

**Faculté des Sciences et Techniques
Settat**

THÈSE DE DOCTORAT

Pour l'obtention de grade de Docteur en Sciences et Techniques

Formation Doctorale : Physique et Sciences de l'Ingénieur

Spécialité : Ingénierie Biomédicale et Traitement d'Image

Sous le thème

**Développement d'un système d'aide à la
détection précoce des masses
mammaires par optimisation
multiobjectif**

Présentée par :

Hamid EL MALALI

Soutenue le : 14/07/2021

A la Faculté des Sciences et Techniques de Settat devant le jury composé de:

Pr. Ahmed MOUHSEN	PES	Université Hassan 1er, FST-Settat	Président
Pr. Rabha ALLAOUI	PES	Université Ibn Tofail, EST Kenitra	Rapporteuse
Pr. Nouredine EL BARBRI	PH	Université Sultan Moulay Slimane, ENSA Khouribga	Rapporteur
Pr. Larbi EL ABDELLAOUI	PES	Université Hassan 1er FST-Settat	Rapporteur
Pr. Mohamed LAMHAMDI	PH	Université Hassan 1er FST-Settat	Examineur
Pr. Azeddine MOUHSEN	PES	Université Hassan 1er FST-Settat	Examineur
Pr. Abdelhadi ASSIR	PA	Université Hassan 1er FST-Settat	Invité
Pr. Mohammed HARMOUCHI	PES	Université Hassan 1er FST-Settat	Directeur de Thèse

Remerciement

A l'issue de la rédaction de cette thèse je n'aurais jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes dont la générosité, la bonne humeur et l'intérêt manifestés à l'égard de ma recherche m'ont permis de progresser durant les moments les plus difficiles.

Je voudrais tout d'abord remercier grandement mon directeur de thèse, Monsieur le professeur **Mohammed HARMOUCHI** de la FST de Settat, pour toute son aide. Je suis ravi d'avoir travaillé en sa compagnie car outre son appui scientifique, il a toujours été là pour me soutenir et me conseiller le long de l'élaboration de cette thèse.

J'adresse de chaleureux remerciements à mon co-encadrant de thèse, Monsieur le professeur **Abdelhadi ASSIR** de la FST de Settat, d'avoir consacré beaucoup de son temps pour la concrétisation de ce travail. Ses conseils avisés et son écoute ont été prépondérants pour la bonne réussite de cette thèse.

Mes remerciements vont également à Monsieur **VIKRANT BHATEJA**, Professeur associé au département de Génie électronique et des communications (ECE), Shri Ramswaroop Memorial Group of Professional Colleges (SRMGPCs), Lucknow, Inde, et également Doyen (académique et IQAC) dans le même collège, pour sa collaboration scientifique qui constitue une composante importante de l'enrichissement et le développement de la recherche scientifique.

Je suis très honoré à remercier de la présence à mon jury de thèse et je tiens à remercier et exprimer ma gratitude à :

- Monsieur le professeur **Ahmed MOUHSEN** de la FST-Settat, d'avoir accepté de présider ce jury et consacré une partie de son temps pour juger cette thèse.

- Madame la professeur **Rabha ALLAOUI** de l'EST de Kénitra d'avoir accepté de rapporter et jugé ce travail de thèse.

- Monsieur le professeur **Noureddine EL BARBRI**, de l'ENSA de Khouribga d'avoir accepté de rapporter et jugé ce travail de thèse.

- Monsieur le professeur **Larbi EL ABDELLAOUI** de la FST de Settat, d'avoir accepté de rapporter cette thèse. C'est grâce à lui que j'ai admiré le domaine de traitement d'image.

- Monsieur le professeur **Mohamed LAMHAMDI** de la FST Settat pour son aide, ses encouragements ainsi que pour son acceptation d'être examinateur de ce travail de thèse.

- Monsieur le professeur **Azeddine MOUHSEN** de la FST de Settat, et Directeur du Laboratoire d'accueil Rayonnement-Matière & Instrumentation. Je voudrais lui témoigner toute ma reconnaissance pour sa disponibilité, sa rigueur et ses conseils. Je le remercie encore d'avoir accepté d'évaluer mon travail en qualité d'examineur.

Enfin mes plus vifs remerciements vont à toutes les personnes avec qui j'ai partagé mes études et notamment ces années de thèse et ma pensée est dirigée vers les membres du Laboratoire Rayonnement-Matière & Instrumentation pour leur aide et leur bonne humeur.

Résumé

Cette thèse a pour objectif de développer un système "DAOe" d'aide à la détection précoce des masses mammaires dont l'introduction de l'optimisation dans les différentes étapes de prétraitement et de segmentation, a permis d'augmenter considérablement la localisation précise desdites lésions.

Nous avons commencé notre travail à délimiter la région d'intérêt "ROI" de l'image origine par élimination des artefacts radio-opaques et la suppression du muscle pectoral présent sur les vues médio-latéral-oblique "MLO" du mammogramme, du fait que la présence de ces artefacts dégrade la qualité de l'image et par conséquent son analyse. La région d'intérêt ainsi obtenue a subi une amélioration optimale de sa qualité via un modèle conçu basé sur l'exploitation locale des propriétés de la fonction sigmoïde gérée par un algorithme d'optimisation multiobjectif. Par ailleurs, l'amélioration optimale de la qualité de l'image est caractérisée par l'optimisation des paramètres d'évaluation "IQA" qui la caractérise à savoir le contraste "EME", la contenance en bord "EC", la luminosité "AMBE" et les informations contenues dans l'image origine "FSIM". Il s'agit en fait d'augmenter le contraste et la contenance en bord de l'image tout en conservant sa luminosité et son contenu. La région d'intérêt améliorée subit enfin une segmentation dans le but d'isoler les zones suspectes de l'image en utilisant la méthode de k-means. D'abord et afin d'éviter le problème de la première initialisation aléatoire de l'algorithme k-means nous lui avons associé l'algorithme d'optimisation par la méthode des algorithmes génétiques. Par la suite nous avons complété le processus de segmentation par la méthode de croissance de régions afin de se débarrasser des pixels aberrants suite au caractère disjoint des classes résultantes. La localisation et les dimensions des zones suspectes ainsi obtenues sont comparables à celles fournies par la base de données Mini-MIAS. Ajoutant à cela que les valeurs des paramètres d'évaluation de la qualité des images obtenues par notre modèle sont meilleures que celles obtenues par les techniques usuelles basées aussi bien sur l'égalisation d'histogramme que sur l'optimisation. Les résultats obtenus confirment la performance de notre système à détecter et à localiser les masses mammaires d'une manière précise et à un stade prématuré.

Mots clefs :

Cancer du sein, masse mammaire, mammographie, détection précoce, prétraitement, segmentation, qualité d'image, région d'intérêt, zone suspecte, artefacts, optimisation mono-objectif, optimisation multiobjectif, modélisation, conception, algorithme k-means, algorithme génétique, croissance de régions.

Abstract

The objective of this thesis is to develop a "CADe" system to assist in the early detection of breast masses and their precise localization. The introduction of optimization in the different steps of pre-processing and segmentation, allows an accurate localization of breast masses. The introduction of optimization in the different steps of preprocessing and segmentation, allows an accurate localization of breast masses. In the beginning, we started with delimiting the region of interest "ROI" from the original image by eliminating radio-opaque artifacts and removing the pectoral muscle present on the medial-lateral-oblique view "MLO" of the mammogram, because the presence of these artifacts degrades the quality of the image and therefore its interpretation. The quality of the obtained ROI is optimally improved using the designed model based on the local exploitation of the properties of the sigmoid function managed by a multiobjective optimization algorithm. Indeed, the optimal improvement of the image quality is characterized by the optimization of the image quality assessment "IQA" such as contrast "EME", edge content "EC", brightness, and information conservation using "AMBE" and "FSIM" respectively. The aim is to increase the contrast and the edge content of the image while preserving its brightness and its information. The enhanced ROI is finally segmented to isolate the suspicious areas of the image using the k-means algorithm. To avoid the problem of the first random initialization of the k-means algorithm, the centers of the clusters are determined by optimizing them by means of genetic algorithm. Subsequently, the segmentation process is completed by the region growing method to remedy the problem of disjoint clusters. The location and the dimensions of the suspect areas obtained are compared to those provided by the Mini-MIAS database. Also, the values of the "IQA" parameters obtained by using our model are better than those obtained by current models. These results confirm the performance of our system to detect and localize breast masses in an accurate way and at a premature stage.

Keywords:

Breast cancer, breast mass, mammography early detection, preprocessing segmentation, image quality, region of interest, suspect area, artifacts, single-objective optimization, multiobjective optimization, modeling, design, k-means algorithm, genetic algorithm, region growth.

Table des matières

REMERCIEMENT.....	I
RESUME	II
ABSTRACT	III
TABLE DES MATIERES.....	IV
LISTES DES FIGURES	VIII
LISTE DES TABLEAUX.....	X
LISTE DES ALGORITHMES	XI
LISTE DES ACRONYMES	XII
INTRODUCTION GENERALE	1
CHAPITRE 1 CANCER DU SEIN ET MAMMOGRAPHIE	4
1. ANATOMIE DU SEIN	4
2. TUMEURS MAMMAIRES.....	5
2.1. GENERALITES SUR LES TUMEURS	5
2.2. CANCER DU SEIN	6
2.3. SYMPTOMATOLOGIE CLINIQUE DU CANCER DU SEIN :	8
2.4. FACTEURS DE RISQUE LIES AU CANCER DU SEIN	8
2.5. TYPE DE TUMEURS DU SEIN.....	10
2.5.1. Formes bénignes.....	10
2.5.2. Formes malignes.....	11
3. DEPISTAGE ET DIAGNOSTIC DU CANCER DU SEIN	12
3.1. ACTES PRELIMINAIRES	12
3.2. MAMMOGRAPHIE	13
3.2.1. Incidences standards	14
3.2.2. Incidences complémentaires.....	15
3.2.3. Tomosynthèse mammaire ou mammographie 3D.....	15
3.3. ECHOGRAPHIE MAMMAIRE.....	16
3.4. IMAGERIE PAR RESONANCE MAGNETIQUE IRM.....	17
3.5. EXAMENS DE CONFIRMATION DU CANCER DU SEIN	18
4. CLASSIFICATION DES LESIONS MAMMAIRES	19
4.1. CLASSIFICATION DE WOLFE	19
4.2. CLASSIFICATION TNM.....	19
4.3. CLASSIFICATION LE GAL	20
4.4. CLASSIFICATION BIRADS (BREAST IMAGING REPORTING AND DATA SYSTEM).....	20
4.4.1. Classification des seins suivant leur composition	21
4.4.2. Caractères pathologiques des lésions du sein	21
5. CONCLUSION.....	24
CHAPITRE 2 DESCRIPTION ET OPTIMISATION DU SYSTEME D'AIDE A LA DETECTION ET AU DIAGNOSTIC ASSISTE PAR ORDINATEUR "DAO" DES MASSES MAMMAIRES	25
INTRODUCTION.....	25
PARTIE 1 : DESCRIPTION DU SYSTEME D'AIDE A LA DETECTION ET AU DIAGNOSTIC "DAO" DES MASSES MAMMAIRES ET CARACTERISTIQUES DE L'IMAGE NUMERIQUE.....	25
1. DIFFERENTES ETAPES DU SYSTEME "DAO"	25
2. IMAGE NUMERIQUE : CARACTERISTIQUES ET BASE DE DONNEES DES IMAGES UTILISEES.....	26
2.1. IMAGE BINAIRE	28

2.2.	IMAGE EN NIVEAU DE GRIS.....	28
2.3.	RESOLUTION D'UNE IMAGE.....	28
2.4.	CONTRASTE D'UNE IMAGE.....	28
2.5.	FORMAT D'UNE IMAGE.....	29
2.6.	BASE DE DONNEES MINI-MIAS.....	30
3.	TECHNIQUES DE PRETRAITEMENT DES IMAGES MAMMOGRAPHIQUES.....	30
3.1.	TECHNIQUES BASEES SUR LA MODIFICATION DE L'HISTOGRAMME.....	31
3.1.1.	<i>Egalisation de l'histogramme (HE)</i>	31
3.1.2.	<i>Egalisation bi-histogramme pour la préservation de la luminosité (BBHE)</i>	32
3.1.3.	<i>Égalisation adaptative d'histogramme à contraste limité (CLAHE)</i>	32
3.2.	TECHNIQUES D'AMELIORATION POINT A POINT.....	33
3.2.1.	<i>Transformation logarithmique</i>	33
3.2.2.	<i>Transformation exponentielle</i>	34
3.2.3.	<i>Transformation par loi de puissance</i>	35
3.2.4.	<i>Transformation sigmoïde</i>	36
4.	TECHNIQUES DE SEGMENTATION DES IMAGES MAMMOGRAPHIQUES.....	37
4.1.	TECHNIQUES BASEES REGIONS.....	39
4.1.1.	<i>Méthode de croissance de région</i>	40
4.1.2.	<i>Méthode de décomposition fusion</i>	40
4.2.	METHODES BASEES PIXELS.....	41
4.2.1.	<i>Seuillage</i>	41
A.	Seuillage global.....	41
B.	Seuillage multiple.....	42
C.	Seuillage adaptatif.....	42
4.2.2.	<i>Méthodes basées regroupement des pixels (classification)</i>	42
A.	Méthode de classification K-moyennes (K-means).....	43
B.	Méthode de classification C-moyennes Flou (Fuzzy C-mens).....	44
4.3.	METHODES BASEES SUR LE CONTOUR.....	44
A.	Méthode basée gradient.....	45
B.	Modèles déformables.....	46
PARTIE 2 : OPTIMISATION DES SYSTEMES D'AIDE AU DIAGNOSTIC "DAO" DES MASSES MAMMAIRES.....		48
1.	CONCEPTS FONDAMENTAUX.....	48
2.	OPTIMISATION MONO-OBJECTIF.....	49
3.	OPTIMISATION MULTIOBJECTIF.....	50
3.1.	FRONT DE PARETO.....	51
3.2.	ETUDE DES RELATIONS ENTRE LES OBJECTIFS D'UN PROBLEME.....	52
3.2.1.	<i>Relations globales</i>	52
A.	Relation d'indépendance.....	52
B.	Relation d'harmonie.....	52
C.	Relation du conflit.....	53
3.2.2.	<i>Etude des relations locales et composites entre les objectifs d'un problème multiobjectif</i>	53
3.2.3.	<i>Détermination des objectifs significatifs</i>	54
3.2.4.	<i>Méthode de la somme pondérée</i>	55
4.	ALGORITHMES GENETIQUES.....	55
4.1.	TERMINOLOGIE ET CONCEPTS DE BASE.....	56
4.2.	CODAGE DES VARIABLES.....	57
4.2.1.	<i>Codage binaire standard</i>	57
4.2.2.	<i>Codage binaire gray</i>	58
4.2.3.	<i>Codage réel</i>	58
4.3.	EVALUATION ET SELECTION DES CHROMOSOMES.....	58
A.	Sélection par rang.....	59
B.	Sélection par roulette.....	59
C.	Sélection par tournoi :.....	59

D.	Reste stochastique sans remplacement	59
4.4.	REPRODUCTION ET RECOMBINAISON GENETIQUE	60
E.	Opérateur de croisement	60
F.	Opérateur de mutation	61
4.5.	CRITERE D'ARRET.....	62
5.	CONCLUSION	62
CHAPITRE 3 PROPOSITION D'UN MODELE POUR L'AMELIORATION OPTIMALE DES IMAGES MAMMOGRAPHIQUES63		
1.	INTRODUCTION.....	63
2.	IMPERFECTIONS DES IMAGES MAMMOGRAPHIQUES ET AMELIORATIONS.....	64
2.1.	PROBLEMES LIES AUX IMAGES MAMMOGRAPHIQUES	64
2.2.	AMELIORATION DE LA QUALITE DES MAMMOGRAMES : CONTEXTE DE L'APPROCHE PROPOSEE	64
2.3.	DESCRIPTION ET EXPRESSIONS DES PARAMETRES D'EVALUATION "IQA"	66
2.3.1.	Mesure effective de l'amélioration : EME (Effective Measure of Enhancement)	66
2.3.2.	Contenance de bord : EC (Edge Content)	66
2.3.3.	Erreur moyenne absolue de luminosité : AMBE (Absolute Mean Brightness Error)	67
2.3.4.	Mesure d'indice de similarité des caractéristiques : FSIM (Feature Similarity Index Measure).....	67
2.3.5.	Mesure de l'entropie d'une image	68
2.3.6.	Rapport signal-bruit de crête : PSNR (Peak Signal-to-Noise Ratio)	69
3.	EVALUATION DES PARAMETRES "IQA" DES IMAGES TEST PAR LES TECHNIQUES USUELLES	69
3.1.	TECHNIQUES USUELLES BASEES SUR LA MODIFICATION DE L'HISTOGRAMME.....	69
3.1.1.	Présentation et analyse des paramètres d'évaluation "IQA"	70
3.1.2.	Présentation et analyse des images traitées	71
3.2.	TECHNIQUE USUELLE BASEE SUR L'OPTIMISATION ("TUBO").....	72
3.2.1.	Présentation du modèle d'optimisation usuel	72
3.2.2.	Présentation et analyse des paramètres d'évaluation "IQA" du modèle "TUBO"	74
3.2.3.	Présentation et analyse des images traitées	75
3.3.	COMPARAISON DES DEUX TECHNIQUES	76
4.	MODELES D'AMELIORATION PROPOSES A OPTIMISATION MONO- ET MULTIOBJECTIF.....	76
4.1.	MODELE PROPOSE A OPTIMISATION MONO-OBJECTIF	79
4.1.1.	Présentation et analyse des paramètres d'évaluation "IQA"	80
4.1.2.	Présentation et analyse des images traitées	81
4.1.3.	Comparaison des résultats des deux modèles mono-objectif proposé et "TUBO"	82
4.2.	MODELE PROPOSE A OPTIMISATION MULTIOBJECTIF	83
4.2.1.	Formulation mathématique du problème	83
4.2.2.	Transformation du problème d'optimisation multiobjectif en mono-objectif	85
A.	Front de Pareto des images de test	85
B.	Etude des relations globales entre les objectifs	87
C.	Etude des relations locales et composites entre les objectifs	88
D.	Etude de la signifiante des objectifs du problème	90
E.	Réduction des objectifs du problème.....	91
F.	Fonction objectif et algorithme du modèle proposé	92
4.2.3.	Résultats de l'application du modèle à optimisation multiobjectif	93
A.	Résultats d'application du modèle proposé aux images test	93
B.	Résultats d'application du modèle proposé à toute la base de données Mini-MIAS	95
4.2.4.	Comparaison des résultats des deux modèles (mono et multiobjectif)	96
5.	CONCLUSION.....	97
CHAPITRE 4 CONCEPTION D'UN SYSTEME AUTOMATISE D'AIDE A LA DETECTION "DAOE" DES MASSES MAMMAIRES		
.....		98
1.	INTRODUCTION.....	98
2.	PRESENTATION DU SYSTEME D'AIDE A LA DETECTION "DAOE"	98
3.	ETAPE DE PRETRAITEMENT DE L'IMAGE ORIGINE.....	99

3.1.	DELIMITATION AUTOMATIQUE DE LA REGION D'INTERET DU MAMMOGRAME	100
3.1.1.	<i>Elimination des artefacts</i>	100
A.	Méthodes et techniques utilisées.....	100
B.	Résultats de l'étape de l'élimination des artefacts.....	102
3.1.2.	<i>Identification et Elimination du muscle pectoral</i>	102
A.	Identification de l'orientation et ré-orientation du sein.....	103
B.	Elimination du muscle pectoral	104
C.	Résultats de l'élimination du muscle pectoral.....	106
3.2.	AMELIORATION DE LA QUALITE DES "ROI"	108
4.	SEGMENTATION DE LA REGION D'INTERET	109
4.1.	ALGORITHME K-MEANS : LES AVANTAGES.....	109
4.2.	ALGORITHME K-MEANS : LES INCONVENIENTS ET REMEDIATION	110
4.2.1.	<i>Remédiation au problème lié à l'initialisation</i>	110
4.2.2.	<i>Remédiation au problème de disjonction de clusters</i>	111
4.3.	RESULTATS DE LA SEGMENTATION DES ROI.....	112
5.	CONCLUSION	114
	CONCLUSION GENERALE ET PERSPECTIVES	116
	ANNEXES.....	119
	BIBLIOGRAPHIE :	126

Listes des figures

FIGURE 1. 1. COUPE SAGITTALE DE LA PAROI THORACIQUE ET GLANDE MAMMAIRE.	4
FIGURE 1. 2. ENVAHISSEMENT DES CELLULES TUMORALES DU TISSU SAIN AVOISINANT.	5
FIGURE 1. 3. DEBUT D'UN CANCER (A) COUPE SAGITTALE DU SEIN (B) DEBUT D'UN CANCER LOBULAIRE (C) DEBUT D'UN CANCER CANALAIRE.	7
FIGURE 1. 4. DIFFERENTS QUADRANTS POUR LE REPERAGE DES LESIONS DU SEIN.	7
FIGURE 1. 5. PRINCIPAUX ELEMENTS D'UN MAMMOGRAPHE.....	13
FIGURE 1. 6. ILLUSTRATIONS DES INCIDENCES STANDARDS EN MAMMOGRAPHIE : (A) INCIDENCE CRANIO CAUDALE (CC), (B) INCIDENCE MEDIO LATERALE OBLIQUE (MLO) ET (C) INCIDENCE DE PROFIL.	14
FIGURE 1. 7. PRINCIPE DE LA TOMOSYNTHESE MAMMAIRE.....	15
FIGURE 1. 8. PRINCIPE DE L'ECHOGRAPHIE MAMMAIRE.....	16
FIGURE 1. 9. PRINCIPE D'IMAGERIE PAR RESONANCE MAGNETIQUE.....	18
FIGURE 1. 10. CLASSIFICATION BIRADS DU SEIN SELON SA COMPOSITION (A) ENTIEREMENT GRAISSEUX, (B) SEIN DE DENSITE FIBROGLANDULAIRE DISPERSEE (C) SEIN DE DENSITE HETEROGENE (D) SEIN EXTREMEMENT DENSE.....	21
FIGURE 1. 11. LES DIFFERENTES FORMES DE MASSES SELON LE STANDARD BIRADS 5 : (A) RONDE, (B) OVALE, (C) ET (D) IRREGULIERE.....	22
FIGURE 1. 12. DIFFERENTS TYPES DE CONTOUR DES MASSES SELON LE BI-RADS 5 : (A) CONTOUR CIRCONSCRIT, (B) CONTOUR MASQUE, (C) CONTOUR MICRO-LOBULE, (D) CONTOUR INDISTINCT ET (E) CONTOUR SPICULE.....	22
FIGURE 2. 1. LES DIFFERENTES ETAPES DU SYSTEME D'AIDE A LA DETECTION ET AU DIAGNOSTIC "DAO".....	26
FIGURE 2. 2. MAMMOGRAMMES DE DIFFERENTES CLASSES PRESENTANT DES LESIONS (A) MDB028 SEIN GRAISSEUX AVEC UNE TUMEUR MALIGNE (B) MDB001 SEIN DENSE AVEC UNE TUMEUR BENIGNE (C) MDB193 SEIN EXTREMEMENT DENSE AVEC UNE TUMEUR BENIGNE.....	31
FIGURE 2. 3. COURBE DE LA TRANSFORMATION LOGARITHMIQUE.....	34
FIGURE 2. 4. COURBE DE LA TRANSFORMATION EXPONENTIELLE.....	35
FIGURE 2. 5. COURBE DE LA TRANSFORMATION PAR LOI DE PUISSANCE.....	36
FIGURE 2. 6. COURBE DE LA TRANSFORMATION SIGMOÏDE.....	37
FIGURE 2. 7. DIFFERENTES REGIONS CARACTERISANT LE MAMMOGRAMME MDB012.....	38
FIGURE 2. 8. CLASSIFICATION DES TECHNIQUES DE SEGMENTATION.....	39
FIGURE 2. 9. PRINCIPE DE SEGMENTATION PAR LA METHODE DE LA CROISSANCE DE LA REGION (A) IMAGE ORIGINE, (B) INITIALISATION, (C) CROISSANCE (D) REGION ISOLEE.....	40
FIGURE 2. 10. PRINCIPE DE SEGMENTATION PAR LA METHODE DE LA CROISSANCE DE LA REGION (A) IMAGE ORIGINE, (B) ET (C) DIVISION DES REGIONS (D) FUSION DE DEUX REGIONS.....	41
FIGURE 2. 11. MASQUES DE CONVOLUTION POUR LA DETECTION DES CONTOURS (PREWITT, SOBEL, SCHARR).....	45
FIGURE 2. 12. DETECTION DE BORDS PAR APPLICATION DES MASQUES DE CONVOLUTION ; (A) "ROI" MDB184, (B) MASQUE PREWITT, (C) MASQUE SOBEL.....	45

FIGURE 2. 13. SOLUTIONS NON DOMINEES OU FRONT DE PARETO D'UN PROBLEME A DEUX OBJECTIFS (A) CAS D'UNE MINIMISATION, (B) CAS D'UNE MAXIMISATION.....	51
FIGURE 2. 14. SCHEMA ILLUSTRANT LA CARTE DE REGIONS DANS LE CAS DE QUATRE OBJECTIFS.....	54
FIGURE 2. 15. LES DIFFERENTS NIVEAUX D'ORGANISATION D'UN ALGORITHME GENETIQUE	56
FIGURE 2. 16. STRUCTURE DE FONCTIONNEMENT D'UN ALGORITHME GENETIQUE.....	57
FIGURE 2. 17. CODAGE REEL D'UN CHROMOSOME DANS LE CAS DE L'OPTIMISATION D'UNE FONCTION A N VARIABLES.	58
FIGURE 2. 18. CROISEMENT ENTRE DEUX CHROMOSOMES ; (A) A UN SEUL POINT, (B) A DEUX POINTS.	60
FIGURE 2. 19. MUTATION D'UN ALLELE D'UN CHROMOSOME A LA CINQUIEME POSITION.....	61
FIGURE 3. 1. LES IMAGES TEST AMELIOREES PAR LES TECHNIQUES A MODIFICATION D'HISTOGRAMME : (A) IMAGE ORIGINE, (B) PAR HE (C) PAR BBHE, (D) PAR CLAHE	71
FIGURE 3. 2. LES IMAGES TEST AMELIOREES PAR LE MODELE A OPTIMISATION USUEL.	76
FIGURE 3. 3. EFFETS DE MODIFICATION DES PARAMETRES DE LA FONCTION SIGMOÏDE. (A) FONCTION SIGMOÏDE USUELLE ; (B) MODIFICATION DU CENTRE A ; (C) MODIFICATION DE LA BASE B ; (D) MODIFICATION DE LA LARGEUR (B).....	78
FIGURE 3. 4. IMAGES TEST AMELIOREES PAR LES MODELES A OPTIMISATION MONO-OBJECTIF :.....	82
FIGURE 3. 5. REPARTITION DES SOLUTIONS DANS LES CARTES DE REGIONS DES IMAGES D'ETUDE; (A) MDB209; (B) MDB210; (C) MDB211; (D) MDB212.	89
FIGURE 3. 6. RESULTATS D'AMELIORATION DES IMAGES TEST : (A) IMAGE ORIGINE, (B) PAR LE MODELE PROPOSE A OPTIMISATION MONO-OBJECTIF (C) PAR LE MODELE PROPOSE A OPTIMISATION MULTIOBJECTIF.....	95
FIGURE 4. 1. ORGANIGRAMME DU SYSTEME CONÇU D'AIDE A LA DETECTION "DAOE"	99
FIGURE 4. 2. MAMMOGRAMME MDB012 AVEC DIFFERENTS ARTEFACTS POUVANT INFECTER UN MAMMOGRAMME	99
FIGURE 4. 3. ETAPES D'ISOLATION DU PROFIL DU SEIN DU RESTE DE L'IMAGE (A) IMAGE ORIGINE MDB012, (B) IMAGE BINAIRE, (C) OBJET PLUS LARGE DE L'IMAGE, (D) PROFIL DU SEIN.	101
FIGURE 4. 4. RESULTAT DE L'ETAPE D'ELIMINATION DES ARTEFACTS RADIO-OPAQUES.....	102
FIGURE 4. 5. DEUX MAMMOGRAMMES D'ORIENTATIONS DIFFERENTES ; (A) DE DROITE VERS LA GAUCHE, (B) DE GAUCHE VERS LA DROITE	103
FIGURE 4. 6. INVERSEMENT DE L'ORIENTATION DU MAMMOGRAMME MDB271 ; (A)IMAGE ORIGINE ORIENTEE DE DROITE VERS LA GAUCHE, (B) MEME IMAGE ORIENTEE DE GAUCHE VERS LA DROITE.....	104
FIGURE 4. 7. DELIMITATION AUTOMATIQUE DE LA "ROI" AVEC REDUCTION DE TAILLE (MDB010) (A) IMAGE ORIGINE SANS ARTEFACT, (B) IMAGE AUTOMATIQUEMENT ROGNEES, (C) "ROI" DETECTEE.....	104
FIGURE 4. 8. DELIMITATION AUTOMATIQUE DE LA "ROI" SANS REDUCTION DE TAILLE (MDB271) (A) IMAGE ORIGINE SANS ARTEFACT, (B) IMAGE AUTOMATIQUEMENT ROGNEES, (C) "ROI" DETECTEE.....	106
FIGURE 4. 9. RESULTAT DE L'ETAPE D'ELIMINATION DU MUSCLE PECTORAL. (A) IMAGE ORIGINE, (B) IMAGE SANS ARTEFACT ORIENTEE DE GAUCHE A DROITE, (C) ELIMINATION DU MUSCLE PECTORAL.....	107
FIGURE 4. 10. IMAGE OBTENUES APRES AMELIORATION DES "ROI".....	109
FIGURE 4. 11. RESULTATS DE LA SEGMENTATION DES ROI; (A) ROI AMELIOREE (B) SEGMENTATION PAR K-MEANS, (C) ZONES SUSPECTES ISOLEES	113

Liste des tableaux

TABLEAU 1. 1 CLASSIFICATION DE L'ACR DES LESIONS MAMMAIRES SUIVANT LE STANDARD BIRADS.....	23
TABLEAU 3. 1. PARAMETRES D'EVALUATION DES IMAGES TRAITEES PAR LES TECHNIQUES HE, BBHE ET CLAHE.....	70
TABLEAU 3. 2. LES VECTEURS DE DECISION DES IMAGES TEST TRAITEES PAR LE MODELE D'OPTIMISATION USUEL.....	74
TABLEAU 3. 3. PARAMETRES D'EVALUATION "IQA" PAR LA TECHNIQUE USUELLE A OPTIMISATION.....	75
TABLEAU 3. 4. VECTEURS DE DECISION DES IMAGES TEST TRAITEES PAR NOTRE MODELE PROPOSE A OPTIMISATION MONO-OBJECTIF. 80	
TABLEAU 3. 5. PARAMETRES D'EVALUATION DES IMAGES TEST TRAITEES PAR NOTRE METHODE BASE SUR L'OPTIMISATION MONO-OBJECTIF.....	81
TABLEAU 3. 6. SOLUTIONS NON DOMINEES DE L'IMAGE MDB209.....	86
TABLEAU 3. 7. SOLUTIONS NON DOMINEES DE L'IMAGE MDB210.....	86
TABLEAU 3. 8. SOLUTIONS NON DOMINEES DE L'IMAGE MDB211.....	86
TABLEAU 3. 9. SOLUTIONS NON DOMINEES DE L'IMAGE MDB212.....	86
TABLEAU 3. 10. LES COEFFICIENTS DE CORRELATION DE KENDALL POUR CHAQUE PAIRE D'OBJECTIFS DES IMAGES TEST.....	87
TABLEAU 3. 11. INTERVALLES D'EVOLUTIONS DES OBJECTIFS.....	91
TABLEAU 3. 12. MOYENNES DES OBJECTIFS DES IMAGES TEST POUR DIFFERENTES VALEURS DE Ω_1 ET Ω_2	92
TABLEAU 3. 13. VECTEURS DE DECISION DES IMAGES TEST PAR LE MODEL D'OPTIMISATION MULTIOBJECTIF.....	94
TABLEAU 3. 14. PARAMETRES D'EVALUATION "IQA" DES IMAGES TEST PAR LE MODELE MULTIOBJECTIF.....	94
TABLEAU 3. 15. MOYENNES DES DIFFERENTS PARAMETRES D'EVALUATION DES IMAGES DE LA BASE DE DONNES MINI-MIAS PAR NOTRE MODELE A OPTIMISATION MULTIOBJECTIF.....	96
TABLEAU 4. 1. LES VECTEURS DE DECISION DES "ROI" AMELIOREES PAR LE MODELE MULTIOBJECTIF.....	108
TABLEAU 4. 2. LES PARAMETRES D'EVALUATION DE L'AMELIORATION DE LA QUALITE DES" ROI" PAR LE MODELE MULTIOBJECTIF ..	108
TABLEAU 4. 3. VECTEURS DE DECISION DES "ROI" PERMETTANT UNE MEILLEURE SEGMENTATION.....	113
TABLEAU 4. 4. VALEURS EVALUEES DES COORDONNEES DU CENTRE (x, y) ET DU RAYON R DU CERCLE CIRCONSCRIT DE LA TUMEUR EN PIXELS.....	114
TABLEAU 4.5. ECART ENTRE LES VALEURS EVALUEES DES COORDONNEES DU CENTRE (x, y) ET DU RAYON R DU CERCLE CIRCONSCRIT DE LA TUMEUR EN PIXELS ET LEURS VALEURS FOURNIES PAR LA BASE DE DONNEES MINI-MIAS.....	114

Liste des algorithmes

ALGORITHME 3. 1. ALGORITHME D'OPTIMISATION DE LA QUALITE DES MAMMOGRAMES PAR LA TECHNIQUE USUELLE BASEE SUR L'OPTIMISATION	74
ALGORITHME 3. 2. ALGORITHME DE TRAITEMENT D'UNE IMAGE PAR LA TRANSFORMATION LOCALE BASEE SUR LA FONCTION SIGMOÏDE PROPOSEE DE PARAMETRE B, A ET B	79
ALGORITHME 3. 3. ALGORITHME D'AMELIORATION DE LA QUALITE DES MAMMOGRAMES PAR OPTIMISATION MONO-OBJECTIF.....	80
ALGORITHME 3. 4. ALGORITHME D'AMELIORATION DE LA QUALITE DES MAMMOGRAMES PAR OPTIMISATION MULTIOBJECTIF	84
ALGORITHME 3. 5. ALGORITHME DU MODELE PROPOSE POUR L'AMELIORATION DE LA QUALITE DES MAMMOGRAMES PAR OPTIMISATION MULTIOBJECTIF	93
ALGORITHME 4. 1. ALGORITHME D'ELIMINATION DES ARTEFACTS RADIO-OPAQUES	101
ALGORITHME 4. 2. ALGORITHME D'ELIMINATION DU MUSCLE PECTORAL AVEC REDUCTION DE TAILLE.....	105
ALGORITHME 4. 3. ALGORITHME D'ELIMINATION DU MUSCLE PECTORAL SANS REDUCTION DE TAILLE.....	106
ALGORITHME 4. 4. ALGORITHME DE SEGMENTATION PAR LA TECHNIQUE DE K-MEANS	110
ALGORITHME 4. 5. . ALGORITHME DE SEGMENTATION DE LA ROI PAR COUPLAGE DE L'ALGORITHME K-MEANS ET L'ALGORITHME GENETIQUE	112

Liste des acronymes

ACR	American College of Radiology
AG	Algorithme génétique
AMBE	Absolute Mean Brightness Error
AMBE _n	Normalized Absolute Mean Brightness Error
BBHE	Brightness Preserving Bi-Histogram Equalization
BIRADS	Breast Imaging Reporting and Data System
BRCA1	Breast cancer 1
BRCA2	Breast cancer 2
CLAHE	Contrast Limited Adaptive Histogram Equalization
DAO	Détection et diagnostic assisté par ordinateur
DAO _e	Détection assisté par ordinateur
DAO _x	Diagnostic assisté par ordinateur
DICOM	Digital imaging communication in medicine
EC	Edge content
EME	Effective Measure of Enhancement
FSIM	Feature Similarity Index Measure
HE	Histogram Equalization
IQA	Image Quality Assessment
IRM	Imagerie par résonance magnétique
MLO	Médio-latéral-oblique
NEMA	National Electrical Manufacturers Association
OMS	Organisation Mondiale de la Santé
PAM	Plaque aréolo-mamelonnaire
PSNR	Peak Signal-to-Noise Ratio
QIE	Quadrant inférieur externe
QII	Quadrant inférieur interne
QSE	Quadrant supérieur externe
QSI	Quadrant supérieur interne
ROI	Region of interest
RX	Rayon X
TMN	Tomosynthèse mammaire numérique
TUBO	Technique usuelle basée sur l'optimisation
UICC	Union Internationale Contre le Cancer

Introduction générale

Un cancer (ou 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 menacé. Il existe généralement une classification du cancer selon les organes touchés et le cancer du sein constitue l'un des problèmes majeurs de la santé publique partout dans le monde. Au cours de l'année 2020, les statistiques de l'Organisation Mondiale de la Santé (OMS) ont classé le cancer du sein en tête de tous les cancers recensés à l'échelle mondiale avec 2.261.419 incidences suivies de celui du poumon avec 2.206.771 [1]. Au Maroc, pour la même année 2020, le cancer du sein est le premier cancer féminin fréquent et mortel suivi de celui du col utérin. Il en résulte que le taux de mortalité reste en hausse du fait que dans la plupart des pays les femmes ne sont prises en charge qu'après un stade avancé de la maladie, ce qui diminue leurs chances de guérison et de leur survie. D'où la détection précoce du cancer du sein s'avère nécessaire et urgente. Son objectif est de diagnostiquer des signes précurseurs et symptômes du cancer avant que le mal ne soit à un stade avancé, ce qui permet un traitement plus léger et efficace [2], [3].

Actuellement, la mammographie et malgré le problème d'irradiation et de confort constitue la modalité d'imagerie médicale recommandée par les oncologistes pour la détection précoce du cancer du sein vu son accessibilité et sa capacité de détection des tumeurs de petite taille même avant qu'elles ne soient palpables. D'autres modalités d'imagerie sont employées et peuvent être complémentaires à savoir l'échographie mammaire, l'imagerie par résonance magnétique (IRM) et la tomosynthèse mammaire numérique (TMN) [3]–[6].

En mammographie un bon diagnostic est fortement lié à l'analyse correcte des images mammographiques qui constitue un challenge aux radiologues puisque cette tâche est fortement corrélée à leurs degrés d'expertise et surtout à la qualité des images à analyser. Malheureusement, de nombreux facteurs conduisent à la dégradation de la qualité des images mammographiques dont les facteurs liés à la structure du sein lui-même et d'autres d'ordre humains. En effet, les seins de classe BIRADS 3 et BIRADS 4 sont caractérisés par une grande densité de la glande mammaire par rapport au reste des tissus mammaires, leurs mammogrammes correspondants sont alors moins contrastés et présentent des bords flous suite à la faible différence d'atténuation des RX entre les lésions probables et les tissus sains avoisinants [7]–[9]. Quant aux facteurs humains, la limitation du système visuel humain et le nombre élevé des mammogrammes à analyser pendant les campagnes de dépistage constituent les principales causes influant négativement la précision de la détection et la classification des lésions

mammaires. Dans ce sens des études cliniques ont montré qu'environ 30 % des lésions ne sont pas détectées lors des dépistages du cancer du sein [10]–[13] dont les mammogrammes sont classifiés en tant que faux négatifs, ce qui favorise le passage des patientes à la phase de métastase où tous traitements restent sans efficacité. Avec l'avènement de la digitalisation des instruments et des pratiques, l'implication de la vision artificielle par l'utilisation des systèmes d'aide à la détection et au diagnostic assisté par ordinateur "DAO" a amélioré nettement la détection des zones suspectes ainsi que la prédiction de leurs sévérités. En effet des études cliniques ont montré que l'introduction des systèmes d'aide "DAO" peut améliorer la détectabilité du cancer du sein d'environ 21% et réduire de 5 à 15% les lésions classifiées faux négatifs [10], [14]–[16].

D'après la littérature le système "DAO" se compose d'un système d'aide à la détection "DAOe" qui permet la détection des zones probablement suspectes suivi d'un système d'aide au diagnostic "DAOx" qui prédit la sévérité des zones suspectes détectées [17], [18]. Un ensemble de techniques de traitement d'images numériques sont à la base du fonctionnement de ces systèmes. Plusieurs travaux de recherche se sont investis à développer des systèmes plus performants pour une meilleure détection et un meilleur diagnostic possibles, et ce dans ce contexte s'inscrit l'objectif de notre travail. En effet notre étude s'est focalisée sur la phase de détection vu son intérêt grandissant. Il s'agit de concevoir un système de détection précoce des masses mammaires totalement automatisé et à haute résolution. Ce travail est précédé par l'élaboration d'un modèle permettant l'amélioration optimale de la qualité des images mammographiques. Ce modèle sera intégré au sein de notre système de détection qui sera appelé à améliorer la qualité des images prétraitées. Notre approche pour le développement de ce modèle s'est basée sur les travaux de l'auteur A. Gandhamal et ses collaborateurs [8] du fait que la technique proposée a montré une bonne efficacité au niveau de l'amélioration de la qualité des images médicales bien qu'elle ne permette pas le contrôle du processus de l'amélioration des images. Les résultats fournis par notre modèle concernant l'amélioration de la qualité des images, seront comparés à ceux fournis par les méthodes usuelles (basées aussi bien sur la modification de l'histogramme que sur l'optimisation) d'une part, et à ceux fournis par la technique de Gandhamal et al.

La conception de notre système de détection DAOe passe par la conception des deux étapes dont il est composé prétraitement et segmentation. Il s'agit durant la phase de prétraitement à déterminer automatiquement la région d'intérêt "ROI" de l'image origine à partir des vues médio-latéral-oblique "MLO" des mammogrammes par élimination de toute information inutile. Cette région "ROI" subira par la suite une amélioration de sa qualité par optimisation à l'aide

du modèle proposé. Il s'agit en fait d'augmenter le contraste et la netteté de l'apparence des bords délimitant les différentes zones de la région ROI, tout en conservant autant que possible la luminosité et les informations de l'image origine [26], [27]. Par la suite vient l'étape de segmentation qui consiste à isoler avec une grande précision la zone suspecte du reste de la ROI en utilisant des méthodes qui ne nécessitent aucune initialisation de la part de l'utilisateur.

Le modèle réalisé ainsi que le système de détection conçu sont appliqués sur des images mammographiques issues de la base de données Mini-MIAS. Les résultats fournis par notre système de détection seront comparés à ceux fournis par cette base.

Le présent rapport synthétise le travail que nous avons effectué. En plus d'une introduction et d'une conclusion ce rapport est organisé en 4 chapitres comme suit :

- Le premier chapitre donne une présentation générale sur l'anatomie et le cancer du sein et décrit les différentes techniques d'imagerie (en particulier la mammographie) permettant son exploration.
- Le deuxième chapitre aborde l'image numérique et ses caractéristiques d'une part, et décrit les différentes étapes du système d'aide à la détection et au diagnostic assisté par ordinateur "DAO" permettant le traitement automatique de l'image d'autre part. L'intérêt particulier est donné à la description des différentes techniques utilisées durant les étapes de prétraitement et de segmentation. Ce chapitre aborde aussi les procédés d'optimisations globales intégrées dans ces étapes afin de détecter, localiser et isoler les zones suspectes d'une manière précise.
- Le troisième chapitre est consacré à la proposition d'un modèle permettant d'améliorer la qualité d'une image mammographique par optimisation multiobjectif.
- Le quatrième chapitre est consacré à la conception d'un système automatisé permettant la détection précoce des masses mammaires tout en exploitant le modèle proposé au chapitre 3.

Chapitre 1

Cancer du sein et mammographie

Ce chapitre a pour objectif de présenter un aperçu général concernant le cancer du sein et les techniques de son exploration. Au préalable seront présentés l'anatomie du sein, les principaux signes anatomopathologiques et physiopathologiques en cas d'atteinte, les facteurs de risque augmentant la probabilité de l'occurrence du cancer du sein. Par la suite les techniques d'exploration du cancer seront décrites et plus particulièrement la mammographie qui constitue l'examen de référence pour une détection précoce. Enfin sera traité un ensemble de standards de classification des lésions mammaires permettant d'unifier leurs descriptions et les conduites à tenir pour chaque cas.

1. Anatomie du sein

Le sein est une glande exocrine d'origine ectodermique, il occupe la partie antéro-supérieure du thorax, de part et d'autre du sternum en avant des muscles pectoraux. Il s'étend depuis la partie inférieure de la deuxième côte jusqu'au sixième cartilage costal et transversalement, du bord externe du sternum à la ligne axillaire antérieure. Son poids varie selon la morphologie de la femme, son état de grossesse ou de lactation. Son poids varie de 200 g chez une jeune fille et peut atteindre 500 g voire 900 g chez une femme allaitante. Le sein a une forme hémisphérique de taille d'environ 12 cm en hauteur et largeur. Les deux mamelons sont distants d'environ 20 cm. Généralement les seins sont asymétriques [17], [18]. Le sein est constitué de la glande mammaire composée d'une vingtaine environ de lobes séparés par un tissu adipeux, un autre tissu de soutien constitué de vaisseaux sanguins et de fibres et de graisse. Chaque lobe de la glande mammaire est constitué de plusieurs lobules qui secrètent du lait et de canaux galactophores qui l'acheminent vers le mamelon pendant la période de l'allaitement (figure 1.1) [3], [19].

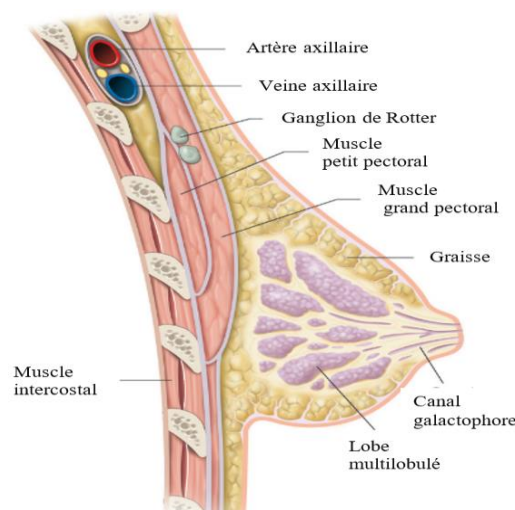


Figure 1. 1. Coupe sagittale de la paroi thoracique et glande mammaire.

La glande mammaire se développe de point de vue forme et fonction sous l'effet des hormones sexuelles (les œstrogènes et la progestérone) secrétées au niveau des ovaires. En effet les œstrogènes permettent le développement des seins pendant la période de la puberté et jouent un rôle primordial le long de la grossesse (assouplissement des tissus, augmentation du volume sanguin nécessaire à l'alimentation du bébé, etc.) ; quant à la progestérone, cet hormone joue un rôle significatif dans le cycle menstruel notamment la préparation de l'utérus à une grossesse probable [2]. Le sein est parcouru par divers réseaux, dont le réseau de vaisseaux sanguins composé d'un réseau sous-dermique, un réseau pré-glandulaire et un réseau anastomotique intra glandulaire permettant l'alimentation de la glande mammaire et le réseau des vaisseaux lymphatiques qui aide à côté des ganglions lymphatiques à combattre les infections et à éliminer les déchets tels que les cellules mortes et autres débris [18].

2. Tumeurs mammaires

2.1. Généralités sur les tumeurs

Autrefois, le terme « tumeur » désignait toute augmentation de volume localisé aboutissant à une déformation d'un organe ou d'une partie du corps (figure 1.2). En effet, la croissance d'un tissu ou d'un organe est conditionnée par des processus complexes permettant la régulation des différentes étapes de la vie d'une cellule : la prolifération, la différenciation, la sénescence et la mort cellulaire programmée.

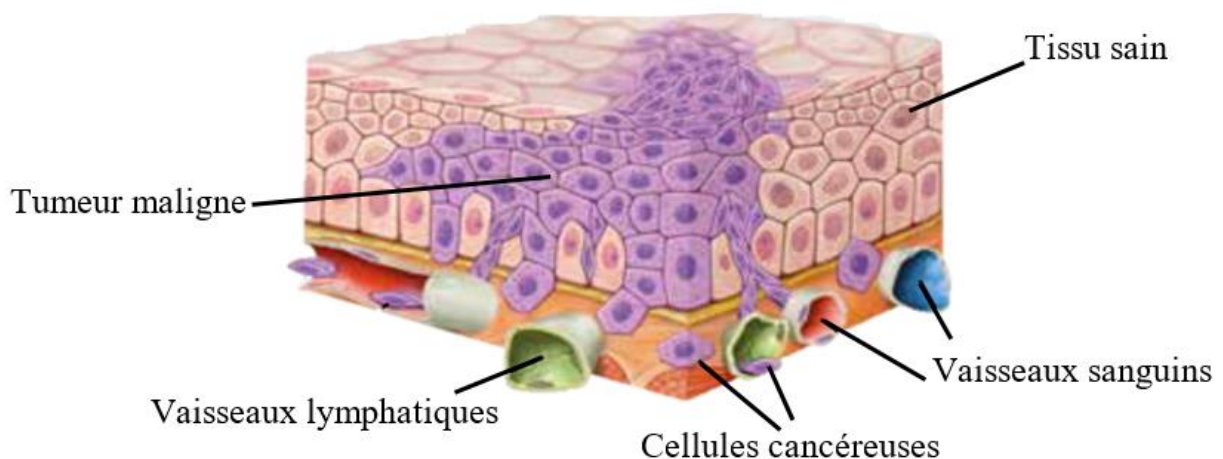


Figure 1. 2. Envahissement des cellules tumorales du tissu sain avoisinant.

Au sein d'un tissu, l'équilibre entre ces processus est à l'origine de l'homéostasie tissulaire. Actuellement, les termes « tumeur, néoplasme ou néoplasie » désignent une prolifération cellulaire excessive qui aboutit à une masse tissulaire ressemblant plus ou moins au tissu normal homologue adulte ou embryonnaire, ayant tendance à persister et à croître. Le processus de prolifération se poursuit même après la disparition du stimulus qui lui a donné naissance ce qui témoigne l'autonomie biologique du processus.

2.2. Cancer du sein

Le tissu tumoral est constitué de cellules tumorales dites aussi cellules prolifératives anormales et d'un tissu de soutien appelé stroma tumoral formé de cellules et de substance extracellulaire dans laquelle est située la vascularisation tumorale. D'une manière générale les cellules du stroma ne présentent pas d'anomalies génétiques des cellules tumorales [4], [20]. Généralement, les tumeurs sont classées en fonction de l'organe dont elles sont dérivées (sein, foie, poumon, prostate...), puis en fonction du type histologique qui correspond à la cellule normale qui lui a donné naissance [20].

Le sein est parmi les organes les plus touchés par le cancer. En effet, le cancer du sein est le plus fréquent chez la femme [3], [21]. Au cours de la vie, les seins ne cessent de se transformer suite aux fluctuations hormonales durant le cycle menstruel, la grossesse, l'allaitement, ainsi qu'au moment de la ménopause où le tissu glandulaire et conjonctif diminue dans le sein au profit du tissu adipeux. Lorsqu'un cancer du sein se développe, les cellules atteintes peuvent conserver leurs récepteurs hormonaux et être donc stimulées dans leur croissance par les hormones sexuelles féminines [4], [22].

Lorsqu'une ou plusieurs cellules du sein commencent à se multiplier de façon incontrôlée, elles finissent par donner naissance à une tumeur. Quand celle-ci est maligne, on parle de cancer du sein, ou carcinome mammaire [21].

Le cancer du sein peut se développer dans différents endroits du sein. Les types les plus fréquents (95 %) sont des adénocarcinomes qui se développent à partir des cellules épithéliales de la glande mammaire. Les adénocarcinomes naissent le plus souvent à partir des cellules des canaux et plus rarement à partir des cellules des lobules [19]. Selon le type de cellules à partir desquelles ils se forment on distingue différents types de cancers du sein :

- Carcinome lobulaire infiltrant (figure 1.3 (b))
- Carcinome canalaire infiltrant (le plus fréquent) (figure 1.3 (c))
- Autres types : les carcinomes tubuleux, mucineux, micropapillaires, adénoïdes kystiques...

Il existe d'autres tumeurs que les adénocarcinomes primitifs, mais elles sont très rares (lymphome, sarcome, métastase mammaire)[19], [21].

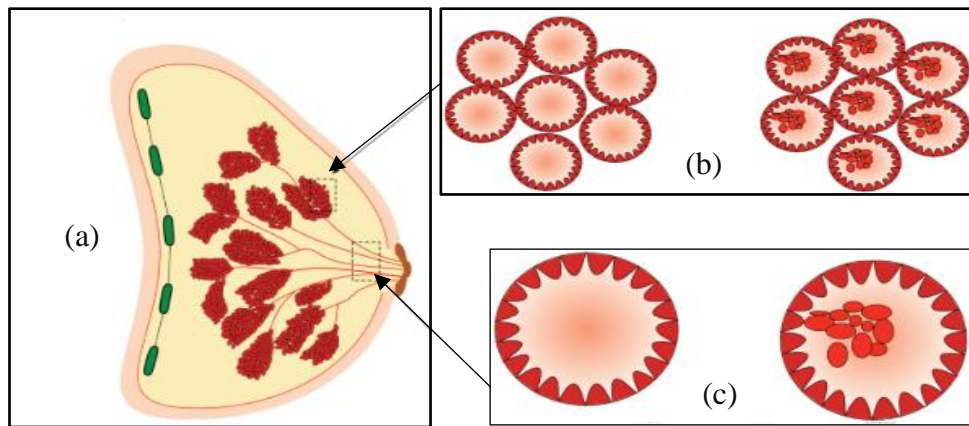


Figure 1. 3. Début d'un cancer (a) coupe sagittale du sein (b) début d'un cancer lobulaire (c) début d'un cancer canalaire.

Au début de l'apparition un adénocarcinome, les cellules cancéreuses sont d'abord moins nombreuses et elles se localisent aux canaux ou aux lobules du sein : ce type s'appelle un carcinome in situ. Avec le temps et en l'absence de tout traitement, la tumeur peut franchir la membrane basale et infiltrer le tissu qui entoure les canaux et les lobules; il s'agit alors de cancer ou carcinome infiltrant ou invasif [19]. Aussi, les cellules cancéreuses peuvent se détacher de la tumeur et être transportées par les vaisseaux sanguins ou les vaisseaux lymphatiques pour atteindre d'autres parties du corps loin du sein (les ganglions axillaires, les ganglions lymphatiques, le foie, les poumons...) [3], [19].

Pour des raisons de traitement, la localisation de l'emplacement des tumeurs est primordiale. A cet effet, une localisation standard consiste à diviser le sein en quatre secteurs appelés quadrants : le quadrant supérieur externe (QSE), le quadrant supérieur interne (QSI), le quadrant inférieur externe (QIE) et le quadrant inférieur interne (QII) (figure 1.4) [4].



Figure 1. 4. Différents quadrants pour le repérage des lésions du sein.

Les statistiques ont montré que plus de la moitié des tumeurs du sein apparaissent dans le QSE du fait que cette partie est plus riche en tissu glandulaire, suivie de QSI et QIE avec presque

une même proportion ; le QII et la région mamelonnaire (PAM) sont plus souvent les moins touchés.

2.3. Symptomatologie clinique du cancer du sein :

Un symptôme d'une maladie est toute manifestation anormale provoquée par cette maladie. Durant les premiers stades de développement d'un cancer du sein, la symptomatologie est frustrante. La maladie est d'une manière générale sans douleur mais il peut exister chez 1 à 15% de cas. C'est à un stade avancé que la symptomatologie liée à cette maladie peut se manifester comme suit [2]–[4], [22] :

- **Une boule dans le sein** : c'est l'un des signes le plus fréquent des cancers mammaires 65 à 80% des cas. Cette masse en général est non douloureuse le plus souvent, de consistance dure et présente des contours irréguliers. Elle apparaît par ailleurs comme fixée dans le sein.
- **Des ganglions durs au niveau de l'aisselle** : l'existence d'une ou plusieurs masses dures à l'aisselle signifie parfois qu'un cancer du sein s'est propagé aux ganglions axillaires. Les ganglions restent toutefois sans douleur.
- **Des modifications de la peau du sein et du mamelon** : traduisent par des altérations cutanées et aréolo-mamelonnaires, une ulcération, une inflammation ou une rétractation de la peau au niveau d'un des seins, un changement de coloration, écoulement unipore sérosanguin au niveau du mamelon, un eczéma ou un érythème de l'aréole ou du mamelon.
- **Un changement de la taille ou de la forme du sein** : des symptômes inflammatoires du cancer du sein peuvent être ressentis au niveau du sein, tels qu'une rougeur, un œdème ou une chaleur importante.

Autres symptômes :

A défaut de détection et de traitement du cancer du sein à un stade précoce, la tumeur peut grossir et se propager vers d'autres parties du corps, entraînant ainsi d'autres symptômes dits tardifs, tels que : des douleurs osseuses, des nausées, une perte d'appétit, une perte de poids et une jaunisse ; un essoufflement, une toux et une accumulation de liquide autour des poumons (épanchement pleural), des maux de tête, une vision double et une faiblesse musculaire [22].

2.4. Facteurs de risque liés au cancer du sein

Jusqu'à notre jour, les causes principales du cancer du sein ne sont encore connues, il est donc impossible de prévenir la maladie. Cependant, certains facteurs peuvent favoriser le développement d'un cancer du sein. Parmi ces facteurs de risques : l'âge, le sexe, les antécédents personnels et familiaux de cancer du sein, les prédispositions génétiques[3].

- **Facteur lié à l'âge** : l'âge est l'un des facteurs de risque majeur favorisant le développement du cancer du sein. En effet, les statistiques montrent qu'environ de 80% des cancers du sein sont diagnostiqués chez les femmes âgées de plus de 50 ans [3], [4].
- **Facteur lié au sexe** : l'homme n'est pas loin d'être atteint par le cancer du sein même avec une probabilité faible. Les statistiques montrent que ce type de cancer ne dépasse pas 1% parmi tous les cancers du sein diagnostiqués. Le cancer du sein masculin est favorisé d'une manière générale par les antécédents familiaux de cancer du sein chez le sujet, la prédisposition génétique liée essentiellement à une mutation du gène BRCA2 (responsable de 15% de ce type de cancer), le syndrome de Klinefelter, une exposition antérieure aux rayonnements ionisants au niveau du thorax [3].
- **Facteur lié aux antécédents personnels** : le développement d'un cancer dans l'un des ovaires à un âge inférieur à 40 ans constitue un sur-risque de développer un cancer du sein ultérieurement. Le rapport d'incidence standardisé de cancer du sein varie entre 2.8 et 7.4 par comparaison avec la population générale. Dans ce cas, une analyse génétique est obligatoire afin de savoir si la patiente a hérité une mutation BRCA1 ou BRCA2. Quand c'est le cas, le sujet doit être objet d'une surveillance afin de minimiser le risque [3], [23].
- **Facteur lié aux antécédents familiaux** : le risque de développement d'un cancer du sein augmente chez les femmes ayant des cas du cancer du sein dans leurs familles. Selon les statistiques 20 à 30% des cancers diagnostiqués sont enregistrés chez les femmes ayant des antécédents familiaux [23]. Ce risque augmente avec l'augmentation du nombre d'antécédents, leur rang (1^{er} degré, 2^{ème} degré), leur précocité de l'âge, atteinte d'un membre d'un cancer bilatéral ou un parent homme par le cancer du sein [3], [23].
- **Facteur lié aux prédispositions génétiques** : Ce risque est lié principalement aux mutations génétiques causant de 5 à 10% de la totalité des cancers des seins diagnostiqués [3], [23]. Les gènes responsables sont BRCA1 sur le chromosome 17 et BRCA2 sur le chromosome 13 [3]. Ces deux gènes contribuent à la réparation des lésions subies par l'ADN. Une mutation au niveau de ces gènes altère cette fonction de réparation en favorisant l'augmentation du risque d'atteindre un cancer du sein et de l'ovaire. En effet, 7 à 8 femmes sur 10 porteuses de cette mutation pourront développer un cancer du sein au cours de leur vie [23]. Il existe d'autres mutations génétiques capables d'augmenter le risque de cancer du sein : TP53 (syndrome de Li-Fraumeni), CHEK2/ATM (ataxie-télangiectasies), PTEN (syndrome de Cowden), STK11 (syndrome de Peutz-Jeghers) ; mais ces mutations sont plus rares [3], [23]. Les femmes

porteuses d'une mutation génétique doivent faire objet d'une prise en charge spécifique par une consultation d'oncogénétique et doter d'une surveillance et/ou la chirurgie prophylactique [3].

- **Autres facteurs de risque du cancer du sein :** ce sont des facteurs de risque modérément ou faiblement classés. Parmi ces facteurs on cite : la densité mammaire élevée (la glande mammaire excède 75% de la totalité du sein), la prise d'une contraception orale pendant une longue période, traitement hormonal substitutif à la ménopause, atteinte du diabète de type 2, l'obésité, la consommation régulière de l'alcool, le mode alimentaire surtout la faible consommation des fruits et le manque de l'exercice physique [3], [23].

2.5. Type de tumeurs du sein

Divers lésions sont liées au sein dont les tumeurs qui peuvent être soit bénignes ou malignes. L'existence d'un nodule, une ulcération ou une inflammation ne signifie pas automatiquement un cancer du sein. Toutefois, une consultation sans retard est obligatoire [3]. En effet, trois nodules palpables sur quatre sont bénignes [4].

2.5.1. Formes bénignes

Ces tumeurs ont un développement local et séparées des tissus avoisinants où elles sont nées. Dans certains cas, ces tumeurs sont encapsulées ce qui facilite leur ablation qui est généralement sans récurrence. Les tumeurs ont des formes circonscrites bien limitée, elles sont composées des cellules régulières ressemblant à la structure du tissu initial et ne finissent pas par des métastases [20]. Parmi ces lésions nous citons :

- **Les calcifications mammaires :** ce sont des dépôts de calcium dans le tissu mammaire. On en distingue deux types : les macrocalcifications et les microcalcifications.

Les macrocalcifications ont une taille grossière, cette forme est plus fréquente chez les femmes âgées de plus de 50 ans [3], [4], [22].

- **Le fibroadénome :** cette lésion est sous forme d'une petite masse de tissus fibreux et de tissu glandulaire qui touche les jeunes femmes âgées de moins de 25 ans. Il s'agit d'une tumeur bénigne et rarement associée à un cancer du sein [3], [4], [22].

- **Le kyste :** il s'agit d'un sac rempli de liquide sécrété par les lobules lorsqu'un de ses canaux excréteurs est bouché [4] . Il est très fréquent chez les femmes âgées de 35 à 50 ans. Il est rarement associé à un cancer du sein mais lorsqu'ils sont multiples dans le sein une surveillance s'avère importante pour contrôler leur évolution [3], [4], [22].

- **La mastose ou fibrose kystique du sein :** cette maladie est née suite à des modifications du tissu épithéliale et conjonctive associées principalement aux facteurs suivants [3], [21] :

- existence des kystes suite à la dilatation des canaux galactophores ;

- une hyperplasie épithéliale canalaire ;
- une adénose suite à une prolifération de lobules.
- **L'hyperplasie mammaire** : c'est une pathologie liée à une prolifération soit des cellules des canaux galactophores ou de lobules ce qui tapisse ces derniers [3]. Cette lésion est considérée comme proche du carcinome in situ. Le risque de développer un cancer du sein est très élevé [21].
- **L'écoulement mamelonnaire** : l'écoulement anormal du mamelon peut être provoqué par divers maladies autre que le carcinome mammaire qui sont : une mastite due à une infection bactérienne, un papillome intracanales, un fibroadénome, une dilatation des canaux galactophores, un trouble endocrinien thyroïdien ou hypophysaire ou un effet indésirable d'un médicament (antidépresseurs ou antihypertenseurs) [3].

2.5.2. Formes malignes

A l'encontre des tumeurs bénignes, les tumeurs malignes sont mal limitées et non encapsulées. De point de vue histologique, elles sont moins semblables au tissu d'origine. Leur croissance est très rapide et envahissent les tissus avoisinants. Elles peuvent migrer vers d'autres endroits du corps provoquant ainsi des métastases. Leur exérèse complète est difficile, la récurrence d'une tumeur maligne après son exérèse supposée complète est possible [20]. On distingue différents types de tumeurs malignes :

- **Le cancer canalaire in situ** : Ce type constitue le commencement du cancer du sein. Les cellules cancéreuses sont localisées au niveau des canaux galactophores et n'ont pas encore franchi la membrane basale qui les entoure. Il est souvent révélé par la présence de microcalcifications. Ils représentent environ 80 % des cancers du sein [3], [4].
- **Le cancer lobulaire in situ** : il constitue aussi le noyau du cancer du sein ayant pris naissance au niveau des lobules. Ce type est rare il constitue 5 à 15% de tous les cancers du sein [4]. Certaines données de la littérature récente tendent à classer le carcinome lobulaire in situ comme étant un facteur de risque de cancer et non comme un précurseur du cancer [3].
- **Le cancer ou adénocarcinome infiltrant** : c'est un stade évolué des cancers in situ. Les cellules cancéreuses infiltrent le tissu avoisinant des canaux galactophores ou des lobules. Les cancers infiltrants sont dans la majorité de cas d'origine canalaire (huit cancers infiltrants sur dix sont d'origine canalaire) [3], [4], [19].
- **Les adénocarcinomes métastasés** : Si aucun traitement efficace pendant les deux premières phases n'est appliqué, les cellules cancéreuses poursuivent leur évolution en se

propageant via les vaisseaux sanguins ou lymphatiques vers d'autres parties du corps. Le processus de prolifération rapide continue en détruisant les tissus sains avoisinants [3], [4], [19].

- **Les microcalcifications** : ce sont des minuscules dépôts de calcium au niveau des canaux galactophores. Leurs degrés de malignité augmentent lorsqu'elles apparaissent isolées ou regroupées en grappes. Quand c'est le cas, une analyse histologique est nécessaire pour s'assurer s'il s'agit d'un cancer [3], [4], [21], [23].

3. Dépistage et diagnostic du cancer du sein

Prévenir du cancer n'est encore possible vu que les étiologies favorisant son incidence ne sont pas encore connues [23]. La détection précoce du cancer reste le seul moyen permettant l'efficacité de son traitement. Le dépistage est une action collective de santé publique minutieusement préparée permettant l'identification d'une maladie ou d'une anomalie non connues chez des sujets considérés comme indemnes. Au Maroc, la population cible du dépistage du cancer du sein est constituée de femmes âgées de 49 à 69 ans [2]. La détection du cancer du sein peut commencer depuis la femme elle-même en procédant à la palpation pour identifier l'existence d'un nodule, suivi de l'examen mammographique réalisé dans le cadre d'un dépistage du cancer du sein. Ces actes sont systématiques même en l'absence de tout symptôme ou lors de l'apparition d'un ou plusieurs symptômes (nodule palpable, modifications cutanées et aréolo-mamelonnaires...) [3], [4]. A cet égard, divers examens visant à identifier toute anomalie probable sont à effectuer dont la mammographie constitue l'examen de référence [3]–[5], [23]. D'une manière générale, l'examen mammographique peut être accompagné des autres examens complémentaires tels que l'échographie mammaire afin de collecter plus d'informations concernant les anomalies, ou des examens de confirmation tels que l'IRM ou l'analyse histologique. Dans ce qui suit nous présenterons succinctement ces techniques ainsi que quelques actes préliminaires qui les précèdent.

3.1. Actes préliminaires

Ces actes peuvent se résumer à une inspection et entretien qui ont pour but la détermination des différents troubles ou symptômes pouvant indiquer la présence d'une anomalie, ainsi de connaître le mode de vie et les antécédents familiaux chez une femme. Une palpation minutieuse des seins à la main doit être pratiquée afin de déceler toute modification ou nodule dans le sein. L'examen doit inclure les aires ganglionnaires au niveau du creux de l'aisselle ainsi qu'au niveau de la clavicule et du sternum.

La palpation peut donner une indication sur la nature d'un nodule [3], [4]. Ce dernier peut être associé à un cancer du sein lorsqu'il est dur, régulier, indolore et il apparaît fixe dans le sein. Il

est d'autant plus suspect s'il est irrégulier mené d'un écoulement du mamelon ou une adénopathie axillaire sus- ou sous- claviculaire [3].

3.2. Mammographie

La mammographie est la technique radiologique conventionnelle la plus exigeante dont l'appareillage est constitué d'un tube à RX, d'un système de compression, d'anti-diffusion et, d'un récepteur d'image.

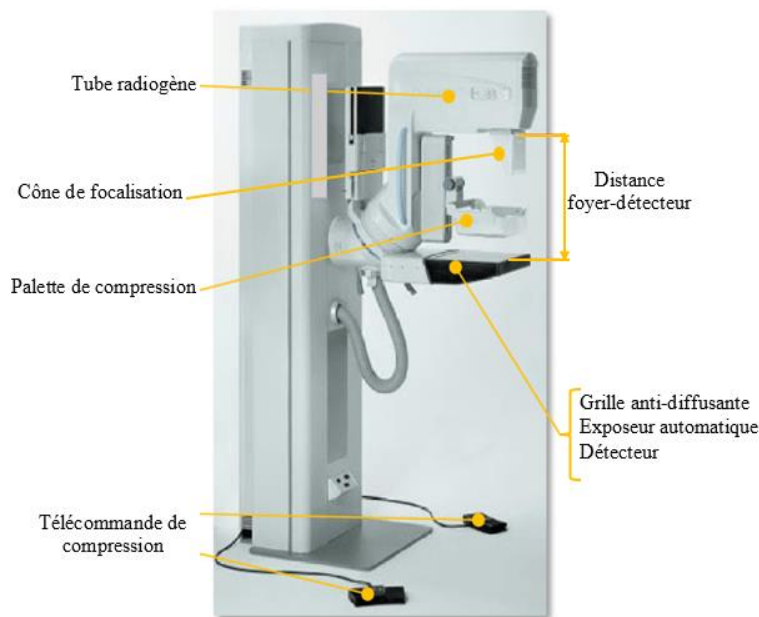


Figure 1. 5. Principaux éléments d'un mammographe.

Cet examen permet une exploration de la structure interne du tissu mammaire en faisant traverser chaque sein, soumis à une compression, par une faible dose de rayons X afin de mettre en évidence les lésions auxquelles il est atteint [4]–[6], [23]. En mammographie numérique, le détecteur couple écran-film analogique est remplacé par un détecteur d'image électronique (la conversion des rayons X en signal électronique peut être directe ou indirecte). Les clichés doivent être optimisés en termes de résolution spatiale, de contraste et de bruit. La mammographie constitue un examen de référence de la plupart des anomalies et des changements pathologiques liés au sein [3], [4]. Elle peut être prescrite dans le cadre d'un dépistage organisé proposé aux femmes cibles (mammographie de dépistage) [3] ou dans le cas de signes cliniques (mammographie diagnostique). Cet examen permet la détection des tumeurs de petite taille même avant qu'elles ne soient palpables [3]–[6].

Un des inconvénients de la mammographie conventionnelle provient de la superposition des tissus mammaires sur l'image projetée. Ce phénomène est d'autant plus important et gênant pour l'interprétation de l'examen que les seins ont une texture radiologique plus dense. Pour y remédier l'exploration du sein se réalise sous différentes incidences [24], [25] ou par utilisation

de la tomosynthèse mammaire qui est une technique relativement récente d'imagerie en 3D en matière de mammographie.

3.2.1. Incidences standards

Ce sont les incidences fréquemment utilisées dans le but de la détection d'un cancer du sein, généralement on parle de trois incidences :

- L'incidence de face connues sous le nom cranio caudale (CC), sous cette incidence, la partie externe du sein doit être favorisée puisque en aucun cas on ne peut explorer la totalité des quadrants internes et externes (figure 1.6 (a)) [6], [25], [26].
- l'incidence oblique externe appelée aussi Médio Latérale Oblique (MLO) acquise d'une manière oblique (30–70°), sous cette incidence, le parenchyme mammaire est exposé en totalité y compris le muscle pectoral, les ganglions auxiliaires peuvent être aussi visualisés. L'exposition d'une grande partie du muscle pectoral indique que le sein est bien couvert et correctement positionné dans le champ d'exploration (figure 1.6 (b))[6], [25], [26]
- L'incidence de profil externe qui est d'une importance primordiale dans l'exploration d'un foyer de micro-calcification. Elle est considérée comme étant la seule projection pouvant montrer le caractère déclive ou cupuliforme, des calcifications.

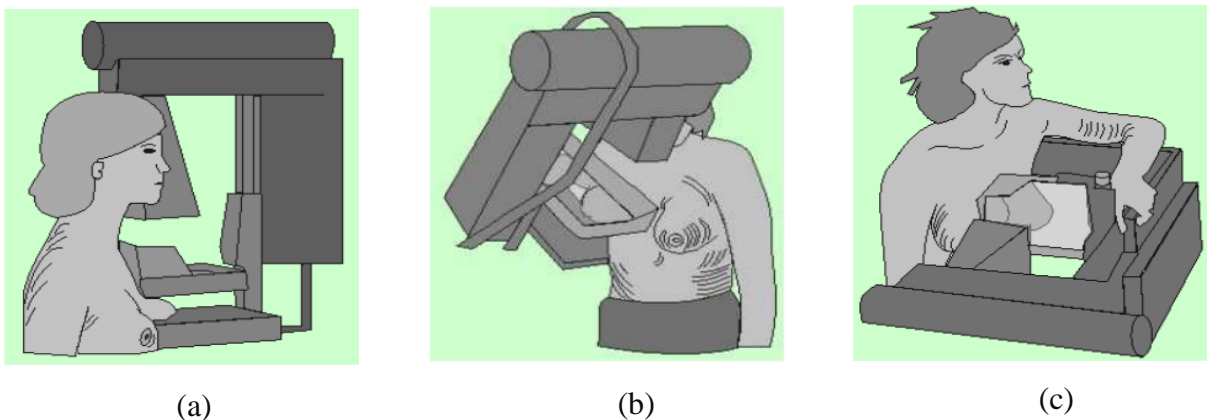


Figure 1. 6. Illustrations des incidences standards en mammographie : (a) incidence cranio caudale (CC), (b) incidence médio latérale oblique (MLO) et (c) incidence de profil.

L'examen mammographique de base consiste à explorer chaque sein au moins sous deux incidences [3], [6]. L'incidence médio latérale oblique couplée avec l'incidence de face constituent les incidences de référence pour la détection des lésions lors du dépistage du cancer du sein. Dans certain cas, la spécificité d'une lésion exige d'autres investigations et qui nécessitent d'imager le sein sous d'autres incidences complémentaires.

3.2.2. Incidences complémentaires

Les incidences complémentaires servent à approfondir l'exploration de l'exploration du sein. Elles peuvent être indispensables lorsque le radiologue cherche plus d'informations concernant une lésion détectée. En effet selon la localisation de la lésion dans le sein une incidence peut être plus informative qu'une autre. L'objectif est de bien classer la sévérité de la lésion, afin d'élaborer une bonne conduite à tenir [24], [25].

- L'incidence face externe ayant pour but de privilégier l'exploration soit des quadrants externes lorsque la lésion est profonde.
- L'incidence face interne permettant l'exploration des quadrants internes lorsque la lésion détectée est située au périphérique du sein.
- L'incidence profil interne qui est d'une très grande utilité dans l'exploration des lésions au niveau des quadrants internes (figure 1.6 (c))
- Incidence tangentielle ayant pour objectif la confirmation du caractère superficiel d'une lésion.

Afin de garantir un examen précis sans recourt aux incidences complémentaires, la tomosynthèse mammaire peut être pratiquée.

3.2.3. Tomosynthèse mammaire ou mammographie 3D

La tomosynthèse mammaire numérique (TMN) ou mammographie tridimensionnelle est une technique dont le principe est similaire à celui de la mammographie 2D. Dans cette technique, le tube à rayons X décrit un arc circulaire variant de 15 à 50 degrés (selon l'appareil) autour du sein comprimé et le détecteur électronique pouvant ainsi générer de 9 à 25 images radiographiques (figure 1.7) [5], [27], [28].

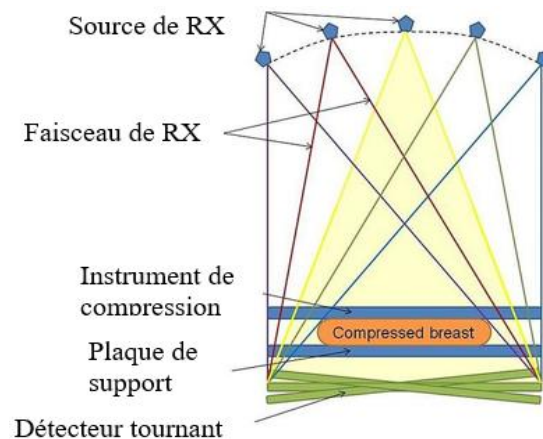


Figure 1. 7. Principe de la tomosynthèse mammaire

Plusieurs incidences sont alors captées à partir de la vue cranio caudale (axe tête - pieds) jusqu'à la vue médio latérale oblique (vue latérale angulée). Le système informatique lié à l'appareil permet de réaliser une analyse tomographique du sein afin de transformer les données acquises en coupes successives de faible épaisseur 0.5 à 1mm [28]. L'objectif de la technique est de pallier les inconvénients de la mammographie conventionnelle 2D que ce soit numérique ou analogique, notamment la superposition des tissus mammaires sur l'image projetée, réduire l'aspect suspect des tissus normaux et faciliter la distinction et la classification des lésions mammaires [27], [28]. Les images produites par la TMN sont caractérisées par une meilleure apparition des contours, ce qui permet une meilleure distinction entre les lésions bénignes et malines [29]. L'examen TMN peut être utilisé en dépistage ou en diagnostic du cancer du sein. Au Canada, l'agence canadienne des médicaments et des technologies de la santé a constaté d'après sept vastes études réalisées sur 10.000 femmes concernant le dépistage du cancer du sein, que la TMN peut réduire le nombre de rappels pour les examens complémentaires tels que l'échographie, l'IRM, ou des biopsies inutiles comparativement au dépistage par la mammographie bidimensionnelle 2D [28]. D'autres études indiquent que la tomosynthèse permet la détection des tumeurs de très petite taille comparativement à la mammographie conventionnelle [28], [30].

3.3. Echographie mammaire

L'échographie mammaire est un examen basé sur l'utilisation des ondes ultrasonores pour produire des images internes du sein en temps réel. Cet examen est souvent réalisé en complément d'une mammographie lorsque celle-ci a mis en évidence une anomalie.

L'examen échographique est à balayage manuel en utilisant une sonde formée d'une série de transducteurs fonctionnant alternativement en émetteurs et récepteurs des ultrasons à haute fréquence (figure 1.8).

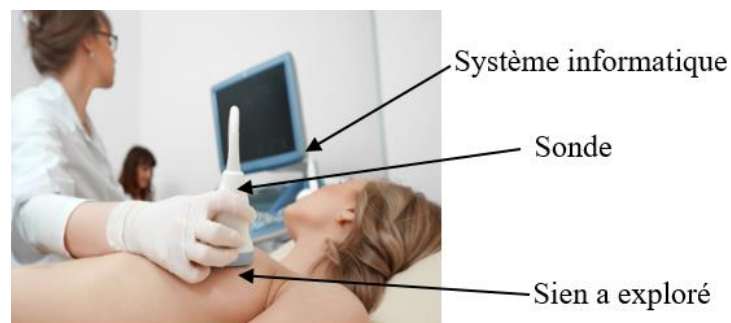


Figure 1. 8. Principe de l'échographie mammaire

Les nouvelles générations d'échographie mammaire utilisent des sondes multifréquences large bande. Les fréquences avec lesquelles ce type d'examen est réalisé sont comprises entre 7 et 13

MHz choisis selon la morphologie du sein ainsi que le type de la lésion à analyser. Cette gamme de fréquence élevée 10 à 13 MHz permet l'amélioration de la résolution en contraste chose permettant l'analyse des lésions infra centimétrique notamment dans un sein lépromateux [25], [31]. Cependant, la pénétration des ultrasons est insuffisante (3 à 4 cm) ce qui rend difficile la détection des lésions profondes.

L'examen échographique peut être préconisé en première intention pour les femmes enceintes pour qui les rayons X utilisés pour la mammographie sont contre indiqués, ainsi que dans certains cas où la densité des seins ne permet pas d'obtenir des mammogrammes de qualité (jeunes femmes ou adolescentes, sein de classe BIRADS 3 ou 4), ou dans le cas d'un sein enflammé [2], [3]. Au contraire, cet examen est considéré comme complémentaire ou de seconde intention dans le cas ou on cherche une caractérisation plus précise sur la nature d'une lésion palpée ou repérée par la mammographie (exemple caractérisation de la nature liquide ou solide d'un kyste). Signalant enfin que le radiologue s'aide le plus souvent de l'échographie qui le guide en temps réel pour réaliser un prélèvement au bon endroit [3], [23], [25].

L'échographie est un examen opérateur-dépendant, ce qui peut engendrer une multitude d'analyse et d'interprétation chez une même patiente [25], [31]. Cependant, la reproductibilité d'un tel examen doit être améliorée afin de permettre la relecture ainsi que la comparaison satisfaisante des examens. Par conséquent, la production des documents de qualité au niveau support ainsi que la qualité intrinsèque des images est nécessaire. De l'autre côté, l'examen doit être fait minutieusement, d'abord la glande mammaire doit être étalée harmonieusement sur la paroi thoracique, puis la sonde explore la totalité des quatre quadrants, les sillons sous-mammaires, les creux axillaires, la partie rétroaréolaire, et l'espace intermammaire du haut vers le bas et du bas vers le haut tout en maintenant un bon contact sonde-peau. Cet examen est bilatéral afin de permettre la comparaison, le sein peut être exploré en coupes sagittales ou en coupes radiaires autour du mamelon [25], [32].

3.4. Imagerie par résonance magnétique IRM

La prescription de l'IRM est systématique dans le cas où la patiente possède des antécédents soit d'irradiation thoracique médicale par une dose élevée soit familiale du cancer du sein ou présence d'une mutation des gènes BRCA1 ou BARCA2. Cet examen est d'ordre optionnel lors de l'apparition d'une anomalie radiologique à la mammographie, dans le cas d'un sein de densité mammaire de classe BIRADS 4 ou devant un écoulement unipore séro-sanglant [3].

Le principe de l'IRM est fondé sur les propriétés magnétiques d'un atome, l'hydrogène, contenu, à des degrés variables, dans tous les tissus du corps humain.

Sous l'action d'un champ magnétique puissant \vec{B}_0 (1 à 3T) les protons des atomes d'hydrogène du corps sont stimulés simultanément, ce qui a pour effet la précession des moments magnétiques de tous les atomes d'hydrogène autour de \vec{B}_0 avec une vitesse angulaire $\vec{\omega}_0 = \gamma_H \cdot \vec{B}_0$, γ_H étant la constante gyromagnétique de l'atome d'hydrogène. Par la suite les atomes sont mis en résonance en leur faisant subir une excitation par un champ radiofréquence dont la fréquence est égale à la fréquence de précession des protons. À l'arrêt de la stimulation, le retour à l'état initial est accompagné d'une émission d'une onde électromagnétique caractéristique des tissus. La détection et l'analyse de cette onde par un système informatique permet la reconstruction des images [5], [33]. L'IRM est un examen sans contre-indication connue, il est non irradiant, non invasif, reproductible permettant d'obtenir des coupes fines d'une résolution élevée dans tous les plans de l'espace.

Concernant l'exploration des lésions mammaires, l'examen porte sur l'exploration de tout le volume mammaire sans aucune compression du sein, l'utilisation d'une antenne surfacique spécifique augmente le confort des patientes (figure 1.9) [5].

L'exploration peut porter sur chaque sein séparément ou sur les deux seins simultanément pour comparaison [25]. La sensibilité de cet examen peut atteindre 99% avec une possibilité d'un nombre important de faux positifs (15% à 35%) ce qui rend cet examen non préconisé en première intention [5], [25], [31].

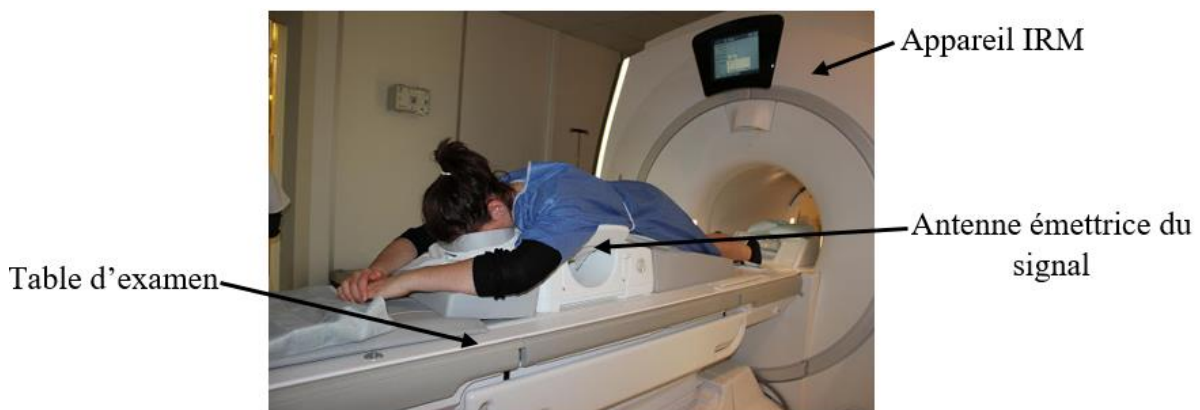


Figure 1. 9. Principe d'imagerie par résonance magnétique.

3.5. Examens de confirmation du cancer du sein

Afin de confirmer la sévérité d'une lésion détectée et classifiée à partir d'un moyen de détection radiologique (mammographie, échographie, IRM...), un prélèvement d'un échantillon de tissu suspect est nécessaire pour une analyse histologique pour que les médecins aient des informations précises sur le type du cancer et ses caractéristiques. L'examen anatomopathologique a pour but de confirmer d'abord le diagnostic, s'assurer de la localisation

ou la propagation des cellules cancéreuses vers d'autres parties du corps, choisir le traitement le mieux adopté [3], [34]. Selon le type de la lésion, le prélèvement est fait sous guidage de la palpation lorsque la lésion est palpable, dans le cas contraire le prélèvement est fait sous guidage de l'échographie ou de la mammographie. Dans le cas des microcalcifications le guidage se fait avec la mammographie pour un prélèvement d'une macro-biopsie [3].

4. Classification des lésions mammaires

Le besoin d'unifier et de standardiser la description des lésions mammaires d'une manière claire et précise, plusieurs classifications ont été adoptées. Ces standards servent de base pour la classification des lésions détectées en fonction de la suspicion de leurs caractères pathologiques. C'est en fonction de cette classification préliminaire que la conduite à tenir est déterminée. L'étude morphologique des lésions et de la composition mammaire était à la base de plusieurs classifications dont on cite : la classification de Wolf, la classification TNM, la classification Le Gal et la classification BIRADS.

4.1. Classification de Wolfe

De point de vue anatomique, un sein est composé principalement de : la glande, les fibres et la graisse avec des proportions variables. La proportion de tissu fibroglandulaire par rapport au tissu graisseux détermine si le sein est graisseux (lorsque cette proportion est faible) ou dense (lorsque cette proportion est élevée). Wolfe en 1976 était le premier à relier le risque d'atteinte d'un cancer du sein et la densité de ce dernier. La classification de Wolfe consiste en quatre catégories selon l'apparence du parenchyme mammaire en mammographie [31], [35], [36].

- La catégorie N1 correspond à un sein de densité normale et à risque minimal ;
- La catégorie P1 correspond à un sein graisseux, présentant des canaux proéminents derrière l'aréole ou dans le quadrant axillaire supérieur n'occupant pas plus de 25 % de la poitrine et à risque minimal ;
- La catégorie P2 correspond à un sein dont le réseau de canaux plus proéminent que P1, un quart ou plus du sein est occupé, le risque est élevé ;
- La catégorie DY correspond à un sein dense avec la possibilité d'une diminution des canaux proéminents, le risque est élevé.

4.2. Classification TNM

L'Union Internationale Contre le Cancer (UICC) a créé un standard de classification du cancer du sein dit TNM (Tumor, Nodes, Metastasis). La classification TNM permet de classer les tumeurs selon leur taille, l'envahissement ganglionnaire et /ou la présence de métastases (voir Tableau en annexe 1) [2], [4].

- La lettre T accompagnée d'une description par un chiffre ou un chiffre et une lettre (0, 1, 2, 3, 4, 4-a, 4-b, 4-c, 4-d) désigne sa taille ainsi que son extension.
- La lettre N indique l'atteinte ou non des ganglions (Nodes). L'ajout des chiffres (0, 1, 2,3) décrit la présence et la nature de l'atteinte.
- La lettre M accompagnée des chiffres (0 ou 1) indique l'absence ou la présence d'une Métastase

4.3. Classification Le Gal

Cette classification est conçue par Le Gal appartenant au service de radiodiagnostic de l'institut Curie à Paris. La classification dite de Le Gal décrit cinq types morphologiques pouvant renseigner les médecins sur la sévérité croissante des microcalcifications qui peuvent touchées le sein [26], [31].

- Type 1 : Microcalcifications de forme annulaires, arciformes ou polyédriques ; le risque de cancer du sein quasi nul.
- Type 2 : Microcalcifications de forme rondes avec des tailles variables, le risque de cancer est d'environ 22%.
- Type 3 : Microcalcifications poussiéreuses ; le risque de malignité est de 36%.
- Type 4 : Microcalcifications à caractère irrégulier dont le risque de cancer atteint 56%.
- Type 5 : Microcalcifications vermiculaires ou branchées, le risque de carcinome étant très élevé, il est de l'ordre de 90%.

Cependant, nombreuses sont les lésions mammaires pouvant être à l'origine de malignité mais ne sont pas pris en compte dans cette étude dont on cite : les masses et l'étude des distorsions architecturales.

4.4. Classification BIRADS (Breast Imaging Reporting and Data System)

C'est le standard utilisé actuellement par les radiologues dans la description des lésions mammaires et la planification de la conduite à tenir par la suite. Ce standard est établi par l'American College of Radiology (ACR) dont l'objectif de combler les lacunes des classifications antérieures [26], [31]. La première version de ce standard a été créée en 1993 et elle a été objet de quatre réformes en 1995, 1998, 2003 et 2013[7]. La classification BIRADS est un système normalisé permettant la catégorisation de la composition du sein ; la caractérisation des pathologies mammaires, leur manifestation en mammographie, échographie et IRM. La cinquième version intègre aussi la caractérisation des lésions en cas d'utilisation de d'autres nouvelles technologies telles que l'élastographie par onde de cisaillement, le degré de suspicion du caractère pathologique de chaque lésion et la conduite à tenir devant chaque

situation [7], [26], [31].

4.4.1. Classification des seins suivant leur composition

La catégorisation d'un sein est liée à sa composition anatomique par la comparaison de la quantité relative de la graisse et du tissu fibroglandulaire. Les pourcentages du tissu fibroglandulaire auxquels est liée la catégorisation du sein sont abrogés dans la nouvelle version vu la difficulté de l'évaluation subjective de ce pourcentage. Les catégories sont au nombre de quatre [7]:

- (A) Le sein est presque entièrement graisseux (figure 10 (A)) ;
- (B) Existence des zones dispersées de densité fibroglandulaire (figure 10 (B)) ;
- (C) Le sein présente une densité hétérogène. les petites masses mammaires peuvent être masquée (figure 10 (C)) ;
- (D) Le sein est extrêmement dense, la sensibilité de la détection par la mammographie diminue (figure 10 (D)).

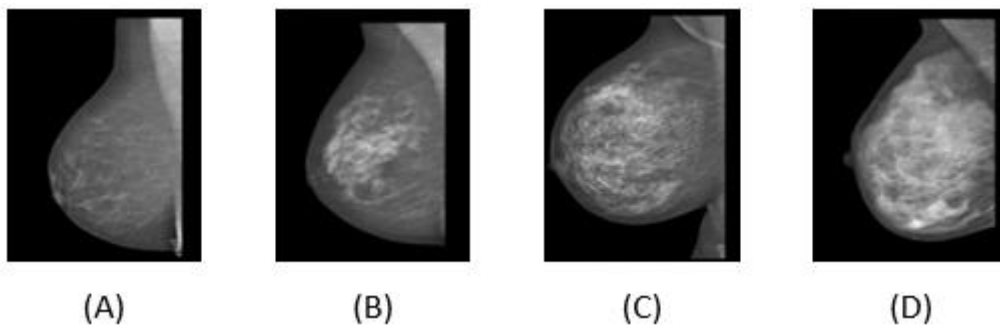


Figure 1. 10. Classification BIRADS du sein selon sa composition (A) entièrement graisseux, (B) sein de densité fibroglandulaire dispersée (C) sein de densité hétérogène (D) sein extrêmement dense.

4.4.2. Caractères pathologiques des lésions du sein

L'analyse d'un mammogramme suivant le standard BIRADS consiste à sa classification dans l'une des catégories fixées par ce standard. Cette classification est faite sur la base d'un ensemble d'informations recueillies à partir des images médicales issues de différentes modalités d'imagerie (mammographie, échographie, IRM...). En mammographie, le standard intègre plusieurs types pathologiques pouvant être source de malignité tels que les microcalcifications, les macrocalcifications, les distorsions architecturales, les masses, la forme des lésions, leur aspect, leur texture, leur densité... [7], [26], [31], [37]. Les masses constituent les lésions les plus difficiles à distinguer vu leurs grandes ressemblances avec le tissu fibroglandulaire normal [9]. Une masse est une lésion importante ayant un caractère tridimensionnel, elle doit apparaître au moins sous deux incidences différentes en mammographie. Selon la cinquième édition du standard BIRADS, une masse est caractérisée

par [7], [26], [31], [37] :

- **La forme** qui peut être :
 - Ronde lorsque la masse est quasi sphérique figure 1.11 (a) ;
 - Ovale dans le cas où la lésion a une allure elliptique figure 1.11 (b) ;
 - Irrégulière lorsque la lésion est d'une forme aléatoire figure 1.11 (b) et (c).

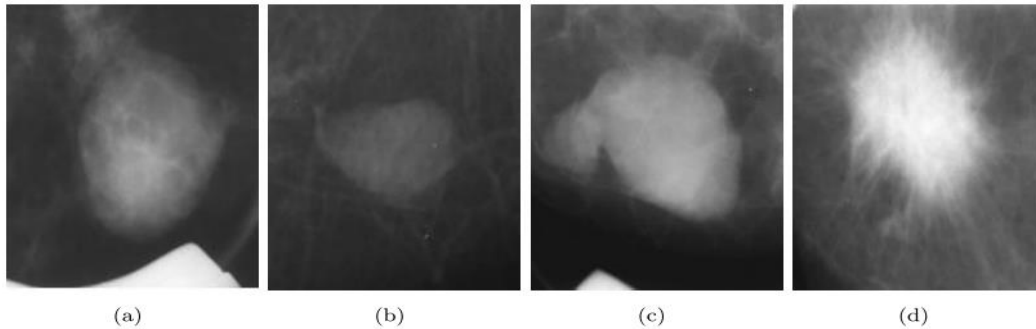


Figure 1. 11. Les différentes formes de masses selon le standard BIRADS 5 : (a)Ronde, (b) Ovale, (c) et (d) Irrégulière

- **Le contour** qui peut être :
 - Circonscrit, il caractérise une masse dont les bords délimitant la lésion et le tissu avoisinant sont bien définis sur au moins 75% de sa surface (figure 1.12 (a)) ;
 - Masqué, dans le cas où le tissu avoisinant cache en partie plus de 25% des bords délimitant la lésion (figure 1.12 (b)) ;
 - Microlobulé, lorsque la lésion est délimitée par des bords présentant de petites ondulations (figure 1.12 (c)) ;
 - Indistinct, c'est le cas contraire de la situation où le contour est circonscrit, le contour de la lésion est mal défini (figure 1.12(d)) ;
 - Spiculé, dans ce cas le contour de la masse mammaire est discrètement irrégulier, dentelé avec des spicules extrêmement courts et fins (figure 1.12 (e))

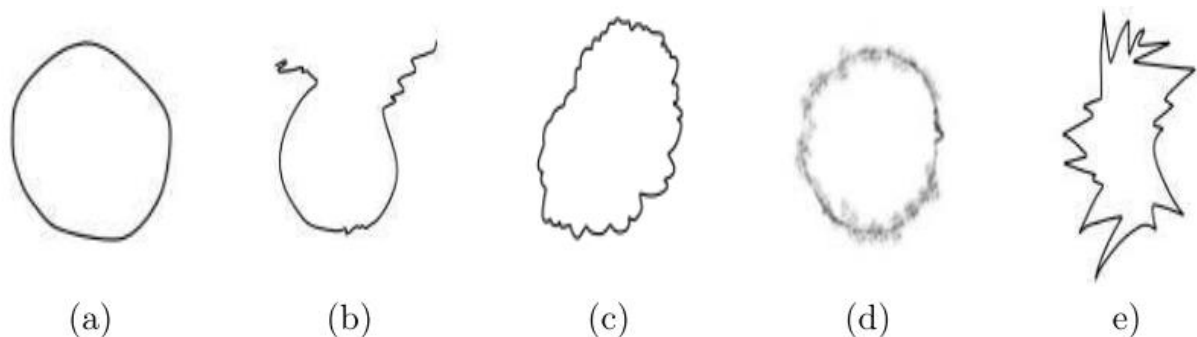


Figure 1. 12. Différents types de contour des masses selon le BI-RADS 5 : (a) contour circonscrit, (b) contour masqué, (c) contour micro-lobulé, (d) contour indistinct et (e) contour spiculé.

- **La densité** qui est une propriété liée à la capacité d'atténuation des Rayon X par la lésion

en comparaison avec le tissu fibroglandulaire normal. Plus sa densité augmente plus sa valeur prédictive positive (VPP) augmente : la densité peut être : adipeuse $VPP=0$; faible, isodense $VPP = 22\%$ ou forte $VPP = 70\%$.

L'analyse des caractéristiques (forme, texture, contour, densité...) fournies par l'image permet aux radiologues d'apprécier la valeur prédictive positive (VPP) et de la classifier par la suite dans la catégorie convenable. En mammographie, les catégories fixées par ce standard [2], [3], [7], [31], [37] sont résumées dans le tableau ci-dessous :

Tableau 1. 1 Classification de l'ACR des lésions mammaires suivant le standard BIRADS

Catégorie	Niveau de suspicion et conduite à tenir
ACR0	L'image mammographique ne permet pas la classification. D'autres investigations complémentaires sont nécessaires (clichés localisés, agrandissement, incidence radiographique particulière, échographie...).
ACR1	Mammographie normale, risque du cancer 0% ; une deuxième lecture est recommandée si cette classification est confirmée, la femme est invitée à refaire l'examen après deux ans.
ACR2	Mammographie montrant des lésions sûrement bénignes risque du cancer 0%. Cette catégorie englobe les opacités rondes, les macrocalcifications, les fibroadénomes ou kystes typiques, les microcalcifications annulaires ou arciformes, semi-lunaires, sédimentées, rhomboédrique. Une seconde lecture est recommandée pour la confirmation de la classification. Aucune surveillance ou examens complémentaires ne sont nécessaires. La femme est invitée à refaire l'examen mammographique après deux ans.
ACR3	Les images montrent des lésions probablement bénignes ($0\% < VPP < 2\%$). Les lésions appartenant à cette classe sont les opacités bien circonscrites, rondes, ovales non calcifiées, non liquidiennes en échographie ; microcalcifications rondes ou punctiformes, régulières ou pulvérulentes, peu nombreuses, en petits amas ronds. La deuxième lecture n'est pas nécessaire ; la patiente doit bénéficier d'un suivi spécifique, une surveillance à court terme est nécessaire.
ACR4	Cette classe est subdivisée en trois catégories dans la nouvelle version BIRADS. Une vérification histologique est obligatoire. La procédure du diagnostic et de prise en charge thérapeutique en fonction des résultats doit commencer. Les images mammographiques appartenant à cette classe présentent des lésions suspectes telles que les opacités non liquidiennes rondes ou ovales aux contours lobulés, ou masqués ; les distorsions architecturales ; les asymétries de densité localisé à limites convexes ; les microcalcifications irrégulières, polymorphes ou granulaires, peu nombreuses... <ul style="list-style-type: none"> ▪ 4A correspond à une anomalie faiblement suspecte ($2\% < VPP \leq 10\%$) ▪ 4B correspond à une anomalie modérément suspecte ($10\% < VPP \leq 50\%$) ▪ 4C correspond à une anomalie modérément suspecte ($50\% < VPP < 95\%$)
ACR5	Les images mammographiques montrent la présence des lésions très suspectes ($VPP \geq 95\%$) telles que les microcalcifications vermiculaires, arborescentes, ou microcalcifications irrégulières, polymorphes ou granulaires, nombreuses et groupées, les opacités mal circonscrites aux contours flous et irréguliers, les opacités spéculées à centre dense. La seconde lecture est inutile ; une analyse histologique est nécessaire pour confirmer le caractère malin de la lésion afin de planifier la prise en charge thérapeutique.
ACR6	Cancer prouvé à la biopsie $VPP=100\%$. Une planification de la thérapie convenable est nécessaire

5. Conclusion

Dans ce chapitre, un ensemble de notions relatives au cancer du sein sont exposées. Après avoir présenté succinctement l'anatomie de la glande mammaire, une étude détaillée des tumeurs, leurs aspects anatomiques ainsi que leurs classifications selon leurs sévérités en bénigne ou maligne sont présentés. Ensuite pour des raisons de prévention, un ensemble de facteurs de risque augmentant la probabilité d'atteindre un cancer du sein est énuméré. Aussi, une énumération d'un ensemble de symptômes pouvant indiquer l'existence d'une anomalie est donnée afin de planifier des examens cliniques de confirmation ou d'infirmer. Dans ce sens un ensemble de techniques d'imagerie médicales permettant l'exploration du sein est présentée dont la mammographie constitue l'examen de référence, vu sa capacité à détecter les petites lésions, son accessibilité et son coût raisonnable. Le chapitre a évoqué aussi les différents standards de classification adoptés par les radiologues et les oncologues afin d'unifier leurs langages, leurs descriptions ainsi que leurs conduites à tenir devant un tel cas. Parmi ces classifications on a détaillé la classification TMN, celle de Le Gal, et le standard BIRADS qui est considéré actuellement le standard de référence en termes de description et classification des lésions mammaires.

Chapitre 2

Description et Optimisation du Système d'aide à la Détection et au Diagnostic Assisté par Ordinateur "DAO" des masses mammaires

Introduction

Ce chapitre est subdivisé en deux parties. La première sera consacrée à la description générale des différentes étapes du système d'aide au diagnostic "DAO" des masses mammaires. Ce système "DAO" est composé à son tour de deux systèmes, le système "DAOe" (qui renferme les étapes de prétraitement et segmentation) servant à la détection de la zone suspecte et à son isolement, et le système "DAOx" permettant son diagnostic (bénigne ou maligne). Notre étude s'est limitée à l'étude de la phase de détection vu son intérêt grandissant. Ainsi la description des différentes techniques employées durant le prétraitement et la segmentation sont évoquées dans ce chapitre. La seconde partie de ce chapitre consiste à présenter le concept de l'optimisation globale sous ses deux aspects : mono-objectif et multiobjectif afin de les intégrer ultérieurement dans les phases de prétraitement et de segmentation dans le but d'augmenter la précision de la détectabilité des zones probablement suspectes. Cela aura un impact direct sur la précision du diagnostic.

Partie 1 : Description du système d'aide à la détection et au diagnostic "DAO" des masses mammaires et caractéristiques de l'image numérique

Le système d'aide au "DAO" (détection/diagnostic) est un ensemble de programmes informatiques qui exploite les techniques de traitement d'images numériques et d'outils d'aide à la décision pour extraire les informations utiles et les analyser automatiquement. Cela permettra aux cliniciens une seconde lecture concernant l'état pathologique d'un mammogramme afin de prendre la bonne décision. Cette dernière est considérée comme étant le centre de l'acte médical. En effet, sur la base de cette décision que tous les actes qui en découlent sont planifiés.

1. Différentes étapes du système "DAO"

Deux catégories de systèmes d'aide à la prise de décision se présentent dans le cas de l'assistance médicale des radiologues à la détection précoce du cancer du sein à savoir les systèmes d'aide à la détection "DAOe" et les systèmes d'aide au diagnostic "DAOx". Les systèmes "DAOe" ont pour but l'extraction et la localisation de la zone probablement suspecte dans une image mammographique dont le radiologue précise son état de normalité, de bénignité ou de malignité. Tandis que les systèmes "DAOx" complètent le processus commencé par les "DAOe" dans le but d'identifier l'état pathologique de la zone suspecte précédemment extraite [17], [18].

La figure 2.1 illustre les différentes étapes de chacun des systèmes "DAOe" et "DAOx" qui fonctionnent complémentaires afin de bien localiser la zone suspecte et de prédire sa sévérité.

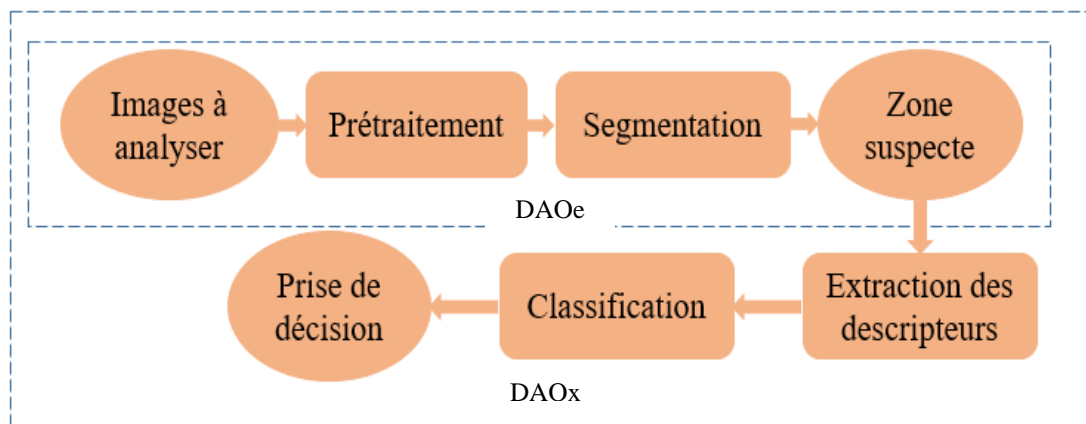


Figure 2. 1. Les différentes étapes du système d'aide à la détection et au diagnostic "DAO"

L'acquisition des mammogrammes en format numérique ou leurs numérisations est à la base de la création des systèmes d'aide au "DAO". En fait, c'est grâce à ce format numérique qu'ils peuvent être manipulés comme un signal numérique à deux dimensions [38], [39]. L'analyse automatique des mammogrammes passe par quatre étapes principales. La première, consiste au prétraitement de l'image pour délimiter la région d'intérêt ("ROI") et améliorer sa qualité. Afin de se débarrasser de toute information inutile et de rendre la "ROI" plus informative et efficacement interprétable. Cette étape est d'une très grande importance dans la précision des résultats des étapes postérieures [11], [40]–[42]. La seconde étape consiste à la segmentation de la "ROI" améliorée en un ensemble de régions afin d'isoler la zone suspecte du reste de l'image. La troisième étape est consacrée à l'extraction des caractéristiques de la zone suspecte sous forme de descripteurs mathématiques. C'est sur la base de ces descripteurs que l'état pathologique de la zone probablement suspecte est déterminé dans la dernière étape [43]–[45][46].

Notre travail portera sur la détection des masses mammaires autrement dit sur les deux étapes prétraitement et segmentation des images mammographiques. Ces deux étapes constituent la base de système "DAO" du fait que leurs précisions aboutira à une meilleure précision du diagnostic dans les étapes qui suivent [9], [11], [40]–[42].

2. Image numérique : caractéristiques et base de données des images utilisées

Une image est une représentation d'une scène réelle, transformée par des capteurs et stockée sur un support (film, papier, support magnétique...) [47], [48]. Dans le domaine médical, une image est une représentation imagée d'un ensemble d'informations anatomiques ou fonctionnelles d'une partie du corps humain. Sa formation est due à l'exploitation d'un

ensemble de phénomènes physiques spécifiques tels que la différence d'atténuation des rayons X, la réflexion des ultrasons, la récupération du signal de résonance magnétique. En mammographie, la formation de l'image est basée sur les différences d'atténuation des RX par les différents tissus d'un sein irradié. Les rayons transmis constituent le signal à numériser pour construire l'image [49]–[52]. En effet, les signaux acquis à partir des dispositifs d'imagerie médicale sont de nature continue, leurs numérisations nécessitent trois opérations essentielles l'échantillonnage, la quantification et le codage. L'échantillonnage d'une image correspond à une discrétisation spatiale par sa décomposition en une matrice de taille n lignes m colonnes, un élément unitaire de l'image est appelé pixel (picture element). Le nombre et la taille des pixels ont une grande influence sur la qualité de l'image. Plus le nombre de pixels augmente plus leurs tailles diminuent. Par conséquent, l'image peut atteindre une résolution spatiale élevée. La quantification quant à elle détermine le nombre de valeurs entières différentes pouvant être attribuées à chaque pixel. Plus l'échelle de quantification est étendue, meilleur sera le résultat en termes de discrimination en contraste. Les valeurs correspondantes sont codées en nombres binaires composés d'une série de 0 et 1. Un pixel peut contenir un scalaire représentant une intensité (niveau de gris, couleur indexée) ou un vecteur représentant une couleur vectorielle en ce point de l'image. En augmentant le nombre de ces valeurs la précision et la fidélité de représentation de données seront alors élevées [49], [52]. Une image numérique est modélisée par une fonction f qui varie dans un plan (x, y) définie comme suit :

$$\begin{cases} f: [1, M] \times [1, N] \rightarrow [0, I_{max}] \\ (x, y) \rightarrow f(x, y) \end{cases} \quad 2.1$$

M, N représentent respectivement les nombre de lignes et de colonnes ; $f(x, y)$ l'intensité de l'image du pixel de coordonnées (x, y) ; I_{max} l'intensité maximale que peut prendre un pixel. Dans ce qui suit, nous présenterons quelques caractéristiques de l'image numérique nécessaires à ses différents traitements numériques. Le type de l'image, sa résolution, son contraste ainsi que son format sont les principaux attributs abordés.

La quantification détermine ainsi le nombre de valeurs entières différentes pouvant être attribuées aux échantillons. On parle aussi de « profondeur », qui s'exprime en bits (binary digit). Plus l'échelle de quantification n'est étendue, meilleur sera le résultat en termes de discrimination en contraste.

2.1. Image Binaire

C'est une image codée sur un seul bit. Seulement deux états sont possibles pour chaque pixel 0 ou 1 qui correspondent respectivement à la couleur noire et blanche [49], [53].

2.2. Image en niveau de gris

Pour ce type d'images, seulement le niveau de l'intensité lumineuse est codé. Si L est le nombre de bits sur lequel l'intensité est codée alors on distingue 2^L niveaux possibles. Par convention, la valeur zéro représente le noir (intensité lumineuse nulle) et la valeur $2^L - 1$ représente le blanc (intensité lumineuse maximale), les autres valeurs correspondent au dégradé du niveau de gris variant entre le noir et le blanc. Certains systèmes d'acquisition permettent d'acquérir des images avec une intensité codée sur 12, 14 ou 16 bits, permettant de représenter jusqu'à 4096, 16384 ou 65536 valeurs différentes respectivement. Ce nombre élevé des niveaux a pour but de pouvoir détecter des variations plus faibles [54], [55]. Le codage de l'intensité sur un octet (8bits) est le plus répandu, le nombre de nuance de gris est alors 256 (0 à 255) niveaux [49].

2.3. Résolution d'une image

La résolution d'une image constitue l'un des critères qualitatifs principaux d'une image. Elle correspond à sa capacité de présenter deux structures spatialement très proches et ayant des intensités physiques de leurs signaux caractéristiques proches. Ceci revient à définir deux types de résolutions : la résolution spatiale et la résolution en contraste. Cette dernière est liée principalement à l'échelle de quantification (le nombre de bits sur lequel un pixel est codé). Quant à la résolution spatiale d'une image, il est nécessaire de définir le concept de la définition de l'image qui se distingue de sa résolution spatiale. Si la définition d'une image est le nombre total de pixels qu'elle contient $NL \times NC$, sa résolution est liée à la surface de son affichage ou de son impression. La résolution d'une image est donc liée à la densité des pixels par unité de longueur dans les deux sens x et y de l'image. Elle est mesurée en points par pouce (ppp ou dpi (dote per Inch)). En minimisant les distances du support physique d'affichage ou d'impression, la densité des pixels augmente ce qui entraîne l'augmentation de la résolution spatiale [49], [56].

2.4. Contraste d'une image

Il s'agit d'une caractéristique intrinsèque à une image donnée. Cette grandeur est beaucoup étudiée et utilisée pour évaluer soit qualitativement soit quantitativement la qualité des images. Cependant aucune définition universelle du contraste n'est adoptée. D'une manière générale, le contraste réfère au rapport entre la luminance locale et la luminance moyenne ou la luminance

du fond. Dans la littérature, divers définitions ainsi que de formules de quantification du contraste ont été proposées dont nous citons [57], [58]:

- Le contraste de Michelson qui a introduit les valeurs de la luminance maximale L_{max} et minimale L_{min} dans l'image afin de quantifier le contraste. Cette définition a connu un grand succès en psychophysique, l'expression du contraste est donnée par l'expression :

$$C_M = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \quad (2.3)$$

- Le contraste de Weber-Fechner qui a introduit une autre notion en considérant que l'image est composée d'un fond de luminance L uniforme et d'un objet donnant naissance à une incrémentation de luminance ΔL . Cette proposition met en évidence la relation entre l'incrémentation ΔL nécessaire afin de rendre visible l'objet et la luminance L du fond. L'expression du contraste de Weber-Fechner est donnée par :

$$C_{W-F} = \frac{\Delta L}{L} \quad (2.4)$$

2.5. Format d'une image

Un format numérique est une norme conventionnelle utilisée pour stocker et transporter les données sous une forme numérique appelée fichier. Ladite norme permet principalement la communication facile de ces données entre divers ressources informatiques telles que des programmes informatiques des logiciels et des machines. Nombreux sont les informations à stocker autres que les données de l'image elles-mêmes : le type de donnée, l'auteur, la date, les conditions d'acquisition... Les formats se distinguent par leurs profondeurs (1bit, 8bits, ...), l'usage d'une compression ou non et leurs fréquences d'échantillonnage [52]. On peut distinguer principalement entre les formats classiques utilisés depuis l'arrivée de la numérisation (dont on cite les formats suivants : pbm, pgm, ppm, tiff, jpeg) et le format DICOM créé spécialement pour la manipulation et la gestion des images médicales.

La norme DICOM (Digital Imaging COmmunication in Medicine) a été développée en 1993 par l'American College of Radiology (ACR) en association avec "National Electrical Manufactures Association (NEMA)". En effet, au cours des années 80 du siècle dernier, le domaine d'imagerie médicale a connu une grande évolution essentiellement au niveau des systèmes d'acquisition de données. Le besoin à l'archivage, la connectivité et l'interopérabilité des équipements médicaux ont imposé aux radiologues et aux fabricants de déployer des efforts en commun pour répondre à cette exigence. Le résultat était le développement de ce nouveau format. Bien que la norme DICOM soit un format d'image, elle est aussi un protocole d'échange de données par réseau ou par fichier. En 1999 ce format est imposé comme le format standard

à utiliser dans le domaine de l'imagerie médicale. Comme les autres formats, la norme DICOM propose une quantification sur 16 bits. Selon les modalités, 12 à 14 bits seront consacrés au codage de la profondeur des images. Le reste de bits peut être utilisé aux autres données d'ordre techniques et administratifs (nom et prénom du patient, type d'examen et ses conditions...) [49].

2.6. Base de données Mini-MIAS

Il s'agit d'une base de données publique publiée par l'organisation des groupes de recherche sur la mammographie au royaume uni. Les images sont obtenues par la numérisation des mammogrammes par un scanner Joyce-Lobel micro-densitomètre qui a une résolution de 50×50 micromètre. Chaque pixel est codé sur 8 bits (256 niveau de gris allons de 0 à 255) avec une résolution spatiale de 1024×1024 pixels. La base de données Mini-MIAS contient 322 images mammographiques en incidence médio-latéral-oblique "MLO" portant le nom mdbxxx (avec $x=001$ à 322) stocker en format "portable graymap .pgm". Les images de cette base de données sont réparties en 207 images normales, 38 images contiennent des masses et 169 contiennent d'autres anomalies. Pour chaque image, les radiologues ont décrit le type de la pathologie, sa localisation en donnant les coordonnées du centre de la pathologie ainsi que le rayon du cercle circonscrit autour de la lésion.

3. Techniques de prétraitement des images mammographiques

Le prétraitement d'une image numérique fait référence à sa transformation par le biais des algorithmes de traitement du signal numérique en une image plus informative et efficacement interprétable [59], [60]. L'image résultante sert comme une meilleure entrée aux prochaines étapes des systèmes permettant son analyse automatique.

D'une manière générale, les images mammographiques souffrent d'un mauvais contraste et des bords flous entre leurs régions ce qui augmente la difficulté de la discrimination entre les lésions et le tissu sain avoisinant. Ceci est dû généralement à la structure du sein qui peut avoir une grande densité de la glande mammaire par rapport au reste du tissu mammaire (BIRADS 3 et 4). Il en résulte alors une faible différence d'atténuation des RX entre les lésions probables et le tissu avoisinant. La figure 2.2 montre trois mammogrammes de classes différentes qui présentent des lésions. Comme le montre la figure 2 (a) la lésion du mammogramme mdb028 (BIRADS 1) est nettement visible, tandis que les lésions des mammogrammes mdb001 (BIRADS 3) et mdb193 (BIRADS 4) ne sont pas facilement décelables par une analyse subjective.

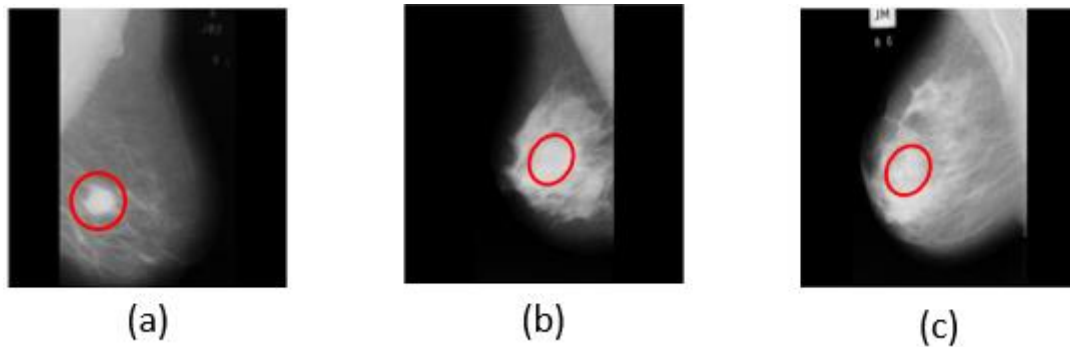


Figure 2. 2. Mammogrammes de différentes classes présentant des lésions (a) mdb028 sein gras avec une tumeur maligne (b) mdb001 sein dense avec une tumeur bénigne (c) mdb193 sein extrêmement dense avec une tumeur bénigne.

L'amélioration de la qualité des images mammographiques consiste alors à modifier leurs apparences visuelles afin de mettre en évidence leurs détails. Cette modification est accompagnée d'un changement d'un ensemble de leurs caractéristiques telles que le contraste, la luminosité, la contenance en bords...etc. En effet l'amélioration de ces attributs n'améliore pas non seulement son aspect visuel, mais aussi l'opération de sa segmentation en un ensemble de régions distinctes [8], [61], [62].

Dans la littérature, divers techniques d'amélioration de la qualité des images sont à distinguer et qui peuvent être classées sur la base de plusieurs critères tels que leurs espaces utilisés (spatial, fréquentiel ou hybride) et leurs principes de fonctionnement (modification de l'histogramme, point à point...). Dans ce qui suit, une description de quelques techniques basées sur la modification de l'histogramme sera présentée. Ces techniques sont très utilisées dans l'amélioration des images y compris les images médicales [63]–[67]. Elles sont aussi utilisées comme techniques de comparaison afin de tester la performance des nouveaux algorithmes [68]. Ensuite il sera présenté un ensemble de techniques de transformation point à point dont la transformation sigmoïde présente l'avantage d'intégrer simultanément les caractéristiques de certaines transformations ayant des effets opposés (transformation exponentielle et logarithmique). En effet, c'est sur la base de ces avantages que sera fondé notre modèle d'amélioration de la qualité des images mammographiques.

3.1. Techniques basées sur la modification de l'histogramme

3.1.1. Egalisation de l'histogramme (HE)

C'est une méthode ayant pour objectif de redistribuer d'une manière équitable les niveaux de gris d'une image à améliorer. Le principe consiste à reconstruire la carte de distribution des niveaux de gris en réaffectant pour chaque pixel une nouvelle valeur afin d'avoir un histogramme aplati et étendu sur toute la dynamique de l'image ($[0, L - 1]$ où $L = 2^n$, n est le nombre de bits sur lequel les pixels de l'image sont codés). Le mappage de chaque niveau de

gris est basé sur la fréquence de son occurrence, qui est représentée par l'histogramme normalisé de l'image à améliorer [66], [69]. Ceci a pour effet de garder certaines caractéristiques visuelles de l'image.

Malgré le rehaussement résultant, la méthode présente plusieurs inconvénients qui se résument d'un côté à une sur-amélioration de l'image résultante, du fait que la technique attribue à l'image améliorée la totalité de la dynamique $[0, L - 1]$, de l'autre côté une perte d'informations locales est observée vu que la technique ne tient en compte que de l'information globale fournie par l'histogramme. [52][70]. L'algorithme de la technique est donné en annexe 2 (a).

3.1.2. Egalisation bi-histogramme pour la préservation de la luminosité (BBHE)

La technique BBHE est une version modifiée de la technique HE. Elle est développée afin de combler certaines de ses inconvénients, spécialement celui de sur-luminance. Son principe consiste à décomposer l'image à améliorer en deux sous images en se basant sur sa luminance. La première sous image est formée des pixels ayant une intensité inférieure ou égale à la valeur moyenne des intensités de tous les pixels de l'image (luminance), la deuxième est formée du reste des pixels dont l'intensité est supérieure à la luminance. L'algorithme procède alors à l'égalisation des histogrammes des deux sous images séparément. La préservation de luminosité consiste à attribuer à la première sous image la dynamique qui s'étend du minimum de l'image origine jusqu'à sa moyenne et attribuer à la seconde la dynamique qui s'étend de la moyenne jusqu'au maximum de l'image origine [66], [70], [71]. L'algorithme de la technique est donné en annexe 2 (b).

3.1.3. Égalisation adaptative d'histogramme à contraste limité (CLAHE)

Bien que la technique BBHE ait amélioré la technique HE de point de vue conservation de la luminosité de l'image origine, les deux techniques HE et BBHE ne prennent en compte que de l'information globale du contraste ce qui engendre une perte d'informations locales et une sur-amélioration dans certaines zones de l'image [8]. La technique CLAHE est introduite alors pour combler ces lacunes. Le principe de la technique consiste d'abord à décomposer l'image en plusieurs sous images non superposables (tuiles) de même taille. La technique fixe ensuite une limite de la hauteur de l'histogramme et l'expansion des niveaux de gris à ne pas dépasser avant de calculer la fonction de distribution cumulative qui sera utilisée pour l'égalisation de l'histogramme de chaque sous image [66], [70]–[73]. La taille des blocs et la limite du contraste fixée ont pour effet d'empêcher la sur-amélioration et l'amplification du bruit dans les zones homogènes. En effet ces deux paramètres permettent de contrôler la qualité de l'image

améliorée. Cette méthode a été développée pour améliorer la qualité des images médicales souffrant d'un faible contraste [72], [74],[75]. Malgré les améliorations introduites par rapport aux versions antérieures, la technique CLAHE produit également une sur-amélioration qui entraîne la perte de certaines informations locales [73]. L'algorithme de la technique est donné en annexe 2 (c).

3.2. Techniques d'amélioration point à point

Ces techniques, largement utilisées pour l'amélioration de la qualité des images dans le domaine spatial, permettent d'exploiter les propriétés de certaines fonctions pour modifier l'aspect visuel ainsi que certains attributs des images. L'utilisation des fonctions dont les valeurs de sortie sont supérieures à leurs correspondantes en entrée permet l'éclaircissement d'une image. Au contraire, l'utilisation des fonctions dont les valeurs d'entrée sont supérieures à leurs correspondantes en sortie entraîne l'assombrissement de l'image. Le principe de ces techniques est de faire correspondre pour chaque valeur d'un niveau de gris d'un pixel de l'image d'entrée une nouvelle valeur dans l'image de sortie. La forme générale de ces transformations est exprimée comme suit [50], [52], [61] :

$$I_s(i, j) = T(I_e(i, j)) \quad (2.5)$$

Avec $I_e(i, j)$ la valeur de l'intensité du pixel de coordonnées (i, j) dans l'image origine ; $I_s(i, j)$ la nouvelle valeur de l'intensité du même pixel dans l'image de sortie ; T la transformation qui fait correspondre pour chaque valeur $I_e(i, j)$ de l'image d'entrée une nouvelle valeur $I_s(i, j)$ dans l'image de sortie.

Ces techniques ont l'avantage d'avoir la possibilité de contrôler la modification des attributs tels que la luminosité et la dynamique de l'image en agissant sur la courbe représentant la transformation [8].

3.2.1. Transformation logarithmique

C'est une transformation non linéaire basée sur la fonction logarithme népérien ou chaque valeur du niveau de gris de l'intervalle de l'image origine I_e est remplacée par sa valeur logarithmique correspondante dans l'intervalle de l'image améliorée I_s par application de l'expression suivante [52] :

$$I_s(i, j) = c. \ln(1 + (e^\sigma - 1). I_e(i, j)) \quad (2.6)$$

$I_e(i, j)$ et $I_s(i, j)$ sont respectivement les intensités du pixel de coordonnées (i, j) avant et après la transformation ;

$c = \frac{L-1}{\ln(1+(e^\sigma-1)I_{max})}$ la constante de mise en échelle des valeurs de niveau de gris de l'image de sortie. Le paramètre σ est introduit pour contrôler l'allure de la courbe de la transformation et qui reflète le degré d'éclaircissement de l'image origine [52].

Comme le montre la figure 2.3, cette transformation permet de faire correspondre à un intervalle étroit des niveaux de gris faibles de l'image d'entrée un large intervalle dans l'image de sortie ; en même temps qu'un large intervalle des niveaux de gris élevés est transformé en un intervalle étroit.

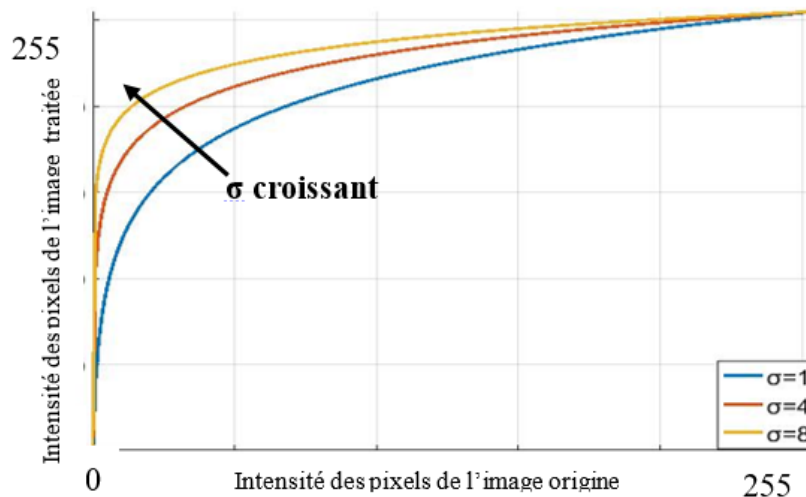


Figure 2. 3. Courbe de la transformation logarithmique

D'une manière générale cette transformation permet d'éclaircir une image (ou des zones sombres d'une image) afin d'augmenter la discrimination des détails qu'elle contient. Ceci est dû au fait que la transformation permet d'augmenter considérablement la dynamique des niveaux de gris de faible intensité. [50], [52].

3.2.2. Transformation exponentielle

C'est une transformation non linéaire basée sur l'exploitation des propriétés de la fonction exponentielle. Elle a des effets inverses de la transformation logarithmique. L'intensité de chaque pixel de l'image améliorée I_s correspond à l'exponentielle de sa valeur I_e dans l'image d'entrée. L'expression de la transformation est comme suit [52]:

$$I_s(i, j) = c. ((1 + \alpha)^{I_e(i, j)} - 1) \quad (2.7)$$

Où c la constante de la mise en échelle de l'intervalle de l'image de sortie ;

$(1 + \alpha)$ représente la base de la transformation exponentielle ;

Cette forme de base est utilisée afin de pouvoir contrôler l'allure de la transformation et par la suite son effet sur l'image.

Comme le montre la figure 2.4, la transformation exponentielle contrairement à la transformation logarithmique, fait correspondre à un large intervalle de niveau de gris de l'image origine un intervalle étroit de faibles valeurs de niveau de gris dans l'image améliorée. En même temps qu'un petit intervalle d'intensité élevé de l'image d'entrée est transformé en un large intervalle dans l'image de sortie. Ceci aura comme résultat l'augmentation de la discrimination des détails contenus dans les zones claires. D'une manière générale la transformation exponentielle permet alors d'assombrir une image ou des zones claires d'une image [50], [52].

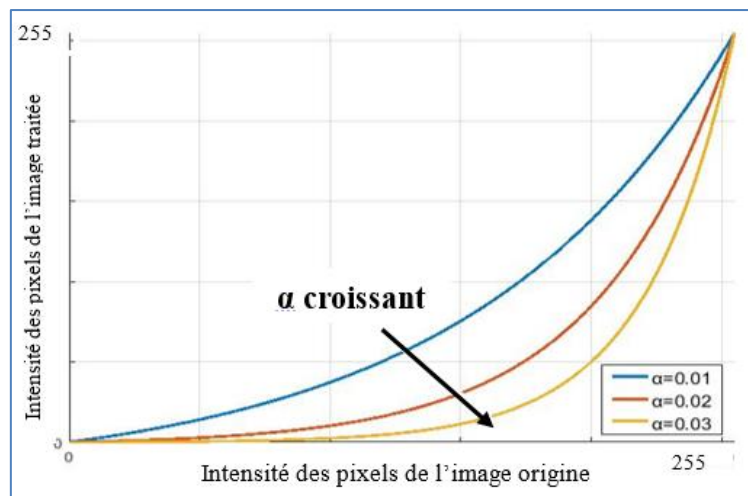


Figure 2. 4. Courbe de la transformation exponentielle

3.2.3. Transformation par loi de puissance

L'amélioration des images par la loi de puissance nommée aussi la correction γ est donnée par l'expression mathématique suivante :

$$I_s(i, j) = c. (I_e(i, j))^\gamma \quad (2.8)$$

Dans le cas où $\gamma > 1$, Cette transformation a un effet similaire à celui de la transformation exponentielle. En effet, la dynamique d'un intervalle étroit des intensités élevées est augmentée, la visibilité des détails des zones claires est améliorée, en même temps qu'un assombrissement est appliqué à l'image origine suite à la compression d'un large intervalle des niveaux de gris de faible intensité (figure 2.5).

Dans le cas où $\gamma < 1$ la transformation a un effet identique à la transformation logarithmique. Une amélioration des détails des zones sombres de l'image est engendrée suite à l'augmentation de l'étendu des niveaux de gris de faible intensité muni d'une compression d'un petit intervalle des niveaux de gris d'intensité (figure 2.5). La valeur de γ contrôle l'arc d'assombrissement ou d'éclaircissement dans les deux cas [50], [52].

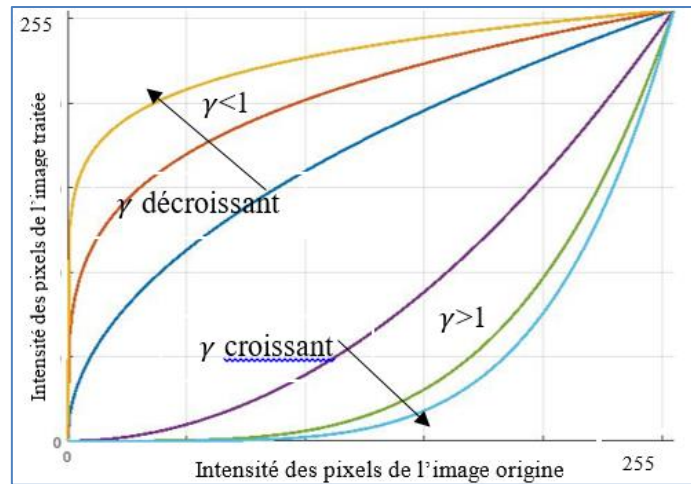


Figure 2. 5. Courbe de la transformation par loi de puissance

3.2.4. Transformation sigmoïde

La transformation sigmoïde est basée sur l'exploitation d'une fonction continue non linéaire définie sur l'intervalle $[0, 1]$ vers $[0, 1]$ ayant une courbe en forme de la lettre «S». Cette transformation fait correspondre pour chaque valeur normalisée des niveaux de gris d'une image d'entrée une nouvelle valeur normalisée de niveau de gris dans l'image de sortie. Une dé-normalisation de la valeur de niveau de gris de l'image de sortie est nécessaire afin de retrouver la valeur finale correspondante à une valeur normalisée de l'image d'entrée. L'expression mathématique de la transformation sigmoïde est comme suit [8], [76], [77]: :

$$s = \frac{1}{1 + e^{-\frac{r-\alpha}{\beta}}} \quad (2.9)$$

s est la valeur de l'intensité normalisée de chaque pixel de l'image améliorée qui correspond à la valeur de l'intensité normalisée r de chaque pixel de l'image origine. α et β sont respectivement le centre et la largeur de la transformation sigmoïde.

L'intensité normalisée des niveaux de gris de l'image d'entrée est donnée par :

$$r = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (2.10)$$

Où I est l'intensité du pixel en cours ; I_{max} et I_{min} sont respectivement les intensités maximale et minimale de l'image origine.

Le calcul des intensités normalisées des niveaux de gris de l'image de sortie est donnée par :

$$I_s = s \cdot (I_{max} - I_{min}) + I_{min} \quad (2.11)$$

Avec I_s l'intensité du pixel résultante de l'image de sortie ; I_{max} et I_{min} sont respectivement l'intensité maximale et minimale de l'image origine.

D'après la figure 2.6, la transformation sigmoïde a l'avantage d'intégrer à la fois les propriétés des transformations permettant l'assombrissement d'une image (transformation exponentielle et loi de puissance avec $\gamma > 1$) pour les niveaux de gris bas, et les propriétés des transformations permettant l'éclaircissement d'une image (transformation logarithmique et loi de puissance avec $\gamma < 1$) pour les niveaux de gris élevés. La transformation entraîne un rapprochement (compression) simultané des niveaux de gris de faibles intensités situés dans la partie 1 et des niveaux de gris d'intensités élevées situés dans la partie 2 (figure 2.6). Ainsi, la transformation permet d'augmenter l'écart (gap) entre les deux parties. Il en résulte alors une différence de densité optique permettant une meilleure apparence de détails de l'image ainsi que de bords délimitant ses différentes régions.

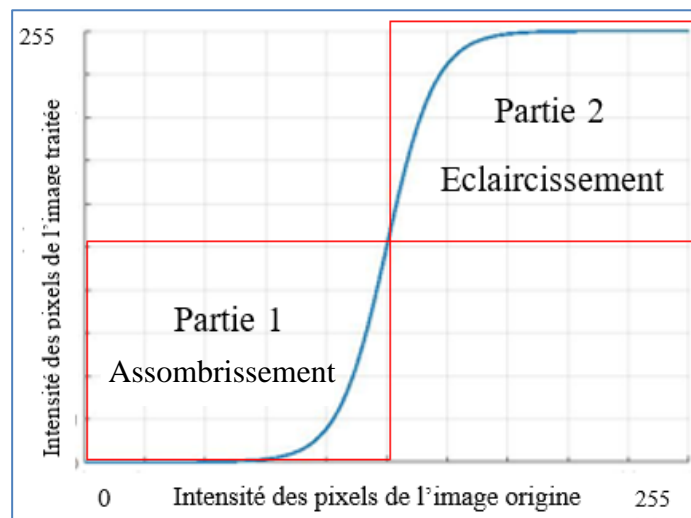


Figure 2. 6. Courbe de la transformation sigmoïde

L'exploitation de cette transformation pour améliorer le contraste des images médicales est d'une très grande importance. En effet la technique permet d'améliorer le contraste des images tout en conservant l'intervalle de niveau de gris des images originales. Cependant, l'utilisation de la transformation globale de la fonction sigmoïde conduit à une image trop améliorée où les caractéristiques sensibles à la luminosité seront perdues à cause de la très grande différence entre les valeurs minimale et maximale de l'image originale [8].

Dans notre travail, nous avons proposé une exploitation locale et adaptative de cette transformation. L'objectif étant de trouver pour chaque image mammographique tous les paramètres convenables de la transformation locale conduisant à une meilleure amélioration.

4. Techniques de segmentation des images mammographiques

La segmentation consiste au partitionnement d'une image en un ensemble de régions significatives, le plus souvent pour distinguer les objets ou les régions d'intérêt du reste de

l'image [51]. Dans le domaine médical, la segmentation permet d'une part, de délimiter et de séparer les différentes structures anatomiques, d'autre part d'éliminer toute information inutile. Le partitionnement d'une image I est basé sur sa décomposition en un ensemble de régions $S = \{R_1, R_2, \dots, R_n\}$ non chevauchantes et homogènes par rapport à un prédicat H déterminé au préalable (intensité, texture, couleur...) tels que [78]:

$$\begin{cases} R_i \neq \emptyset \quad \forall i \in [1 \dots n] \\ R_i \cap R_j = \emptyset \quad \forall i, j; i \neq j \\ I = \cup R_i \quad i \in [1 \dots n] \\ H(R_i) = \text{vraie} \quad \forall i \in [1 \dots n] \\ H(R_i, R_j) = \text{Faux}, \forall i \neq j \end{cases} \quad (2.12)$$

Dans le cas d'une image mammographique, le profil du sein constitue la "ROI" qui peut contenir la lésion figure 2.7. La segmentation de celle-ci se déroule selon deux processus, d'abord il s'agit d'isoler le profil du sein du reste de l'image et par la suite la détermination de la zone probablement suspecte.

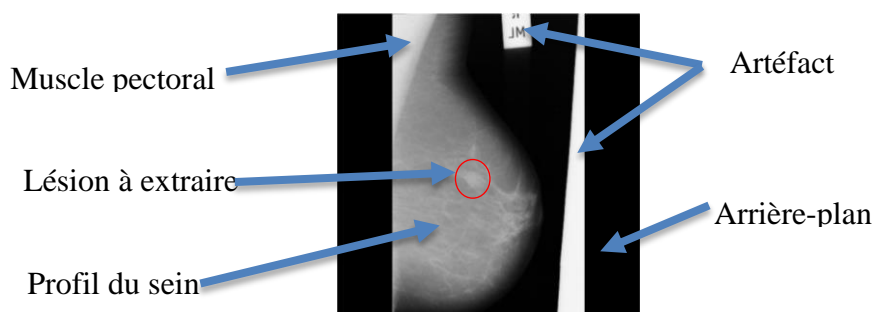


Figure 2. 7. Différentes régions caractérisant le mammogramme mdb012

Dans la littérature, différentes techniques de segmentation se présentent dont aucune ne peut être considérée comme étant standard ou peut fonctionner pour toutes les images, du coup, la segmentation est encore considérée comme étant un problème mal posé [79]–[81]. En faisant référence aux critères mis en œuvre dans le processus de partitionnement des images, les techniques de segmentation peuvent être classées en trois catégories : les techniques basées pixels, les techniques basées contours et les techniques basées régions (figure 2.8).

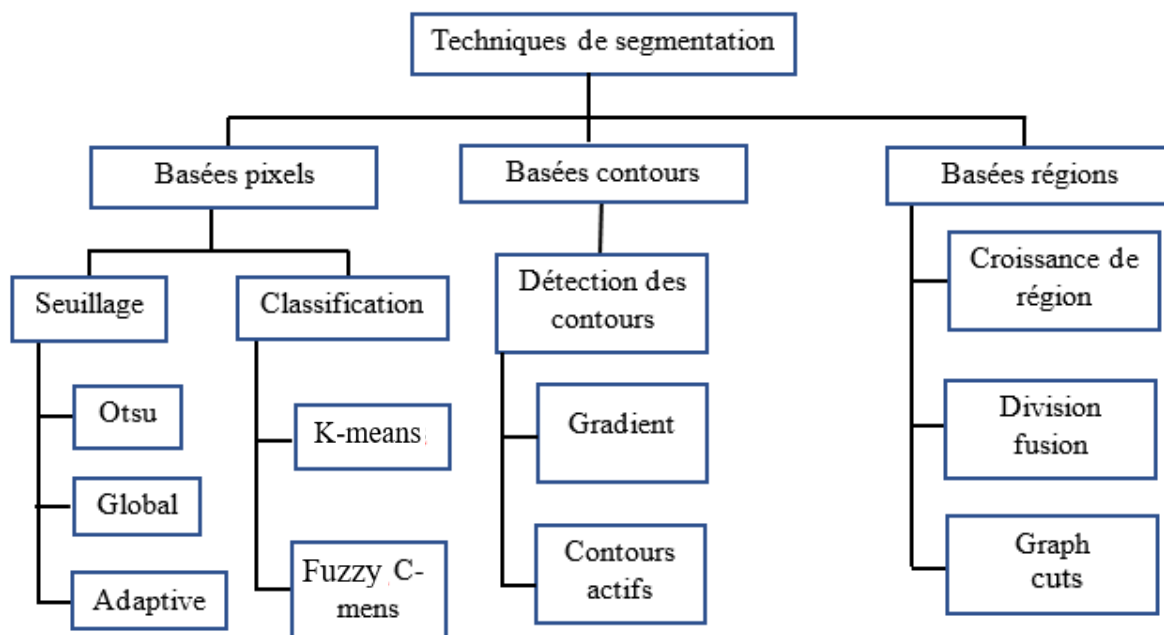


Figure 2. 8. Classification des techniques de segmentation

Dans la littérature, différentes techniques de segmentation d’images ont été largement utilisées pour segmenter les images médicales. Ceci est dû au fait que leurs principes de fonctionnement s’adaptent bien aux propriétés de ces images. Ci-après la description des principales techniques pouvant être utilisées dans la segmentation des images médicales à savoir les techniques de seuillage, les techniques de classification des pixels, la technique de segmentation par la croissance de région et les techniques de segmentation basées sur les contours actifs.

4.1. Techniques basées régions

Ces techniques ont pour objectif de segmenter une image donnée en se basant sur un critère d’homogénéité qui doit être choisi au préalable. Par conséquent la détermination des critères qui précisent les conditions d’homogénéité entre les pixels est la pièce maîtresse de ces techniques. En réalité, une région correspond à un ensemble de pixels connexes et partageant des propriétés similaires. Elle peut correspondre à un objet ou une de ses parties. Parmi les techniques fondées sur le concept de la région on cite les techniques de croissance de la région, les techniques de division fusion des régions et leurs combinaisons.

4.1.1. Méthode de croissance de région

La technique de croissance de région est parmi les techniques interactives permettant la segmentation d'une image donnée. Le démarrage du processus de partitionnement de l'image nécessite la sélection d'un pixel d'amorçage appelé aussi graine (figure 2.9 (b)). A partir de cette graine et en faisant référence aux conditions d'homogénéité définies au préalable, les pixels du voisinage de la graine sont testés et ajoutés à la région dans le cas où ces conditions sont satisfaites (figure 2.9 (c)). L'opération est répétée d'une manière itérative pour tout pixel qui vient d'être attribué à la région. Ainsi, la région croît jusqu'à ce qu'aucun pixel ne puisse être ajouté (figure 2.9 (d)). A nouveau, une nouvelle graine doit être choisie pour commencer la croissance d'une nouvelle région. Ce processus se poursuit jusqu'à ce que toute l'image soit partitionnée [82]–[84] . (Le pseudo- algorithme de la méthode en annexe 3(a))

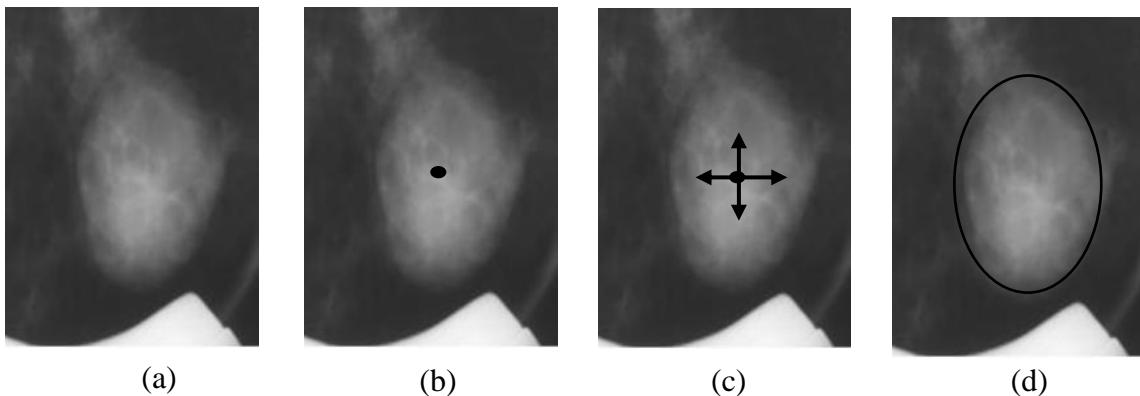


Figure 2. 9. Principe de segmentation par la méthode de la croissance de la région (a) image origine, (b) initialisation, (c) croissance (d) région isolée.

4.1.2. Méthode de décomposition fusion

La technique division /fusion a été proposée par Pavlidis et Horowitz en 1974 [85]. C'est une approche permettant la détection directe et automatique des régions [86]. L'algorithme fonctionne d'une manière hybride et se déroule en deux étapes : la première consiste à tester la condition d'homogénéité sur l'image entière puis la diviser en quatre quadrants en utilisant la structure de l'arbre quaternaire lorsque la condition d'homogénéité n'est pas vérifiée (figure 2.10 (b)). Le processus d'évaluation division se poursuit d'une manière récursive sur tous les quadrants résultants de la même manière que l'image entière (figure 2.10 (c)) [86], [87]. En effet, cette division hiérarchisée peut engendrer plusieurs régions adjacentes qui peuvent partager le même prédicat de l'homogénéité. D'où la nécessité de leur fusion afin qu'elles soient intégrées dans une même région (figure 2.10 (d)). L'algorithme évalue alors le critère d'homogénéité pour les régions adjacentes et fusionne en une seule celles ayant vérifié le

critère d'homogénéité [82], [85], [88], [89]. L'algorithme de la technique est donné en annexe 3 (b).

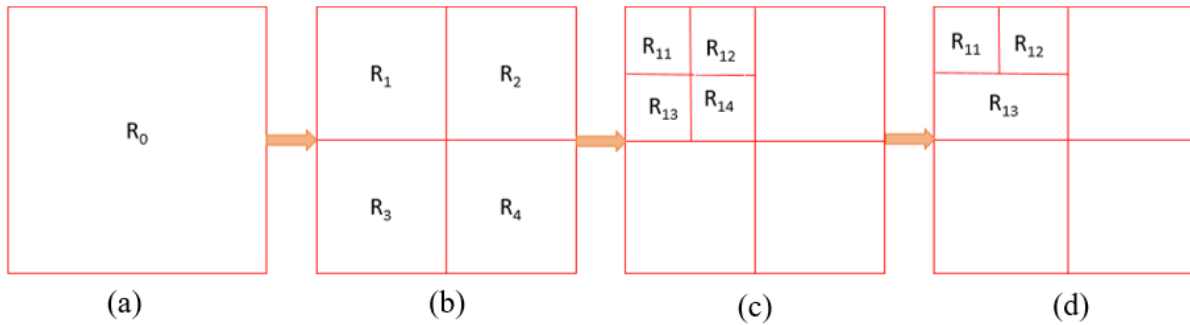


Figure 2. 10. Principe de segmentation par la méthode de la croissance de la région (a) image origine, (b) et (c) division des régions (d) fusion de deux régions.

4.2. Méthodes basées pixels

4.2.1. Seuillage

Le seuillage est une méthode simple et rapide permettant la segmentation des images. Parmi ses principaux avantages sa simplicité d'implémentation et son efficacité dans les systèmes en temps réel. Son principe consiste à partitionner une image en un ensemble de régions en se basant seulement sur les intensités des pixels sans faire recours à aucune autre information telle que l'information spatiale de l'image. Ceci rend la méthode sensible au bruit et à des inhomogénéités d'intensité de l'image. Le seuillage convient bien lorsque les objets à isoler présentent une clarté par rapport au fond de l'image. Le challenge de ces méthodes réside au niveau de la localisation automatique des seuils à utiliser pour séparer les objets les uns des autres [88]–[90]. Dans la littérature on peut différencier entre trois types de seuillage : le seuillage global, le seuillage multiple et le seuillage adaptatif [78], [83], [89].

A. Seuillage global

Cette méthode a pour but de créer une image binaire à partir d'une image donnée. Elle convient bien pour une image dont l'histogramme est bimodal. Le seuil " th " peut être déterminé interactivement à partir de l'histogramme, sa valeur étant celle qui sépare les deux modes ou automatiquement à partir de l'histogramme normalisé de l'image, en calculant le seuil permettant de maximiser la variance interclasse ou bien qui minimise la variance intra-classe (méthode d'Otsu) [91]. Les pixels dont l'intensité inférieure au seuil seront attribués au fond de l'image avec une intensité égale à zéro (noire). Quant au reste de pixels dont l'intensité supérieure au seuil constitue l'objet recherché, l'intensité des pixels vaut 1 (blanc) [82], [91]. L'expression $I_b(x, y)$ de l'image binaire résultante a pour expression :

$$I_b(x, y) = \begin{cases} 1 & \text{Si } I(x, y) \geq th \\ 0 & \text{Sinon} \end{cases} \quad (2.16)$$

Avec $I(x, y)$ l'intensité du pixel ayant les coordonnées cartésiennes (x, y) ; th étant le seuil.

B. Seuillage multiple

Dans le cas où l'image contient plusieurs objets, le seuillage global ne peut isoler les différents objets de l'image. L'histogramme correspondant à ce type d'image est multimodal. D'où la nécessité de la partitionner en diverses classes C_i afin d'isoler efficacement son contenu. Si on considère que l'image doit être partitionnée en k classes, il est nécessaire de déterminer $k - 1$ seuils. Le partitionnement de l'image sera comme suit :

$$\left\{ \begin{array}{l} C_1 \leftarrow I(x, y) < th_1 \\ C_i \leftarrow \text{si } th_{i-1} \leq I(x, y) \leq th_i \\ \dots \\ C_k \leftarrow \text{si } th_{k-1} \leq I(x, y) \end{array} \right. \quad (2.17)$$

La détermination des seuils peut se faire d'une manière interactive à partir de l'histogramme sachant que chacune de ses vallées correspond à un seuil [90]. Cependant, leur détermination d'une manière automatique a beaucoup séduit les chercheurs. En effet, divers méthodes utilisant d'autres critères autres que l'information du niveau de gris ont été appliquées afin d'aboutir à une séparation efficace entre les objets et l'arrière-plan. Parmi ces critères, on cite l'entropie de Kapur, la variance entre classes et l'entropie de Tsalli. Ces critères sont utilisés comme fonctions objectifs à optimiser afin de garantir un meilleur partitionnement de l'image [92].

C. Seuillage adaptatif

Il s'agit d'utiliser un seuil qui s'ajuste avec chaque partie de l'image. Il est utilisé dans le cas où le fond de l'image n'est pas uniforme ou l'objet à isoler présente une hétérogénéité du niveau de gris. Le seuillage global ne peut alors fonctionner sur toute l'image. L'une des solutions proposées à ce problème est de subdiviser l'image en un ensemble de sous images pour lesquelles on peut déterminer facilement le seuil permettant d'isoler la partie correspondante à l'objet de son fond. L'ensemble des résultats seront alors fusionnés pour retrouver l'objet entier recherché. Pour la détermination des seuils locaux, les mêmes techniques que précédemment peuvent être appliquées en traitant chaque sous image comme une image distincte [82].

4.2.2. Méthodes basées regroupement des pixels (classification)

Le fonctionnement des algorithmes de regroupement est similaire à celui des techniques de classification, la différence réside du fait que les algorithmes de regroupement ne nécessitent pas de données d'entraînement. Le regroupement est une technique qui consiste à regrouper un ensemble de données en un ensemble de classes appelées aussi clusters. Dans le domaine de segmentation d'images, la technique consiste à partitionner l'ensemble des pixels d'une image en un ensemble de classes. Un bon partitionnement est celui caractérisé par une forte similarité

intra-classe (pixels d'une même classe) d'une part, de l'autre part, une forte dissimilarité interclasse (pixels des classes différents) [83], [89], [93]. La similarité est une propriété déterminée en termes d'une distance telle que la distance euclidienne. Une classe est caractérisée par son centroïde qui représente la moyenne des niveaux de gris de tous ses pixels et sa variance. Deux grandes catégories sont à la base du principe de regroupement :

- Les méthodes hiérarchiques basées sur le concept d'arbre qui est un partitionnement récursif dont, la racine de l'arbre représente l'ensemble des pixels de l'image et les nœuds internes représentent les classes ;
- Les méthodes basées sur le partitionnement, leur principe consiste à affecter chaque pixel à une classe qui lui convient en se basant sur son intensité. Le partitionnement se fait d'une manière itérative afin d'augmenter à chaque passage la similarité intra-classe et la dissimilarité interclasse, ceci revient à optimiser une fonction objectif qui représente cette condition.

La classification peut être dure ou douce. Une classification est dite dure lorsqu'elle ne permet à un pixel d'appartenir qu'à une seule classe. A l'encontre, dans le cas d'une classification douce, la division n'est plus stricte. Les pixels sont divisés en classes sur la base d'une appartenance partielle, c'est-à-dire qu'un pixel peut appartenir à toutes les classes sur la base d'un degré d'appartenance calculé à partir de sa distance par rapport au centroïde de chaque classe [89], [93].

A. Méthode de classification K-moyennes (K-means)

Cet algorithme a été développé par Mac Queen en 1967 [94]. Il s'agit de l'un des algorithmes les plus connues en classification non supervisée à frontières dures. Il permet la classification de données en k classes même dans le cas où aucune connaissance a priori n'est connue. Le principe de partitionnement consiste à affecter chaque élément de données à la classe qui lui convient. Le processus de partitionnement se fait d'une manière itérative en assignant chaque donnée au cluster qui lui est proche en se basant sur le critère de similarité caractérisant les données. Suite à la migration de quelques données vers les classes dont leurs centroïdes sont les plus proches, les centroïdes qui occupent les barycentres des clusters doivent être recalculés. Ce processus est répété jusqu'à l'invariabilité des centroïdes. Le meilleur partitionnement est celui qui minimise la fonction objectif suivante :

$$SSE = \sum_{j=1}^K \sum_{p_i \in Cl_j} (p_i(x, y) - C_j)^2 \quad (2.18)$$

Où SSE est la somme des erreurs quadratiques, $p_i(x, y)$ est le pixel de coordonnées (x, y) et appartient à la classe Cl_j de centre C_j .

Les étapes de l'algorithme k-moyennes sont [82], [94], [95] :

- 1- Détermination du nombre de classes K auxquelles les données seront partitionnées
- 2- Générer aléatoirement l'ensemble de K centroïdes. Chacun pour un cluster
- 3- Calculer pour chaque pixel les distances Euclidiennes qui les séparent des différents centroïdes par l'expression suivante :

$$d_i = |p(x, y) - C_i| \quad (2.19)$$

- 4- Affecter chaque pixel au cluster dont le centroïde est le plus proche au pixel en question.
- 5- Calculer la nouvelle position de chaque centroïde en utilisant la formule suivante :

$$C_i^n = \frac{1}{N_i} \sum_{j=1}^{N_i} p_j(x, y) \quad (2.20)$$

C_i^n le nouveau centroïde calculé, N_i le nombre de pixels contenus dans le cluster K_i

- 6- Répéter l'étape 3 et 4 jusqu'à l'invariabilité des données.

(Le détail et le pseudo-algorithme de la méthode en annexe 5(b))

B. Méthode de classification C-moyennes Flou (Fuzzy C-mens)

La méthode Fuzzy C-means "FC+M" est une méthode non supervisée qui a pour but de partitionner les données en C classes homogènes. Dans le cas de la segmentation des images, l'algorithme attribue à chaque pixel de l'image un degré d'appartenance correspondant à chaque classe. Ce dernier est calculé à partir de la distance (en niveau de gris) qui le sépare du centroïde de la classe en question. Plus un pixel est plus proche à un centre d'une classe donnée, plus la probabilité de son affectation à cette classe augmente. Cette souplesse de possibilité d'appartenir à plusieurs classes reflète la réalité physique de l'échantillon étudié face aux divers problèmes fréquentés au niveau de l'acquisition des images médicales (bruit, volume partiel, hétérogénéité des champs...) [96], [97].

4.3. Méthodes basées sur le contour

Un contour correspond à une discontinuité d'intensité dans l'image dans une direction donnée. Généralement, un objet correspond à une région délimitée par des contours fermés. La détection de cet objet consiste à détecter ses contours. En effet ce concept constitue le principe que notre système visuel utilise pour comprendre un contenu sémantique d'une image. Les techniques de segmentation basées contour sont des approches ayant pour principe la détection des discontinuités contenues dans l'image qui correspond à un changement plus ou moins brusque du niveau de gris. Contrairement aux techniques basées régions et pixels, les techniques basées contours présentent un avantage du fait que les régions résultantes de ces techniques peuvent renfermer des pixels présentant des variations significatives [90], [98]. Deux grandes

catégories de détection de bords à distinguer : les méthodes de détection du gradient et les modèles déformables.

A. Méthode basée gradient

Cette méthode fait appel à des noyaux ou masques de convolution afin de détecter les bords contenus dans une image. Un masque de convolution est une petite matrice permettant d'approximer la dérivée vectorielle d'une image en un pixel donnée. L'objectif étant de localiser l'ensemble des pixels présentant des discontinuités des intensités. En effet ces derniers correspondent aux pixels appartenant aux bords. Dans la littérature, divers détecteurs de bords ont été développés. La figure 2.11 montre trois exemples de masques Prewitt, Sobel et Scharr dont chacun est composé d'un détecteur dans le sens horizontal P_x , S_x , Sc_x et un autre dans le sens vertical P_y , S_y , Sc_y [8], [99]–[101].

Prewitt	Sobel	Scharr
$P_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$Sc_x = \begin{bmatrix} 3 & 0 & -3 \\ 10 & 0 & -10 \\ 3 & 0 & -3 \end{bmatrix}$
$P_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$	$S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$	$Sc_y = \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix}$

Figure 2. 11. Masques de convolution pour la détection des contours (Prewitt, Sobel, Scharr)

La figure 2.12 montre le résultat de l'application des masques de convolution Prewitt et Sobel à une région d'intérêt extraite de l'image mammographique mdb184. Le résultat montre que les contours détectés sont mal localisés et non fermés. Pour terminer l'opération de la segmentation, il faut utiliser des techniques de reconstruction de contours par interpolation. Les modèles déformables constituent alors un moyen permettant de combler ce problème par la détection des contours fermés.

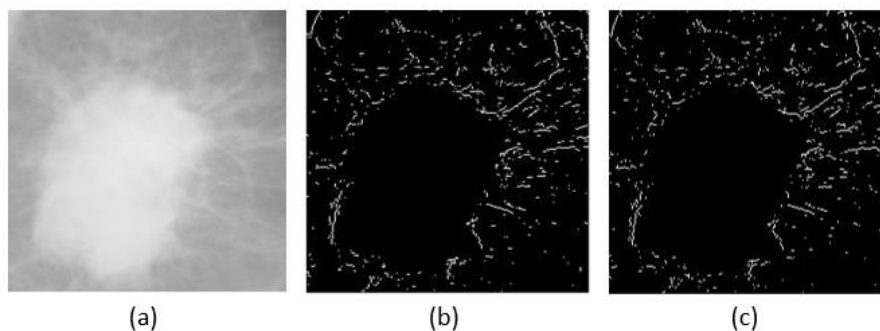


Figure 2. 12. Détection de bords par application des masques de convolution ; (a) "ROI" mdb184, (b) masque Prewitt, (c) masque Sobel

B. Modèles déformables

Le modèle déformable est un outil doté de la capacité de se déformer et se déplacer automatiquement au sein d'une image jusqu'à atteindre les frontières de l'objet à extraire. Bien que les modèles déformables soient fondés sur les modèles élastiques Burr 1981, la communauté s'est accordée de les attribuer à l'équipe de Kass, Witkin et Terzopoulos en 1988 [102]. Ils englobent les contours actifs dans le cas des images 2D et les surfaces actives dans le cas des images 3D. L'évolution de ces contours est gérée par l'interaction entre les propriétés physiques de rigidité et d'élasticité qui leurs sont attribuées d'une part et les gradients engendrés par la variation de l'intensité des pixels de l'image de l'autre part. Le système qui est caractérisé par une équation de l'énergie évolue alors d'une manière itérative afin de la minimiser en faisant déplacer le contour actif vers le bord le plus proche en prenant sa forme [88], [103]. Ces modèles ont été l'objet de plusieurs améliorations afin de surmonter les problèmes rencontrés tels que la nécessité de leur initialisation tout près des bords, la difficulté d'accès aux petites cavités, la détection des objets proches les uns des autres. Deux grandes catégories à distinguer : les modèles déformables paramétriques et les modèles déformables géométriques.

- **Les modèles déformables paramétriques**

Ces modèles utilisent des courbes paramétriques pour représenter la forme du modèle. Ces courbes sont définies dans le domaine spatial de l'image. Au cours de leurs évolutions, les déformations sont déterminées par la géométrie, la cinématique, la dynamique, ainsi que d'autres contraintes dans le cas où elles existent. Sur le plan mathématique, les contours actifs sont des "splines" dont leurs états, positions et dynamiques sont déterminés par une énergie. L'évolution des courbes actives est gérée par analogie avec la physique d'un problème de minimisation de l'énergie. Une courbe paramétrique E_{courbe} est définie comme la somme de trois termes d'énergie :

$$E_{courbe} = E_{interne} + E_{externe} + E_{contrainte} \quad (2.14)$$

Avec $E_{interne}$, $E_{externe}$ sont respectivement l'énergie interne et externe résultantes des forces internes et externes, $E_{contrainte}$ est l'énergie due aux contraintes imposées à la courbe.

Ce modèle présente certains inconvénients dont on cite :

- la courbe doit être paramétrée à nouveau après un certain nombre d'itérations suite au changement topologique subi au cours de son évolution ;
- la position initiale du modèle doit être placée suffisamment près de la limite souhaitée pour éviter la convergence vers de mauvaises limites [83], [104].

- **Les modèles déformables géométriques**

Ces modèles constituent une alternative pour y pallier à certains inconvénients des modèles paramétriques. Leur principe s'appuie sur l'utilisation d'une représentation des formes par niveaux (Osher et Sethian, 1988) afin de prévenir contre les problèmes issus des changements topologiques au cours de l'évolution des courbes actives. Dans ce modèle le contour actif est intégré dans une fonction de niveau dimensionnel supérieur tout en considérant que le contour étant le niveau zéro. Ceci a pour objectif de suivre le niveau zéro de la fonction au lieu de suivre les points discrets du contour afin de traiter facilement ses changements topologiques.

Dans le but d'amener le contour actif aux frontières de l'objet à isoler, les fonctions de vitesse doivent être définies correctement. Malladi et al (1993), Malladi et Sethian (1996) et Caselles et al ont utilisé les informations du gradient pour définir la fonction de la vitesse en prenant en compte l'influence de la courbure pour garder le contour lisse. Au voisinage des limites de l'objet à isoler l'augmentation de la magnitude du gradient entraîne une diminution de la vitesse de la courbure qui s'annule lorsqu'elle se superpose au contour de l'objet. Le choix du critère d'arrêt est décisif pour que la courbure s'arrête au bon endroit. Le bruit et le flou des contours dans certains cas entraînent la difficulté de l'arrêt. Le contour actif continue sa convergence jusqu'à sa disparition après un certain nombre d'itérations. Afin d'améliorer le modèle, Kichenassamy et al. (1996) et Siddiqi et al. (1998) ont introduit respectivement l'élément de la force de bord, et de surface. La modélisation de segmentation de cette version des algorithmes dite aussi géodésique est ramenée à un problème d'optimisation pour trouver la courbe de distance minimale dans l'image qui correspond aux limites de la structure recherchée. L'équation décrivant le déplacement de la courbe est dérivée d'une fonction énergétique [103]. Parmi les modèles géométriques les plus populaires et améliorés celui proposé par Chan et Vese (CV) (1999) et Chan et Vese (2002). Le modèle de Chan-Vese est une version simplifiée du modèle énergétique. L'avantage le plus appréciable de ce modèle est qu'il peut obtenir une limite de points discrets, ce qui est très important dans le cas où les objets d'intérêt sont représentés par des groupes de pixels discrets et n'ont pas de limites clairement définies, ceci est dû au fait que le modèle n'utilise pas l'information du gradient de l'image. Le principe de l'algorithme pour l'extraction de l'objet désiré est de minimiser simultanément les variations d'intensité à l'intérieur et à l'extérieur du contour [105]. Le principe du modèle Chan-Vese est de minimiser la fonctionnelle d'énergie suivante :

$$F(C_1, C_2, C) = v \cdot \text{longueur}(C) + \eta \cdot \text{Aire}(C) + \lambda_1 \int_{\text{intérieur } C} |I_0(x, y) - C_1|^2 dx dy + \lambda_2 \int_{\text{extérieur } C} |I_0(x, y) - C_2| dx dy \quad (2.15)$$

$I_0(x, y)$ est l'image à segmenter ; $v, \eta, \lambda_1, \lambda_2$ sont respectivement les poids de chaque terme ; C_1, C_2 sont respectivement les moyennes des régions interne et externe du contour C .

Concernant notre conception du système d'aide au "DAOe" des masses mammaires, deux objectifs essentiels à réaliser. Le premier consiste en l'automatisation de tous les processus de ce système, dans ce sens, au niveau de la phase de segmentation, nous avons choisi d'utiliser des techniques ne nécessitant aucune initialisation (k-means) et des techniques pouvant être initialisées automatiquement (seuillage global et la technique de la croissance de la région). Le second consiste à optimiser les résultats de chaque étape afin d'aboutir à un bon diagnostic. Dans cette optique nous allons présenter le concept de l'optimisation ainsi que les procédures de son intégration dans les différentes étapes de ces systèmes.

Partie 2 : Optimisation des systèmes d'aide au diagnostic "DAO" des masses mammaires

Dans cette partie, un ensemble de concepts d'ordre général de l'optimisation globale sera présenté. Sur le plan mathématique, l'optimisation est définie comme étant un outil qui consiste à minimiser ou maximiser une fonction. Selon la nature de cette fonction, l'optimisation peut être mono-objectif ou multiobjectif. Un ensemble de notions qui sont liées à ces deux concepts seront présentés. On termine par la description de l'algorithme génétique qui est le solveur choisi dans notre étude. L'ensemble de ces outils seront exploités dans les chapitres suivants afin d'optimiser les résultats de l'amélioration de la qualité et de la segmentation des images mammographiques à analyser.

1. Concepts fondamentaux

L'optimisation est une démarche systématique qui se compose d'un ensemble d'actes complémentaires et nécessaire à la résolution d'un problème donné. La première consiste à la définition précise du problème, essentiellement les objectifs souhaités, les paramètres influençant sa performance ainsi que les contraintes qui les relient. La seconde a pour objectif de caractériser le problème par une équation mathématique (mono-objectif) ou plusieurs équations (multiobjectif). La dernière étape consiste à la résolution du problème en utilisant un algorithme d'optimisation [106], [107]. Le développement informatique et des techniques d'intelligences artificielles qui ont caractérisé les dernières décennies a conduit à un énorme développement des procédés d'optimisation mathématiques qui peuvent aujourd'hui prendre en compte des dizaines de paramètres de conception. La complexité de ces problèmes et leurs grandes tailles rendent leurs résolutions difficiles en utilisant les méthodes déterministes conventionnelles. Les méthodes stochastiques basées sur des processus aléatoires dans l'exploration de l'espace de recherche prennent alors le pas sur les techniques déterministes dans la résolution de ce type de problème. Le principe consiste à rechercher et à comparer des

solutions possibles et réalisables jusqu'à l'obtention d'une meilleure solution [106], [108]–[110].

Le traitement d'un problème en tant qu'un problème d'optimisation nécessite les éléments essentiels suivants :

- Les paramètres influençant le problème dites variables de décision ou de conception. Ils constituent l'espace de décision et englobe l'ensemble des solutions possibles.
- La fonction "objectif" qui représente le cœur de l'opération d'optimisation. C'est sur sa base que le problème est formulé en un problème d'optimisation. D'une manière générale, la définition de cette fonction est primordiale car la pertinence des solutions apportée au problème en dépend fortement.
- Un algorithme d'optimisation qui a pour but de chercher dans l'espace de décision la ou les solutions pertinentes au problème.

2. Optimisation mono-objectif

L'optimisation mono-objectif est un acte qui a pour but la recherche d'une solution optimale à une fonction scalaire. Cette dernière est généralement une formulation mathématique d'un problème faisant intervenir un ensemble de paramètres qui doivent être choisis minutieusement dans un espace relativement grand afin d'atteindre le maximum de sa performance (objectif souhaité). Un problème d'optimisation mono-objectif de dimension n peut être formulé par l'expression suivante :

$$\begin{cases} \text{Minimiser/Maximiser } f(X) \\ X = \{x_1, \dots, x_n\} \\ h_i(X) = 0, & i = 1 \dots m \\ g_j(X) \leq 0, & j = 1 \dots k \\ X_{min} \leq X \leq X_{max} & i = 1, \dots, n \end{cases} \quad (2.21)$$

La fonction $f(X)$ à valeurs dans \mathbb{R} est la fonction mathématique caractéristique du problème mono-objectif à maximiser ou à minimiser selon le cas [111].

$X = \{x_1, \dots, x_n\} \in \mathbb{R}^n$ est un vecteur à n composantes représentant les variables de décision du problème.

$h_i(X)$ et $g_j(X)$ sont respectivement des contraintes d'égalité et d'inégalité que doit vérifier le vecteur $X \in [X_{min} X_{max}]$ afin qu'il puisse être solution du problème. Ces contraintes peuvent exister, le problème est dit alors «problème d'optimisation mono-objectif avec contraintes» ; dans le cas où ces contraintes n'existent pas le problème est dit «problème d'optimisation mono-objectif sans contraintes». L'ensemble des régions de l'espace de recherche où les contraintes de conception sont vérifiées est nommé espace réalisable ou domaine admissible [112]–[114].

La résolution de ce type de problème en utilisant un algorithme d'optimisation est facile vu que la recherche aboutit à une seule solution qui est l'optimum global de la fonction $f(X)$. La difficulté réside au niveau de la modélisation du problème par une seule équation [106], [107]. En effet cette tâche n'est pas toujours facile du fait que la modélisation d'un problème n'est pas unique. D'une manière générale, la formulation d'un problème par une seule équation peut entraîner la perte de certaines de ses caractéristiques de ses performances fixées à l'avance. En effet, la plupart des problèmes du monde réel nécessite l'optimisation de plusieurs fonctions (objectifs) afin de répondre à toutes les exigences imposées [108]. Ce type de problèmes s'intègre dans la catégorie des problèmes d'optimisation multiobjectif.

3. Optimisation multiobjectif

L'optimisation multiobjectif dite aussi optimisation multicritères est une alternative permettant de combler les lacunes de la modélisation d'un problème par l'optimisation mono-objectif. Elle a pour but d'optimiser (minimiser ou maximiser) simultanément plusieurs fonctions, en utilisant les mêmes variables de décision. L'ensemble des fonctions à optimiser constitue un vecteur à optimiser d'où l'appellation problème d'optimisations vectorielles [115]. L'optimisation multiobjectif offre plusieurs degrés de libertés afin de mieux caractériser mathématiquement un problème, elle peut être formalisée mathématiquement de la manière suivante :

$$\left\{ \begin{array}{l} \text{Minimiser/Maximiser } F(X) = (f_1(X), f_2(X), \dots, f_l(X)) \\ X = \{x_1, \dots, x_n\} \\ h_i(X) = 0, \quad i = 1 \dots m \\ g_j(X) \leq 0, \quad j = 1 \dots k \\ X_{min} \leq X \leq X_{max} \quad i = 1, \dots, n \end{array} \right. \quad (2.22)$$

Les fonctions $f_i(X)$ représentent les objectifs caractéristiques du problème, cette formulation facilite la description de la performance désirée d'une telle conception.

De même que les problèmes d'optimisation mono-objectif $X = \{x_1, \dots, x_n\} \in \mathbb{R}^n$ est le vecteur représentant les variables de décision du problème.

$h_i(X)$ et $g_j(X)$ sont respectivement des contraintes d'égalité et d'inégalité qui ne doivent pas être violées par chaque combinaison des variables de décision afin qu'elle soit retenue comme solution au problème. Bien que l'optimisation multiobjectif permette une meilleure caractérisation mathématique d'un problème, la difficulté majeure réside au niveau de la solution elle-même. En effet, la résolution d'un tel problème en utilisant un algorithme d'optimisation aboutit à un ensemble de solutions nommées solutions non dominées ou front de Pareto. Ceci est dû au fait que les objectifs sont contradictoires et que l'optimisation d'un

paramètre entraîne une détérioration de l'autre. Le choix de la solution à retenir est confié au concepteur qui doit prévoir un compromis entre les objectifs [106], [110].

3.1. Front de Pareto

Le problème d'optimisation multiobjectif intègre un ensemble d'objectifs conflictuels entre eux. La résolution de ce problème en utilisant un algorithme d'optimisation multiobjectif aboutit dans la plupart de cas à une multitude de solutions dont on ne peut trouver une solution pour laquelle tous les objectifs correspondent à leur état optimal simultanément. La solution du problème correspond alors à un ensemble de points de l'espace pour lesquels il est impossible

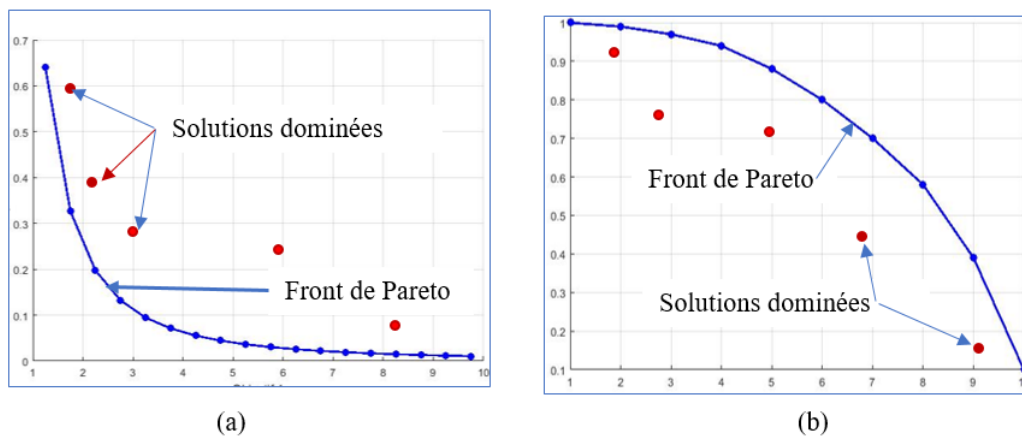


Figure 2. 13. Solutions non dominées ou front de Pareto d'un problème à deux objectifs (a) cas d'une minimisation, (b) cas d'une maximisation

d'améliorer un objectif sans détériorer les autres. On parle alors des solutions non dominées ou front de Pareto. La figure 2.13 illustre la position des solutions dominées et non dominées (front de Pareto) d'un problème d'optimisation à deux objectifs dans le cas d'une minimisation (figure 2.13(a)) et dans le cas d'une maximisation (figure 2.13(b)). Dans le cas de plus de deux objectifs le front de Pareto correspond à une surface [106], [110],[113].

Sur le plan mathématique dans le cas d'une minimisation une solution X_0 est une solution non dominée s'il n'existe aucune solution X pour laquelle $f_i(X) < f_i(X_0) \forall i \in [1, l]$.

L'implication de plusieurs objectifs rend difficile voire impossible le choix d'une solution unique au problème. D'où la grande difficulté de la résolution de ce type du problème [116]–[119]. Cependant, divers techniques ont été proposées pour simplifier et approcher la recherche d'une solution unique à ce problème par la possibilité de sa transformation en un problème d'optimisation mono-objectif. Une étude des relations globale d'harmonie, de conflit et d'indépendance suivie de l'étude des relations locales et composites entre les objectifs est nécessaire. Cette étude a pour objectif de n'intégrer dans le problème que les objectifs significatifs conflictuels [120].

3.2. Etude des relations entre les objectifs d'un problème

Son objectif est d'étudier la possibilité de réduire le nombre d'objectifs sans altérer les caractéristiques du problème. En effet la réduction de sa dimensionnalité permet de réduire les difficultés de calcul et d'augmenter l'approximation du front de Pareto [116], [121], [122].

3.2.1. Relations globales

Dans un problème multiobjectif, pour chaque paire d'objectifs trois types de relations globales à distinguer : l'indépendance, l'harmonie et le conflit. L'étude et la compréhension de ces relations permet une meilleure formulation et résolution du problème. La détermination de ces relations est basée sur le calcul du coefficient de corrélation de Kendall qui est une mesure statistique entre le nombre de solutions concordantes et celles discordantes [117], [122]. Son expression est donnée par :

$$k = \frac{C - D}{C + D} \quad (2.23)$$

k est le coefficient de Kendall, C le nombre de solutions en concordance et D le nombre de solutions en discordance. Ce coefficient est à valeurs dans l'intervalle $[-1, 1]$.

A. Relation d'indépendance

Dans un problème d'optimisation multiobjectif, quelques objectifs peuvent être indépendants les uns des autres. Dans ce cas l'optimisation d'un des objectifs n'affecte jamais l'optimisation des autres. L'optimisation de ces objectifs peut être réalisée séparément en n'introduisant que les variables de décision influençant chacun. En effet, l'optimisation distincte de ces objectifs permettra l'amélioration de leurs performances. Dans le cas de la relation d'indépendance, le coefficient de Kendall est quasi nul [117], [118], [120], [122].

B. Relation d'harmonie

Dans ce cas les objectifs ne sont plus indépendants. L'amélioration d'un objectif entraîne l'amélioration des autres. D'une manière générale, les objectifs en harmonie admettent une solution unique. Dans un problème d'optimisation multiobjectif, l'implication des objectifs en harmonie n'est qu'une redondance qui augmente la complexité et le temps de calcul du problème. La conservation de l'un de ces objectifs et l'omission des autres n'altère en aucun cas les caractéristiques du problème. Dans le cas d'une relation d'harmonie le coefficient de Kendall est positif. Lorsque ce coefficient dépasse 0.5 on dit que les objectifs sont fortement en harmonie [117], [118], [120], [122].

C. Relation du conflit

Dans le cas d'une relation du conflit les objectifs sont dépendants. L'amélioration d'un objectif entraîne une détérioration des autres. C'est en fait le problème majeur qui est à la base du développement des techniques d'optimisation multiobjectif. Ces objectifs nécessitent alors leur implication dans une seule formulation afin de résoudre le problème multiobjectif en question. On parle du conflit entre deux objectifs lorsque leurs coefficients de Kendall est négatif. Lorsque ce coefficient est inférieur à -0.5 on dit que les objectifs sont fortement en conflit [117], [118], [120], [122].

3.2.2. Etude des relations locales et composites entre les objectifs d'un problème multiobjectif

Dans un problème d'optimisation multiobjectif, la connaissance des relations globales entre les objectifs ne suffit pas pour décrire le caractère relationnel entre les objectifs sur tout l'espace objectif. Autrement dit que même si quelques objectifs sont globalement en conflit, on peut trouver des zones dans l'espace objectif où ils ont de bonnes ou acceptables valeurs. On parle alors de l'existence de zones ou de surfaces de compromis entre les objectifs [120].

R. Pinheiro et al ont proposé dans la référence [120] une méthode permettant l'approfondissement de la compréhension des relations locales et composites entre les différents objectifs. La méthode exploite le tableau de Karnaugh utilisé pour simplifier une fonction logique à partir de sa table de vérité pour construire une carte de répartition des solutions dans l'espace objectif (carte des régions). Le principe de la méthode consiste à déterminer d'abord pour chaque objectif le seuil pour lequel on juge que la solution est bonne ou mauvaise. Ce seuil peut être déterminé à partir des connaissances issues de la littérature, ou fixé par le concepteur lui-même. L'utilisation de la moyenne des solutions non-dominées pour chaque objectif est souvent utilisée dans le cas où aucune connaissance a priori n'est connue. Ensuite répartir le nombre ou le pourcentage des solutions correspondant à chaque région de la carte figure 2.14 [120].

	f_2^\uparrow		f_2^\downarrow		
	r_0	r_1	r_3	r_2	f_3^\uparrow
f_4^\uparrow	r_4	r_5	r_7	r_6	f_3^\downarrow
	r_{12}	r_{13}	r_{15}	r_{14}	
f_4^\downarrow	r_8	r_9	r_{11}	r_{10}	f_3^\uparrow
	f_1^\uparrow	f_1^\downarrow		f_1^\uparrow	

Figure 2. 14. Schéma illustrant la carte de régions dans le cas de quatre objectifs

La figure 2.14 représente la carte de régions destinée à la répartition des solutions d'un problème à quatre objectifs. Sur la carte des régions chaque valeur 0 du code gray est remplacée par la flèche haut \uparrow et qui indique que l'objectif en question a une bonne valeur. De même chaque valeur 1 du code gray est remplacée par la flèche bas \downarrow et qui indique que l'objectif en question a une mauvaise valeur. Pour illustrer l'opération de répartition des solutions nous devons mettre dans la région r_0 qui correspond à la valeur (0000, $\uparrow\uparrow\uparrow\uparrow$) le nombre de solutions dont tous les objectifs ont de bonnes valeurs. Lorsque les objectifs sont en conflit cette région ne contiendra aucune solution. D'une manière générale cette carte permet de faciliter l'identification des objectifs ayant à la fois de bonnes valeurs ainsi que de repérer les objectifs présentant des surfaces de compromis [120].

3.2.3. Détermination des objectifs significatifs

Dans un problème d'optimisation multiobjectif, un objectif est dit significatif si l'intervalle dont il évolue est suffisamment large. A l'encontre, un objectif ayant un intervalle étroit est dit non significatif. En effet, en se basant sur les seuils d'acceptation des solutions de chaque objectif, il est évident que les valeurs d'un objectif significatif peuvent être classées facilement en bonnes ou mauvaises vu l'écart qui peut exister entre les valeurs. Cependant dans le cas des objectifs non significatifs les écarts entre les valeurs est faible de façon à ce que l'on peut négliger. Par conséquent les objectifs non significatifs peuvent être ignorés durant le processus d'optimisation car il se peut qu'ils prennent des valeurs convenables vu l'étroitesse de leurs intervalles [120], [123], [124].

Les objectifs significatifs et conflictuels constituent les objectifs caractéristiques du problème. Leur détermination ainsi que la connaissance des compromis existant entre eux permettent de transformer efficacement le problème multiobjectif à l variables en un problème monoobjectif. La méthode de la somme pondérée est habituellement employée pour effectuer ce type de transformation [107].

3.2.4. Méthode de la somme pondérée

Cette méthode est appelée aussi méthode d'agrégation des objectifs. Elle permet de transformer un problème multiobjectif en un problème mono-objectif. Son principe consiste à attribuer pour chaque objectif caractéristique du problème un coefficient de pondération qui reflète son importance dans le problème.

Le problème d'optimisation multicritères de la fonction $F(X) = \{f_1(X), \dots, f_l(X)\}$ est ramené à un problème d'optimisation mono-objectif dont la fonction objectif est donnée par la fonction : $G(X) = \sum_{i=1}^n w_i \cdot f_i(X)$

avec X le vecteur de décision, $G(X)$ la fonction objectif du problème mono-objectif résultant, n le nombre des objectifs caractéristiques restant, w_i le poids correspondant à l'objectif f_i avec $\sum_{i=1}^n w_i = 1$.

Les contraintes entre les objectifs doivent être gardées dans le cas où le problème multiobjectif est avec contrainte [110], [125]–[127].

La nouvelle fonction objectif résultante est à utiliser alors avec un algorithme d'optimisation mono-objectif afin de déterminer une solution unique au problème. Dans notre travail, l'algorithme génétique est l'algorithme d'optimisation utilisé afin de résoudre le problème d'optimisation de l'amélioration de la qualité des images mammographiques ainsi que leur segmentations afin d'en isoler la zone suspecte. Cet algorithme sera détaillé dans le reste de ce chapitre.

4. Algorithmes génétiques

Les algorithmes génétiques ("AG") appartiennent à la famille des techniques stochastiques utilisés pour résoudre les problèmes de l'optimisation globale. Les premières versions des AGs ont été créés par John Holland (1975) puis les travaux de David Goldberg (1989) qui les ont largement enrichies [108], [128]. Leur théorie est fondée sur les mécanismes de l'évolution naturelle des êtres vivants et de la génétique proposée par Charles Darwin en 1860. En effet, lors du processus de recherche d'une solution à un problème, les individus (parents) qui sont mieux adaptés à leur environnement survivent et peuvent participer au processus de reproduction afin de donner naissance à d'autres nouveaux individus (enfants). Au cours de la reproduction, un échange du patrimoine génétique entre les parents est pratiqué par imitation des mécanismes de recombinaison génétique. Ce processus a pour objectif l'enrichissement et la diversification des configurations de l'espace de recherche afin de déterminer l'individu le plus adapté au problème [108], [128], [129].

4.1. Terminologie et concepts de base

La résolution du problème d'optimisation en utilisant la méthode des AGs consiste à déterminer la combinaison convenable parmi les paramètres du problème permettant l'optimisation de la fonction objectif traduisant le problème. L'ensemble de toutes les combinaisons possible constituent L'environnement ou espace de recherche. Dans le cas où le problème est avec contraintes, l'ensemble des combinaisons vérifiant les contraintes s'appelle l'espace réalisable ou l'ensemble des solutions permises. L'AG utilise le codage des paramètres de la fonction à optimiser au lieu des paramètres eux-mêmes. La figure 2.15 illustre la hiérarchie d'organisation des solutions. Une solution réalisable est nommée individu. Ce dernier est composé d'un ou d'un ensemble de chromosomes. Ces derniers sont formés d'un ensemble de gènes qui sont une forme codée des paramètres du problème. En effet, La forme codée des paramètres constitue le génotype d'un chromosome tandis que le phénotype désigne les valeurs des paramètres réels de la fonction objectif qui lui sont associées [108], [128].

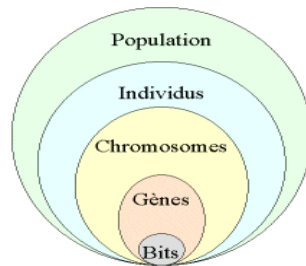


Figure 2. 15. Les différents niveaux d'organisation d'un algorithme génétique

L'AG démarre le processus de la recherche de la solution d'un problème donné à partir de la génération aléatoire d'une population initiale composée d'un ensemble d'individus. D'une manière itérative, l'AG fait l'évoluer par application de ses opérateurs spécifiques qui sont : la sélection des individus les mieux adaptés, leur croisement afin de remplacer les individus moins adaptés et la mutation qui a pour but d'introduire de nouveaux gènes qui n'ont pas d'existence dans les diverses générations. Le processus continue jusqu'à la convergence vers la solution comme le montre la figure 2.16.

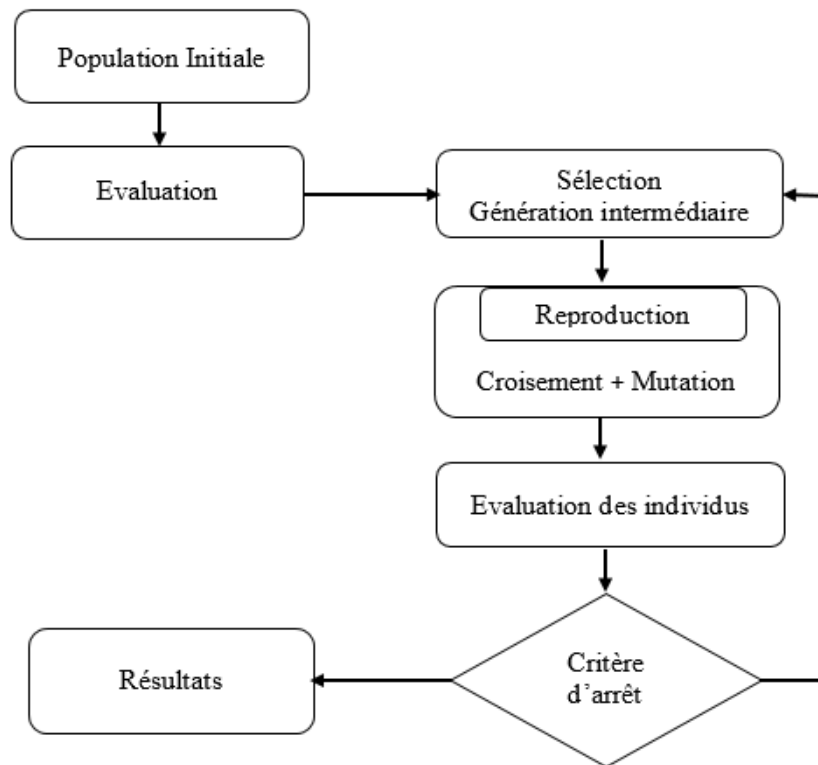


Figure 2. 16. Structure de fonctionnement d'un algorithme génétique.

4.2. Codage des variables

Les algorithmes génétiques ne manipulent pas directement les variables de décision du problème à résoudre. Ils procèdent d'abord à leur codage avant d'aborder le processus de recherche de la solution. Le choix du codage convenable est primordial. En effet c'est en fonction du codage que la structure des chromosomes est déterminée et que l'implémentation des opérateurs de transformation en dépend. Dans la littérature, plusieurs types de codages sont utilisés dont on cite le codage binaire, le codage gray et le codage réel qui sont adaptés à divers problème d'optimisation [108] [112]–[114].

4.2.1. Codage binaire standard

Ce codage était le premier à être utilisé dans le domaine des AGs. Ce type consiste à utiliser une série de 0 et 1 afin de coder une solution (chromosome). Chaque paramètre du problème doit être codé par le plus petit mot binaire qui est bien adapté à sa représentation. Un décodage qui consiste à convertir le mot binaire à la valeur du paramètre équivalent est nécessaire afin de pouvoir évaluer l'adaptation des chromosomes au problème. Le grand avantage de ce type de codage est la simplicité de l'application des opérateurs de croisement et de mutation. Cependant, ses principaux inconvénients résident au niveau de l'augmentation du temps de calcul pour les problèmes de grande taille d'une part, de l'autre part la grande dissimilarité entre deux solutions voisines exemple (7 (0111) et 8 (1000)) mais elles diffèrent de quatre bits

(distance de Hamming). En effet, cette grande dissimilarité peut être derrière la non convergence de l'algorithme vers l'optimum globale [108], [130].

4.2.2. Codage binaire gray

Afin de combler les problèmes causés par le codage binaire standard, essentiellement la difficulté liée au franchissement des "falaises de Hamming", le code gray peut être utilisé. En effet l'une des propriétés de ce type de codage est que pour deux valeurs adjacentes codées sur le même nombre de bits, ces deux valeurs ne diffèrent que par un seul bit (distance de Hamming =1). En utilisant ce type de codage, le passage entre deux valeurs réelles voisines devient beaucoup plus facile puisqu'il suffit de modifier un seul bit dans le chromosome. Ceci aura pour effet d'éviter le piègeage dans les optimaux locaux [107], [130].

4.2.3. Codage réel

Dans ce cas, un chromosome est un vecteur dont les composantes sont présentées par les paramètres du problème à optimiser. Si on suppose que l'on cherche d'optimiser une fonction de n variable $f(x_1, x_2, \dots, x_{n-1}, x_n)$, un chromosome sera représenté par la figure 17

L'avantage de ce type de codage réside dans la rapidité de l'algorithme à la convergence vers un bon résultat. En effet, l'étape de décodage des variables n'est plus nécessaire dans ce cas [130].

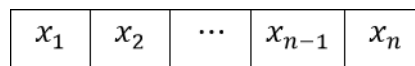


Figure 2. 17. Codage réel d'un chromosome dans le cas de l'optimisation d'une fonction à n variables.

4.3. Evaluation et sélection des chromosomes

La population initiale constitue la matière première à partir de laquelle la solution optimale du problème sera créée. En fait, chaque chromosome est considéré comme solution probable au problème. L'algorithme doit mesurer le degré de l'adaptation de chaque chromosome au problème. La fonction objectif qui est la formulation mathématique du problème à résoudre permet de quantifier cette adaptation. En effet tous les chromosomes seront décodés, chaque variable va prendre sa valeur correspondante à sa forme codée pour calculer la valeur (où les valeurs) de la fonction objectif. C'est à la lumière de cette valeur nommée aussi coût ou fitness que l'algorithme décide l'aptitude d'un chromosome à être sélectionné et survivre afin de contribuer à la génération des descendants plus adaptés à la situation. Divers stratégies de sélection sont utilisées par cette approche.

A. Sélection par rang

Cette méthode consiste à attribuer à chaque chromosome un rang en se basant sur sa valeur de fitness, de façon à ce que le rang augmente avec l'augmentation de l'adéquation du chromosome au problème. Ainsi la probabilité de sélection d'un individu (i) est calculée de manière suivante :

$$P_i = \frac{Rang(Ch_i)}{\sum_{j=1}^N Rang(Ch_j)} \quad (2.24)$$

Avec P_i la probabilité de sélectionner le chromosome Ch_i ; N le nombre de chromosomes.

B. Sélection par roulette

Cette méthode tire son existence à partir de la roue de loterie. Chaque individu lui est attribué une portion de la roue en fonction de la valeur de sa fitness. Par analogie après entrainement de la roue, cette dernière s'arrête en indiquant le chromosome sélectionné. Les individus les mieux adaptés au problème ont statistiquement une grande chance de survivre. Cette méthode donne aussi une chance de l'exploitation du patrimoine génétique des moins adaptés. La probabilité de survie de chaque individu est calculée cette fois ci en utilisant les valeurs de fitness de chaque individu.

$$P_i = \frac{f(Ch_i)}{\sum_{j=1}^N f(Ch_j)} \quad (2.25)$$

$f(Ch_i)$ est la fitness du chromosome Ch_i .

C. Sélection par tournoi :

Cette méthode comme son nom l'indique, met en jeu une compétition entre les individus d'un sous-groupe choisis parmi les chromosomes de la population. Le principe consiste à effectuer un tirage avec remise, comparer leurs valeurs de fitness, par la suite choisir le meilleur parmi eux. Le processus est répété jusqu'à l'obtention du nombre de parents à conserver. Cette méthode a aussi l'avantage de garder la diversité du patrimoine génétique vue qu'elle garde la chance de la survie des individus moins adaptés. En effet ils peuvent participer plusieurs fois à plusieurs compétitions.

D. Reste stochastique sans remplacement

Cette méthode combine la sélection déterministe et la sélection par roulette. D'abord en fonction de la fitness de chaque individu, un nombre de futurs parents est sélectionné. Le principe consiste à calculer pour chaque individu son nombre de copies donné par

$$N_i = E \left(\frac{f_i}{f_m} \right) \quad (2.26)$$

Où N_i le nombre de copie de l'individu (i) ; f_i sa fitness ; f_m la fitness moyenne de toute la population.

Le nombre des individus qui seront sélectionnés est $N_{s=\sum_{i=1}^N N_i}$, d'une manière générale peu d'individus peuvent être sélectionnés dans cette phase, alors le reste est complété d'une manière aléatoire en utilisant la sélection par la roulette. L'évaluation d'un individu est donnée par le reste stochastique $r_i = f_i - E(f_i)$.

4.4. Reproduction et recombinaison génétique

Le principe de recherche de la solution optimale consiste à faire évoluer la population initiale jusqu'à ce que l'on trouve le meilleur individu. Cette évolution est régit par deux opérateurs principaux qui sont le croisement et la mutation. En effet ces opérateurs garantissent la diversification des individus ainsi qu'une exploration intelligente de l'espace de recherche.

E. Opérateur de croisement

Durant le processus d'évolution de l'algorithme génétique, l'opérateur de croisement imite l'opération de la reproduction chez les êtres vivants. Le croisement est effectué entre deux chromosomes dits parents donnant naissance à deux nouveaux chromosomes qui s'appellent enfants. Ces derniers vont remplacer les chromosomes rejetés jugés inaptes à répondre au besoin du décideur et compléter la population. Cet opérateur permet l'échange du patrimoine génétique entre deux parents choisis aléatoirement afin d'enrichir la diversité de la population. Les chromosomes descendants peuvent hériter les bons gènes de leurs parents et deviennent plus adaptés que leurs ancêtres. Dans la littérature, divers méthodes de croisement existent dont on cite le croisement en un point, ou en multiples points (figure 2.18).

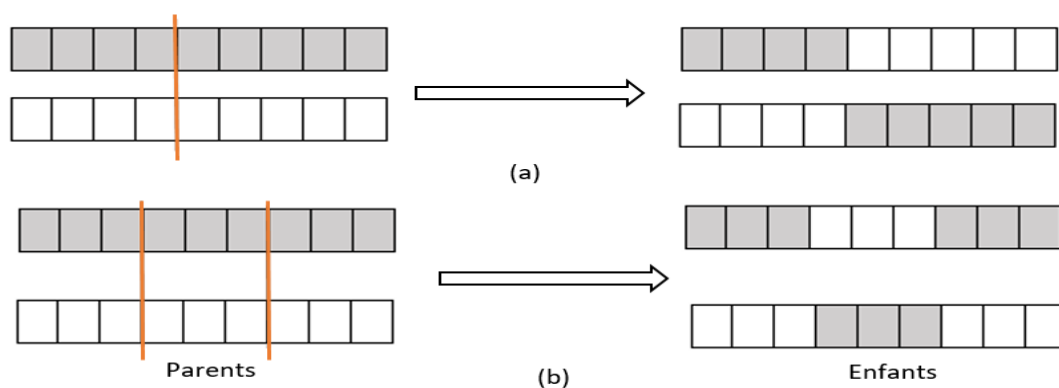


Figure 2. 18. Croisement entre deux chromosomes ; (a) à un seul point, (b) à deux points.

Ce type de croisement peut être appliqué pour tous type de codage réel ou binaire. Cependant dans le cas de codage réel un autre croisement dit croisement de barycentrique qui peut être appliqué de la façon suivante :

$$\begin{cases} E_1(i) = \beta \cdot P_1(i) + (1 - \beta) \cdot P_2(i) \\ E_2(i) = (1 - \beta) \cdot P_1(i) + \beta \cdot P_2(i) \end{cases} \quad (2.27)$$

Où E_1 et E_2 sont les enfants résultants du croisement des parent P_1 et P_2 ; β est un coefficient de pondération aléatoire. Il doit être adapté au domaine d'extension des gènes. Dans la littérature β peut prendre des valeurs dans l'intervalle $[0, 1]$ ou dans $[-0.5, 1.5]$ si on désire générer des valeurs à l'extérieur des deux gènes en cours [112], [131].

F. Opérateur de mutation

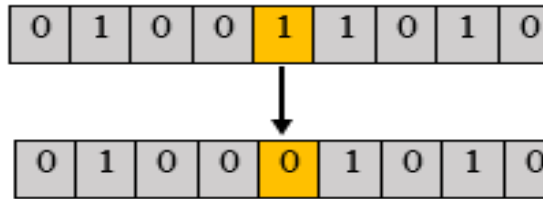


Figure 2. 19. Mutation d'un allèle d'un chromosome à la cinquième position

Par analogie avec la mutation biologique caractérisée par une modification de l'information génétique survenue dans le génome d'une cellule, l'opérateur de mutation en algorithmes génétiques a comme effet de modifier aléatoirement la valeur d'un gène au niveau d'un chromosome. Cet opérateur a pour objectif d'introduire une diversité de gènes au niveau des chromosomes. En effet, la mutation peut introduire des caractères génétiques non existants dans les parents. D'où la capacité de l'algorithme à balayer efficacement l'espace de décision et d'éviter le piègeage dans un optimum local. Dans le cas d'un codage binaire la mutation entraîne le changement du bit 0 par 1 et vice versa figure 2.19.

De même pour le codage réel l'algorithme génère aléatoirement le chromosome et plus précisément le gène qui va subir la mutation. Par analogie avec la mutation binaire le gène en question changera de valeur par une autre dans son intervalle de définition.

Au cours de l'évolution de l'algorithme, la dernière génération ne subit pas de mutation, de même, il est préférable que le meilleur individu ne subira pas de mutation afin de garder l'élite de chaque itération [108], [126], [128].

Lorsqu'il s'agit d'un problème d'optimisation avec contraintes, la mutation de certains paramètres du problème ne doit pas violer les contraintes existantes entre eux. Dans ce cas le processus de mutation des paramètres est géré par une fonction qui contrôle l'adaptabilité et la faisabilité d'une telle mutation. Ce mode de mutation conduit à un brassage chromosomique important qui doit être contrôlé par des procédures de sélection extinctrices pour supprimer les "mauvais" mutants engendrés [107], [132].

4.5. Critère d'arrêt

Le critère d'arrêt d'un algorithme génétique est la condition qui permet de stopper le processus de recherche de la solution optimale et maintenir le meilleur individu dans le cas de l'optimisation mono-objectif ou les meilleurs individus dans le cas de l'optimisation multiobjectif comme solution du problème en question. Ce critère peut être le nombre de génération, le temps d'évolution, l'évolution faible de la valeur de la fonction de fitness après un nombre déterminé d'itérations [108], [133].

5. Conclusion

Dans ce chapitre nous avons mis le point sur quelques techniques et méthodes pouvant être utilisées dans la conception et l'optimisation des phases de prétraitement et de segmentation d'un système d'aide à la détection "DAOe" des masses mammaires.

Dans la phase de prétraitement, un ensemble de techniques permettant l'amélioration de la qualité des images mammographiques à examiner sont présentées et détaillées dont les techniques usuelles basées sur la modification de l'histogramme et d'autres transformations point à point. Dans notre travail, la technique basée sur la transformation point à point par la fonction sigmoïde sera employée, vu les avantages qu'elle présente.

Quant à la phase de segmentation, un ensemble de techniques connues dans la littérature, issues de différentes catégories, ont été présentées dont nous avons retenu les techniques suivantes : la technique basée sur le seuillage global, celle basée sur la croissance de régions et la technique basée sur la classification des pixels k-means. Ce choix est basé sur le fait que le processus de démarrage de partitionnement de l'image par ces techniques est automatique ou peut être automatisé facilement.

Par ailleurs, l'intégration de l'outil de l'optimisation dans le processus du fonctionnement des deux étapes, prétraitement et segmentation, est d'une importance grandissante dans l'amélioration de la qualité des images traitées ainsi que dans leurs segmentations. Dans cette optique, un ensemble de notions liées au concept ainsi qu'à la résolution de l'optimisation globale sous ses deux aspects mono-objectif et multiobjectif ont été développés.

Cet ensemble de méthodes et techniques nous seront utiles par la suite pour la conception d'un nouveau modèle d'amélioration optimale de la qualité des images mammographiques.

Proposition d'un modèle pour l'amélioration optimale des images mammographiques

1. Introduction

Ce chapitre est consacré à la proposition d'un nouveau modèle permettant une meilleure amélioration de la qualité des images mammographiques par optimisation de leurs paramètres caractéristiques tout en évitant les problèmes de la sur-amélioration des images traitées par les ex-modèles à optimisation [1], [2]. Il s'agit d'un modèle exploitant les propriétés de la fonction sigmoïde à l'échelle locale et, géré par un algorithme d'optimisation multiobjectif. Ce modèle permet d'optimiser simultanément divers paramètres de l'évaluation de la qualité de l'image EME, EC, AMBE et FSIM (qui traduisent respectivement l'adéquation du contraste, la contenance de bords, la conservation de la luminosité et la conservation des informations des images originales). A noter que ces paramètres sont appelés les paramètres "IQA" (Image Quality Assessment).

Pour concevoir ledit modèle nous avons commencé à proposer d'abord un modèle à optimisation mono-objectif vu sa simplicité. Pour ce faire on s'est basé sur un modèle usuel à optimisation [134]–[137] par changement de sa fonction de transformation par la fonction sigmoïde (équation 3.13) tout en conservant sa fonction objectif (équation 12). Quoique ce modèle à optimisation monoobjectif ait amélioré nettement les résultats des ex-modèles à optimisation, sauf qu'il s'avère possible d'être amélioré davantage par optimisation multiobjectif suite à la recherche faite. Ce choix connaît une complexité qui réside dans la détermination d'une solution unique au problème. Néanmoins, dans le cadre d'une approche simplifiée, il est possible de rechercher une solution unique en transformant le problème d'optimisation multiobjectif en un problème d'optimisation mono-objectif. Nombreuses sont les techniques permettant cette transformation parmi lesquelles nous avons choisi la méthode de la somme pondérée des objectifs significatifs contradictoires. Ci-après nous trouverons les études réalisées pour la conception des deux modèles mono- et multi-objectif (description, analyse, résultats, comparaison). Les études sont réalisées sur des images mammographiques test issues de la base de données Mini-MIAS. Les résultats fournis par notre modèle (qualité des images et les paramètres d'évaluation "IQA") seront confrontés à ceux fournis par les techniques usuelles (aussi bien celles basées sur la modification d'histogramme que celle basée sur l'optimisation). Pour cette raison avant de proposer nos modèles nous avons commencé à évaluer les paramètres "IQA" des images test fournis par les techniques usuelles.

Mais avant tout quelles sont les imperfections liées aux images mammographiques pour qu'elles soient traitées afin d'améliorer leurs qualités ?

2. Imperfections des images mammographiques et améliorations

2.1. Problèmes liés aux images mammographiques

Il est d'autant très connu que la mammographie ne fournit pas de bon contraste ni entre les tissus glandulaires normaux du sein et les tissus malins avoisinants, ni entre les lésions cancéreuses et le fond d'image, en particulier dans le cas des seins denses [9], [63], [138]–[141]. Selon la loi d'atténuation, plus le tissu est épais, moins les photons le traversent. Cela signifie que lorsque le faisceau des RX traverse les tissus glandulaires normaux du sein et les tissus malins dans les tissus mammaires denses, la différence d'atténuation entre les deux tissus sera faible et par conséquent il y aura un faible contraste entre les tissus glandulaires normaux et les tissus malins [9]. Un autre problème lié aux images mammographiques réside dans l'existence des bords flous en raison de l'effet de volume partiel et de l'intensité altérée dans les mêmes régions de tissu en raison du champ déformé [8], [9]. L'un des autres problèmes qui accompagne l'acquisition des mammogrammes est le bruit qui se produit lorsque la luminosité de l'image n'est plus uniforme dans les zones qui représentent les mêmes tissus. Ceci est supposé être dû à une distribution non uniforme des photons. Ce problème peut être réduit en augmentant le temps d'exposition du sein, mais pour des raisons d'irradiation, les radiologues préfèrent de le diminuer au détriment de l'augmentation du bruit quantique [9]. Ainsi, les mammogrammes sont altérés par l'existence des artéfacts radio-opaques et du muscle pectoral qui influent négativement sur leurs analyses automatiques [142]–[144]. L'ensemble de ces problèmes constitue la difficulté majeure qui empêche la précision de l'analyse subjective de l'image ainsi que sa segmentation [8]. Afin de détecter correctement les anomalies présentes dans les mammogrammes on aura recours à leurs traitements en faisant appel à des techniques d'amélioration de la qualité d'images.

2.2. Amélioration de la qualité des mammogrammes : contexte de l'approche proposée

D'une manière générale, l'amélioration de la qualité des images mammographiques a pour but de rehausser le contraste et de renforcer l'apparition des bords entre les différentes structures anatomiques de l'image sans altérer autant que possible les informations sensibles aux changements de luminosité ainsi que les informations globales contenues dans l'image origine [145], [146].

Les techniques récentes basées sur l'optimisation [147] introduisent souvent des paramètres caractéristiques permettant le contrôle du processus de traitement de l'image à améliorer afin d'aboutir à une meilleure image améliorée. Chose qui peut être prouvée par l'optimisation des

paramètres d'évaluation de la qualité des images "IQA" [59], [73], [145], [147]–[150]. Les travaux récents ont introduit les paramètres EC, EME, AMBE et FSIM comme paramètres capables d'évaluer objectivement la qualité de l'image. Parmi ces travaux on cite en particulier :

- La technique proposée par N.H. Chan et al (2019) "*An Improved Enhancement Technique for Mammogram Image Analysis : A Fuzzy Rule-Based Approach of Contrast Enhancement*" [141].
- La technique proposée par E.Daniel et al. (2016) "*Optimum wavelet based masking for the contrast enhancement of medical images using enhanced cuckoo search algorithm*" [69].
- La technique proposée par A. Gandhamal et ses collaborateurs en 2017 intitulée "*Local Gray Level S-Curve Transformation – A Generalized Contrast Enhancement Technique for Medical Images*" [8].

C'est sur cette dernière technique de Gandhamal que notre approche s'est basée pour les raisons suivantes :

- la technique a ciblé uniquement les images médicales ;
- la technique est basée sur la fonction sigmoïde ;
- la technique a employé quatre paramètres EC, EME, AMBE et FSIM pour évaluer la qualité des images alors que les autres techniques en ont employé moins ;
- Les images traitées ont été bien améliorées.

Bien que la technique de Gandhamal ait montré son efficacité au niveau de l'amélioration de la qualité d'un ensemble de modalités des images médicales, sauf qu'elle ne permet pas le contrôle du processus de l'amélioration des images. Par ailleurs cette méthode traite toutes les images de la même manière (toutes les images sont traitées par la même fonction), or il s'avère d'après notre étude que l'adaptation de la transformation pour chaque image pourrait améliorer considérablement les résultats de cette technique. En effet, c'est en fonction des niveaux de gris de chaque image et de leurs arrangements spatiaux, qu'on peut trouver la combinaison idéale de la subdivision en blocs et la fonction sigmoïde pouvant optimiser l'amélioration de la qualité des mammogrammes.

Notre modèle sera appliqué sur des images mammographiques issues de la base de données Mini-MIAS. Les paramètres d'évaluation (*EME, EC, AMBE, FSIM*) de la qualité des images traitées par notre modèle seront comparés à ceux résultants :

- des méthodes usuelles (les méthodes basées sur la modification de l'histogramme et sur l'optimisation) ;
- de la méthode de Gandhamal et al.[8].

2.3. Description et expressions des paramètres d'évaluation "IQA"

Dans le domaine spatial, divers paramètres destinés à l'évaluation objective de la qualité des images sont dérivés de la mesure du contraste de Michelson ou de la loi de Weber-Fechner [145]. Ces paramètres peuvent être classés selon l'utilisation ou non d'une image de référence ces paramètres peuvent être :

- avec référence complète lorsque l'image de référence est disponible,
- avec référence réduite lorsque l'image de référence existe partiellement et que quelques caractéristiques extraites servent comme information qui aide à l'évaluation ;
- sans référence lorsqu'il n'y a pas d'image de référence. C'est ce qu'on appelle aussi l'évaluation aveugle de la qualité [59], [150].

Dans ce qui suit nous trouverons une brève description et les expressions des différents paramètres d'évaluation de la qualité des images mammographiques.

2.3.1. Mesure effective de l'amélioration : **EME** (Effective Measure of Enhancement)

Le paramètre *EME* appartient à la catégorie des paramètres de l'évaluation aveugle. Il est utilisé pour quantifier et évaluer le degré de l'amélioration ou de dégradation du contraste d'une image. Ce paramètre est calculé pour l'image origine et l'image traitée et la comparaison permet d'en déduire si le contraste de l'image a subi une amélioration ou une dégradation. Le paramètre *EME* est calculé selon l'expression ci-après après avoir divisé l'image en un ensemble de blocs $B(i, j)$ non superposables [8], [59], [99], [145], [151] :

$$EME = \frac{1}{N_h \cdot N_v} \sum_{i=1}^{N_h} \sum_{j=1}^{N_v} 20 \log \left(\frac{I_{max}(i, j)}{I_{min}(i, j)} \right) \quad (3.1)$$

Avec N_h et N_v sont respectivement les nombres de blocs dans les sens horizontal et vertical ; $I_{max}(i, j)$ et $I_{min}(i, j)$ sont respectivement les intensités maximale et minimale dans chaque bloc $B(i, j)$.

2.3.2. Contenance de bord : **EC** (Edge Content)

Le paramètre *EC* est aussi parmi les paramètres de l'évaluation aveugle. Il est utilisé afin d'estimer la contenance des bords d'une image. Un bord correspond à des changements locaux importants dans une image en donnant naissance généralement aux limites entre deux régions distinctes [8], [99], [152]. Plus la contenance en bords augmente meilleure sera la discrimination entre les structures différentes dans l'image. Le calcul de ce paramètre est basé sur la détection du gradient du niveau de gris dans l'image (voir chapitre 2 §3). Le paramètre *EC* est alors la norme du gradient des intensités des niveaux de gris, il est donné par l'expression :

$$EC = \frac{1}{M \cdot N} \sum_x \sum_y |G(x, y)| \quad (3.2)$$

$$|G(x, y)| = \sqrt{G_x^2 + G_y^2} = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial I(x, y)}{\partial x} \\ \frac{\partial I(x, y)}{\partial y} \end{bmatrix} \quad (3.3)$$

Avec M et N la taille de l'images, $I(x, y)$ intensité de niveau de gris du pixel de coordonnée (x, y) .

G_x et G_y représentent les gradients des niveaux de gris dans les sens horizontal et vertical. Ils peuvent être calculés en utilisant l'un des détecteurs de bords tels que l'opérateur de Prewitt ou celui de Sobel [8], [99].

2.3.3. Erreur moyenne absolue de luminosité : **AMBE** (Absolute Mean Brightness Error)

C'est un paramètre de la catégorie référence complète. Il est utilisé pour évaluer le degré de distorsion de la luminance entre les images origine et traitée. La valeur du paramètre *AMBE* est donnée par la différence des moyennes de luminosité entre les images origine et traitée telle que :

$$AMBE = |E(I_e) - E(I_s)| \quad (3.4)$$

Où $E(I_e)$ est $E(I_s)$ sont respectivement les valeurs moyennes des intensités de l'image origine et l'image traitée [59], [99], [153].

A noter qu'une valeur tendant vers zéro de ce paramètre indique une bonne préservation des informations sensibles à la luminosité.

Le paramètre *AMBE* peut être donné aussi sous une forme normalisée "*AMBE_n*" selon l'expression ci-après [15] :

$$AMBE_n = \frac{1}{1 + AMBE} \quad (3.5)$$

Le paramètre *AMBE* normalisé "*AMBE_n*" appartient donc à l'intervalle [0, 1]. Plus la valeur de ce paramètre ne tend vers 1 meilleure est la qualité de l'amélioration.

2.3.4. Mesure d'indice de similarité des caractéristiques : **FSIM** (Feature Similarity Index Measure)

Il s'agit d'un paramètre "IQA" de catégorie référence complète basée sur la structure contrairement aux autres paramètres qui se sont basés sur les pixels. Ce paramètre est fondé sur le fait que le système visuel humain (SVH) comprend une image principalement en fonction de ses caractéristiques de bas niveau plutôt que de haut niveau. En effet, le paramètre *FSIM* est

conçu de façon à atteindre une cohérence beaucoup élevée avec les évaluations subjectives (vision humaine). La description mathématique du paramètre FSIM fait appel à deux caractéristiques qui sont : la congruence de phase (Phase Congruency "PC") et l'amplitude du gradient (Gradient Magnitude "GM"). La congruence de phase est l'une des méthodes de détection des caractéristiques d'une image. En effet, les caractéristiques d'une image sont perçues aux points où la congruence de phase est maximale postule le modèle énergétique local de détection des caractéristiques d'une image (Morrone et al. 1986, 1987) [154]. Le paramètre *FSIM* est calculé selon l'expression suivante :

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\sum_{x \in \Omega} PC_m(x)} \quad (3.6)$$

Où Ω est l'ensemble formé du domaine spatial de l'image, S_L est la similarité globale entre l'image origine et l'image améliorée, PC_m est le maximum de la congruence de phase entre l'image origine et l'image traitée [148], [149], [155], [156]. Les fondements et le calcul de ce paramètre sont développés en annexe 4.

2.3.5. Mesure de l'entropie d'une image

Une image numérique est un signal bidimensionnel dont chaque pixel est une variable aléatoire discrète représentant les caractéristiques particulières d'un signal lumineux. La moyenne de la quantité d'information apportée par cette variable est quantifiée par l'entropie qui lui est associée. Cette quantification a été introduite en 1945 par Shannon qui a associé pour chaque source modélisée par une variable aléatoire discrète X la notion de l'entropie comme moyenne de la quantité d'information apportée.

En effet, une image en niveau de gris est formée d'un ensemble de pixels P à valeurs entières dans l'intervalle $[0, 255]$ (image codée sur 8 bits). Si on désigne par $p(n_i)$ la probabilité pour qu'un niveau de gris en pixel P soit n_i . La quantité d'informations $I(n_i)$ est donnée par [51], [59], [63]:

$$I(n_i) = \log_2 \frac{1}{p(n_i)} \quad (3.7)$$

L'entropie de chaque pixel de l'image $H(P)$ est définie alors par :

$$H(P) = - \sum_{i=0}^{255} p(n_i) \log_2 p(n_i) \quad (3.8)$$

L'entropie est mesurée en bits et par convention " $0 \log_2 0 = 0$ ". Le nombre $255 = 2^8 - 1$ est la valeur maximale de niveaux de gris dans l'hypothèse où l'image est codée sur 8 bits. plus le nombre de niveaux de gris augmente plus l'entropie de l'image augmente [63], [157], [158].

2.3.6. Rapport signal-bruit de crête : PSNR (Peak Signal-to-Noise Ratio)

Ce paramètre est couramment utilisé pour mesurer les performances des algorithmes concernant l'amplification ou non du bruit. Un PSNR plus élevé indique une meilleure performance antibruit. L'expression de ce paramètre est définie comme suit [66], [137], [156], [159], [160]:

$$PSNR(I_e, I_s) = 10 \times \log_{10} \left(\frac{(L - 1)^2}{MSE(I_e, I_s)} \right) \quad (3.9)$$

Avec $L = 2^n$ est la dynamique de l'image où n est le nombre de bits sur lequel l'image est codée ; $MSE(I_e, I_s)$ la moyenne de l'erreur quadratique calculée par l'expression suivante :

$$MSE(I_e, I_s) = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (I_e(i, j) - I_s(i, j))^2 \quad (3.10)$$

M et N sont respectivement les nombres de lignes et de colonnes de l'image ; $I_e(i, j)$ et $I_s(i, j)$ sont respectivement les intensités des pixels de coordonnées cartésiennes (i, j) des images origine et traitée.

A signaler que le paramètre FSIM est publié par son auteur sur le site <https://www4.comp.polyu.edu.hk/~cslzhang/IQA/FSIM/FSIM.htm>. Les paramètres PSNR et l'entropie sont intégrés dans l'environnement Matlab. Les paramètres inexistant dans l'environnement Matlab tels que (EC, EME et AMBE) je les ai programmés.

3. Evaluation des paramètres "IQA" des images test par les techniques usuelles

Les techniques usuelles utilisées dans le traitement des images mammographiques ce sont de deux types celles basées sur la modification de l'histogramme d'une part, et celles basées sur l'optimisation d'autre part. Notre but c'est de comparer les résultats de notre modèle à proposer à ces techniques à travers la quantification de leurs différents paramètres d'évaluation "IQA".

Concernant les techniques d'optimisation notre modèle sera comparé au modèle d'optimisation usuel (le modèle le plus utilisé) [134]–[137].

Les évaluations des différentes techniques, y compris notre modèle, sont menées sur les images test mdb209, mdb210, mbd211 et mdb212 extraites de la base de données Mini-MIAS, déjà utilisées dans le travail de E. Daniel et al. [69].

3.1. Techniques usuelles basées sur la modification de l'histogramme

Le but de cette partie est d'étudier les effets de traitement des images test par les trois variantes de la technique basée sur la modification d'histogramme à savoir HE, BBHE et CLAHE. A cet effet ces techniques sont appelées par leurs codes dans l'environnement Matlab "R2015a". Les résultats obtenus (les paramètres d'évaluation "IQA" et les images) avant et après le traitement sont récapitulés dans le tableau 3.1 et la figure 3.1.

3.1.1. Présentation et analyse des paramètres d'évaluation "IQA"

Le tableau 3.1 présente les valeurs des paramètres d'évaluation "IQA" des différentes images test traitées par les différentes techniques à base d'histogramme.

Tableau 3. 1. Paramètres d'évaluation des images traitées par les techniques HE, BBHE et CLAHE

Image		EC	EME	AMBE	FSIM	PSNR	Intervalle du niveau de gris
mdb209	Origine	1.345	6.781	0	100%	--	[0, 228]
	HE	1.604	0.929	102.834	85.7%	7.710	[105, 255]
	BBHE	1.939	2.784	25.539	90.7%	19.101	[0, 255]
	CLAHE	4.297	5.679	20.672	79.6%	17.330	[1, 255]
mdb210	Origine	1.326	6.355	0	100%	--	[0, 235]
	HE	1.637	0.926	97.110	85.7%	8.192	[101, 255]
	BBHE	2.151	2.624	26.265	88.3%	18.256	[0, 255]
	CLAHE	4.564	5.822	18.029	76.5%	17.244	[1, 255]
mdb211	Origine	1.119	4.786	0	100%	--	[0, 227]
	HE	0.792	0.367	112.561	86.8%	6.64	[138, 255]
	BBHE	1.459	1.682	27.036	91.2%	17.101	[0, 255]
	CLAHE	3.543	4.493	6.858	80.9%	18.36	[0, 255]
mdb212	Origine	0.954	5.267	0	100%	--	[0, 225]
	HE	0.933	0.4494	114.291	84.1%	6.579	[134, 255]
	BBHE	1.371	1.725	24.636	90.0%	17.782	[0, 255]
	CLAHE	3.464	4.625	8.421	80.3%	18.69	[0, 255]

D'après ce tableau on remarque que les techniques basées sur la modification de l'histogramme :

- ont amélioré le paramètre EC (contenance de bords) pour toutes les images, sauf pour les images mdb211 et mdb212 où la valeur de EC a été diminuée par la technique HE.
- ont détérioré le contraste de toutes les images traitées vu la diminution du paramètre EME pour ces images.
- De point de vue conservation de la luminosité, les résultats montrent que les techniques basées sur la modification de l'histogramme ne parviennent pas à respecter cette exigence. Comme cela est montré par les valeurs du paramètre AMBE qui sont élevées même trop pour la technique HE.

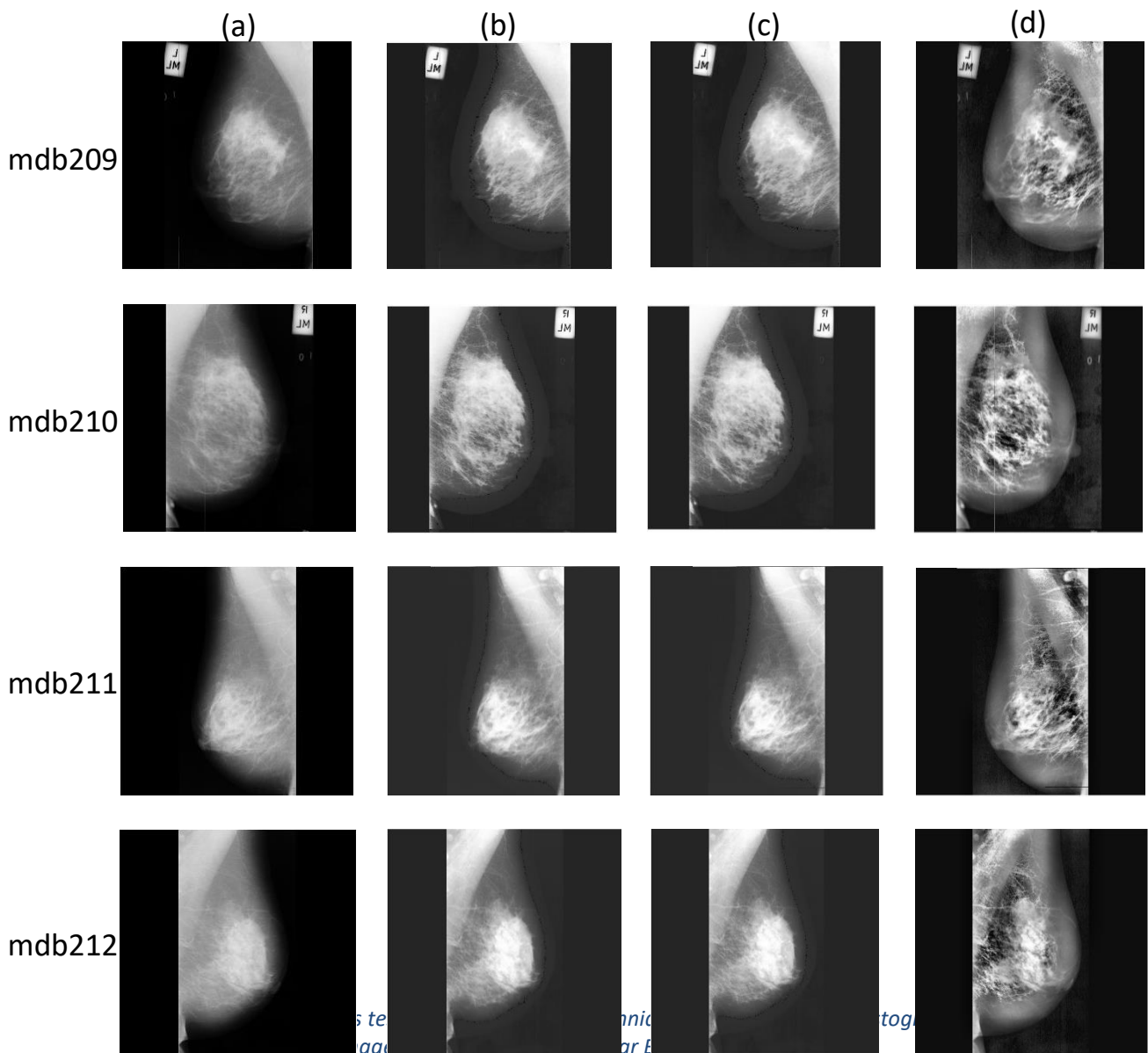
- Concernant la conservation des informations de l'image origine, ces techniques ont montré leurs limitations vu que le paramètre FSIM n'a pas dépassé 90 % dans les meilleurs cas.

- Concernant l'intervalle du niveau de gris la technique HE a décalé cette intervalle vers les intensités élevées. Les deux techniques BBHE et CLAHE ont modifié cet intervalle de façon à occuper la totalité de l'intervalle des niveaux de gris [0,255].

Les techniques CLAHE et BBHE amplifient moins le bruit que la technique HE du fait que la moyenne du paramètre PSNR élevé (environ 18) est inférieure à celle HE (PSNR 7)

3.1.2. Présentation et analyse des images traitées

La figure 3.1 ci-dessous présente les photos des images test avant et après traitement par les techniques d'histogramme. L'inspection visuelle de cette figure montre que les images traitées Figure 3.1 (b-c-d)) sont trop brillantes. Ceci peut être justifié par la valeur élevée du paramètre



AMBE (Tableau 3.1). Aussi, une analyse minutieuse des images originales et leurs correspondantes traitées par la technique CLAHE figure 3.1 (d), montre que certains détails sont perdus dans les images traitées et cela reste en concordance avec la faible valeur du paramètre FSIM (qui ne dépasse pas 81 %) (Tableau 3.1). Bref la qualité des images traitées est en concordance avec les valeurs des paramètres "IQA" évalués.

En résumé, les techniques à modification d'histogramme ne permettent pas une meilleure amélioration de la qualité des images. D'où le recours au développement des techniques pouvant être plus performantes. Dans ce sens et d'après notre étude il s'avère que les techniques dont le processus d'amélioration est contrôlable aboutissent à des résultats meilleurs. En effet différentes méthodes ont exploité les techniques de l'intelligence artificielle afin de guider les algorithmes vers les résultats optimaux dont l'optimisation globale fait partie [161].

3.2. Technique usuelle basée sur l'Optimisation ("TUBO")

Puisque notre modèle est à base d'optimisation, il s'avérerait judicieux de se baser sur un des modèles d'optimisation les plus utilisés (modèle usuel). En effet, les techniques d'amélioration à base d'optimisation permettent le contrôle des paramètres "IQA" de l'image en cours de traitement. Ainsi ces techniques permettent d'orienter et d'ajuster le processus d'amélioration de la qualité des images durant le traitement [31].

La qualité de chaque image est améliorée en sélectionnant les paramètres optimaux du contrôle des techniques utilisées. Cette sélection ne peut se faire que sur la base d'une fonction objectif (à optimiser) dont le processus d'amélioration évalue un ensemble d'images jusqu'à déterminer la meilleure. Dans le cas des images en niveau de gris, l'amélioration de leurs qualités consiste à manipuler la distribution des niveaux de gris au voisinage de chaque pixel. Cette manipulation est faite par application d'une fonction de transformation appelée technique de rehaussement local, l'image ainsi formée est évaluée par une fonction dite fonction objectif afin de lui attribuer un score. La meilleure image améliorée est celle qui a un bon score.

3.2.1. Présentation du modèle d'optimisation usuel

Dans la littérature le modèle d'optimisation de l'amélioration de la qualité d'images le plus utilisé et qui est exploité jusqu'à nos jours, utilise la fonction de transformation locale donnée par l'équation suivante [136] :

$$I_s(i, j) = \frac{k \cdot D}{\sigma(i, j) + b} [I_e(i, j) - c \cdot m(i, j)] + m(i, j)^a \quad (3.11)$$

Avec $D = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N I_e(i, j)$ représente la moyenne globale de l'image d'entrée où M et N représentent respectivement les nombres de lignes et de colonnes.

$m(i, j) = \frac{1}{n \times n} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} I_e(i, j)$ la moyenne locale du voisinage du pixel en cours avec ($n = 3, 5, 7, \dots$).

$\sigma(i, j) = \sqrt{\frac{1}{n \times n} \sum_{x=1}^n \sum_{y=1}^n (I_e(x, y) - m(i, j))^2}$ est l'écart type local du pixel de coordonnées (i, j) par rapport à son voisinage de taille $n \times n$.

Le problème d'amélioration du contraste est ramené alors à déterminer les paramètres de décision $a \in [0, 1.5], b \in \left[0, \left(\frac{D}{2}\right)\right], c \in [0, 1], k \in [0.5, 1.5]$ qui permettent de maximiser la fonction objectif $F(I_s)$ donnée par l'équation suivante :

$$F(I_s) = \log(\log(E(I_s))) \cdot \frac{n_{edgels}(I_s)}{M \times N} \cdot H(I_s) \quad (3.12)$$

La fonction objectif $F(I_s)$ évalue la qualité de l'image traitée I_s ;

$E(I_s)$ est la somme des intensités des pixels de bords calculée par l'application de l'un des détecteurs de bords (Sobel, Prewitt,...) ;

$n_{edgels}(I_s)$ est le nombre de pixels de bords de l'image traitée ;

$H(I_s)$ est l'entropie de l'image traitée.

Divers travaux ont exploité ce modèle afin d'améliorer la qualité des images y compris celles issues du domaine médical dont nous citons : Apurba Gorai et al (2009) [162], Sanjay Agrawal et al (2012) [163], Nirmal Singh et al (2013) [134], Dharendra Pal Singh et al (2014) [164], Milad Maleki et al (2015) [161]. K. Sowjanya et al (2017) [136] et N. Mohanapriya et al (2019) [137]. Ces deux derniers auteurs ont montré dans leurs travaux que l'utilisation de ce modèle conduit à une sur-amélioration des images traitées. Dans ce sens nous avons contribué à l'évaluation de la performance dudit modèle par la programmation (sous Matlab 2015Ra) de sa transformation locale (Equation 3.11) et de sa fonction objectif (Equation 3.12).

La boîte à outils d'optimisation de *MATLAB R2015a* est exploitée pour gérer le processus d'optimisation afin de déterminer le vecteur de décision (a, b, c, k) selon l'algorithme suivant :

Algorithme 3. 1. Algorithme d'optimisation de la qualité des mammogrammes par la technique usuelle basée sur l'optimisation

Algorithme 3.1 : Amélioration de la qualité des mammogrammes par la technique usuelle basé sur l'optimisation

Données : Image

Vers l'algorithme génétique

- Générer la population initiale 50 chromosomes (a, b, c, k)
- Améliorer l'image **Image** par chaque chromosome (Transformation 3.11)
- Calculer la fitness de chaque individu fonction objectif (Equation 3.12)

Répéter

- Sélection des individus (stochastique uniforme)
- Croisement (point unique)
- Mutation (adaptive possible)
- Améliorer l'image **Image** par chaque chromosome (Transformation 3.11)
- Calculer la fitness de chaque individu fonction objectif (Equation 3.12)

Jusqu'à ce que la population converge ou critère d'arrêt est vrai (nombre d'itération =100)

- Calculer la matrice image correspondante au meilleur vecteur (a, b, c, k)
 - Afficher l'image améliorée
-

3.2.2. Présentation et analyse des paramètres d'évaluation "IQA" du modèle "TUBO"

L'application de l'algorithme 3.1 aux images mammographiques afin d'améliorer leurs qualités a abouti à la détermination des vecteurs de décision récapitulés dans le tableau 3.2. L'amélioration de la qualité d'une image, est obtenue par son traitement avec la transformation de rehaussement local donnée par l'équation 3.11, en utilisant pour chaque image son vecteur de décision du tableau 3.2, obtenu par application de l'algorithme 3.1.

Tableau 3. 2. Les vecteurs de décision des images test traitées par le modèle d'optimisation usuel

	a	b	c	k
mdb209	0.0651	0.0175	0.8142	1.4999
mdb210	0.2214	0.0033	0.9106	1.5000
mdb211	0.9481	0.4093	0.9492	1.4999
mdb212	0.5404	0.0009	0.9519	1.4999

Le tableau 3.3 ci-dessous présente les valeurs des paramètres d'évaluation "IQA" de la technique usuelle basée sur l'optimisation qu'on vient de noter "TUBO" qui représente un modèle de référence pour notre modèle à optimisation monoobjectif.

Tableau 3. 3. Paramètres d'évaluation "IQA" par la technique usuelle à optimisation

Image Test		EC	EME	AMBE	FSIM	PSNR	Intervalle du niveau de gris
mdb209	Origine	1.345	6.781	0	100%	--	[0, 228]
	TUBO*	4.484	21.707	19.438	95.98%	17.72	[0, 255]
mdb210	Original	1.326	6.355	0	100%	--	[0, 235]
	TUBO	8.057	29.248	12.108	94.09%	20.826	[0, 255]
mdb211	Original	1.119	4.786	0	100%	--	[0, 227]
	TUBO	4.668	18.140	11.23	96.51%	21.91	[0, 255]
mdb212	Original	0.954	5.267	0	100%	--	[0, 225]
	TUBO	9.1	33.89	2.99	93%	23.785	[0, 255]

* TUBO : Technique Usuelle Basée sur l'Optimisation citées par plusieurs références [136], [137], [161]-[164], Les résultats du tableau 3.3 montrent que la technique basée sur l'optimisation augmente considérablement les valeurs des deux paramètres EC et EME pour toutes les images de test en comparaison avec les techniques basées sur la modification de l'histogramme. Cette augmentation est interprétée comme étant une sur-amélioration des images résultantes. En effet cette sur-amélioration est justifiée par la valeur élevée du paramètre AMBE ainsi que la modification de la dynamique des images résultantes de façon à occuper la totalité de l'intervalle des niveaux de gris [0, 255] qui sont signe de la brillance des images traitées. Cela confirme les résultats annoncés par les auteurs K. Sowjanya et al (2017) [136], et N. Mohanapriya et al (2019) [137] concernant la sur-amélioration des images traitées par ce modèle.

3.2.3. Présentation et analyse des images traitées

La figure 3.2 met en évidence la qualité des images mammographiques de test mdb209, mdb210, mdb211 et mdb212 obtenues par la technique basée sur l'optimisation. L'inspection visuelle de la figure 3.2 (b), montre que les images traitées par cette technique sont très brillantes avec apparence des détails qui peuvent ne pas exister dans l'image origine. Ceci est dû à ce que la valeur du paramètre AMBE est très élevée. Ainsi la méthode introduit des nouveaux niveaux de gris inexistant dans l'image origine suite au changement de la dynamique de l'image origine qui a occupé l'intervalle entier [0, 255]. En outre, une minutieuse analyse des images traitées montre l'existence d'un bruit qui n'est pas bien apparent dans les images origines. Ceci montre que malgré la valeur du PSNR relativement élevée pour cette technique par rapport aux techniques basées sur la modification de l'histogramme, elle reste insuffisante afin de produire des images bien améliorées.

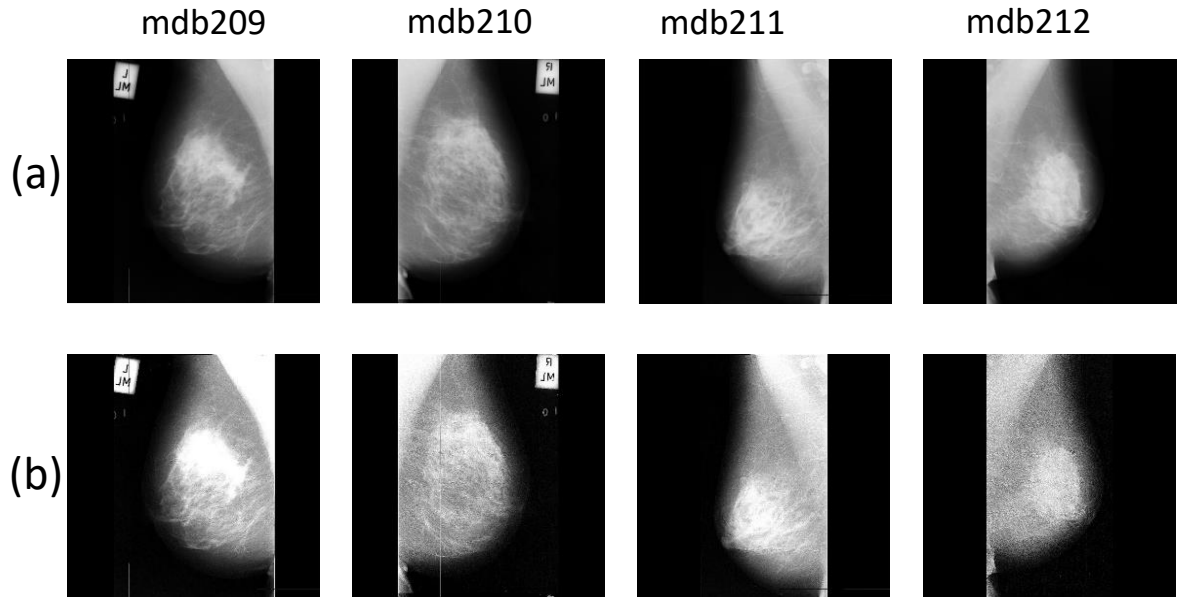


Figure 3. 2. Les images test améliorées par le modèle à optimisation usuel.

En conclusion nous avons bien vérifié que cette technique bien qu'elle donne des résultats satisfaisants à l'échelle du contraste, elle conduit à une sur-amélioration des images traitées. Ces résultats sont en concordance avec les résultats de certains auteurs [136],[137].

3.3. Comparaison des deux techniques

D'après les tableaux 3.1 et 3.3 on constate que :

- La technique usuelle basée sur l'optimisation améliore les paramètres EC et EME jusqu'à la sur-amélioration des images traitées. Cette technique ne parvient pas à conserver la luminosité de l'image origine ; au contraire sa capacité à conserver les informations de l'image d'entrée est bonne.
- Les techniques basées sur l'histogramme ont des effets variés sur le paramètre EC tandis qu'elles ont tous détérioré le paramètre EME. Ces techniques ne parviennent pas à conserver ni la luminosité ni les informations de l'image origine. De plus les techniques à histogramme amplifient plus le bruit que la technique usuelle basée sur l'optimisation.

En résumé la qualité des images traitées par la technique usuelle basée sur l'optimisation est meilleure que la qualité des images traitées par les techniques à histogramme. Quoique le modèle usuel à optimisation donne des images sur-améliorées dans certains cas.

4. Modèles d'amélioration proposés à optimisation mono- et multiobjectif

Afin de combler les lacunes présentées par les techniques usuelles d'une part, et la technique de Gandhamal (sur laquelle s'est basée notre modèle) d'autre part, nous avons procédé à la conception d'un modèle permettant une amélioration optimale de la qualité des images

mammographiques. C'est-à-dire un modèle permettant une bonne amélioration du contraste et un rehaussement adéquat des bords contenus dans l'image, tout en conservant au mieux la luminosité de l'image origine ainsi que les informations qu'elle renferme. Il s'agit alors d'un problème d'optimisation se traduisant par la maximisation simultanée des paramètres "IQA" (EME, EC, FSIM et AMBEn) (AMBEn = AMBE normalisé). Cela peut être résolu en adoptant soit une approche d'optimisation mono-objectif soit une approche d'optimisation multiobjectif. Notre proposition, consiste alors à exploiter les propriétés avantageuses de la fonction sigmoïde du fait que la fonction est composée de deux zones (avant et après le point d'inflexion) ayant des propriétés opposées (figure 3.3 (a)). Dans la 1^{ère} zone la fonction sigmoïde possède les propriétés des fonctions permettant l'assombrissement d'une image tandis que dans la 2^{ème} zone possède les propriétés des transformations permettant l'éclaircissement d'une image. L'application locale de cette propriété, (c'est à dire dans une zone restreinte de l'image) permet de transformer l'intervalle large des faibles niveaux de gris en un intervalle étroit (effet de la partie inférieure de la transformation). Simultanément elle permet de transformer l'intervalle large des niveaux de gris élevés en un intervalle étroit (effet de la partie supérieure de la transformation). Ainsi un éloignement (gap) entre les deux intervalles est observé. Cette différence de densité optique engendrée a pour effet d'augmenter la capacité de la discrimination entre les petits détails de l'image d'une part, et l'apparition des bords entre les différentes zones de l'image d'autre part.

En outre, comme la transformation sigmoïde considérée est définie de l'intervalle $[0, 1]$ vers $[0, 1]$, cette transformation conservera alors la dynamique de l'image origine. Ce qui a pour effet d'éviter toute sur-amélioration des images traitées.

Afin d'augmenter l'adaptabilité de la transformation sigmoïde à la particularité de chaque mammogramme, nous avons reformulé l'expression générale de la fonction sigmoïde traduite par l'équation 2.9 (chapitre 2) en lui attribuant une base variable. l'expression de la nouvelle fonction sigmoïde proposée devient [165], [166]:

$$s = \frac{1}{1 + B^{-\frac{r-\alpha}{\beta}}} \quad (3.13)$$

s étant l'intensité normalisée du pixel de l'image améliorée qui correspond à l'intensité normalisée r du même pixel de l'image origine. B , α et β sont respectivement la base $B \in [1, 10]$, le centre et la largeur de la transformation sigmoïde. L'attribution de la base variable à la transformation sigmoïde a pour objectif de donner à la technique un autre degré de liberté durant le processus de recherche de la fonction convenable pour chaque image.

Le principe du modèle consiste à subdiviser l'image en un ensemble de blocs sur lesquels la transformation par une fonction sigmoïde convenable est appliquée suivant l'algorithme 3.2. La combinaison du nombre de blocs et la fonction convenable permettent une meilleure amélioration pour chaque image. Les paramètres de décision de la technique à optimiser sont alors le nombre de blocs (dans les deux sens horizontal et vertical NBH et NBV) et les paramètres de la fonction sigmoïde B , α et β . La figure 3.4 montre l'effet des paramètres de la fonction sigmoïde sur la courbe de la transformation et par conséquent sur l'image traitée.

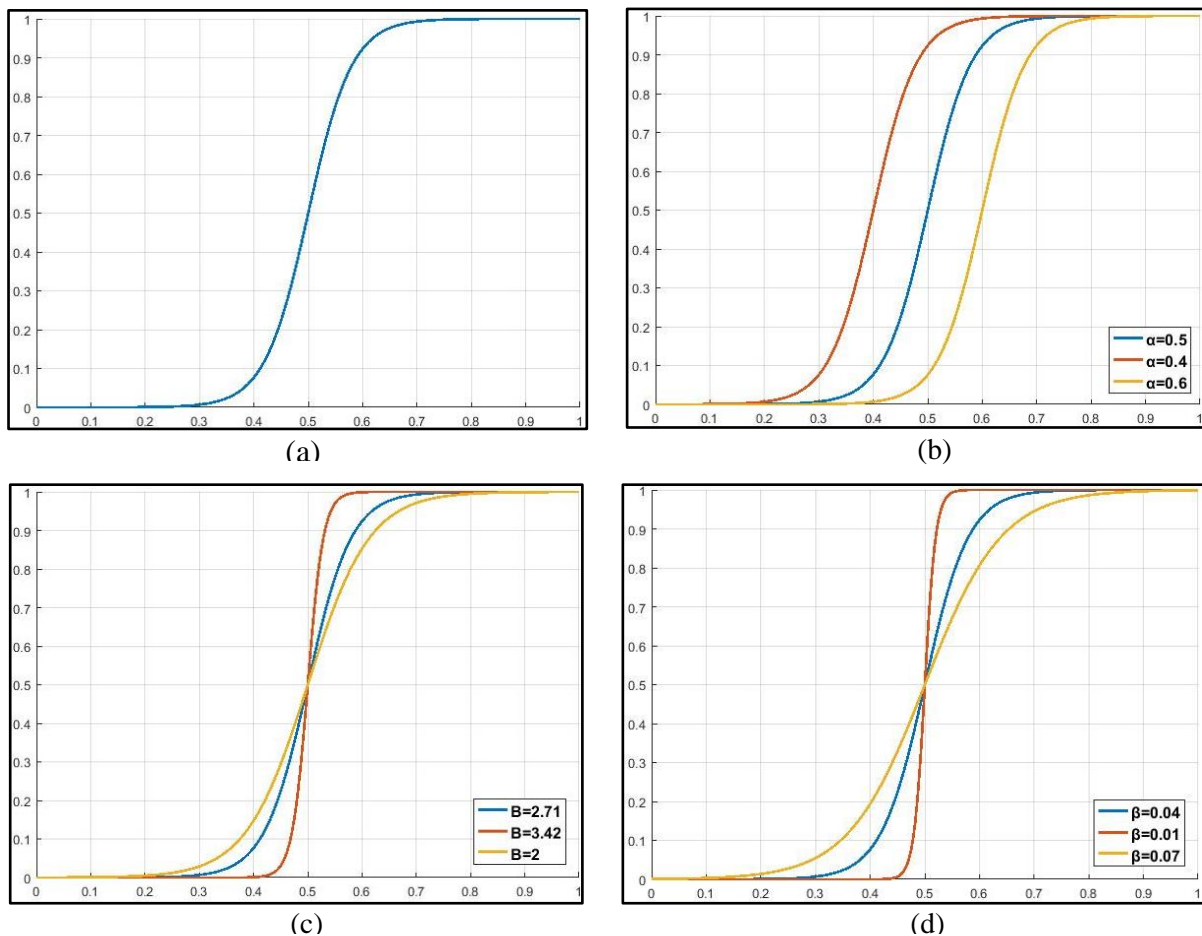


Figure 3. 3. Effets de modification des paramètres de la fonction sigmoïde. (a) fonction sigmoïde usuelle ; (b) modification du centre α ; (c) modification de la base B ; (d) modification de la largeur (β).

D'après les figures ci-dessus on en déduit :

- le changement du centre α de la transformation (figure 3.3(b)) permet de favoriser soit l'assombrissement d'une partie de l'image au détriment de l'éclaircissement soit le contraire. Ainsi, lorsque $\alpha > 0.5$ l'intervalle des niveaux de gris à assombrir est plus large que celui des niveaux de gris à éclaircir. Au contraire, lorsque $\alpha < 0.5$, l'éclaircissement des parties de l'image est favorisé par rapport à leurs assombrissements (un intervalle large des niveaux de gris est à éclaircir au détriment d'un intervalle étroit des niveaux de gris à assombrir).

- le changement de la base B ainsi que celui du paramètre β ont un effet sur le degré de la compression pratiquée sur les niveaux de gris de l'image traitée. Les figures 3.3 (c) et (d) montrent que la compression pratiquée aux intervalles des niveaux de gris de l'image augmente lorsque la base B de la sigmoïde augmente ou si sa largeur β diminue et vice versa.

L'algorithme suivant montre la transformation locale par la transformation sigmoïde.

Algorithme 3.2. Algorithme de traitement d'une image par la transformation locale basée sur la fonction sigmoïde proposée de paramètre B , α et β

Algorithme 3.2 : Transformation locale par la fonction sigmoïde

Données : Image, $NBH, NBV, B, \alpha, \beta$

- 1- Subdiviser l'image en $NBH \times NBV$ blocs
- 2- **Pour** chaque bloc faire
 - Normaliser les niveaux de gris dans l'intervalle $[0, 1]$ (Eq 2.11)

$$r = \frac{I - I_{min}}{I_{max} - I_{min}}$$
 - Calculer la nouvelle valeur normalisée par la transformation (Eq 3.13)

$$s = \frac{1}{1 + B \frac{r - \alpha}{\beta}}$$
 - Dénormaliser le niveau de gris s dans l'intervalle $[I_{min}, I_{max}]$ (Eq 2.12)

$$I_s = s \cdot (I_{max} - I_{min}) + I_{min}$$

Fin pour

Résultats

- 3- Afficher l'image améliorée
-
-

4.1. Modèle proposé à optimisation mono-objectif.

Notre but consiste à concevoir un modèle permettant l'amélioration optimale de la qualité des images mammographiques à optimisation mono-objectif vu sa simplicité.

Le modèle proposé s'est inspiré du modèle de la technique usuelle basée sur l'optimisation "TUBO" traitée précédemment. En effet notre modèle utilise la fonction sigmoïde que nous avons établie (équation 3.13) comme transformation locale et l'équation (3.12) du modèle "TUBO" comme fonction objectif d'évaluation des images au cours du processus de l'optimisation.

Le problème se ramène alors à la recherche du vecteur de décision (ensemble des paramètres de la transformation locale) $NBH, NBV, B, \alpha, \beta$ permettant de maximiser la fonction objectif.

Le processus de la recherche de la solution optimale est géré par la méthode des algorithmes génétiques. La boîte à outils de l'optimisation de *MATLAB R2015a* est exploitée à cet effet.

Le processus de recherche de la solution optimale se fait suivant l'algorithme 3.3 ci-après :

Algorithme 3.3 : Amélioration de la qualité des mammogrammes par le modèle mono-objectif proposé

Données : Image mammographique

Vers l'algorithme génétique

- Générer la population initiale 50 chromosomes ($NBH, NBV, B, \alpha, \beta$)
- Améliorer l'image **Image** par chaque chromosome (Transformation 3.13)
- Calculer la fitness de chaque individu fonction objectif (Equation 3.12)

Répéter

- Sélection des individus (stochastique uniforme)
- Croisement (point unique)
- Mutation (adaptive possible)
- Améliorer l'image **Image** par chaque chromosome (Transformation 3.13)
- Calculer la fitness de chaque individu fonction objectif (Equation 3.12)

Jusqu'à ce que la population converge ou critère d'arrêt est vrai (nombre d'itération =100)

- Calculer la matrice image correspondante au meilleur vecteur ($NBH, NBV, B, \alpha, \beta$)
 - Afficher l'image améliorée
-
-

Le modèle proposé est appliqué sur des images test de la base de données Mini-MIAS. Les résultats obtenus sont comparés avec ceux de la technique usuelle basée sur l'optimisation "TUBO".

4.1.1. Présentation et analyse des paramètres d'évaluation "IQA"

Le traitement des images test par l'application de l'algorithme 3.3 nous ont permis la détermination des vecteurs de décision récapitulés dans le tableau 3.4.

Tableau 3.4. Vecteurs de décision des images test traitées par notre modèle proposé à optimisation mono-objectif

Image test	NBH	NBV	B	α	β
mdb209	52	51	8.746	0.496	0.00013
mdb210	53	58	6.430	0.472	0.0033
mdb211	61	52	9.396	0.425	0.0015
mdb212	52	65	4.941	0.497	0.00081

D'après ce tableau on constate que chaque image mammographique est caractérisée par son vecteur de décision permettant son amélioration optimale. Ce qui confirme l'idée proposée qui consiste à traiter chaque image par sa transformation locale adéquate, à l'encontre de l'idée adoptée par Gandhamal et al. [8] où toutes les images mammographiques sont améliorées en utilisant les mêmes paramètres de la transformation locale.

Les valeurs des vecteurs de décision du tableau 3.4 nous ont permis d'aboutir aux résultats des paramètres d'évaluation des images test traitées par notre méthode basé sur l'optimisation mono-objectif.

Tableau 3. 5. Paramètres d'évaluation des images test traitées par notre méthode basé sur l'optimisation mono-objectif

Image		EC	EME	AMBE	FSIM	PSNR	Intervalle niveau de gris
mdb209	Origine	1.345	6.781	0	100%	--	[0, 228]
	Modèle mono-objectif	2.286	9.236	0.271	91.7%	29.80	[0, 228]
mdb210	Origine	1.326	6.355	0	100%	--	[0, 235]
	Modèle mono-objectif	2.323	9.539	0.546	92.4%	30.67	[0, 235]
mdb211	Origine	1.119	4.786	0	100%	--	[0, 227]
	Modèle mono-objectif	1.901	6.528	1.193	94.6%	31.63	[0, 227]
mdb212	Origine	0.954	5.267	0	100%	--	[0, 225]
	Modèle mono-objectif	1.716	7.556	0.198	95.1%	32.78	[0, 225]

Il est clair que notre modèle proposé basé sur l'optimisation mono-objectif a amélioré tous les paramètres "IQA" sans atteindre le degré de sur-amélioration. En effet, et comme le montre les résultats du tableau ci-dessus on en déduit :

- une nette amélioration du contraste des images traitées du fait que la moyenne du paramètre EC est passée d'environ 1,3 à une moyenne de 2,29 pour les images mdb209 et mdb210, et d'environ 1 à environ 1,8 pour les images mdb211 et mdb212.
- une meilleure contenance des bords des images traitées du fait que la valeur du paramètre EME est augmentée pratiquement de 5,795 à une moyenne de 8,214.
- les images traitées ont conservé la luminosité et les informations de l'image origine du fait que la méthode proposée est caractérisée par une faible valeur du paramètre AMBE (d'ordre 0,55 en moyenne) et une valeur relativement moins bonne du paramètre FSIM (93,5 %) pour toutes les images traitées.

De point de vue amplification du bruit, comme le montre les résultats du tableau 3.5, notre modèle parvient à ne pas amplifier le bruit du fait que la valeur du paramètre PSNR est élevé.

4.1.2. Présentation et analyse des images traitées

La figure 3.4 illustre les résultats de traitement des images de test en utilisant le modèle mono-objectif proposé. Les images de la figure 3.4 (a) représentent les images originales à améliorer, celles de la figure 3.4 (b) sont les images traitées en utilisant le modèle usuel et les images de la figure 3.4 (c) présentent les images traitées par notre modèle mono-objectif proposé. D'après ces figures on constate que les images produites par notre modèle sont plus proéminentes,

fidèles aux images origines avec des détails plus clairs. Ceci est dû principalement à l'amélioration adéquate du contraste et de contenance en bord avec une conservation de luminosité et de la dynamique de l'image origine (résultats quantitatifs).

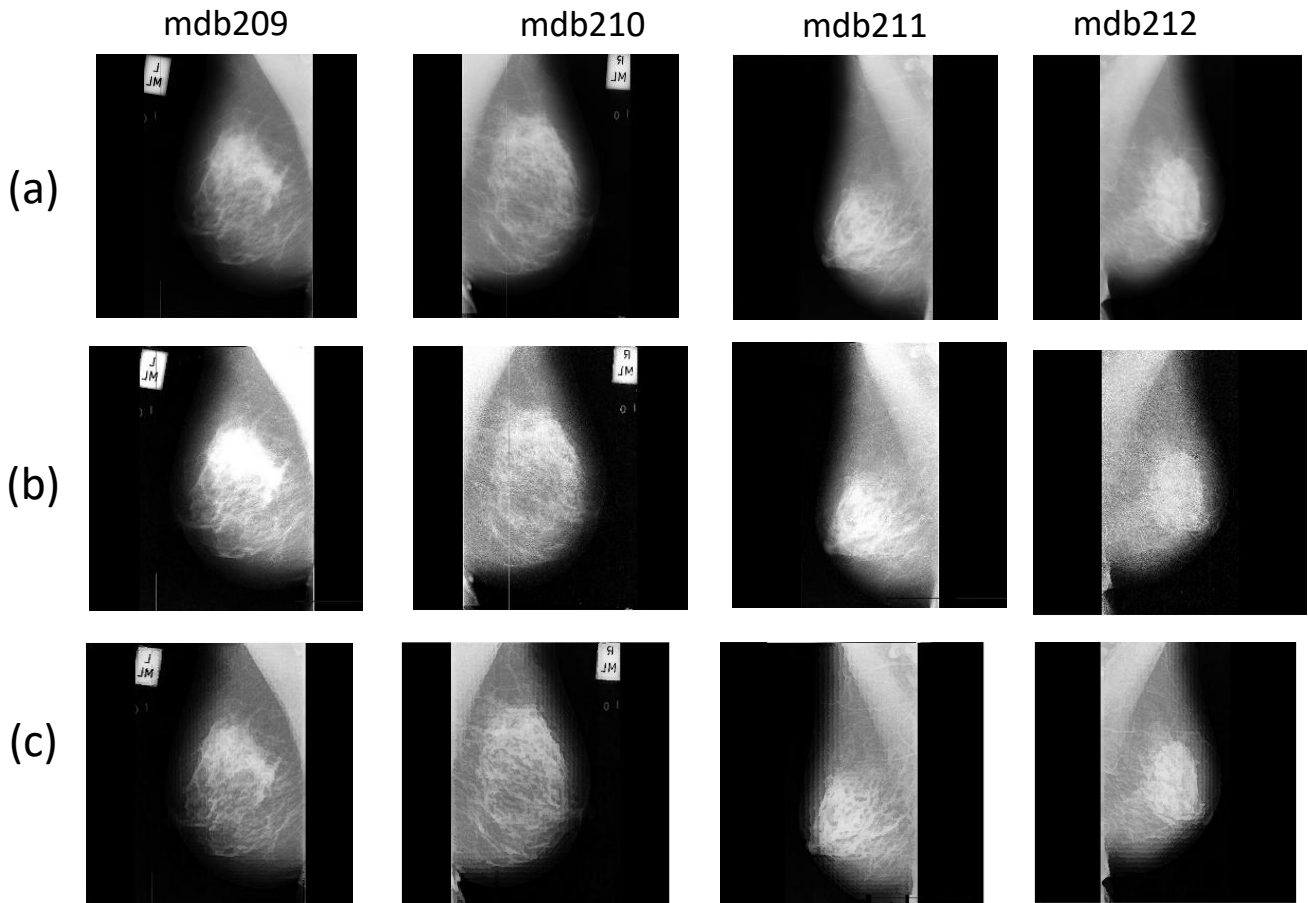


Figure 3. 4. Images test améliorées par les modèles à optimisation mono-objectif :
 (a) Image origine, (b) par le modèle à optimisation usuel (c) par notre modèle proposé

4.1.3. Comparaison des résultats des deux modèles non-objectif proposé et "TUBO"

D'après les résultats des tableaux 3.3 et 3.5 notre modèle proposé améliore les deux paramètres EC et EME sans atteindre la sur-amélioration vu la faible valeur du paramètre AMBE et la conservation de la dynamique de l'image origine. Aussi, notre modèle amplifie moins le bruit contrairement à la technique usuelle basée sur l'optimisation ("TUBO") qui engendre des sur-améliorations des images traitées. Ceci peut être mis en évidence par l'inspection visuelle des images traitées par les deux techniques. En effet les images engendrées par notre modèle (figure 3.4 (c)) sont fidèles aux images origines (figure 3.4 (a)) avec une clarté visuelle de leurs détails. Au contraire du modèle usuel qui produit des images brillantes dont beaucoup de détails sont altérés, avec apparition du bruit qui n'est pas bien apparent au niveau des images origines vu la valeur faible du paramètre PSNR (figure 3.4 (b)).

Bien que la méthode mono-objectif proposée soit capable d'améliorer tous les paramètres d'évaluation et qu'elle présente sa supériorité par rapport aux techniques usuelles (HE, BBHE et CLAHE) et au modèle d'optimisation usuel, certains travaux récents ont montré des résultats intéressants dont les valeurs de certains paramètres dépassent les nôtres. En effet :

- La technique proposée par A. Gandhamal et al. dans la référence [8] est caractérisée par une meilleure conservation des informations de l'image. La moyenne du paramètre FSIM pour toute la base de données Mini MIAS a dépassé 98%.

- La technique proposée par E. Daniel et al. dans la référence [69] a pu conserver davantage la luminosité des mêmes images après leurs améliorations. La moyenne du paramètre AMBE est de l'ordre de 0.159.

Chose qui nous a poussé à améliorer davantage la solution mono-objectif proposée. Ainsi après étude réalisée il s'avérait possible de déterminer une nouvelle fonction objectif et à résoudre le problème de l'amélioration de la qualité des images mammographiques en tant qu'un problème d'optimisation multiobjectif.

4.2. Modèle proposé à optimisation multiobjectif

4.2.1. Formulation mathématique du problème

La modélisation multiobjectif est une alternative permettant une modélisation réelle d'un problème donné. Dans notre cas une meilleure amélioration de la qualité des images mammographiques est caractérisée respectivement par un bon contraste, une meilleure apparence de bords et une bonne conservation de luminosité et d'information de l'image origine. Ceci se traduit par la maximisation simultanée de toutes les composantes du vecteur objectif $(EME, EC, AMBE_n, FSIM)$.

Notre problème de l'amélioration de la qualité des images mammographiques se ramène alors à déterminer le vecteur de décision $X = (NBH, NBV, B, \alpha, \beta)$ de la transformation locale donnée par l'équation 3.13 permettant de maximiser la fonction objectif déterminée comme suit [166] :

$$\begin{cases} f: X \rightarrow F \\ f(X) \rightarrow (EC(I_m(X)), EME(I_m(X)), AMBE_n(I_m(X)), FSIM(I_m(X))) \end{cases} \quad (3.14)$$

Où $(I_m(X))$ est l'image traitée par la transformation locale donnée par l'équation 3.13 en utilisant le vecteur de décision X .

A nouveau pour résoudre ce problème, nous avons fait appel à la méthode des algorithmes génétiques. La boîte à outils de l'optimisation multiobjectif de *MATLAB R2015a* est exploitée afin de déterminer les vecteurs de décision $NBH, NBV, B, \alpha, \beta$ permettant de maximiser simultanément tous les objectifs de la fonction objectif 3.14. Le pseudo algorithme de la méthode est :

Algorithme 3. 4. Algorithme d'amélioration de la qualité des mammogrammes par optimisation multiobjectif

Algorithme 3.4 : Algorithme d'optimisation

Données : Image mammographique

Vers l'algorithme génétique (Multi-objectif)

Nombre maximum de génération max_gen

Générer la population initiale (50 chromosomes($X = (NBH, NBV, B, \alpha, \beta)$))

Calculer les images $I_m(X)$

Calculer les vecteurs objectifs des individus

($EC(I_m(X)), EME(I_m(X)), AMBE_n(I_m(X)), FSIM(I_m(X))$)

Répéter

Répéter

Identifier les individus non dominés

Attribuer une fitness pour chaque individu

Mettre à jour le front de Pareto

Jusqu'à la classification de la population

Sélection (Tournoi)

Croisement (point unique)

Mutation (adaptive possible)

Génération=Génération +1

Jusqu'à (Génération \geq max_gen ou pas de changement du front)

Résultat : front de Pareto

Il est à noter que la solution d'un problème modélisé en tant que problème d'optimisation multiobjectif, n'est plus une solution unique vu le caractère contradictoire qui peut exister entre ses objectifs. Cependant, dans le but d'automatiser le choix d'une solution adéquate sur le Front de Pareto, divers techniques peuvent être utilisées. Ces dernières permettent de transformer un problème multiobjectif en son équivalent mono-objectif. Cependant une étude détaillée du comportement des objectifs du problème est nécessaire afin de faire cette transformation sans crainte de perdre les caractéristiques du problème [120].

4.2.2. Transformation du problème d'optimisation multiobjectif en mono-objectif.

La multiplicité des solutions optimales sur le front de Pareto et la difficulté voire l'impossibilité de leur attribuer un ordre, rend difficile le choix d'une solution adéquate au problème. Cependant, dans le cadre d'une approche simplifiée, il est possible de rechercher une solution unique en transformant le problème d'optimisation multiobjectif en un problème d'optimisation mono-objectif. A cet égard, divers techniques peuvent être utilisées afin d'effectuer cette transformation. Dans notre cas, nous avons choisi la méthode de la somme pondérée des objectifs contradictoires afin d'effectuer cette transformation. Néanmoins, une étude des relations globales du conflit, d'harmonie ou d'indépendance d'une part, et la compréhension des relations locales et composites ainsi que la signifiante des objectifs du problème d'autre part, sont nécessaires pour convertir le problème multiobjectif en un problème mono-objectif, sans crainte de perdre certaines de ses caractéristiques [107], [167].

Notre étude a été effectuée d'abord sur les quatre images de test mdb209, mdb210, mdb211 et mdb212 puis généralisée sur d'autres images de la base de données Mini-MIAS.

L'application de l'algorithme 3.4 aux images de test a abouti à un ensemble de solutions non dominées appelées aussi front de Pareto pour chaque image. C'est à base de ces solutions que notre étude pourra être accomplie.

A. Front de Pareto des images de test

Le front de Pareto correspond à un ensemble de solutions dont les états de tous les objectifs ne correspondent pas à leurs états optimaux, et que toute tentative d'améliorer l'un des objectifs entraîne la détérioration des autres. L'algorithme 3.4 est appliqué aux images test afin de déterminer les solutions optimales du problème. Les tableaux 3.6, 3.7, 3.8 et 3.9 constituent les solutions non dominées ou front de Pareto des images test.

Tableau 3. 6. Solutions non dominées de l'image mdb209

EME	EC	AMBE _n	FSIM
8.132	2.104	0.999	0.970
9.881	2.237	0.678	0.941
9.814	2.249	0.726	0.946
7.505	1.660	0.815	0.994
6.335	1.396	0.890	0.998
7.119	1.606	0.949	0.994
9.122	2.083	0.609	0.975
6.207	1.361	0.959	0.998
6.560	1.451	0.802	0.998
7.780	1.763	0.664	0.990
6.955	1.557	0.823	0.996
8.873	1.950	0.494	0.978
8.299	1.923	0.561	0.986
8.437	1.789	0.353	0.981
8.401	2.113	0.809	0.970
6.667	1.425	0.626	0.998
9.350	2.104	0.450	0.968
9.735	2.205	0.449	0.948
9.536	2.024	0.424	0.953
7.076	1.506	0.583	0.996

Tableau 3. 7. Solutions non dominées de l'image mdb210

EME	EC	AMBE _n	FSIM
6.328	1.377	0.895	0.999
6.340	1.449	1.000	0.998
9.221	2.154	0.311	0.943
9.459	2.322	0.764	0.932
6.700	1.574	0.851	0.996
9.463	2.320	0.699	0.932
8.430	2.040	0.391	0.962
8.977	2.072	0.349	0.953
7.281	1.782	0.560	0.988
6.437	1.485	0.963	0.997
7.963	2.099	0.779	0.968
7.885	1.865	0.604	0.982
7.445	1.705	0.926	0.992
6.858	1.643	0.989	0.993
8.591	1.991	0.451	0.968
8.064	1.904	0.722	0.978
8.846	2.204	0.743	0.962
7.588	1.761	0.828	0.990
6.993	1.660	0.626	0.994
9.015	2.132	0.331	0.955

Tableau 3. 8. Solutions non dominées de l'image mdb211

EME	EC	AMBE _n	FSIM
4.471	1.197	0.770	0.998
4.278	1.199	0.999	0.998
6.546	1.900	0.746	0.958
6.499	1.859	0.953	0.963
4.833	1.205	0.464	0.996
6.925	1.872	0.801	0.956
5.267	1.309	0.373	0.993
5.488	1.432	0.431	0.990
5.712	1.421	0.313	0.985
4.664	1.351	0.567	0.994
6.131	1.774	0.915	0.981
6.784	1.884	0.770	0.956
4.195	1.173	0.991	0.998
5.839	1.653	0.667	0.984
6.322	1.793	0.591	0.977
5.109	1.321	0.554	0.995
5.996	1.696	0.943	0.976
6.545	1.910	0.874	0.969

Tableau 3. 9. Solutions non dominées de l'image mdb212

EME	EC	AMBE _n	FSIM
4.809	1.004	1.000	0.998
7.736	1.677	0.699	0.964
7.697	1.691	0.748	0.964
4.800	1.021	0.980	0.998
5.280	0.979	0.906	1.000
5.864	1.273	0.783	0.993
6.743	1.425	0.489	0.986
5.573	0.976	0.569	0.999
7.699	1.690	0.755	0.963
5.121	1.114	0.927	0.998
6.603	1.368	0.457	0.990
7.590	1.632	0.470	0.966
6.507	1.492	0.645	0.987
7.451	1.527	0.350	0.972
6.367	1.535	0.810	0.986
6.045	1.326	0.681	0.994
6.633	1.385	0.529	0.989
7.236	1.517	0.622	0.981
7.378	1.555	0.364	0.971
6.947	1.447	0.594	0.984

Les résultats des tableaux ci-dessus montrent que les fronts de Pareto correspondant aux images mdb209, mdb210 et mdb212 sont composés de vingt solutions possibles et dix-huit solutions possibles pour l'image mdb211. L'analyse des solutions de chaque image, montre qu'il n'existe aucune solution où tous les objectifs sont meilleurs ou tous les objectifs sont mauvais simultanément. D'où le caractère contradictoire global entre les objectifs du problème. La difficulté majeure de ce problème réside donc du choix d'une solution adéquate au problème. Il est nécessaire alors d'instaurer des compromis entre les objectifs selon leurs importances dans le problème afin de pouvoir déterminer la solution à retenir.

B. Etude des relations globales entre les objectifs

La nature des relations globales entre les objectifs est déterminée par le coefficient de Kendall k qui nous permet de comprendre les relations d'harmonie, d'indépendance ou du conflit qui relie chaque paire d'objectifs. A partir des solutions non dominées des images de test (Tableaux 3.6, 3.7, 3.8 et 3.9), les coefficients de corrélation de Kendall sont déterminés en utilisant l'équation 2.23 (Chapitre 2 §3). Le tableau 3.10 donne les valeurs de ce coefficient pour chaque paire d'objectifs.

Tableau 3. 10. Les coefficients de corrélation de Kendall pour chaque paire d'objectifs des images test

Paires d'objectifs	mdb209	mdb210	mdb211	mdb212	Moyenne de k
<i>EC</i> - <i>EME</i>	0.8211	0.8842	0.8431	0.7579	0.8211
<i>EC</i> - <i>AMBEn</i>	-0.2421	-0.4947	0.1111	-0.2211	-0.2421
<i>EC</i> - <i>FSIM</i>	-0.8947	-0.9178	-0.8431	-0.9014	-0.8947
<i>EME</i> - <i>AMBEn</i>	-0.4211	-0.5474	0.0327	-0.4211	-0.4211
<i>EME</i> - <i>FSIM</i>	-0.8842	-0.9284	-0.8954	-0.8374	-0.8842
<i>AMBEn</i> - <i>FSIM</i>	0.3053	0.5464	-0.1111	0.2614	0.2505

Les résultats mentionnés dans ce tableau montrent :

- Que les objectifs EC et EME sont fortement en harmonie. En effet, la valeur du coefficient de corrélation de Kendall dépasse 0,5 pour toutes les images testées, sa moyenne est de 0,821.
- Que les objectifs AMBEn et FSIM sont en harmonie. Les valeurs du coefficient de corrélation de Kendall sont positives pour les images de test mdb209, mdb210 et mdb212, et une valeur pratiquement nulle pour l'image mdb211. La valeur moyenne de ce coefficient la AMBEn-FSIM est de l'ordre 0,251.
- En revanche, l'objectif FSIM est fortement en conflit avec chacun des objectifs EC et EME ; la valeur du coefficient de corrélation de Kendall est inférieure à -0.88 dans chaque cas.
- Le reste des combinaisons des objectifs, EC, AMBEn et EME, AMBEn sont en conflit, la moyenne des coefficients de corrélation de Kendall sont respectivement de -0,242 et -0,421.

C. Etude des relations locales et composites entre les objectifs

Le coefficient de Kendall k ne peut pas nous renseigner sur l'existence ou non d'une surface de compromis entre les objectifs contradictoires bien qu'il permette de déterminer la nature des relations globales qui existe entre deux objectifs [120].

L'étude des relations locales et composites entre les objectifs du problème permet de localiser les surfaces de compromis dans le cas où elles existent entre les objectifs. Pour mener cette étude, la détermination d'un seuil pour distinguer les bonnes solutions et les moins bonnes est nécessaire. La détermination des seuils peut se faire à partir des connaissances du domaine du problème, résultat de littérature ou d'une manière empirique (moyenne d'un objectif sur toutes les solutions non dominées par exemple) [120].

- Le paramètre AMBE : pour ce paramètre le meilleur seuil dont on dispose est le résultat publié par E. Daniel et al [69] qui est de l'ordre de 0.159. Le seuil qu'on a fixé pour ce paramètre est 0.10.

- Pour les objectifs EC, EME et FSIM ; pour chaque image et pour chaque objectif, le seuil est fixé à la moyenne des valeurs de chaque objectif, calculée à partir des solutions non dominées données par les tableaux 3.6, 3.7, 3.8, 3.9.

Après la détermination des seuils des objectifs, pour chaque solution parmi les solutions des données par les quatre tableaux des solutions non dominées, on peut déterminer l'état de chaque objectif (bon ou mauvais). Ce qui permet de faire la répartition des solutions dans la carte de régions pour chaque image selon la carte décrite au chapitre 2 (figure 2.14 § 7).

La figure 3.5 montre la répartition des solutions dans la carte des régions de toutes les images. Chaque case représente un état des objectifs. A titre d'exemple la première case en haut à gauche, correspond à l'état où tous les objectifs ont de bonne valeur, aucune des solutions ne vérifie cet état pour toutes les images de test.

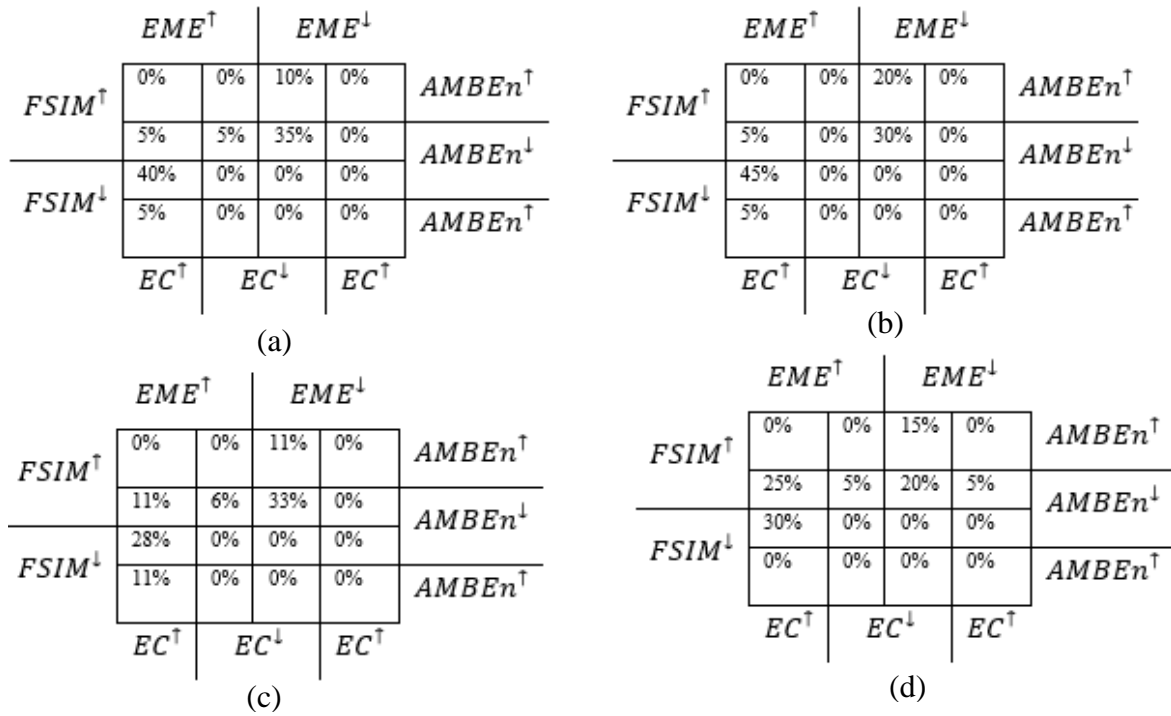


Figure 3. 5. Répartition des solutions dans les cartes de régions des images d'étude; (a) mdb209; (b) mdb210; (c) mdb211; (d) mdb212.

L'analyse de la distribution des solutions sur les cartes de régions de chaque image montre que la répartition des solutions est similaire pour les quatre images. Comme le montre la figure 3.5 l'ensemble des solutions pour toutes les images sont réparties sur les mêmes régions de la carte de régions. Aussi il est à noter que les solutions ne sont pas réparties équitablement sur toutes les régions, ainsi que la majorité de régions sont vides (au moins 10/16). Ceci prouve que les objectifs ne sont pas indépendants et qu'on ne peut pas trouver des solutions pour toutes les combinaisons possibles selon l'état (bon ou mauvais) des objectifs [42]. Les relations locales et composites entre les objectifs sont définies comme suit :

Cas de quatre objectifs : Selon les quatre cartes de répartition des solutions de la figure 3.5, les régions r_0 et r_{15} (voir la carte de région chapitre 2, figure 2.14 § 7) sont vides pour toutes les images. Par conséquent, il n'y a pas de solutions où toutes les valeurs des objectifs sont bonnes ou mauvaises. Ce résultat prouve que les objectifs sont globalement contradictoires. Donc aucune surface de compromis pour laquelle les quatre objectifs ont de bonnes valeurs. Par conséquent, la résolution de ce problème nécessite d'établir un compromis entre les objectifs.

Cas de trois objectifs : Les cartes de répartition des solutions montrent :

- que les solutions dont trois objectifs ont une bonne valeur n'existent que lorsque les deux objectifs EC et EME sont combinées soit avec AMBEn ou avec FSIM.

- Pour la combinaison EC, EME et FSIM les solutions existent pour toutes les images testées dans la région 4 de la Figure 3.5, ce qui montre l'existence d'une surface de compromis locale pour ce triplet,
- Pour la combinaison EC, EME et AMBEn les objectifs ayant de bonne valeur ne sont disponibles que pour mdb209 et mdb211 en région r_8 Figure 3.5 (a) et (c), ce qui prouve que dans certains cas, une surface de compromis locale peut être trouvée entre ces trois objectifs. -
- Pour les autres combinaisons (EC, AMBEn, FSIM) et (EME, AMBEn, FSIM), les régions r_1 et r_2 de la figure 3.5 sont vides ce qui montre qu'aucune solution avec trois bonnes valeurs n'est envisagée pour ces combinaisons.

Cas de deux objectifs : D'après les cartes de répartition figure 3.5,

- les deux objectifs EC et EME ont au moins 50% de solutions avec de bonnes valeurs pour toutes les images de test en région r_4 , r_8 , r_{12} de la figure 3.5. Ce résultat confirme la forte relation d'harmonie reliant ces deux objectifs.
- Concernant les objectifs AMBEn et FSIM, plus de 10 % de bonnes solutions pour toutes les images de test existe dans la région r_3 de la carte, ce qui prouve à nouveau l'harmonie de ces objectifs.
- Pour l'objectif FSIM combiné soit avec EC soit avec EME, même si les objectifs sont deux à deux globalement contradictoires (résultats de l'étape 1), des solutions avec de bonnes valeurs pour ces objectifs existent pour toutes les images testées, ce qui montre l'existence d'un compromis local entre les paires d'objectifs EC, FSIM et EME, FSIM.
- Pour la combinaison de l'objectif AMBEn soit avec les objectifs EC ou EME, les cartes de répartition montrent l'existence de bonnes solutions pour mdb209 et mdb211 région r_8 figure 3.5, cependant, aucune bonne solution concernant les images mdb210 et mdb212. Par conséquent, la surface de compromis n'est pas toujours présente pour les combinaisons d'objectifs EC, AMBEn et EME, AMBEn. Ceci confirme le résultat du conflit global qui existe entre ces paires d'objectifs.

L'étude de ces relations locales permet de comprendre le comportement réel des objectifs sur tout l'espace objectif. En effet bien que deux objectifs soient globalement contradictoires on peut trouver une zone où ils présentent une surface de compromis. Ceci aidera le concepteur dans l'étape de la réduction des objectifs.

D. Etude de la signifiante des objectifs du problème

Dans un problème d'optimisation multiobjectif, un objectif est dit significatif si l'intervalle dont il évolue est plus large (un grand écart entre les valeurs maximales et minimales)[120]. Les intervalles d'évolution de chacun des objectifs du problème sont déduits des solutions non dominées présentées dans les tableaux 3.6, 3.7, 3.8 et 3.9. Le tableau 3.11 montre l'ensemble

de ces intervalles en précisant la valeur minimale et maximale de chaque objectif pour chaque image.

Tableau 3. 11. Intervalles d'évolutions des objectifs

		mdb209	mdb210	mdb211	Mdb212	Moyenne
EC	max	2.249	2.322	1.910	1.691	2.043
	min	1.362	1.377	1.174	0.976	1.222
	largeur	0.888	0.945	0.736	0.714	0.821
EME	max	9.882	9.463	6.926	7.736	8.502
	min	6.208	6.328	4.196	4.800	5.383
	largeur	3.674	3.135	2.730	2.936	3.119
AMBEn	max	1.000	1.000	1.000	1.000	1.000
	min	0.353	0.311	0.314	0.350	0.332
	largeur	0.647	0.689	0.686	0.650	0.668
FSIM	max	0.999	0.999	0.999	1.000	0.999
	min	0.942	0.932	0.956	0.963	0.948
	largeur	0.057	0.067	0.043	0.036	0.051

D'après le tableau 3.8, l'intervalle de l'objectif FSIM est trop étroit par rapport aux intervalles du reste des objectifs du problème pour toutes les images. La moyenne de cet intervalle est d'environ 0,05 tandis que les moyennes des autres objectifs sont de l'ordre de 0,821, 3,119 et 0,668 pour EC, EME et AMBEn respectivement. Par conséquent, les objectifs EC, EME et AMBEn peuvent être considérés comme objectifs significatifs du problème, contrairement à l'objectif FSIM qui est un objectif non significatif. Cette étude a pour objectif de ne garder que les objectifs significatifs. En effet, les valeurs d'un objectif non significatif sont très proches les unes des autres d'où la possibilité de les omettre sans crainte d'altérer les propriétés du problème.

E. Réduction des objectifs du problème

L'objectif de l'étude effectuée est de réduire le nombre des objectifs du problème multiobjectif sans crainte de modifier ou d'altérer ses caractéristiques. Dans notre cas, les objectifs EC et EME sont fortement en harmonie. Par conséquent l'optimisation de l'un entraîne l'optimisation de l'autre. Donc un objectif peut être réduit sans aucun problème. Le paramètre EME est choisi à être omis en raison de son calcul qui est local chose qui augmente son temps de calcul. De même, les deux objectifs AMBEn et FSIM sont en harmonie d'une part, d'autre part l'objectif FSIM est un objectif non significatif du problème vu son intervalle d'évolution qui est étroit ; aussi d'après l'étude des relations locales et composites effectuée une surface de compromis locale peut être trouvée entre l'objectif FSIM et les objectifs EC et EME. Par conséquent, le paramètre FSIM peut être aussi omis du problème[120].

Le problème d'optimisation multiobjectif à quatre objectifs EC, EME, AMBEn et FSIM est réduit alors à un problème multicritère à deux objectifs EC et AMBEn entre lesquels il faut concevoir un compromis afin de transformer le problème multiobjectif en son équivalent mono-objectif.

F. Fonction objectif et algorithme du modèle proposé

D'après l'étude menée, notre problème qui consistait à maximiser à la fois les quatre objectifs EC, EME, AMBEn et FSIM est réduit à un problème multicritère à deux objectifs qui sont EC et AMBEn, et qui conserve les caractéristiques de la première modélisation.

La fonction objectif du problème mono-objectif équivalent est représentée par une somme pondérée deux objectifs restant, son expression est donnée par :

$$F(X) = w_1 \cdot EC(I_m(X)) + w_2 \cdot AMBE_n(I_m(X)) \quad (3.15)$$

w_1 et w_2 sont des coefficients de pondération tels que $w_1 + w_2 = 1$, $X = (NBH, NBV, B, \alpha \text{ et } \beta)$ le vecteur de décision, $I_m(X)$ l'image traitée par la transformation locale en utilisant le vecteur de décision X .

Les mêmes images tests étudiées (mdb209, mdb210, mdb211 et mdb212) sont utilisées pour déterminer les coefficients de pondération w_1 et w_2 . L'étude consiste à attribuer aux coefficients de pondération (w_1, w_2) de l'équation (3.15) des valeurs possibles suivantes : $(0.9, 0.1), (0.8, 0.2) \dots, (0.1, 0.9)$ et de déduire la paire qui présente un bon compromis. Le tableau (3.12) présente les résultats de cette étude sous forme de moyenne des quatre objectifs pour toutes les images.

Tableau 3. 12. Moyennes des objectifs des images test pour différentes valeurs de w_1 et w_2

	EC	EME	AMBE	FSIM
Origine	1,186	5,797	0	1,000
(0.9, 0.1)	2,071	8,044	0,180	0,957
(0.8, 0.2)	2,060	8,075	0,128	0,956
(0.7, 0.3)	2,027	7,582	0,002	0,954
(0.6, 0.4)	2,052	7,900	0,103	0,955
(0.5, 0.5)	2,052	7,900	0,103	0,955
(0.4, 0.6)	2,048	7,864	0,084	0,955
(0.3, 0.7)	2,045	7,812	0,073	0,955
(0.2, 0.8)	2,049	7,869	0,091	0,955
(0.1, 0.9)	2,049	7,861	0,088	0,955

Comme le montre les résultats du tableau 3.12, la paire (0.8, 0.2) présente un bon compromis entre les quatre objectifs. La fonction objectif définitive du problème consiste à remplacer les coefficients de pondération w_1 et w_2 par leurs valeurs 0.8 et 0.2 respectivement. Par conséquent le pseudo algorithme d'amélioration du contraste des images mammographiques est comme suit :

Algorithme 3. 5. Algorithme du modèle proposé pour l'amélioration de la qualité des mammogrammes par optimisation multiobjectif

Algorithme 3.5 : Algorithme d'optimisation multiobjectif de la qualité des

Donnée : Image mammographique

Vers l'algorithme génétique mono-objectif

Générer la population initiale de 50 individus ($X = (NBH, NBV, B, \alpha, \beta)$)

Calculer les images $I_m(X)$ (3.13)

Calculer la fitness de chaque individu équation Eq (3.15) .

Répéter

Sélection (stochastique uniforme)

Croisement (point unique)

Mutation (adaptive possible)

Calculer les images $I_m(X)$

Calculer la fitness de chaque individu de la nouvelle génération (3.15) avec :

$$w_1 = 0.8, w_2 = 0.2$$

Jusqu'à ce que la population converge (ou critère d'arrêt vérifié)

Calculer L'image du meilleur vecteur($NBH, NBV, B, \alpha, \beta$)

Résultat : afficher l'image améliorée

4.2.3. Résultats de l'application du modèle à optimisation multiobjectif

Notre modèle basé sur l'optimisation multiobjectif a été appliqué d'abord aux images de test. Cette opération nous permettra de comparer sa performance avec le modèle basé sur l'optimisation mono-objectif de point de vue quantitatif et qualitatif ainsi qu'avec la technique de E. Daniel et al. Ensuite sera appliqué à toute la base de données Mini-MIAS afin de comparer ses résultats quantitatifs ceux obtenus par A. Gandhamal et al. en référence [8].

A. Résultats d'application du modèle proposé aux images test

- **Présentation et analyse des paramètres d'évaluation "IQA"**

L'application de l'algorithme 3.5 aux images de test a abouti à la détermination des vecteurs de décisions présentés dans le tableau 3.13. La transformation des images de test par la

transformation sigmoïde (Equation 3.13) en utilisant le vecteur de décision relatif à chaque image, suivi du calcul des paramètres "IQA" des images traitées permet d'obtenir les résultats quantitatifs du tableau 3.14 et les résultats qualitatifs de la figure 3.16.

Tableau 3. 13.Vecteurs de décision des images test par le model d'optimisation multiobjectif

	NBH	NBV	B	α	β
mdb209	76	83	9.5219	0.5064	0.0299
mdb210	103	85	9.8113	0.5065	0.0236
mdb211	52	86	8.3361	0.5062	0.0126
mdb212	98	84	6.6174	0.5048	0.0177

Tableau 3. 14. Paramètres d'évaluation "IQA" des images test par le modèle multiobjectif

image		EC	EME	AMBE	FSIM	PSNR	Intervalle niveau de gris
mdb209	Origine	1.345	6.781	0	100%	--	[0, 228]
	Modèle multiobjectif	2.280	9.280	0.170	95.10%	32.99	[0, 228]
mdb210	Origine	1.326	6.355	0	100%	--	[0, 235]
	Modèle multiobjectif	2.292	8.521	0.002	94.31%	33.18	[0, 235]
mdb211	Origine	1.119	4.786	0	100%	--	[0, 227]
	Modèle multiobjectif	1.941	7.712	0.125	96.5%	33.75	[0, 227]
mdb212	Original	0.954	5.267	0	100%	--	[0, 225]
	Modèle multiobjectif	1.727	7.712	0.215	96.43%	35.87	[0, 225]

D'après les résultats du tableau 3.14 on conclut que le modèle proposé à optimisation multiobjectif permet d'améliorer tous les paramètres "IQA" du problème. En effet, tous les paramètres "IQA" mis en œuvre dans cette étude ont des valeurs satisfaisantes. Comme le montre le tableau :

- Le contraste et l'apparence de bord sont convenablement améliorés, du fait que le modèle augmente les valeurs des paramètres EME et EC.
- La luminosité, le contenu informatif et la dynamique des images origines sont maintenus après l'amélioration. Ceci est traduit par la faible valeur du paramètre AMBE qui a une moyenne de l'ordre de 0.128 qui est inférieure à la valeur engendrée par la technique proposée par E. Daniel, et al. [69] prise comme meilleure valeur dans notre étude.
- La dynamique des images améliorées est invariante du fait que l'image améliorée conserve exactement l'intervalle du niveau de gris de l'image origine.
- Les informations de l'image origine sont maintenues d'après les valeurs élevées du paramètre FSIM qui ont dépassé les valeurs du modèle mono-objectif.
- L'amplification du bruit est faible du fait que la valeur du paramètre PSNR est la meilleure parmi toutes les techniques examinées dans ce travail.

- **Présentation et analyse des images traitées**

D'après la figure 3.6 on remarque que les images traitées par notre modèle multi-objectifs proposés figure 3.6 (c) sont bien contrastées et plus proéminentes avec des détails plus clairs, en plus de leurs fidélités aux images origines figure 3.6 (a). Cela est dû principalement à l'amélioration adéquate par lesdites méthodes du contraste et de contenance de bord avec une conservation de luminosité et de la dynamique de l'image origine.

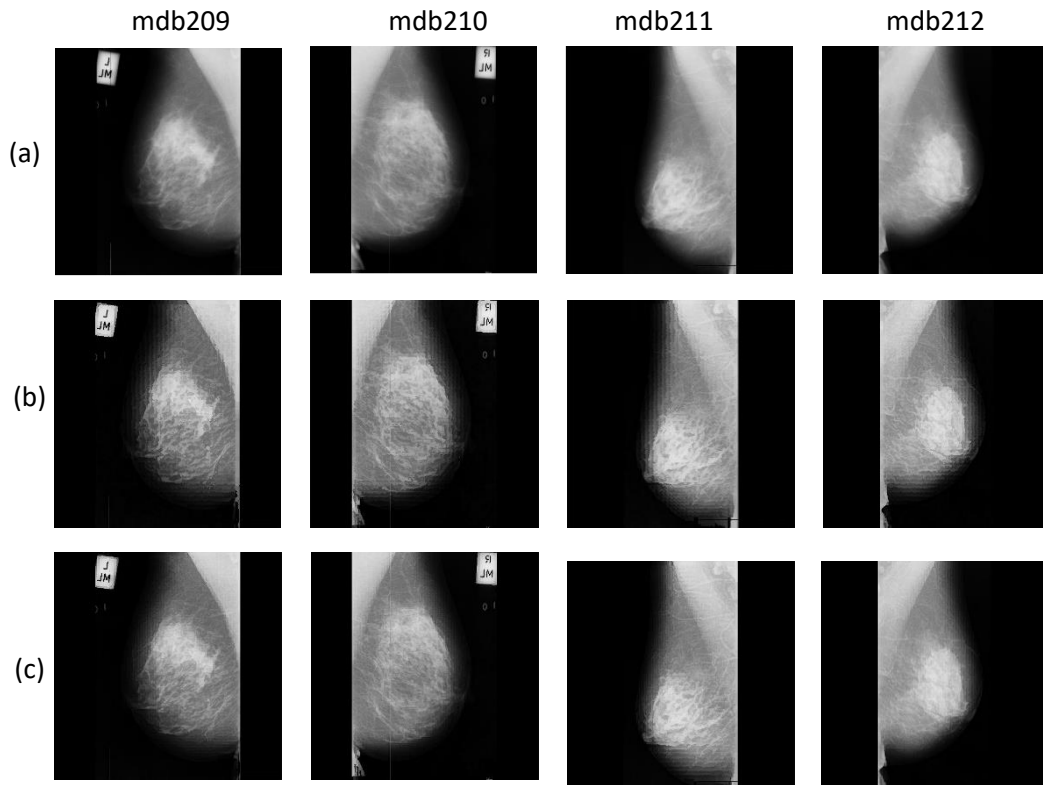


Figure 3. 6. Résultats d'amélioration des images test : (a) Image origine, (b) par le modèle proposé à optimisation mono-objectif (c) par le modèle proposé à optimisation multiobjectif

B. Résultats d'application du modèle proposé à toute la base de données Mini-MIAS

Afin de valider notre modèle d'optimisation de l'amélioration de la qualité des images mammographiques qui est basé sur la transformation sigmoïde et gérer par un algorithme d'optimisation multiobjectif, nous l'avons appliqué à toutes les images mammographiques de la base de données Mini-MIAS afin d'améliorer leur qualité. Les résultats obtenus sont comparés avec ceux publiés par A. Gandhamal et al. dans la référence [8]. Le tableau 3.12 récapitule les résultats de toutes les méthodes sous forme de moyenne de chaque paramètre "IQA" évalué dans cette étude pour toutes les images.

Tableau 3. 15. Moyennes des différents paramètres d'évaluation des images de la base de données Mini-MIAS par notre modèle à optimisation multiobjectif

	EC	EME	AMBE	FSIM
Images origines	1.222	5.523	0	100%
Gandhamal et al [19]	1.805	6.703	4.239	98.90%
Modèle proposé	2.114	7.530	0.106	95.20%

Comme on peut le constater à partir du tableau 3.15, le modèle proposé dépasse largement la technique de comparaison. En effet, notre modèle est caractérisé par une capacité d'améliorer à l'optimale le contraste et la contenance en bords des images évaluées tout en maintenant intacte leur luminosité.

- Le paramètre EME qui évalue l'amélioration en contraste est passé de 5.523 à 7.530 contre 6.703 pour la méthode de Gandhamal et al. en référence [19]

- Le paramètre EC permettant d'évaluer la contenance en bords a passé de 1.222 à 2.114 contre 1.805 pour la méthode de Gandhamal et al.

Le paramètre AMBE qui évalue la conservation de luminosité n'a pas dépassé 0.106 contre 4.239.

- Concernant le paramètre FSIM la méthode de Gandhamal et al dépasse notre modèle sa valeur étant 98.90% contre 95.20% pour la nôtre. Cette supériorité est justifiée par le caractère contradictoire des objectifs EME et EC d'une part et de FSIM d'autre part. En effet toute augmentation au niveau de EME et EC va entraîner une détérioration de FSIM. Cependant vu les valeurs des objectifs dans leurs globalités la valeur du paramètre FSIM reste une valeur acceptable.

4.2.4. Comparaison des résultats des deux modèles (mono et multiobjectif)

D'après les tableaux 3.5 et 3.14 on constate que le modèle basé sur l'optimisation multiobjectif présente un très bon compromis entre tous les objectifs du problème en le comparant avec le modèle monoobjectif proposé auparavant. Comme montrent les résultats des tableaux 3.5 et 3.14, les paramètres EC et EME ont des valeurs satisfaisantes en même temps que le reste des paramètres AMBE, FSIM et PSNR ont été améliorés en comparaison avec le modèle monoobjectif proposé qui présente à son tour sa supériorité par rapport aux techniques usuelles. Ainsi l'inspection visuelle des images traitées par les deux modèles figure 3.6 montrent leur fidélités aux images origines avec des détails clairs. Par ailleurs les images engendrées par le modèle basé sur l'approche d'optimisation multiobjectif figure 3.6 (c) contiennent moins d'artefacts que les images engendrées par le modèle monoobjectif. Par conséquent la qualité des images traitées par le modèle multiobjectif est meilleure que celle obtenue par le modèle monoobjectif et cela reste en concordance avec les résultats quantitatifs.

5. Conclusion

D'après les résultats auxquels nous nous sommes parvenus, on peut conclure que notre modèle proposé basé sur la transformation sigmoïde de l'équation (3.13) : $s = \frac{1}{1 + B^{-\frac{r-\alpha}{\beta}}}$, et géré par l'algorithme d'optimisation multiobjectif dont la fonction objectif est donnée par l'équation (3.15) : $F(X) = \alpha.EC(I_m(X)) + (1 - \alpha).AMBE_n(I_m(X))$, avec $\alpha = 0.8$ dans notre cas, présente un outil bien adapté à la spécificité des images mammographiques. Il peut être considéré comme un outil performant pour l'optimisation de l'amélioration de la qualité desdites images. Ce modèle sera exploité pour la conception d'un système automatisé permettant la détection précoce des masses mammaires (voir le chapitre suivant).

Conception d'un système automatisé d'aide à la détection "DAOe" des masses mammaires

1. Introduction

Ce chapitre constitue le vif de notre travail. Il s'agit de présenter et de décrire l'ensemble des techniques et méthodes mises en œuvre pour la conception d'un système d'aide automatisé à haute résolution permettant la détection et la localisation précoce des masses mammaires. L'ensemble des techniques et méthodes de prétraitement et de segmentation développées au 2^{ème} chapitre d'une part, et le modèle conçu pour l'amélioration optimale des images mammographiques développé au 3^{ème} chapitre d'autre part, seront à la base de la conception de notre système d'aide à la détection "DAOe". Ce dernier est composé de deux étapes principales : la première est l'étape de prétraitement qui vise à délimiter la région d'intérêt "ROI" de l'image origine et améliorer sa qualité ; la seconde étape c'est la segmentation qui consiste à segmenter la "ROI" améliorée afin d'isoler la zone suspecte (masse) du reste du mammogramme. Notre système d'aide à la détection "DAO" est appliqué à un ensemble d'images de la base de données Mini-MIAS dont les images mdb010, mdb028, mdb134, mdb184, mdb271 constituent un échantillon afin d'évaluer son efficacité. Les résultats obtenus confirment la précision de l'extraction de la zone suspecte à partir de la région d'intérêt.

2. Présentation du système d'aide à la détection "DAOe"

Comme cela a été déjà signalé un système d'aide au "DAO" que ce soit un système d'aide à la détection "DAOe" ou un système d'aide au diagnostic "DAOx", est un ensemble de programmes informatiques qui exploitent un ensemble de techniques de traitement d'images numériques et d'outils d'aide à la décision pour extraire les informations utiles du mammogramme et les analyser automatiquement. Cela permettra aux cliniciens de prendre la bonne décision suite à la seconde lecture de l'état pathologique du mammogramme.

Notre étude s'est contentée de l'étude de détection vu son intérêt grandissant. Notre but est de concevoir un système capable de détecter les zones suspectes (masses) automatiquement et à un stade prématuré. La figure 4.1 représente l'organigramme de ce système nommé "DAOe". Les différents étapes, leur succession ainsi que leurs processus permettant le fonctionnement automatisé du système sont présentées [167], [168]. Dans ce qui suit nous allons détailler chaque étape en présentant les différentes techniques et méthodes qui y sont incorporées.

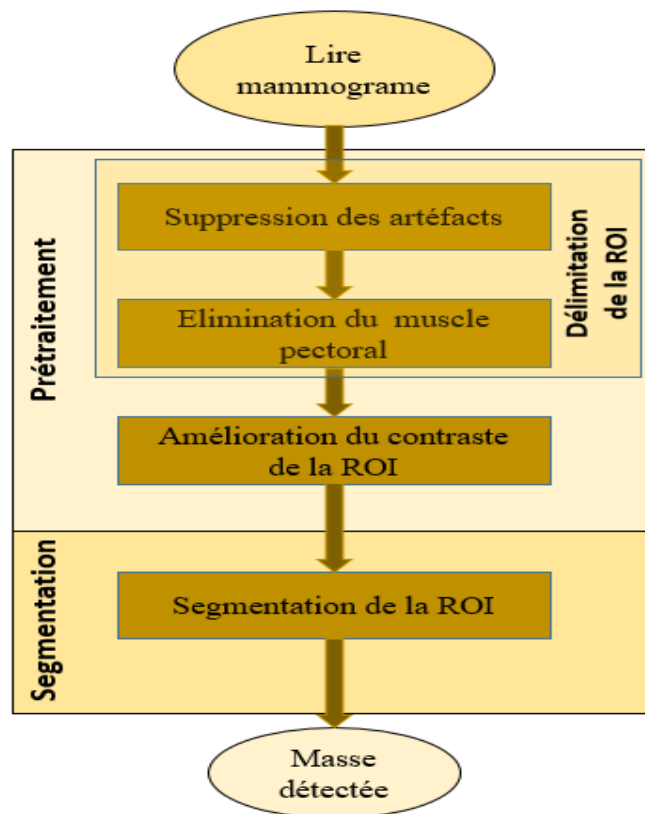


Figure 4. 1. Organigramme du système conçu d'aide à la détection "DAOe"

3. Etape de prétraitement de l'image origine

En plus du mauvais contraste et des bords flous dont souffrent les images mammographiques [65], [66] ; ces dernières renferment souvent divers artéfacts tels que les artéfacts radio-opaques de l'étiquetage (qui peuvent apparaître en haut des images mammographiques) et les artéfacts liés au bruit comme les montre la figure 4.2 [169]. Par ailleurs, les mammogrammes acquis sous l'incidence médio-latéral-oblique MLO contiennent en plus du profil du sein, le muscle pectoral qui apparaît à gauche ou à droite de la partie supérieure du mammogramme. Cela constitue aussi un artéfact ayant des effets néfastes sur l'analyse automatique de ce type d'images (figure 4.2) [6]. D'où la nécessité d'éliminer d'abord tous les artéfacts et d'obtenir ainsi la région d'intérêt "ROI". Cette région subit une amélioration de sa qualité avant qu'elle soit segmentée.

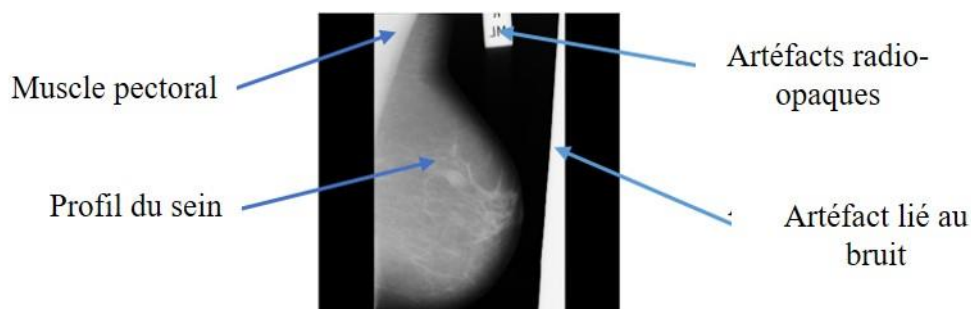


Figure 4. 2. Mammogramme mdb012 avec différents artéfacts pouvant interférer un mammogramme

3.1. Délimitation automatique de la région d'intérêt du mammogramme

3.1.1. Elimination des artefacts

Dans la plupart de cas, l'arrière-plan des images mammographiques comporte des artefacts radio-opaques tels que des étiquettes et des marqueurs. D'une manière générale, ils occupent des régions plus au moins petites en haut de l'image. Leur caractère clair constitue des informations très nuisibles aux systèmes d'aide au diagnostic. Leur prise en compte dans l'analyse automatique des images peut fausser les résultats de l'analyse en augmentant le nombre des faux positifs. Par conséquent, la segmentation et la suppression de tout détail inutile dans les mammogrammes à analyser est d'une très grande importance afin d'augmenter la précision du diagnostic [169]. En effet, le but de la segmentation des images médicales est d'automatiser la délimitation des structures anatomiques ou d'autres régions d'intérêt présentes dans l'image médicale [169].

A. Méthodes et techniques utilisées

La suppression automatique des artefacts consiste à utiliser des informations permettant de les distinguer et de les caractériser de l'objet principal (sein) afin de les isoler les uns des autres. Dans le cas des images mammographiques, le profil du sein correspond à l'objet le plus large de l'image. C'est en fait sur la base de cette information que la méthode utilisée est fondée.

D'abord l'image mammographique qui est une image en niveau de gris, est convertie en une image binaire en utilisant la technique de seuillage global. Sur le masque en noir et blanc ainsi formé (figure 4.3 (b)), l'ensemble des objets détectés sont étiquetés afin de les repérer. Ensuite l'objet le plus large qui correspond au profil du sein est conservé, le reste est attribué à l'arrière-plan figure 4.3 (c). Cette opération est suivie d'un traitement morphologique permettant de réduire les distorsions et les irrégularités engendrées par des pixels isolés en région cible (zone blanche) et en région fond (zone noire) figure 4.3 (c). Le principe est d'appliquer une correction pour chaque voisinage d'un pixel à huit connexes 3×3 ; si la valeur d'un pixel est différente de celle des cinq pixels au moins de son voisinage alors le pixel doit changer de valeur de 0 à 1 ou inversement.

Une autre correction est apportée au bord du profil du sein qui consiste à appliquer une érosion à l'aide d'un élément structurant de forme de disque de rayon 5 pixels, suivie d'une dilatation avec le même élément structurant. Cette opération a pour but d'éliminer les distorsions et les irrégularités au voisinage des bords du profil du sein. Le masque formé après ces opérations,

est appliqué à l'image origine afin d'obtenir l'image sans artéfacts figure 4.3 (d). Le pseudo-algorithme de l'étape permettant la suppression des artéfacts est donnée par l'algorithme 4.1.

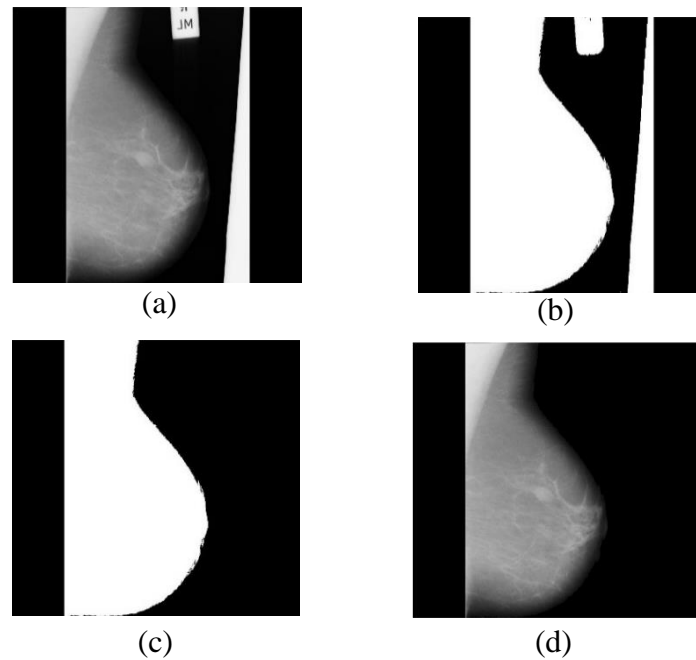


Figure 4. 3. Etapes d'isolation du profil du sein du reste de l'image (a) image origine mdb012, (b) image binaire, (c) objet plus large de l'image, (d) profil du sein.

Algorithme 4. 1. Algorithme d'élimination des artéfacts radio-opaques

Algorithme 4.1 : Algorithme pour l'élimination des artéfacts d'un mammogramme.

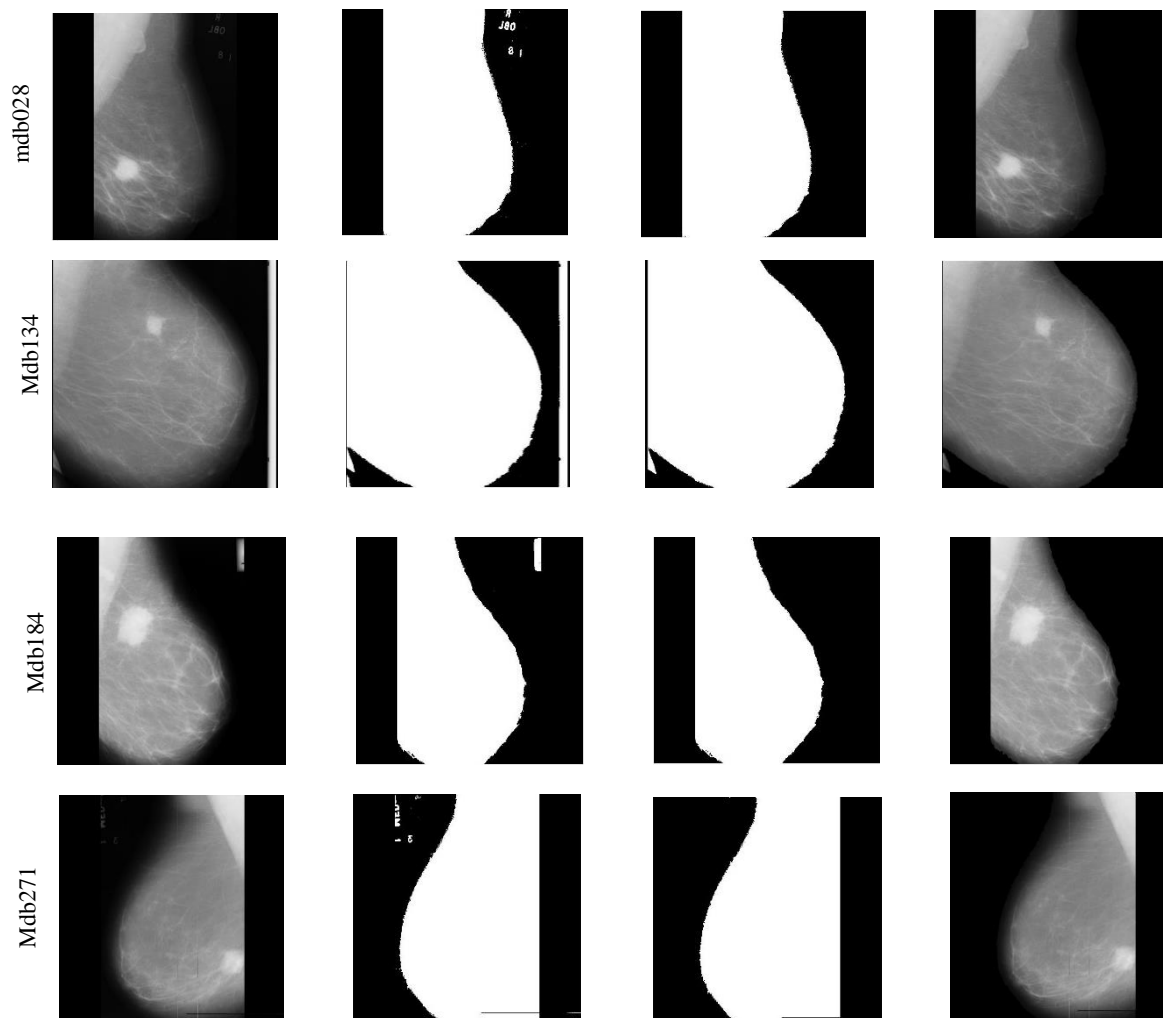
Donnée : Image mammographique avec artéfacts

- 1- Appliquer un seuillage global ($T = 18$) (transformer l'image en niveau de gris en image binaire).
- 2- Étiqueter les objets dans l'image binaire.
- 3- Sélectionner l'objet le plus grand
- 4- Réduire les distorsions (élimination des pixels isolés, correction des valeurs des pixels par rapport à leurs voisinages)
- 5- Eroder l'image binaire sélectionnée en utilisant un élément structurant en forme de disque plat de rayon 5.
- 6- Dilater l'image binaire en utilisant le même élément de structuration.
- 7- Multiplier l'image binaire résultante par l'image originale.

Résultats Image mammographique sans artéfacts

B. Résultats de l'étape de l'élimination des artefacts

L'algorithme 4.1 basé sur le seuillage global est utilisé pour éliminer les artefacts susceptibles d'être présents dans une image mammographique. Cet algorithme procède à transformer l'image en niveau de gris en une image binaire. Les objets détectés sont alors étiquetés dont le profil du sein correspond à l'objet le plus grand. Les images mdb028, mdb134, mdb184 et mdb271 sont des échantillons parmi les images évaluées :



Comme le montre la figure 4.4, les artefacts radio-opaques sont correctement éliminés des images prétraitées. Cette étape est suivie de l'élimination du muscle pectoral qui constitue un autre artefact naturellement lié à l'image.

Figure 4. 4. Résultat de l'étape d'élimination des artefacts radio-opaques

3.1.2. Identification et Elimination du muscle pectoral

Le muscle pectoral apparaît dans l'un des deux coins supérieurs gauche ou droit de l'image mammographique. Bien que l'exposition d'une grande partie du muscle pectoral en vue MLO d'un examen mammographique soit préférable, afin de garantir que le parenchyme mammaire est correctement positionné dans le champ d'exploration, le caractère plus brillant de ses pixels suite à sa forte atténuation des RX est néfaste sur l'analyse automatique de ces images. D'où sa

classification parmi les artefacts pouvant infecter non seulement la segmentation mais aussi l'extraction des caractéristiques et le processus de classification, ce qui entraîne un taux élevé de faux positifs [6], [169]–[173]. D'où l'élimination de cet artefact à partir des vues MLO des mammogrammes est cruciale. Pour ce faire nous avons fait appel à la segmentation par la méthode de la croissance région. Vu que cette technique nécessite la sélection d'une graine qui déclenche le processus de partitionnement, nous avons prévu une étape intermédiaire permettant d'orienter toutes les vues MLO de façon à ce que le sein soit orienté de gauche vers la droite afin de faciliter le processus de l'initialisation de sa segmentation.

A. Identification de l'orientation et ré-orientation du sein

Dans une vue MLO d'un mammogramme, le profil du sein peut être orienté soit de droite vers la gauche ou de gauche vers la droite. Lorsque le muscle pectoral occupe le coin supérieur droit, le sein est orienté de droite vers la gauche figure 4.5 (a), tandis que le sein est orienté de gauche vers la droite lorsque le muscle pectoral occupe le coin supérieur gauche figure 4.5 (b).

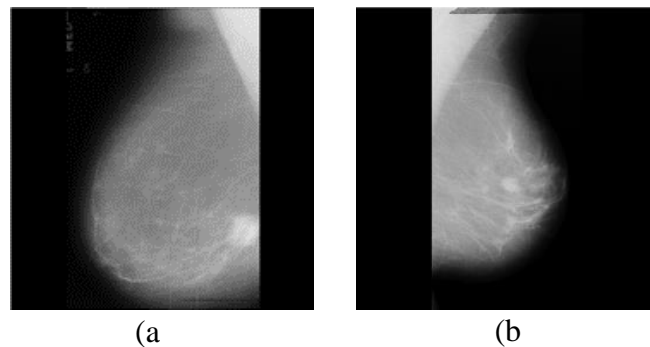


Figure 4. 5. Deux mammogrammes d'orientations différentes ;
(a) de droite vers la gauche, (b) de gauche vers la droite

La première étape consiste à déterminer l'orientation du sein dans l'image origine en divisant l'image verticalement en deux sous images de même taille ; les moyennes des niveaux de gris de chaque sous image sont calculées. La partie ayant la plus grande moyenne est celle qui contient le sein en grande partie y compris le muscle pectoral.

La deuxième étape consiste à inverser l'orientation des images orientées de droite vers la gauche (figure 4.6) pour unifier l'orientation de tous les mammogrammes pendant cette étape. De cette manière l'automatisation du processus d'initialisation de la segmentation par la technique de croissance de la région peut être effectuée aisément.

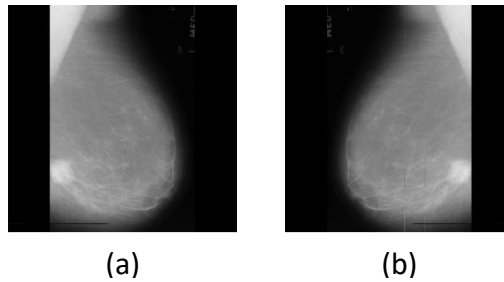


Figure 4. 6. Inversement de l'orientation du mammogramme mdb271 ; (a) image origine orientée de droite vers la gauche, (b) même image orientée de gauche vers la droite

B. Elimination du muscle pectoral

Les étapes précédentes assurent que le muscle pectoral occupera le coin supérieur gauche de l'image résultante quelle que soit l'orientation du sein dans l'image d'entrée. La sélection de la graine pour l'amorçage du processus de la segmentation par la méthode de croissance de la région, du muscle pectoral peut être faite alors d'une manière automatique selon deux algorithmes : le premier se déroule avec réduction de la taille de l'image tandis que le deuxième se déroule avec conservation de la taille de l'image [167], [168].

- **Elimination du muscle pectoral avec réduction de taille de l'image**

L'algorithme de cette méthode consiste à éliminer toutes les colonnes noires de tous les côtés du profil du sein. Dans l'image résultante le muscle pectoral sera adjacent au coin supérieur gauche (figure 4.7 (b)). Le choix de la graine d'amorçage est fait par ses coordonnées cartésiennes sachant que le coin supérieur gauche constitue l'origine. Le processus de segmentation continue jusqu'à ce que le muscle pectoral soit isolé et ajouté à l'arrière-plan de l'image. Cette méthode peut être utilisée lorsque la taille de l'image résultante ainsi que la localisation de l'anomalie sont sans intérêt. Le pseudo-algorithme de la méthode est donné par l'algorithme 4.2 [167].

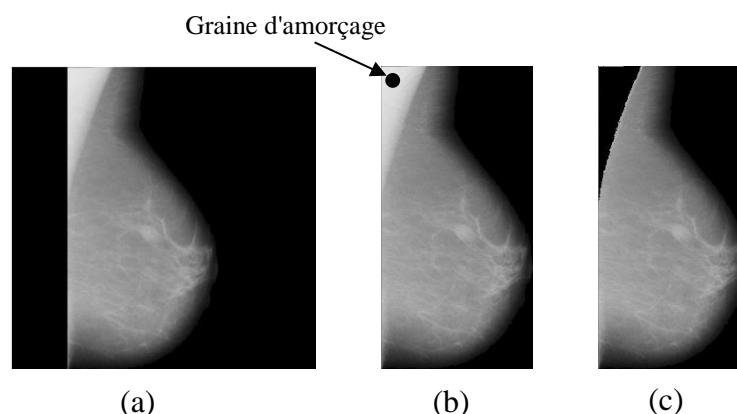


Figure 4. 7. Délimitation automatique de la "ROI" avec réduction de taille (mdb010)
(a) image origine sans artéfact, (b) image automatiquement rognées, (c) "ROI" détectée

Algorithme 4.2 : Algorithme pour l'élimination du muscle pectoral avec réduction de taille de l'image.

Données : Image mammographique traitée par l'algorithme 1 (image sans artéfacts)

- 1- Déterminer l'orientation de l'image mammographique
- 2- Si l'image est orientée de droite à gauche
 |
 Inverser l'image
 Fin si
- 3- Eliminer tous les colonnes nulles de la matrice image
- 4- Sélectionner la graine pixel (10, 10)
- 5- Exécutez l'algorithme de segmentation basé croissance de la région

Résultats : Afficher la ROI

- **Elimination du muscle pectoral sans réduction de taille de l'image**

Afin d'éliminer le muscle pectoral du mammogramme tout en permettant la localisation de la zone suspecte dans l'image mammographique, nous avons procédé à une initialisation du processus de la segmentation du muscle pectoral par la méthode de la croissance de la région sans recours à la réduction des dimensions de l'image. D'après le positionnement du profil du sein dans la vue MLO, le muscle pectoral est toujours adjacent à la face supérieure de l'image figure 4.8 (a). Aussi suite aux opérations de vérification et de réorientation des vues orientées de droite vers la gauche, le muscle pectoral occupe le côté gauche de l'image figure 4.8 (b). Le principe est similaire à celui adopté précédemment il consiste à fixer une ligne parmi les premières lignes de l'image sur laquelle on cherche à rencontrer le premier pixel ayant une intensité élevée. Ceci est dû au fait que le muscle pectoral est adjacent au bord supérieur de l'image, et peut s'étendre jusqu'à la mi-hauteur de l'image. Par ailleurs, le muscle pectoral est caractérisé par la brillance de ses pixels vu sa forte capacité d'atténuation des RX. Le pseudo-algorithme de la méthode est donné par la figure 4.11 [168].

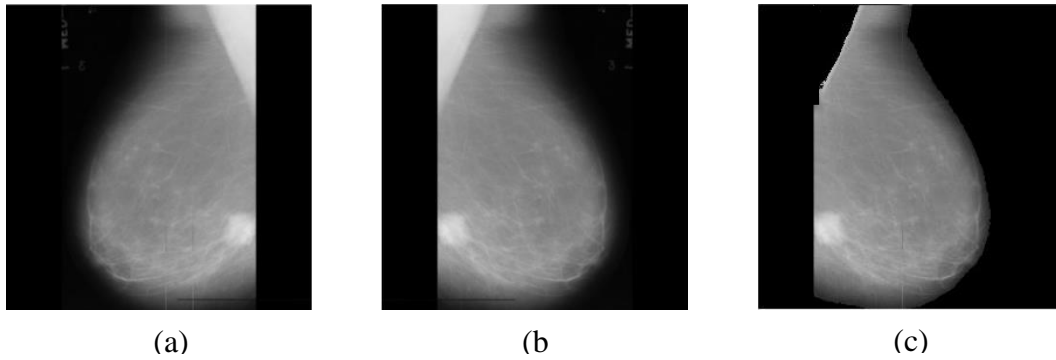


Figure 4. 8. Délimitation automatique de la "ROI" sans réduction de taille (mdb271) (a) image origine sans artéfact, (b) image automatiquement rognées, (c) "ROI" détectée

Algorithme 4. 3. Algorithme d'élimination du muscle pectoral sans réduction de taille

Algorithme 4.3 : Algorithme pour l'élimination du muscle pectoral sans réduction de taille d'image

Données : Image mammographique traitée par l'algorithme 1 (image sans artéfacts)

- 1- Déterminer l'orientation de l'image mammographique
- 2- Si l'image est orientée de droite à gauche
 - Inverser l'image
- Fin si**
- 3- Retrouver puis sélectionner comme graine le premier pixel $P(10, x)/I(10, x) \geq 100$
- 4- Exécutez l'algorithme de segmentation basé croissance de la région

Résultats : Afficher la ROI

C. Résultats de l'élimination du muscle pectoral

La localisation de la zone suspecte est parmi les objectifs à accomplir par notre système d'aide à la détection "DAOe". D'où le choix de la segmentation du muscle pectoral par la méthode de la croissance de la région en utilisant l'algorithme 4.3 qui permet la conservation des dimensions de l'image. La figure 4.9 montre les résultats de l'étape d'élimination du muscle pectoral appliquée aux images mdb010, mdb028, mdb134, mdb184, mdb271.

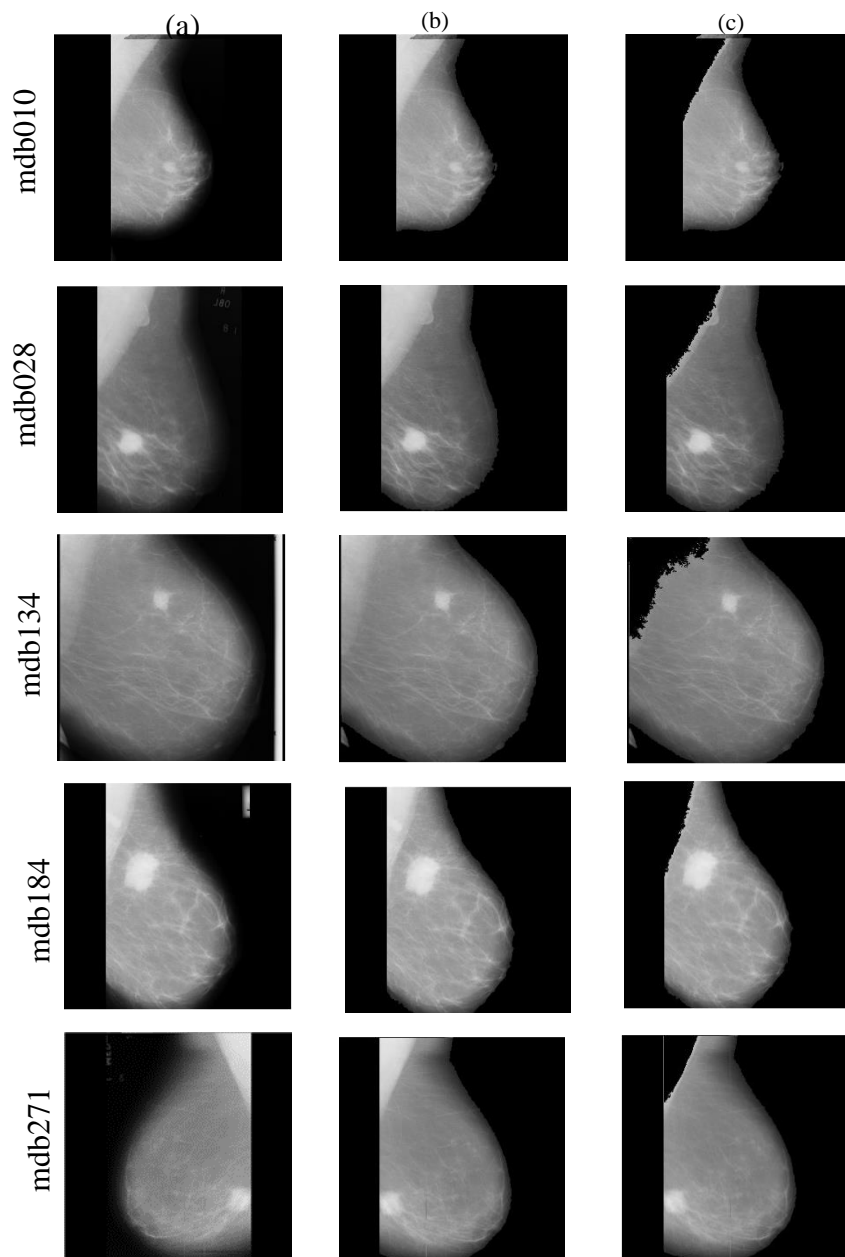


Figure 4. 9. Résultat de l'étape d'élimination du muscle pectoral. (a) image origine, (b) image sans artéfact orientée de gauche à droite, (c) élimination du muscle pectoral.

On constate d'après cette figure que le muscle pectoral est correctement éliminé des mammogrammes. En outre l'algorithme vérifie l'orientation de l'image, la réoriente dans le cas où l'image mammographique est orientée de droite à gauche image mdb271 figure 4.9 (b). La sélection de la graine pour démarrer le processus de segmentation est faite correctement figure 4.9 (c).

3.2. Amélioration de la qualité des "ROI"

Après avoir délimité la région d'intérêt qui est le résultat des étapes précédentes, vient l'étape de l'amélioration de sa qualité. Dans cette étape nous employons notre modèle d'amélioration de la qualité des images mammographiques (Chapitre 3) basé sur l'optimisation multiobjectif. Les résultats quantitatifs (les vecteurs de décision et les paramètres IQA) sont donnés par les tableaux 4.1 et 4.2, quant aux images (avant et après traitement) sont présentées sur la figure 4.14.

Tableau 4. 1. Les vecteurs de décision des "ROI" améliorées par le modèle multiobjectif

Image	NBH	NBV	B	α	β
Mdb010	76	77	9.997	0.501	0.0054
Mdb028	83	78	9.451	0.532	0.0042
Mdb134	80	88	9.981	0.500	0.0055
Mdb184	91	70	8.959	0.501	0.00032
Mdb271	74	112	9.600	0.498	0.0005

Le tableau 4.1 présente l'ensemble des vecteurs de décision (NBH, NBV, B, α , β) qui représentent les paramètres de la transformation locale de notre modèle d'amélioration de la qualité des mammogrammes. Ces vecteurs permettent l'optimisation de l'amélioration de la qualité des "ROI" avant de procéder à leurs segmentations. L'application de la transformation locale en utilisant le vecteur de décision relatif à chaque région d'intérêt conduit aux résultats du tableau 4.2 qui donne les valeurs des paramètres "IQA" des régions d'intérêt.

Tableau 4. 2. Les paramètres d'évaluation de l'amélioration de la qualité des "ROI" par le modèle multiobjectif

"ROI"		EC	EME	AMBE	FSIM
Mdb010	Origine	0.958	6.991	0	100%
	Améliorée	1.291	7.161	0.077	99.02%
Mdb028	Origine	1.143	5.651	0	100%
	Améliorée	1.521	6.444	0.190	98.63%
Mdb134	Origine	1.855	5.152	0	100%
	Améliorée	2.532	5.667	0.310	98.19%
Mdb184	Origine	1.245	4.752	0	100%
	Améliorée	1.694	5.945	0.180	98.54%
Mdb271	Origine	1.187	4.641	0	100%
	Améliorée	1.607	6.218	0.170	98.62%

D'après les résultats du tableau toutes les régions d'intérêt "ROI" sont bien améliorées du fait que les paramètres EME et EC responsables respectivement de l'évaluation de l'amélioration du contraste et de la contenance de bord sont augmentés. Leurs valeurs moyennes ont subi des augmentations respectivement de 5.437 à 6.287 et de 1.277 à 1.729. Cette amélioration est accompagnée d'une bonne conservation de la luminosité d'une part, et de la conservation d'informations contenues dans l'image d'autre part. Ceci s'est traduit respectivement par les faibles valeurs du paramètre AMBE et les valeurs élevées du paramètre FSIM. En effet, la moyenne du paramètre AMBE responsable de l'évaluation de la conservation de luminosité n'a

pas dépassé 0.185, quant au paramètre FSIM traduisant la conservation des informations a dépassé 98% en moyenne.

Ces résultats quantitatifs sont confirmés par les profils des images traitées (figure 4.10). Comme le montre la figure 4.10 (b) les images correspondantes aux ROI traitées sont bien contrastées, plus proéminentes et fidèles aux régions d'intérêt sources.

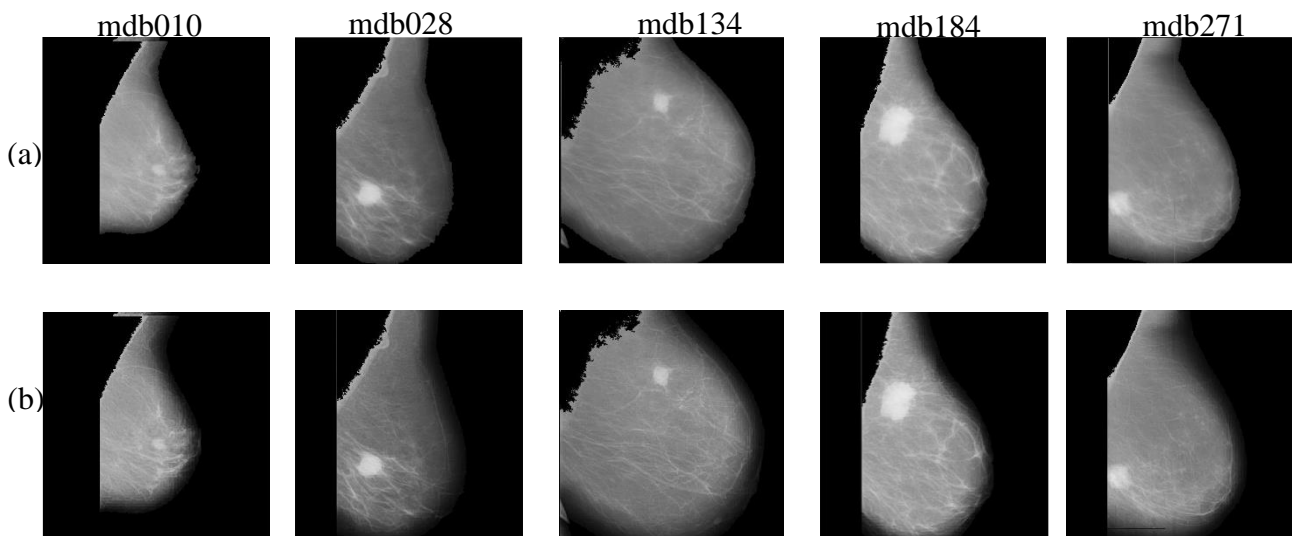


Figure 4. 10. Image obtenues après amélioration des "ROI"

4. Segmentation de la Région d'intérêt

Après avoir délimité automatiquement la "ROI" qui est constituée seulement du parenchyme mammaire, vient l'étape décisive qui a pour but d'isoler la zone probablement suspecte (masse) du reste de la "ROI". Pour l'identification de la zone suspecte d'une manière automatisée, nous avons choisi l'algorithme de regroupement de données K-means pour la segmentation de la "ROI" en trois classes correspondantes dans le sens ascendant des niveaux de gris à l'arrière-plan de l'image, les tissus adipeux et celle des tissus fibro-glandulaires du parenchyme mammaire [174], [175]. Cet algorithme ne nécessite aucune initialisation ou intervention de l'utilisateur. Bien que cet algorithme présente des avantages certains, à l'opposé il possède quelques inconvénients mais qui restent remédiables.

4.1. Algorithme k-means : les avantages

L'algorithme k-means est un algorithme de classification non supervisée de données. Il convient bien à la segmentation des images médicales vu plusieurs avantages qu'il offre. En effet l'algorithme peut être appliqué à des données de grandes dimensions et ne nécessite aucune initialisation ou une autre information pour l'opération de partitionnement. Parmi d'autres avantages de l'algorithme sa facilité d'implémentation et le temps d'exécution raisonnable [176], [177]. L'algorithme 4.4 illustre le principe adopté par cette technique.

Algorithme 4.4 : Algorithme de segmentation par la technique k-means

Données

Nombre de classes (K)

Ecart maximal acceptable pour le déplacement des centroïdes ε

Générer aléatoirement K centroïdes $\{C_1, \dots, C_k\}$ des clusters

Répéter

Pour chaque pixel de l'image

 Calculer les distances $d_i \leftarrow |p(x, y) - C_i|$

 Affecter le pixel $p(x, y)$ au cluster dont le centroïde est le plus proche

Fin pour

 Recalculer les centroïdes de tous les clusters

$$C_i^{(n)} = \frac{1}{N_i} \sum_{j=1}^{N_i} p_j(x, y)$$

Jusqu'à ce que $\sup(|C_i^n - C_i^{n-1}|) < \varepsilon$

Résultats : Afficher le résultat de partitionnement

4.2. Algorithme k-means : les inconvénients et remédiation

Quoique les avantages qu'il présente l'inconvénient majeur de l'algorithme k-means réside au niveau de la dépendance des résultats finals de la segmentation de la première étape de l'initialisation des centroïdes des clusters ou classes (deux initialisations différentes aboutissent à deux résultats de segmentations différents [97]–[99]). L'autre inconvénient il ignore l'information spatiale des pixels, d'où la possibilité de production des clusters qui peuvent être disjoints [97]–[99], [178]–[180]. Pour remédier à ces problèmes nous avons procédé à coupler l'algorithme principal k-means avec d'autres algorithmes qui lui permettront d'améliorer les résultats finals de la segmentation.

4.2.1. Remédiation au problème lié à l'initialisation

Concernant le problème de l'initialisation des centroïdes des clusters, la solution apportée consiste à faire appel au processus de l'optimisation afin de déterminer les centres adéquats permettant une meilleure segmentation pour chaque image. En effet, l'ensemble des centres des clusters permettant un meilleur partitionnement de l'image est celui qui maximise simultanément la dissimilarité interclasse et la similarité intra-classe. Ceci est traduit par la minimisation de la fonction objectif donnée par l'équation 2.18 (§4 du chapitre 2).

Sur le plan pratique nous avons introduit l'algorithme génétique afin de déterminer les centres des clusters (μ_1, μ_2, μ_3) qui permettent d'obtenir une meilleure segmentation. Les paramètres de l'algorithme génétique utilisé sont comme suit [167], [168] :

- Le codage réel des paramètres de décision (μ_1, μ_2, μ_3) ;
- Le nombre initial de la population est de 50 chromosomes maintenus constant ;
- La sélection *stochastique uniforme* ;
- Le croisement à *point unique* ;
- La mutation *adaptative possible*
- La fonction objectif à minimiser est l'équation 2.18 (§4 du chapitre 2).
- Le critère d'arrêt est une combinaison entre le changement faible de la fonction objectif pendant 50 itérations et le nombre de générations est fixé à 100.

4.2.2. Remédiation au problème de disjonction de clusters

Le standard BIRADS stipule qu'une masse correspond à une forme ronde ou ovale apparente au moins sous deux incidences [30], [41], [181], d'où le caractère jointif d'une masse mammaire. En effet la suppression des artéfacts radio-opaques et de muscle pectoral qui peuvent exister dans un mammogramme, est un moyen qui prévient l'algorithme de les inclure parmi la zone suspecte vu leurs caractères brillants dans l'image. Au niveau de la "ROI", afin d'éliminer les pixels aberrants qui peuvent nuire la précision de l'analyse automatique, l'algorithme procède à calculer le centre de l'objet détecté qui servira comme graine pour segmenter à nouveau l'image traitée en utilisant la méthode de la croissance de la région. Cette dernière opération a pour objectif d'ajouter les pixels aberrants à l'arrière-plan de l'image. Le pseudo-algorithme de la technique permettant d'isoler la zone suspecte du reste de la "ROI" est donné par la figure 4.14.

Algorithme 4.5 : Algorithme couplage AG-k-means pour la détection de la zone suspecte

Donnée : ROI améliorée

1. Vers l'algorithme génétique

- Générer la population initiale (50 chromosomes de (μ_1, μ_2, μ_3))
- Calcul de la fitness de la population initiale Equation (2.18)

Répéter

- Sélection (stochastique uniforme)
- Croisement (point unique)
- Mutation (adaptative)
- Calcul de la fitness des nouvelles générations Equation (2.18)

Jusqu'à ce que la population converge

Résultat intermédiaire (μ_1, μ_2, μ_3)

2. Vers l'algorithme K-means

- Pour chaque pixel, calculer les distances euclidiennes d_1, d_2, d_3 Equation (2.19)
- Attribuer chaque pixel au centroïde le plus proche μ_1, μ_2, μ_3 .
- Regrouper en fonction de la plus petite distance

Résultat de segmentation

- Calculer le centre de l'objet

3. Vers l'algorithme de croissance de région

- Segmenter l'image en utilisant le centre comme graine

Résultat afficher la zone suspecte

4.3. Résultats de la segmentation des ROI

Il s'agit maintenant de segmenter la "ROI" améliorée afin d'isoler la zone probablement suspecte du reste de la "ROI". Cette opération est réalisée par l'algorithme 4.14 qui couple l'algorithme k-means et l'algorithme génétique. Les résultats quantitatifs et qualitatifs obtenus sont présentés dans les tableaux 4.3, 4.4 et 4.5, et la figure 4.15.

Le tableau 4.3 montre les vecteurs de décision des régions évaluées après l'opération de l'optimisation des centres des classes de chaque "ROI" améliorée (résultats de la première partie de l'algorithme 4.4). Comme le montre les résultats du tableau, chaque région d'intérêt a ses propres centres de classes qui permettent une meilleure segmentation. L'utilisation de ces centres dans le processus de partitionnement de chaque "ROI" a abouti aux images présentées par la figure 4.11 et les résultats quantitatifs récapitulés dans les tableaux 4.4 et 4.5

Tableau 4. 3. Vecteurs de décision des "ROI" permettant une meilleure segmentation

Image	μ_1	μ_2	μ_3
Mdb010	9	114	243
Mdb028	8	143	236
Mdb134	9	105	136
Mdb184	8	110	243
Mdb271	9	113	244

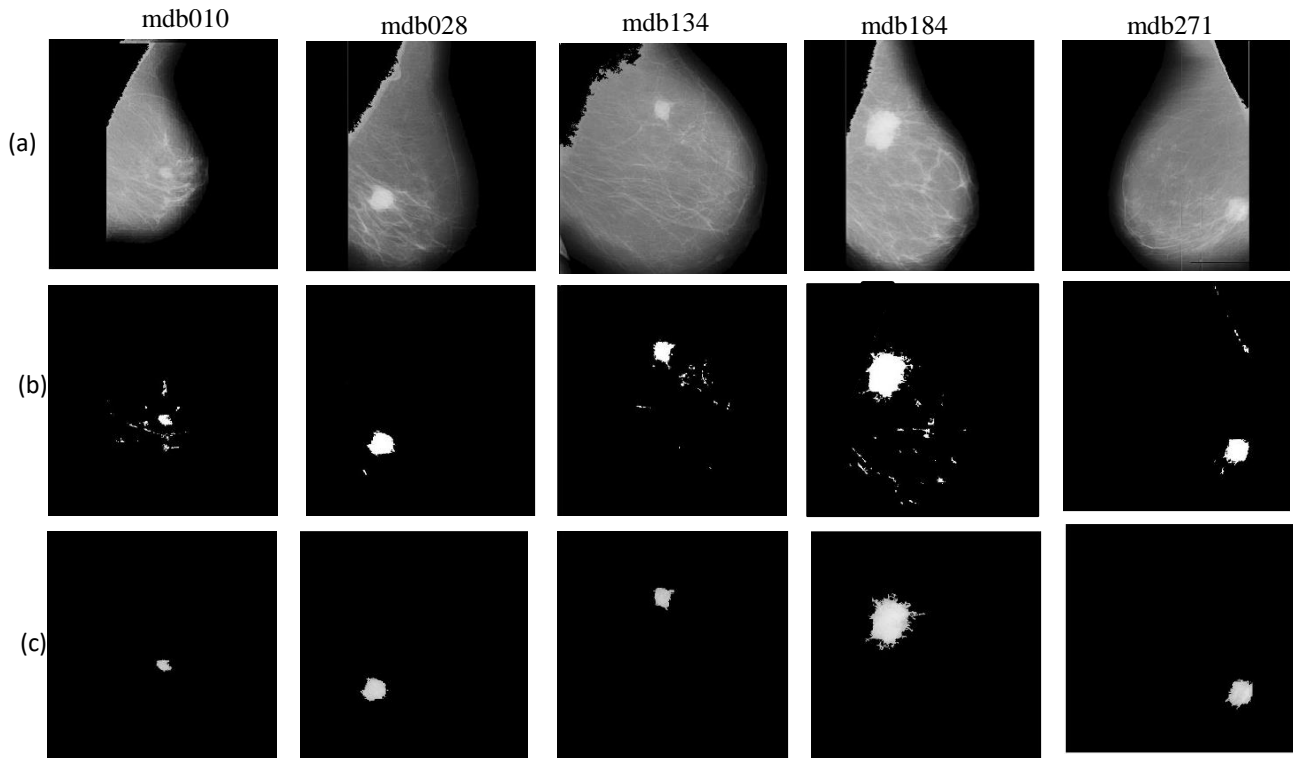


Figure 4. 11. Résultats de la segmentation des ROI; (a) ROI améliorée (b) segmentation par k-means, (c) zones suspectes isolées

Comme le montre la figure 4.11 (b) la segmentation en utilisant l'algorithme k-means peut engendrer des classes disjointes. Un ensemble de pixels aberrants de même ordre de l'intensité de la zone suspecte apparaissent. C'est la raison pour laquelle on a couplé à nouveau l'algorithme k-means avec l'algorithme de la segmentation par la croissance de la région. Le résultat de ce dernier couplage montre son efficacité à isoler complètement la zone suspecte du reste de l'image.

Afin d'évaluer la qualité de la segmentation, l'algorithme détermine le centre et le rayon du cercle circonscrit qui délimite la zone suspecte. L'origine du repère étant le coin inférieur gauche de l'image (repère fixé dans la base de données Mini-MIAS). Le tableau 4.4 montre les valeurs évaluées des coordonnées cartésiennes du centre de la masse (x, y) et du rayon R du cercle circonscrit de la tumeur en pixels.

Tableau 4. 4. Valeurs évaluées des coordonnées du centre (x, y) et du rayon R du cercle circonscrit de la tumeur en pixels

Image	x	y	R
Mdb010	524	425	34
Mdb028	338	317	54
Mdb134	469	729	49
Mdb184	352	625	110
Mdb271	784	269	67

Le tableau 4.5 présente les écarts entre les valeurs évaluées des coordonnées du centre (x, y) et du rayon R du cercle circonscrit de la masse en pixels et leurs valeurs fournies par la base de données Mini-MIAS.

Tableau 4.5. Ecart entre les valeurs évaluées des coordonnées du centre (x, y) et du rayon R du cercle circonscrit de la masse en pixels et leurs valeurs fournies par la base de données Mini-MIAS

Image	$ x - x_c $	$ y - y_c $	$ R - R_t $
Mdb010	1	0	1
Mdb028	0	3	2
Mdb134	0	1	0
Mdb184	0	1	4
Mdb271	0	1	1

D'après le tableau, on constate que les écarts sont faibles ce qui montre que les zones suspectes sont correctement isolées de la "ROI". En effet, le plus grand écart relatif aux abscisses est 1 pixel, et 3 pixels pour les ordonnées. Concernant le rayon, le plus grand écart correspond au rayon du cercle circonscrit de la masse de l'image mdb184 qui n'a pas dépassé 4 pixels. Ces résultats prouvent la grande précision de notre système d'aide à la détection "DAOe".

5. Conclusion

En guise de conclusion, le système d'aide à la détection automatisée des masses mammaires qu'on vient de proposer aboutit à des résultats très satisfaisants. Ceci est dû à la bonne conception de ses étapes prétraitement et segmentation. Concernant le prétraitement la coopération des méthodes proposées a permis de délimiter et améliorer correctement la région d'intérêt "ROI". En premier lieu l'élimination des artéfacts radio-opaques est réalisé par la méthode de seuillage global, suivie de l'élimination du muscle pectoral en utilisant la segmentation par la méthode de la croissance de la région. La qualité de la "ROI" ainsi déterminée est améliorée en utilisant notre modèle d'amélioration de la qualité des images mammographiques par optimisation multiobjectif. Enfin la "ROI" améliorée est segmentée en utilisant un algorithme couplant l'algorithme k-means à l'algorithme génétique. L'ensemble des résultats obtenus permettent la localisation parfaite de la zone suspecte dans l'image mammographique.

Le système d'aide à la détection des masses mammaire conçu, pourra être utilisé comme une

base importante et précise pour la conception d'un système d'aide au diagnostic assisté par ordinateur "DAOx", par intégration de l'une des techniques des "machines learning" ou de "Deeplearning" afin de prédire automatiquement la sévérité des zones suspectes détectées.

Conclusion générale et perspectives

Les travaux de cette thèse entre dans le cadre général des traitements informatiques des images médicales, et plus particulièrement le traitement des images mammographiques depuis que les statistiques confirment que le cancer du sein représente une menace prépondérante pour la vie de la femme. Le seul moyen de vaincre cette maladie est sa détection précoce avant que le mal ne soit à un stade avancé où tout traitement reste sans efficacité. C'est dans ce sens que s'inscrivait l'objectif de notre travail ; il s'agissait en effet de développer un système automatisé de haute résolution d'aide à la détection précoce des masses mammaires "DAOe", qui servira par la suite comme une base solide pour la conception d'un système d'aide au diagnostic. En effet le succès de tout système d'aide au diagnostic "DAO" utilisant l'apprentissage automatique dépend fortement de l'efficacité de l'étape de prétraitement. Bien que les sensibilités soient considérablement augmentées par l'incorporation des techniques d'apprentissage profond, celles-ci ne peuvent remplacer le prétraitement. Cela est évident car les caractéristiques ne peuvent pas être extraites efficacement (pour les classifieurs) si elles ne sont pas clairement visibles dans les mammogrammes à analyser. D'où, la nécessité d'une meilleure amélioration de la qualité de ces images, chose traduite par l'optimisation des paramètres d'évaluation "IQA" de la qualité de l'image qui sont le contraste EME, la contenance en bord EC tout en conservant la luminosité et le contenu de l'image origine évalués respectivement par AMBE et FSIM.

Notre système de détection DAOe se compose de deux étapes essentielles prétraitement et segmentation. D'abord le prétraitement a pour objectif de rendre l'image à traiter plus informative et efficacement exploitable par la délimitation de la région d'intérêt "ROI" ainsi que par l'amélioration de sa qualité. Par la suite vient l'étape segmentation pour isoler la zone probablement suspecte du reste de la "ROI".

Durant la phase de prétraitement la délimitation de la région ROI est réalisée à partir de la vue médio-latérale-oblique de l'image médicale (mammogramme) à analyser par élimination successive des artefacts radio-opaques et du muscle pectoral. Durant l'élimination des artefacts radio-opaques nous avons utilisé la technique de seuillage global qui fonctionne d'une manière automatique. Cette opération est suivie de la suppression des pixels aberrants que ce soit de la partie du fond de l'image (partie noire de l'image binaire) ou de la partie de l'objet (partie blanche de l'image binaire). Cette suppression est pratiquée sur la base d'un traitement morphologique de l'image par application d'une érosion suivie d'une dilatation par l'élément structurant qui est un disque de rayon 5 pixels. Ensuite l'algorithme enchaine vers l'élimination du muscle pectoral. Afin d'accomplir cette tâche correctement, la vérification de l'orientation du sein dans l'image (de gauche vers la droite et vice versa) est effectuée, l'image est réorientée dans le cas où l'image est orientée de droite vers la gauche. Cette opération intermédiaire a pour

but de rendre le muscle pectoral apparent dans la partie supérieure gauche de l'image afin de faciliter la sélection du pixel graine qui démarre le processus de la segmentation par la méthode de croissance de la région. En effet, le muscle pectoral s'étend dès le bord supérieur de l'image et peut atteindre la mi-hauteur de l'image. La graine est choisie sur une ligne fixe parmi les premières lignes en cherchant à rencontrer le premier pixel ayant une intensité élevée.

Après élimination des artefacts radio-opaques et de muscle pectoral, la qualité de l'image résultante est améliorée davantage avant qu'elle subisse la segmentation pour isoler la zone suspecte. Cette amélioration est effectuée grâce à un modèle que nous avons proposé (chapitre 3). Il s'agit d'un modèle basé sur l'optimisation multi-objectif, où la fonction sigmoïde que nous avons établie constitue sa transformation locale et, l'équation représentant le compromis entre les paramètres d'évaluation EC et $AMBEn$ (qui représentent respectivement les paramètres d'évaluation de la contenance en bord et la conservation de la luminosité de l'image) constitue sa fonction objectif. En effet, l'expression de la fonction sigmoïde établie (équation 3.13) est déduite de l'expression classique de la fonction sigmoïde [9] en modifiant la base e par une base variable B dans l'intervalle $]1, 10]$. Cette modification avait pour objectif d'ajouter un autre degré de liberté à la fonction sigmoïde en plus des paramètres α et β afin d'obtenir facilement la sigmoïde la plus adaptée pour chaque image (la nouvelle fonction sigmoïde obtenue devient caractériser alors par un triplet (B, α, β)). La transformation locale a donc pour paramètres $(NBH, NBV, B, \alpha, \beta)$ qui sont les vecteurs de décision à optimiser (NBH et NBV) représentent respectivement les nombres de blocs dans le sens horizontal et vertical). Le principe de la méthode proposée consiste à déterminer pour chaque image le nombre de blocs ($NBH \times NBV$) sur lequel l'image doit être partitionnée et les paramètres (B, α, β) adéquats de la fonction sigmoïde, et ce en utilisant l'optimisation par la méthode des algorithmes génétiques. Ensuite les niveaux de gris de chaque bloc sont traités par la fonction sigmoïde afin de constituer l'image améliorée qui optimise la fonction objectif du problème. Le but du problème était la maximisation simultanée du vecteur objectif $(EC, EME, AMBEn, FSIM)$; où les composantes de ce vecteur représentent respectivement la contenance en bord, l'amélioration en contraste, la conservation de luminosité et la conservation d'information de l'image origine. Le problème modélisé en tant qu'un problème d'optimisation multiobjectif admet un ensemble de solutions non dominées. Ainsi nous avons transformé ce problème multiobjectif en son équivalent monoobjectif afin d'automatiser la détermination d'une solution unique. Afin que le problème conserve ses caractéristiques, une étude des relations globales, locales et composites d'une part et, de la signifiante des objectifs d'autre part a été menée. Cette étude nous a permis une réduction des objectifs et de ne retenir que deux objectifs significatifs conflictuels EC et $AMBEn$ entre lesquels nous avons déterminé un compromis. Finalement la fonction objectif

est exprimée en tant que somme pondérée des deux objectifs EC et AMBEn dont les poids sont respectivement 0.8 et 0.2 (équation 3.15)

Après avoir délimité et amélioré la région d'intérêt ROI vient l'étape de sa segmentation consistant à isoler la zone probablement suspecte du reste de la ROI. La technique utilisée est celle de classification des pixels k-means puisque elle ne nécessite aucune initialisation ou une autre connaissance à priori sur les données (pixels). Nous avons couplé cette technique avec un algorithme d'optimisation mono-objectif pour déterminer les centres des classes afin de remédier au problème de l'initialisation aléatoire de ces centres. Vu que la segmentation par l'algorithme k-means engendre des classes disjointes nous avons procédé à la segmentation par croissance de régions pour éliminer les pixels aberrants. Le centre du plus grand objet est utilisé comme graine afin de démarrer la segmentation permettant par la suite d'isoler parfaitement la zone suspecte.

Notons enfin que le modèle réalisé (concernant l'amélioration optimale de la qualité des images mammographiques) et le système de détection conçu "DAOe" sont testés et validés par leurs applications sur les images de la base de données Mini-MIAS. Les résultats fournis par le modèle (la qualité des images et leurs paramètres d'évaluation) sont nettement meilleurs par rapport à ceux évalués par les méthodes usuelles (les méthodes basées sur la modification de l'histogramme et sur l'optimisation) d'une part, et à ceux obtenus par la méthode de Gandhamal et al [9] sur laquelle notre méthode s'est basée, d'autre part. Quant au système "DAOe" conçu, il a montré son efficacité du fait que toutes les régions d'intérêt sont efficacement extraites, améliorées et segmentées. Aussi les dimensions des zones suspectes obtenues par notre système sont comparables à celles fournies par la base de données Mini-MIAS. L'ensemble confirme la performance de notre système.

Perspectives :

A la suite des travaux de cette thèse, et dans le cadre de prolongement des travaux effectués, divers perspectives sont à envisager. Ainsi il est d'autant intéressant de développer le système d'aide au diagnostic "DAOx" la poursuite du système d'aide à la détection "DAOe" qu'on vient de concevoir afin de prédire la sévérité des lésions détectées. Ceci nécessite la caractérisation des zones suspectes par des descripteurs décrivant au mieux leurs états de malignité ou de bénignité et, d'utiliser par la suite l'une des technologies d'intelligence artificielle telles que les machines learning ou l'utilisation des approches basées sur l'apprentissage profond (Deep learning) dont l'apprentissage automatique fondé sur une extraction non-supervisée des descripteurs. Finalement le système détection/diagnostic qui sera réalisé pourra être utilisé comme un module complémentaire intégré à la mammographie afin de fournir aux cliniciens la localisation de la lésion probable ainsi que son état de gravité à un stade prématuré.

T : Tumeur	
Correspond au terme « Tumeur » : La taille de la tumeur, son degré d'infiltration dans les tissus avoisinants (envahissement local) sont précisées.	
T0	Aucun signe de la tumeur/ pas de tumeur palpable
Tis	Tumeur in situ ; modification tissulaire non invasive, limitée aux cellules superficielles d'un canal galactophore.
T1	Tumeur jusqu'à 2 cm ; la taille de la tumeur peut être précisée avec les lettres minuscules a, b et c. Exemple : pT1a = jusqu'à 0,5 cm.
mi	carcinome micro-invasif (infiltration visible uniquement au microscope).
T2	tumeur de 2 à 5 cm. T3
T3	tumeur de plus de 5 cm.
T4	tumeur de toute taille ayant envahi (= infiltré) la paroi thoracique et/ou la peau ou de nature inflammatoire ; l'emploi de lettres minuscules de a à d permet de préciser l'anomalie. Exemple : pT4d = carcinome inflammatoire (voir
M	multifocal/multicentrique : présence de plusieurs foyers tumoraux dans un sein.
r	récidive : réapparition de la tumeur.
N : Ganglions lymphatiques régionaux (Nodes)	
Métastases ganglionnaires avec précision de l'anomalie	
N0	pas de tumeur dans les ganglions lymphatiques régionaux.
N1	cellules cancéreuses dans 1 à 3 ganglions lymphatiques
N2	cellules cancéreuses dans 4 à 9 ganglions lymphatiques
N3	cellules cancéreuses dans 10 ganglions ou plus.
sn	ganglion sentinelle atteint ou non du ganglion sentinelle
Mi	micrométastase : petites métastases microscopiques (0,2 mm à 0,2 cm) dans les ganglions lymphatiques
L1, V1	présence de cellules cancéreuses dans les vaisseaux lymphatiques (L) ou sanguins (V, invasion veineuse).
Métastases	
Propagation de cellules tumorales vers d'autres organes éloignés	
M0	absence de métastases à distance.
M1	présence de métastases ; une précision peut être indiquée : (OSS = osseuses ; PUL = pulmonaires ; HEP = hépatiques
mi	micrométastase (métastase de taille microscopique)

Annexe 2 (a)

La technique d'égalisation de l'histogramme

Considérons une image de taille $m \times n$ où m et n représentent respectivement le nombre de ligne et de colonne de l'image ; l'intensité de chaque pixel qui occupe la position (i, j) étant $I = I(i, j)$ avec $I(i, j) \in \{I_0, I_1, \dots, I_{L-1}\}$, $N = m.n$. Le pseudo algorithme de la méthode est :

Pseudo code de la technique (HE)

1. Calcul de l'histogramme $h(i)$ pour chaque $i \in [0, L - 1]$
 2. Normalisation de l'histogramme $h_n(i) = \frac{h_i}{N}$
 3. Calcul de la densité de de probabilité normalisée pour chaque niveau de gris I
 - a. $D(i) = \sum_{j=0}^i h_n(j)$
 4. Transformation des niveaux de gris de l'image $T(I(i, j)) = D(I(i, j)) \times (L - 1)$
-
-

Annexe 2 (b)

Pseudo code de la technique (BBHE)

1. Calculer $I_{inf}(i, j) = \{I_0, I_1, \dots, I_m\}$ et $I_{sup}(i, j) = \{I_{m+1} + 1, I_{m+2}, \dots, I_{L-1}\}$
 2. Calculer l'histogramme de chaque sous image
 3. $D_{inf}(i) = \sum_{j=1}^i h_n(j) \quad j \in \{I_0, I_1, \dots, I_m\}$
 4. $D_{sup}(i) = \sum_{j=X_{m+1}}^{i=L-1} h_n(j) \quad j \in \{I_{m+1}, I_{m+2}, \dots, I_{L-1}\}$
 5. $T_{inf}(I(i, j)) = I_0 + (I_m - I_0) \times D_{inf}(I(i, j))$
 6. $T_{sup}(I(i, j)) = I_{m+1} + (I_{L-1} - X_{m+1}) \times D_{sup}(I(i, j))$
 7. $T(I(i, j)) = T_{inf}(I(i, j)) \cup T_{sup}(I(i, j))$
-
-

Pseudo code de la technique (CLAHE)

Données Image origine, N_{dec} limite normalisée de découpage de l'histogramme

1. Redimensionner l'image I_c de $(M \times M)$
2. Décomposer l'image en n blocs, $n = \frac{M \times M}{m \times m}$
3. $H_n = \text{histogramme}(n)$, % histogramme de n blocs
4. Calculer N_{gris} % nombre de niveau de gris dans le bloc
5. Calculer $N_{moy} = \frac{N_X \times N_Y}{N_{gris}}$ % la moyenne des pixels par niveau de gris dans le bloc
6. Calcul de la limite de découpage $L_{Dec} = N_{Dec} \times N_{moy}$
7. Découpage de l'histogramme H_n selon la valeur de L_{Dec} . Pour chaque niveau de gris dépassant L_{Dec} soit $N_{\Sigma L_{Dec}}$ le nombre de pixels découpés
Calculer $N_{p_{Dec}} = \frac{N_{\Sigma L_{Dec}}}{N_{gris}}$ % distribution des pixels découpés
8. I_c Image obtenue après égaliser l'histogramme découpé et redistribué conformément à l'algorithme de HE
9. I_{cf} Image finale obtenue par lissage des voisinages des blocs par interpolation bilinéaire, et redimensionner l'image pour retrouver les dimensions initiales

Résultat I_{cf} image améliorée en utilisant la technique CLAHE

Algorithmel : Segmentation par la technique de croissance de la région

Données : Déterminer l'ensemble des Graines $G = \{G_1, \dots, G_n\}$

Déterminer le critère d'homogénéité des régions $H(G_i, P)$

Sortie : les régions $(R_1 \dots, R_m)$

$i \leftarrow 0$

Tant que $G \neq \emptyset$ **faire**

$R_i \leftarrow \emptyset$

Ajouter $G_i \leftarrow R_i$

Supprimer G_i de G

$R_{i_{centre}} \leftarrow G_i$

Pour chaque *pixel* $\in R_i$

Voisinage \leftarrow le voisinage spatiale du pixel

Pour chaque pixel P du voisinage faire

Si $H(G_i, P \in V(G_i))$ est vrai

Ajouter P à la région R_i

Supprimer P de G

Mettre à jour $R_{i_{centre}}$

Fin Si

Fin pour

Fin pour

$i \leftarrow i + 1$

Fin Tant que

Résultats : Afficher le résultat de partitionnement

 Algorithme 2 : Segmentation par la technique Division/ Fusion

Données : Image I de taille $M \times M$ ($M \leftarrow 2^m$)

Le critère d'homogénéité H

La taille minimale de la région $T_{min} = 2^m$

$R_0 = I$ (la première région est l'image entière)

Tant que R_i n'est pas homogène $H(R_i) = faux$, et la taille de $R_i > T_{min}$ **faire**

Diviser récursivement R_i en quatre quadrants égaux

Si $H(R_j \cup R_k) = Vrai$ pour toutes régions de même niveau et de même parent

Fusionner spatialement R_j et R_k

Fin si

Fin tant que

Pour toutes régions adjacentes R_j et R_k de différents niveaux

Si $H(R_j \cup R_k) = Vrai$

Fusionner spatialement R_j et R_k

Fin si

Fin pour

Résultats : Afficher le résultat de partitionnement

Annexe 4

Mesure d'indice de similarité des caractéristiques (Feature Similarity Index Measure (FSIM))

- La phase de congruence

En chaque point x d'une image $2D$ PC est définie par :

$$PC_{2D}(x) = \sum_j \frac{E_{\theta_j}(x)}{\varepsilon + \sum_n \sum_j A_{n,\theta_j}(x)} \quad (8)$$

Avec $A_{n,\theta_j}(x)$ est l'amplitude locale d'ordre n et d'orientation θ au point x , ε est une petite constante positive.

$$A_{n,\theta_j}(x) = \sqrt{I_{e_{n,\theta_j}}(x)^2 + I_{s_{n,\theta_j}}(x)^2} \quad (9)$$

$E_{\theta_j}(x)$ est l'énergie locale du point x selon la direction θ_j ; $E_{\theta_j}(x) = \sqrt{F_{\theta_j}(x)^2 + H_{\theta_j}(x)^2}$

$F_{\theta_j} = \sum_n I_{e_{n,\theta_j}}(x)$ et $H_{\theta_j} = \sum_n I_{s_{n,\theta_j}}(x)$. I_e et I_s sont respectivement l'image origine et l'image améliorée.

- L'amplitude du gradient :

Elle peut être calculée en utilisant le produit de convolution entre les images et l'un des masques de détection du gradient : Opérateur de Sobel, Opérateur de Prewitt et celui de Scharr.

L'amplitude du gradient est donnée par : $G = \sqrt{G_x^2 + G_y^2}$ ou G_x et G_y sont respectivement le gradient dans le sens horizontal et vertical par application de l'un des opérateurs indiqué en haut.

- Les composantes du paramètre FSIM :

On en déduit alors les deux composantes de la mesure FSIM entre $I_e(x)$ et $I_s(x)$; une relative à PC et l'autre relative à GM . La mesure de similarité pour PC_1 et PC_2 est définie comme suit :

$$S_{PC}(x) = \frac{2PC_1(x).PC_2(x) + T_1}{PC_1^2(x) + PC_2^2(x) + T_1}$$

T_1 est une constante positive utilisée afin d'augmenter la stabilité de S_{PC} , elle dépend de l'intervalle au quelle appartient les valeurs PC .

De même, pour la composante GM la mesure de similarité pour $G_1(x)$ et $G_2(x)$ est définie comme suit :

$$S_G(x) = \frac{2G_1(x).G_2(x) + T_2}{G_1^2(x) + G_2^2(x) + T_2}$$

Où T_2 est une constante positive qui dépendant de l'intervalle au quelle appartient les valeurs de GM .

Ensuite, la similarité relative à PC et celle relative à GM sont combinés pour obtenir la similarité $S_L(x)$ de $I_e(x)$ et $I_s(x)$. Elle est définie comme suit :

$$S_L(x) = [S_{PC}(x)]^\alpha \cdot [S_G(x)]^\beta$$

Où α et β sont deux paramètres utilisés pour ajuster l'importance relative des caractéristiques des composante PC ou GM . $\alpha = \beta = 1$ est souvent utilisable.

Etant donné que le HVS est sensible aux structures à phase congruente, une valeur élevée de PC à une position x donnée implique que cette dernière aura un impact élevé sur le HSV. L'utilisation des positions où les valeurs PC sont élevées dans l'évaluation de la similarité entre les deux images $I_e(x)$ et $I_s(x)$ sera plus significatif. Par conséquent l'expression de PC sera :

$$PC_m(x) = \max(PC_1(x), PC_2(x))$$

L'expression $PC_m(x)$ est utilisée afin de pondérer l'importance de $S_L(x)$ dans la similarité globale entre les deux images $I_e(x)$ et $I_s(x)$. L'expression du paramètre $FSIM$ entre deux images est alors :

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\sum_{x \in \Omega} PC_m(x)}$$

Où Ω est l'ensemble du domaine spatial de l'image.

Bibliographie :

- [1] world health Organization, “Cancer Today.” <https://gco.iarc.fr/today/online-analysis-multi-bars> (accessed Mar. 04, 2021).
- [2] Ministère de la Santé du Maroc, “Guide de détection précoce Des cancers Du Sein et du Col de l’utérus,” pp. 18–25, 2011, [Online]. Available: http://contrelecancer.ma/site_media/uploaded_files/Guide_de_detection_preücoce_des_cancers_du_sein_et_du_col_de_luteürus.pdf.
- [3] Haute Autorité de Santé, *Dépistage et prévention du cancer du sein Mini-synthèse Description anatomique des seins*. 2015.
- [4] S. Lanz and I. Fleischmann, *Le cancer du sein Carcinome mammaire*. 2015.
- [5] R. Vairavan, O. Abdullah, P. B. Retnasamy, Z. Sauli, M. M. Shahimin, and V. Retnasamy, “A Brief Review on Breast Carcinoma and Deliberation on Current Non Invasive Imaging Techniques for Detection,” *Curr. Med. Imaging Former. Curr. Med. Imaging Rev.*, vol. 15, no. 2, pp. 85–121, 2017, doi: 10.2174/1573405613666170912115617.
- [6] S. Sapate and S. Talbar, *An Overview of Pectoral Muscle Extraction Algorithms Applied to Digital Mammograms*. 2016.
- [7] D. A. Spak, J. S. Plaxco, L. Santiago, M. J. Dryden, and B. E. Dogan, “BI-RADS® fifth edition: A summary of changes,” *Diagn. Interv. Imaging*, vol. 98, no. 3, pp. 179–190, 2017, doi: 10.1016/j.diii.2017.01.001.
- [8] A. Gandhamal, S. Talbar, S. Gajre, A. F. M. Hani, and D. Kumar, “Local gray level S-curve transformation -- A generalized contrast enhancement technique for medical images,” in *Computers in Biology and Medicine*, 2017, vol. 83, pp. 120–133, doi: 10.1016/j.combiomed.2017.03.001.
- [9] S. Z. Ramadan, “Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review,” *J. Healthc. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/9162464.
- [10] M. J. G. Calas, B. Gutfilen, and W. C. de A. Pereira, “cad and mammography: Why use this tool?,” *Radiol. Bras.*, vol. 45, no. 1, pp. 46–52, 2012, doi: 10.1590/S0100-39842012000100011.
- [11] W. E. Fathy and A. S. Ghoneim, “A deep learning approach for breast cancer mass detection,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 1, pp. 175–182, 2019, doi: 10.14569/IJACSA.2019.0100123.
- [12] M. S. Mabrouk, H. M. Afify, and S. Y. Marzouk, “Fully automated computer-aided

- diagnosis system for micro calcifications cancer based on improved mammographic image techniques,” *Ain Shams Eng. J.*, vol. 10, no. 3, pp. 517–527, 2019, doi: 10.1016/j.asej.2019.01.009.
- [13] J. Dheebea, N. Albert Singh, and S. Tamil Selvi, “Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach,” *J. Biomed. Inform.*, vol. 49, pp. 45–52, 2014, doi: 10.1016/j.jbi.2014.01.010.
- [14] N. I. R. Yassin, S. Omran, E. M. F. El Houby, and H. Allam, “Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review,” *Comput. Methods Programs Biomed.*, vol. 156, pp. 25–45, 2018, doi: 10.1016/j.cmpb.2017.12.012.
- [15] J. M. Seely and T. Alhassan, “Screening for breast cancer in 2018—what should we be doing today?,” *Curr. Oncol.*, vol. 25, no. June, pp. S115–S124, 2018, doi: 10.3747/co.25.3770.
- [16] B. Vikrant *et al.*, *Haralick Features-Based Classification of Mammograms Using SVM*, vol. 1. Springer Singapore, 2018.
- [17] C. éditorial pédagogique de L’UVMaF, “Anatomie De La Glande Mammaire,” *Rev. Fr. Gynecol. Obstet.*, vol. 79, no. 3, pp. 163–169, 1984.
- [18] F. A, C. B, D. E, and L. L, *Chirurgie du cancer du sein.*, 2011th ed., vol. 110, no. 4. Elsevier Masson SAS, 1990.
- [19] *Les traitements des cancers du sein, collection Guide patients*, vol. 27, no. 146. 2013.
- [20] P. Hubert, “Généralités sur les tumeurs 26/04/2018,” 2018.
- [21] C. F. des Pathologistes, “Item 309 (ex item 159) – Tumeurs du sein,” vol. 308, 2013.
- [22] Institut National du Cancer, “Accélérons les progrès face aux cancers,” 2016. <https://www.e-cancer.fr/Comprendre-prevenir-depister/Qu-est-ce-qu-un-cancer> (accessed Aug. 04, 2020).
- [23] R. Pr, “LES CANCERS DU SEIN Recherche Prévention et promotion du dépistage Actions pour les personnes malades Société et politiques de santé (SPS),” 2018, [Online]. Available: <https://www.ligue-cancer.net/sites/default/files/brochures/cancers-sein-2018-01.pdf>.
- [24] L. Cambier, “Les incidences complémentaires en mammographie,” *Imag. la Femme*, vol. 16, no. 1, pp. 4–12, 2006, doi: 10.1016/s1776-9817(06)73022-x.
- [25] R. Physiquetechnologie and D. E. S. Mammographes, “Techniques d ’ exploration radiologique du sein (mammographie , échographie , imagerie par résonance magnétique),” 2001.
- [26] U. D. Evry and V. A. L. D. Essonne, “Description et classification des masses

- mammaires pour le diagnostic du cancer du sein Imen Cheikhrouhou To cite this version : HAL Id : tel-00875976 Docteur de l ' Universit ' Description et classification des masses mammaires pour le diagnostic du canc,” 2013.
- [27] “Tomosynthèse mammaire Considérations pour l ' utilisation en routine clinique.”
- [28] ACMTS, “Notes sur les technologies de la santé en émergence,” 2017.
- [29] C. Dromain, “Nouvelles techniques d ' imagerie : à qui s ' adressent-elles ?,” 2003.
- [30] RSNA, “Rsna 2017,” 2017. [Online]. Available: <https://rsna2017.rsna.org/>.
- [31] Y. B. E. N. Youssef, “Etude des Techniques de Classification Automatique des Tumeurs du Sein sur les Images Mammographiques,” 2018.
- [32] A. Athanasiou and C. Balleyguier, “Nouveautés techniques en échographie mammaire,” pp. 247–254, 2007.
- [33] CEA, “Livret Pédagogique L ' Imagerie Médicale,” *Livret pédagogique*, 2017.
- [34] Élisabeth Luporsi and L. Leichtnam-Dugarin, “Comprendre le Cancer du sein,” 2007.
- [35] S. V. Destounis, A. Santacroce, and A. Arieno, “Update on breast density, risk estimation, and supplemental screening,” *Am. J. Roentgenol.*, vol. 214, no. 2, pp. 296–305, 2020, doi: 10.2214/AJR.19.21994.
- [36] T. Carlile *et al.*, “Breast Cancer Prediction and the Wolfe Classification of Mammograms,” *JAMA J. Am. Med. Assoc.*, vol. 254, no. 8, pp. 1050–1053, 1985, doi: 10.1001/jama.1985.03360080062030.
- [37] H. Berment, V. Becette, M. Mohallem, F. Ferreira, and P. Chérel, “Les masses en mammographie : quelles lésions anatomopathologiques sous-jacentes ?,” *J. Radiol. Diagnostique Interv.*, vol. 95, no. 2, pp. 126–135, 2014, doi: 10.1016/j.jradio.2013.12.006.
- [38] I. N. Bankman, *Handbook of Medical Image Processing and Analysis*. 2009.
- [39] M. S. Nixon and A. S. Aguado, *Feature Extraction and Image Processing*, vol. 4, no. 1. 2016.
- [40] H. Behzadi-khormouji *et al.*, “Deep learning, reusable and problem-based architectures for detection of consolidation on chest X-ray images,” *Comput. Methods Programs Biomed.*, vol. 185, p. 105162, 2020, doi: 10.1016/j.cmpb.2019.105162.
- [41] Z. Sha, L. Hu, and B. D. Rouyendegh, “Deep learning and optimization algorithms for automatic breast cancer detection,” *Int. J. Imaging Syst. Technol.*, no. December 2019, pp. 1–12, 2020, doi: 10.1002/ima.22400.
- [42] M. A. Ebrahimighahnavieh, S. Luo, and R. Chiong, “Deep learning to detect Alzheimer’s disease from neuroimaging: A systematic literature review,” *Comput. Methods Programs Biomed.*, vol. 187, p. 105242, 2020, doi: 10.1016/j.cmpb.2019.105242.

- [43] R. Rabidas, A. Midya, J. Chakraborty, and W. Arif, "A Study of Different Texture Features Based on Local Operator for Benign-malignant Mass Classification," *Procedia Comput. Sci.*, vol. 93, no. September, pp. 389–395, 2016, doi: 10.1016/j.procs.2016.07.225.
- [44] D. Saraswathi and E. Srinivasan, "Performance analysis of mammogram CAD system using SVM and KNN classifier," *Proc. Int. Conf. Inven. Syst. Control. ICISC 2017*, pp. 1–5, 2017, doi: 10.1109/ICISC.2017.8068653.
- [45] A. A. Kayode, N. O. Akande, A. A. Adegun, and M. O. Adebisi, "An automated mammogram classification system using modified support vector machine," *Med. Devices Evid. Res.*, vol. 12, pp. 275–284, 2019, doi: 10.2147/MDER.S206973.
- [46] A. Jalalian *et al.*, "FOUNDATION AND METHODOLOGIES IN COMPUTER-AIDED DIAGNOSIS SYSTEMS FOR BREAST CANCER DETECTION," *EXCLI J.*, vol. 16, pp. 113–137, 2017.
- [47] Gonzales, "Digital image processing using MATLAB," *Igarss 2014*, no. 1, pp. 1–5, 2014.
- [48] A. McAndrew, "An Introduction to Digital Image Processing with MATLAB," *Notes SCM2511 Image Process.*, pp. 1–264, 2004.
- [49] D. Jean-Philippe, M. Elisabeth, and C. Zorn, *GUIDE des technologies de l'imagerie médicale et de la radiothérapie Quand la théorie éclaire la pratique.* .
- [50] R. C. Gonzalez and R. E. Woods, *digital-image-processing-2ed--2002.pdf*. 2002.
- [51] G. DOUGHERTY, *Digital Image Processing for Medical Applications*. 2009.
- [52] A. K. Jain, *Fundamentals of Digital Image Processing*, vol. 14, no. 8. 1989.
- [53] MathWorks, *Image Processing Toolbox For Use with M ATLAB. Computation, Visualization, Programming*. 2002.
- [54] M. Q. Khairuzzaman, *The Essential Guide to Image Processing*, vol. 4, no. 1. 2016.
- [55] B. Image, *Biosignal and Biomedical Image*. 2004.
- [56] E. Durand and E. Blondiaux, *Caractéristiques d'une image médicale*, 2017th ed. Elsevier Masson SAS.
- [57] A. Chetouani, "Aladine Chetouani Vers un syst `eme d `evaluation de la qualit `e d `image multi-crit `eres," 2011.
- [58] N. HALLTIERE, D. AUBERT, and M. JOURLIN, "Mesure du contraste local dans les images, application à la mesure de distance de visibilité par caméra embarquée," *Trait. Du Signal*, vol. 23, no. 2, pp. 145–158, 2006.
- [59] J. V.L and R. Gopikakumari, "IEM: A New Image Enhancement Metric for Contrast and Sharpness Measurements," vol. 79, no. 9, pp. 1–9, 2013, doi: 10.5120/13766-1620.

- [60] S. Kumari and P. Lehana, “Image Content Enhancement of Natural Images Using Genetic Algorithm,” *Phys. Sci. Int. J.*, vol. 4, no. 9, pp. 1244–1259, 2014, doi: 10.9734/psij/2014/6891.
- [61] A. Chourasiya and N. Khare, “A Comprehensive Review Of Image Enhancement Techniques,” *Int. J. Innov. Res. Growth*, vol. 8, no. 6, pp. 8–13, 2019, doi: 10.26671/ijirg.2019.6.8.101.
- [62] N. A. Al-saiyd and S. M. Talafha, “Mammographic Image Enhancement Techniques-A Survey,” *Int. J. Comput. Sci. Inf. Technol. Secur. (IJCSITS)*, vol. 5, no. February, pp. 200–206, 2015.
- [63] J. Dabass, S. Arora, R. Vig, and M. Hanmandlu, “Mammogram Image Enhancement Using Entropy and CLAHE Based Intuitionistic Fuzzy Method,” *2019 6th Int. Conf. Signal Process. Integr. Networks, SPIN 2019*, pp. 24–29, 2019, doi: 10.1109/SPIN.2019.8711696.
- [64] N. Hassan, S. Ullah, N. Bhatti, H. Mahmood, and M. Zia, “The Retinex based improved underwater image enhancement,” *Multimed. Tools Appl.*, 2020, doi: 10.1007/s11042-020-09752-2.
- [65] W. G. Flores and W. C. de A. Pereira, “A contrast enhancement method for improving the segmentation of breast lesions on ultrasonography,” *Comput. Biol. Med.*, vol. 80, no. November 2016, pp. 14–23, 2017, doi: 10.1016/j.combiomed.2016.11.005.
- [66] K. Akila, L. S. Jayashree, and A. Vasuki, “Mammographic Image Enhancement Using Indirect Contrast Enhancement Techniques – A Comparative Study,” *Procedia Comput. Sci.*, vol. 47, pp. 255–261, 2015, doi: 10.1016/j.procs.2015.03.205.
- [67] S. Shashi B. Rana, “A Review of Medical Image Enhancement Techniques for Image Processing,” *Int. J. Curr. Eng. Technol.*, vol. 5, no. 2, pp. 1282–1286, 2011, doi: 10.14741/ijcet/22774106/5.2.2015.121.
- [68] A. Gandhamal, S. Talbar, S. Gajre, A. F. M. Hani, and D. Kumar, “A generalized contrast enhancement approach for knee MR images,” *2016 Int. Conf. Signal Inf. Process. IConSIP 2016*, 2017, doi: 10.1109/ICONSIP.2016.7857499.
- [69] E. Daniel and J. Anitha, “Optimum wavelet based masking for the contrast enhancement of medical images using enhanced cuckoo search algorithm,” *Comput. Biol. Med.*, vol. 71, pp. 149–155, 2016, doi: 10.1016/j.combiomed.2016.02.011.
- [70] S. Gupta, “Review of Different Local and Global Contrast Enhancement Techniques for a Digital Image,” vol. 100, no. 18, pp. 18–23, 2014.
- [71] P. V. Oak and P. R. S. Kamathe, “Contrast Enhancement of brain MRI images using histogram based techniques,” vol. 1, no. 3, pp. 90–94, 2013.

- [72] S. Lal and M. Chandra, "Efficient algorithm for contrast enhancement of natural images," *Int. Arab J. Inf. Technol.*, vol. 11, no. 1, pp. 95–102, 2014.
- [73] B. C. Patel and G. R. Sinha, "Gray level clustering and contrast enhancement (GLC–CE) of mammographic breast cancer images," *CSI Trans. ICT*, vol. 2, no. 4, pp. 279–286, 2015, doi: 10.1007/s40012-015-0062-z.
- [74] J. Ma, X. Fan, S. X. Yang, X. Zhang, and X. Zhu, "Contrast Limited Adaptive Histogram Equalization Based Fusion for Underwater Image Enhancement," no. March, pp. 1–27, 2017, doi: 10.20944/preprints201703.0086.v1.
- [75] A. Papadopoulos, D. I. Fotiadis, and L. Costaridou, "Improvement of microcalcification cluster detection in mammography utilizing image enhancement techniques," *Comput. Biol. Med.*, vol. 38, no. 10, pp. 1045–1055, 2008, doi: 10.1016/j.combiomed.2008.07.006.
- [76] V. Bhateja, M. Misra, and S. Urooj, "Unsharp masking approaches for HVS based enhancement of mammographic masses : A comparative evaluation," *Futur. Gener. Comput. Syst.*, 2017, doi: 10.1016/j.future.2017.12.006.
- [77] M. Kciuk, K. Chwastek, K. Kluszczyński, and J. Szczygłowski, "A study on hysteresis behaviour of SMA linear actuators based on unipolar sigmoid and hyperbolic tangent functions," *Sensors Actuators, A Phys.*, vol. 243, pp. 52–58, 2016, doi: 10.1016/j.sna.2016.02.012.
- [78] . B. B., "a Comparative Study on Classification of Image Segmentation Methods With a Focus on Graph Based Techniques," *Int. J. Res. Eng. Technol.*, vol. 03, no. 15, pp. 310–315, 2014, doi: 10.15623/ijret.2014.0315060.
- [79] V. Chaurasia and C. Agarwal, "A Survey Of K-Means And GA-KM The Hybrid Clustering Algorithm," *Int. J. Sci. Technol. Res.*, vol. 3, no. 6, pp. 119–122, 2014.
- [80] *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*. 2011.
- [81] S. De, S. Bhattacharyya, S. Chakraborty, and P. Dutta, "Image Segmentation: A Review," vol. 1, no. 4, pp. 29–40, 2016, doi: 10.1007/978-3-319-47524-0_2.
- [82] A. Norouzi *et al.*, "Medical image segmentation methods, algorithms, and applications," *IETE Tech. Rev. (Institution Electron. Telecommun. Eng. India)*, vol. 31, no. 3, pp. 199–213, 2014, doi: 10.1080/02564602.2014.906861.
- [83] M. S. Fasihi and W. B. Mikhael, "Overview of Current Biomedical Image Segmentation Methods," *Proc. - 2016 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2016*, pp. 803–808, 2017, doi: 10.1109/CSCI.2016.0156.
- [84] T. Blaschke and C. Burnett, "Remote Sensing Image Analysis: Including the Spatial

- Domain,” *Remote Sens. Image Anal. Incl. Spat. Domain*, no. January, 2005, doi: 10.1007/1-4020-2560-2.
- [85] B. F. Zohra and B. Mohamed, “L ’ Interprétation des Images Médicales Cérébrales : Segmentation par L ’ approche Split et Merge.”
- [86] L. Rundo *et al.*, “Combining split-and-merge and multi-seed region growing algorithms for uterine fibroid segmentation in MRgFUS treatments,” *Med. Biol. Eng. Comput.*, vol. 54, no. 7, pp. 1071–1084, 2016, doi: 10.1007/s11517-015-1404-6.
- [87] J. Martínez, F. F. Rivera, J. C. Cabaleiro, D. L. Vilariño, T. F. Pena, and D. Miranda B., “A rule-based classification from a region-growing segmentation of airborne lidar,” *Image Signal Process. Remote Sens. XXII*, vol. 10004, p. 100040F, 2016, doi: 10.1117/12.2240750.
- [88] D. Withey and Z. Koles, “A review of medical image segmentation: methods and available software,” *Int. J. Bioelectromagn.*, vol. 10, no. 3, pp. 125–148, 2008.
- [89] D. Kaur and Y. Kaur, “Various Image Segmentation Techniques: A Review,” *Int. J. Comput. Sci. Mob. Comput.*, vol. 3, no. 5, pp. 809–814, 2014.
- [90] D. D. Patil, S. G. Deore, and S. Bhusawal, “Medical Image Segmentation: A Review,” *Ijcsmc*, vol. 2, no. 1, pp. 22–27, 2013.
- [91] A. Kefali and T. Sari, “Evaluation de plusieurs techniques de seuillage d’ images de documents arabes anciens,” no. 1, pp. 123–134, 2009.
- [92] S. Pare, A. Kumar, G. K. Singh, and V. Bajaj, “Image Segmentation Using Multilevel Thresholding: A Research Review,” *Iran. J. Sci. Technol. - Trans. Electr. Eng.*, vol. 44, no. 1, 2020, doi: 10.1007/s40998-019-00251-1.
- [93] . R. S. C. B., “Mammogram Image Segmentation Using Rough Clustering,” *Int. J. Res. Eng. Technol.*, vol. 02, no. 10, pp. 66–77, 2013, doi: 10.15623/ijret.2013.0210009.
- [94] Hartono, O. E., and A. D., “Determining a cluster centroid of k-means clustering using genetic algorithm,” *Int. J. Comput. Sci. Softw. Eng.*, vol. 4, no. 6, pp. 160–164, 2015.
- [95] Z. Guellil and L. Zaoui, “Proposition d’une solution au problème d’initialisation cas du k-means,” *CEUR Workshop Proc.*, vol. 547, 2007.
- [96] Z. Hakima and M. Abdelouahab, “Clustering par fusion floue de données appliqué à la segmentation d’images IRM cérébrales,” *CEUR Workshop Proc.*, vol. 547, pp. 1–12, 2007.
- [97] A. N. Benaichouche, H. Oulhadj, and P. Siarry, “Improved spatial fuzzy c-means clustering for image segmentation using PSO initialization, Mahalanobis distance and post-segmentation correction,” *Digit. Signal Process. A Rev. J.*, vol. 23, no. 5, pp. 1390–1400, 2013, doi: 10.1016/j.dsp.2013.07.005.

- [98] H. Zhang, J. E. Fritts, and S. A. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Comput. Vis. Image Underst.*, vol. 110, no. 2, pp. 260–280, 2008, doi: 10.1016/j.cviu.2007.08.003.
- [99] S. Gupta and R. Porwal, "Appropriate Contrast Enhancement Measures for Brain and Breast Cancer Images," vol. 2016, no. 1, 2016, doi: <http://dx.doi.org/10.1155/2016/4710842>.
- [100] B. G. S. Ramesh Jain, Rangachar Kasturi, "MACHINE VISION Ramesh Jain,Rangachar Kasturi,Brian G.Schunck." McGraw-Hill Science/Engineering/Math, 1995.
- [101] M. Prince, "A novel Approach to Gradient Edge Detection Available Online at www.jgrcs.info A NOVEL APPROACH TO GRADIENT EDGE DETECTION," no. October 2014, 2015.
- [102] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988, doi: 10.1007/BF00133570.
- [103] Z. Ma, J. M. R. S. Tavares, R. N. Jorge, and T. Mascarenhas, "A review of algorithms for medical image segmentation and their applications to the female pelvic cavity," *Comput. Methods Biomech. Biomed. Engin.*, vol. 13, no. 2, pp. 235–246, 2010, doi: 10.1080/10255840903131878.
- [104] G. Tsechpenakis, "Deformable Model-Based Medical Image Segmentation Gavriil," in *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, 2011, vol. 1, pp. 33–68, doi: 10.1007/978-1-4419-8195-0.
- [105] D. Mandal, A. Chatterjee, and M. Maitra, "Particle Swarm Optimization Based Fast Chan-Vese Algorithm for Medical Image Segmentation," *Stud. Comput. Intell.*, vol. 704, pp. 49–74, 2017, doi: 10.1007/978-3-662-54428-0_4.
- [106] Yann Collette and P. Siarry, *Optimisation multiobjectif*, vol. 91. 2002.
- [107] B. Sareni, "Méthodes d ' optimisation multimodales associées à la modélisation numérique en électromagnétisme To cite this version : HAL Id : tel-00139202 Bruno SARENI METHODES D ' OPTIMISATION MULTIMODALES," 2007.
- [108] C. M. Anderson-Cook, *Practical Genetic Algorithms (2nd ed.): Randy L. Haupt and Sue Ellen Haupt*, vol. 100. 2005.
- [109] G. Allaire, *Analyse Numerique et Optimisation*. 2005.
- [110] K. Deb and ., *Multi-Objective Optimization Using Evolutionary Algorithms*. Wiley, Chichester, 2001.
- [111] S. R. Singiresu and C. Gables, *Engineering Optimization Theory and Practice*. 2020.
- [112] mbureller, "Thèse Hanaa_Hachimi.pdf."
- [113] R. Zaghdoud, "Hybridation d ' algorithme génétique pour les problèmes des véhicules

intelligents autonomes : applications aux infrastructures portuaires de moyenne taille
Radhia Zaghoud To cite this version : HAL Id : tel-01319833,” 2016.

- [114] O. Hajji, “Contribution au développement de méthodes d’optimisation stochastique, Application à la conception des dispositifs électrotechniques,” 2003.
- [115] T. Illés and G. Lovics, “Approximation of the whole Pareto optimal set for the vector optimization problem,” *Acta Polytech. Hungarica*, vol. 15, no. 1, pp. 127–148, 2018, doi: 10.12700/APH.15.1.2018.1.8.
- [116] D. Brockhoff and E. Zitzler, “Are All Objectives Necessary? On Dimensionality Reduction in Evolutionary Multiobjective Optimization,” *Parallel Probl. Solving from Nat. - PPSN IX, 9th Int. Conf.*, pp. 533–542, 2006.
- [117] R. C. Purshouse and P. J. Fleming, “Conflict, Harmony, and Independence: Relationships in Evolutionary Multi-criterion Optimisation,” *Evol. Multi-Criterion Optim. Second Int. Conf. EMO 2003*, pp. 16–30, 2003.
- [118] D. Brockhoff, “On Objective Conflicts and Objective Reduction in Multiobjective Optimization Motivation (Approximation of) efficient set Generating,” *Gene Expr.*, no. 243, pp. 1–41, 2006.
- [119] H. Sato, K. Tomita, and M. Miyakawa, “Preferred Region Based Evolutionary Multi-objective Optimization Using Parallel Coordinates Interface,” *Proc. - 2015 3rd Int. Symp. Comput. Bus. Intell. ISCBI 2015*, no. 1, pp. 33–38, 2016, doi: 10.1109/ISCBI.2015.13.
- [120] R. L. Pinheiro, D. Landa-Silva, and J. Atkin, “Analysis of Objectives Relationships in Multiobjective Problems Using Trade-Off Region Maps,” *Proc. 2015 Genet. Evol. Comput. Conf. - GECCO '15*, pp. 735–742, 2015, doi: 10.1145/2739480.2754721.
- [121] K. Deb and D. K. Saxena, “On Finding Pareto-Optimal Solutions Through Dimensionality Reduction for Certain Large-Dimensional Multi-Objective Optimization Problems EMO for Many Objectives,” *Kanpur Genet. Algorithms Lab.*, no. KanGAL Report No. 2005011, pp. 3353–3360, 2005, doi: 10.1242/jcs.079046.
- [122] A. López Jaimes, C. A. Coello Coello, and D. Chakraborty, “Objective reduction using a feature selection technique,” *Proc. 10th Annu. Conf. Genet. Evol. Comput. - GECCO '08*, p. 673, 2008, doi: 10.1145/1389095.1389228.
- [123] X. Guo, Y. Wang, and X. Wang, “Using objective clustering for solving many-objective optimization problems,” *Math. Probl. Eng.*, vol. 2013, no. 2, 2013, doi: 10.1155/2013/584909.
- [124] T. Tusar and B. Filipic, “Visualization of Pareto Front Approximations in Evolutionary Multiobjective Optimization: A Critical Review and the Prosection Method,” *IEEE*

- Trans. Evol. Comput.*, vol. 19, no. 2, pp. 225–245, 2015, doi: 10.1109/TEVC.2014.2313407.
- [125] G. Chiandussi, M. Codegone, S. Ferrero, and F. E. Varesio, *Comparison of multi-objective optimization methodologies for engineering applications*, vol. 63, no. 5. Elsevier Ltd, 2012.
- [126] N. Nedjah and L. de M. Mourelle, “Evolutionary multi-objective optimisation: a survey,” *Int. J. Bio-Inspired Comput.*, vol. 7, no. 1, p. 1, 2015, doi: 10.1504/ijbic.2015.067991.
- [127] Y. Dujardin and I. Chad, “Solving multi-objective optimization problems in conservation with the reference point method,” pp. 1–16, 2018.
- [128] S. Cagnoni, E. Lutton, and G. Olague, *Genetic and Evolutionary Computation for Image Processing and Analysis*. .
- [129] D. A. Coley, *An introduction to genetic algorithms for scientists and engineers*. 1998.
- [130] P. Sabatier, “Méthodologie de conception de contrôleurs intelligents par l’approche génétique- application à un bioprocédé,” 2009.
- [131] J. Alliot, “Jean-Marc Alliot 27 octobre 1998,” 1998.
- [132] C. Mathworks, “Global Optimization Toolbox User ’ s Guide R 2018 b,” 2018.
- [133] S. Hashemi, S. Kiani, N. Noroozi, and M. E. Moghaddam, “An image contrast enhancement method based on genetic algorithm,” *Pattern Recognit. Lett.*, vol. 31, no. 13, pp. 1816–1824, 2010, doi: 10.1016/j.patrec.2009.12.006.
- [134] N. Singh and K. V. P. Singh, “Parameter Optimization In Image Enhancement Using PSO,” no. 05, pp. 84–90, 2013.
- [135] A. Horé and D. Ziou, “Image quality metrics: PSNR vs. SSIM,” *Proc. - Int. Conf. Pattern Recognit.*, pp. 2366–2369, 2010, doi: 10.1109/ICPR.2010.579.
- [136] K. Sowjanya and P. R. Kumar, “Gray level image enhancement using nature inspired optimization algorithm : An objective based approach,” vol. 13, no. 1, pp. 66–80, 2017.
- [137] N. Mohanapriya and B. Kalaavathi, “Adaptive image enhancement using hybrid particle swarm optimization and watershed segmentation,” *Intell. Autom. Soft Comput.*, vol. 25, no. 4, pp. 663–672, 2019, doi: 10.31209/2018.100000041.
- [138] Z. Suhail, E. R. E. Denton, and R. Zwiggelaar, “Classification of micro-calcification in mammograms using scalable linear Fisher discriminant analysis,” *Med. Biol. Eng. Comput.*, vol. 56, no. 8, pp. 1475–1485, 2018, doi: 10.1007/s11517-017-1774-z.
- [139] M. G. Ertosun and D. L. Rubin, “Probabilistic visual search for masses within mammography images using deep learning,” *Proc. - 2015 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2015*, pp. 1310–1315, 2015, doi: 10.1109/BIBM.2015.7359868.
- [140] S. Anand, J. Murugachandavel, K. Valarmathi, A. Mano, and N. Kavitha, “Contrast

- Enhancement of Mammograms and Microcalcification Detection,” *Int. J. Recent Technol. Eng.*, vol. 8, no. 4, pp. 3926–3932, 2019, doi: 10.35940/ijrte.f2473.118419.
- [141] N. H. Chan, K. Hasikin, and N. A. Kadri, “An Improved Enhancement Technique for Mammogram Image Analysis: A Fuzzy Rule-Based Approach of Contrast Enhancement,” *Proc. - 2019 IEEE 15th Int. Colloq. Signal Process. its Appl. CSPA 2019*, no. March, pp. 202–206, 2019, doi: 10.1109/CSPA.2019.8696016.
- [142] M. Mustra, M. Grgic, and R. M. Rangayyan, “Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms,” *Med. Biol. Eng. Comput.*, vol. 54, no. 7, pp. 1003–1024, 2016, doi: 10.1007/s11517-015-1411-7.
- [143] M. Moghbel, C. Y. Ooi, N. Ismail, Y. W. Hau, and N. Memari, “A review of breast boundary and pectoral muscle segmentation methods in computer-aided detection/diagnosis of breast mammography,” *Artif. Intell. Rev.*, vol. 53, no. 3, pp. 1873–1918, 2020, doi: 10.1007/s10462-019-09721-8.
- [144] B. Mughal, N. Muhammad, M. Sharif, T. Saba, and A. Rehman, “Extraction of breast border and removal of pectoral muscle in wavelet domain,” *Biomed. Res.*, vol. 28, no. 11, pp. 5041–5043, 2017.
- [145] K. Panetta, A. Samani, and S. Agaian, “Choosing the Optimal Spatial Domain Measure of Enhancement for Mammogram Images,” *Int. J. Biomed. Imaging*, vol. 2014, pp. 1–8, 2014, doi: 10.1155/2014/937849.
- [146] Z. Yao, “Image Enhancement Based on Bi-Histogram Equalization with Non-parametric Modified Technology,” 2016, doi: 10.1109/ICPADS.2016.160.
- [147] M. C. Daikha, “Amélioration et Évaluation de la qualité des images,” 2015.
- [148] X. M. and D. Z. Lin Zhang, Lei Zhang, “Correspondence FSIM : A Feature Similarity Index for Image,” *IEEE Trans. IMAGE Process.*, vol. 20, no. 8, pp. 2378–2386, 2011, doi: 10.1109/TIP.2011.2109730.
- [149] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: A feature similarity index for image quality assessment,” *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, 2011, doi: 10.1109/TIP.2011.2109730.
- [150] Y. Y. . Al-najjar and D. C. . Soong, “Comparison of Image Quality Assessment: PSNR, HVS, SSIM, UIQI,” *Int. J. Sci. Eng. Res.*, vol. 3, no. 8, pp. 1–5, 2012, [Online]. Available: <http://www.ijser.org/researchpaper%5CComparison-of-Image-Quality-Assessment-PSNR-HVS-SSIM-UIQI.pdf>.
- [151] K. Panetta, Y. Zhou, S. Agaian, and H. Jia, “Nonlinear unsharp masking for mammogram enhancement,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 6, pp. 918–928, 2011, doi: 10.1109/TITB.2011.2164259.

- [152] B. G. S. Ramesh Jain, Rangachar Kasturi, “Ramesh Jain, Rangachar Kasturi, Brian G. Schunck - Machine vision (1995, McGraw-Hill Science_Engineering_Math).pdf.” McGraw-Hill Science Engineering Math, 1995.
- [153] H.-S. Gan *et al.*, “Medical Image Contrast Enhancement Using Spline Concept: Data from the Osteoarthritis Initiative,” *J. Med. Imaging Heal. Informatics*, vol. 4, no. 4, pp. 511–520, 2014, doi: 10.1166/jmihi.2014.1294.
- [154] M. C. Morrone and R. A. Owens, “Feature detection from local energy,” *Pattern Recognit. Lett.*, vol. 6, no. 5, pp. 303–313, 1987, doi: 10.1016/0167-8655(87)90013-4.
- [155] Z. Wang, Z. Li, W. Lin, and C. Liu, “Image quality assessment based on improved feature similarity metric,” *APSIPA ASC 2011 - Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. 2011*, pp. 442–446, 2011.
- [156] U. Sara, M. Akter, and M. S. Uddin, “Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study,” *J. Comput. Commun.*, vol. 07, no. 03, pp. 8–18, 2019, doi: 10.4236/jcc.2019.73002.
- [157] S. Anand and S. Gayathri, “Mammogram image enhancement by two-stage adaptive histogram equalization,” *Optik (Stuttg.)*, vol. 126, no. 21, pp. 3150–3152, 2015, doi: 10.1016/j.ijleo.2015.07.069.
- [158] T. X. Pham, P. Siarry, and H. Oulhadj, “A multi-objective optimization approach for brain MRI segmentation using fuzzy entropy clustering and region-based active contour methods,” *Magn. Reson. Imaging*, vol. 61, no. May, pp. 41–65, 2019, doi: 10.1016/j.mri.2019.05.009.
- [159] H. Singh and V. Singh, “A Comparative Analysis on Histogram Equalization Techniques for Medical Image Enhancement,” *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 7, no. 6, pp. 364–370, 2017, doi: 10.23956/ijarcse/v7i6/0269.
- [160] C.-M. Chen *et al.*, “Automatic Contrast Enhancement of Brain MR Images Using Hierarchical Correlation Histogram Analysis,” *J. Med. Biol. Eng.*, vol. 35, no. 6, pp. 724–734, 2015, doi: 10.1007/s40846-015-0096-6.
- [161] M. Maleki, “Image Enhancement by using the intelligent optimization algorithm with a combination of bacteria and Particle Swarm Optimization Introduction :,” vol. 2, no. 7, pp. 1–9, 2015.
- [162] A. Gorai and A. Ghosh, “Gray-level image enhancement by particle swarm optimization,” *2009 World Congr. Nat. Biol. Inspired Comput. NABIC 2009 - Proc.*, no. 1, pp. 72–77, 2009, doi: 10.1109/NABIC.2009.5393603.
- [163] S. Agrawal and R. Panda, “An efficient algorithm for gray level image enhancement using cuckoo search,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif.*

- Intell. Lect. Notes Bioinformatics*), vol. 7677 LNCS, pp. 82–89, 2012, doi: 10.1007/978-3-642-35380-2_11.
- [164] D. Pal Singh and A. Khare, “Evolutionary Image Enhancement Using Multi-Objective Genetic Algorithm,” *Int. J. Image, Graph. Signal Process.*, vol. 6, no. 1, pp. 61–67, 2014, doi: 10.5815/ijigsp.2014.01.09.
- [165] H. El malali, A. Assir, M. Harmouchi, M. Rattal, A. Lyazidi, and A. Mouhsen, “Adaptive Local Gray Level Transformation based on Variable S-curve for Contrast Enhancement of Mammogram Images,” in *Embedded Systems and Artificial Intelligence*, 2020, vol. 1076, pp. 671–679, doi: https://doi.org/10.1007/978-981-15-0947-6_63.
- [166] H. El Malali, A. Abdelhadi, B. Vikrant, M. Azeddine, and M. Harmouchi, “A Contrast Enhancement Model for X-Ray Mammograms Using Modified Local s-Curve Transformation Based on Multi-Objective Optimization,” *IEEE Sens. J.*, no. d, pp. 1–13, 2020, doi: 10.1109/JSEN.2020.3028273.
- [167] R. C. Purshouse and P. J. Fleming, “Evolutionary Multi-Criterion Optimisation,” doi: 10.1007/b106458.
- [168] H. El malali, A. Assir, M. Harmouchi, A. Lyazidi, and A. Mouhsen, “Fully Automatic Computer Aided Detection of Breast Cancer based on Genetic Algorithm Optimization,” doi: 10.1109/IRASET48871.2020.9092186.
- [169] H. El Malali *et al.*, “Automatic mammogram image breast abnormality detection and localization based on the combination of k-means and genetic algorithms methods,” *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 1.5 Special Issue, pp. 76–83, 2020, doi: 10.30534/ijatcse/2020/1291.52020.
- [170] J. Nagi, S. A. Kareem, F. Nagi, and S. K. Ahmed, “Detection Using Digital Mammograms,” *IEEE EMBS Conf. Biomed. Eng. Sci.*, no. December, pp. 87–92, 2010.
- [171] I. L. Aroquiaraj and K. Thangavel, “Pectoral Muscles Suppression in Digital Mammograms using Hybridization of Soft Computing Methods,” *CoRR*, vol. abs/1401.0, no. 2, pp. 161–168, 2014, [Online]. Available: <http://arxiv.org/abs/1401.0870>.
- [172] K. S. Camilus, V. K. Govindan, and P. S. Sathidevi, “Pectoral muscle identification in mammograms,” *J. Appl. Clin. Med. Phys.*, vol. 12, no. 3, pp. 215–230, 2011, doi: 10.1120/jacmp.v12i3.3285.
- [173] S. Sreedevi and E. Sherly, “A novel approach for removal of pectoral muscles in digital mammogram,” *Procedia Comput. Sci.*, vol. 46, no. Ict 2014, pp. 1724–1731, 2015, doi: 10.1016/j.procs.2015.02.117.
- [174] P. Juneja, E. J. Harris, and P. M. Evans, “Investigation of the accuracy of breast tissue

- segmentation methods for the purpose of developing breast deformation models for use in adaptive radiotherapy,” *J. Phys. Conf. Ser.*, vol. 489, no. 1, 2014, doi: 10.1088/1742-6596/489/1/012030.
- [175] J. A. Rosado-Toro *et al.*, “Automated Breast Segmentation of Fat and Water MR Images Using Dynamic Programming,” *Acad. Radiol.*, vol. 22, no. 2, pp. 139–148, 2015, doi: 10.1016/j.acra.2014.09.015.
- [176] H. P. Ng, S. H. Ong, K. W. C. Foong, P. S. Goh, and W. L. Nowinski, “Medical image segmentation using k-means clustering and improved watershed algorithm,” *Proc. IEEE Southwest Symp. Image Anal. Interpret.*, vol. 2006, pp. 61–65, 2006, doi: 10.1109/ssiai.2006.1633722.
- [177] N. M. Zaitoun and M. J. Aqel, “Survey on Image Segmentation Techniques,” *Procedia Comput. Sci.*, vol. 65, no. Iccmit, pp. 797–806, 2015, doi: 10.1016/j.procs.2015.09.027.
- [178] M. Paulinas and A. Ušinskas, “A survey of genetic algorithms applications for image enhancement and segmentation,” *Inf. Technol. Control*, vol. 36, no. 3, pp. 278–284, 2007, doi: 10.1.1.120.4391.
- [179] A. V. S. Kumar, “Parallel Implementation of Genetic Algorithm using K-Means Clustering,” vol. 1455, no. 06, pp. 1450–1455, 2012.
- [180] N. Dhanachandra, K. Manglem, and Y. J. Chanu, “Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm,” *Procedia Comput. Sci.*, vol. 54, pp. 764–771, 2015, doi: 10.1016/j.procs.2015.06.090.
- [181] F. A. Zeiser *et al.*, “Segmentation of Masses on Mammograms Using Data Augmentation and Deep Learning,” *J. Digit. Imaging*, 2020, doi: 10.1007/s10278-020-00330-4.