de dépistage dans la population générale, a une place essentielle dans l'exploration par imagerie, des femmes à risque génétique. Cette méthode se base sur l'utilisation d'un aimant (constituant le champ magnétique) et d'ondes de radiofréquences. Son principe consiste à faire vibrer de façon imperceptible, les atomes d'hydrogène du corps humain. Placés dans un champ magnétique puissant, tous les atomes d'hydrogène s'orientent vers la même direction. Ils sont excités par des ondes radio, durant une très courte durée. On dit qu'ils sont mis en résonance. A l'arrêt de cette stimulation, les atomes restituent l'énergie accumulée en produisant un signal. Ce dernier est enregistré et traité sous forme d'image, par un système informatique. La figure 1.3 présente l'image IRM des seins gauche et droit de la figure 1.2.



(a)

(b)

« (a) lésion maligne et (b) lésion bénigne »

Figure 1.3 : exemple d'IRM mammaire bi-sein [7]

L'IRM permet une étude dans tous les plans de l'espace et l'obtention d'images, de grande qualité. De plus, elle fournit une image en haute résolution qui permet une étude approfondie de la pathologie. Cependant, cette technique est très coûteuse et appliquée uniquement aux patients, non porteurs de prothèses métalliques. Ainsi, son utilisation est restreinte pour apporter des précisions, concernant une anomalie déjà détectée en mammographie ou en échographie [7].

1.4.4 La mammographie

a. La radiographie

La radiologie est inventée en 1895 par le physicien allemand Wilhelm Conrad Röntgen [9], auteur des premières radiographies d'intérêt médical. Son principe repose sur l'utilisation des rayons X. Ces derniers traversent le corps humain, mais sont plus ou moins absorbés par les tissus en fonction de leur densité. Ils impriment un film photographique placé à l'opposé du patient telle une pellicule photo. Il est possible d'opacifier certaines structures creuses (appareil digestif, articulation, etc.) en injectant un produit de contraste, opaque aux rayons X. Pour réduire les doses de rayons émises lors des examens, les films radiographiques sont aujourd'hui, souvent remplacés par des détecteurs électroniques, qui numérisent directement les images et sont plus sensibles que les plaques. La radiographie est fréquemment utilisée en orthopédie, en rhumatologie et en orthodontie pour, étudier les traumatismes osseux, les déformations du squelette ou les implantations dentaires. Elle permet également d'observer des anomalies sur certains organes comme des infections

bactériennes ou virales ou encore, des tumeurs au niveau des poumons ou des seins (mammographie).

b. Principe de la mammographie [9][10][11]

Une mammographie (ou mastographie) est un examen radiologique du sein. Il permet d'obtenir des images de l'intérieur du sein, à l'aide des rayons X pour, détecter ainsi, certaines anomalies. La quantité de rayons X 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 rayons X reçus, même si des mammographies de dépistage sont répétées. Une mammographie est pratiquée dans deux circonstances: dans le cadre d'un dépistage ou d'un diagnostic.

- **Dépistage:** les femmes se soumettent de façon systématique à une mammographie, en vue du dépistage précoce du cancer du sein. On cherche des signes du développement d'un cancer du sein, même en l'absence de symptômes. La figure 1.4 indique un des résultats possibles.

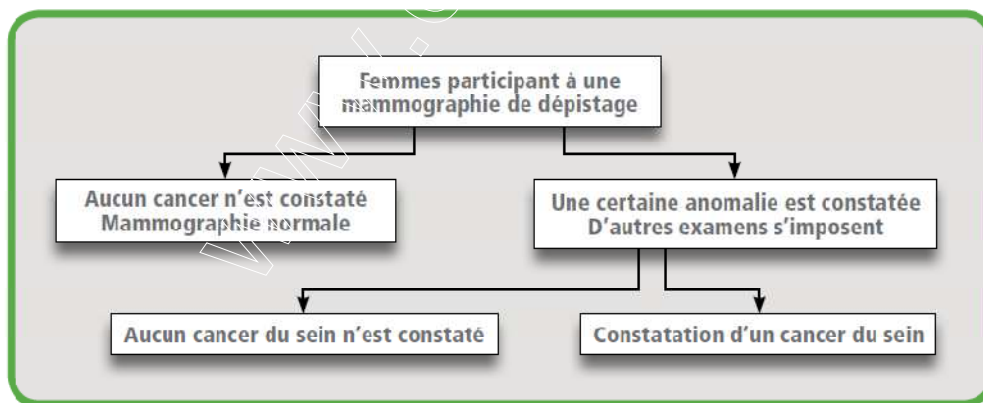


Figure 1.4: principe de la mammographie de dépistage [11]

- **Diagnostic:** il se pratique généralement pour vérifier la présence d'un cancer du sein, après avoir constaté une grosseur ou tout autre signe ou symptôme comme de la douleur, l'épaississement de la peau, un écoulement au mamelon ou la modification de la taille, ou de la forme d'un sein. Il peut

aussi servir d'examen complémentaire, si une mammographie de dépistage révèle une anomalie.

c. Mammographie analogique [8] [9]

La mammographie analogique utilise comme détecteur le couple écran-film. Le film photographique fut le premier détecteur à être utilisé en radiographie, dès la découverte des rayons X. Il est sensible à la lumière et aux rayons X, dans une moindre mesure. Il contient une émulsion contenant des cristaux de d'halogénure d'argent (souvent de bromure d'argent). Ces cristaux, soumis aux photons, se dissocient en ions par effet photolytique, créant ainsi une image latente. C'est un système d'imagerie indirect. L'image latente est transformée en image réelle après plusieurs étapes se déroulant dans l'obscurité ou sous une lumière inactinique. La révélation est réalisée en plongeant le film dans une solution basique, qui réduit les ions argent positifs en argent métallique. La fixation de l'image est obtenue, en plongeant le film dans une solution acide permettant de stopper ces réactions de réduction. Après lavage et rinçage du film pour éliminer les différents réactifs, les zones du film les plus irradiées contiennent l'argent métallique et sont les plus opaques à la lumière. Les zones non-irradiées du film sont transparentes et apparaissent blanches si, on le place sur un négatoscope. Avec l'arrivée des nouveaux détecteurs, cette habitude a été conservée. Ainsi, en radiographie, les images sont présentées de façon à ce que les zones les plus exposées soient noires et les zones les moins exposées soient blanches. Pour améliorer la sensibilité du film aux rayonnements très pénétrants que sont les rayons X ou gamma, il est couplé à des écrans renforçateurs, disposés de part et d'autre du film. Ils sont constitués de sels fluorescents qui convertissent les rayons X en photons lumineux. Le couple écran film est disposé à l'abri de la lumière, dans une cassette qui est placée derrière l'objet à radiographier. Le couple écran-film possède une courbe sensitométrique d'allure sigmoïde ce qui oblige à exposer ce détecteur à une quantité précise de photons (latitude d'exposition) pour obtenir un contraste satisfaisant.

Le couple-écran film, seul détecteur analogique, est resté longtemps une référence en radiographie du fait de son excellente résolution spatiale et de sa bonne sensibilité. Néanmoins, il est de moins en moins utilisé, au profit des systèmes de

détection numériques qui permettent de délivrer des doses moins importantes au patient tout en conservant une qualité d'image suffisante à un diagnostic. La mammographie conventionnelle (ou analogique, ou « classique ») produit une image imprimée sur un film argentique (un cliché).

Les limites du couple écran-film sont [11]:

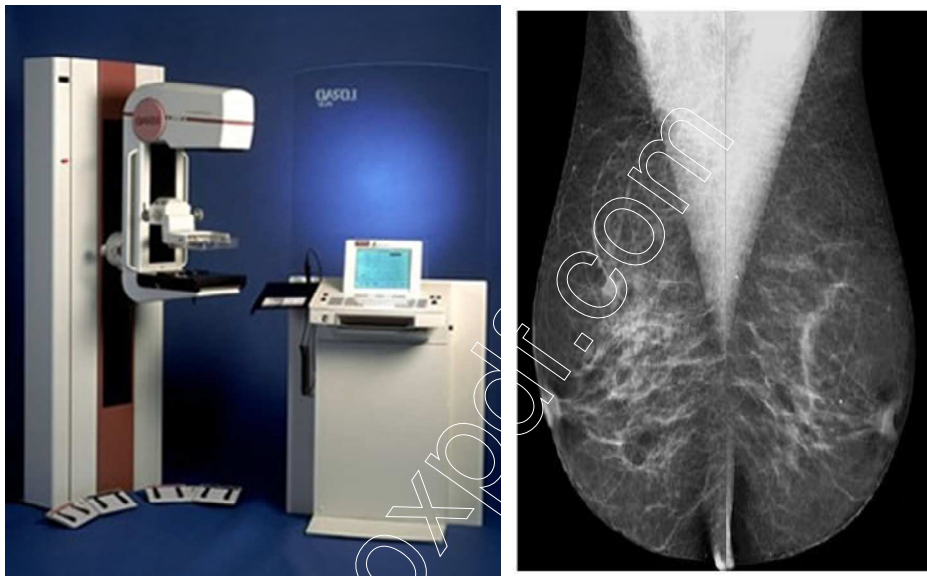
- une faible latitude d'exposition;
- la granularité du film, rend le film plus ou moins sensible, en variant la résolution spatiale. Plus le grain est fin, moins l'émulsion est sensible. Inversement, plus le grain est grand, plus l'émulsion est sensible mais ceci s'accompagne, d'une perte de résolution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

d. Mammographie numérique

En mammographie numérique, on remplace le capteur analogique (couple écran-film) par un capteur numérique. La chaîne de la mammographie numérique, est plus complexe que la chaîne analogique. Elle intègre des éléments électroniques susceptibles, de se détériorer au cours du temps. L'image finale est le résultat d'une acquisition et de divers traitements, avec sa visualisation sur un écran d'ordinateur, ou sur un film imprimé par reprographe laser.

Il y a trois types de détecteurs numériques (les systèmes à balayage, les écrans radioluminescents à mémoire et les capteurs numériques plein champ) qui utilisent un mode indirect de conversion des rayons X en charges électriques avec; une première étape de conversion des photons X en photons lumineux ou un mode direct de conversion des rayons X en charges électriques. Avec ces détecteurs, l'image numérique obtenue est différente, de l'image analogique parce que son contraste, son niveau de bruit et son contenu fréquentiel, sont différents ce qui peut augmenter l'influence du bruit anatomique. De plus, ils peuvent être modifiés après l'acquisition de l'exposition. Pour cette raison, elle doit être lue sur des consoles de diagnostic qui pour des raisons pratiques, doivent avoir au minimum 2 écrans.

La détection radiologique des lésions du sein exige des images de grande précision, qui peuvent être obtenues seulement, avec le mammographe (figure 1.5.a). Le mammographe possède entre autres un dispositif de compression du sein, qui permet à la fois d'obtenir des images de haute résolution et de diminuer les doses de rayonnement utilisées. Le mammographe dispose aussi des accessoires nécessaires, pour réaliser les incidences complémentaires [10].



a : Le mammographe

b : clichés associés

Figure 1.5 : mammographie numérique(MN) et les clichés associés [10]

e. Corrélation entre l'anatomie et les images mammographiques

L'image mammographique est le résultat d'atténuation d'un faisceau de rayons X traversant les différents tissus mammaires. L'atténuation de ce faisceau dépend essentiellement de la composition des tissus traversés. En effet, la graisse est considérée comme une zone radio transparente vu qu'elle a une densité physique très légère. De ce fait, elle apparaît très sombre sur un cliché mammographique. En revanche, les zones radio opaques apparaissent claires et correspondent au tissu fibroglandulaire et au calcium qui est le composant essentiel des lésions mammaires. Pour les matières prédominantes dans le sein, le tableau 1.1 présente la correspondance entre les composants du tissu

mammaire, la radio opacité et l'aspect sur le cliché mammographique. En rassemblant les informations concernant l'anatomie et la radio transparence, on peut confirmer que l'aspect général d'une mammographie est sombre, alors que les zones contenant des microcalcifications ou des masses (composées de calcium) sont plus claires [12].

Composant	Atténuation radiologique	Aspect sur mammographie
graisse	radio transparent	très sombre
eau	légèrement radio opaque	sombre
tissu conjonctif	radio opaque	claire
calcium	très radio opaque	très claire

Tableau 1.1 : atténuation radiologique des composants mammaires [12]

f. Types de la densité mammaire

A la ménopause, les seins sont souvent radio transparents. Les images tumorales s'y inscrivent, avec une netteté très particulière. La classification « BI-RADS » de « l'American College of Radiology » [13] permet de classer les seins en quatre groupes (tableau 1.2) pour lesquels, le risque d'avoir un cancer en raison du caractère hétérogène du tissu fibroglandulaire, augmente lorsque l'on passe de la catégorie 1 à la catégorie 4.

Type	Type de la densité mammaire
1	Le sein est presque entièrement graisseux
2	Il y a des opacités fibroglandulaires dispersées
3	Le tissu mammaire est dense et hétérogène
4	Le tissu mammaire est extrêmement dense

Tableau 1.2 : classification BI-RADS de la densité mammaire [13]

1.6. Les pathologies mammaires [10][14]

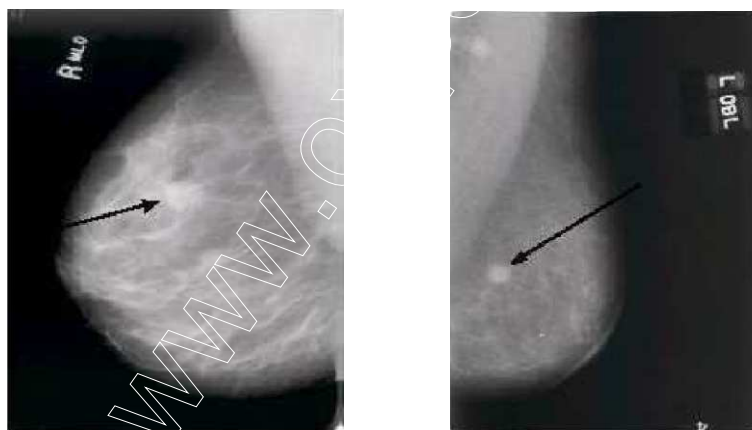
On peut classer les anomalies en trois catégories :

- 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**.

1.6.1 Les masses

Une masse ou une opacité est une anomalie focale de densité visible, sur au moins 2 incidences.

Etant plus denses que le tissu mammaire adjacent, les masses apparaissent blanches sur les mammogrammes (figure 1.6).



a: masse bénigne

b: masse maligne

Figure 1.6: exemple de masses mammographiques [10]

a. L'analyse des opacités mammaires

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

- **La taille:** elle est variable de quelques millimètres, à plusieurs centimètres. Cependant la taille ne prédit pas le caractère malin, sauf sur des clichés successifs lorsque l'on voit la taille augmenter régulièrement.

- **La forme :** les masses mammaires peuvent avoir la forme ronde, ovale, lobulée, irrégulière, ou distorsion architecturale (figure 1.7).
- Ronde: il s'agit d'une masse sphérique, circulaire ou globuleuse.
 - Ovale: elle présente une forme elliptique (ou en forme d'œuf).
 - Lobulée: la forme de la masse présente une légère ondulation.
 - Irrégulière: cette appellation est réservée aux masses, dont la forme est aléatoire.
 - Distorsion architecturale: cette forme inclut des rayons convergeant, vers un point avec une rétraction focale.

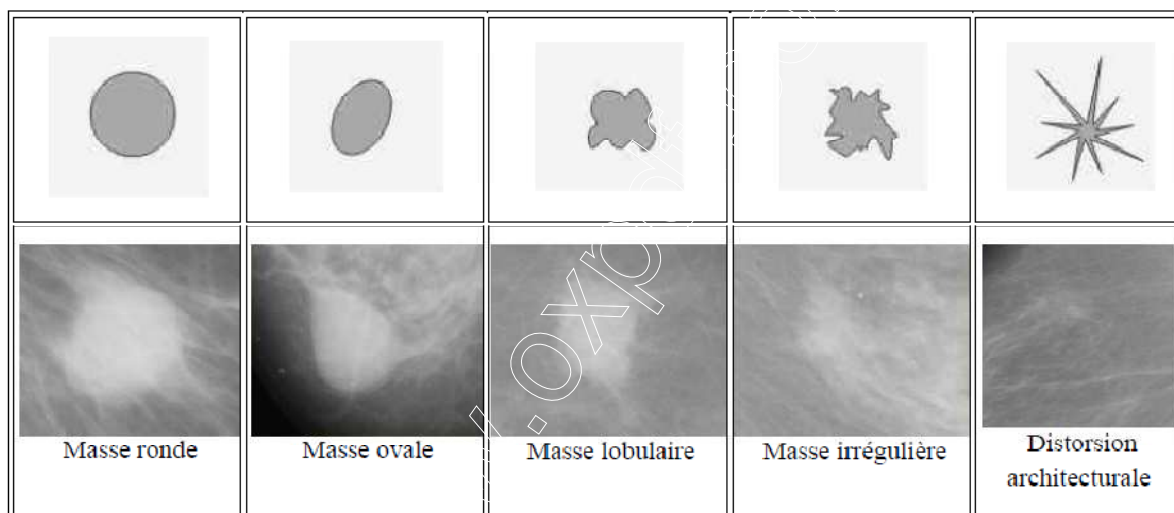


Figure 1.7: formes possibles d'une masse mammaire [10]

- **Le contour :** 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. 8**).
- Circonscrit: il s'agit d'une transition brusque entre la lésion et le tissu environnant. Le contour est net et bien défini. Pour qu'une masse soit qualifiée de circonscrite, il faut qu'au moins 75% de son contour soit nettement délimité.
 - Micro-lobulé: dans ce cas, de courtes dentelures du contour créent de petites ondulations.

- Masqué: un contour masqué est un contour qui est caché par le tissu normal adjacent.
- Mal définis « Indistinct » : dans ce cas, le contour mal défini, peut correspondre à une infiltration.
- Spiculé: la masse est caractérisée par des lignes radiaires, prenant naissance sur le contour de la masse. Ces lignes radiaires sont appelées, les spicules.

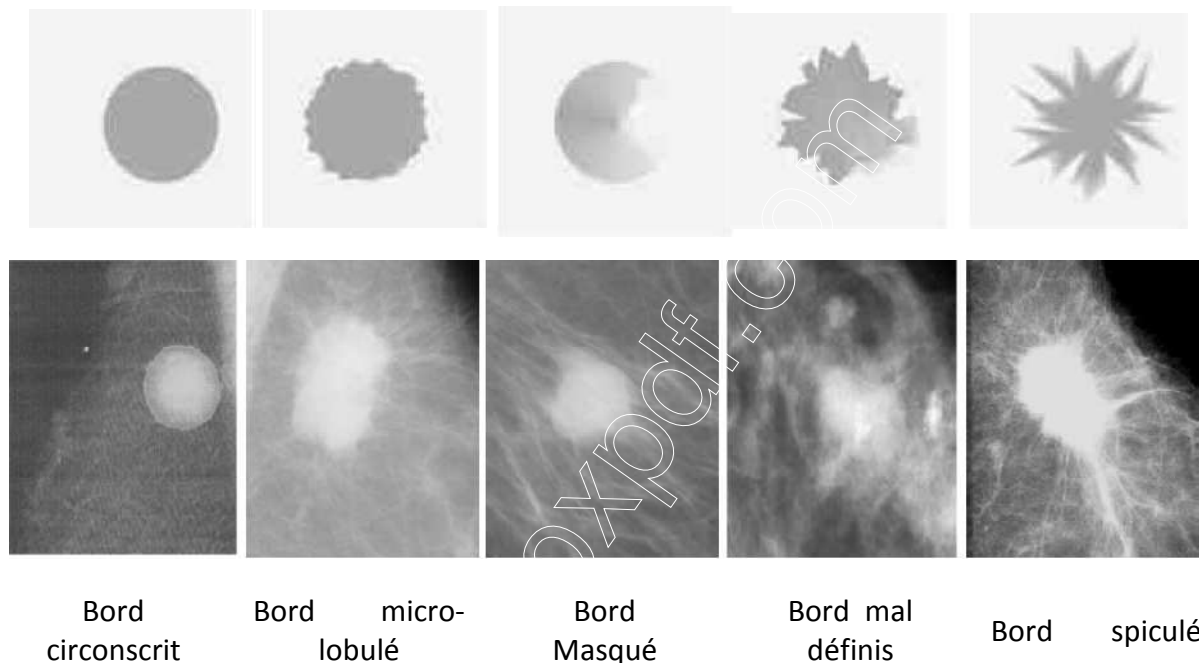


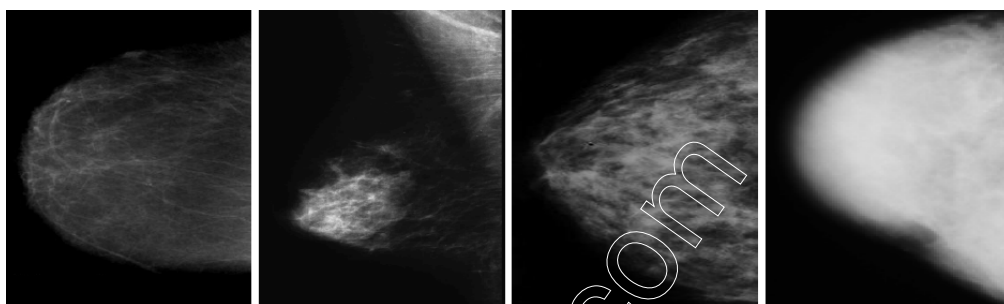
Figure 1.8: contours possibles d'une masse mammaire [10]

- **La densité:** la plupart des cancers du sein présentent une densité supérieure, ou égale au tissu conjonctivo-glandulaire. Il est rare (mais non impossible) qu'un cancer du sein, soit de densité inférieure. Les cancers du sein n'incluent pas de tissu adipeux, à moins que l'infiltration par le processus néoplasique n'ait incorporé un éventuel lobule graisseux. En outre, l'hypothèse d'un liposarcome est très improbable.

La densité (figure 1.9) peut être appréciée en utilisant la terminologie suivante:

- Radio-transparent, graisseux, adipeux: ce terme est utilisé pour toute lésion contenant de la graisse, tel qu'un kyste huileux, un lipome, une galactocèle, ainsi que des lésions mixtes telles qu'un hamartome ou adéno-lipome.

- Faible densité: il s'agit d'un tissu hypodense par rapport au tissu conjonctivo-glandulaire mais, sans inclusion de tissu adipeux.
- Densité moyenne: c'est un tissu isodense c'est à dire comparable au tissu conjonctivo-glandulaire.
- Forte densité: c'est un tissu hyperdense par rapport au tissu conjonctivo-glandulaire.



Radio-transparent Faible densité Densité moyenne Forte densité

Figure 1.9 : différentes densités mammaires [10]

b. les critères de bénignité et de malignité des opacités

Le tableau suivant indique les principaux critères de bénignité et de malignité des opacités.

Opacités	
Critères de malignité	Critères de bénignité
- Âge plus de 50 ans -Apparition récente -Augmentation de taille -Haute densité -Microcalcifications -Contours partiellement effacés (flous ou irréguliers) -Spicules	-Âge moins de 35 ans -Stabilité -Multiplicité -Faible densité et lacunes graisseuses -Macrocalcifications -Contours nets -Liseré clair

Tableau 1.3 : critères de bénignité et de malignité des opacités [14]

I.6.2 Les calcifications mammaires

Les calcifications mammaires sont des dépôts de calcium qui se forment dans le tissu du sein. Elles n'ont aucun lien avec la quantité de calcium absorbée au cours de l'alimentation

ou obtenue, par l'intermédiaire de compléments alimentaires. Les calcifications mammaires sont assez courantes et la plupart, ne sont pas associées au cancer. Afin de s'en assurer, le radiologue étudie leur taille, leur forme et leur disposition à l'aide d'une mammographie sur laquelle, elles apparaissent souvent sous forme de petits points blancs. Certaines de leurs caractéristiques, comme une forme irrégulière ou certains regroupements, peuvent être suspectes [4][10].

Il existe deux types de calcifications : les macrocalcifications et les microcalcifications.

a. Les macrocalcifications : Les macrocalcifications (figure 1.11) sont des dépôts grossiers de calcium dans le sein. Elles apparaissent comme de gros points blancs ou des tirets sur une mammographie. Elles sont plus fréquentes chez les femmes âgées, de plus de 50 ans. Elles sont souvent associées à des modifications bénignes qui se produisent, dans le sein et qui sont liées au vieillissement des artères du sein, à d'anciennes lésions, à une inflammation ou à des masses telles qu'un fibroadénome. C'est la raison pour laquelle, lorsque ces macrocalcifications sont découvertes, le radiologue ne recommande pas systématiquement de biopsie [25].

b. Les microcalcifications : les microcalcifications (figure 1.12) sont de minuscules dépôts de calcium dans le sein. Leur présence signifie parfois que l'activité de certaines cellules du sein, est accrue. Une cellule plus active absorbe en effet davantage de calcium que celle qui l'est moins. Les microcalcifications peuvent faire suspecter un cancer du sein, notamment lorsqu'elles apparaissent isolées ou regroupées en grappes à la mammographie. Dans ce cas, le radiologue recommande de réaliser une biopsie afin de vérifier, qu'il ne s'agit pas d'un cancer.

En fonction de leur taille, on peut classer les calcifications en microcalcifications ou macrocalcifications. Le diagnostic dépend de leur position dans le sein, de leur arrangement géométrique (linéaire, en forme d'agrégats, etc.) et du nombre de microcalcifications dans un amas, ainsi que de la comparaison avec une mammographie antérieure. Il se base sur les analyses suivantes:

- Analyse de la forme: les calcifications de forme régulière, rondes ou ovales sont des tumeurs bénignes alors que les calcifications irrégulières de formes variées, sont souvent associées avec une pathologie maligne.

- Analyse de la taille: les calcifications de moins de 0.5 mm, sont souvent associées à des cancers, celles de plus de 2 mm sont des lésions bénignes.
- Analyse du nombre: les petites calcifications nombreuses sont suspectes; un nombre au-dessus de 4 microcalcifications impose, une interprétation plus prudente.
- Analyse de l'orientation: les amas en forme triangulaire ou losangique orientée vers le mamelon, indiquent une nature maligne.
- Analyse de la morphologie: cette analyse permet le plus souvent de séparer les microcalcifications bénignes et malignes. Les calcifications arrondies ou ovales, uniformes dans leur taille et leur forme, sont probablement bénignes. A l'inverse, celles qui sont irrégulières dans leur taille, ressemblant par leur hétérogénéité à des 'débris de verre', sont probablement malignes.

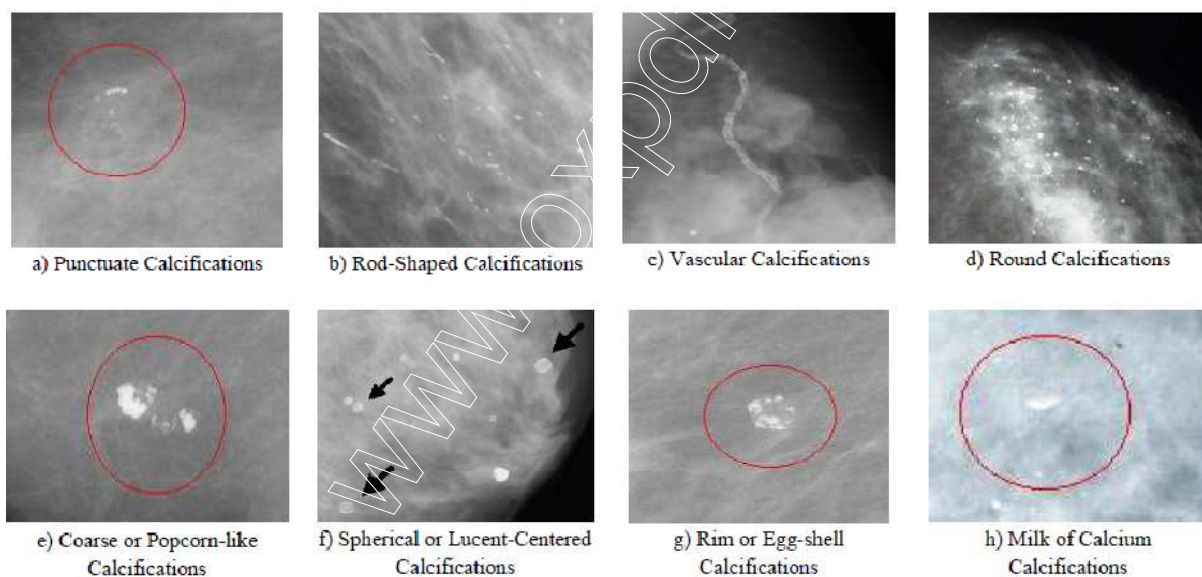


Figure 1.10 : calcifications bénignes [10]

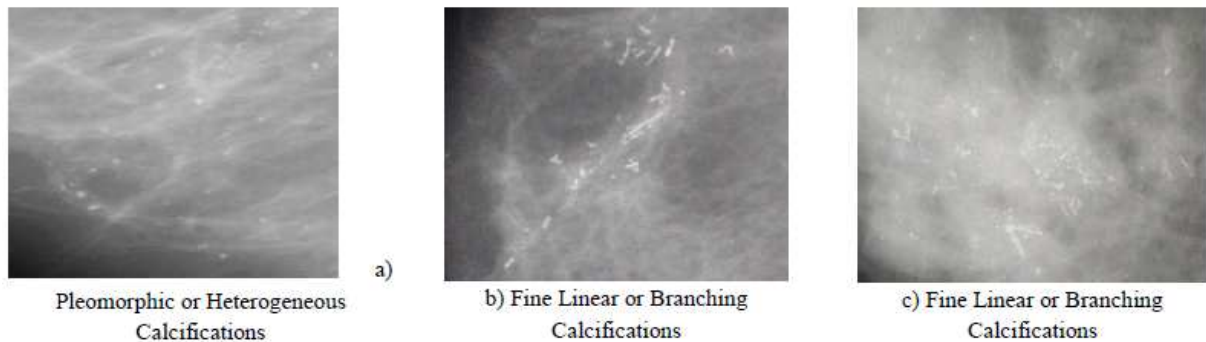


Figure 1.11 : calcifications malignes [10]

1.6.3 Asymétrie de densité [10]

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.

1.7 Conclusion

Dans ce chapitre, nous avons introduit quelques notions générales concernant l'anatomie du sein et, l'imagerie mammaire. Nous avons ensuite parcouru les principales pathologies qui peuvent atteindre les tissus du sein, en passant par l'examen de base qu'est la mammographie. Le chapitre suivant, introduit les techniques de prétraitement et de segmentation de l'image, adoptées dans le cas de notre étude.

Imagerie Mammographique

2.1 Introduction

Le traitement d'image est né de l'idée de la nécessité de remplacer l'observateur humain par la machine. La médecine est un des domaines qui n'a pu échapper, à ce souffle numérique que ce soit dans la chirurgie, le dépistage, le diagnostic et, la pratique thérapeutique.

Le but du traitement des images médicales, est d'extraire à partir des images acquises, les informations utiles au diagnostic, de révéler des détails difficiles à percevoir à l'œil nu, tout en évitant la création d'artefacts, faussement informatifs. Pour cela le traitement fait appel à des outils, des algorithmes, qui permettent d'agir sur l'image numérisée. L'un des processus fondamentaux dans la chaîne de traitement d'image est la segmentation [15].

Nous aborderons dans ce chapitre les principales méthodes de la segmentation, accompagnées de brèves descriptions et quelques illustrations, pour situer l'approche à adopter, dans le cadre de la détection des masses et des calcifications mammographiques.

2.2 Système de détection automatique

Un système de détection automatique de l'information pertinente produite par les images médicales de manière générale, se compose de deux étapes: une dédiée au prétraitement qui permet le rehaussement du contraste et, l'autre à la détection des lésions. Chacun de ces modules, peut se décomposer comme une étape de marquage, suivie d'une prise de décision [15][16].

La figure 2.1, résume les étapes du traitement d'une image mammographique. Elles se composent des modules suivants:

- une étape de prétraitement qui sert à améliorer la qualité de l'image;
- une étape de segmentation qui permet d'extraire les lésions mammographiques;
- une étape de caractérisation des lésions suivant des attributs spécifiques à la forme, et la texture;
- une étape de classification pour une prise de décision.

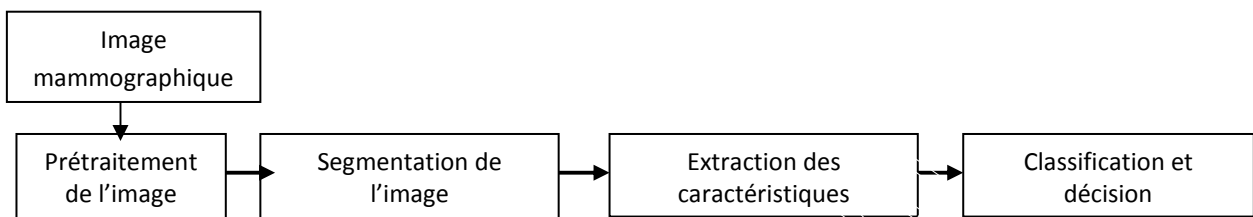


Figure 2.1: système de détection de pathologies mammographiques [16]

2.3 Prétraitement des images

Le prétraitement est une étape, essentielle dans l'amélioration de la qualité de l'image mammographique. 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 [16].

Une approche couramment utilisée en traitement d'images, consiste à modifier 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 [15].

Le but du prétraitement d'image est d'éliminer, le plus possible, les informations non pertinentes dues au bruit résultant de l'acquisition de l'image et par conséquent, faciliter l'extraction des informations utiles à l'analyse. Un rehaussement du contraste est appliqué aux images mammographiques, afin d'accentuer, les caractéristiques d'une image pour accroître la séparabilité des régions.

Une approche couramment utilisée pour l'amélioration du contraste, est la modification globale de l'histogramme. Cette méthode consiste à réaffecter les valeurs d'intensité des

pixels afin de rendre la nouvelle répartition des intensités, plus uniforme. Ceci peut être réalisé par l'égalisation d'histogramme ou par l'expansion de la dynamique [15].

Toutefois, ces transformations ont l'inconvénient de faire ressortir le bruit dans les images déjà bruitées, de ce fait on peut dire que les transformations globales, d'amélioration du contraste, ne peuvent être adaptées aux informations locales dans une image mammographique. Pour cette raison, il s'avère nécessaire de mettre au point des approches locales, permettant de préparer le cliché mammographique, à une segmentation judicieuse. Dans notre étude, nous utilisons un ajustement de l'image par une coopération entre l'expansion de la dynamique et, les opérations de la morphologie mathématique; pour mettre en évidence les calcifications. Le contraste des masses, est amélioré, via l'expansion de la dynamique.

2.3.1 Expansion de la dynamique [16]

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

De cette façon, si les valeurs de l'histogramme sont très proches les unes des autres, l'étirement va permettre de fournir une meilleure répartition afin de rendre les pixels clairs encore plus clairs et les pixels foncés proches du noir. Il est ainsi possible d'augmenter le contraste d'une image. Une image trop foncée pourra devenir plus « visible».

La courbe tonale correspondant à un étalement d'histogramme est de la forme suivante :

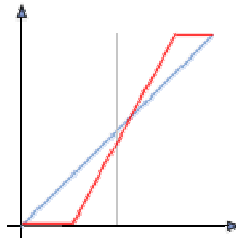


Figure 2.2: courbe de modification

L'expansion de la dynamique, consiste en la normalisation d'histogramme, suivant une transformation qui affine le niveau de gris des pixels de telle sorte que, l'image utilise toute la dynamique de représentation. Cette transformation est basée sur l'équation suivante:

$$f_{new}[x, y] = (f[x, y] - N_{min}) \cdot \frac{2^D - 1}{N_{max} - N_{min}} \quad 2.1$$

Avec : D : la dynamique

N_{min} : La valeur minimale du niveau de gris

N_{max} : La valeur maximale du niveau de gris dans l'image.

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

2.3.2 Morphologie mathématique [17][18]

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

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

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

a. Élément structurant

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

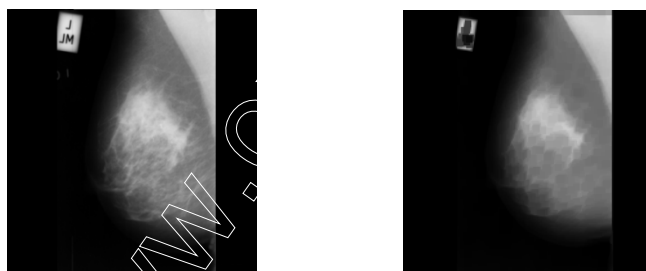
b. Erosion

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

L'ensemble des positions x correspondant à une réponse positive, forme le nouvel ensemble Y, appelé érodé de X par B, noté $X \ominus B$ (voir Figure 2.3).

Cet ensemble satisfait l'équation suivante :

$$X \ominus B = \{x \text{ dans } X : B \subset X\} \quad 2.2$$



a) Image mammographique initiale b) Image mammographique érodée

Figure 2.3: érosion d'une image mammographique

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

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

c. Dilatation

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

L'ensemble des points de l'image (Figure 2.4) correspondant aux réponses positives forme le nouvel ensemble Y des dilatés de X, noté comme suit:

$$y = x \oplus B = \{ \text{dans } 1 : B \cap X \neq \emptyset \} \quad 2.3$$



a) Image mammographique initiale **b)** Image mammographique dilatée

Figure 2.4: dilatation d'une image mammographique

d. Ouverture

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

$$OB(X) = (X \ominus B) \oplus B \quad 2.4$$

L'ensemble ouvert est plus régulier et moins riche en détails que l'ensemble X initial (Figure 2.5). La transformation par ouverture adoucit les contours, coupe les isthmes étroits, supprime les petites îles et les caps étroits.



a) Image mammographique initiale **b)** Image mammographique avec ouverture

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

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

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

e. Fermeture

La fermeture est l'opération "inverse" de l'ouverture, c'est à dire que l'on applique tout d'abord une dilatation puis une érosion (toujours en gardant le même élément structurant). Le résultat est nommé le fermé de X par B, noté comme suit :

$$F^B(X) = (X \oplus B) \ominus B \quad 2.5$$

Un ensemble fermé, est également moins riche en détails que l'ensemble initial (Figure 2.6). La transformation par fermeture, bouche les canaux étroits, supprime les petits lacs et les golfes étroits.



a) Image mammographique initiale b) Image mammographique avec fermeture

Figure 2.6 : fermeture d'une image mammographique par un élément structurant circulaire

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

- boucher les trous plus petits que B,
- conserver souvent la taille et la forme,
- ne pas conserver nécessairement la topologie,
- souder les formes proches en particulier.

f. Chapeau haut de forme

On appelle *transformation chapeau haut-de-forme* (ou *top hat*) la différence entre l'image originale et l'ouvert de cette image, ou entre le fermé de l'image et l'image originale. Le

premier type de chapeau haut-de-forme est appelé *chapeau haut-de-forme blanc*, car il permet de détecter ce que l'ouverture a fait disparaître, c'est-à-dire les pics ou parties claires de l'image originale. Le second type du chapeau haut-de-forme, dual du premier, est appelé *chapeau haut-de-forme noir*, car il détecte les vallées, ou parties sombres d'une image.

Le chapeau haut de forme noir est représenté par l'équation suivante :

$$WTH(x) = f(x) - f_{Bn}(x) \tag{2.6}$$

Le chapeau haut de forme blanc :

$$BTH(x) = f_{Bn}(x) - f(x) \tag{2.7}$$

Où B est l'élément structurant,

n représente la taille de l'élément structurant.

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

2.4 Segmentation de l'image

2.4.1 Principe de la segmentation

La segmentation des images constitue le cœur de tout système de vision. C'est une étape importante dans le processus d'analyse des images. La segmentation est un traitement de bas niveau, qui consiste à créer une partition d'une image **A** en sous-ensembles **R_i**, appelés régions [15], tels que:

$$\left. \begin{aligned} \forall i R_i &\neq \emptyset \\ \forall i, j; i \neq j R_i \cap R_j &= \emptyset \\ A &= \bigcup_i R_i \end{aligned} \right\} \dots\dots\dots \tag{2.8}$$

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 [15].

Il n'y a pas de méthode unique de segmentation. Selon **Cocquerez [15]**, le choix d'une technique est lié à la nature de l'image, aux opérations situées en aval de la segmentation, aux primitives à extraire, aux contraintes d'exploitation. Du fait de cette diversité, il est difficile de définir une méthode universelle de segmentation qui convienne à tous les cas. Cependant, en s'inspirant des systèmes de la vision humaine, la segmentation fait référence, aux notions de différence et de similarité. Il en découle deux approches courantes: l'approche « **frontière** » et, l'approche « **région** ». L'approche région, s'attache à faire apparaître des régions homogènes selon un critère d'homogénéité (niveaux de gris, couleur ou texture). C'est une approche, qui fait référence à des regroupements de régions, ayant des caractéristiques similaires; alors que, l'approche frontière tente de trouver des contours ou frontières de régions, présentant une variation rapide du même critère, ou, une discontinuité entre les propriétés de deux ensembles, de points connexes. Ces deux approches, sont duales en ce sens, qu'une région définit une ligne par son contour et, qu'un contour fermé, définit une région. Elles amènent, cependant à des algorithmes différents et, ne fournissent pas, en général, les mêmes résultats. Certaines méthodes, récentes tentent de faire coopérer plusieurs méthodes de segmentation, afin de combiner les avantages de chaque méthode ou d'exploiter leurs complémentarités.

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**).

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

- L'approche région,
- L'approche contour,
- L'approche coopérative.

2.4.2 Approche région [15][19][20]

Les méthodes de cette approche s'intéressent au contenu de la région (des informations locales ou globales). Elles consistent à partitionner une image, en un ensemble de régions connexes, en se reposant sur la recherche de zones de l'image possédant des attributs similaires, comme le niveau de gris, la texture, ou la couleur.

Le choix d'une méthode est lié, à la nature des images et à la forme des primitives à extraire.

Il existe principalement quatre approches de la segmentation en régions:

- la segmentation par division de régions,
- la segmentation par fusion de régions,
- la segmentation par division-fusion de régions,
- la segmentation par croissance de régions.

a. Segmentation par division (split)

Le principe de ces méthodes; consiste à tester la validité du critère de manière globale sur une large zone de l'image. Ainsi, si le critère est validé, l'algorithme s'arrête. Sinon, on décompose la zone considérée en zones plus petites et on applique, la méthode jusqu'à obtention d'une zone qui valide le critère. Ce qui produit un quaternaire (quad-tree), correspondant à une structure hiérarchisée, des régions qui répondent au critère d'homogénéité.

Cette structure de données est un arbre quaternaire qui permet de stocker l'image à plusieurs niveaux de résolution. On part d'une région initiale qui est l'image tout entière. Si cette image vérifie un critère d'homogénéité, l'algorithme s'arrête. Sinon, on découpe cette région en quatre parties de même taille et on lance la procédure récursivement dans ces quatre parties. La région initiale va être stockée comme un nœud dans un graphe et les sous parties comme des fils de ce nœud. La figure 2.7.a montre une image en noir et blanc 8x8 et le découpage correspondant à chaque niveau. La structure d'arbre associée à ce découpage est illustrée dans la figure 2.7.b.

Le quadtree est une représentation de l'image exploitant la récursivité du pavage carré. Il ne s'applique que sur des images carrées dont le coté est une puissance de 2 (de taille $2^N \times 2^N$). Cette représentation correspond donc, à la structure rigide du pavage carré.

L'approche de la segmentation par division de régions consiste à diviser l'image originale, en régions homogènes au sens d'un critère donné.

Ce processus est récursif et considère que la région initiale, correspond à l'image à analyser. Si une région ne respecte pas un prédicat d'homogénéité, elle est divisée en quatre sous régions de taille égale. Chaque sous région est ensuite analysée. L'algorithme récursif s'arrête, lorsque toutes les régions respectent le prédicat d'homogénéité.

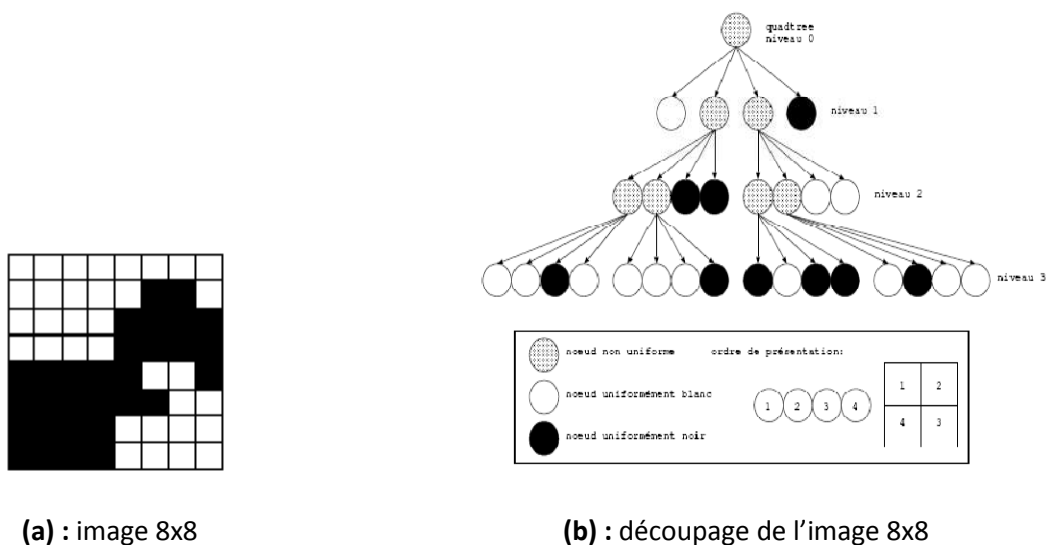


Figure 2.7 : quadtree sur une image 8x8 [15]

b. Segmentation par fusion de régions

L'approche de fusion de régions est principalement fondée, sur l'analyse d'un graphe d'adjacence de régions qui analyse une image pré-segmentée, constituée d'un ensemble de régions.

C'est une structure de données constituée d'un graphe non orienté dont chaque nœud représente une région et chaque arête, représente une adjacence entre deux régions. La figure 2.9 représente, à titre d'exemple, le graphe d'adjacence de régions obtenu à partir de l'image de la figure 2.8.

L'analyse du graphe d'adjacence de régions permet de fusionner, des régions d'une image sursegmentée. Le procédé consiste à fusionner deux nœuds, reliés par une arête à condition qu'ils respectent un critère de fusion.

Les méthodes d'analyse des graphes d'adjacence de régions, se distinguent selon l'ordre de parcours des différents nœuds du graphe et selon, les critères de fusion. Comme pour la croissance de régions appliquée à l'image, l'ordre de balayage du graphe d'adjacence de régions influe sur les résultats de la fusion.

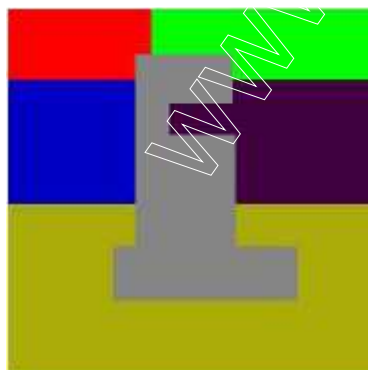


Figure 2.8 : image originale [15]

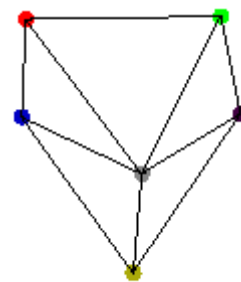


Figure 2.9 : graphe d'adjacence [15]

Cette méthode consiste à construire un état initial où chaque pixel de l'image constitue une région élémentaire, puis à fusionner progressivement les régions connexes qui satisfont un certain critère d'homogénéité, Pour obtenir un ensemble de régions homogènes. Comme exemple de critère simple, on peut considérer l'écart entre les valeurs moyennes des pixels

dans deux régions. Si cet écart est inférieur à un certain seuil, on peut fusionner les deux régions en une seule. L'ordre dans lequel sont réalisées les fusions influe naturellement sur le résultat. Il convient de commencer par fusionner les régions les plus "proches" selon le critère considéré. Le résultat est une image où tous les pixels d'une région sont marqués de la même valeur. Le plus naturel est de prendre pour chaque région la valeur moyenne de ses pixels dans l'image d'origine. Leurs inconvénients majeurs c'est que le résultat final « l'image segmentée » dépend fortement du choix des régions de départ.

c. Segmentation par division fusion

Le processus est une approche mixte mêlant la fusion de région avec un processus de division. Cette technique se compose des 2 phases suivantes:

- Découper itérativement (figure 2.10) l'image jusqu'à avoir des blocs contenant exclusivement des pixels similaires : on commence par définir un bloc de la taille de l'image, puis on examine le contenu de ce bloc. Si le bloc est homogène alors on arrête la décomposition. Sinon, on découpe le bloc en 4 sous-blocs et on examine le contenu de chaque sous-bloc. Ainsi de suite jusqu'à ce qu'il n'y ait plus besoin, de décomposer les blocs. Le résultat obtenu est donc un ensemble jointif, de blocs de différentes tailles qui recouvrent entièrement l'image [15].

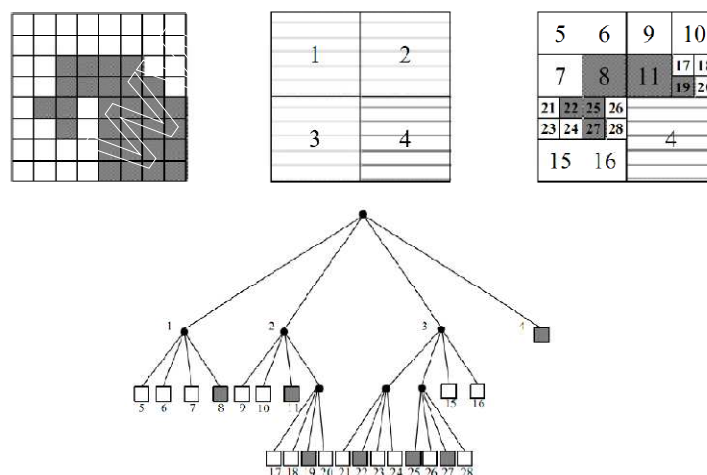


Figure 2.10 : découpage itérative de l'image [15]

- Regrouper les blocs voisins (figure 2.11) s'ils sont similaires: cette étape a pour objectif d'identifier les régions, qui composent l'image en regroupant les blocs jointifs et similaires. Il faut tout d'abord définir le critère de similarité, entre blocs. Si l'union de ces deux régions vérifie, le critère d'homogénéité, alors, on fusionne les régions. La principale difficulté de cette approche, réside bien sûr dans le parcours de l'ensemble, de tous les couples de régions voisines.

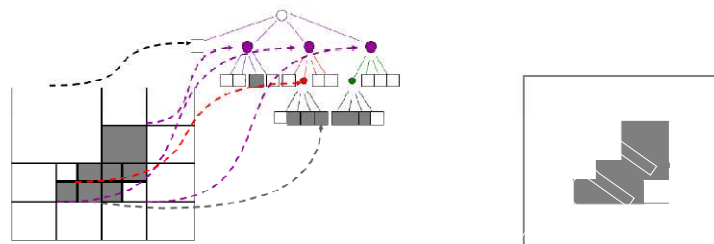


Figure 2.11 : regroupement des blocs voisins similaires [15]

d. Segmentation par croissance de régions

Ce type de segmentation consiste à faire croître des régions en y ajoutant, successivement les pixels adjacents qui satisfont un critère d'homogénéité. La croissance s'arrête, lorsque tous les pixels ont été traités.

L'étape initiale consiste à sélectionner les germes des régions qui correspondent généralement à un pixel. Puis, les régions sont construites en y ajoutant successivement les pixels qui leur sont connexes et qui, vérifient un critère de similarité.

La plupart des algorithmes de croissance de régions, sont régis par un parcours de l'image selon le balayage de gauche à droite et de haut en bas.

L'avantage de la croissance de régions est de préserver, la forme de chaque région de l'image [15].

La figure 2.12, illustre un exemple d'une segmentation d'image, en croissance de régions.

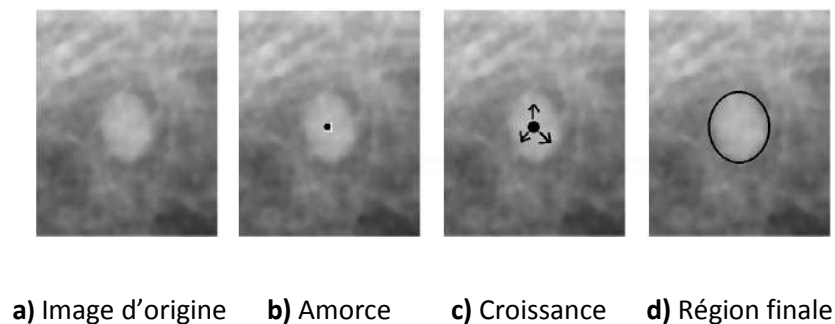


Figure 2.12: segmentation d'une image par croissance de régions

Parmi les avantages de cette technique, nous pouvons citer :

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

Cependant, il existe plusieurs inconvénients:

- L'influence du choix des germes initiaux et du critère d'homogénéité, sur le résultat de la segmentation.
- Une mauvaise sélection des germes ou un choix du critère de similarité, mal adapté peuvent entraîner des phénomènes de sous-segmentation ou de sur-segmentation.
- Il peut y avoir des pixels qui ne peuvent, pas être classés.

2.4.3 Segmentation par contours [15][19]

L'approche contour est l'approche duale de l'approche région: elle consiste à rechercher les frontières des régions caractérisées par une discontinuité visuelle. Ces méthodes sont basées sur la recherche de variations locales significatives. Largement étudiée pour les images en niveaux de gris, la recherche de contours repose principalement sur le calcul des dérivées.

Les méthodes de segmentation basées sur l'approche contour, ont pour objectif de trouver les lieux de fortes variations du niveau de gris. Une multitude de méthodes ont été

développées. Ces dernières s'appuient sur la détection des discontinuités dans l'image et, peuvent être divisées en trois classes :

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

Dans le cas de notre application, nous nous intéressons aux modèles déformables.

a. Les méthodes dérivatives

Les méthodes dérivatives sont les plus utilisées dans la détection des transitions d'intensités. Ce sont des méthodes locales qui balayent l'image, avec un masque définissant la zone d'intérêt. Elles consistent à calculer la dérivée en chaque point, de l'image afin de mettre en évidence les variations des niveaux de gris. On peut classer les méthodes dérivatives en deux groupes; selon qu'on utilise la dérivée première (approche gradient ou la dérivée seconde (approche Laplacien) (figure 2.13).

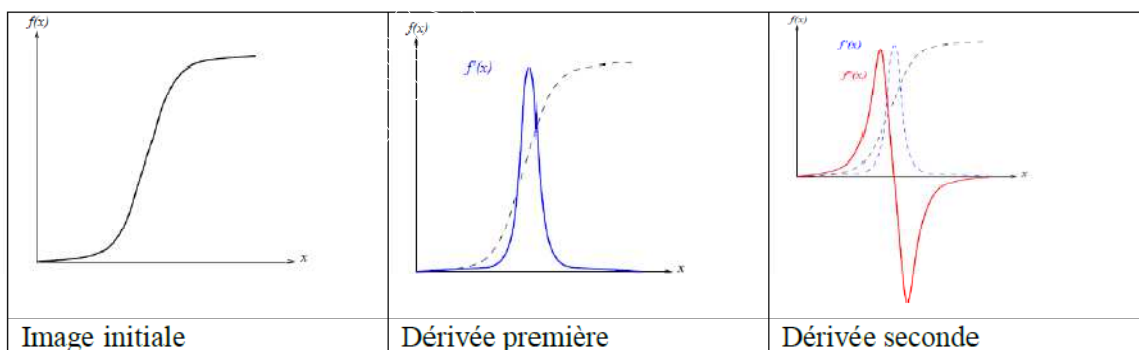


Figure 2.13 : contour et ses dérivées [19]

- **L'approche gradient** : ce type de détecteur se base sur la première dérivée de l'image I en chacun de ces pixels; dans les deux directions horizontale et verticale. Un pixel de contours aura une amplitude et une direction.

La valeur du gradient est ainsi disponible en tout pixel de l'image permettant d'effectuer une recherche des maxima locaux. Ceux-ci correspondent aux passages par zéro de la dérivée seconde, dans la direction du gradient ou encore aux pixels contours recherchés.

- **L'approche Laplacien** : Ce type de détection de contour se base sur la dérivée seconde de l'image. Les pixels de contour sont situés aux passages par zéro du laplacien. Contrairement au gradient, le laplacien permet d'obtenir des contours fermés et d'un pixel d'épaisseur, par contre il a l'inconvénient d'être plus sensible au bruit que le gradient.

b. Les méthodes analytiques

Un autre type d'approche plus récente repose sur la définition de critères d'optimalité de la détection de contours. Ces critères débouchent sur des filtres de lissage optimaux. Canny [20] a proposé en 1986 une étude théorique de la détection de contours dans laquelle, il a formalisé trois critères que doit valider un détecteur de contour: une bonne détection liée à la robustesse au bruit, une bonne localisation dépendant de la précision de la localisation du point contour et l'unité pour que pour chaque contour, il y ait une seule réponse.

Notons que, à l'exception des méthodes par passage par zéro du Laplacien, la plupart des détecteurs de contours fournissent des contours ouverts, c'est à-dire qu'ils ne séparent pas les composantes de l'image dans les objets topologiquement distincts. Pour aboutir à une segmentation correcte, il faut leur adjoindre une opération de fermeture de contour ou utiliser les contours actifs.

c. Les modèles déformables

Les algorithmes de segmentation fondés sur les modèles déformables ont l'avantage, par rapport aux méthodes dérivatives, de fournir des contours ou surfaces fermés [21]. Parmi ces méthodes, on trouve les contours actifs et les ensembles de niveaux.

➤ **Les contours actifs « Snakes »**

La méthode des contours actifs est une réponse originale au problème de la segmentation, proposée par M. Kass, A. Witkin, et D. Terzopoulos [22]. Un contour actif est une courbe, évoluant d'une manière dynamique sur une image ; à partir d'une forme initiale vers les contours d'un objet d'intérêt. Cette évolution est réalisée de manière itérative sous l'action de deux forces: une force interne issue de la courbe elle-même et permettant, de gérer la régularité du contour actif et une force externe, calculée à partir des données de l'image [23].

Le principe de cette méthode (figure 2.14), pour la segmentation d'une région d'intérêt, consiste à initialiser le contour de la région et à le déplacer pour épouser au mieux les frontières de la région d'intérêt (figure 2.15). Cette évolution dynamique est appliquée sous la contrainte de conserver certaines caractéristiques du contour, telles que la courbure, la répartition des points, ou d'autres contraintes liées à la disposition des points. L'application de cette approche sur des images difficiles à segmenter [24], a montré que la segmentation par les contours actifs est une technique qui possède des avantages remarquables en termes de structure obtenue (courbe fermée), de qualité de contour et de rapidité de convergence des résultats lorsque l'initialisation est proche du contour de la région d'intérêt.

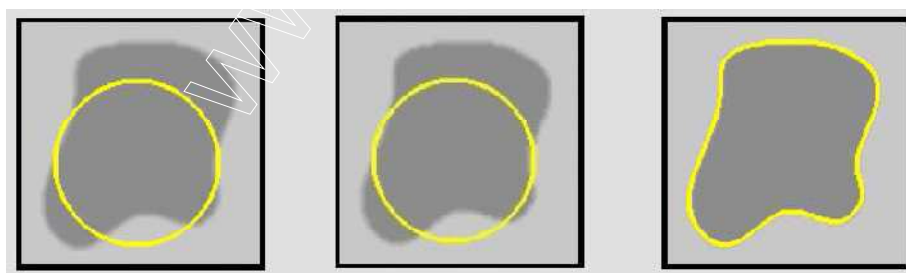


Figure 2.14 : principe du contour actif [22]

La figure 2.15 illustre un exemple d'application des contours actifs sur une image mammographique.

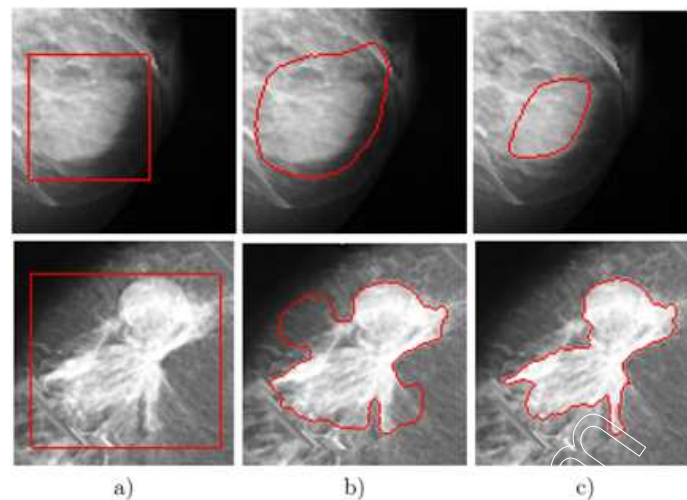


Figure 2.15 : segmentation d'une masse circonscrite (première ligne) et d'une masse spiculée [24] (seconde ligne) (a : initialisation, b : propagation, c : convergence vers le contour de la lésion)

➤ **Les ensembles de niveaux « levels set »**

La méthode des ensembles de niveaux (figure 2.16) est une méthode de simulation numérique utilisée pour l'évolution des courbes et des surfaces dans les domaines discrets. C'est une représentation variationnelle de contours qui évoluent dans le temps, mais différemment des contours actifs [25]. En effet, les ensembles de niveaux peuvent changer de topologie si les contours l'imposent. Par exemple, changer de connexité (un contour simple peut par exemple donner naissance à plusieurs contours fils qui vont évoluer séparément).

Pour permettre son évolution ultérieure, la fonction ψ dépend du temps:

$$\begin{aligned} \psi : D \times \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, t) &\rightarrow \psi(x, t) \end{aligned}$$

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

$$\Gamma = \{x \mid \psi(x, t) = 0\} \quad 2.9$$

Un choix habituel pour l'initialisation de la fonction $\psi(x)$ est la distance euclidienne signée, qui sépare, x du contour Γ . Le signe dépend de la position du point x à l'intérieur ou, à l'extérieur de Γ . Sur la figure 2.16, la région R_{in} est l'ensemble des points pour lesquels ψ est positive :

$$R_{in} = \{x | \psi(x, t) \geq 0\} \tag{2.10}$$

La fonction des ensembles de niveaux (figure 2.16) peut-être initialisée, à partir de la distance euclidienne comme suit :

$$\psi \begin{cases} 0 & \text{si } x \in \Gamma \\ \inf_{y \in D} \|x - y\| & \text{si } x \in R_{in} \\ \inf_{y \in D} \|x - y\| & \text{si } x \text{ n'appartient pas à } R_{in} \end{cases} \tag{2.11}$$

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

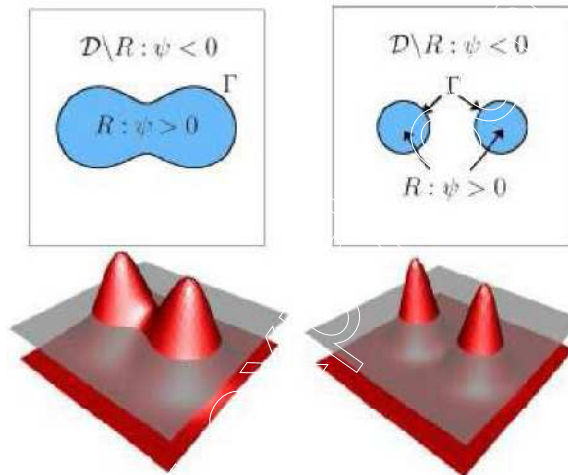


Figure 2.16 : fonction d'ensembles de niveaux [2(3)]

La figure 2.17 présente un exemple d'application des ensembles de niveaux, sur une image synthétique.

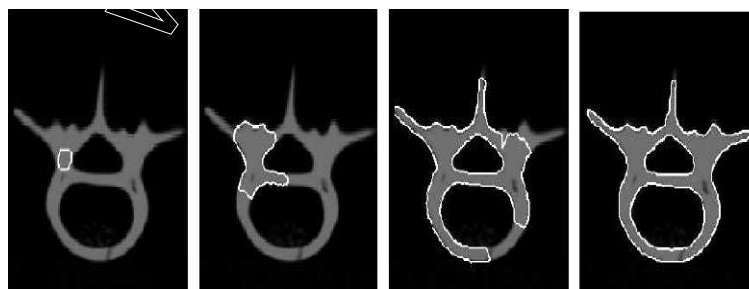


Figure 2.17: exemple d'application des Level sets sur une image synthétique [25]

La méthode des ensembles de niveaux présente de nombreux avantages [25]:

- Adaptation au changement de topologie: plusieurs objets peuvent être segmentés à la fois prise en compte des changements de topologie automatique.
- Extension à la 3 D simple.
- Utilisation des méthodes numériques connues pour calculer les dérivées.
- Facilité de calcul des grandeurs géométriques intrinsèques de l'interface (normale, courbure).

On peut cependant relever les inconvénients suivants :

- Le temps de calcul.
- Nécessité de recalculer la fonction distance par rapport au niveau zéro.
- La vitesse v n'est définie que pour le niveau zéro de la fonction Level Set (Extension des vitesses à tout le domaine).
- Implémentation complexe et coûteuse.

L'avantage principal de cette méthode est la possibilité de gérer automatiquement le changement de topologie, de la courbe en évolution.

❖ **Modèle de Chan et Vese [26]**

Pour ce type de modèle déformable, Chan & Vese utilisent un descripteur région qui est en fonction de la moyenne d'intensité. Un tel critère sert à segmenter la région homogène, car on cherche la région dont les pixels, ont une intensité proche de la moyenne de la région. Trois descripteurs sont pris en compte: descripteur de l'intérieur, de l'extérieur et du contour.

Soit I une image, C un contour fermé qui sépare deux régions l'une à l'intérieur (Ω_{int}) et l'autre à l'extérieur de C (Ω_{ext}) (figure 2.18).

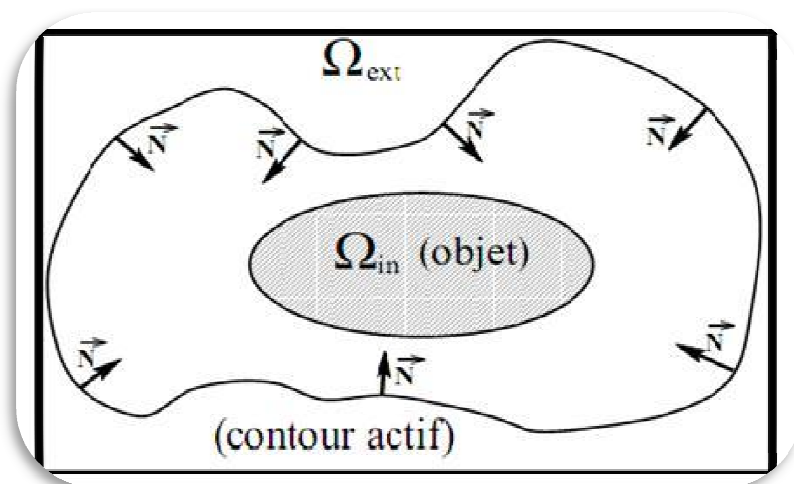


Figure 2.18 : domaines composant l'image [26]

C_1 et C_2 : L'intensité de I à l'intérieur et l'extérieure de C , respectivement. On définit les fonctionnelles citées ci-dessus par :

$$E_{int}(c) = \int_{intérieure\ de\ c} |I(x, y) - c_1|^2 dx \tag{2.12}$$

$$E_{ext}(c) = \int_{extérieure\ de\ c} |I(x, y) - c_2|^2 dx \tag{2.13}$$

$$Length(C) = L\{\phi = 0\} = \int_{\Omega} \delta(\phi(x, y)) |\nabla\phi(x, y)| dx dy \tag{2.14}$$

Avec $\delta(x)$: est la fonction de Dirac

Le modèle de Chan et Vese est le suivant:

$$E(c) = \mu(\text{longueur}(c) + v \cdot \text{l'aire}(l'intérieur\ de\ c) + \lambda_1 E_{int}(c) + \lambda_2 E_{ext}(c) \tag{2.15}$$

Avec :

$$\text{aire}(l'intérieur(c)) = A\{\Phi > 0\} = \int_{\Omega} H(\Phi(x, y)) dx dy \tag{2.16}$$

Avec :

$H(x)$: est la fonction de Heaviside

On intègre les équations (2.12), (2.13), (2.14) et (2.16) dans (2.15), on aura l'énergie totale :

$$E(c) = \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy + v \int_{\Omega} H(\Phi(x, y)) dx dy + \lambda_1 \int_{interieur\ de\ c} |I - c_1|^2 dx + \lambda_2 \int_{exterieur\ de\ c} |I - c_2|^2 dx \quad (2.17)$$

Il est clair que les régions citées ci-dessus peuvent se définir par :

$$\begin{cases} C = \{(x, y) \in \Omega : \phi(x, y) = 0\} \\ interieur(C) = \{\phi(x, y) \in \Omega : \phi(x, y) > 0\} \\ exterieur(C) = \{\phi(x, y) \in \Omega : \phi(x, y) < 0\} \end{cases} \quad (2.18)$$

En introduisant la fonction du Level Set on peut écrire:

$$\int_{E_{int}(c)} |I(x, y) - c_1|^2 dx dy = \int_{\phi > 0} |I(x, y) - c_1|^2 dx dy = \int_{\Omega} |I(x, y) - c_1|^2 H(\phi(x, y)) dx dy \quad (2.19)$$

$$\int_{E_{ext}(c)} |I(x, y) - c_2|^2 dx dy = \int_{\phi < 0} |I(x, y) - c_2|^2 dx dy = \int_{\Omega} |I(x, y) - c_2|^2 H(\phi(x, y)) dx dy \quad (2.20)$$

- L'ensemble $\{X \in R^2, \phi(x) > 0\}$ représente l'intérieur du contour.
- L'ensemble $\{X \in R^2, \phi(x) < 0\}$ représente l'extérieur du contour.
- H : fonction de Heaviside : c'est une fonction détectrice de chacune des deux phases.
- C_1 et C_2 : sont deux constantes pour lesquelles l'image est approchée dans chacune des phases.

L'objectif est donc de minimiser la fonctionnelle énergétique :

$$\begin{aligned} \text{Min}(E(c)) = & \text{Min}(\mu \int_{\Omega} \delta(\phi(x,y)) |\nabla \phi(x,y)| dx dy + v \int_{\Omega} H(\phi(x,y)) dx dy + \\ & \lambda_1 \int_{\Omega} |I(x,y) - c_1|^2 dx dy + \lambda_2 \int_{\Omega} |I(x,y) - c_2|^2 dx dy) \end{aligned} \quad (2.21)$$

Pour calculer les intensités moyennes, c_1 et c_2 on peut suivre :

$$c_1 = \frac{\int_{\Omega} I(x,y) H(\phi(x,y)) dx dy}{\int_{\Omega} H(\phi(x,y)) dx dy} \quad (2.22)$$

$$c_2 = \frac{\int_{\Omega} I(x,y) (1-H(\phi(x,y))) dx dy}{\int_{\Omega} (1-H(\phi(x,y))) dx dy} \quad (2.23)$$

V : est un terme de pondération entre la régularisation et l'attache aux données. Cette fonctionnelle est minimisée, à l'aide des équations d'Euler, Lagrange et du gradient.

L'algorithme du modèle de Chan & Vese est donné suivant les étapes ci-dessous :

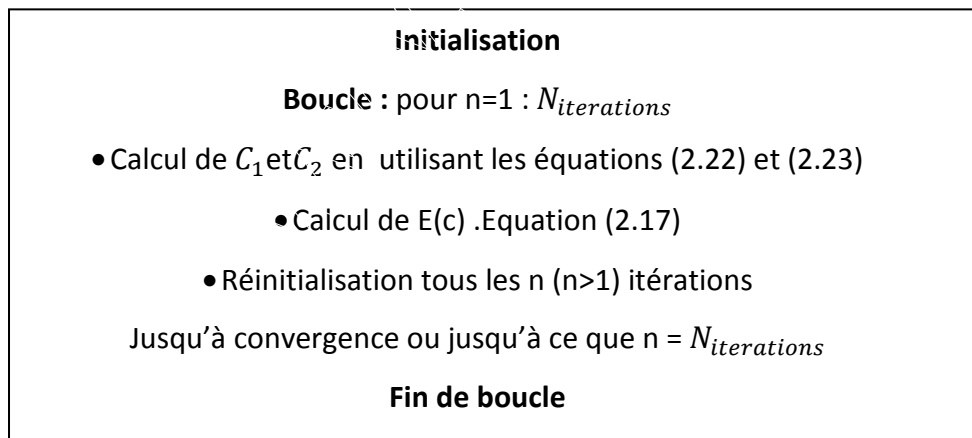


Figure 2.19 : algorithme de Chan et Vese [26]

2.4.4 Segmentation par coopération [27]

Il existe plusieurs approches de segmentation d'images, chacune agit de manière différente et utilise des attributs différents. De plus, chaque méthode ayant ses avantages et ses limites d'utilisation selon le problème à résoudre.

Les méthodes coopératives apportent une meilleure segmentation des images, puisque les algorithmes émergents de la coopération prennent en compte les différentes caractéristiques de ces images. Il existe en grande partie deux types de coopération; une coopération dite « séquentielle » dans laquelle il existe une exploitation successive du résultat d'une approche pour guider une autre. Une autre coopération dite « itérative » est fondée sur une définition mutuelle de contraintes et qui, permet une émergence de la solution.

- Coopération séquentielle : les approches par coopération séquentielle sont principalement fondées sur des principes de focalisations successives et de corrections de résultats intermédiaires. L'objectif est de réduire progressivement la difficulté du problème en le décomposant en sous-parties de plus en plus fines.
- Coopération itérative : dans les approches par coopération itérative, plusieurs modules de segmentation s'exécutent simultanément. L'objectif est de construire la segmentation progressivement, en optimisant à chaque étape les résultats obtenus par chaque méthode mise en jeu. Il existe relativement peu de travaux proposant un tel type de coopération pour la segmentation.

2.5 Bref parcours de quelques travaux sur la détection des pathologies mammaires

Les approches menées pour la détection des pathologies mammaires sont diverses et variées. On retrouve des adaptations de toutes les méthodes du traitement d'images mettant en œuvre, l'analyse multirésolution [28], la segmentation par croissance de régions adaptative [29][30] [31], La détection par la morphologie mathématique mise en œuvre par la ligne de partage des eaux dans les travaux de Vachier [32], pour la détection

des masses, l'extraction des structures mammaires, en se basant sur les ensembles de niveaux dans la recherche de Boujelben et al [33], etc..

L'équipe d'imagerie pour l'aide à la décision du laboratoire LATSI, s'intéresse à la détection de tumeurs mammaires, issues des images mammographiques et cytologiques et; consacre ses travaux, à l'analyse topologique et texturale pour la réalisation d'un système d'aide au diagnostic, suivant une classification connexionniste [34], par les arbres de décision [35] ainsi que par les SVM [36] [37] et les algorithmes génétiques en hybridation avec les réseaux de neurones [38][39]. L'objectif de ce travail, est de réaliser, une classification coopérative entre les algorithmes génétiques et, le modèle du perceptron multicouche, pour la discrimination des signes malins et bénins des masses et des calcifications mammaires.

2.6 Méthodologie adoptée pour la détection des masses et des calcifications

2.6.1 Cas des calcifications

La détection des calcifications est complexe et subtile, vue leurs tailles et leurs formes. L'approche menée pour la détection des calcifications se fait sur la zone d'intérêt, expertisée par le radiologue. L'image, subit un rehaussement du contraste, suivant l'opération du chapeau haut de forme, que l'on superpose à l'image d'origine, à partir de laquelle, on extrait les calcifications, suivant le modèle de Chan et Vese.

L'organigramme de la figure 2.20 illustre les étapes, élaborées pour la détection des calcifications.

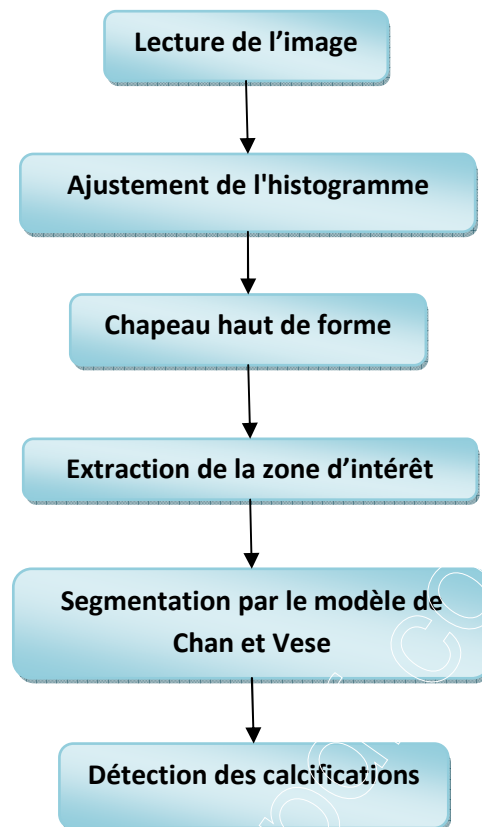


Figure 2.20 : organigramme d'extraction des calcifications

2.6.2 Cas des masses

La détection de masses, se fait suivant la coopération séquentielle, entre l'image accentuée par l'expansion de la dynamique et, les ensembles de niveaux, réalisés autour du modèle de Chan et Vese [26][33]. L'information fournie par cette coopération, permet d'éliminer les faux segments et de gagner en temps de traitement.

L'organigramme de la figure 2.21, illustre les étapes, que nous avons élaborées pour la détection des masses.

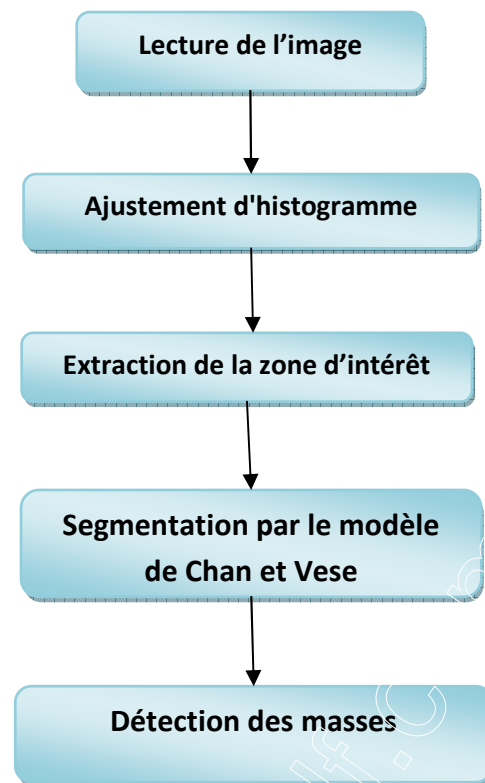


Figure 2.21 : organigramme d'extraction des masses

2.7 Conclusion

Dans ce chapitre, nous avons parcouru quelques méthodes de segmentation d'image, pour concevoir notre système de détection des masses et des calcifications. Nous nous sommes intéressées, plus précisément aux modèles déformables, pour une précision des résultats de la segmentation.

La segmentation d'image est une étape incontournable et fondamentale dans le traitement de l'image médicale. Ses résultats conditionnent fortement l'étape de l'interprétation si bien, qu'une bonne segmentation assure une bonne caractérisation. Le chapitre qui suit, est consacré à la caractérisation en se basant sur les descripteurs texturaux et morphologiques, ainsi qu'à la classification neuro-génétique des pathologies mammaires.

connexionniste en Imagerie Mammographique

3.1 Introduction

Ce chapitre introduit l'extraction de caractéristiques d'images. Par ce terme, nous entendons l'ensemble des méthodes permettant d'extraire des informations relatives à la texture et à la morphologie, pour caractériser le contenu structurel de l'image. Cette description, nous permet de nous orienter, vers la classification hybride entre le perceptron multicouche et les algorithmes génétiques, étape finale de notre système.

3.2. Analyse de la texture

L'objectif de l'analyse à l'échelle d'un motif ou d'un élément de texture, est de prendre en compte l'organisation des variations spatiales dans l'image en plus; de l'intensité lumineuse. La notion de texture en traitement d'image, est d'autant plus délicate à aborder que les définitions que l'on en donne, sont multiples. Ces définitions, sont souvent liées à un aspect particulier mais sont rarement génériques et, la quantité importante d'approches utilisées pour l'analyse des textures, témoigne de l'absence d'une définition précise. L'objet de ce chapitre, est d'aborder, la notion de texture, pour expliciter les descripteurs étudiés. Bien que la notion de texture soit naturelle pour l'être humain, elle résiste depuis Longtemps à toute tentative de définition. On peut s'en approcher en disant qu'une texture est une zone de l'image qui, présente certaines caractéristiques d'homogénéité qui la font apparaître comme une zone unique. Nous pouvons aussi la décrire comme étant, un ensemble de primitives de taille et de forme variables, présentant une organisation spatiale particulière [40][41].

Une des raisons principales qui fait que le problème d'analyse et de caractérisation de texture soit encore un problème ouvert est qu'il n'existe pas de définition universellement admise qui permette de caractériser la notion de texture. De ce fait, la texture est définie et étudiée de manières différentes en fonction de la façon dont on la caractérise. Nous citons ici quelques définitions proposées dans la littérature:

- ❖ Bonnevey [41] [43], la texture est définie comme étant une combinaison de deux types d'information:
 - Une information locale autour de chaque pixel, (statistiques, géométriques, sémantiques, etc).
 - une information globale qui correspond à la distribution spatiale de l'information locale (fenêtre de visualisation).

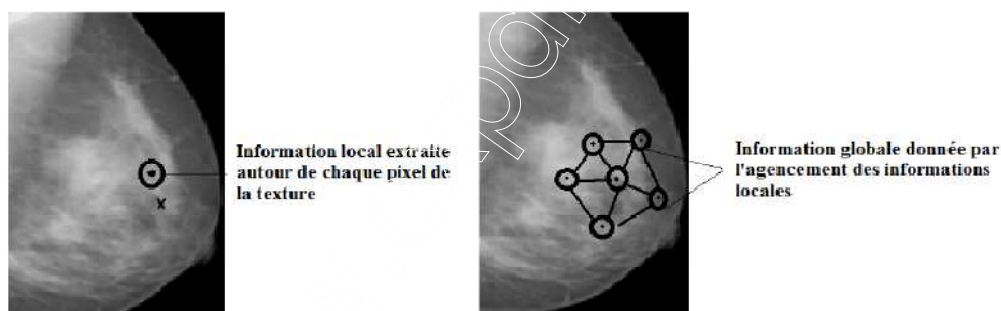


Figure 3.1 : information locale et information globale [43]

- ❖ Haralick [40] a décrit la notion de texture comme suit : « Une image texturée est décrite par le nombre et les types de ses primitives tonales ainsi que leurs orientations spatiales. Elle ne peut pas être analysée sans une vue de la primitive tonale de référence. Au fur et à mesure que la résolution augmente, on observe une texture fine puis une texture grossière.

3.2.1 Types de texture

Dans la littérature, on trouve plusieurs familles de textures qu'on peut classer en plusieurs catégories :

- **Textures aléatoires (micro-textures)**: les textures aléatoires sont désordonnées tout en apparaissant, globalement homogènes. Elles se distinguent par un aspect plus fin. Cette catégorie de texture a fait l'objet de nombreux travaux de recherche fondés sur les méthodes d'analyse statistique [41].

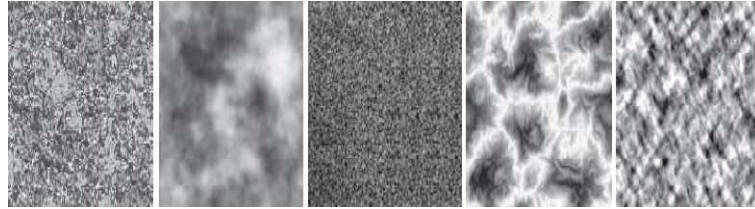


Figure 3.2 : exemple de textures aléatoires [41]

- **Textures structurales (macro-textures)**: on peut considérer les textures structurales comme étant la répétition spatiale de motifs élémentaires, appelé texton, suivant des règles de directions et de placement. Nous pouvons différencier les textures parfaitement périodiques, appelées textures déterministes et les textures dont, les primitives ont subi quelques transformations, appelées textures observables. L'analyse de ce type de texture se fait en utilisant, les méthodes d'analyses structurales [41].

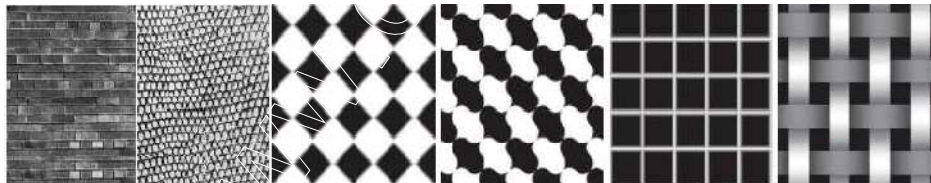


Figure 3.3: exemple de textures structurales [41]

- **Textures directionnelles (textures observables)**: les textures directionnelles ne sont pas totalement aléatoires et ne présentent pas, d'éléments structurants de base. Elles se caractérisent essentiellement, par une certaine orientation [41].

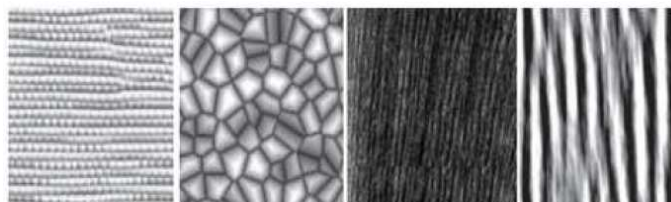


Figure 3.4 : exemple de textures directionnelles [41]

3.2.2 Méthodes d'Analyse des textures

L'objectif principal de l'analyse de la texture est la détermination de signatures texturales permettant une meilleure caractérisation de l'image autrement dit c'est la formalisation des descriptifs de la texture par des paramètres mathématiques qui serviraient à l'identifier. L'absence d'une définition formelle et précise de la texture a conduit à l'élaboration d'une multitude de méthodes d'analyse de textures. Le choix d'une méthode de caractérisation de textures est étroitement lié à l'application visée et à la nature de la texture à discriminer. Une multitude de méthodes, de variantes et de combinaisons de méthodes ont déjà été proposées dans la littérature. On les distingue :

a. Les méthodes structurelles

Les méthodes structurelles tiennent compte de l'information structurelle et contextuelle d'une forme et sont particulièrement bien adaptées aux textures macroscopiques. Les étapes d'analyse sont d'abord l'identification des éléments constitutifs, puis la définition des règles de placement. Les deux structures les plus importantes sont les structures de graphe et les structures syntaxiques [41].

b. Les méthodes statistiques

Du point de vue des méthodes statistiques, la texture est considérée comme la réalisation d'un processus stochastique stationnaire. Des paramètres statistiques sont estimés pour chaque pixel de l'image. Suivant la modalité des images à étudier, la signature la plus discriminante de la texture, est à rechercher soit dans des méthodes qui exploitent directement les propriétés statistiques de la texture (matrices de cooccurrences, matrice de longueurs de plages, matrice de voisinage, fonction d'auto corrélation, modèle de Markov, modèle autorégressif, modèles issues de la morphologie mathématique), soit dans des méthodes qui exploitent les propriétés statistiques à partir d'un plan transformé dans lequel on réécrit l'image de texture (densité spectrale, méthode des extrema locaux, méthodes de transformation de Fourier, Karhunen Loeve, Walsh Hadamard ou avec des filtres numériques) [41].

c. Les méthodes basées sur l'étude des formes

Ces méthodes se trouvent au croisement de la reconnaissance des formes, de la caractérisation de défauts et de l'analyse macro-texturale. Les régions texturales de l'image épousent des formes particulières et, peuvent être caractérisées par des paramètres dits de formes [41].

d. Les méthodes spatio-fréquentielles

Dans ces méthodes, les représentations spatio-fréquentielles préservent à la fois les informations globales et locales. Elles sont donc bien adaptées aux signaux quasi périodiques. En effet, les textures sont des signaux quasi périodiques qui ont, une énergie fréquentielle localisée. Ces méthodes permettent de caractériser la texture à différentes échelles [41].

e. Les méthodes fractales

Du point de vue de la synthèse d'image, les méthodes fractales sont à part, car elles permettent de synthétiser des images très proches de la réalité. En analyse de texture, la dimension fractale, qui est une mesure du degré d'irrégularité d'un objet, décrit une certaine propriété de la texture. Le modèle fractal est basé, essentiellement sur l'estimation par des méthodes spatiales de la dimension fractale de la surface, représentant les niveaux de gris de l'image [41].

3.2.3 La matrice de cooccurrence [42][43]

Elle estime des propriétés des images relatives, à des statistiques du second ordre. Une matrice de cooccurrence, mesure la probabilité d'apparition des paires de valeurs de pixels situés à une certaine distance, dans l'image. Elle est basée sur le calcul de la probabilité $P(i, j, d, \theta)$ qui représente le nombre de fois où un pixel de niveau de couleur i , apparaît à une distance relative d , d'un pixel de niveau de couleur j et selon une orientation θ donnée. Les directions angulaires θ classiquement utilisées, sont 0, 45, 90 et 135 degrés. Les relations de voisinage entre pixels, nécessaires au calcul des matrices, sont illustrées dans la figure

(3.5). Par exemple, les plus proches voisins de 'x' selon la direction $\theta=135$ degrés, sont les pixels 4 et 8.

La distance « d » permet d'avoir une description significative de la périodicité de la texture et l'angle « θ » permet d'évaluer la direction de texture. Cette matrice décrit les régularités observables dans les niveaux de gris des pixels d'une région.

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

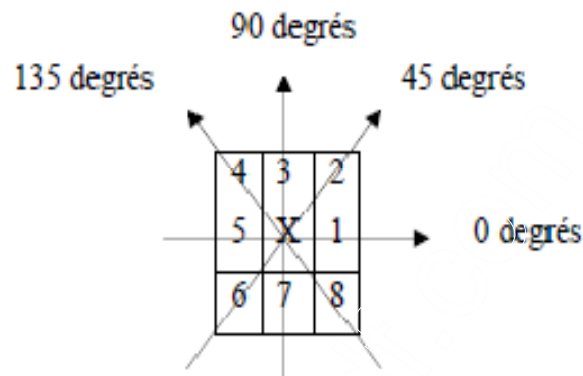


Figure 3.5: plus proches voisins de x selon 4 directions [42]

Les matrices de cooccurrence contiennent une masse d'information difficilement manipulable, c'est pour cela qu'elle n'est pas utilisée directement. De ce fait, quatorze indices définis par Haralick [41], qui correspondent à des caractères descriptifs de la texture peuvent être calculés, à partir de ces matrices.

Selon différentes études en imagerie mammographique [43], nous définissons les 6 paramètres considérés, comme les plus pertinents, dans le cas de notre travail.

a. L'entropie

Ce paramètre (équation 3.1) fournit un indicateur du désordre que présente, une texture. L'entropie atteint de fortes valeurs lorsque la texture, est complètement aléatoire (sans structure apparente). mesure la complexité de l'image. Lorsque les valeurs de la matrice de cooccurrences sont presque toutes égales, l'entropie est élevée. Elle permet ainsi de caractériser le degré de granulation de l'image. En effet, plus la valeur de l'entropie est grande et plus la granulation est grossière. L'entropie est faible si on a souvent les mêmes couples de pixels. C'est un indicateur de désordre. 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.

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

b. La variance

La variance (équation 3.2) mesure l'hétérogénéité de la texture. Elle augmente lorsque les niveaux de gris différent, de leur moyenne. La variance est indépendante du contraste.

$$VAR = \sum_i \sum_j \left((i - \mu)^2 P_{ij}(d, \theta) \right) \quad 3.2$$

c. Le contraste

Il peut être définie, comme le rapport entre les parties les foncées et les parties les plus claires d'une image. La valeur de ce paramètre (équation 3.3), est d'autant plus élevée, que la texture présente un fort contraste.

$$CST = \sum_i \sum_j \left((i, j)^2 P_{ij}(d, \theta) \right) \quad 3.3$$

d. L'énergie

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

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

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

e. La corrélation

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

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

f. Homogénéité

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

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

Ces indices, bien que corrélés, réduisent l'information contenue dans la matrice de cooccurrence et permettent une meilleure discrimination des textures. Le choix du vecteur de déplacement, et de la taille de la fenêtre du voisinage sur laquelle s'effectue la mesure sont les paramètres sur lesquels repose la réussite de la méthode. La difficulté à surmonter lors de l'application de cette technique réside justement dans le choix de ces paramètres car ils varient selon le type d'images et de textures [42][43].

3.3 Analyse de la forme [44][45][46]

La notion de la forme est très importante puisqu'elle nous permet d'identifier, les objets qui nous entourent. L'analyse de la forme n'est considérée réussie que si elle permet de décrire les objets de façon similaire à la perception humaine. Les signes de malignité dans le cas des masses et des microcalcifications sont différents et, indépendants. Le degré de malignité des microcalcifications, dépend de leur nombre dans un cluster, du contour que forme leur cluster (régulier, rond, ovale) et de leur forme (polymorphes, annulaires, arborescentes, arciformes, semi-lunaires, rhomboédriques, sédimentées). Alors que le degré de malignité dans le cas des masses augmente essentiellement, avec l'élévation de la complexité du contour.

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

Les caractéristiques que nous avons adoptées, ce sont celles qui nous paraissent les plus pertinentes, pour modéliser numériquement les régions d'intérêt mammaires. Elles permettent ainsi, d'identifier les pathologies mammaires, en se basant sur les descripteurs morphologiques.

Parmi ces descripteurs, nous allons étudier seulement quelques descripteurs géométriques, ces derniers permettent de caractériser l'aspect de la forme d'une façon assez globale et ils sont généralement indépendants du domaine d'application comme l'aire et le périmètre.

3.3.1 L'aire

Parmi les descripteurs de forme les plus répandus, on cite l'aire de la masse. Ce paramètre est calculé à partir, du nombre de pixels contenus dans une lésion.

3.3.2 La compacité

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

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

3.3.3 Le nombre de régions

Le nombre de régions, est comptabilisé à partir de l'image segmentée.

3.3.4 Les moments invariants de Hu [47]

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

A partir des moments géométriques, Hu [47] a proposé un ensemble de sept moments invariants aux translations, rotations et changement d'échelle. Les moments offrent un cadre théorique puissant pour résoudre des problèmes rencontrés, dans plusieurs applications d'imagerie y compris l'imagerie médicale.

Hu [36] a défini sept descripteurs invariants suivant ci-dessous:

$$\begin{aligned}
 \Phi_1 &= \eta_{20} + \eta_{02} \\
 \Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11} \\
 \Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{30} - 3\eta_{21})^2 \\
 \Phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\
 \Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
 & \quad (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
 \Phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
 \Phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]
 \end{aligned} \tag{3.8}$$

Ces descripteurs sont engendrés des moments;

$$\eta_{pq} = \frac{m'_{pq}}{m'^n_{00}} \tag{3.9}$$

où $n = (p + q) / 2$.

Avec m'_{pq} , représentant les moments centraux :

$$m'_{pq} = \sum_x \sum_y (x - x_g)^p (y - y_g)^q f(x, y) \tag{3.10}$$

Le centre de (x_g, y_g) représente le centre de gravité de l'image dont les coordonnées, sont définies par:

$$x_g = \frac{m_{10}}{m_{00}}, \quad y_g = \frac{m_{01}}{m_{00}} \tag{3.11}$$

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

Nous considérons dans le cadre de notre projet, la moyenne des descripteurs invariants.

3.4 Introduction à la classification supervisée d'images

La classification est considérée comme la dernière étape dans un système de diagnostic assisté par ordinateur (DAOx). Elle exploite le résultat de la description (qui lui-même exploite le résultat de la segmentation) pour pouvoir décider de la nature pathologique des régions d'intérêt.

Les méthodes de classification [48] [49], permettent de regrouper des objets, en classes d'objets plus homogènes. Les objets regroupés, ont des caractéristiques communes. Ils sont similaires mais, se distinguent clairement des objets, des autres classes.

La mise en œuvre d'une procédure de classification, ayant pour objectif de classer automatiquement des objets, comporte généralement deux phases fondamentales:

- une phase d'apprentissage dont le but, est de déterminer un espace de représentation des données et, de rechercher les paramètres discriminants capables, de caractériser chaque classe d'objets ;
 - une phase de reconnaissance au cours de laquelle, on attribue à une classe, chacun Des objets inconnus, dans l'espace de représentation déterminé durant l'apprentissage.
- La classification peut être supervisée ou non supervisée.

Les méthodes de classification, ont pour but d'identifier les classes auxquelles appartiennent des objets, à partir de certains traits descriptifs. Elles s'appliquent à un grand nombre d'activités humaines et conviennent en particulier, au problème de la prise de décision automatisée. La procédure de classification sera extraite, automatiquement à partir d'un ensemble d'exemples. Un exemple, consiste en la description d'un cas avec la classification correspondante. Un système d'apprentissage doit alors, à partir de cet ensemble d'exemples, extraire une procédure de classification. Il s'agit à cet effet, d'extraire une règle générale à partir des données observées. La procédure générée, devra classer correctement les exemples de l'échantillon et, avoir un bon pouvoir prédictif, pour classer correctement de nouvelles descriptions.

Les méthodes utilisées pour la classification sont nombreuses, citons: la classification par les réseaux de neurones, les machines à supports vecteurs (SVM), etc.

Nous étudions dans ce mémoire la classification connexionniste, qui montre de bonnes performances, dans la résolution de problèmes réels, tels que; la catégorisation de textes ou le diagnostic médical et, ce même sur des ensembles de données de très grandes dimensions. Dans le cas de notre projet, nous considérons l'étude de la classification supervisée.

La classification supervisée, demande à l'utilisateur d'instruire le système, en désignant des zones de l'image, comme étant des échantillons représentatifs des classes à extraire. La classification est donc, précédée d'un apprentissage. Pour instruire le système, il faut avoir une bonne connaissance du terrain observé. Pour cela, il faut disposer d'une vérité terrain. C'est un outil pour réaliser un bon apprentissage, ainsi que pour valider une classification. Il est suffisant de disposer d'une vérité terrain sur une portion de l'image où les types de chaque région, sont représentés [48][49].

La classification supervisée, s'organise en quatre étapes:

- **espace des états** : on extrait de l'image, des critères jugés suffisamment discriminants pour la classification. Cet ensemble définit un espace, dont la dimension est égale au nombre de critères extraits;
- **zones d'apprentissage** : on définit sur l'image des zones d'apprentissage, qui sont des régions de l'image que l'on juge représentatives d'une certaine classe;
- **apprentissage** : il consiste, pour une classe donnée, à rassembler l'ensemble des pixels situés dans les zones d'apprentissage et, d'étudier la répartition de leurs vecteurs associés dans l'espace des états.
- **classifieur**: il utilise l'information issue de l'apprentissage, pour attribuer à chaque point, ou à chaque région de l'image, une catégorie ou une classe

3.5 Réseaux de Neurones [48-49][50-51]

On considère généralement qu'un réseau de neurones, est constitué d'un grand ensemble d'unités (ou neurones), ayant chacune une petite mémoire locale. Ces unités, sont reliées par des canaux de communication (les connexions, aussi appelées synapses d'après le terme biologique correspondant), qui transportent des données numériques. Les unités, peuvent uniquement agir sur leurs données locales et, sur les entrées qu'elles reçoivent par leurs connexions.

3.5.1 Neurone biologique

Le cerveau est composé de plusieurs cellules que l'on peut diviser en deux grands groupes : d'un côté, les neurones et de l'autre, l'ensemble des cellules de support que l'on nomme 'neuroglia'. Il n'y a que les neurones qui soient des éléments actifs dans le cerveau, les cellules gliales ne servent qu'à supporter physiquement les neurones. Elles agissent aussi comme isolateur électrique, elles aident au métabolisme et, fournissent des protéines aux neurones (cf. Figure 3.7).

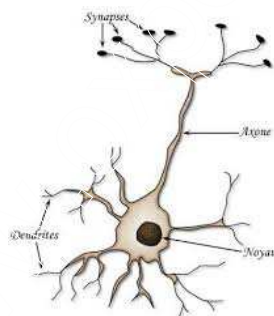


Figure 3. 6 : neurone biologique [49]

3.5.2 Neurone artificiel

Un réseau de neurones artificiel, est un ensemble de neurones interconnectés. Le neurone est un élément contenant plusieurs entrées et, une sortie. Chaque entrée x_i est pondérée par un poids w_{ki} , avant d'être sommé pour produire v_k , k étant un indice, indiquant Le numéro du neurone. Le résultat v_k est soumis à, une fonction de transfert $\varphi(.)$ non linéaire (cf. Figure 3.8).

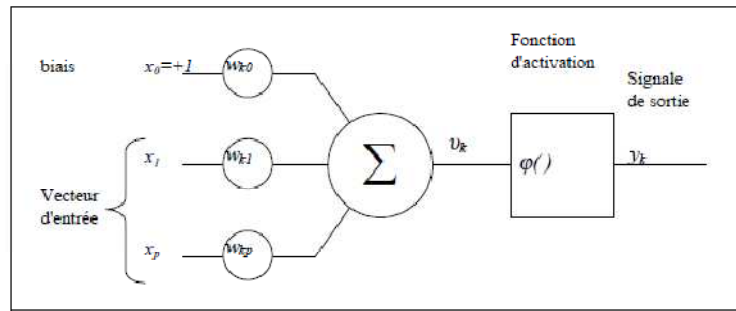


Figure 3.7 : représentation schématique d'un neurone artificiel [49]

3.5.3 Principe de fonctionnement [49]

L'information ne se transmet que dans un seul sens: des dendrites vers l'axone. Le neurone va donc recevoir des informations, venant d'autres neurones, grâce à ses dendrites. Il va ensuite y avoir une sommation, au niveau du corps cellulaire; de toutes ces informations et via, un signal électrique. Le résultat de l'analyse, va transiter le long de l'axone, jusqu'aux terminaisons synaptiques. Il va y avoir, libération des neurotransmetteurs (médiateurs chimiques) dans la fente synaptique. Le signal électrique, ne pouvant pas passer la synapse, les neurotransmetteurs permettent donc le passage des informations, d'un neurone à un autre. Les neurotransmetteurs excitent ou inhibent, le neurone suivant et, peuvent ainsi générer ou interdire, la propagation d'un nouvel influx nerveux. Les synapses possèdent une sorte de «mémoire» qui, leur permet d'ajuster leur fonctionnement. En fonction de leur «histoire», c'est-à-dire de leur activation répétée ou non, entre deux neurones, les connexions synaptiques vont se modifier: c'est le mécanisme d'apprentissage.

Il existe plusieurs types de réseaux de neurones, nous étudions dans le cadre de ce mémoire, le perceptron multicouche.

3.5.4 Le modèle du perceptron [49] [50] [51]

a. Perceptron simple

Le perceptron simple, permet de classifier des éléments compris dans deux catégories différentes; à condition que ces deux catégories, soient linéairement séparables. Par la suite, il est également possible de prédire, à quelle catégorie appartient un élément.

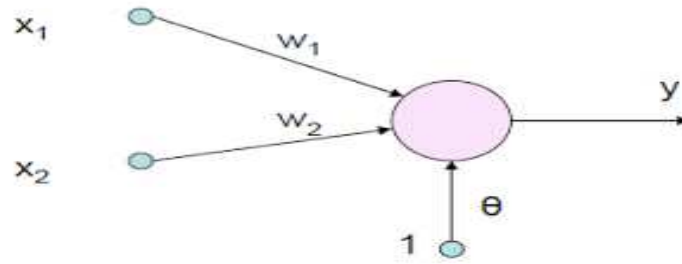


Figure 3.8 : perceptron simple [50]

b. Perceptron multicouche (PMC)

Malgré son nom quelque peu barbare, le perceptron multicouche, est sans doute le plus simple et le plus connu des réseaux de neurones. Il s'agit d'un classifieur linéaire de type réseau neuronal formel organisé en plusieurs couches, au sein desquelles une information, circule de la couche d'entrée vers la couche de sortie, uniquement. Chaque couche, est constituée d'un nombre variable de neurones, les neurones de la couche de sortie correspondant toujours aux sorties du système. La structure du (PMC) est relativement simple (cf. Figure 3.9):

- ✓ **Les nœuds d'entrée :** la première couche, est appelée couche d'entrée. Elle recevra les données source, que l'on veut utiliser pour l'analyse. Dans le cas de l'aide au diagnostic médical, cette couche recevra les symptômes. Sa taille est donc directement déterminée, par le nombre de variables d'entrées.
- ✓ **Les nœuds cachés :** La seconde couche est une couche *cachée*, en ce sens qu'elle n'a qu'une utilité intrinsèque pour le réseau de neurones et, n'a pas de contact direct avec l'*extérieur*. Les fonctions d'activations, sont en général non linéaires sur cette couche, mais il n'y a pas de règle à respecter. Le choix de sa taille n'est pas implicite et doit être, ajusté. En général, on peut commencer par une taille moyenne des couches d'entrée et de sortie mais ce n'est pas toujours, le meilleur choix. Il sera souvent préférable, pour obtenir de bon résultats, d'essayer le plus de tailles possibles.

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

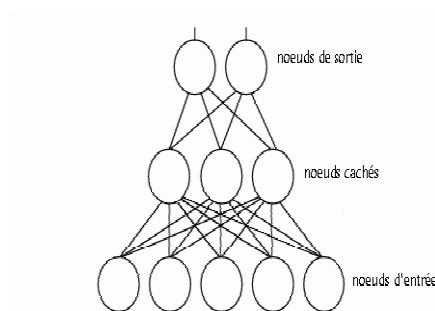


Figure 3.9: perceptron multicouche [50]

3.5.5 Fonction d'activation [49]

Comme indiqué précédemment, un perceptron multicouche (MLP) est un réseau de neurones avec une architecture « **feed-forward** » et des connexions unidirectionnelles complètes entre les différentes couches. Toutefois, ce n'est pas suffisant pour définir les propriétés d'un réseau. Outre les architectures de réseaux, les neurones d'un réseau possèdent des fonctions d'activation, qui vont transformer les signaux émis par les neurones de la couche précédente à l'aide, d'une fonction mathématique. Cette fonction f , est la fonction d'activation elle-même, qui, peut profondément influencer sur la performance du réseau. Il est donc important de bien choisir ce type de fonction d'activation des neurones, dans un réseau de neurones. Les neurones d'entrée, ne possèdent généralement aucune fonction d'activation.

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

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

- ✓ Les paramètres sont attachés à la non-linéarité du neurone et ils interviennent directement, dans la fonction f .

✓

3.5.6 Fonctionnement d'un réseau de neurones

Les réseaux de neurones artificiels sont construits sur une architecture semblable, en première approximation, à celle du cerveau humain. Le réseau reçoit les informations sur une couche réceptrice de « neurones », traite ces informations avec ou sans l'aide d'une ou plusieurs couches « cachées » contenant un ou plusieurs neurones et produit un signal (ou plusieurs) de sortie. Chaque neurone, qu'il appartienne à la première couche (réceptrice), aux couches cachées ou à la couche de sortie, est lié aux autres neurones par des connexions (similaires aux synapses du cerveau) auxquelles sont affectés des poids (eux-mêmes assimilables aux potentiels synaptiques) [49].

a. Calcul des poids synaptiques

La rétropropagation (backpropagation) est une méthode de calcul des poids pour un réseau à apprentissage supervisé qui consiste à minimiser, l'erreur quadratique de sortie. La rétropropagation consiste à rétropropager l'erreur commise par un neurone à ses synapses et aux neurones, qui y sont reliés. Pour les réseaux de neurones, on utilise habituellement la rétropropagation du gradient de l'erreur, qui consiste à corriger les erreurs selon l'importance des éléments qui ont justement participé à la réalisation de ces erreurs : les poids synaptiques qui contribuent à engendrer une erreur importante se verront modifiés de manière plus significative que les poids qui ont engendré une erreur marginale [50].

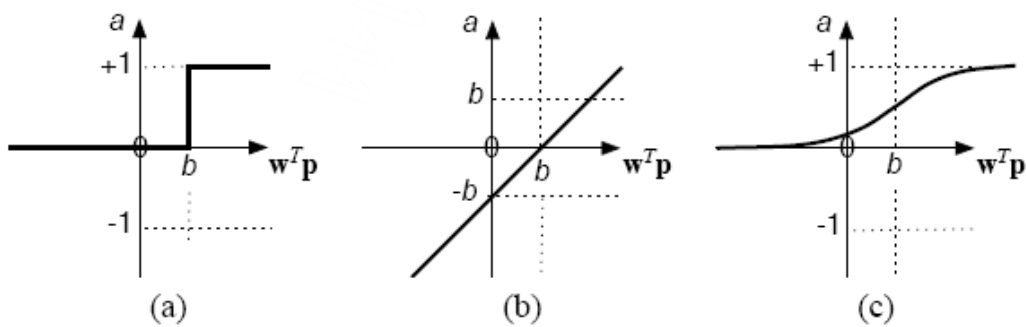
b. Fonctions de transfert

Une fonction de transfert calcule la valeur de l'état du neurone. C'est cette valeur qui sera transmise aux neurones avals. Il existe de nombreuses formes possibles pour la fonction de transfert. Différentes fonctions de transfert pouvant être utilisées comme fonction d'activation du neurone sont énumérées au tableau.3.1.

Nom de la fonction	Relation d'entrée/sortie	Icône	Nom Matlab
seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlim
seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlims
linéaire	$a = n$		purelin
linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$		satlin
linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$		satlins
linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$		poslin
sigmoïde	$a = \frac{1}{1+\exp^{-n}}$		logsig
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
compétitive	$a = 1$ si n maximum $a = 0$ autrement		compet

Tableau 3.1 : fonctions d'activation du neurone [50]

Les plus courantes sont les fonctions seuil (en anglais «hard limit»), linéaire et sigmoïde (figure 3.9).



(a) neurone à seuil ; (b) neurone linéaire ; (c) Neurone sigmoïde

Figure 3.10 : Fonctions de transfert [50]

3.5.7 Avantages et inconvénients

Parmi les avantages des réseaux, nous citons :

- ✓ précision du classifieur (si bien paramétré);
- ✓ apprentissage automatique des poids;
- ✓ possibilité de faire le parallélisme (les éléments de chaque couche peuvent fonctionner en parallèle);
- ✓ résistance aux pannes (si un neurone ne fonctionne plus, le réseau ne se perturbe pas).

Les inconvénients se résument suivant;

- ✓ la complexité de l'architecture du réseau;
- ✓ paramètres difficiles à interpréter (boîte noire);
- ✓ difficulté du paramétrage, surtout pour le nombre de neurones, dans la couche cachée.

3. 6 Introduction aux algorithmes génétiques [52][53][54]

Les algorithmes génétiques sont des algorithmes d'exploration fondés sur les mécanismes de la sélection naturelle et de la génétique. Ils utilisent à la fois les principes de la survie des structures les mieux adaptées, et les échanges d'information pseudo-aléatoires, pour former un algorithme d'exploration qui possède certaines des caractéristiques de l'exploration humaine. Ils sont une transcription directe du processus naturel d'évolution au monde de l'informatique. Les possibilités espérées des mécanismes d'évolution ont conduit quelques chercheurs des années 1950 à vouloir les simuler pour les appliquer à l'ingénierie. Mais ces travaux n'ont pas été probants en raison des connaissances insuffisantes, à l'époque, de la génétique naturelle et aussi en raison des faibles performances des calculateurs alors disponibles. Mais depuis le début des années 60, la quantité de travaux sur les algorithmes évolutionnaires suit une croissance exponentielle. En effet, les AGs ont des propriétés qui en font des candidats de choix pour résoudre des problèmes d'optimisation dans lesquels la taille de l'espace de recherche est importante, où les paramètres interagissent de manière

complexe et où très peu d'informations sur la fonction à optimiser sont disponibles. Ces propriétés ont motivé leur application en vision et traitement d'images [52].

3.6.1 La fonction d'évaluation [52]

Elle prend en argument l'individu et lui, associe un coût de performance. Cette fonction doit guider l'algorithme vers l'optimum en réalisant implicitement une pression de sélection dans cette direction. De plus, l'efficacité d'un AG s'appuie pour une grande part, sur la qualité de cette fonction. La mise au point d'une bonne fonction d'adaptation doit respecter plusieurs critères qui se rapportent à sa complexité et à la satisfaction, des contraintes du problème. Lorsque la fonction d'adaptation apparaît excessivement complexe, consommant une importante puissance de calcul, la recherche d'une approximation plus simple est souhaitable.

➤ *Qualité, fitness, d'un individu*

Le calcul de la qualité d'un individu est essentiel aux algorithmes génétiques. Cette fonction donne, en valeur numérique, la qualité d'un individu. C'est selon cette valeur numérique que sont calculées, les chances de sélection de cet individu. Les algorithmes génétiques étant une technique d'optimisation, cherchent la qualité maximale, donc l'optimisation de la fonction de qualité.

3.6.2 Les opérateurs génétiques

Ces opérateurs, définissent la manière dont les individus, se recombinent et s'agencent pendant la phase de reproduction. Les opérateurs qu'on retrouve le plus souvent sont: la sélection, le croisement et la mutation.

a. Sélection: cette opération est fondée sur le principe d'adaptation de chaque individu d'une population à son environnement, suivant la théorie de la sélection naturelle introduite par Charles Darwin [53]. Ainsi, seuls les individus les plus aptes à répondre à certains critères seront sélectionnés à survivre et à se multiplier. La fonction d'évaluation de l'adéquation des individus évalue chaque chromosome de la population, et seuls ceux dont la qualité sera suffisante seront conservés d'une génération à une autre. D'autres individus prendront la place des plus mauvais. N' individus seulement sur N ($N > N'$) sont remplacés dans la population pour obtenir la nouvelle génération, N : étant la taille de la population. Le nombre de

remplacements ayant été fixé arbitrairement par l'utilisateur. Les (N-N') individus qui survivent sont parmi les meilleurs. Il s'agit d'une stratégie dite élitiste. Les N' autres individus de la génération suivante sont obtenus suite à une sélection en fonction de la valeur d'adaptation de chaque individu et application des opérateurs de reproduction.

Dans la littérature nous trouvons plusieurs méthodes de sélection. Dans ce qui suit, nous allons décrire trois méthodes:

b. La roulette: à chacun des individus de la population est associé un secteur d'une roue. L'angle du secteur étant proportionnel à la qualité de l'individu qu'il représente. Les tirages des individus sont ainsi pondérés par leur qualité. D'une manière logique, les meilleurs individus ont plus de chance d'être sélectionnés.

c. La sélection par le rang: il faut trier la population en fonction de la qualité des individus puis leur attribuer à chacun un rang. Les individus de moins bonne qualité obtiennent un rang faible (à partir de 1). La suite de la méthode, consiste uniquement en l'implémentation d'une roulette basée sur les rangs des individus. L'angle de chaque secteur de la roue sera proportionnel au rang de l'individu qu'il représente.

d. La sélection par tournoi: un tournoi consiste en une rencontre entre plusieurs individus, pris au hasard dans la population. Le vainqueur du tournoi est l'individu de meilleure qualité.

e. Croisement: le croisement consiste à combiner deux individus quelconques (dits parents) pour en ressortir deux autres individus (dits enfants) pas forcément meilleurs que les parents. Il existe plusieurs variantes de cet opérateur (figures 3.11 et 3.12), mais en général il consiste à couper en un ou plusieurs points deux individus (aux mêmes endroits dans les deux individus) et à échanger les parties situées entre ces points. Ces derniers sont générés aléatoirement. Une autre variante est le croisement uniforme qui utilise une chaîne de bits générée aléatoirement et de même longueur que les individus; appelée « masque ». Les gènes des individus initiaux sont échangés en fonction de cette chaîne aléatoire, lorsque le bit correspondant vaut ' 1 '. Le croisement uniforme (figure 3.13) de permutations a les avantages de la simplicité et d'une bonne efficacité.

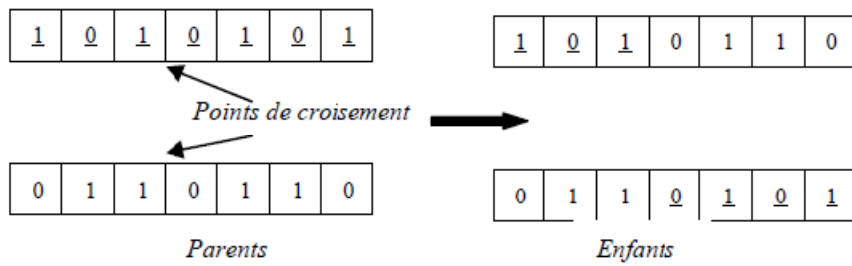


Figure 3.11 : exemple du croisement à un point [54]

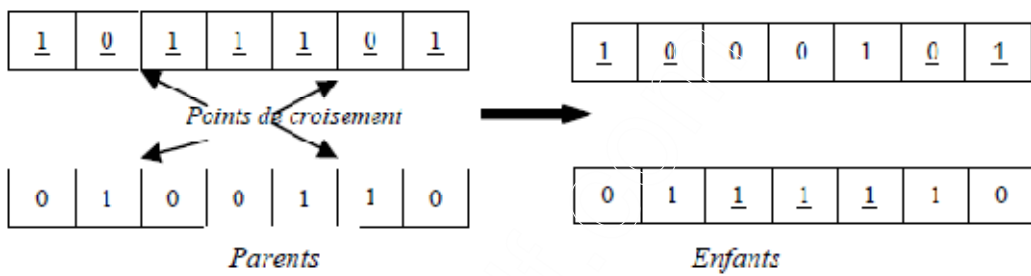


Figure 3.12 : exemple du croisement à deux points [54]

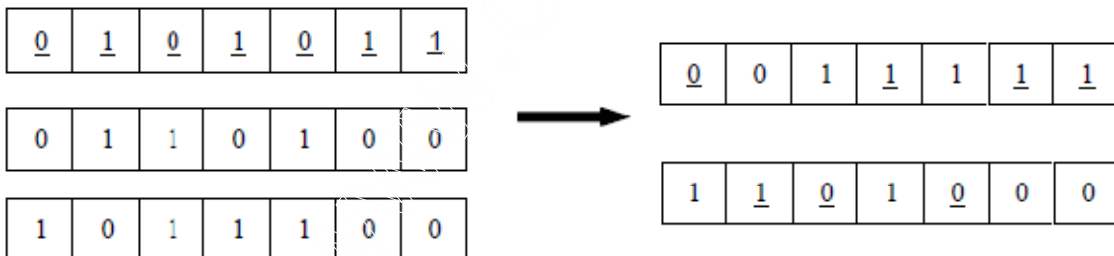


Figure 3.13 : exemple du croisement uniforme [54]

f. Mutation: cet opérateur agit sur un individu. Il consiste à choisir d'une manière aléatoire un ou plusieurs gènes (bits) et à, modifier leurs valeurs. Chaque bit d'un individu a une probabilité, pour qu'il subisse une mutation.

3.6.3 Critère d'arrêt

Les étapes d'évaluation reproduction/sélection sont appliquées d'une manière itérative sur la population, jusqu'à satisfaction d'une condition Fin ; cette condition est soit un nombre maximum de générations, soit une valeur de fitness minimale, soit une convergence vers la meilleure solution. Au cours de son évaluation, la population tend à converger, c.-à-d., que les individus tendent à se ressembler de plus en plus. Quand la population s'est uniformisée en grande partie, les individus fournissent une bonne approximation d'un optimum du problème. Si cet optimum n'est pas toujours l'optimum global, c'est généralement un optimum local proche (en qualité) de celui-ci.

3.6.4 Principe des algorithmes génétiques

Le principe des algorithmes génétiques est simple (figure 3.14). Il est basé sur trois phases:

- a. **La genèse:** il s'agit de l'initialisation aléatoire d'individus pour la population de la première génération.
- b. **La reproduction:** il s'agit de l'évolution des individus de la génération courante vers la suivante.
 - ✓ la sélection des individus reproducteurs.
 - ✓ le croisement génétique de ces individus pour la création de nouveaux individus.
 - ✓ la mutation de certains individus pour que le pool génétique ne s'affaiblisse pas.
 - ✓ l'évaluation des individus par le calcul de leur fitness.
- c. **Recueil du meilleur individu:** recherche de l'individu le plus adapté, selon les critères souhaités.

Les sous étapes 2 et 3 se répètent autant de fois qu'il y a besoin de générations (itérations de l'algorithme) pour satisfaire, un critère d'arrêt. Celui-ci est défini avant que le processus ne commence. La solution est alors représentée par le meilleur individu de la dernière génération [54].

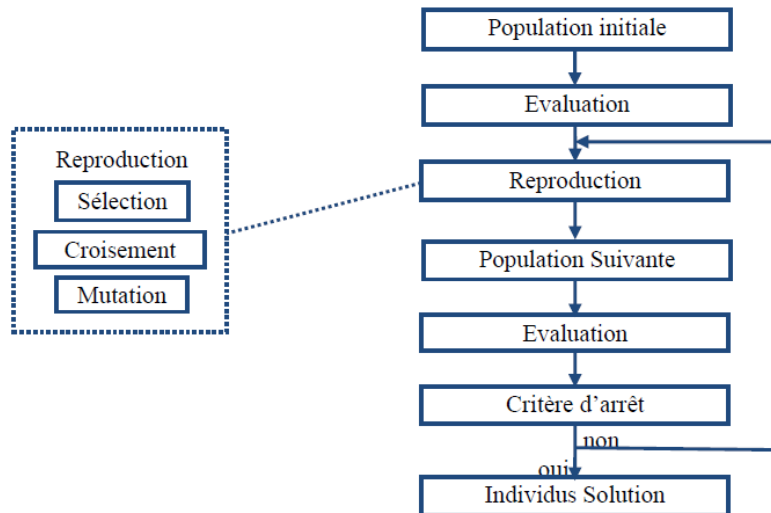


Figure 3.14 : schéma du principe des algorithmes génétiques [54]

3.7 Classification neuro-génétique [55]

On sait que ce qui caractérise un réseau de neurones, ce sont ses poids (W_i) et ses biais (B_i), c'est cela qui relie toutes les entrées à tout le neurone, jusqu'à la couche de sortie.

Dans un RN classique, on intègre les entrées dans la couche primaire et, l'apprentissage se fait avec un certain nombre d'itérations et de tests. L'arrêt s'effectue, lorsque l'erreur quadratique moyenne est basse et, donc le taux de réussite augmente. Cependant, le défaut est, à chaque fois, qu'on recommence l'entraînement, on a un résultat différent; c'est donc un processus aléatoire, qui ne donne pas toujours un très bon résultat, surtout quand les classes sont difficilement séparables.

Avec la classification neuro-génétique (figure 3.15) adoptée dans notre projet, lorsqu'on spécifie le nombre de neurones, le réseau est automatiquement créé. Les biais et les poids sont initialisés aléatoirement.

À la première itération: on fait une comparaison des entrées avec les sorties et, on voit si les biais et les poids sont ajustés. Évidemment, comme ils ont été initialisés aléatoirement, ils sont obligatoirement faux, donc, c'est là que les AG prennent leur importance, avec l'évaluation de ses individus (W_i , B_i) on peut effectuer l'opération de sélection, croisement, mutation et, une nouvelle fois, l'évaluation. À partir de là, une nouvelle population est créée et, à nouveau évaluée. Le réseau est de nouveau, évolué suivant les opérations de sélection,

croisement et, mutation et, ainsi de suite, avec des dizaines, voire des centaines d'itérations.

A la fin, on obtient les poids et biais optimums, ce qui nous donne un réseau de neurone proche de la perfection, sans avoir à faire une multitude d'entraînements, sans intérêt.

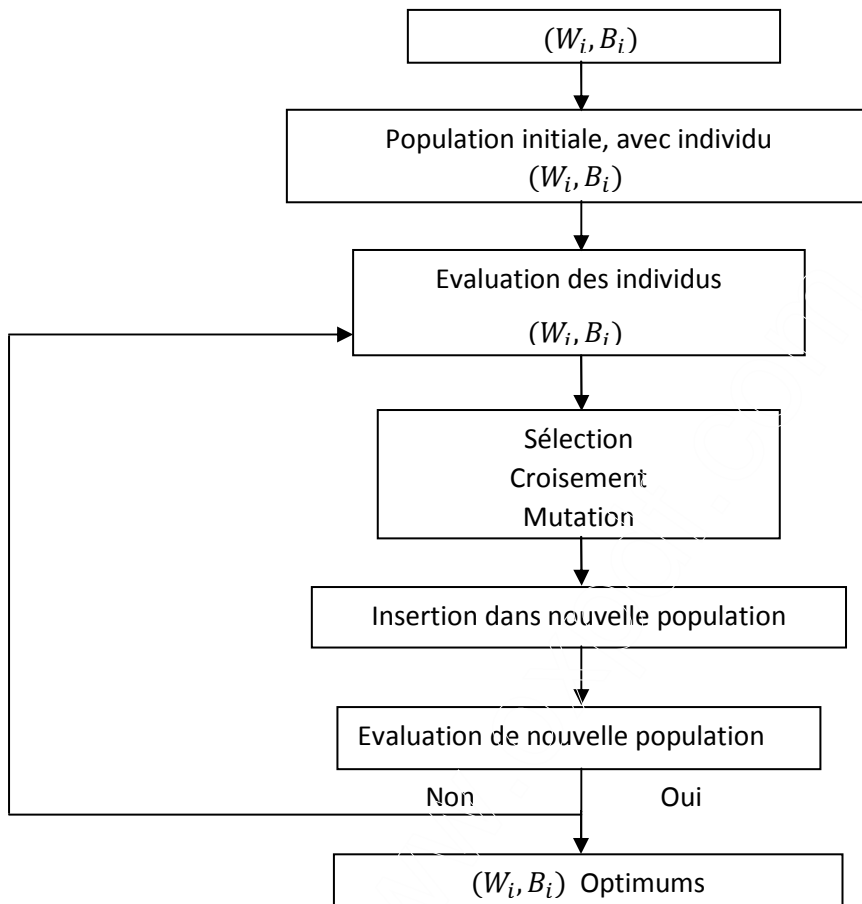


Figure 3.15: organigramme de génération des poids et des biais [55]

La classification consiste alors, à prendre en compte, les individus du perceptron évalués par l'algorithme génétique, afin de classifier les masses et les calcifications, suivant des cas bénins et malins. Dans ce cadre, deux classifieurs neuro-génétiques, sont réalisés, permettant de discriminer les masses et les calcifications.

3.8 Conclusion

Ce chapitre a fait l'objet d'une analyse texturale, qui se base sur les méthodes statistiques, pour évaluer différents paramètres qui nous informent sur la distribution des niveaux de gris

ainsi que l'aspect des pathologies mammaires. Une description morphologique, a également été prise en compte suivant la compacité, la surface et d'autres attributs, afin d'identifier le changement de la forme, dans le cas des images mammographiques. La discrimination de masses et de calcifications, est réalisée à partir des vecteurs d'attributs caractéristiques par une combinaison, entre le perceptron multicouche et l'algorithme génétique.

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

Introduction générale

Le cancer du sein est devenu le cancer le plus fréquent chez la femme algérienne. Il touche le plus souvent des mères de famille. Avec plus de 1000 nouveaux cas diagnostiqués [1] par an (et probablement autant qui doivent rester méconnus); il devient un réel problème de santé publique qui menace la société toute entière du fait, qu'il affecte et bouleverse des familles entières en les privant encore trop souvent de leur noyau maternel et ce, en dépit des remarquables progrès qui ont lieu en matière de traitement. Il devient tellement fréquent que l'on a l'impression d'assister à une réelle épidémie.

Certes, le cancer du sein résulte de l'intervention de multiples facteurs et son dépistage à un stade infraclinique, demeure le moyen le plus sûr et le plus efficace de réduire la fréquence des formes évoluées et donc des formes les plus graves. Le dépistage permet de saisir la maladie à un stade où, elle est encore parfaitement maîtrisable et, totalement curable. De nombreuses études scientifiques attestent que le risque d'essaimage métastatique est lié à la taille de la tumeur [1]. Dans ce cadre, les malades dépistées à temps guérissent définitivement de leur maladie dans la grande majorité des cas.

En effet, la mammographie est l'outil le plus utilisé pour une détection précoce de la maladie, car elle permet de produire des images de haute résolution du sein. Elle permet de mettre en évidence des anomalies telles que les masses et les calcifications qui peuvent traduire, des lésions malignes. Cependant tous les radiologues reconnaissent la difficulté de l'interprétation des mammographies qui s'accroît encore par le type de tissu, du sein examiné.

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

Notre objectif est alors, d'extraire les différentes structures composant l'aspect pathologique du sein via une segmentation par les modèles déformables, de les caractériser

selon une analyse texturale et morphologique et, de classifier les différentes pathologies rencontrées selon, le réseau de neurones en coopération avec les algorithmes génétiques, pour l'aide au diagnostic clinique.

Ce mémoire est rédigé suivant l'organisation suivante ;

- Le premier chapitre introduit le contexte mammographique ainsi, que quelques notions sur le cancer du sein et l'imagerie mammaire.
- Le second chapitre est consacré aux principales techniques du prétraitement et de la segmentation d'images, comme, l'analyse d'histogramme, les opérateurs de la morphologie mathématique et les contours actifs.
- Le troisième chapitre appréhende l'analyse morphologique et texturales et, spécifie les paramètres choisis pour la caractérisation des lésions mammographiques ainsi que la classification connexionniste et son association avec les algorithmes génétiques pour la prise de décision.
- Le quatrième chapitre est consacré à la présentation générale de la méthodologie adoptée ainsi qu'à, son application sur des images mammographiques d'une base de données réelles.

Chapitre 4

Mise en œuvre du système sur

les images mammographiques et discussion des résultats

4.1 Introduction

L'objectif de ce chapitre, est de présenter l'application développée, les traitements illustrés par des images mammographiques dans le but de l'extraction et de la classification des pathologies mammaires issues de la base de données MIAS (Mammographic Image Analysis Society) [56], suivant le perceptron multicouche amélioré par les processus génétiques. 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 images étudiées, comprennent, des masses ainsi que des calcifications. Elles sont réparties suivant, des cas bénins et malins. La zone suspecte, est identifiée par des radiologues experts [56].

4.2 Environnement de travail

Notre système a été développé sous environnement Windows 7 édition intégrale à l'aide de l'outil de programmation Matlab de la version 8.1(R1013a).

4.2.1. Matériel utilisé

L'implémentation de nos système a été réalisée grâce à un PC portable de type hp Intel(R) Core(TM) i3-2370M CPU @ 2.40GHz 2.40 GHz, et 4 Go RAM.

4.2.2 Présentation du langage utilisé MATLAB

MATLAB "MATrix LABoratory" est un logiciel commercial de calcul interactif, un langage de programmation extrêmement simple à utiliser. Il comprend de nombreuses fonctions graphiques, un système puissant d'opérateurs s'appliquant à des matrices, des algorithmes et numériques. L'aspect modulaire est l'un des plus grands atouts de MATLAB: l'utilisateur peut lui-même définir ses propres fonctions, en regroupant des instructions MATLAB dans un fichier portant le suffixe ".m". La syntaxe est bien plus abordable que dans les langages classiques et devrait, éliminer les réticences habituelles des programmeurs débutants pour écrire des fonctions. En termes de vitesse d'exécution, les performances sont inférieures à celles obtenues avec un langage de programmation classique. L'emploi de MATLAB devrait être restreint à des problèmes peu gourmands en temps calcul, mais dans la plupart des cas, il présente une solution élégante et rapide à mettre en œuvre [57].

Matlab fonctionne suivant deux modes:

- **Le mode interactif:** exécute des instructions intégrées, par l'utilisateur dans la fenêtre de commande de matlab (le workspace);
- **le mode exécutif:** exécute le programme ou script, qui sont des fichiers regroupant, une suite d'instructions (les fichiers .m).

a. Fonctionnalités du Matlab

Une image Matlab est une matrice bidimensionnelle de valeurs entières ou réelles. Les principales fonctions de traitement d'images sous Matlab, se trouvent dans la boîte à outils (toolbox) image processing (traitement d'images). L'aide sur cette boîte à outils est obtenue, en tapant help images en ligne de commande de Matlab. L'aide sur une commande particulière est ensuite, obtenue en tapant help suivi du nom de la commande.

Il existe plusieurs commandes et fonctions de MATLAB que l'on pourra trouver dans les différentes boîtes à outils (toolbox) qui viennent compléter, le logiciel de base. On présente ci-dessous, quelques commandes intéressantes:

1. Gestions des commandes et des fonctions

- help : aide.
- lookfor : recherche d'une entrée dans le help.
- which : localise les fonctions et fichiers.
- demo : lance la démonstration.
- info : information sur MATLAB et The MathWorks.

2. Gestion des variables et de l'espace de travail

- who : affiche les variables courantes.
- whos : affiche les variables courantes, format long.
- save : sauve l'espace de travail sur disque.
- load : restaure l'espace de travail à partir du disque.
- clear : efface les variables et fonctions de la mémoire.

b. Présentation de la boîte à outil GUIDE

Les versions actuelles de Matlab permettent, de construire des interfaces directement avec la souris grâce au GUIDE (Graphical User Interface Development Environment). Cet outil est capable de construire, des applications de très haut niveau.

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 (figure 4.1).

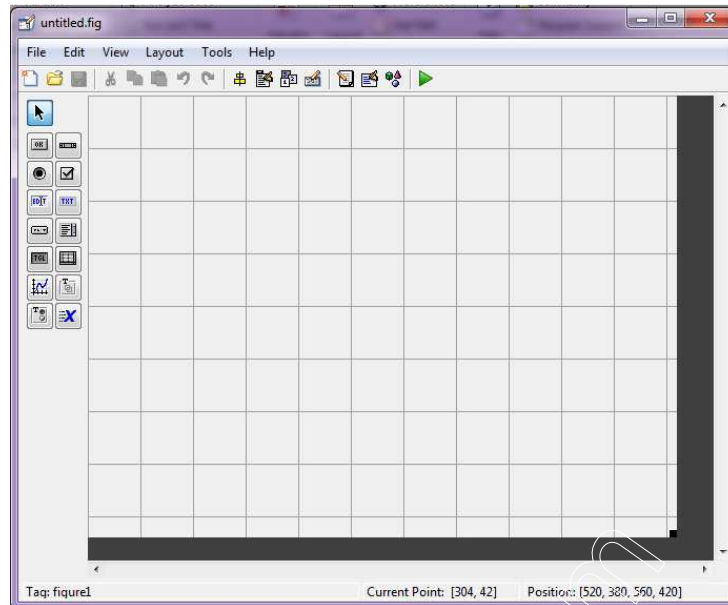


Figure 4.1: fenêtre principale du guide

4.3 Présentation du système développé

Le système (figure 4.2) que nous avons développé a pour objectif final, de classifier les masses et les calcifications mammographiques moyennant un classifieur neuro-génétique (NG), qui utilise le perceptron multicouche dont, les paramètres neuronaux sont ajustés par les algorithmes génétiques, pour réduire les confusions. L'identification de ces zones s'est faite, par les étapes suivantes:

1. rehaussement du contraste par, ajustement de l'histogramme et chapeau haut de forme, afin de faciliter la détection des masses et des calcifications sur le mammographe.
2. Segmentation par le modèle de Chan- Vese, afin de détecter les régions d'intérêt.
3. Caractérisation texturale et morphologique des masses et des calcifications.
4. Classification neuro-génétique **NG par** un cascavage entre deux classifieurs les réseaux de neurones et les algorithmes génétiques, pour discriminer les cas malins et bénins des masses et des calcifications ».

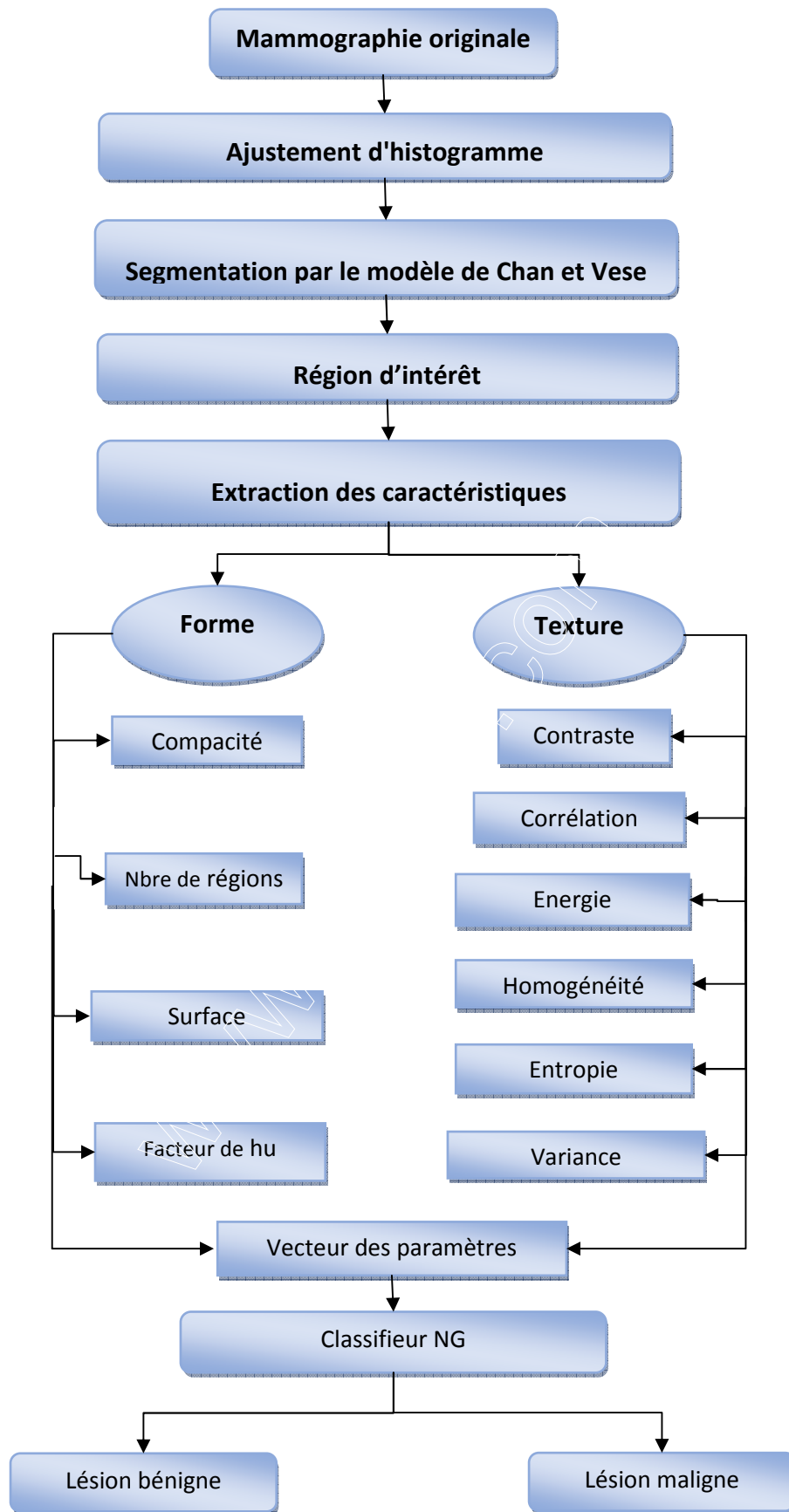


Figure 4.2: système d'aide à la décision mammographique "NGMAMM "

4.3.1 La base d'image

Dans ce travail, nous utilisons la base d'image mini-Mias (Mammography Image Analysis Society). MIAS est une organisation des groupes de recherches sur les mammographies du royaume uni (United Kingdom) qui a développé une base de données de clichés mammographiques numériques avec une résolution spatiale de 1024*1024 pixels pour chaque image, cette base contient 322 images sont répartie en 207 images normales, 85 images contient des masses et 30 images contient des calcifications.

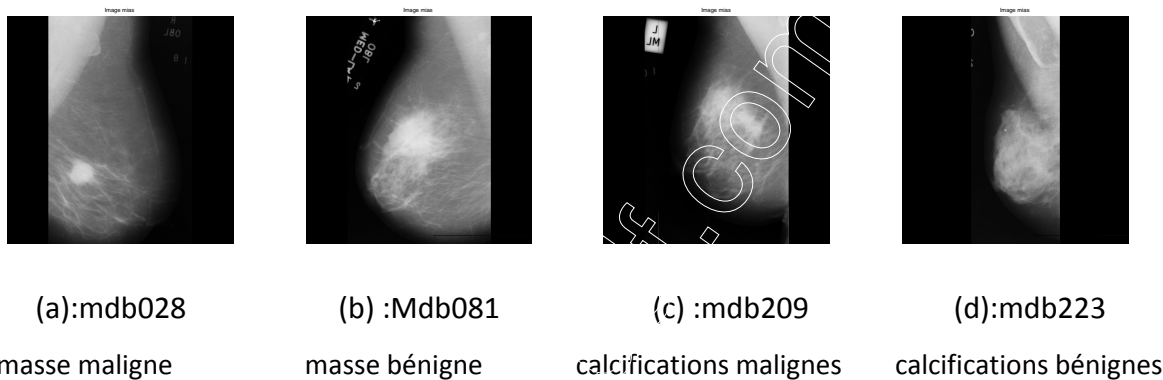


Figure 4.3: image mammographiques de la base MIAS [56]

4.3.2 Interface du système

L'interface principale de notre application, le permet plusieurs traitements suivant les divers menus selon la figure 4.4.

- Fichier: ce module permet d'ouvrir les images.
- Prétraitement d'image: opération d'ajustement de l'histogramme et traitement de l'image par chapeau haut de forme.
- La région d'intérêt : extraction de la zone d'intérêt selon les coordonnées du radiologue.
- Segmentation: exécution de l'algorithme de la segmentation des modèles déformables.
- Caractérisation: extraction des caractéristiques des régions d'intérêt suivant les paramètres de texture et de forme.
- Classification: classification des types des tumeurs.

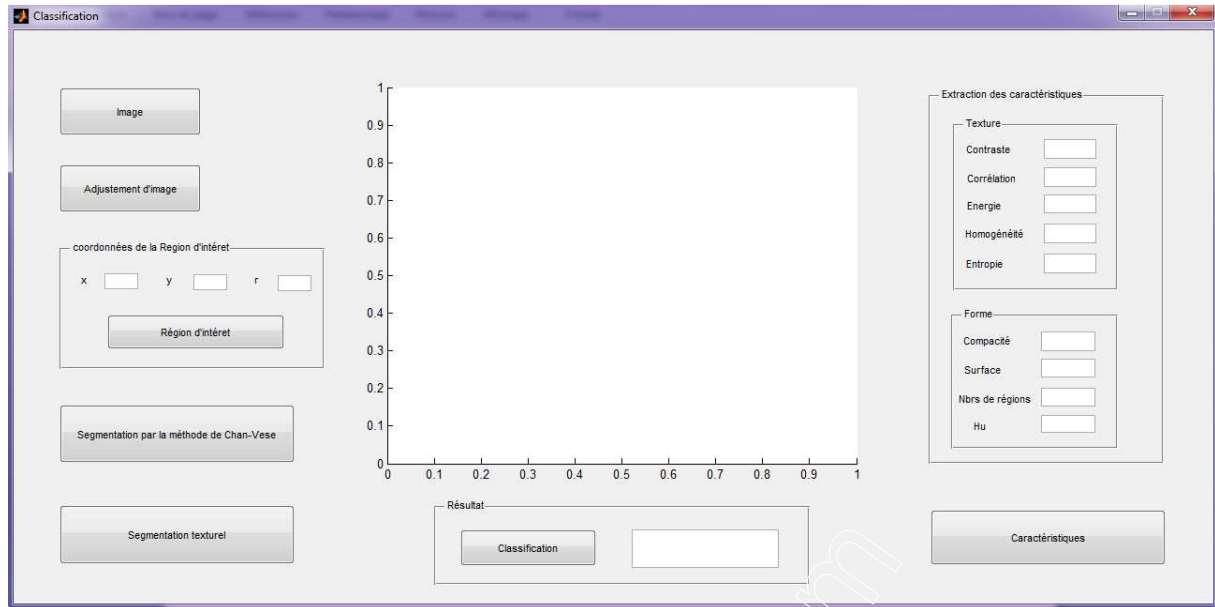


Figure 4.4: interface de l'application "NGMAMM"

4.4. Résultats et interprétation

L'application 'NGMAMM' parcourt l'analyse texturale et morphologique des pathologies mammographiques, en vue d'une identification de ces données, pour l'aide à décision.

4.4.1 Détection des masses

La détection des masses s'est faite, par ajustement d'histogramme pour l'obtention de la forme de la masse et la segmentation, par le modèle de Chan et Vese pour l'isolation de la région.

La taille de la masse identifiée par un médecin spécialiste a une importante influence sur le choix du nombre d'itérations. Les figures 4.5 et 4.6, illustrent respectivement la détection de masses bénignes et malignes.

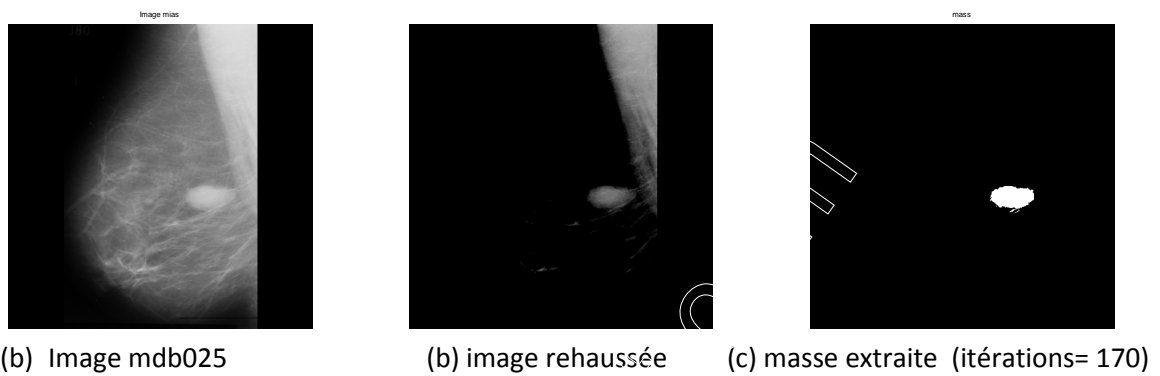


Figure 4.5: détection des masses bénignes par le modèle de Chan et Vese

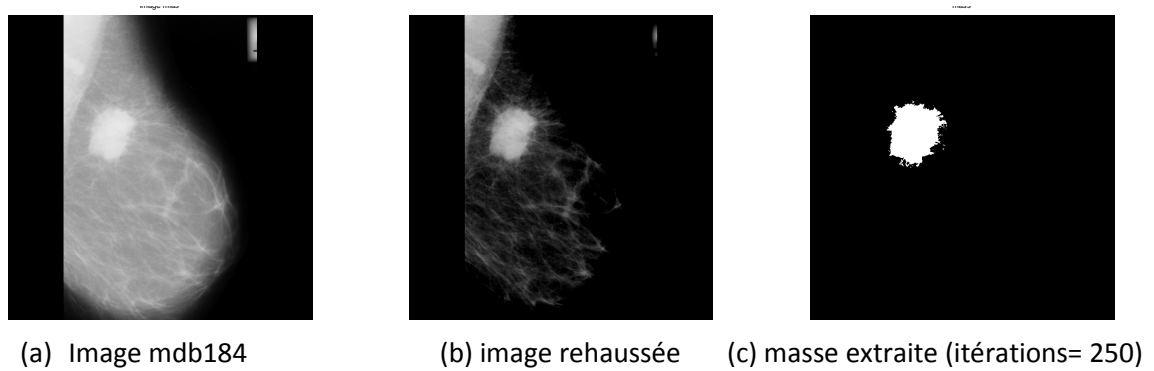
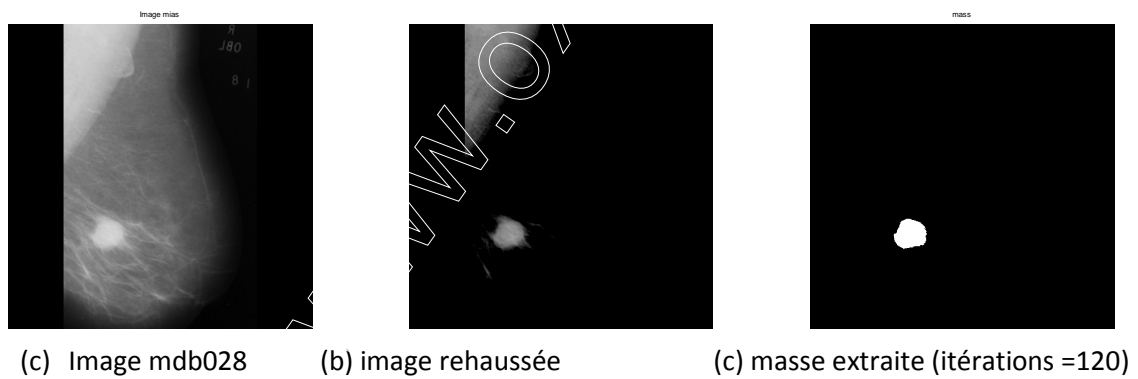


Figure 4.6: détection des masses malignes par le modèle de Chan et Vese

Le nombre d'itérations à une corrélation directe avec la taille de la masse de sorte, qu'il augmente quand la surface est plus importante et diminue, lorsque la surface est moins importante.

4.4.2. Détection des calcifications

La détection des calcifications s'est faite, suivant les opérations du chapeau haut de forme de la morphologie mathématique pour rehausser les calcifications par rapport aux autres structures et par la segmentation par le modèle de Chan et Vese, pour les extraire.

L'élément structurant utilisé pour le chapeau haut de forme correspond à un disque avec une taille de 5X5, car c'est la taille moyenne, d'un élément de base d'une calcification.

Les figures ci-dessous illustrent la détection des calcifications bénignes et malignes.

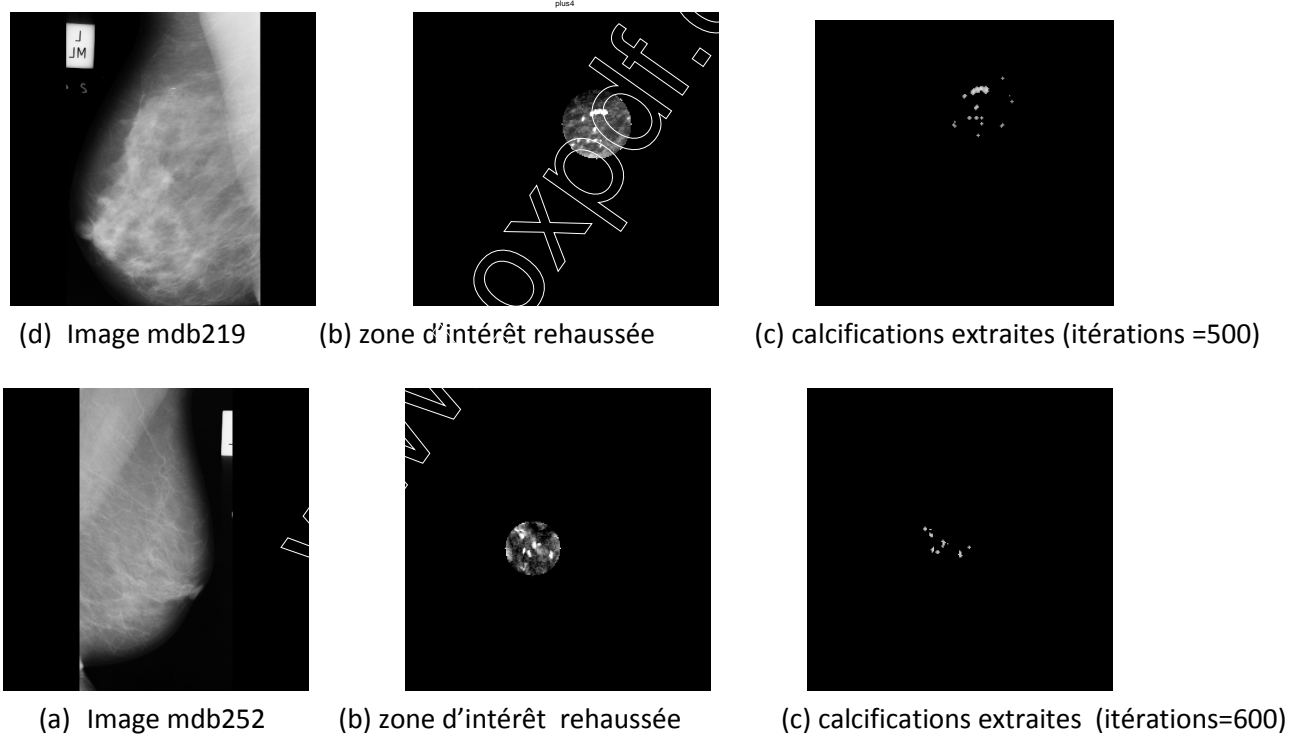


Figure 4.7: détection des calcifications bénignes par le modèle de Chan et Vese

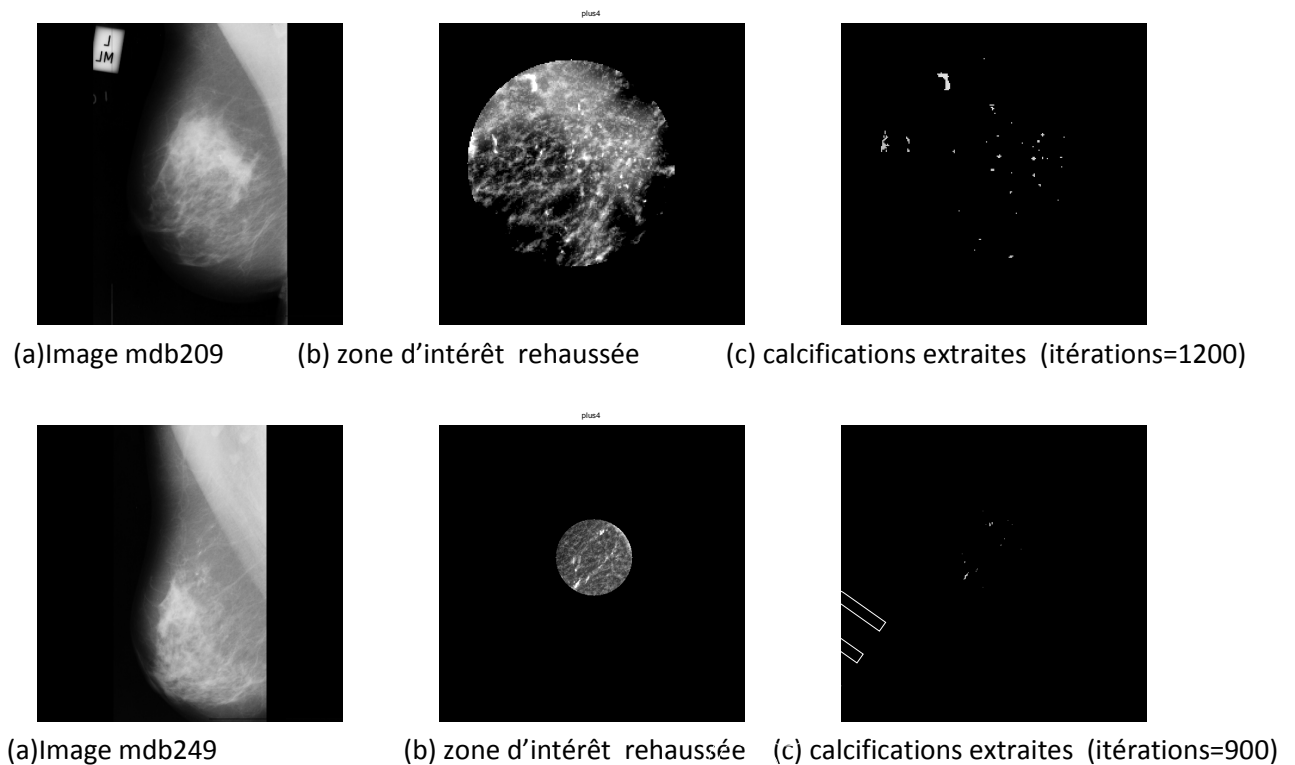


Figure 4.8: détection des calcifications malignes par le modèle de Chan et Vese

Dans le cas des calcifications, le nombre d'itérations pour la détection du contour actif, augmente en fonction de la taille zone d'intérêt.

4.5 Extraction des caractéristiques

La base mammographique MIAS répartie, suivant les masses malignes et bénignes ainsi que, les calcifications malignes et bénignes, a été analysée en explorant six paramètres texturaux (« contraste, corrélation, énergie, homogénéité et, entropie) et quatre attributs morphologiques (compacité, nombre de régions, surface, moment de hu). L'objectif, est de réaliser une base de connaissances via, cette description. Les tableaux suivants, représentent l'évaluation des paramètres utilisés pour la caractérisation, de chaque classe.

	mdb002	mdb005a	mdb005b	mdb010	mdb012	mdb013
Contraste	0,0089	0,0017	0,002	0,0033	0,0026	0,0025
Corrélation	0,9889	0,9739	0,9752	0,9793	0,9766	0,9718
Énergie	0,9771	0,997	0,9965	0,9943	0,9954	0,9975
Homogénéité	0,9996	0,9999	0,9999	0,9998	0,9999	0,9999
Variance	2.2448e+21	2.3377e+21	2.3356e+21	2.3253e+21	2.3302e+21	2.3402e+21
Entropie	0,1494	0,0228	0,0257	0,0417	0,0327	0,0184
Compacité	0,7155	0,6402	0,7507	0,7115	0,6656	0,6885
Surface	12009	1542	1777	2910	2368	1277
Nbre de régions	1	1	1	1	1	1
Hu	0,0245	0,0248	0,0236	0,0233	0,0247	0,0291

	mdb017	mdb019	mdb021	mdb025	mdb030	mdb032
Contraste	0,0051	0,0051	0,0043	0,0036	0,0047	0,0052
Corrélation	0,9854	0,9843	0,9828	0,9884	0,9739	0,9879
Énergie	0,989	0,9882	0,9913	0,9896	0,9948	0,987
Homogénéité	0,9996	0,9996	0,9997	0,9997	0,9999	0,9996
Variance	2.3001e+21	2.2968e+21	2.3110e+21	2.3034e+21	2.3276e+21	2.2908e+21
Entropie	0,0732	0,0808	0,0615	0,0695	0,0358	0,0901
Compacité	0,7625	0,6768	0,6843	0,5595	0,4993	0,4738
Surface	5720	6088	4483	5376	2648	6738
Nbre de régions	1	1	1	1	1	1
Hu	0,0232	0,0239	0,0246	0,0325	0,0261	0,0293

Tableau 4.1: paramètres des images de la classe des masses bénignes

	mdb023	mdb028	mdb058	mdb072	mdb075	mdb090
Contraste	0,0027	0,0049	0,0025	0,0034	0,0014	0,0057
Corrélation	0,9776	0,9899	0,9786	0,9684	0,9671	0,9864
Énergie	0,9955	0,9858	0,9957	0,9968	0,998	0,9879
Homogénéité	0,9998	0,9997	0,9998	0,9999	0,9999	0,9997
Variance	2.3308e+21	2.2852e+21	2.3318e+21	2.3370e+21	2.3427e+21	2.2950e+21
Entropie	0,033	0,0951	0,033	0,024	0,0156	0,0843
Compacité	0,9296	0,8986	0,9596	0,8994	0,9412	0,9148
Surface	2301	7410	2195	1610	989	6299
Nbre de régions	1	1	1	1	1	1
Hu	0,0229	0,0234	0,0228	0,0265	0,0248	0,0235

	mdb095	mdb102	mdb105	mdb110	mdb111	mdb115
Contraste	0,003	0,005	0,018	0,0055	0,012	0,0109
Corrélation	0,9797	0,9833	0,9933	0,9752	0,9924	0,9944
Énergie	0,9957	0,9915	0,9441	0,9926	0,9552	0,9306
Homogénéité	0,9999	0,9999	0,9996	0,9997	0,9995	0,9998
Variance	2.3316e+21	2.3119e+21	2.0923e+21	2.3173e+21	2.1434e+21	2.0346e+21
Entropie	0,0318	0,0592	0,2916	0,0503	0,2721	0,4569
Compacité	0,9347	0,9288	0,9063	0,8512	0,8038	0,8887
Surface	2225	4399	29907	3778	23809	36995
Nbre de régions	1	1	1	1	1	1
Hu	0,0232	0,0228	0,0227	0,0305	0,0241	0,0228

Tableau 4.2: paramètres des images de la classe des masses malignes

	mdb218	mdb219	mdb222	mdb223b	mdb226a	mdb226b
Contraste	0,000971	0,0005994	0,0019	0,000261	0,0003742	0,0024
Corrélation	0,8841	0,5722	0,884	0,4095	0,8613	0,782
Énergie	0,9996	0,9999	0,9993	1	0,9998	0,9995
Homogénéité	1	1	0,9999	1	1	0,9999
Variance	2.3502e+21	2.3517e+21	2.3487e+21	2.3518e+21	2.3511e+21	2.3493e+21
Entropie	0,003	0,00042109	0,0052	0,0002044	0,0015	0,0039
Compacité	1,7453	3,4985	3,6317	1,0186	1,068	6,8072
Surface	173	21	318	10	81	234
Nbre de régions	1	1	4	2	1	3
Hu	0,0337	0,2197	0,0735	0,2863	0,0376	0,2037

Tableau 4.3: paramètres des images de la classe des calcifications bénignes

	mdb209	mdb211	mdb213	mdb231	mdb238	mdb239
Contraste	0,0103	0,0011	0,0041	0,0029	0,0006835	0,0069
Corrélation	0,63	0,5743	0,7485	0,6692	0,5666	0,8169
Énergie	0,9988	0,9999	0,9992	0,9994	0,9999	0,9987
Homogénéité	0,9997	1	0,9998	0,9999	1	0,9998
Variance	2.3461e+21	2.3514e+21	2.3479e+21	2.3491e+21	2.3516e+21	2.3456e+21
Entropie	0,0081	0,0007953	0,006	0,004	0,00050344	0,0094
Compacité	24,1349	2,9134	12,1707	11,2369	2,2834	12,9869
Surface	458	42	345	227	26	582
Nbre de régions	45	8	13	16	6	22
Hu	53,1441	2,1764	11,2201	883,4603	44,8802	0,4673

Tableau 4.4: paramètres des images de la classe des calcifications malignes

L'extraction de paramètres texturaux et morphométriques, nous permet de constituer des vecteurs caractéristiques de chacune des formes étudiées, en vue d'une classification neuro-génétique, pour une prise de décision.

À cet effet Nous montrons, l'influence de la surface, l'entropie, et la compacité sur les masses et les calcifications, suivant les graphes suivantes.

➤ Influence de la surface sur la masse

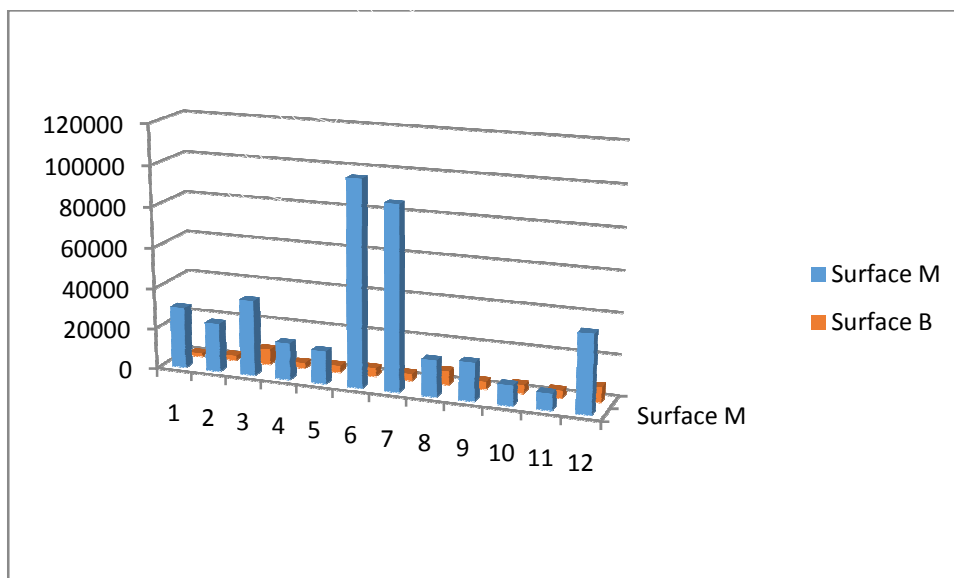


Figure 4.9: influence de la surface sur les masses

➤ **influence de l'entropie sur les masses**

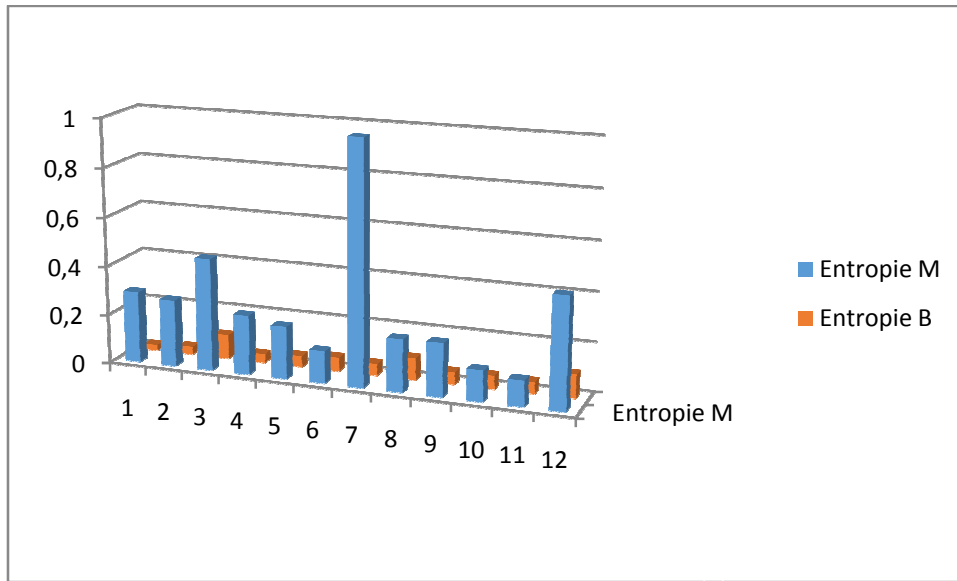


Figure 4.10: influence de l'entropie sur les masses

➤ **influence de la compacité sur les masses**

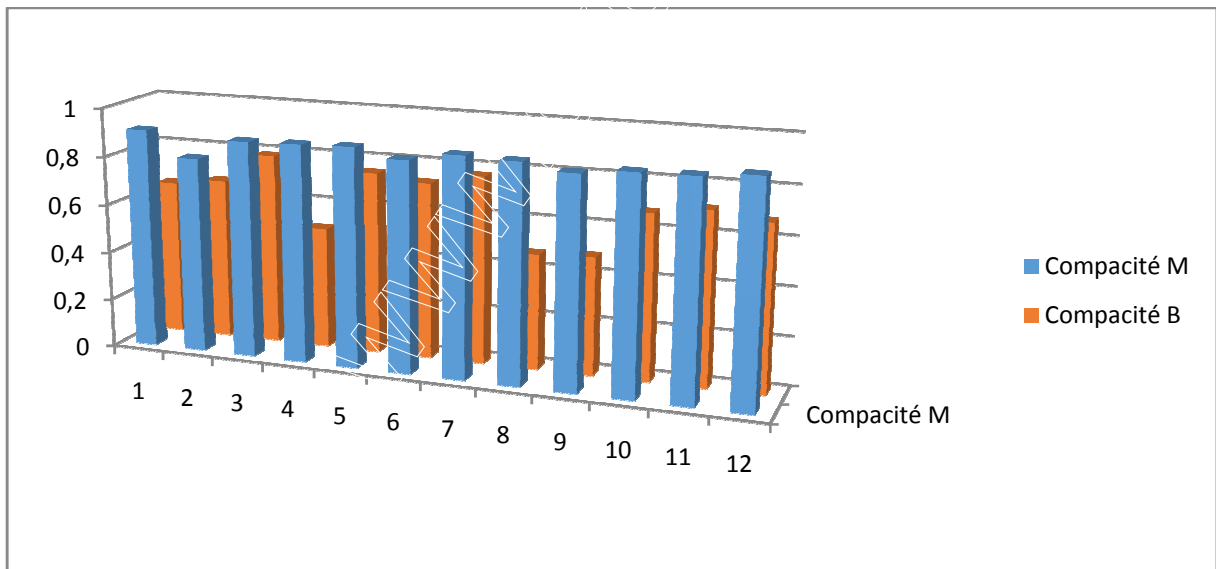


Figure 4.11: influence de la compacité sur les masses

➤ **Influence de la surface sur Les calcifications**

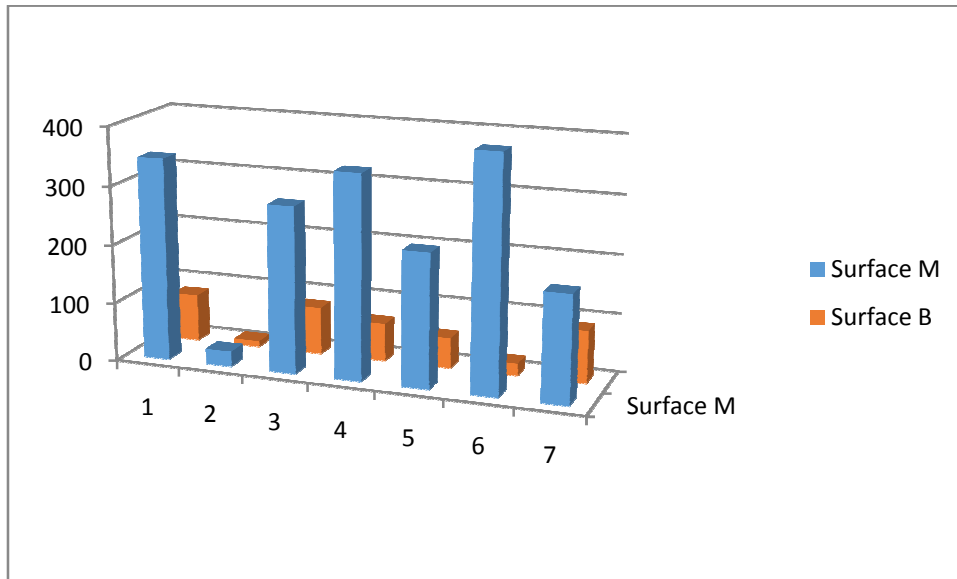


Figure 4.12: influence de la surface sur les calcifications

➤ **influence de l'entropie sur les calcifications**

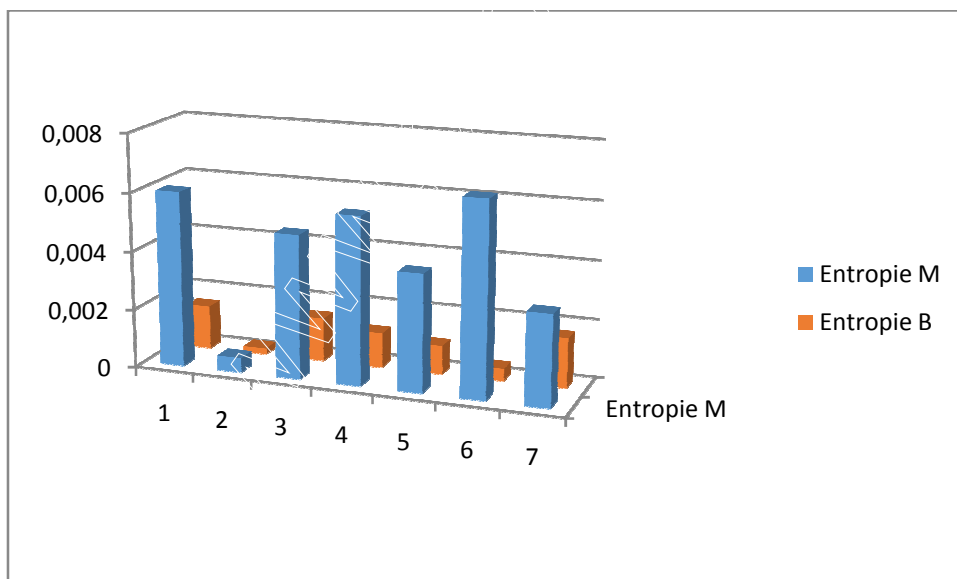


Figure 4.13: influence de l'entropie sur les calcifications

➤ influence de la compacité sur les calcifications

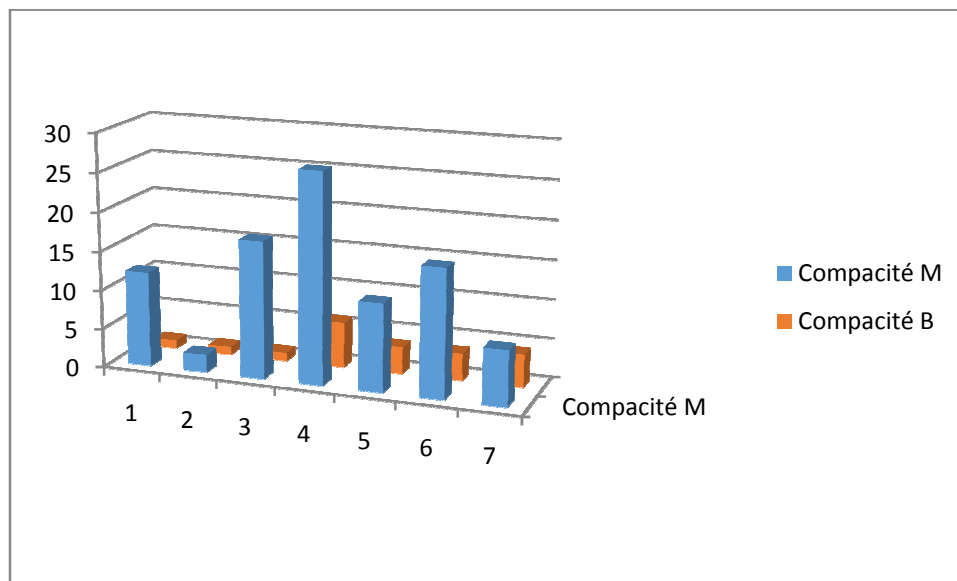


Figure 4.14: influence de la compacité sur les calcifications

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

4.6. Classification des pathologies mammographiques

L'objectif de ce travail est de déterminer la nature des masses et, des calcifications mammographiques. Nous utilisons le modèle du perceptron multicouche en cascading avec les algorithmes génétiques, pour classifier les différentes pathologies mammaires étudiées. Dans ce cas, deux classifieurs sont réalisés, l'un pour les masses et le second, pour les calcifications.

Le nombre de neurones cachés dépend principalement du problème étudié et des résultats de l'expérimentation, sur différentes combinaisons, de ces paramètres. On effectue l'apprentissage avec les deux tiers de la base d'image et, un tiers pour le test.

Nous considérons, la couche d'entrée correspondant aux caractéristiques extraites, une couche cachée et une couche de sortie pour produire les deux classes étudiées, pour chacun des deux classifieurs (l'un pour les masses et l'autre, pour les calcifications).

Les paramètres du classifieur hybride utilisés sont:

- $C = 70$: le paramètre de régularisation (similaire à SVM)
- $\text{Max_indiv} = 1500$: la population initiale
- $\text{Pression} = 5$: la pression sélective [1 - 8]
- $\text{Score} = 1$: ensemble élitisme (1 : oui, 0 : pas)
- $\text{Max_resol} = 12$: le nombre de résolutions de l'algorithme génétique
- $\text{Iter_max} = 10$: le nombre de générations par la résolution

L'apprentissage est l'étape essentielle dans les systèmes de classification. Les résultats corrects (c'est-à-dire les valeurs que l'on désire que le réseau obtienne en sortie) sont fournis au réseau, si bien que celui-ci peut ajuster ses poids de connexions pour les obtenir. Après l'apprentissage, le réseau est testé en lui donnant seulement les valeurs d'entrée mais pas les sorties désirées, et en regardant si le résultat obtenu est proche du résultat désiré.

Pour atteindre le meilleur réseau de neurone, on joue sur le nombre de neurones de la couche cachée. Les configurations que nous avons pris pour le déroulement de l'apprentissage du classifieur « NG » sont les suivants.

« Les meilleurs individus » représentés en rouge, sont initialisés à 1500 individus et, selon les besoins du réseau de neurones, ce nombre peut changer. Comme nous avons considéré 10 neurones (après des tests empiriques) dans la couche cachée pour l'apprentissage des calcifications et 8 neurones pour l'entraînement des masses, ce nombre diminue rapidement, juste après les premières résolutions. Même si à chaque résolution, ils sont initialisés à 1500, ce chiffre s'affaiblit; c'est tout le principe de l'apprentissage.

La fonction fitness moyenne, est représentée en bleu, elle correspond à la « santé » des individus; le but est d'avoir les meilleurs individus possibles pour notre réseau.

Le nombre de générations, représente le nombre de fois que les étapes « mutation et croisement » sont faites sur les individus; 11 dans notre cas, on se retrouve avec un nombre $N+1$ de résolution, qui est égale donc à 12. Le déroulement de l'apprentissage, s'effectue avec 120 itérations.

a. Cas des masses

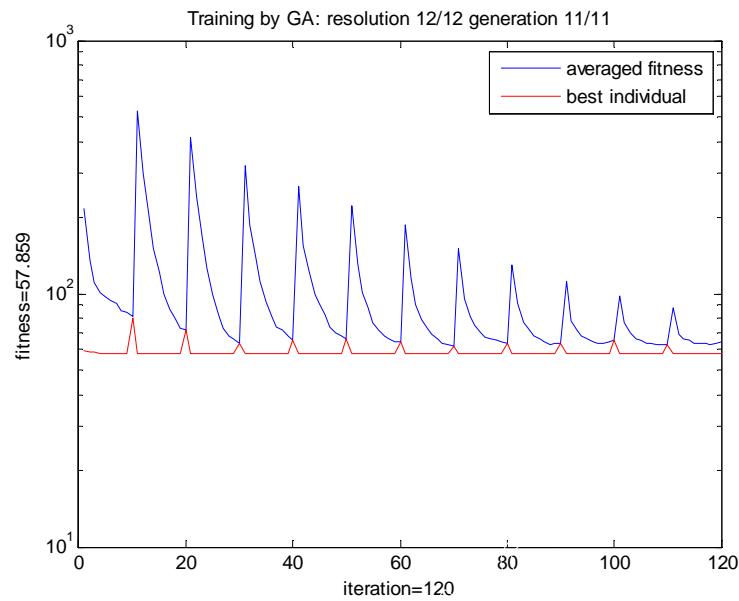


Figure 4.15 : entraînement des masses pour 4 neurones dans la couche cachée

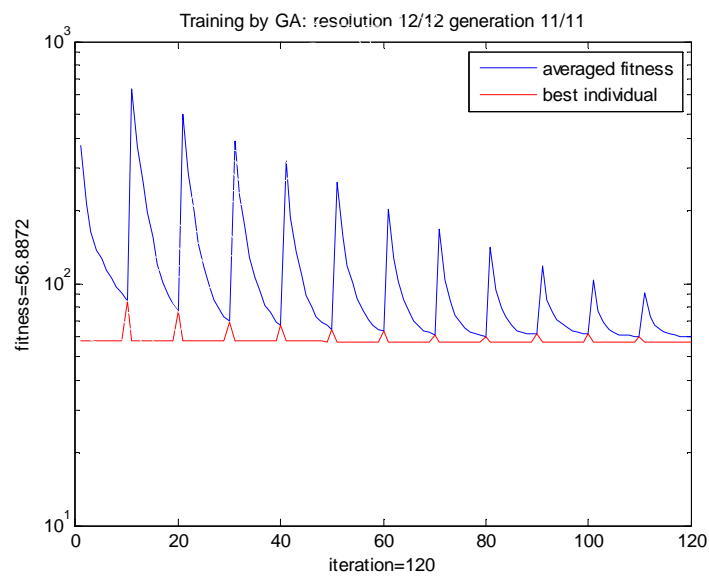


Figure 4.16 : entraînement des masses pour 8 neurones dans la couche cachée

b. Cas de Calcifications

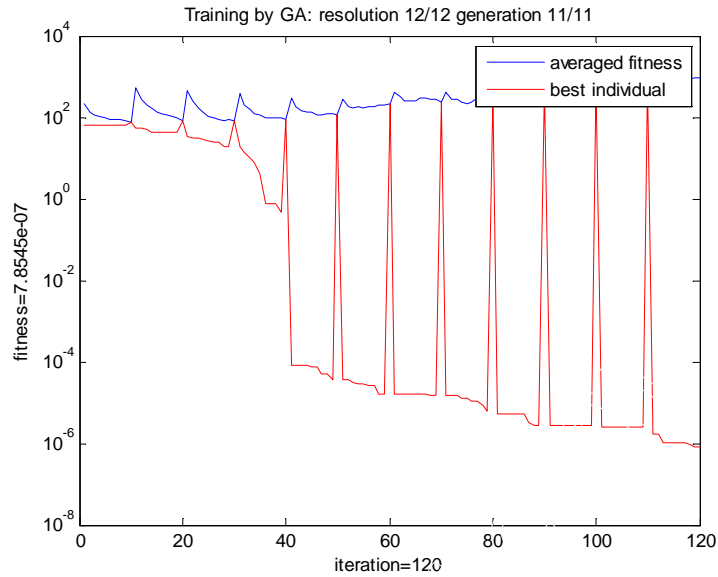


Figure 4.17 : entraînement des calcifications pour 4 neurones dans la couche cachée

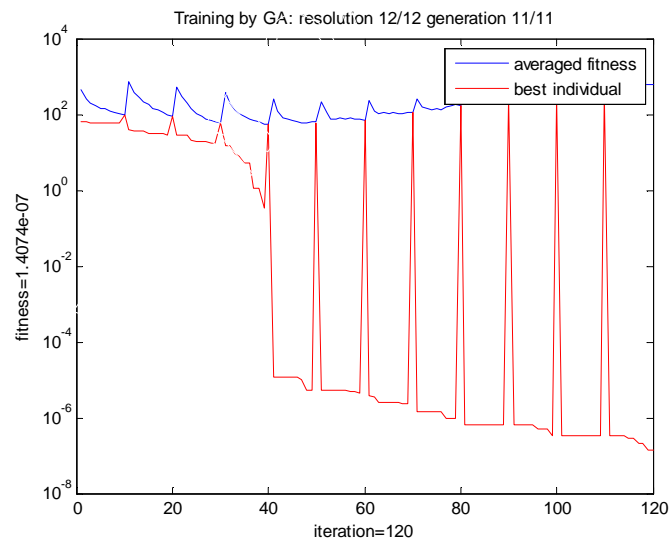


Figure 4.18 : entraînement des calcifications par les AG pour 10 neurones dans la couche cachée

La stabilité des algorithmes génétiques, est obtenue pour 10 neurones dans la couche cachée pour les calcifications et au nombre de 8, pour les masses.

4.7. Evaluation des résultats

Dans cette étape nous faisons le test pour chaque classe. Les résultats que nous avons obtenus sont efficaces, pour notre classifieur hybride. Nous avons évalué le taux de reconnaissance, suivant la matrice de confusion montrée dans la figure 4.17.

	MB	MM	CB	CM	ALL
Classe MB	100%	00%	00%	00%	100%
Classe MM	00%	100%	00%	00%	100%
Classe CB	00%	00%	100%	00%	100%
Classe CM	00%	00%	00%	100%	100%
ALL	100%	100%	100%	100%	100%

Figure 4.19 : matrice de confusion pour la catégorie des tests

La matrice de confusion est un moyen d'évaluer l'efficacité du classifieur utilisé. Les images étudiées, sont toutes reconnues dans leurs classes respectives puisque le taux de classification est de 100%.

Les figures ci-dessous illustrent des exemples de classification à partir de l'application 'NGMAMM'.

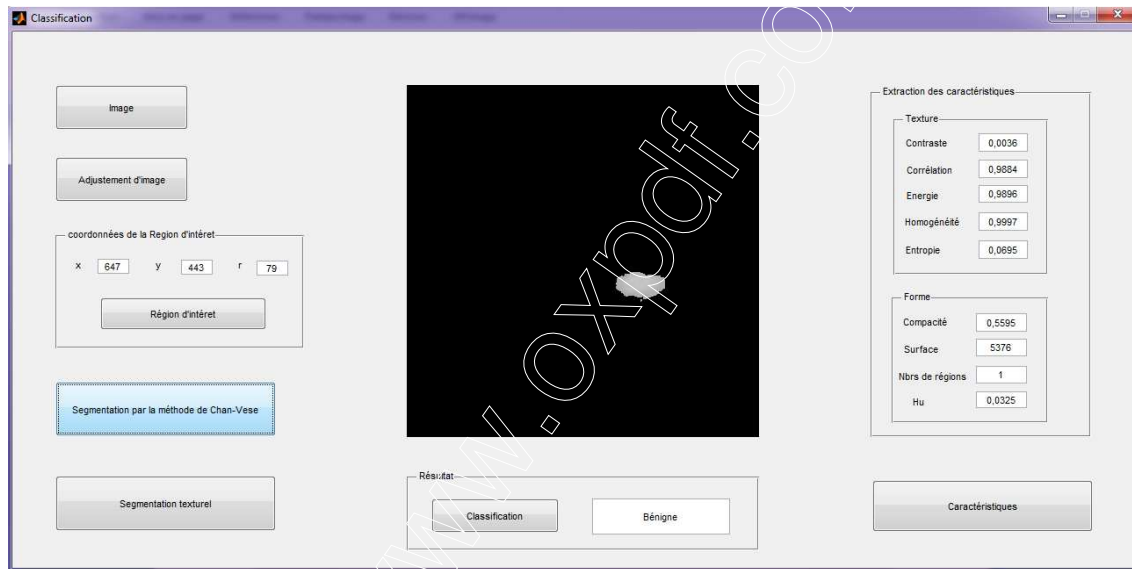
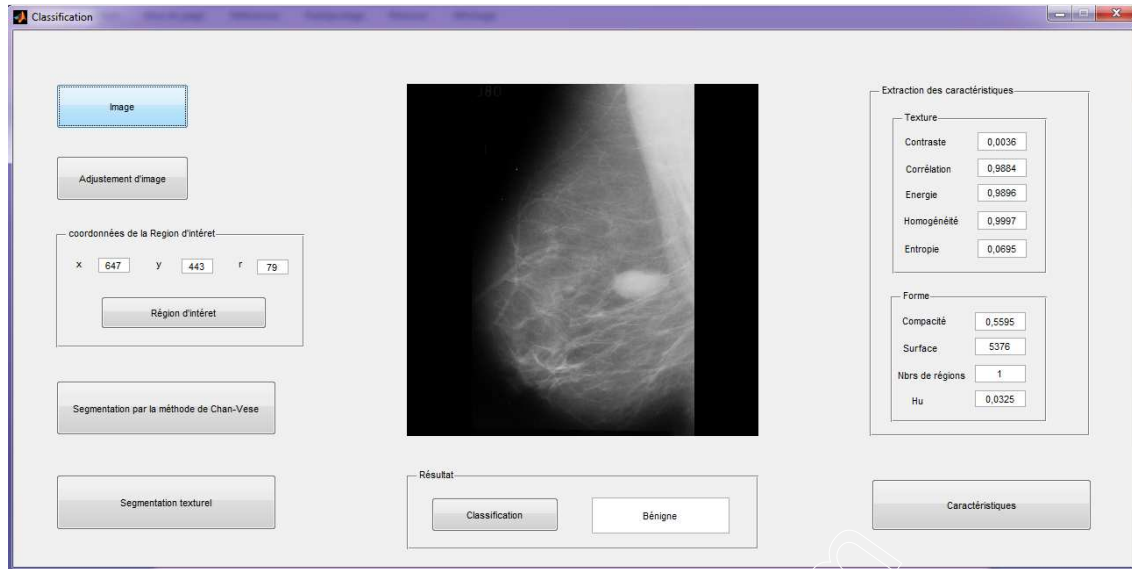


Figure 4.20: classification d'une masse bénigne

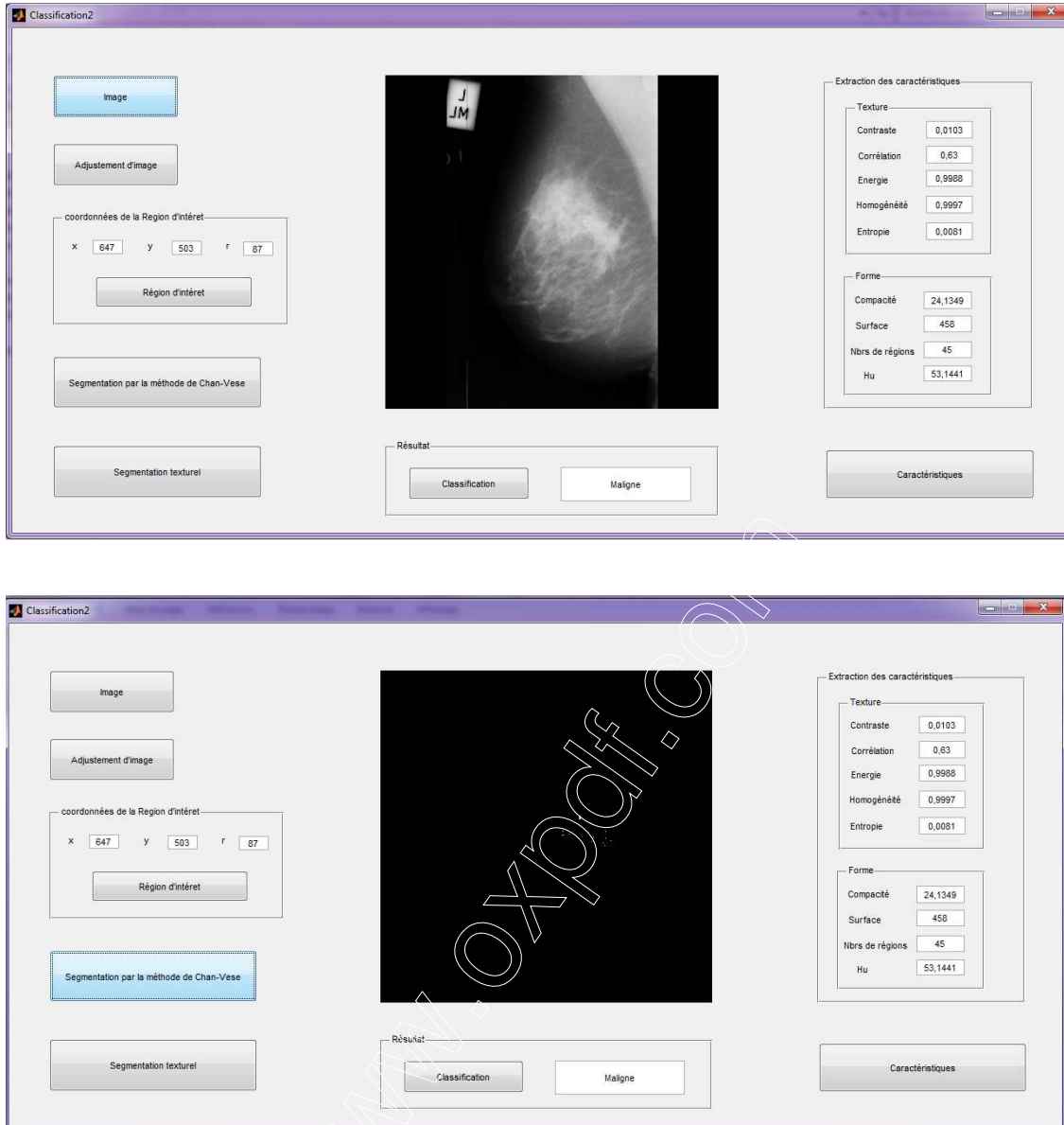


Figure 4.21: classification de calcifications malignes

4.8 Conclusion

Pour les différentes pathologies mammaires la détection des tumeurs est toujours difficile à cause de la variété des types de cancers (masses, microcalcifications) et la variabilité des tissus d'une patiente à une autre. Nous constatons que lorsque nous nous consacrons uniquement aux zones suspectes, le système NGMAMM, offre une classification neuronale précise des masses et des calcifications. Ce qui permet d'établir une interprétation sur les différents résultats obtenus, pour les différentes pathologies mammaires étudiées. Dans le

même cadre, la segmentation, appliquée sur la zone choisie par le clinicien, indique uniquement les structures pathologiques, suivant l'ajustement des paramètres. Nous arrivons ainsi, à mieux distinguer les différentes formes des masses et des calcifications dans les images traitées. Ceci nous confirme davantage, que l'analyse de l'image médicale nécessite, l'aide du clinicien pour établir en premier, la zone suspecte.

Cette étude a confirmé la performance du classifieur hybride, dans la reconnaissance des pathologies mammaires, suivant un choix judicieux des attributs de caractérisation.