

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

المديرية العامة للبحث العلمي و التطوير التكنولوجي

Direction Générale de la Recherche Scientifique et du Développement Technologique



Projet national de recherche : rapport final

SYSTÈME D'AIDE POUR LE SUIVI ET LE DIAGNOSTIC DES TUMEURS MAMMAIRES (SASDUM)

Chef de projet : MEROUANI Hayet Farida

Affiliation : Laboratoire de Recherche en Informatique (LRI)

Organisme de domiciliation : Université de BADJI Mokhtar Annaba

Organisme pilote : CERIST

Programme national de Recherche : 2011-2013



تقرير نهائي لمشروع البحث Rapport final du projet PNR

I-Identification du projet:

PNR

Technologies de l'Information et de la Communication

1-التعريف بالمشروع

Organisme pilote

CERIST

Domiciliation du projet :

Laboratoire de Recherche en Informatique (LRI), Université de Badji Mokhtar Annaba

Intitulé du projet

عنوان المشروع

Système d'aide pour le suivi et le diagnostic des tumeurs mammaires (SASDUM)

Chercheurs impliqués dans le projet

أعضاء المشروع و المؤسسة المستخدمة

Nom et prénom الاسم و اللقب	Grade الرتبة	Etablissement employeur المؤسسة المستخدمة	Observation
Merouani Hayet Farida	MCA	Université Badji-Mokhtar Annaba	Actif
Beledjhem Nedjmeddine	MA	Hopital Ibn Roched CHU Annaba	Non actif
Tilili Yamina	MCA	Université Badji-Mokhtar Annaba	Actif
Layachi Soumia	MCB	Université Badji-Mokhtar Annaba	Actif
Melouah Ahlem	MCB	Université Badji-Mokhtar Annaba	Actif

Déroulement du projet :

Rappeler brièvement les objectifs du projet et les taches prévues

تذكير مختصر بأهداف المشروع و المهام المسطرة :

L'objectif de ce projet est de concevoir un système d'aide au diagnostic et au suivi du cancer du sein avant et après un traitement par chimiothérapie. La réalisation de ce système passe par plusieurs phases :

1. Phase de conception et de réalisation de la base de données : elle permet l'élaboration d'une base de données pour les patientes de la région d'Annaba.
2. Phase de détection de la tumeur : consiste à développer une approche de détection des zones cancéreuses.
3. Phase de diagnostic : permet de décider si les zones décelées sont malignes ou bénignes.
4. Phase d'indexation : porte sur l'indexation des mammographies.
5. Phase de suivi : assure le suivi de l'évolution de la tumeur lors d'un traitement par chimiothérapie.
6. Phase de validation : vérifie le bon fonctionnement du système en collaboration avec les experts.

Liste des Figures	I
Liste des Tables	III
Introduction Générale	1
Chapitre 1: Présentation du projet	
1. Introduction	3
2. Problématique	3
3. Partenaire socio-économique	4
4. Objectif s du projet	4
5. Démarche à suivre	5
6. Planning des tâches	7
7. Impact prévus	8
8. Conclusion	8
Chapitre 2: Conception et réalisation de la base mammaire	
1. Introduction	9
2. Objectifs de la phase	9
3. Equipe de travail	9
4. Démarche adoptée	10
5. Réalisation	11
6. Résultats obtenus	13
7. Difficultés rencontrées	13
8. Conclusion	13
9. Références bibliographiques	13
Chapitre 3: Détection de la tumeur	
1. Introduction	14
2. Objectifs de la phase	14
3. Equipe de travail	15
4. Démarche adoptée	15
5. Réalisation	16
6. Résultats obtenus	17
7. Conclusion	18
10. Références bibliographiques	18
Chapitre 4: Diagnostic de la tumeur	
1. Introduction	19
2. Objectifs de la phase	19
3. Equipe de travail	20
4. Démarche adoptée	20
5. Réalisation et résultats obtenus	22
6. Conclusion	26
7. Références bibliographiques	26

Chapitre 5: Indexation basée sur les automates cellulaires	
1. Introduction	28
2. Objectifs	28
3. Equipe de travail	29
4. Schéma général de l'approche	29
5. Descripteurs utilisés	30
6. Méthode proposée	31
7. La recherche d'images par le contenu	31
8. Réalisation de SASDUM 4.1	32
9. Analyse des résultats de classification	35
10. Conclusion	36
11. Références bibliographiques	36
Chapitre 6: Indexation basée sur la loi Zipf et Zipf inverse	
1. Introduction	37
2. Equipe de travail	37
3. Démarche adoptée	37
4. Résultats obtenus par la classification	42
5. Conclusion	43
6. Références bibliographiques	43
Chapitre 7: Suivi de l'évolution de la tumeur par une approche non rigide	
1. Introduction	45
2. Problème de la phase de suivie	45
3. Objectif de la phase de suivie	46
4. Suivi basé sur un recalage non rigide	47
4.1. Equipe de travail	47
4.2. Démarche adoptée	47
4.3. Réalisation et résultats obtenus	50
5. Conclusion	52
6. Références bibliographique	52
Chapitre 8: Suivi de l'évolution de la tumeur par une approche rigide	
1. Introduction	54
2. Equipe de travail	54
3. Démarche adoptée	54
4. Réalisation	58
5. Conclusion	59
6. Références bibliographiques	59
Chapitre 9: Récapitulation	
1. Introduction	60
2. Statistiques sur l'évolution du projet	60
3. Statistiques sur les dépenses budgétaires	62
4. Difficultés et obstacles rencontrés	62
5. Retombé scientifique du projet	63
6. Impact social du projet	65
7. Conclusion	65
Conclusion générale	66
Annexes	
• Annexe 1	68
• Annexe 2	69

Chapitre 1: Présentation du projet	
Fig 1.1 Schéma général de la démarche à adopter	6
Chapitre 2: Conception et réalisation de la base mammaire	
Fig 2.1 Diagramme de cas d'utilisation	10
Fig 2.2 Diagramme de classes du modèle orienté objet de la base mammaire	11
Fig 2.3 Architecture générale de SASDUM 1	11
Fig 2.4 Fiche utilisateur de SASDUM1	12
Fig 2.5 Fiches Malade de SASDUM1	12
Fig 2.6 Fiche d'image médicale de SASDUM1	13
Chapitre 3: Détection de la tumeur	
Fig 3.1 Démarche de la phase détection de la tumeur	15
Fig 3.2 Vue de l'application SASDUM 2	16
Fig 3.3 Vue d'un résultat de l'application SASDUM 2	17
Fig 3.4 Résultats de l'application de détection de tumeur de SASDUM 2	17
Chapitre 4: Diagnostic de la tumeur	
Fig.4.1 Architecture du réseau bayésien proposé	21
Fig.4.2. Réseau bayésien modalisant sept maladies des seins	22
Fig.4.3 Table des cas retenus pour l'évaluation	23
Fig.4.4 Exemple d'une table de probabilité conditionnelle	23
Fig.4.5 Fiche d'exploitation pour la saisie des données	24
Fig.4.6 Réseau bayésien avec action de maintenance	25
Chapitre 5: Indexation basée sur les automates cellulaires	
Fig.5.1 Schéma générale du système de classification et d'indexation des mammographies SASDUM4.1	29
Fig.5.2 Distance de calcul des matrices de cooccurrence vs orientation	30
Fig.5.3 Interface du système	33
Fig.5.4 Paramètres d'Haralick extraites de la matrice de co-occurrence	33
Fig.5.5 Exemple d'un motif de texture	33
Fig.5.6 Vecteur descripteur de la texture	33
Fig.5.7 Matrice de similarité	34
Fig.5.8 Classification par l'automate cellulaire élaboré	34
Fig.5.9 Visualisation d'une image maligne	34
Fig.5.10 Requête entrée par l'utilisateur	35
Fig.5.11 Images similaires trouvées	35
Chapitre 6: Indexation basée sur la loi Zipf et Zipf inverse	
Fig.6.1 Organigramme de la fusion des lois de Zipf et de Zipf inverse pour la détection de la zone d'intérêt.	38
Fig.6.2 Détection de la tumeur sur trois images mammaires par les lois de Zipf	39
Fig.6.3 Courbes de Zipf et de Zipf inverse de trios images mammaires obtenus suite au codage des images mammaires par la fusion du codage de l'écart type et le codage des rangs généraux	40
Fig.6.4 Architecture du système de recherché d'image mammaire par la méthode template Matching	41
Fig 6.5 Acquisition de l'image mammaire requête	42
Fig.6.6 Résultat de recherche des images mammaire les plus similaires à l'image mammaire requête	43

Chapitre 7: Suivi de l'évolution de la tumeur par une approche non rigide

Fig.7.1 Principe de recalage	48
Fig.7.2 Principe générale de notre approche	47
Fig.7.3 Détection multi-échelle des points d'intérêt	49
Fig.7.4 Page principale de SASDUM 5.1	50
Fig. 7.5 Résultat de la déformation rigide de SASDUM 5.1	51
Fig. 7.6 Résultat de SASDUM 5.1 pour	51
Fig. 7.7 Résultat de SASDUM 5.1 après application de GTM	51
Fig. 7.8 Résultat de SASDUM 5.1 après application de la déformation non rigide	52
Fig.7.9 Résultat de détection de la tumeur par SASDUM 5.1	52

Chapitre 8: Suivi de l'évolution de la tumeur par une approche rigide

Fig 8.1 Principe Général de SASDUM 5.2	55
Fig 8.2 Principe de découpage de l'image source et de l'image cible	55
Fig 8.3 Information mutuelle de chaque imagerie	57
Fig 8.4 Vue de l'interface de SASDUM5.2	58
Fig 8.5 Résultat du découpage de (a) l'image source et de (b) l'image cible	58
Fig 8.6 Exemple d'un cycle de recalage de SASDUM 5.2	59

Chapitre 9: Récapitulation

Tab.4.1 Taux avant et après maintenance	26
Tab.5.1 Analyse de la classification	35
Tab.6.1 Descripteurs mammaires	41
Tab.9.1 Statistique d'évolution du projet	61
Tab.9.2 Ventilation budgétaire	62

Introduction générale

«Le cancer du sein est en nette progression et il touche une femme sur dix en Algérie et 9000 nouveaux cas sont enregistrés annuellement dans nos différents centres de santé. Le combat contre cette maladie est une affaire qui doit impliquer tout le monde».

Mohamed Benzerga,
(tiré du journal EL WATAN du 9 février 2010)

Au cours des dernières décennies, le cancer du sein est devenu un problème de santé publique. Le cancer du sein est l'une des causes majeures de mortalités chez les femmes. En Algérie, il représente près de 50 % des cancers gynécologiques. L'incidence du cancer du sein reste croissante en Algérie et il n'existe toujours pas des programmes de dépistage organisé. Du fait de son diagnostic tardif, il en résulte souvent un traitement lourd, mutilant et coûteux qui s'accompagne d'un taux de mortalité élevé. Ces constatations montrent l'importance d'une détection précoce et d'un suivi rigoureux lors du traitement de cette maladie.

La mammographie est l'examen radiographique par excellence pour repérer des anomalies dans les seins des femmes. Elle permet grâce à de faibles doses de rayons X l'obtention de clichés permettant de détecter les cellules cancéreuses dans la glande mammaire.

L'imagerie médicale est, certainement, l'un des axes de la médecine qui a plus progressé ces vingt dernières années. L'imagerie médicale est utilisée dans un but préventif, pour le dépistage d'un certain nombre de pathologies, elle permet aussi un suivi très précis de l'évolution d'une maladie en offrant des comparatifs rationnels.

Un système d'aide à la prise de décision dédiée à l'imagerie médicale est un programme numérique qui consiste à assister le clinicien dans la tâche parfois délicate d'interprétation et de classement des images médicales. Ce domaine est représenté par deux applications principales : les systèmes d'aide à la détection (*Computer Aided Detection* ou *CAD*) et les systèmes d'aide au diagnostic (*Computer Aided Diagnosis* ou *CADs*). Les CAD ont pour rôle de détecter et d'attribuer un score à une ou plusieurs tumeurs dans une image. Les CADs différencient les tumeurs malignes et bénignes parmi celles détectées par les CAD ou fournies par un clinicien. L'objectif de ce projet est de concevoir et de réaliser un système qui en plus des fonctions offertes par un CADs il permet de suivre l'évolution d'une tumeur mammaire tout au long d'un traitement par chimiothérapie.

Le but d'un traitement par chimiothérapie est de diminuer le risque d'apparition de métastases à distance. Le dosage et la nature du médicament prescrit dépend du type de la tumeur et de son stade d'évolution. Les spécialistes appliquent cinq niveaux de traitement qu'ils nomment des protocoles. Un protocole d'un niveau supérieur implique un médicament plus puissant ou un dosage plus important.

L'inconvénient majeur de la chimiothérapie est ses effets secondaires. Ces derniers peuvent être néfastes pour quelques patientes. De ce fait, il faut s'assurer, qu'après quelques séances de chimiothérapie, la tumeur répond soit positivement soit négativement au traitement donné. Cette vérification est possible en réétudiant la forme de la tumeur qui peut évoluer selon les schémas suivants:

1. la taille de la tumeur reste inchangée ou elle devient plus importante: dans cette situation, le traitement est vu comme non efficace et doit être modifier;

2. disparition visuelle de la tumeur: ce cas implique une très bonne réponse de la patiente au traitement;
3. réduction et/ou uniformité du périmètre de la tumeur: cette situation permet d'avancer que la patiente répond assez bien au traitement;
4. plusieurs parties internes de la tumeur disparaissent donnant lieu à plusieurs nids sur l'espace de l'ancienne tumeur: comme dans la situation précédente ce cas implique une assez bonne réponse de la patiente au traitement par chimiothérapie.

La tâche principale du projet SASDUM est de détecter les tumeurs mammaires et de les classer parmi les quatre cas cités précédemment. Vu la difficulté de cette tâche nous avons décomposé ce projet en plusieurs phases. En d'autres termes, le principe « diviser pour régner » est la démarche adoptée par ce projet. Ainsi, le problème de base est décomposé en sous-problèmes plus simples à résoudre. Chaque sous-problème, qui correspond à une phase du projet, est pris en charge par une équipe de travail. La coordination entre les différentes équipes est assurée par le chef de projet qui parallèlement contrôle l'évolution du travail. Par accommodation, nous avons nommé le résultat de chaque phase (l'application réalisée) par le nom du projet suivi par le numéro de la phase. Par exemple, l'application développée à la phase 3 sera nommée SASDUM 3 (partie 3 de SASDUM). Toutefois, deux phases : phase 4 et phase 5, ont été à leurs tours décomposées donnant lieu à des solutions partielles de la phase. Une solution partielle est nommée par le nom de la phase suivi par un numéro d'ordre (exemple SASDUM 5.1 pour la solution 1 de la phase 5 de SASDUM)

Outre l'introduction et la conclusion ce rapport est structuré en neuf chapitres.

Le chapitre 1 rappelle le projet et ses étapes comme il a été établi au démarrage de ce projet.

Le chapitre 2 concerne la conception et réalisation de la base mammaire, ou un prototype de la base à été élaboré.

La détection de la zone cancéreuse par la méthode de croissance de région est présentée dans le chapitre 3.

L'utilisation des réseaux bayésiens pour le diagnostic de la tumeur est donnée dans le chapitre 4.

Outre la détection et le diagnostic, l'indexation de la base mammaire est aussi considérée. L'indexation de la base c'est faite selon deux nouvelles méthodes. Celle du chapitre 5 qui applique les automates cellulaires en tant que classifieur. Celle du chapitre 6 qui applique la fusion des lois de puissance.

Une des plus importantes étapes de ce projet est le suivi de l'évolution de la tumeur cancéreuse avant et après chimiothérapie. Ce suivi de la tumeur se fait selon deux différentes méthodes.

Dans le chapitre 7 nous optons pour faire le suivi de l'évolution de la masse mammaire par le recalage non rigide.

Dans le chapitre 8 nous optons d'assurer le suivi de l'évolution d'une masse mammaire par application d'une démarche de recalage rigide.

L'objectif, la démarche adoptée et les résultats obtenus de chaque sous-problème seront détaillés dans le neuvième et dernier chapitre. De ce fait, ce chapitre récapitule le travail accompli, en donnant quelques statistiques du point de vue évolution et dépenses budgétaire, comme nous soulignons son impact social.

Sommaire du chapitre 1

- 1.1. Introduction
 - 1.2. Problématique
 - 1.3. Partenaire socio-économique
 - 1.4. Objectifs du projet
 - 1.5. Démarche à suivre
 - 1.6. Planning des tâches
 - 1.7. Impact prévus
 - 1.8. Conclusion
-

1.1. Introduction

Lorsqu'il y a présence d'une croissance non contrôlée de cellules anormales, il s'agit du cancer. Le cancer se présente habituellement sous la forme d'une petite bosse. Cependant, le volume de la bosse peut augmenter au fil du temps et se propager aux régions environnantes, telle la peau ou les ganglions lymphatiques sous les bras. La tumeur peut également se propager vers les organes tels que le foie, le cerveau, les poumons et les os. Le cancer du sein est une maladie pouvant être traitée et guérie. La détection précoce est l'élément essentiel de la guérison. Si le cancer du sein est dépisté de manière précoce, avant qu'il ne se propage, il peut être guéri.

La mammographie est une radiographie du sein qui a un rôle capital pour le diagnostic des tumeurs mammaires. C'est une exploration simple, mais qui nécessite un appareillage spécialisé et une interprétation rigoureuse. Elle a trois objectifs : illustrer une maladie mammaire bénigne ou maligne, en surveiller l'évolution sous traitement, révéler des lésions qui échappent à l'examen clinique en raison de leur petite taille ou de leur siège. Elle permet aussi un repérage des lésions avant un geste chirurgical : un fil métallique (harpon) est placé sous contrôle de la mammographie et guidera le chirurgien pour enlever cette zone anormale.

Dans ce contexte et depuis près de vingt ans que notre équipe traite ce domaine d'imagerie médicale, ainsi c'est avec cette forte expérience que nous apportons nos connaissances jumelées à celle des médecins gynécologues et radiologistes afin de proposer d'une part un diagnostic rapide et efficace quand au cancer du sein et d'établir d'autre part une base de données propre à la ville d'Annaba.

1.2. Problématique

Les praticiens font face à des données de plus en plus volumineuses, complexes et hétérogènes : les examens se sont diversifiés, la résolution spatiale et temporelle des données augmente, leur complexité s'accroît. L'exploitation des données sans assistance d'un système de traitement s'avère de plus en plus délicate, voir impossible. L'objectif du traitement d'images médicales est de développer des modèles et des algorithmes pour extraire l'information, afin de contribuer à l'analyser, à l'interpréter et au suivi d'une pathologie.

Un des grands problèmes qui reste posé dans le domaine médical est la détection et le suivi de la maladie du cancer du sein. S'il est encore actuellement difficile d'en déterminer ces causes exactes, il est néanmoins possible de le détecter très tôt, et de rendre ainsi le traitement plus efficace. Le traitement dépend du type de cancer et de son stade d'évolution. Il existe 5 catégories de traitements : la chirurgie, la radiothérapie, la chimiothérapie, la thérapie hormonale et la thérapie ciblée.

L'inconvénient majeur de la chimiothérapie est ces effets secondaires. Ces derniers peuvent être néfastes pour quelques patientes. De ce fait, il faut s'assurer, qu'après quelques séances de traitement par chimiothérapie, la taille de la tumeur se réduit progressivement. Autrement, il est impératif de changer de protocole de traitement.

Il est donc, concevable de concevoir un système capable de suivre les transformations des zones cancéreuses au cours d'un traitement par chimiothérapie. Ce suivi permet aux experts d'appliquer le bon protocole de traitement et d'augmenter ainsi, les chances de guérison du malade.

1.3. Partenaire socio-économique

Notre partenaire socio-économique, officiel est Dr.Beledjhem Nedjmeddine : Maitre assistant à l'hôpital Ibn Rochd de Annaba.

Lors de la première année, du démarrage, de ce projet plusieurs réunions ont eu lieu au sein de l'hôpital. La toute première réunion portée essentiellement sur nos objectifs et la manière d'aborder la problématique. La seconde afin de cerner le matériel utiliser, c'est à dire, le mammogram numérique que l'hôpital venait d'acquérir à cette période. Une fois les tâches déterminées, les rencontres avaient lieu entre Dr.Beledjhem et les membres concernés par la tâche proprement dite.

Pr.Aoures Hayet à aussi collaborée dans ce projet (phase de décision) en permettant d'avoir des données cliniques et de valider les probabilités obtenues. En effet, notre étudiante en Post graduation assistée Pr.Aoures lors de ces consultations cliniques afin d'enregistrer les données réelles sur les patientes.

Lorsque on ne pouvait avoir Dr.Beledjhem, nous étions obligés, afin d'avancer, de récupérer des données de chez Dr.Djilani Nour el Houda : Médecin Spécialiste en radiologie de wilaya de Guelma. Ces images nous servaient de tests afin de valider les méthodes d'indexation utilisées.

1.4. Objectifs du projet

Par ce projet de recherche, nous cherchons à atteindre quatre objectifs à savoir :

1. la conception et la réalisation d'une base de données de mammographies pour la ville d'Annaba et extensible à toutes les régions du pays; cette base doit être propre à notre population afin de permettre d'effectuer des études statistiques et de suivre l'évolution de la maladie dans notre pays ;
2. étudier, analyser et critiquer des méthodes de détection de tumeurs existantes dans le but de dégager une stratégie de diagnostic validée par nos experts ;
3. extraire un échantillon (avec anonymat) de la base et le mettre à la disposition de nos chercheurs;
4. élaborer une stratégie de suivi des transformations des zones cancéreuses au cours d'un traitement par chimiothérapie. Ce suivi nous permettra de vérifier si le traitement est bénéfique ou pas.

1.5. Démarche à suivre

L'organigramme de la Fig 1.1. illustre les grandes phases de la démarche à adopter. Six phases sont nécessaires pour atteindre les objectifs de ce projet. Chaque phase inclue un ensemble d'opérations (actions). Il faut noter que certaines de ces phases nécessitent l'intervention des experts pour la collecte des informations ou la validation des résultats. Dans ce qui suit nous détaillerons chaque phase du projet.

- 1. Phase de conception et de réalisation de la base de données :** toute approche de traitement d'images médicales, plus précisément les mammographies, commence par l'élaboration d'une base de données. Les bases disponibles ne sont pas propre à notre population et pauvre en données nécessaire pour effectuer des études statistique ou élaboré un suivi de la maladie dans notre pays, d'ou l'intérêt de cette phase de ce projet. L'élaboration de la base médicale de mammographies n'est pas une tâche facile vue qu'elle demande un temps non négligeable et des efforts considérables. Plusieurs étapes sont nécessaires pour l'élaboration de la base à savoir:
 - collecte des données auprès des experts, ce qui revient à côtoyer fréquemment ces derniers afin de dégager un cahier de charge selon leurs besoins,
 - analyse des données afin de dégager les différents liens existants,
 - conception de la structure de la base qui doit englober toutes les informations nécessaires pour toutes les étapes du système à réaliser,
 - réalisation concrète de la base,
 - validation de la base selon le cahier de charge par les experts.
- 2. Phase de détection de la tumeur (masse) :** cette phase débute par une étude des différents travaux réalisés pour la détection du cancer en utilisant les mammographies. Plusieurs approches sont envisagées : contours actifs, réseaux bayésiens, croissance de région, ...etc. Cette étude et, en collaboration avec les experts, nous permettra soit d'adapter une approche de détection existante, soit de les combiner ou de concevoir, carrément, une nouvelle approche.
- 3. Phase de diagnostic :** déceler une tumeur dans une mammographie n'implique, forcément, pas qu'elle est dangereuse. Uniquement les tumeurs malignes doivent être traitées les autres qui sont dites bénignes ne sont pas source d'inquiétude. L'objectif de cette phase est de décider si la zone cancéreuse décelée est maligne ou bénigne. A cette fin, divers caractéristiques de la tumeur doivent être extraites (en collaboration avec les experts) ; par la suite, un classifieur, comme le plus proche voisin, l'arbre de décision, réseau de neurones ou bien d'autres, est utilisé pour décider si la tumeur est maligne ou pas.
- 4. Phase d'indexation :** le travail consistera à définir des volumes cibles très précis et réutilisables tout au long du traitement, dans le cadre de la radiothérapie du cancer du sein. Les résultats du traitement médical dépendent fortement de la précision du repositionnement du volume cible entre deux séances. Ces deux points mettent en avant la difficulté d'automatiser complètement la planification d'intervention et la nécessité dans un premier temps de requalifier les images par rapport à des volumes de références dont les contours ont été délimités. L'indexation représente une tâche longue et répétitive pour l'humain, surtout avec les bases d'images qui deviennent aujourd'hui de plus en plus grandes. Ainsi, le besoin de méthodes d'indexation et de recherche directement basée sur le contenu de l'image n'est donc plus à démontrer.
L'objectif final du travail sera d'automatiser le traitement des données issues de l'imagerie afin d'améliorer la qualité et la requalification automatique des zones à traiter.

En utilisant une base d'images de mammographies analysées par les cliniciens du CHU Ibn Rochd, les différentes étapes de l'indexation par contenu mises en évidence lors de nos réunions sont les suivantes:

- analyse de la morphologie des masses détectées dans des mammographies,
- extraction d'attributs de l'image qui consiste à calculer une signature résumant le contenu de l'image,
- indexation d'images par le contenu ou la signature calculer par l'étape précédente sera utilisée pour cette indexation,
- une évaluation statistique de l'intérêt de la délimitation automatique quotidienne des volumes d'intérêt par rapport à une délimitation manuelle moins fréquente sera réalisée.

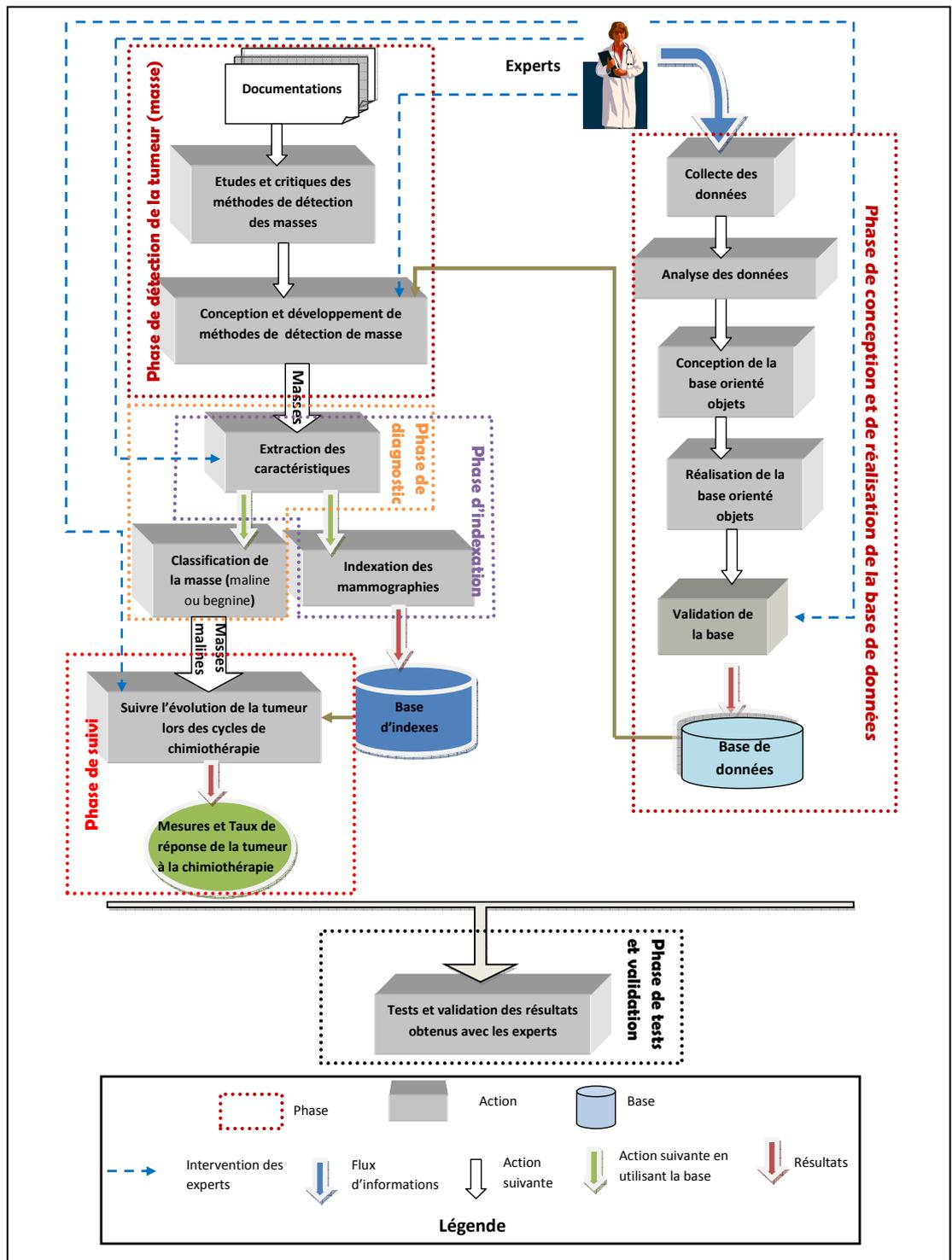


Fig 1.1 Schéma général de la démarche à adopter

5. Phase de suivi : Cette phase consiste à suivre l'évolution de la tumeur après traitement par chimiothérapie. Ce qui revient à comparer les caractéristiques de la tumeur a son stade de détection par celles après traitement. Les caractéristiques sont extraites respectivement des mammographies prise avant et après chimiothérapie. Cette phase inclus plusieurs difficultés. En effet, si une patiente présente plusieurs nids de tumeurs dans ce cas, il est vital de réaliser une projection des deux mammographies afin de correspondre chaque nid de la première mammographie avec celui de la deuxième mammographie. Par conséquent nous seront contraint de faire appel à diverses méthodes à savoir : le calibrage, la détection des points d'intérêts, la mise en correspondance. La réponse de cette phase est une valeur qui indique le taux de réponse de la tumeur au traitement. Cette information est crucial pour le médecin traitent qui en fonction de ce taux décidera de poursuivre le même protocole de traitement ou de le modifier.

6. Phase de tests et de validation : la dernière phase de ce projet consiste à tester et valider le système. Ce dernier s'exercera en parallèle avec le système manuel classique afin de comparer leurs résultats respectifs.

1.6. Planning des tâches

Planning des taches / année				
Tâches	Semestre 1	Semestre 2	Semestre 3	Semestre 4
Phase de conception et de réalisation de la base de données				
1) Collecte des données auprès des experts	←→			
2) Analyse des données	←→			
3) Conception de la base orienté objets	←→			
4) Réalisation de la base orienté objets		←→		
5) Validation de la base		←→		
Phase de détection de la tumeur (masse)				
6) Etudes et critiques des méthodes de détection des masses	←→			
7) Conception et développement de méthodes de détection de masse		←→	←→	
Phase de diagnostic				
8) Extraction des caractéristiques (action commune aux deux phases)			←→	
9) Classification de la masse (maline ou bignine)			←→	
Phase d'indexation				
10) Indexation des images mammographies			←→	
Phase de Suivi				
11) Suivre l'évolution de la tumeur lors des cycles de chimiothérapie			←→	←→
Tests et validation des résultats obtenus avec les experts				
12) Tests et validation des résultats obtenus avec les experts				←→

1.7. Impacts prévus

Les impacts attendus par notre projet de recherche sont :

- **Les retombées sur développement du pays :**

- Apporter une aide précieuse aux médecins en mettant à leur disposition une information vitale qui oriente leurs stratégies de traitement.
- Suivre l'évolution du cancer du sein dans notre pays.
- Avoir une vision sur le taux de guérison et de mortalité à cause de cette maladie.

- **Les retombées sur la formation :**

- Mettre à la disposition des chercheurs une base de données de mammographies propre à notre pays.
- Acquérir des expériences dans le domaine de recherche.
- Enrichir le background des doctorants impliqués dans le projet.
- Concrétiser et valider les travaux de recherche réalisés au sein de l'équipe par des conférences nationales, internationales et des publications.
- Faire soutenir les doctorants dans le respect du délai des thèses

- **les retombées scientifiques :**

Le cancer du sein occupe, aujourd'hui, une place primordiale en cancérologie pour plusieurs raisons. Il représente tout d'abord un défi majeur de santé publique. Il s'inscrit en effet comme le cancer féminin le plus fréquent. Or son dépistage, son diagnostic, son bilan d'extension et son suivi reposent à l'heure actuelle en première intention sur la mammographie et l'échographie mammaire. La mammographie a pour but de déceler au plus tôt des anomalies avant même qu'elles n'aient provoqué des symptômes cliniques. Elle peut permettre, ainsi, de détecter des cancers bien avant qu'ils ne soient palpables. Analyser une mammographie n'est pas tâche facile d'où l'intérêt des CADs (Computer Aided Diagnostic). Par ce projet nous cherchons à apporter un plus en ajoutons la possibilité de suivre les transformations des tumeurs lors d'un traitement par chimiothérapie.

1.8. Conclusion

SASDUM est un projet ambitieux qui cible deux grands objectifs : élaborer une base mammaire propre à la région d'Annaba et, suivre l'évolution du cancer du sein lors d'un traitement par chimiothérapie. Ce projet a été décomposé en plusieurs phases. Chaque phase se penche sur un sous-problème donné. Le logiciel obtenu après intégration des applications résultantes de chaque phase constitue le système final exploitable. La suite de ce rapport détaille chaque phase de SASDUM en commençant par la phase de conception et de réalisation de la base mammaire.

Sommaire du chapitre 2

1. Introduction
 2. Objectifs de la phase
 3. Equipe de travail
 4. Démarche adoptée
 5. Réalisation
 6. Résultats obtenus
 7. Difficultés rencontrées
 8. Conclusion
-

2.1. Introduction

L'imagerie médicale regroupe l'ensemble des techniques utilisées par la médecine pour le diagnostic et le traitement d'un grand nombre de pathologies. Elle a révolutionné la médecine en donnant un accès immédiat et fiable à des informations jusqu'alors invisibles au diagnostic clinique, comme par exemple les caractéristiques anatomiques ou encore, certains aspects du métabolisme des organes. Les techniques d'imagerie médicale ne donnent pas une simple photographie du tissu ou de l'organe étudié mais une représentation visuelle fondée sur des caractéristiques physiques ou chimiques particulières. Toutes ces images acquises doivent être soigneusement étudiées et regroupées pour représenter une référence de comparaison et d'analyse. En d'autres termes, la conception et la réalisation d'une base d'images médical est une phase non discutable pour parler d'un traitement automatique d'images médicales.

2.2. Objectifs de la phase

L'objectif de cette phase peut se résumer en une seule phrase

« Conception et réalisation d'une base mammaires pour la région d'Annaba.»

En réalité cet objectif très simple à formuler incorpore des difficultés importantes. Afin de pouvoir aller progressivement vers sa réalisation nous l'avons répartie en sous objectifs à savoir:

- Collecte des données sur terrain et en collaboration directe avec les spécialistes
- Analyse et organisation des données
- Conception du modèle de la base mammaires
- Discussion avec les experts pour valider le modèle
- Correction du modèle
- Réalisation de la base mammaire
- Validation de la base par les experts
- Saisie des données réelles

2.3. Equipe de travail

Plusieurs étudiants doctorant et en master 2 ont collaboré dans cette phase qui été sous les directives du Dr Layachi S. Ferkous. Benmazou et Hamoud sont les doctorantes qui ont eu des rencontres périodique avec le partenaire socio-économique. Fertikh est l'étudiante en master 2 qui c'est chargée d'implémenter le modèle de la base mammaire conçue.

2.4. Démarche adoptée

En suivant les consignes des experts nous avons opté pour un modèle de base accessible par comptes : compte médecin et compte agent sont les plus important. La figure 2.1 illustre ce principe.

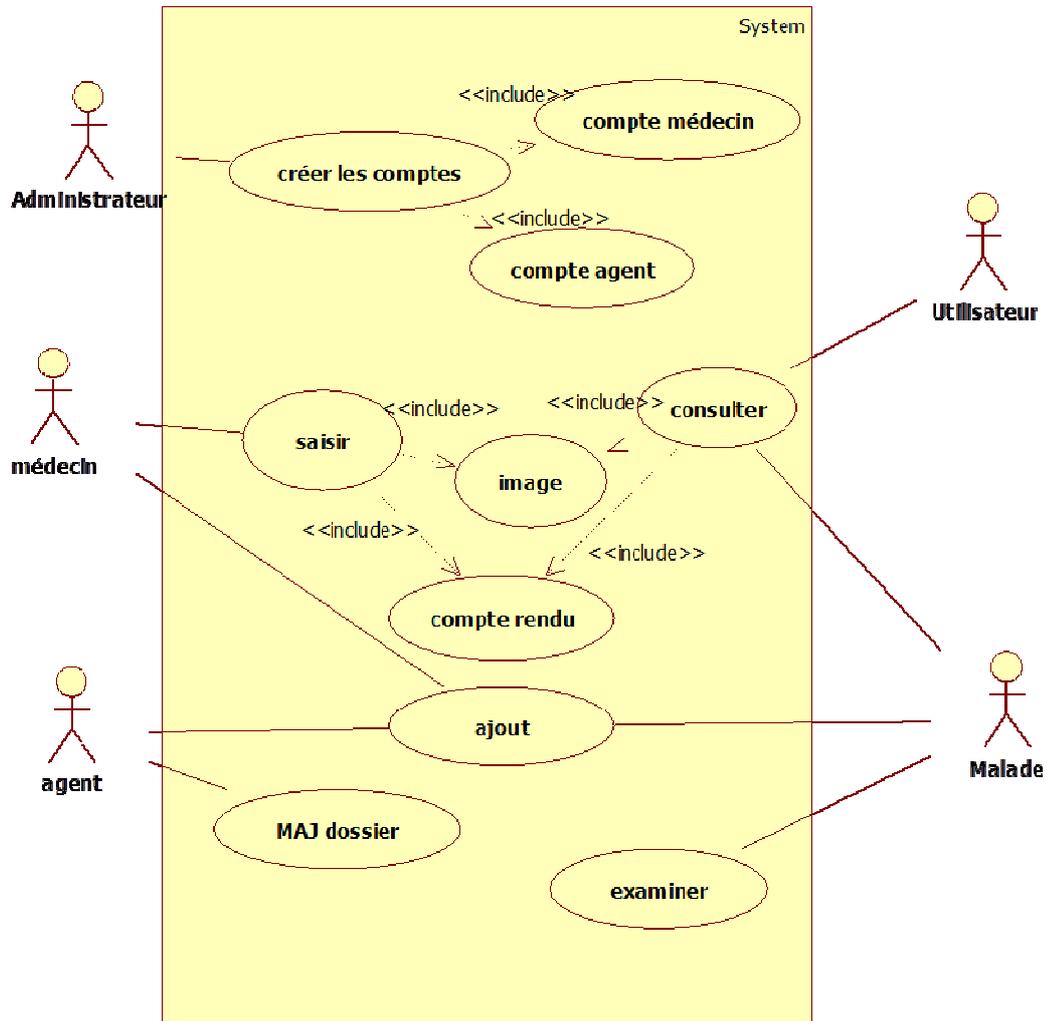


Fig.2.1: Diagramme de cas d'utilisation

Par ailleurs, la base mammaire se base sur le concept orienté objet elle est donc formée d'un ensemble de classes structurées selon le diagramme de la figure 2.2.

Notons que la tâche d'élaboration du diagramme de classes est la tâche la plus importante dans le processus de conception de la base mammaire car, il fourni une vue d'ensemble de la structure de toutes les classes du système avec les relations y afférentes.

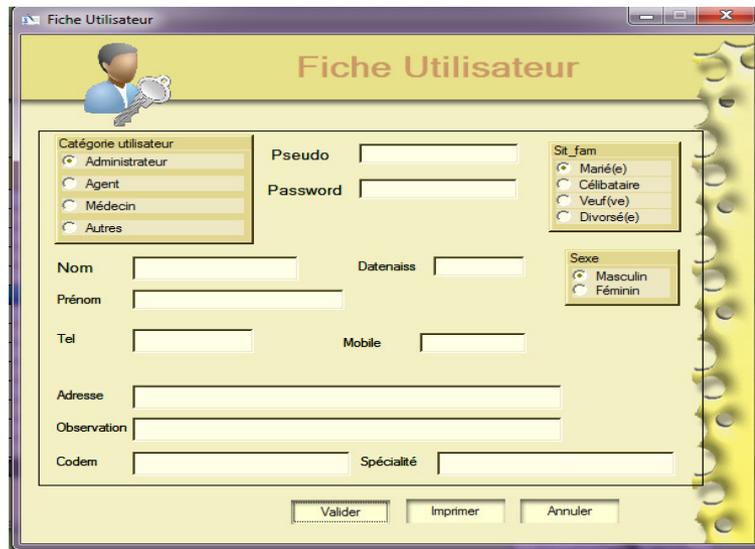


Fig 2.4 Fiche utilisateur de SASDUM1

Concernant les patientes SASDUM1 affiche deux fiches pour chaque malade. La première fiche inclue des informations d'ordre plutôt générale, comme le nom, l'âge l'adresse, ...etc.figure 2.5(a). La deuxième fiche est plus spécifique et cible particulièrement la maladie selon les avis des médecins figure 2.5(b). Bien sùre, afin de respecter la confidentialité des malades les fiches de la figure 2.5 sont remplies d'informations quelconques.



(a)



(b)

Fig 2.5 Fiches Malade de SASDUM1

Enfin SASDUM1 permet de saisir pour chaque malade ces images médicales avec leurs comptes rendus comme le montre la figure 2.6.

La base des images mammaires englobe des images mammaires numérisée; elles ont été acquises au sein de l'hôpital Ibn Rochd à Annaba en Algérie par le radiologue Beledjhem Nedjemeddine. La base des images mammaires est composée de 108 images mammaires : 41 images contenant une tumeur maligne, 49 contenant une tumeur bénigne et 18 sans lésions.

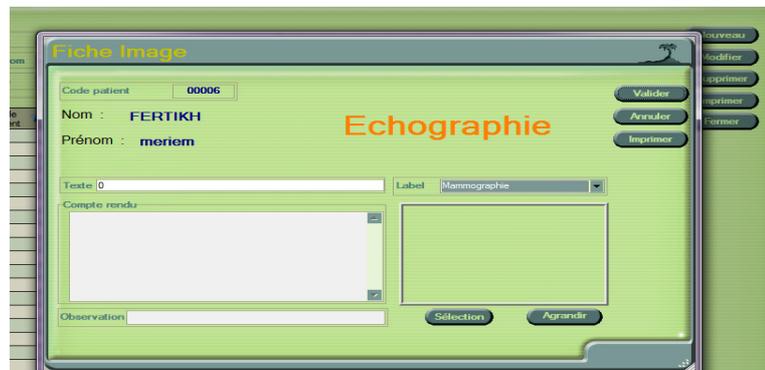


Fig 2.6 Fiche d'image médicale de SASDUM1

2.6. Résultats obtenus

On est dans l'erreur si on annonce qu'on a réalisé une base mammaire pour la ville d'Annaba. Malgré que c'était un objectif de départ. Cependant les difficultés illustrées dans la section suivante ont épuisé notre temps et notre effort dans l'attente de données qui nous arrivées en compte goutte. Ainsi, et pour être exacte nous venons de réaliser un exemplaire très réduit de la base mammaire ciblée au départ.

2.7. Difficultés rencontrée

Nous avons rencontrés plusieurs difficultés qui ont bruité le parcours de cette partie du projet SASDUM. Voici une liste des plus influentes :

- Le nombre d'images fourni par le partenaire est loin de représenter une base de toute la ville d'Annaba.
- Les images sont parfois fournies sans les informations concernant le malade.
- Le partenaire (médecins) ne garde pas la trace du malade, l'image mammographie est toujours chez le patient.
- La difficulté pour les experts de nous donner un peu de leurs temps vu leurs charges de travail.

2.8. Conclusion

La phase de conception et de réalisation de la base mammaire est très importante et elle exige une haute collaboration avec les experts. Cette collaboration est très coûteuse pour ces derniers qui sont hautement sollicité par leur travail. Cette contrainte et divers autres ont constitué des obstacles qui ont ralenti et alourdi la tâche de cette phase. La base mammaire réalisée est une version non complète et d'après les experts demande une révision de plusieurs parties. Nous pensons qu'il fallait plus de temps pour aboutir à un résultat plus satisfaisant.

2.9. Références

- [1] Mullerp-a, Modélisation objet avec UML,eyrolls,1997.
- [2] RTF U,OMG Unified Modeling Langage Spésification , Version 1.3, UML RTF proposed final version OMG June 1997.
- [3] Pierre alain muller, Modelisation objet avec UML, ISBN 2-212-08966, édition eyrolles
- [4] N.Lopez,J.Megueis,E.Pichon , Intégrés UML dans vos projets.
- [5] R. G. Cattell: **The Object Database Management Standard : ODMG-93**, Morgan Kaufmann Publishers, 1994.
- [6] S.R. Dalal and A.A. .Macintosh, «When to stop testmg for large software systems with changing crd », WEE Transactions on software Engineering, Vol. 20, N^o 4, Avril 1994.
- [7] M. Adiba, C. Collet: **Objets et bases de données, le SGBD O2**, Hermes 1993
- [8] American National Standards Institute (ANSI) Database Committee (X3H2): **Database Language SQL3**, J. Melton (Editor), August 1994.
- [9] K. R. Dittrich, U. Dayal, A. P. Buchmann (Eds): **On Object Oriented Database Systems**, Springer Verlag, 1991
- [10] J. Rumbaugh, «Object onented modeling and design », prenticehall,englewo, 1991

Sommaire du chapitre 3

1. Introduction
 2. Objectifs de la phase
 3. Equipe de travail
 4. Démarche adoptée
 5. Réalisation
 6. Résultats obtenus
 7. Conclusion
 8. Référence bibliographiques
-

3.1. Introduction

La phase de détection de la tumeur consiste à cerner la zone suspecte c'est-à-dire la tumeur. A cette fin on doit faire appel à une approche de segmentation. Parmi les approches de segmentation d'image, la segmentation par région occupe un grand espace dans le domaine médical [1-5]. Elle se base sur les propriétés intrinsèques des régions pour extraire une masse donnée. L'intérêt de l'approche de segmentation par croissance de région découle de sa relative simplicité de mise en œuvre et aux résultats intéressants qu'elle peut fournir. Cette méthode qui débute par un ensemble de points initiaux, agrège les pixels suivant un double critère : l'homogénéité et l'adjacence [6,7]. Cette agrégation de pixels est contrôlée par un prédicat de similarité. Le choix de ce prédicat est le point noir de cette approche. En effet, un mauvais choix peut induire à des résultats erronés. Ainsi, il est primordial de bien choisir le prédicat à appliquer.

3.2. Objectifs de la phase

L'algorithme de croissance de régions possède plusieurs avantages mais il possède aussi une grande difficulté qui réside, essentiellement, dans le choix du seuil [8]. Un mauvais seuil peut rendre l'image sur-segmentée ou sous-segmentée et ainsi anéantir toute utilité dans la reconnaissance de forme physiologique. Pour choisir un bon seuil, il existe plusieurs méthodes plus ou moins rigoureuses. La croissance de région la plus simple consiste à agréger deux pixels x et y si leurs intensités $I(x)$ et $I(y)$ sont similaires. Ce critère de similarité $f(\cdot)$ est ensuite utilisé par le prédicat qui la compare à un seuil fixé T . Plus précisément,

$$f(I(x); I(y)) \leq T$$

Cette inégalité définit le contrôle qui autorise ou non la déformation de la carte de segmentation. Par conséquent, les résultats de la segmentation par croissance de région sont directement affectés par le choix du critère de similarité et la valeur du seuil. Ainsi, pour chaque type d'image on doit lui définir un critère de similarité approprié comme on doit lui déterminer la valeur du seuil la plus significative. Plus encore, deux images du même type exigent, parfois, des valeurs de seuil différentes pour obtenir des résultats acceptables. Ainsi, la segmentation par croissance de région est plus performante si elle peut s'adapter à la nature de l'image à traiter. La capacité d'adaptation de la méthode de croissance de région est presque une exigence pour les mammographies qui ont été classées, par des spécialistes, en différentes catégories. Il est donc, intéressant de concevoir une approche de croissance de région qui s'adapte à chaque catégorie de mammographie. L'objectif de ce travail est d'implémenter la partie du système SASDUM qui consiste à détecter les masses dans une mammographie en appliquant une approche de croissance de régions adaptative.

3.3. Equipe de travail

L'équipe de travail est composée de :

- Membre du projet : deux membres du projet sont impliqués dans cette phase à savoir Dr Melouah Ahlem et Ferkous Chokri
- Etudiant chercheur : la partie implémentation était affectée à deux étudiantes en master 2 : Ferhani Imen et Tameur Aicha

3.4. Démarche adoptée

Le processus de croissance de région comprend deux phases : la première pour l'initialisation et la deuxième pour la segmentation. L'initialisation consiste à trouver le point de départ (le germe de la région), dans ce travail le choix du germe est manuel, c'est-à-dire, on doit sélectionner un point qui se trouve dans la région cancéreuse. La segmentation consiste à fixer un seuil T et le comparer avec la valeur de critère de similarité $f(.)$. Cette phase comprend un certain nombre de difficultés. En effet, il est très difficile de trouver le seuil T approprié qui donne une segmentation correcte : un mauvais seuil peut induire à un résultat sur-segmenté ou sous-segmenté et ainsi anéantir toute utilité dans la reconnaissance de forme physiologique.

Partant d'une masse bien cerner par un expert, le système applique l'approche de croissance de région et tente de déterminer les valeurs des paramètres capables de bien localiser la masse selon ses caractéristiques. Par la suite, il cherche les zones suspectes dans une mammographie et l'encercle en appliquant une approche de croissance de régions adaptative.

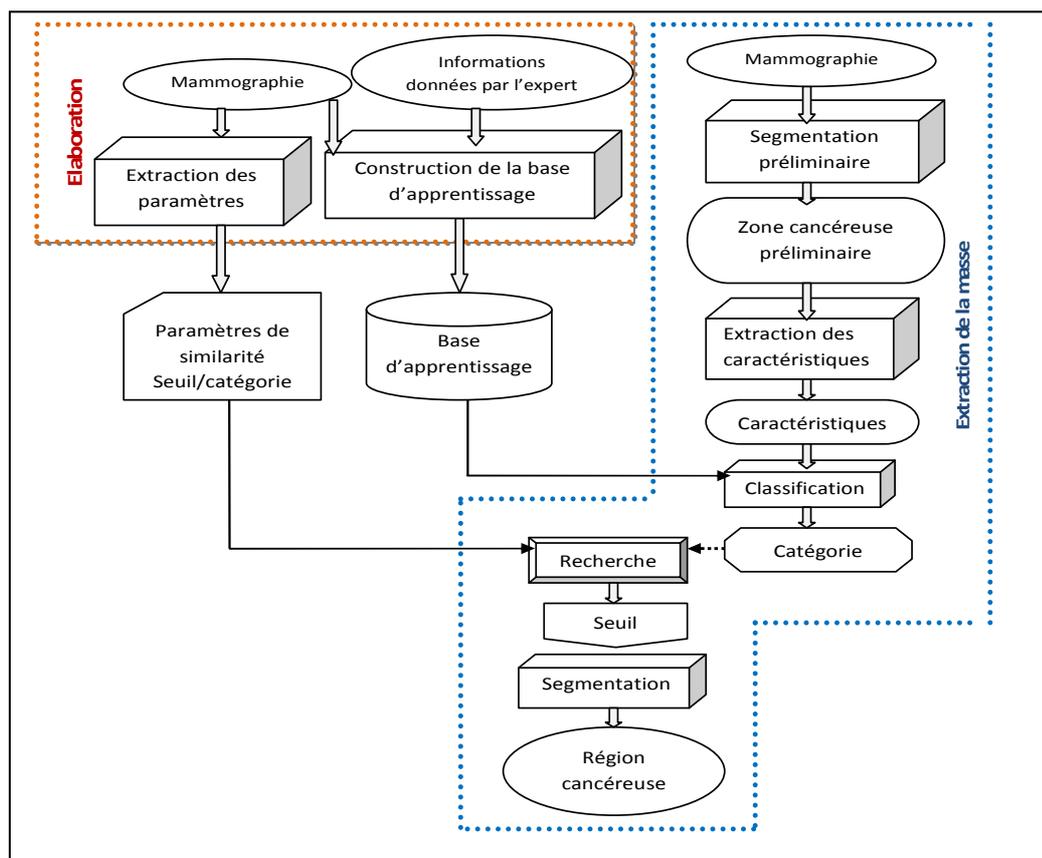


Fig. 3.1: Démarche de la phase détection de la tumeur

L'organigramme de la figure (3.1) illustre les étapes de la démarche de la phase détection de la tumeur. Ce dernier met en évidence deux grandes opérations, la première dite élaboration et la deuxième appelée extraction de la masse.

- *L'élaboration* : est une tâche purement expérimental appliquée une seul fois afin de fixer des paramètres et de construire une base d'apprentissage. L'objectif de **l'extraction des paramètres** est de fixer les paramètres : (i) critère de similarité $f(.)$ (ii) valeur du seuil T qui offre un résultat optimal pour un critère de similarité utilisé. Le but de la **construction de la base d'apprentissage** est de concevoir une base qui doit inclure pour chaque entrée (mammographie) les caractéristiques de la région cancéreuse est la catégorie de l'image à segmenter. Cette dernière information est donnée par l'expert. Par contre l'extraction de la région cancéreuse et le calcul de ses caractéristiques est une tâche du système réalisé. Notons que la tâche d'élaboration consiste à exécuter plusieurs fois la segmentation par croissance de région en partant d'un germe unique mais en modifiant à chaque exécution le critère de similarité et/ou la valeur du seuil.
- *L'extraction de la masse* : est l'étape noyau de ce travail ; elle permet pour une mammographie donnée et en utilisant les résultats de la phase d'élaboration de délimiter la région cancéreuse. Pour une mammographie donnée la démarche proposée commence par la classer dans une des trois catégories. A cette fin, une portion de la zone cancéreuse est cernée par application de l'algorithme de croissance de région. En réalité, la région cancéreuse obtenue est mal cernée mais elle est exploitée pour calculer ces caractéristiques utilisées par la suite par un classifieur afin de déterminer la catégorie du tissu. En déterminant la catégorie du tissu d'une mammographie il sera simple de dégager le critère de similarité à appliquer et la valeur du seuil à utiliser par analogie aux résultats de la phase d'élaboration.

3.5. Réalisation

La partie de SASDUM qui se charge d'assurer la phase de détection de la zone cancéreuse est présentée comme suit :

- Une fenêtre s'affiche permettant à l'expert de charger les mammographies

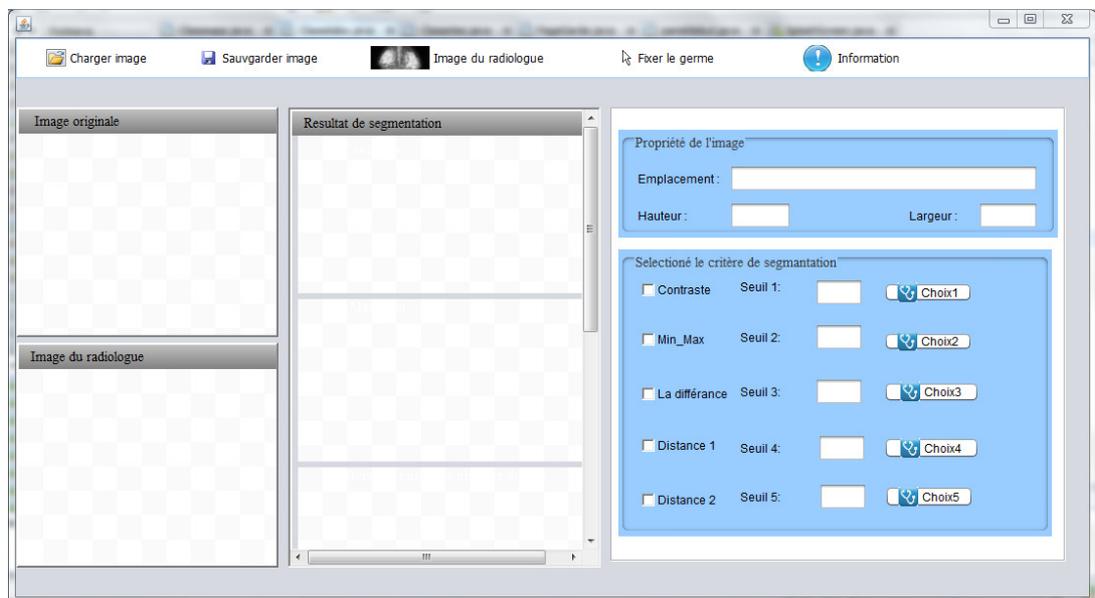


Fig 3.2. Vue de l'application SASDUM 2

Le résultat de la détection de la masse s'affiche à l'expert avec un ensemble de valeurs représentant les caractéristiques calculées et utilisées pour le choix du seuil.

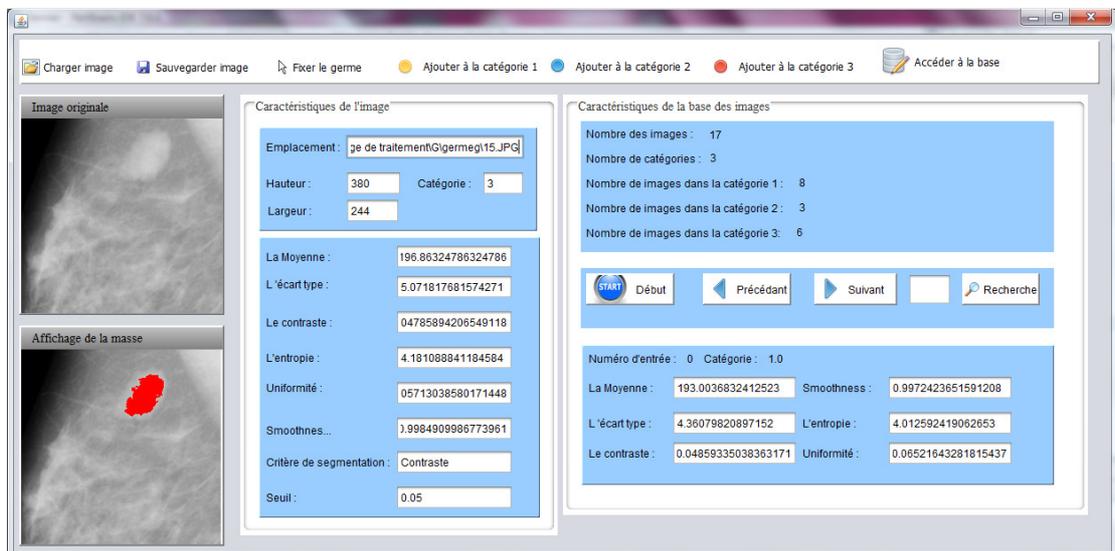


Fig 3.3. Vue d'un résultat de l'application SASDUM 2

3.6. Résultats obtenus

La figure 3.4 présente quelques résultats du système réalisé. Les images (a), (b), (c) et (d) illustrent des mammographies où les masses sont cernées par l'expert. La catégorie de ces images selon, toujours l'expert, est respectivement F, D, G et G. Les images (e), (f), (g) et (h) visualisent les résultats du système réalisé. Ce dernier a classé correctement les trois premières images et il a segmenté les images chacune selon sa catégorie. Les résultats de segmentation sont plus corrects que ceux fournis en utilisant le même seuil pour les différentes images [9-12]. Cependant, la qualité des résultats de segmentation du système réalisé dégrade si ce dernier fournit une catégorie erronée lors du classement. Or cette situation est possible comme le montre la figure 2 pour l'image (h). L'image a été classée dans la catégorie F plutôt que G.

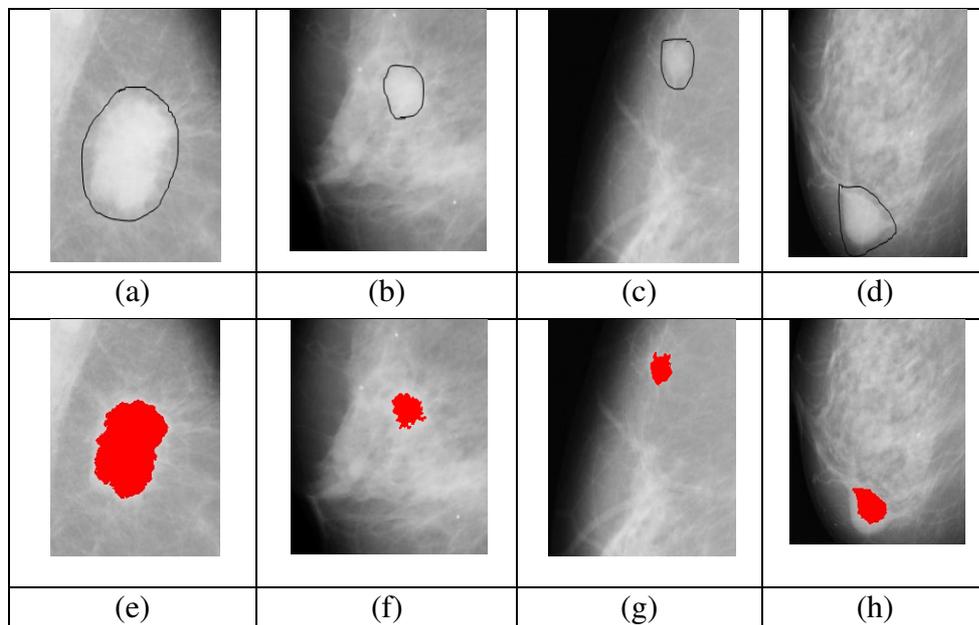


Fig. 3.4 : Résultats de l'application de détection de tumeur de SASDUM 2
 (a),(b),(c) et (d) mammographies avec masses cernées par l'expert
 (e), (f), (g) et (h) résultat de segmentation du système réalisé

3.7. Conclusion

Cette phase du projet propose une démarche d'extraction de masse basée sur la segmentation par croissance de région. L'importance de la démarche proposée réside dans sa capacité de déterminer automatiquement le seuil approprié selon la catégorie de la mammographie. Cette partie réalisée du système SASDUM permet d'améliorer les résultats de segmentation par rapport à l'approche de croissance de région qui utilise une valeur de seuil unique pour toutes les catégories.

3.8. Référence bibliographiques

- [1] Pohlman S, Powell KA, Obuchowski NA, Chilcote W.A, Grundfest-Broniatowski S: Quantitative Classification of Breast Tumors in Digitized Mammograms. Medical Physics, vol. 23, pp. 1337-1346 (1996)
- [2] Sahiner B, Chan HP, Wei D, Petrick N, Helvie MA, Adler DD, Goodsit MM: Image Feature Selection by a Genetic Algorithm: Application to Classification of Mass and Normal Breast Tissue. Medical Physics, vol. 23, pp. 1671-1684 (1996)
- [3] Petrick N, Chan HP, Sahiner B, Helvie MA: Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms. Medical Physics, vol. 26, pp. 1642-1654 (1999)
- [4] Pohle R, Toennies KD: Segmentation of medical images using adaptive region growing. in Medical Imaging, Proceedings of SPIE, vol. 4322, pp. 1337-1346 (2001)
- [5] Shan J, Cheng HD, Wang Y: A novel automatic seed point selection algorithm for breast ultrasound images. Proceedings of International Conference on Pattern Recognition, Tampa, Finland, pp.1-4 (2008)
- [6] Adams R, Bischof L: Seeded Region Growing. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, pp. 641-647 (1994)
- [7] Hojjatoleslami S.A, Kittler J: Region growing: a new approach. IEEE Transactions on Image Processing, vol. 7, pp. 1079-1084 (1998)
- [8] Cao Y, Hao X, Zhu X, Xia S: An adaptive region growing algorithm for breast masses in mammograms. Front. Electr. Electron. Eng. China 2010, vol. 5, pp. 128-136 (2010)
- [9] **A.Melouah**, A.Salhi, "Mass extraction in mammograms by an adaptive region growing technique" 2nd World Conference on Information Technology, 24-27 November 2011, Antalya – Turkey
- [10] **Melouah A.**, "*Toward an adaptive extraction method of masses contours in digitized mammograms*"; the 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA'13), 12-15 May 2013, Algiers, Algeria
- [11] **Melouah A.** "*A novel region growing segmentation algorithm for mass extraction in mammograms*", the fourth International Conference on Computer Science and its Applications *CIIA'2013*) 4-6 May 2013, Saida - Algeria.
- [12] **Melouah A** "*A novel region growing segmentation algorithm for mass extraction in mammograms*"; Modeling Approaches and Algorithms, SCI 488, pp. 95-104. ;Springer International Publishing Switzerland 2013

Sommaire du chapitre 4

1. Introduction
 2. Objectifs de la phase
 3. Equipe de travail
 4. Démarche Adoptée
 5. Réalisation et Résultats obtenus
 6. Conclusion
 7. Référence bibliographiques
-

4.1 Introduction

Les technologies de l'intelligence artificielle ont mené les chercheurs à apercevoir que la construction d'un système expert nécessitait presque toujours la prise en compte de l'incertitude dans le raisonnement ce qui explique d'une part le développement des méthodes probabilistes comme les réseaux bayésiens [1]. Les réseaux bayésiens sont des outils privilégiés pour les problèmes de diagnostic. Au cours de cette dernière décennie, ils ont été utilisés avec succès dans le domaine médical et font l'objet de nombreuses recherches actuellement, avec différents objectifs et de l'évaluation des connaissances. Dans ce travail nous modélisons un système expert qui essaie de simuler le raisonnement des médecins pour l'aide au diagnostic des pathologies les plus fréquentes des seins. La structure du réseau bayésien est évaluée uniquement à partir des avis des experts, elle permet une représentation de connaissances qualitatives et quantitatives exprimant l'incertitude décomposée en quatre niveaux : niveau clinique, niveau biologique, niveau imagerie médicale et niveau diagnostic. Nous utilisons les réseaux bayésiens pour calculer les probabilités à posteriori de ou des causes à partir d'un ensemble de valeurs fixées à priori (probabilité à priori) en utilisant un algorithme d'inférence exacte Pearl. Cette probabilité à priori d'un évènement est estimée par sa fréquence d'apparition dans la base de données qu'on nomme *maximum de vraisemblance* (MV). Une étude statistique à été faite à l'hôpital Ibn Rochd d'ANNABA sous la direction du docteur professeur Hayet AOURES pour accumuler des nouveaux cas pendant l'année 2011. La base de données constituée comprend 100 cas.

4.2 Objectifs de la phase

Dans les applications pratiques, les bases de données sont très souvent incomplètes. Certaines variables ne sont observées que partiellement ou même jamais. La méthode d'estimation de paramètres avec des données incomplètes la plus couramment utilisée est fondée sur l'algorithme itératif *Expectation-Maximisation* (EM) proposé par Dempster. Pour améliorer les performances du système de diagnostic nous avons rajouter au réseau bayésien une partie additionnel qui consiste à développer une nouvelle stratégie de maintenance bayésienne en définissant des variables supplémentaires notées comme des actions de maintenance. Avec cette maintenance nous pourrions mettre à jour le réseau bayésien, car des erreurs s'introduiront que se soit lors de sa construction ou lors des changements apportés par la dynamique du domaine[2, 3].

L'ensemble des données cliniques ou para cliniques est également difficile à synthétiser et il est donc fréquent de tirer des conclusions incertaines qui amèneront des décisions de diagnostics et thérapeutiques qui seront probablement remises en cause. C'est ce qui explique d'une part le développement des méthodes probabilistes non explicatives et d'autre part le développement des systèmes experts qui essayent de simuler les raisonnements des experts et se voulant être plus transparents en permettant d'objectiver les connaissances des experts et le raisonnement suivi. Il n'en

reste pas moins que ce raisonnement des experts est basé sur des inférences prenant en compte l'incertain tant au niveau individuel des patients qu'au niveau de la population [4]. Les recherches actuelles tendent donc à combiner les deux approches en utilisant des raisonnements basés sur les réseaux probabilistes introduits par Pearl sous le terme de 'Belief Networks'.

4.3 Equipe de travail

Membre du projet : Le chef du projet Dr Merouani Hayet Farida est impliqué dans cette phase.

Docteur en gynécologie à l'hôpital Ibn Rochd : Pr. Aoues Hayet

Étudiantes post-graduation (doctorants): Refai Ahlem, Djebbar Akila

Étudiante en master 2: Gasmi Amina

4.4 Démarche adoptée

Les réseaux bayésiens sont préférables à d'autres modèles vu qu'ils possèdent des privilèges comme la possibilité de rassembler et de fusionner des connaissances de diverses natures dans un même modèle. La représentation graphique d'un réseau bayésien est explicite, intuitive et compréhensible par un non-spécialiste, ce qui facilite la validation du modèle. Un réseau bayésien est polyvalent, on peut se servir du même modèle pour évaluer, prévoir, diagnostiquer, ou optimiser des décisions, ce qui contribue à rentabiliser l'effort de construction du réseau bayésien [5, 6].

Les réseaux bayésiens

Un réseau bayésien $B = (G, \theta)$ est défini par :

- $G = (X, E)$, graphe dirigé acyclique dont les nœuds E sont associés à un ensemble de variables aléatoires $X = \{X_1, \dots, X_n\}$.

- $\theta = \{P(X_i | Pa(X_i))\}$, ensemble des probabilités de chaque nœud X_i conditionnellement à l'état de ses parents $Pa(X_i)$ dans G .

- Un ensemble de variables aléatoires $X = \{X_1, \dots, X_n\}$ associées aux nœuds du graphe telles que

$$P(X_1, X_2, \dots, X_n) = \prod_{i \in X} P(X_i | Pa(X_i)) \quad (1)$$

S'il existe un arc $x \in X$ de A vers le nœud B , A est nommé le parent de B . L'ensemble des nœuds parents du nœud X_i est noté par $Pa(X_i)$.

Dans le raisonnement du médecin, plusieurs symptômes identiques peuvent aboutir à des maladies différentes. Le raisonnement médical part d'un examen clinique cognitif qui est un ensemble de données interrogatoires. En effet, à partir des symptômes observés (les signes cliniques), le médecin identifie les facteurs prédisposant et sur la base desquels plusieurs hypothèses seront établies. Pour confirmer ses hypothèses, il doit passer par un examen radiologique (*mammographie*). On sait qu'un examen clinique n'est pas spécifique à 100%. Il reste toujours une incertitude, ce n'est qu'à travers l'examen biologique (*biopsie, cytoponction*) que le médecin arrivera à poser son diagnostic final.

Dans cette phase de SASDUM nous proposons un système de diagnostic basé réseau bayésien, voir figure 4.1

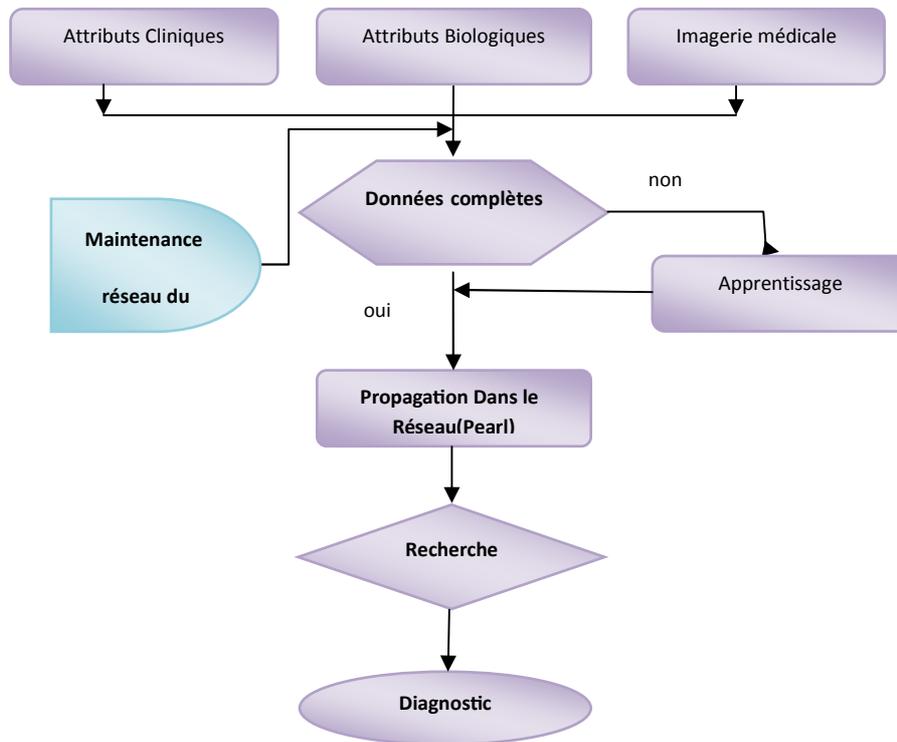


Fig.4.1 Architecture du réseau bayésien proposé

Ainsi la phase de SASDUM diagnostic comprend les fonctionnalités suivantes :

La propagation qui agit par transmission de message (passage de message) proposé par Pearl [7, 8] entre les différents nœuds du réseau. Si l’observation lancée par le processus d’activation est incluse dans la base de données, alors chaque nœud apprend l’information et la fait connaître à l’ensemble du graphe, jusqu’à ce que tout nœud puisse mettre à jour sa probabilité en fonction de l’ensemble de l’information reçue par le graphe. Si l’observation ne figure pas dans la base de données, alors la phase d’apprentissage de paramètres va être déclenchée en utilisant l’algorithme EM.

La recherche consiste à déterminer le résultat du diagnostic à partir d’un ensemble d’entrées au cour du traitement, se seront alors les paramètres cliniques, biologiques, et l’attribut d’imagerie médicale qui déclencheront le réseau à l’état active.

La Maintenance En effet chaque système informatique notamment les systèmes qui essayent de simuler les raisonnements des médecins incertains parfois nécessite en phase de fonctionnement d’être maintenu pour garantir la qualité et la performance. Ainsi, *la phase maintenance* donne une conception dynamique de notre graphe. En littérature il existe trois façons pour construire un réseau bayésien ; automatique, manuelle et hybride. En ce qui nous concerne nous avons opté pour le troisième cas, car la structure du réseau est décrite avec l’aide des médecins et les probabilités sont obtenues à partir d’un apprentissage *statistique* dans notre cas les données sont obtenues en estimant la probabilité d’un évènement par sa fréquence d’apparition dans la base de données cette approche, appelée *maximum de vraisemblance* (MV) [9] nous aurons alors :

$$P(X_i = x_k \mid \text{parent}(X_i) = x_j) = \theta_{i,j,k} = \frac{N_{i,j,k}}{\sum_k N_{i,j,k}}$$

Sachant que :

$N_{i,j,k}$ est le nombre d'événements dans la base de données pour lesquels la variable X_i est dans l'état x_k et ses parents sont dans la configuration x_j . C'est cette approche que nous avons adoptée dans notre application. Nous nous sommes concentrés premièrement sur la réalisation du modèle bayésien avec l'assistance du médecin au moyen d'une discussion qui nous a menées à déterminer les variables représentant les maladies du sein leurs principaux signes, symptômes, les tests, et les examens. Ce processus nous a conduits à un modèle comportant 49 liens et 29 nœuds comme le montre la figure 4.2. Le réseau proposé permet une représentation de connaissances qualitatives et quantitatives exprimant l'incertitude pour le diagnostic des maladies les plus fréquentes des seins et qui est décomposé en niveau : *niveau clinique*, *niveau biologique*, *niveau imagerie médicale* et *niveau diagnostic* chaque niveau est composé d'un ensemble de variables (**discrète ou continue**). Chaque variable est associée à une table de probabilité.

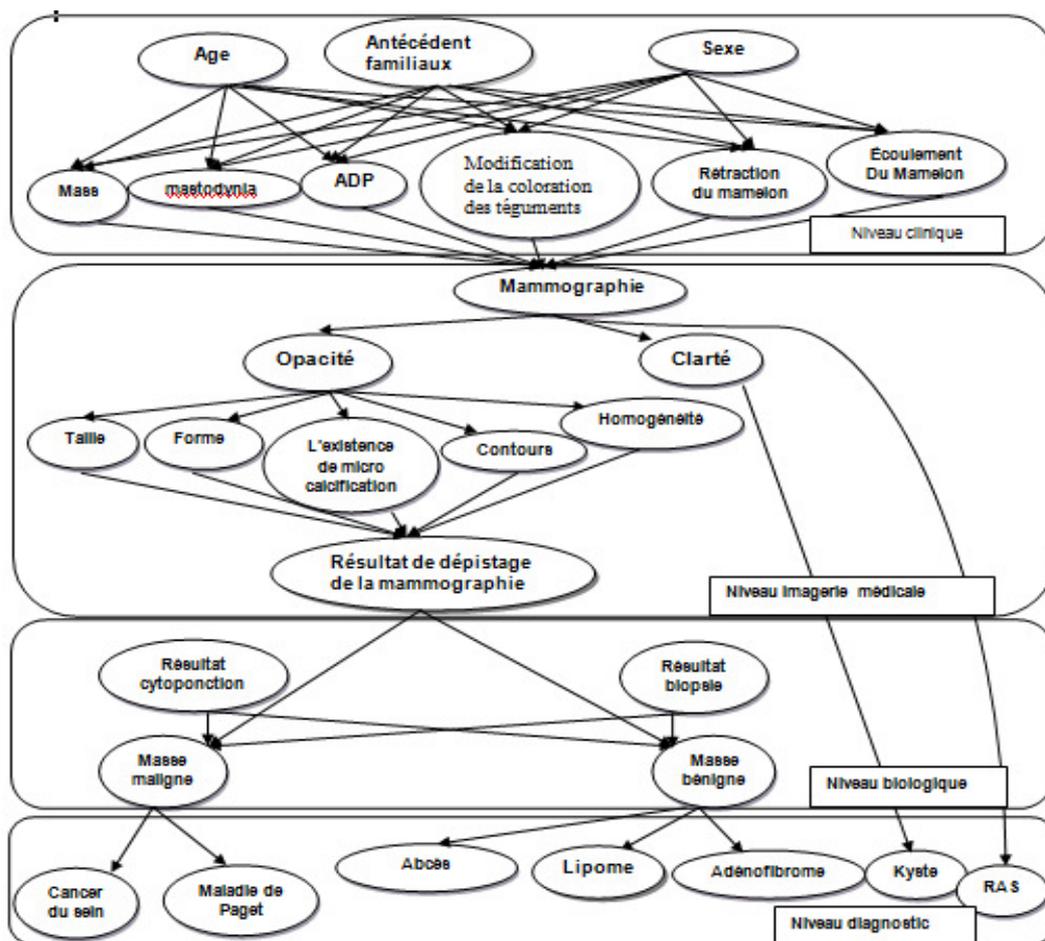


Fig.4.2. Réseau bayésien modalisant sept maladies des seins

4.5 Réalisation et Résultats obtenus

Cependant, dans la pratique, nous nous sommes concentré dans un premier temps à collecter le maximum de dossiers (informations) à l'hôpital Ibn Rochd d'ANNABA, avec l'aide d'un professeur en gynécologie(ou il fallait être présent lors de ces consultations), nous avons réussi à avoir 100 cas à partir desquelles nous avons trouvé les tables à priori et conditionnelles en appliquant le maximum de vraisemblance (MV) et voici une capture sur l'ensemble des cas et exemple d'une table de probabilité conditionnelle retenues après l'application de MV montrer dans la figure 4.3 et figure 4.4.

Age	anté	se	Massi	Mastoc	Adp	Moc	Rétri	Ecou	ma	Of	Clai	R	La t	Mi	Forr	Ci	hoi	C	bi	M	B	Maladie
46	non	f	oui	non	oui	oui	oui	non	oui	oui	non	non	2cm-5	non	Ronde	irrég	Hytér	m	m	oui	non	cancer
46	oui	f	oui	non	non	non	non	non	oui	oui	non	non	<2cm	non	ronde	irrég	homc	m	m	oui	non	cancer
79	non	f	oui	non	non	non	non	non	oui	oui	non	non	2cm-5	non	ronde	régu	Homc	m	m	oui	non	cancer
26	oui	f	oui	non	non	non	non	non	oui	oui	non	non	2cm-5	OUI	ovale	irrég	hytér	m	m	oui	non	cancer
60	non	f	oui	oui	oui	non	oui	non	oui	oui	non	non	2cm-5	OUI	ronde	irrég	hytér	m	m	oui	non	cancer
39	non	f	oui	oui	non	non	non	oui	oui	oui	non	non	2cm-5	non	ovale	irrég	hytér	m	m	oui	non	cancer
44	oui	f	oui	oui	oui	non	non	non	oui	oui	non	non	>5cm	non	Ovale	irrég	Homc	nor	m	oui	non	cancer
52	non	f	oui	non	oui	non	oui	non	oui	oui	non	non	2cm-5	OUI	irrégul	irrég	hytér	m	m	oui	non	cancer
40	non	f	oui	non	oui	non	non	non	oui	oui	non	non	2cm-5	OUI	irrégul	irrég	homc	m	m	oui	non	Cancer
45	non	f	oui	oui	oui	non	non	non	oui	oui	non	non	>5cm	non	Ovale	irrég	Homc	Nor	m	oui	non	cancer
52	non	f	oui	non	oui	non	oui	non	oui	oui	non	non	2cm-5	OUI	irrégul	irrég	hytér	m	m	oui	non	cancer
60	oui	f	oui	oui	oui	non	oui	non	oui	oui	non	non	2cm-5	OUI	ronde	irrég	hytér	m	m	oui	non	cancer
40	non	F	oui	non	oui	non	non	non	oui	oui	non	non	<2cm	OUI	irrégul	irrég	homc	m	m	oui	non	Cancer
45	non	F	oui	oui	non	non	non	oui	oui	non	non	non	<2cm	OUI	ronde	régu	homc	b	b	non	oui	kyste
39	non	F	oui	oui	non	non	non	oui	oui	non	non	non	2-5cm	OUI	ronde	régu	homc	b	b	non	oui	kyste
29	non	F	oui	non	non	non	non	non	oui	oui	non	non	2-5cm	OUI	ronde	régu	homc	b	b	non	oui	adénofib
18	non	F	oui	non	non	non	non	non	oui	oui	non	non	>5cm	OUI	ronde	régu	homc	b	b	non	oui	adénofib
22	oui	F	oui	oui	non	non	non	non	oui	oui	non	non	2-5cm	OUI	ovale	régu	homc	b	b	non	oui	adénofib
21	non	F	oui	non	non	non	non	non	oui	oui	non	non	<2cm	OUI	ronde	régu	homc	b	b	non	oui	adénofib
25	non	F	oui	oui	non	non	non	non	oui	oui	non	non	>5cm	OUI	Ronde	régu	homc	b	b	non	oui	adénofib
17	oui	F	oui	non	non	non	non	non	oui	oui	non	non	<2cm	OUI	Ronde	régu	homc	b	b	non	oui	adénofib
20	non	F	oui	oui	non	non	non	non	oui	oui	non	non	<2cm	OUI	ovale	régu	homc	b	b	non	oui	adénofib
19	non	F	oui	non	non	non	non	non	oui	oui	non	non	2-5cm	OUI	ovale	régu	homc	b	b	non	oui	adénofib

Fig.4.3 : Table des cas retenus pour l'évaluation

L'inférence bayésienne qui consiste à calculer les probabilités de la distribution de probabilités postérieures. Cette partie sert à fournir le résultat final du diagnostic par l'utilisation de l'algorithme d'inférence Pearl [10] qui calcule la probabilité de chaque variable à partir d'un ensemble de valeurs fixées à priori. Cet algorithme passe par deux étapes :

1. *Étape de propagation*

Le principe de cet algorithme est simple, il fonctionne comme suit :

- a. Initialisation les probabilités des variables du réseau.
- b. Faire passer le message π entre les variables du réseau.
- c. Chaque variable reçoit le message doit envoyer le message π qui porte la probabilité du parent vers son fils.
- d. À la réception de message π les récepteurs de la 2ième et 3ième niveau mettent à jour ses probabilités puis ils recalculent le message π pour obtenir la probabilité actuelle par la formule suivante :

$$\pi (X=x) = \sum z P (X=x|Z=z) \pi x (Z=z)$$

- e. Le premier niveau calcule ses probabilités par la formule suivante :

$$\pi (x) = p(x)$$

- f. Pour les variable du 4ième niveau il n'y a pas de message π à envoyer, ils mettent à jour ses probabilité à la lumière du message reçus.

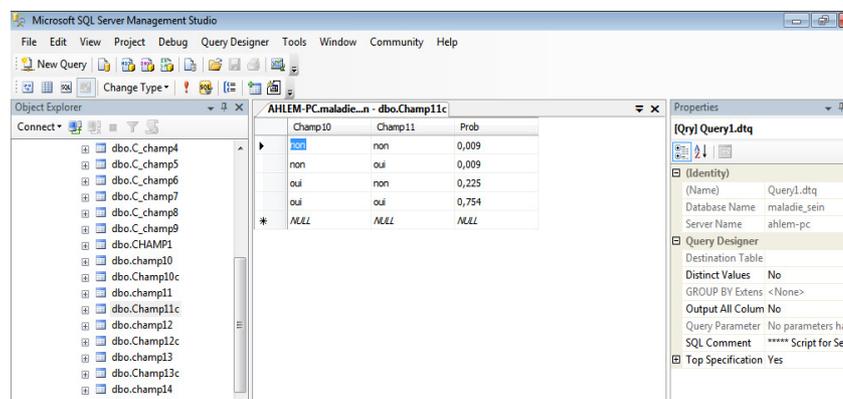


Fig.4.4 : Exemple d'une table de probabilité conditionnelle

2 .Étape de recherche

Le point d'Arrêt de l'étape de propagation représente le point de départ pour l'étape de recherche qui va nous fournir le résultat final du diagnostic. On cherche au niveau du dernier niveau la maladie ayant la probabilité max sachant les observations (les entrées).

$$\text{Arg Max } (p (M=m/e))$$

Ou M=maladie, e : évidence.

Dans notre cas le Réseau bayésien est sans circuit ce qui permet d'utiliser l'algorithme d'inférence proposé par [10]. Avec cet algorithme, il est possible d'inférer la valeur de n'importe quelle variable du graphe, ce qui donne des probabilités d'utilisation plus adaptées. La figure 4.5 montre une fiche d'exploitation avec un résultat de diagnostic.

The screenshot shows a software window titled "Fiche d'exploitation". At the top, it says "veuillez cocher les cases en desous:". Below this, there are three main sections: "Identité de patient(e)", "Examen clinique", and "Examea biologique:". The "Identité de patient(e)" section includes fields for Age (15-35, 35-65, >65), Sexe (Femme, Homme), and Antécédend familial (oui, non). The "Examen clinique" section has several rows of radio buttons for symptoms like "Ecoulement mamelon", "Rétraction mamelon", "Adénopathie", "Opacité", "Clarté", and "Rien à signalé". The "Examea biologique:" section includes "Dipistage" (bénigne, maligne) and "Cytologie" (bénigne, maligne, non concluant). A modal dialog box titled "Project2" is overlaid on the interface, showing the following probabilities: P(cancer)=0,038464, P(lipome)=0,02592, P(abcès)= 0,0091, P(adénofibrome)=0,02, and P(kyste)=0,024. There is an "OK" button in the dialog box.

Fig.4.5 : Fiche d'exploitation pour la saisie des données

3. Etape de maintenance

Une autre étape et pas des moindre, dans cette phase de SASDUM est la maintenance du réseau bayésien. Notre objectif ici est de retenir les variables critiques, afin de prévoir des taches de maintenance sur ces variables du modèle. Ces variables ont été choisies par les experts. Ensuite, les experts ont sélectionnés des taches de maintenance sur les variables et donc sur les composants choisis. Ces actions de maintenances peuvent être vues comme des variables du modèle graphique. ces nœuds se situent en amont des variables sur lesquelles elles agissent. L'autre objectif est de définir des politiques de maintenance visant à observer les changements apporté par la maintenance. Il faut donc retenir des variables critiques, afin de prévoir des taches de maintenance sur ces variables du modèle. Ces variables ont été choisies par les experts. La maintenance du réseau bayésien prend en entrée la collection d'actions données par l'expert du domaine, ainsi que le réseau initial. L'objectif de cette phase est de décider si nous allons intégrer, ou non, les différentes connaissances portées par les règles de maintenance. Ainsi, la définition d'une nouvelle action de notre modèle va nécessiter l'avis d'expert du domaine. L'ensemble des actions a été réfléchi pour faciliter leur interprétation en vue d'une intégration au modèle (ajout, modification ou suppression) d'arcs ou nœuds du graphe. Le cas échéant, il faut redéfinir les tables de probabilité impactées par les modifications, et dans ce cas, il est possible de s'appuyer sur les données et les règles de l'expert du domaine.

Si nous faisons le choix des actions de maintenance suivante :

Nous recensons les actions de maintenance à mettre avec leurs probabilités à priori et conditionnelle sachant que leur type est discret qui sont les suivant : (A : désigne une action de maintenance)

- **A1** : Prendre des pilules, avec probabilité $P(A1)=0.45$.
- **A2** : L'âge d'avoir des enfants, avec probabilité $P(Aa2)=0.22$, $P(Ap2)=0.3$.
- **A3** : obésité, avec probabilité $P(A3)=0.25$.

Le réseau bayésien sera celui de la figure (4.6) avec les actions de maintenance ; l'ajout de telles variables a été réalisé par les experts, ainsi que l'évaluation de leurs probables

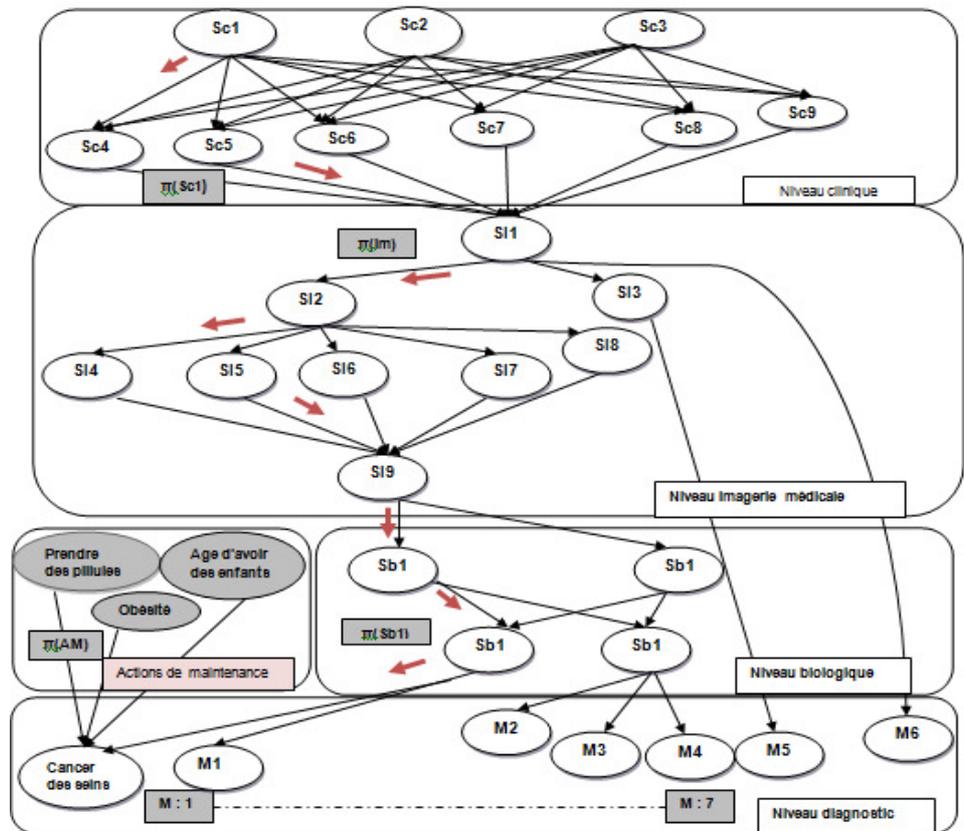


Fig.4.6 Réseau bayésien avec action de maintenance

Les variables sur lesquelles les actions de maintenance [16] sont susceptibles d'agir peuvent être soit des variables d'entrée, soit des variables intermédiaires. Une action de maintenance peut également influencer sur une ou plusieurs variables et pas nécessairement sur des variables de même type (entrée, intermédiaire, sortie). Les trois paramètres que nous avons étudiés dans l'exemple précédent seront ajoutés dans le réseau bayésien résumé dans Figure (4.6). Une fois cette variable importante identifiée, l'aide à la décision prend part via l'intégration des actions de maintenance sur cette variable jugée importante. Pour cela nous considérons les tâches de maintenance comme de nouveaux nœuds de réseau bayésien apportés avec leur probabilité conditionnelle, une nouvelle inférence est faite avec la prise en compte des actions de maintenance.

Le nouveau réseau bayésien avec les actions de maintenance est donné dans Figure (4.6), ou une nouvelle inférence est faite avec les actions de maintenance. [18, 19]

Discussion sur les résultats apportés par la maintenance :

Tab.4.1.Taux avant et après maintenance

	Sans maintenance	Avec maintenance
Les taux résultants	29,94	54,76

Comme exemple nous avons choisi une action de maintenance « prendre des pilules », la validation à apporter des résultats estiment par rapport à l’avis des experts du domaine qui considère les attributs qu’on a ajouté dans notre étude comme des facteurs primordiaux qui bascule la probabilité vers une amélioration. L’ajout a augmenté la probabilité d’avoir un cancer des seins avec un rapport égal à 24.82% décrit dans le tableau (4.1). *Ce qui permet de dire que les actions de maintenance permettent de diminuer le taux d’erreur par l’augmentation de probabilité en donnant un bon diagnostic.*

4.6 Conclusion

Cette étude de SASDUM nous a permis d’établir la formalisation d’un système d’aide au diagnostic des pathologies des seins à base de RB enrichi par la notion de la maintenance qui porte un bénéfice certain pour le RB qui se résume dans l’analyse des résultats. Pour cela nous nous sommes basé sur l’algorithme de Pearl qui permet de calculer les probabilités marginales celles-ci assurant un diagnostic efficace et facile pour le modèle graphique. Ce système est appliqué dans le diagnostic des maladies les plus fréquentes des seins pour aider le médecin dans son processus de diagnostic. Les réseaux bayésien présente une méthodologie de représentation des connaissances permettant d’exprimer les relations entre les variables locales en utilisant la théorie des probabilités pour gérer l’incertitude liée au problème médicale afin d’apporter une vision compréhensible pour un non spécialiste.

4.7 Référence bibliographiques

1. Patrick Naïm, Pierre-Henri Wuillemin, Philippe Leray, Oliver Pourret et Anna Becker, “Réseaux Bayésiens”, 3ème Edition Eyrolles.
2. Philippe Leray « Réseaux bayésiens : apprentissage et modélisation de systèmes complexes », thèse pour obtenir l’habilitation à diriger les recherches.2006.
3. Alain.Delaplace « Approche évolutionnaire de l’apprentissage de structure pour les réseaux bayésiens », thèse pour obtenir le grade de docteur de l’université de tours, 2007.
4. Franck Corset. « Aide à l’optimisation à la maintenance à partir de réseau bayésien et fiabilité dans un contexte doublement censuré », Thèse de doctorat, université de Joseph Fourier, PP 75-86, 2003.
5. Pierre Le Beux, Anita Burgun et Mireille Cléret. « De la méconnaissance à l’expertise », 1994.
6. Naouar Fatih, « Modélisation d’une détection de diagnostic d’une maladie par les réseaux bayésiens possibiliste», Institut Supérieur des Sciences Appliquées et de Technologie de Sousse Université de Sousse Tunisie, 2007.
7. Lobna Bouchaala et Slaheddine Jarbouni, «Application des Réseaux Bayésiens à la Modélisation et la Planification des Réseaux».2009.
8. J.Pearl, «Probabilistic Reasoning in Intelligent Systems». San Mateo, CA: Morgan Kaufmann, 1988.
9. Abdeljabba BEN SALEM « Modèles Probabilistes de Séquences Temporelles et Fusion de Décisions. Application à la Classification de Défauts de Rails et à leur Maintenance » thèse de doctorat le 07 mars 2008, l’université Henri Poincaré – Nancy 1 spécialité Automatique, Traitement du Signal et Génie Informatique. p.1-60.
10. Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, 29:241.
11. David Heckerman. A tutorial on learning with Bayesian network. In Michael I. Jordan, editor, *Learning in Graphical Models*, pages 301–354. Kluwer Academic Publishers, Boston, 1998.
12. Paul Krause. *Learning probabilistic networks*, 1998.

13. Lauritzen, S., Thiesson, B., and Spiegelhalter, D. (1994). Diagnostic systems created by model selection methods: A case study. In Cheeseman, P. and Oldford, R., editors, AI and Statistics IV, volume Lecture Notes in Statistics, 89, pages 143{152. Springer-Verlag, New York.
14. Djebbar Akila, **Hayet Farida Merouani** “Optimizing retrieval phase in CBR through Pearl and JLO algorithms for medical diagnosis.”, in International Journal of Advanced Intelligence Paradigms (IJAIP), Vol. 5, N°3.2013 pp.61-81 **ISSN 1755-0386**, indexed in **INDERSCIENCE**, ProQuest,ULRICH, DOAJ, OpenJ-Gate, Cabell, EI Compendex. <http://www.inderscience.com/ospeers/admin/author/articlelist>.
15. Djebbar Akila, **Hayet Farida Merouani** “Applying BN in CBR Adaptation –Guided Retrieval for Medical Diagnosis”, in International Journal of Hybrid Information Technology, (IJHIT), Vol. 5, N°2, April 2012, pp.41-56. **ISSN 1738-9968**, indexed in **EBSCO**, ProQuest,ULRICH, DOAJ, OpenJ-Gate, Cabell, EI Compendex. <http://www.sersc.org/journals/IJHIT>
16. Refai Ahlem, Djebbar Akila & **H.F.Merouani**, “ RB-Maint: Un Modèle Probabiliste pour la Maintenance d’un Système RàPC” 7 éme Colloque sur l’Optimisation et les Systèmes d’Information, COSI’10, Avril 18-20, 2010, Ouargla, Algérie.
17. Djebbar Akila & **Hayet Farida Merouani** “Retrieval of cases by using a Bayesian Network applied to the medical diagnosis” The Mediterranean Journal of computers and networks , Vol.8, N°2,pp.69-74 April 2012, **ISSN 1744-2397**, indexed in **INSPEC**, **British Library**. www.medjcn.com
18. Ahlam Refai & **H.F.Merouani**, “ Application à l’aide de diagnostic médicale à base ce réseau bayésien dans le cadre de la maintenance” The Third International Conference on Image and Signal processing and their Applications, ISPA’12, 02-04 December 2012, Mostaganem, Algeria.
19. Refai Ahlem & **H.F.Merouani**, “ Maintenance des Réseaux Bayésiens pour le Diagnostic des Pathologies des Seins” Second International Conference on Systems and Information Processing, ICSIP’11, May 15-17, 2011, Guelma, Algérie.

Sommaire du chapitre 5

1. Introduction
 2. Objectifs
 3. Equipe de travail
 4. Schéma général de l'approche
 5. Descripteurs utilisés
 6. Méthode proposée
 7. La recherche d'images par contenu
 8. Réalisation
 9. Analyse des résultats de classification
 10. Conclusion
 11. Référence bibliographiques
-

5.1 Introduction

Les développements récents dans le domaine du traitement de signal et des bases de données offrent tous les éléments nécessaires pour l'extraction, l'indexation et la recherche du contenu visuel des données multimédias, notamment des images. Les techniques d'indexation d'images ont pour objet d'organiser un ensemble de descripteurs (un descripteur étant un vecteur de réels décrivant le contenu visuel d'une image) afin que les procédures de recherche soient performantes en temps de réponse. Cette organisation se traduit généralement par une structuration des descripteurs en petits ensembles et par l'application de stratégies de recherche capable de filtrer toutes les images non pertinentes qui seront évitées (non parcourues) pendant la recherche garantissant ainsi un temps de recherche acceptable par l'utilisateur. Dans ce cadre la recherche consiste à développer et appliquer des outils qui permettent de sélectionner les images les plus pertinentes par leurs contenus. Lors de l'interrogation, un (ou plusieurs) descripteur, est tout d'abord extrait à partir de l'image requête. Ce descripteur requête est ensuite utilisé pour retrouver les descripteurs stockés dans la base qui lui sont les plus proches en terme de similarité. Les descripteurs trouvés permettent d'obtenir les images auxquelles ils sont associés et qui, de fait, sont censés être similaires à l'image requête.

La mise en œuvre d'outil d'indexation et de recherche dans un contexte d'une base d'images de mammographie s'effectue dans le cadre d'un système de Recherche d'images par le contenu (SRIC ou CBIR pour Content Based Image Retrieval).

5.2 Objectifs de la phase d'indexation

L'objectif de la phase d'indexation de SASDUM est la conception et la réalisation d'une technique de recherche par le contenu adaptée aux mammographies. Deux idées proposées par le chef de projet Dr Merouani H.F. ont été exploitées. La première idée basée sur les automates cellulaires est totalement nouvelle. Par contre, la deuxième idée se base sur la fusion des lois Zipf et Zipf inverse qui est aussi une nouvelle approche dans le domaine médicale.

La suite de ce chapitre expose le travail réalisé pour développer SASDUM 4.1 partie de SASDUM pour l'indexation des mammographies basée sur les automates cellulaires. Le chapitre suivant décortique SASDUM 4.2 partie de SASDUM pour l'indexation des mammographies basée sur les lois Zipf et Zipf inverse.

5.3 Equipe de travail

Cette phase est réalisée par deux membres du projet : Dr.Merouani Hayet Farida et Dr.Layachi Soumia et l'étudiante en post-graduation : Melle Benmazou Sarah.

Experts qui ont contribué dans l'élaboration de SASDUM 4.1:

1.Dr.Djilani Nour el Houda :Médecin Spécialiste en radiologie

- Cabinet d'imagerie médicale, échographie radiologie spécialisée mammographie, spécialiste en radiologie et imagerie médicale, wilaya de Guelma.
- Etablissement public hospitalier Ibn Zohr, unité échographie, wilaya de Guelma.

2.Dr.Beledjhem Nedjmeddine : Maitre assistant hospitalo-universitaire

- Etablissement public Ibn Rochd, Annaba.
- Lors de nos réunions nous avons discuté, exploré le thème d'une façon générale, sur le diagnostique du cancer du sein ainsi que les méthodes utilisées, en plus des images octroyées.

5.4 Schéma général de l'approche

Cette architecture résume le principe de SASDUM 4.1. Chaque étape de ce processus sera détaillée séparément dans les prochaines sections.

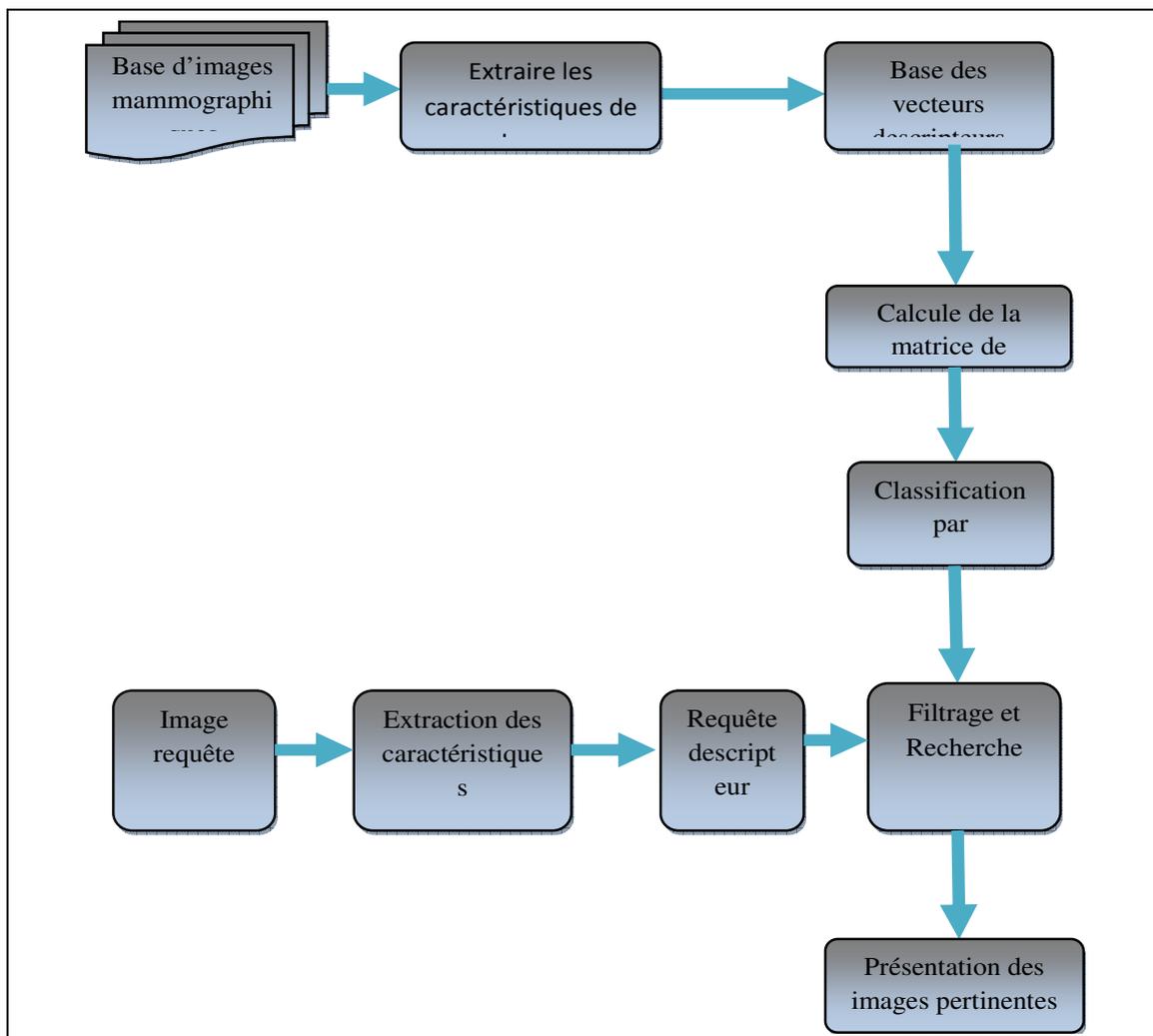


Fig.5.1.Schéma générale du système de classification et d'indexation des mammographies SASDUM4.1

5.5. Descripteurs utilisés

Les descripteurs sont des mesures de caractéristiques de l'image qui doivent être invariantes en rotation, translation et changement d'échelle. Les caractéristiques les plus utilisées sont la couleur, la texture et la forme. Dans notre travail nous nous sommes intéressés aux descripteurs de la texture qui donne une description de la structure locale ainsi qu'une information sur le champ aléatoire qui contrôle les détails à petite échelle. Elle décrit habituellement ce qui est laissé après élimination de l'information de la couleur et de la forme.

La méthode de la matrice de cooccurrence à niveau de gris pour extraire les indices de textures a été implémentée, car elle est très utilisée en recherche d'images, et en générale donne de bons résultats. Rappelons que la matrice de cooccurrence à niveau de gris (GLCM) a été proposée par Haralick [11]. Cette approche est basée sur la probabilité jointe de la distribution des pixels dans l'image. Ou l'élément $p_{d,\theta}(i,j)$ de la matrice de cooccurrence définit la fréquence d'apparition des couples de niveaux de gris i et j pour les couples de pixels séparés par une distance d selon la direction θ . Cette matrice décrit les régularités observables dans les niveaux de gris des pixels d'une région. Le calcul de la matrice de cooccurrence nécessite le choix d'une distance et d'un angle de déplacement. Il a été noté par les chercheurs que la distance d'un pixel combinée avec les distances de 0° , 45° , 90° et 135° donne de bons résultats [11, 12].

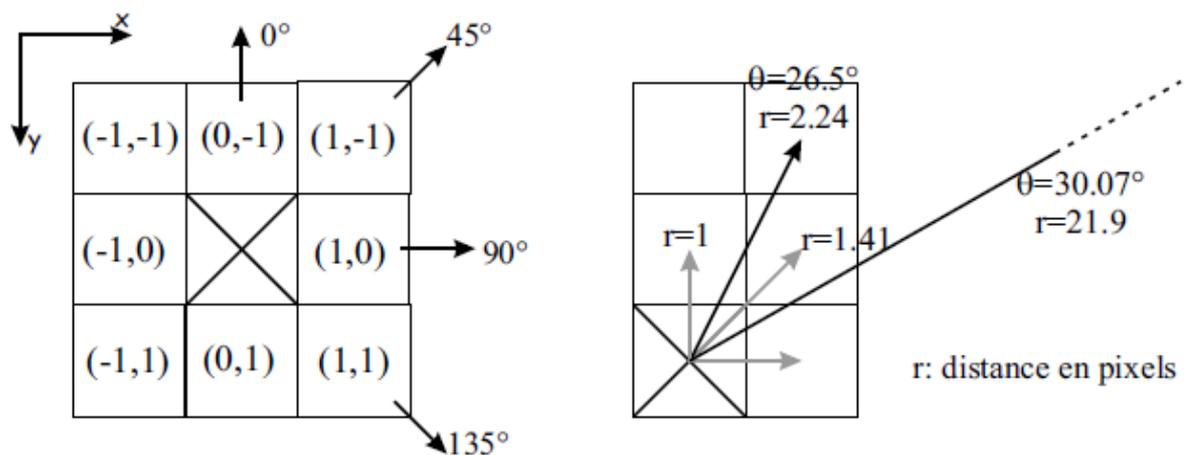


Fig.5.2. Distance de calcul des matrices de cooccurrence vs orientation

Gauche : plus proches voisins. Droite : plus d'orientations implique des distances de calcul plus grandes.

Notre choix s'est porté sur le calcul d'une matrice de cooccurrence pour distance $d = 1$ est direction $\theta = 0^\circ$.

Afin d'estimer la similarité entre les matrices de cooccurrence [13, 14], Haralick a proposé 14 caractéristiques statistiques extraites à partir de cette matrice. Actuellement, seulement cinq caractéristiques les plus appropriées sont utilisées et que nous avons également calculés pour notre application, et qui sont : l'énergie (homogénéité), la corrélation, la variance, le moment inverse de différence et l'entropie.

Avant d'utiliser la méthode proposée de classification basée automates cellulaires [8, 9], il est impératif de mesurer la similarité de notre base d'images de mammographie. Idéalement les mesures doivent être capable de mesurer la similarité sémantique de deux images, dans la pratique elle ne sont capable que de mesurer la similarité visuelles entre ces deux images [7].

De nombreuses définitions de distances ont été proposées, chacune donnant bien entendu des résultats différents, le choix d'une fonction distance est important dans ce domaine. Notre choix s'est porté sur la distance de Manhattan[7]. Cette norme est aussi connue sous le nom city-block est plus appropriée pour mesurer la similarité entre les données multivariées; elle est moins sensible au bruit coloré que la distance Euclidienne.

Considérons deux images I_1 et I_2 indexées par des vecteurs dans

$$R^N: I_1 = \{I_1(i), 1 \leq i \leq N\} \text{ et } I_2 = \{I_2(i), 1 \leq i \leq N\}$$

Le calcul de similarité entre I_1 et I_2 est donnée par la formule suivante :

$$Dist(I_1, I_2) = \sum_{i=1}^N (|I_1(i) - I_2(i)|)$$

Où i représente le nombre totale de paramètres (l'énergie, corrélation, la variance, le moment inverse de différence et l'entropie).

5.6. Methode Proposée

L'automate que nous proposons est un réseau de cellules dans un espace 2D. Cet automate possède 4 états possibles et le rayon de voisinage est une seule cellule, c'est-à-dire que le voisinage utilisé n'est rien d'autre que le voisinage de Moore (8 cellules voisines autour de la cellule elle-même) légèrement modifié [16]. Ainsi, une cellule de l'automate est soit morte, soit vivante, soit isolée soit contient une donnée d'où l'ensemble des états de l'automate est {Morte, Vivante, Isolée, Active}. Une cellule morte contiendra la valeur 0, une cellule vivante contiendra la valeur -1, une cellule isolée contiendra la valeur -2 et une cellule active contiendra une donnée (numéro ou index de l'image dans la base). Nous avons utilisés ces valeurs et surtout la valeur de la cellule vivante (-1) pour ne pas confondre entre une cellule vivante contenant la valeur 1 et une cellule contenant une donnée (numéro ou index de l'image) 1. Donc une cellule contiendra une valeur de l'ensemble $\{-2, -1, 0, 1, 2, \dots, N\}$ où N est le numéro de la dernière image dans la base. Si N est le nombre de documents à classifier, la taille de la grille cellulaire 2D est $m \times m$ sachant que $m = 2 \times (\text{Int}(\text{Racine}(n)) + 1)$ avec Int représentant la partie entière d'un nombre réel, Racine représentant la racine carrée et 2 représente un coefficient empirique pour avoir de la place pour l'organisation spatiale des classe dans la grille. A chaque itération de l'algorithme, les cellules vont changés leurs état selon les règles de transition définies par l'automate cellulaire qui vont tendre à regrouper des états similaires pour les cellules actives (contenant l'index des images). La classification est *recouvrant* (Les données peuvent apparaître plusieurs fois dans la grille)[17, 18].

5.7. La recherche d'images par le contenu

Maintenant nous présentons la méthode d'indexation et recherche d'images par le contenu, contrairement aux autres méthodes notre méthode consiste à utiliser les automates cellulaires pour un but de filtrage des images afin que la recherche soit rapide et pertinente.

L'idée principale repose sur l'utilisation de la classification par AC, lors d'une interrogation, un premier parcours de la carte de l'AC permet de sélectionner les approximations des données les plus pertinentes au vecteur requête au sens de la similarité.

Cette étape n'est pas coûteuse puisque elle est utilisée sur un ensemble de vecteurs sélectionnés, elle permet d'élire rapidement les vecteurs pouvant figurer parmi les plus proches voisins. Ensuite les

vecteurs correspondant sélectionnés dans la première phase sont visités séquentiellement pour calculer la similarité avec le vecteur d'image requête. Ainsi, le nombre d'E/S est réduit et le coût de calcul par rapport à une recherche séquentielle qui elle, analyse la totalité de la base.

Notre algorithme est composé de deux parties principales, la *phase de filtrage* et la *phase de recherche*. La première prépare un petit ensemble d'images candidates pour la deuxième phase, la phase de recherche. Cette dernière effectue une recherche exacte des images similaires [19].

5.7.1. Filtrage

La phase de filtrage est importante car elle détermine l'ensemble des vecteurs candidats qui vont être utilisés lors de la deuxième phase. Si cet ensemble possède trop peu d'images alors cet ensemble ne va pas contenir les vrais plus proches voisins, les résultats ne seront pas satisfaisants. D'un autre côté si l'ensemble des vecteurs est très important dans ce cas trop d'images devront être traitées ce qui aura pour conséquence l'explosion du temps de calcul de la distance. Détaillons maintenant le déroulement de cette phase :

Cette phase contient deux méthodes qui s'exécutent séquentiellement, la première elle détient l'automate cellulaire et les informations nécessaires au fonctionnement du système, elle consiste à affecter la requête à une cellule adéquate de l'AC, après l'affectation de la requête le système fait appel à la deuxième méthode.

La deuxième méthode compare la similarité entre le vecteur d'une cellule (le vecteur de l'image requête) et son voisinage selon une distance (Manhattan). Cette distance est utilisée dans le but d'éliminer certains vecteurs, elle compare la distance de la cellule en question avec une de son voisinage (voisinage de Moore), si au moins une cellule répond au critère alors tous les vecteurs des cellules voisines sont pris comme candidats. Sinon il s'arrête et la requête est affectée à une autre cellule. Ainsi, se ne sont que les cellules voisines, seulement, qui vont être explorées. Le grand avantage c'est d'éviter de faire des calculs inutiles sur des cellules vides et éloignées de l'image requête.

Au terme de cette phase, on aura déterminé l'ensemble des vecteurs candidats.

5.7.2. Recherche

Une fois la phase de filtrage se termine, on applique une recherche séquentielle sur l'ensemble des images candidates et afficherons le résultat.

5.8 Réalisation de SASDUM 4.1

Lors de ce travail nous avons rencontré d'énorme problème pour la collecte des images, ainsi que des difficultés dans le diagnostic de quelques images car le partenaire n'était pas très disponible du fait de son emploi du temps chargé.

Malgré toutes ces difficultés nous avons pu implémenter SASDUM 4.1. La séquence des figures ci-dessus présentent des pages écran des différentes étapes de l'approche d'indexation basée sur les automates cellulaires.

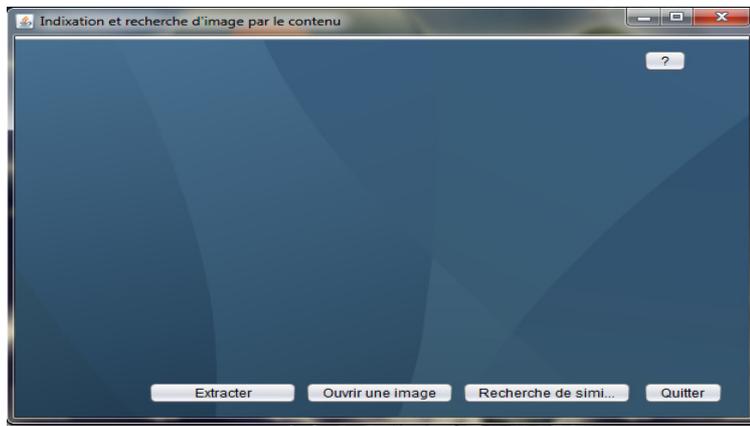


Fig.5.3.Interface du système

Homogenite	corelation	variance	moment des l...	Entropie
1195.98336	231.4186875...	851.8323528	290023.0	19.92274237...
1002.999996	200.7420570...	725.67850509	269665.0	17.45082153...
21.8900738	5.877904499...	10.40238381	40229.0	0.879120722...
104.03133	32.05172879...	91.8943415	84777.0	3.142846717...
196.0664148	45.38177999...	122.4651594	126919.0	4.916297566...
691.89117	128.6154407...	401.2968786	215891.0	12.80349586...
211.2218208	31.12307550...	60.87776454	105111.0	4.758696236...

Fig.5.4.Paramètres d'Haralick extraites de la matrice de co-occurrence



Fig.5.5.Exemple d'un motif de texture

Homogenite	corelation	variance	moment des l...	Entropie
949.9465549...	312.5437800...	480.9721698...	476576.0	16.05145772...

Fig.5.6.Vecteur descripteur de la texture

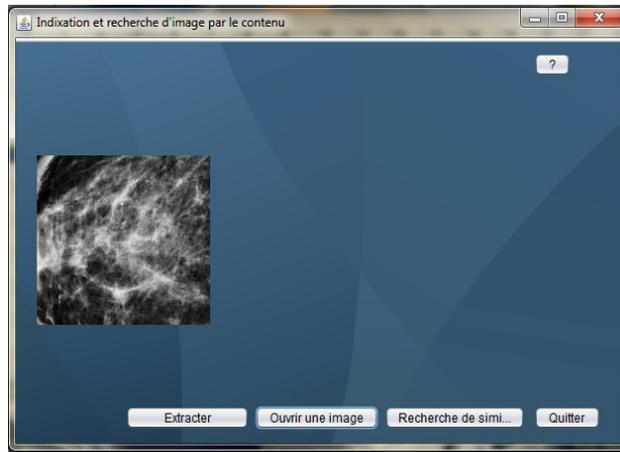


Fig.5.10.Requête entrée par l'utilisateur

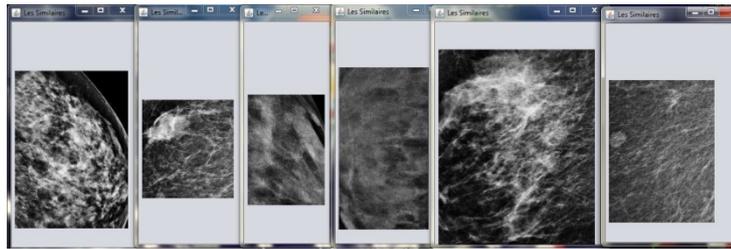


Fig.5.11.Images similaires trouvées

5.9 Analyse des résultats de la classification

L'analyse des résultats de classification va se faire en triant les réponses négatives et positives en fonction de la vraie réponse. Ceci conduit à séparer 4 catégories d'images :

- les images positives et identifiées comme telles : on parle de vrais positifs.
- les images négatives et identifiées comme telles : ce sont les vrais négatifs.
- les images positives classées comme négatives : ce sont les faux négatifs.
- les images négatives classées comme positives : ce sont les faux positifs.

Les effectifs dans ces 4 catégories permettent de calculer des index caractérisant la justesse de la classification, et en particulier :

- la *sensibilité*, est définie comme le pourcentage de cas positifs correctement identifié (la sensibilité mesure la probabilité conventionnelle que le test soit positif lorsque la maladie est présente).

$$Se = VP / (VP + FN)$$

- la *spécificité*, est définie comme le pourcentage de cas négatifs correctement identifiés (la spécificité mesure la probabilité conventionnelle que le test soit négatif lorsque la maladie est absente).

$$Sp = VN / (VN + FP)$$

Le taux de la classification correcte est défini comme suit :

$$CC = (VP + VN) / ((VP + VN) + (FP + FN))$$

Après expérimentation de différentes valeurs de seuil de similarité qui représente la distance entre deux images de la base. Si cette distance est inférieure ou égale au seuil alors les images sont similaires. Nous obtenons les résultats suivants :

Tab.5.1. Analyse de la classification

fenêtre	32 × 32	16 × 16
Sensibilité	89%	84%
Spécificité	84%	81%

Nous avons obtenus différents résultats de classification en fonction de fenêtres, ainsi, le résultat de l'évaluation dépend du seuil choisis. Simplement, choisir un bon seuil implique un bon résultat de classification.

5.10. Conclusion

Cette partie présente les étapes de notre travail, extraction de vecteur de caractéristique d'une image mammographie à partir de la texture en utilisant la matrice de Haralick. A l'aide de l'automate cellulaires, on obtient une classification de nos images mammographie, enfin on applique notre approche de recherche sur les images déjà classées afin de trouver l'ensemble des images similaires à une image requête introduite par l'utilisateur.

5.11. Références bibliographiques

1. Ahmedin,J., Freddie,B. :Global Cancer Statistics. A Cancer Journal for Clinicians, Vol. 61, 69-90(2011)
2. Le guide de la médecine et de la santé en Algérie .www.santemaghreb.com. (2011)
3. Abdalla, A., Deris, S., Zaki,N. :Breast Cancer detection based on statistical features and support vector machine. In 4th international conference on innovation in information technology ,728–730(2007)
4. Halkiotis,S.,Botsis,T.,Rangoussi,M. :Automatic detection of clustered microcalcifications in digital mammograms using mathematical morphology and neural networks Signal Processing, 87, 1559–1568(2007)
5. Manrique,D., Rios,J.,Rodriguez-Paton,A. :Evolutionary system for automatically constructing and adapting radial basis function networks Neurocomputing, 69 (October) ,2268–2283(2006)
6. Ramirez,N., Acosta-Mesa,H., Carillo-Calvert,H.,Nava-Fernandez,I.,Barrientos-Martinez,R. :Diagnosis of breast cancer using Bayesian networks: A case study Computers in Biology and Medicine, 37 ,1553–1564(2007)
7. Tourassi,G.,Haarawood,B.,Singh,S.,Lo,J.,Floyd,C. :Evaluation of information-theoretic similarity measures for content based retrieval and detection of masses in mammograms. Medical Physics,34,140–150(2007)
8. Azzag,H.,Picarougne,F.,Guinot,C.,Venturini,G.:Classification de données par automate cellulaire,(2006)
9. von Neumann,J. : Theory of Self-Reproducing Automata .University of Illinois Press,Urbana,Ill.,(1966)
10. Zuse,K.,Rechnender Raum, tome 1 de Schriften zur Datenverarbeitung (Friedrich Vieweg & Sohn, Braunschweig, 1969)
11. Haralick,R .M., Shanmugam,K., Dinstein,I. :Textural features for image classification, IEEE Trans Sys Man Cyb,610—21(1973)
12. Connors, R.W., Harlow, C.A.: A Theoretical Comparison of Texture Algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence (1980)
13. Weszka, J.S., Dyer, C.R., Rosenfeld, A.: A Comparative Study of Texture measures for Terrain Classification. IEEE Transactions on Systems Man and Cybernetics (1976)
14. Zucker,S.W.,Terzopoulos,D. : Finding structure in co-occurrence matrices for texture analysis. Computer Graphics and Image Processing, 12 :286.308(1980)
15. Beichel,R., Sonka,M. :Computer vision approaches to medical image analysis. Lecture Notes in Computer Science, Springer, 4241, (2006)
16. Benmazou sarah, Layachi Soumia &**H.F.Merouani**, “ Approche Automate Jeu de la Vie-SMA pour la Segmentation des Images IRM Cérébrales” The Second international conference on complex Systems, CISC’11,06-08 December 2011, Jijel, Algeria.
17. Benmazou Sara, **Layachi Soumia** &**Hayet Farida Merouani** “Segmentation in Images based Cellular Automata Reactive Agent Implemented in Netlogo Platform”, in International Journal of Computer Application, Vol.45, N°24, May, 2012, pp.40-43. **ISBN**: 978-93-80866-46-3, **ISSN** 09758887, indexed in Springer, INASP, DOAJ,LUND, BIOMed Central, Editeur. Foundation of Computer Science <http://www.ijcaonline.org>
18. Benmazou Sarah, **Layachi Soumia** & **H.F.Merouani** “Segmentation d’images sur un Couplage Automate-Agent”, ICIST’11, 1st International Conférence on Information Systems and Technologies, 24-27 April 2011, Tebessa, Algeria, pp 525-532.**ISBN : 978-9931-9004-0-5**
19. Benmazou Sarah, **Layachi Soumia** & **H.F.Merouani**, “ Automates Cellulaire SMA pour la Segmentation des images ” First International Conference on Signal, Image, Vision and their Applications, SIVA’11, at Guelma, November 21-24, 2011, Guelma, Algérie.

Sommaire du chapitre 6

1. Introduction
 2. Equipe de travail
 3. Démarche adoptée
 4. Résultats obtenus par la classification
 5. Conclusion
 6. Référence bibliographiques
-

6.1. Introduction

Y.Caron a prouvé que les lois de Zipf et de Zipf inverse peuvent être utilisées pour caractériser la fréquence d'apparition des motifs dans des images en fonction de distributions en loi de puissance [1, 2]. Ces lois permettent la caractérisation de la complexité structurelle des textures dans les images ; cette mesure de complexité pourrait indiquer une région d'intérêt dans les images mammaires. En effet, notre approche mesure la complexité structurelle des textures au sein des régions (imassettes), qui est toujours différente dans les régions portant des tumeurs et dans les régions de tissu sein, permettant la détection des tumeurs, qui consiste en une segmentation séparant les régions portant des tumeurs du reste du tissu du sein pour extraire les tumeurs.

6.2. Equipe de travail

Membres du Projet: Dr.H.F.Merouani, Dr.Y.Tlili

Etudiantes en Post –graduation (Doctorat LMD): S.Benmazou, M.Hamoud

Specialistes :

1.Dr.Djilani Nour el Houda :Médecin Spécialiste en radiologie

- Cabinet d'imagerie médicale, échographie radiologie spécialisée mammographie, spécialiste en radiologie et imagerie médicale, wilaya de Guelma.
- Etablissement public hospitalier Ibn Zohr, unité échographie, wilaya de Guelma.

2.Dr.Beledjhem Nedjmeddine : Maitre assistant hospitalo-universitaire

6.3. Démarche adoptée

La méthodologie proposée dans ce travail est basé sur une fusion de la loi de Zipf et la loi de Zipf inverse pour détecter la tumeur dans les images mammaires [3, 7, 8], ces lois localisent l'emplacement des tumeurs en utilisant une analyse d'imassettes composant le sein. Le processus de détection des tumeurs est illustré par la figure 6.1.

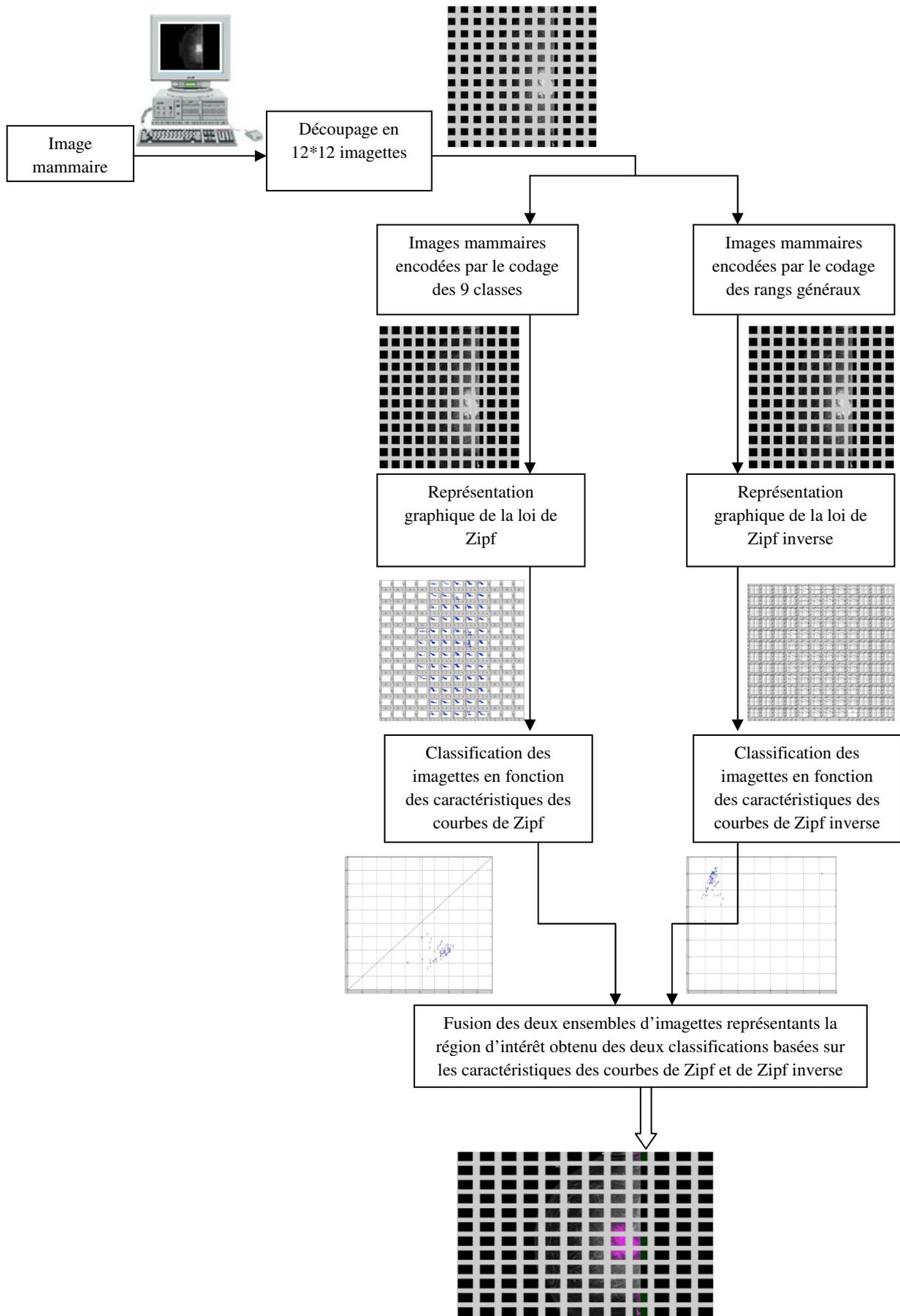


Fig.6.1. Organigramme de la fusion des lois de Zipf et de Zipf inverse pour la détection de la zone d'intérêt.

Le processus de détection comporte quatre étapes [10] :

- Dans la première étape, nous effectuons un découpage de l'image mammaire en imagerie.

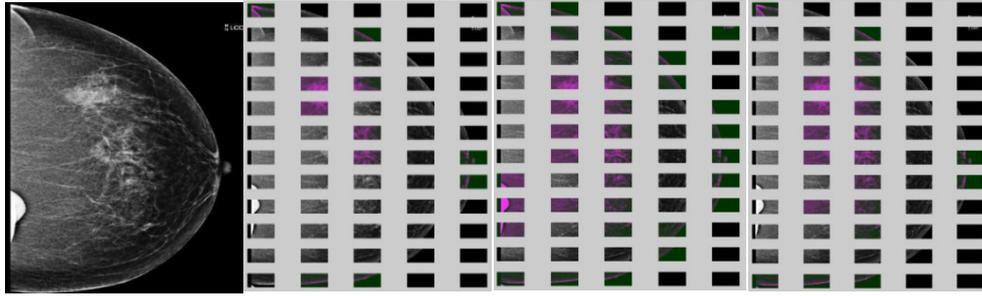


Fig.6.2. Détection de la tumeur sur trois images mammaires par les lois de Zipf: (a) l'image mammaire originale, (b) région d'intérêt obtenue par l'application de la loi de Zipf et le codage des 9 classes, (c) région d'intérêt obtenue par l'application de la loi de Zipf inverse et le codage des 9 classes, (d) région d'intérêt obtenue par la fusion de la loi de Zipf avec le codage des 9 classes et la loi de Zipf inverse avec le codage des rangs généraux.

- La deuxième étape consiste à réaliser la représentation graphique des lois de puissance : Zipf et Zipf inverse. La fusion des lois de Zipf se base sur les deux codages différents mais complémentaires appliquées aux niveaux de gris des images mammaires : le codage des 9 classes [3] et le codage des rangs généraux [7,8], le codage des 9 classes permet de mettre en évidence les principales structures de l'image, mais il ne permet pas d'étudier la texture fine de l'image et il perd des informations sur les petites variations de luminance, qui normalise les textures de l'image. Le codage des rangs généraux a les effets opposés du précédente, à s'avoir, qu'il révèle l'existence de légères variations de luminance, mais nous perdons des informations sur l'amplitude relative des variations. Ces deux codages sont complémentaires, chacun mettant en évidence une caractéristique particulière de l'image, ainsi nous avons pensé à une fusion de ces deux codages sur les images mammaires pour l'extraction d'une région d'intérêt afin d'obtenir de meilleurs résultats [12, 13].
- La troisième étape est une classification séparant les imagerie représentant des tumeurs des imagerie représentant un tissu normal du sein. Cette classification est basée sur les descripteurs texturales extraits des courbes de Zipf et de Zipf inverse, ces courbes sont générées à partir de chaque imagerie de l'image mammaire. Nous pouvons remarquer que les courbes de Zipf et de Zipf inverse de l'image mammaire sans tumeur, l'image mammaire portant une tumeur bénigne et l'image mammaire portant une tumeur maligne obtenue après le codage de la mammographie par la fusion du codage de l'écart-type et le codage des rangs généraux sont différentes. Donc, l'analyse des images mammaires avec les lois de Zipf et de Zipf inverse avec le codage des images mammaires par la fusion du codage de l'écart-type et le codage des rangs généraux peut être efficace pour la distinction entre les tumeurs malignes et les tumeurs bénignes ainsi que le tissu sein n'ayant pas de tumeurs, cela en utilisant des descripteurs calculés à partir des courbes de Zipf et de Zipf inverse de chaque image mammaire. Le descripteur de chaque image mammaire est défini par le tableau 6.1. Notre méthodologie pour la classification des images mammaires à des images mammaires portants des tumeurs bénignes, des tumeurs malignes ou des images mammaires normales sans lésions se base sur Template Matching approche [9, 10], nous utilisons une technique de recherche

pour trouver dans la base de données, les images mammaires qui sont pathologiquement similaires à chaque image mammaire requête.

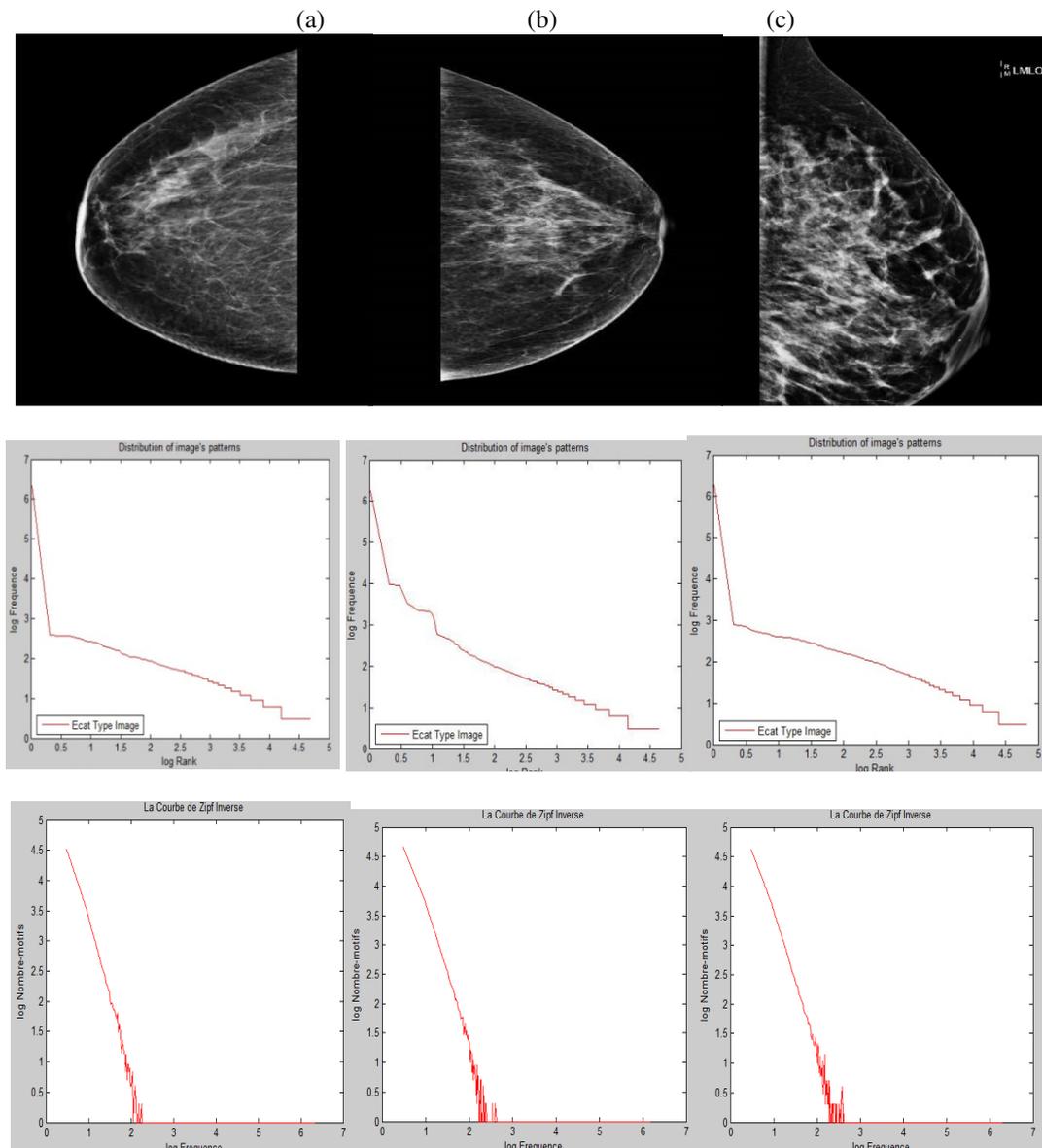


Fig.6.3. Courbes de Zipf et de Zipf inverse de trios images mammaires obtenus suite au codage des images mammaires par la fusion du codage de l'écart type et le codage des rangs généraux : (a) image mammaire sans tumeur, (b) image mammaire portant une tumeur bénigne et (c) image mammaire portant une tumeur maligne.

Nous assignons à la mammographie requête la classe de l'image mammaire la plus similaire. Comme mentionné sur la figure suivante, le système proposé consiste en deux phases : hors ligne, pour l'analyse des images mammaires par les lois de Zipf dans le but d'extraire les vecteurs descripteurs de texture pour chaque image mammaire dans la base de données. D'autre part nous distinguons la phase en ligne; dans cette phase, l'utilisateur intègre une image mammaire requête au système pour identifier son type. Pour chaque nouvelle image mammaire, un vecteur de descripteurs de texture est calculé et l'image mammaire est indexée dans la base de données. Les similarités entre le vecteur descripteur de l'image mammaire requête et les vecteurs descripteurs de chaque image mammaire dans la base de données indexée sont calculés par la distance Euclidienne.

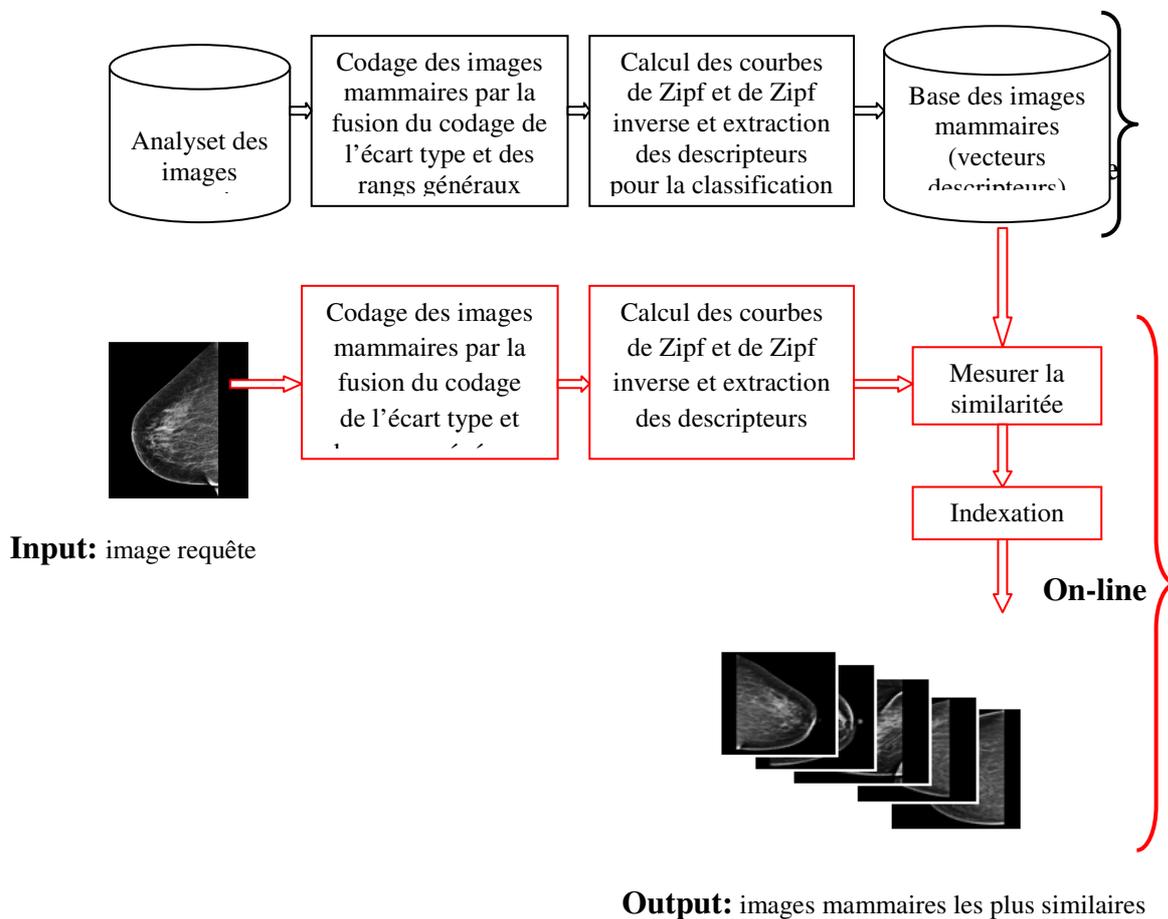


Fig.6.4. Architecture du système de recherché d'image mammaire par la méthode template Matching.

Tab 6.1 : Descripteurs mammaires

<p>L'aire délimitée par la courbe de Zipf</p>	<p>Un premier critère que l'on peut définir est l'aire délimitée par la courbe de Zipf. Soit n le nombre de motifs de la courbe, f_i la fréquence et r_i le rang du motif i, l'aire de la courbe est donnée par la formule (1).</p> $A = \sum_{i=1}^{n-1} \frac{(f_i + f_{i+1})(r_{i+1} - r_i)}{2} \quad (1)$
<p>La pente de la courbe de Zipf et de Zipf inverse</p>	<p>La pente moyenne d'une courbe est le coefficient directeur de la droite des moindres carrés. Elle est donnée par la formule (2).</p> $p = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (2)$
<p>L'entropie de la courbe de Zipf</p>	<p>Il est possible de définir une mesure d'entropie à partir des fréquences d'apparition des motifs de l'image. Deux formulations de l'entropie sont possibles, l'entropie relative aux motifs et l'entropie relative aux fréquences d'apparition.</p> <ul style="list-style-type: none"> ◆ <i>L'entropie relative aux motifs</i> L'entropie relative aux motifs est définie par la formule (3) pour un ensemble de R motifs distincts : $H_w = - \sum_{r=1}^R \frac{f(r)}{T} \log_R \frac{f(r)}{T}$ <p>Dans cette formule, $f(r)$ représente la fréquence du motif de rang r, T représente le nombre total de motifs différents ou non, et on utilise un logarithme de base R.</p> ◆ <i>L'entropie relative à la fréquence</i> L'entropie relative à la fréquence est définie par la formule (4) : $H_f = - \sum_{f=1}^F \frac{I(f)}{R} \log_F \frac{I(f)}{R}$ <p>$I(f)$ représente le nombre de motifs distincts de fréquence f et F le nombre total d'occurrences des motifs dans l'image. On utilise un logarithme à base F afin que l'entropie soit comprise entre 0 et 1</p>

6.4 Résultats obtenus par la classification

Notre système d'aide au diagnostic médical du cancer du sein permet aux radiologues de chercher des mammographies qui sont pathologiquement similaire à une image mammaire requête, une interface avec un format convivial pour fournir une utilisation efficace par le radiologue a été réalisée.

La base de données présentée dans le deuxième chapitre a été utilisé pour les expérimentations, nous avons utilisé 75 mammographies pour extraire leur vecteurs descripteurs composé d'un ensemble de caractéristiques de texture obtenus suite à l'analyse des images mammaires par les lois de puissance : Zipf et Zipf inverse avec comme codage des motifs des images mammaires une fusion du codage de l'écart type et le codage des rangs généraux, nous avons utilisé 33 images mammaires requêtes pour l'étape la recherche . Le radiologue introduit une image mammaire requête et obtient comme résultat les images mammaires les plus similaires à l'image mammaire requête dans la base d'images mammaires ; en fonction d'une mesure de similarité. Nous avons utilisé la distance euclidienne pour mesurer la distance entre les vecteurs descripteurs de deux images mammaires..

Nous avons obtenu un taux de classification de 90%. Nous illustrons dans la figure 6.5 l'acquisition d'une image mammaire requête dans notre système.

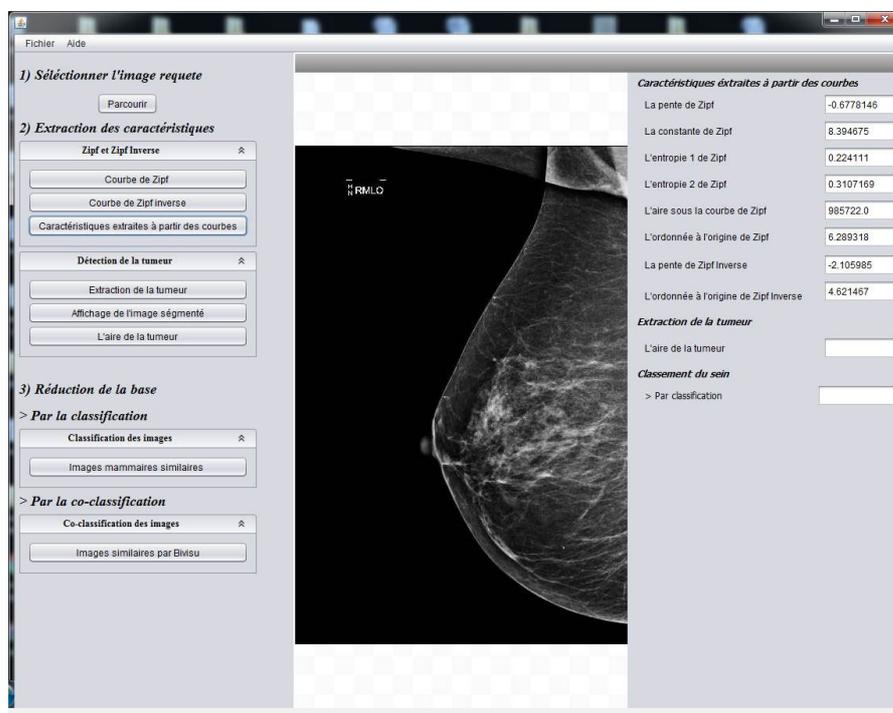


Fig 6.5. Acquisition de l'image mammaire requête

Le résultat de la recherche de l'image mammaire la plus similaire à une image mammaire requête est donné par la figure 6.6 où nous distinguons une bonne classification de l'image mammaire requête vu qu'elle porte une tumeur maligne et que le résultat obtenu de l'image mammaire la plus similaire dans la base des images mammaires est aussi une image mammaire portant une tumeur maligne.

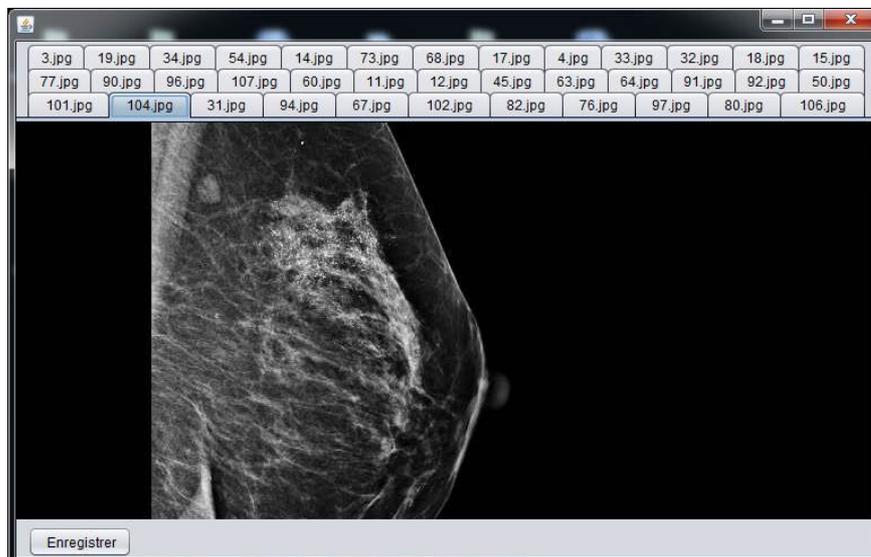


Fig.6.6 Résultat de recherche des images mammaire les plus similaires à l'image mammaire requête

6.5 Conclusion

Notre approche de détection et de classification des masses au sein des mammographies est différente des autres approches qu'on trouve dans la littérature. En effet, nous analysons des régions au sein des mammographies par les lois de puissance: Zipf et Zipf inverse qui génèrent la représentation des régions selon des courbes, à partir desquelles nous extrayons les descripteurs représentant ces régions et permettant la distinction entre les régions constituant la masse et les régions du tissu normal. Ces lois de puissance permettent de caractériser la complexité structurelle de texture dans les images mammaires. En effet, l'analyse de la texture des images mammaires est une technique importante représentant les propriétés de niveau de gris des images mammaires ; sachant que l'anatomie du sein se compose de plusieurs structures linéaires causant des orientations directionnelle de la texture dans les mammographies, d'autre part, les masses provoquent des distorsions et des changements de texture dans le tissu du sein, ce qui signifie que la perturbation de l'homogénéité de la texture du tissu mammaire est un moyen destiné à détecter des distorsions architecturales au sein représentant les masses à travers un ensemble de descripteurs de texture extraites à partir des courbes de Zipf et de Zipf inverse des images mammaires .

Nous avons proposé un système de diagnostic utilisant une classification des images mammaires basée sur la recherche d'image mammaire par la méthode template Matching.

Nous avons obtenu un têt de classification de 90%.

6.6 Références bibliographiques

- [1] Y. Caron, P. Makris, N. Vincent. "A method for detecting artificial objects in natural environments", Proceedings of International Conference on Pattern recognition (ICPR-IAPR) Québec, pp. 600-6003.2002.
- [2] Y. Caron, P. Makris, N. Vincent, "Use of power law models in detecting region of interest", the journal of the Pattern Recognition 40 (2007) 2521 – 2529.
- [3] D. Bi. Segmentation d'images basée sur les statistiques de rangs des niveaux de gris. PhD thesis, Université de Tours, Janvier 1997.
- [4] Hanifi Majdoulayne. Extraction de caractéristiques de texture pour la classification d'images satellites. PhD thesis, Université Toulouse III, Novembre 2009.
- [5] Biomedical Information Technology, 2008, Pages 83-113 Tom Weidong Cai, Jinman Kim, David Dagan Feng.
- [6] M.O. Honda, P.M. de Azevedo Marques, J.A.H. Rodrigues, Content-based image retrieval in mammography: using texture features for correlation with BI-RADS categories, in: H.-O. Peitgen (Ed.), Proceedings of the Sixth International Workshop on Digital Mammography: IWDM 2002, Springer, Bremen, Germany, 2002, pp. 231–233.

- [7] Chia-Hung Wei, Yue Li, Pai Jung Huang. Mammogram retrieval through machine learning within BI-RADS standards. Journal of Biomedical Informatics 44 (2011) 607–614.
- [8] Hamoud Meriem & **H.F.Merouani**, **Détection d'une Zone d'intérêt dans les images par la Loi de Zipf et la Loi de Zipf Inverse**, 1ère Journées Doctorales du Laboratoire d'Informatique d'Oran, JDLIO'11, 31 MAI- 01 Juin 2011, Oran, Algérie. www.univ-oran.dz/rencontres/jdllo/index.html
- [9] Hamoud Mériem, **Hayet Farida Merouani** “Detection of a region of interest in the Image Based on Zipf Laws”, SITIS'11 at The 7th International Conference on Signal Image Technology & Internet Based Systems, Edited by Kokou Yelongnon, Richard Chbeir and Albert Dipanda, Dijon, France, 28 November-1 December 2011, pp.416-421, ISBN:978-0-7695-4635-3/11,2011 IEEE DO10.1109, Indexed in **Scopus, IEEE explore and in ACM DL**.
- [10] Hamoud Meriem & **H.F.Merouani**, “ Détection d'une Zone d'intérêt dans les images basée loi de Zipf” Conférence Maghrébine sur l'Extraction et la Gestion des Connaissances, EGC'11,23-25 Novembre 2011, Tanger, Maroc
- [11] Hamoud Meriem & **H.F.Merouani**, “Computer Aided to medical diagnosis of Mammogram masses by the Biclustering and the power Laws” Workshop sur la Modelisation Mathématique et Controle, Université Badji Mokhtar Annaba, 20-22 Mai 2013, Annaba, Algeria.
- [12] Hamoud Meriem & **H.F.Merouani**, “ Computer Aided Detection of Mammogram masses by the power Laws” Rencontre des Femmes Scinetifiques Mediteraneennes, REFSCIME 2013, Université 209 Aout 1955 Skikda, 15-16 Avril 2013, Skikda, Algeria.

Suivi de l'évolution de la tumeur par une approche non rigide

Sommaire du chapitre 7

1. Introduction
 2. Problème de la phase de suivie
 3. Objectif de la phase de suivie
 4. Suivi basé sur un recalage non rigide
 - 4.1. Equipe de travail
 - 4.2. Démarche adoptée
 - 4.3. Réalisation et résultats obtenus
 5. Conclusion
 6. Références bibliographique
-

7.1. Introduction

Durant la période du traitement par chimiothérapie deux ou plusieurs examens de mammographies sont effectués. Le radiologue analyse les clichés pour localiser les anomalies potentiels puis les comparer pour suivre l'évolution des anomalies détectées. Cette tâche n'est pas aisée à automatiser. L'œil humain effectue quasi-automatiquement des traitements compliqués qu'il faut reproduire.

Notre approche s'inspire du comportement global du radiologue lors d'un examen de mammographie. On commence par une étape d'alignement des deux images (deux mammographies d'une même patiente acquises à des dates différentes), suivi par une phase de détection de la tumeur dans les deux images, et enfin l'évolution est menée au niveau des pixels en comparant les dimensions des tumeurs détectées dans les deux mammographies acquises avant et après traitement.

7.2. Problème de la phase de suivi

La projection d'une réalité 3D sur une image 2D entraîne des différences liées à l'orientation de la prise de vue, et ne permet pas d'analyser une paire de clichés par une simple différence entre images.

Le processus automatique de suivi de l'évolution d'une tumeur du sein passe par une phase de mise en correspondance des deux mammographies, un alignement des composantes présentes sur les clichés afin qu'elles occupent la même position sur les deux images.

Cette technique, appelée recalage, repose sur le calcul d'une fonction de transformation spatiale entre les deux images afin de les superposer à l'optimum leurs critères de ressemblance [1].

Le recalage est un problème classique en vision par ordinateur qui intervient notamment dans de nombreuses tâches d'analyse et de traitement des images médicales. Le but d'un algorithme de recalage d'images est de trouver une transformation géométrique reliant deux ou plusieurs images (Fig 7.1). En d'autre terme, le recalage d'images pourrait se résumer par le déplacement ou la déformation d'une image afin qu'elle ressemble le plus possible à une autre.

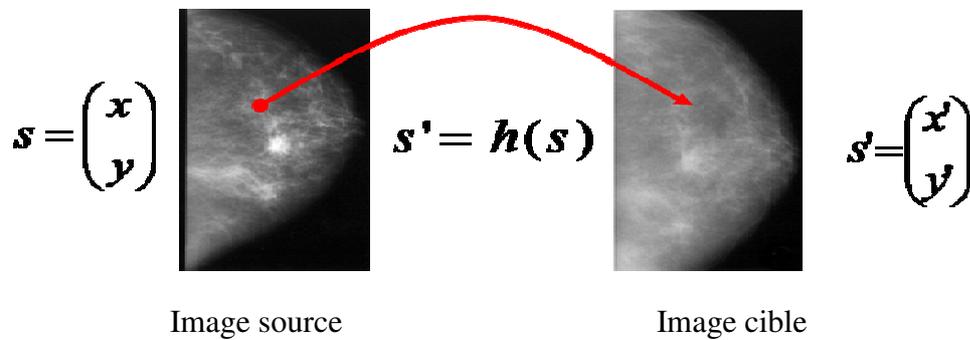


Fig.7.1: Principe de recalage

Les méthodes de recalage se répartissent en deux grandes classes selon le critère de similarité utilisé: [4]

- Les méthodes géométriques: Ces méthodes procèdent en deux étapes. La première est une étape de segmentation, visant à extraire de chacune des images des structures homologues ou amers: il peut s'agir de points, de courbes, ou de surfaces anatomiquement pertinents et présentant une certaine invariance morphologique. Dans un deuxième temps, on recherche la transformation effectuant le meilleur appariement de ces structures au sens d'un critère donné.
- Les méthodes iconiques: Le principe des méthodes iconiques est de chercher une transformation maximisant un critère de similarité au niveau pixel entre deux images.

Comme on peut distinguer deux différentes approches de déformation appliquées aux images recalage rigide, et recalage non rigide.

- Dans le cas de transformation rigide, on considère que les structures présentes dans les images sont fixes les unes par rapport aux autres et ne subissent aucune transformation élastiques : déformations des organes due à la respiration, à la flexibilité des tissus ou à des rotations articulaires [2]. La transformation rigide consiste en l'estimation de translations et de rotations, si on ajoute une homothétie la transformation est affine [3].
- Une transformation non rigide est utilisée dans le cas où il y a une déformation des organes due à la respiration, à la flexibilité des tissus ou à des rotations articulaires [2].

7.3. Objectif de la phase de suivie

L'objectif principal de la phase de suivi est de concevoir et d'implémenter une approche de recalage. Deux approches de recalage sont possibles : rigide et non rigide. Après réflexion et discussion nous avons opté pour étudier les deux approches. Ainsi, deux équipes se sont formées ; une équipe traite le recalage non rigide et l'autre traite le recalage rigide. Par conséquent, le résultat de cette phase de SASDUM est double :

- ✓ SASDUM 5.1 : qui est la partie de SASDUM qui assure le suivi de l'évolution de la masse mammaire par application d'une démarche de recalage non rigide. La suite de ce chapitre détaille cette approche.
- ✓ SASDUM 5.2 : qui est la partie de SASDUM qui assure le suivi de l'évolution d'une masse mammaire par application d'une démarche de recalage rigide. Le huitième chapitre détaille cette approche.

7.4. Suivi basé sur un recalage non rigide

7.4.1. Equipe de travail

Le travail est supervisé par Dr.Melouah, Dr Merouani H.F, Dr Layachi S. et de la doctorante Salhi Amina.

7.4.2. Démarche adoptée

Dans cette approche (Fig.7.2) nous avons utilisé un recalage rigide basé sur la maximisation de l'information mutuel [5], suivi par un recalage non rigide dans lequel nous avons adopté les étapes suivantes:

- Détection et mise en correspondance de points d'intérêt selon le principe de la méthode SIFT[6]
- Elimination des fausses correspondances à l'aide de la méthode GTM [7]
- Calcul d'une transformation non rigide basée sur les splines de plaque mince [8]
- Enfin nous avons utilisé la méthode de segmentation par croissance de région [9] pour détecter la tumeur dans les deux images alignées

Chaque étape de ce processus sera détaillée séparément dans les prochaines sections.

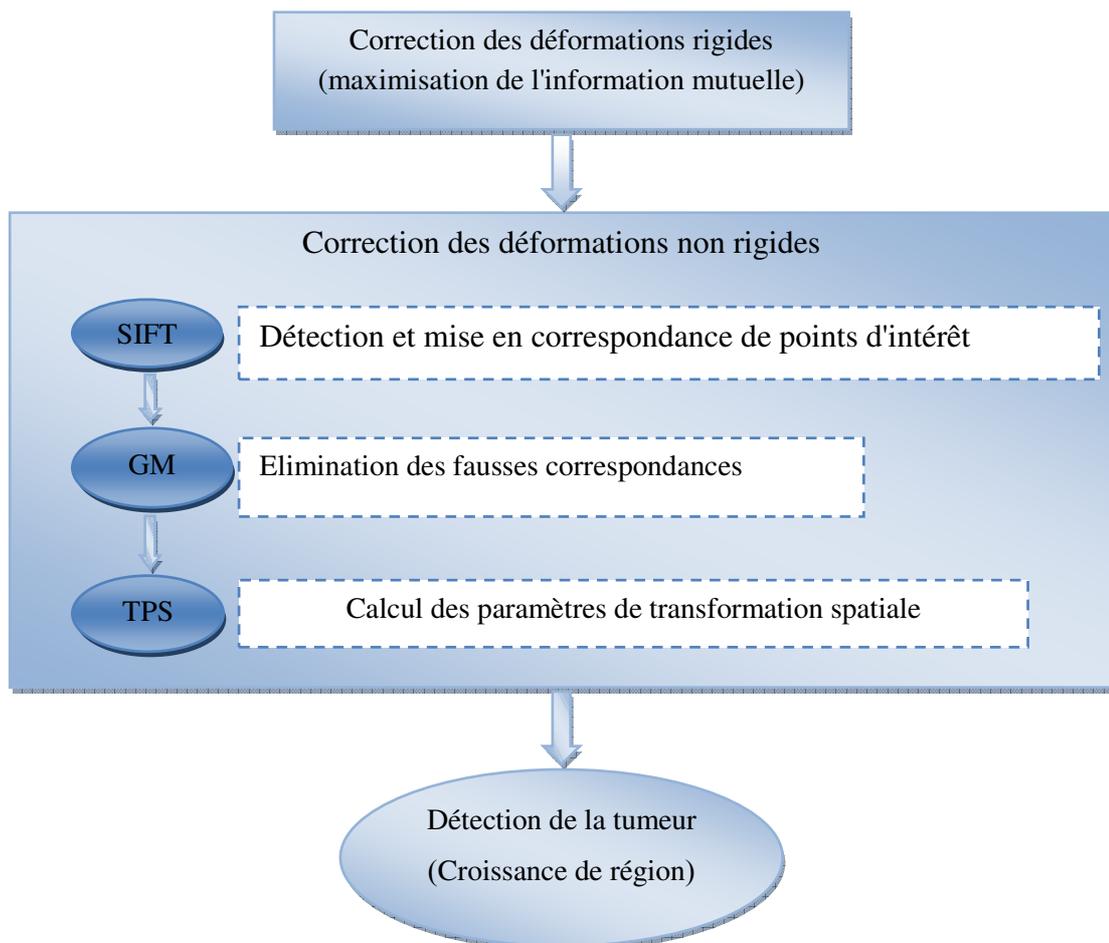


Fig. 7.2: Principe générale de notre approche

a. Correction des déformations rigides

L'information mutuelle est la quantité d'information d'une image contenue dans une seconde image. Ainsi, lorsque l'information mutuelle entre deux images est au maximum, elles sont identiques [5]. Tenter de transformer une image de manière à ce qu'elle ressemble le plus à une seconde image peut donc se faire en maximisant leur information mutuelle.

L'information mutuelle entre les deux images est calculée en utilisant la formule suivante [10]:

$$MI(A, B) = H(A) + H(B) - H(A, B) \quad (1)$$

Les entropies séparées $H(A)$ et $H(B)$ mesurent la complexité des images A et B. L'entropie jointe $H(A, B)$ mesure la quantité d'information que les images A et B apportent en même temps. Si les images A et B sont proches, une image explique bien la seconde, et l'entropie jointe est minimale. L'entropie jointe est plus précisément l'entropie de l'histogramme joint des images A et B. Cet histogramme répertorie la correspondance des intensités des pixels entre deux images (Fig.7.1). Maximiser l'information mutuelle selon cette troisième définition revient donc à dire que le plus d'information est recherchée dans chaque image, tout en cherchant à faire ressembler les images A et B.

b. Correction des déformations non rigides

Les seins sont des organes élastiques, donc toutes les recherches basées sur la présomption qui prend les seins comme des corps rigides sont fondamentalement inappropriés [11,12]. En conséquence, de nombreuses techniques non-rigides ont été proposées [13,14].

Dans notre approche la correction des déformations non rigides se fait en plusieurs étapes, on commence par une phase de détection et mise en correspondance de points d'intérêt, suivi d'une phase d'élimination des fausses correspondances, enfin on calcul la transformation spatiale

- **Détection et mise en correspondance de points d'intérêt : Scale-invariant feature transform (SIFT)**, est un algorithme utilisé pour détecter et identifier les éléments similaires entre différentes images. Il a été développé en 1999 par le chercheur David Lowe [6]. L'étape fondamentale de la méthode proposée par Lowe consiste à calculer ce que l'on appelle les « descripteurs SIFT » des images à étudier. Il s'agit d'informations qui caractérisent le contenu visuel des images de la façon la plus indépendante possible de l'échelle (« zoom » et résolution du capteur), du cadrage, de l'angle d'observation et de l'exposition (luminosité). Il s'agit tout d'abord de détecter sur l'image des zones circulaires intéressantes, centrées autour d'un point-clé et de rayon déterminé appelé facteur d'échelle (voir Fig. 7.3). Celles-ci sont caractérisées par leur unité visuelle et correspondent en général à des éléments distincts sur l'image. Sur chacune d'elles, on détermine une orientation intrinsèque qui sert de base à la construction d'un histogramme des orientations locales des contours, habilement pondéré, seuillé et normalisé pour plus de stabilité. C'est cet histogramme qui, sous la forme d'un vecteur à 128 dimensions (ou valeurs), constitue le descripteur SIFT du point-clé. L'ensemble des descripteurs d'une image établissent ainsi une véritable signature numérique du contenu de celle-ci. Ces descripteurs présentent l'avantage d'être invariants à l'orientation et à la résolution de l'image, et peu sensibles à son exposition, à sa netteté ainsi qu'au point de vue 3D. Une fois le calcul des descripteurs effectué, l'algorithme de mise en correspondance intervient dans le but de rechercher les zones d'une image dite suspecte qui contiennent des éléments visuellement similaires.

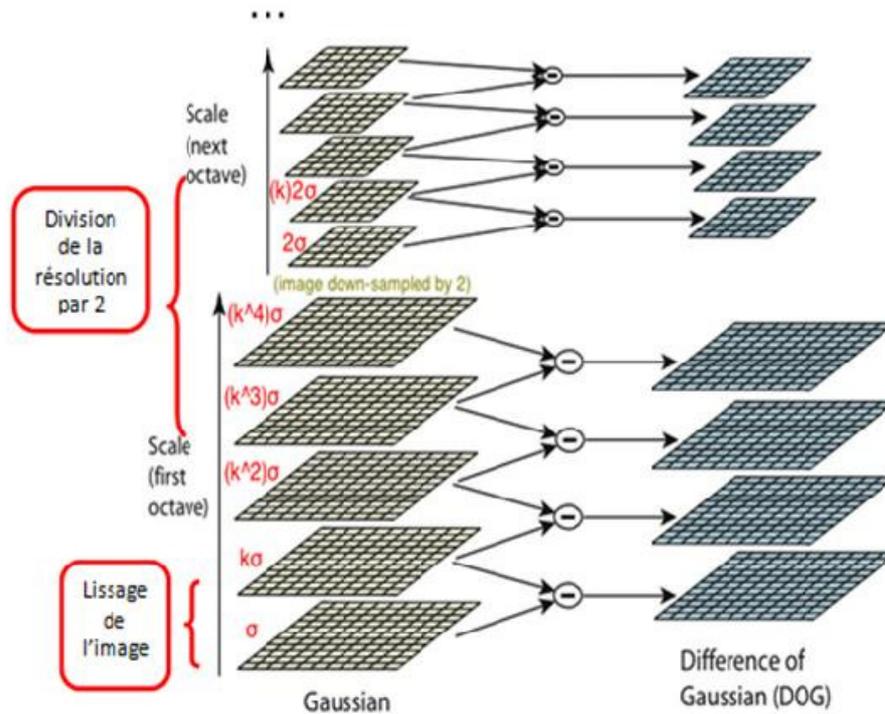


Fig.7.3: Détection multi-échelle des points d'intérêt

- **Elimination des fausses correspondances** : Dans certains cas, l'algorithme SIFT peut générer des fausses correspondances qui peuvent endommager notre transformation plus tard. Pour supprimer ces points nous adoptons un algorithme d'appariement de graphes nommé GTM [15]. Son principe est d'exploiter les relations spatiales cohérentes des points correspondants entre les deux images. GTM repose sur l'hypothèse que la transformation qui se produit entre les deux images est assez lisse, de sorte que les points voisins dans l'image de référence correspondent à des points voisins dans la deuxième image, par exemple: $Q(q_i)$ et $Q'(q'_i)$ de taille N (où q_i correspond à q'_i) sont deux ensembles de points correspondants (résultat de SIFT). Si les résultats sont correctes, les graphes $G(Q)$ et $G(Q')$ sont isomorphes, par ailleurs, la structure de $G(Q)$ est différente de celle de $G(Q')$. Les principales étapes de l'algorithme GTM sont les suivantes:
 - construire le graphe de k plus proche voisin pour les deux images $G_Q(V_Q, E_Q)$ et $G_{Q'}(V_{Q'}, E_{Q'})$
 - construire la matrice d'adjacence $R = |A_Q - A_{Q'}|$, et sélectionner la colonne j^{out} qui donne le nombre maximal de différents bords des deux graphes. Autrement dit, si la colonne j^{out} est conforme à la formule suivante: $j^{out} = \arg \max \sum_1^N R(i, j)$, alors la paire $(V_{j^{out}}, V'_{j^{out}})$ est considéré comme incorrecte
 - supprimer la paire $(V_{j^{out}}, V'_{j^{out}})$, ainsi que leurs correspondances.
 - Diminuer le N , et répétez les étapes ci-dessus jusqu'à $R(i, j) = 0, \forall i, j$.
- **Calcul des paramètres de transformation spatiale** : Pour calculer les paramètres de transformation à partir des points détectés nous avons utilisé la méthode thin plate spline[8]. La méthode de Spline en Plaque Mince ou « thin plate spline » en anglais est une fonction d'interpolation de la forme :

$$Z(s) = Z(x, y) = a_0 + a_1x + a_2y + \sum_{i=1}^n b_i e(s, s_i), \quad s \in D \quad (2)$$

avec $e(s, s_i) = |s - s_i|^2 \ln(|s - s_i|)$, qui minimise:

$$\sum_{i=1}^n [Z(s_i) - Z(s_i)]^2 + \rho \iint \left\{ \left(\frac{\partial^2 Z}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 Z}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 Z}{\partial y^2} \right)^2 \right\} dx dy \quad (3)$$

où ρ est un paramètre de lissage fixé a priori. L'intégrale incluse dans ce critère à minimiser est appelée « l'énergie de flexion ». Elle réfère au concept physique d'énergie de flexion d'une mince plaque de métal. Les surfaces générées par ces splines peuvent donc être comparées à des plaques de métal flexible ajustées de façon à minimiser leurs degrés de courbure tout en passant à proximité des points d'observation.

c. Détection de la tumeur

Cette opération est basée sur la segmentation par croissance de région [9]. Cette dernière exige la détermination de deux paramètres : la position du germe et le critère d'homogénéité.

Vu que la zone de la tumeur est préalablement identifier par un radiologue, le germe est placé manuellement dans la tumeur. Nous avons adopté le critère d'homogénéité suivant, présenté par Shen et all [16]

$$(1 - t) \frac{V_{max} + V_{min}}{2} \leq I(i, j) \leq (1 + t) \frac{V_{max} + V_{min}}{2} \quad (4)$$

7.3.2. Réalisation et Résultats obtenus de SASDUM 5.1

L'implémentation de l'approche proposée a permet de générer SASDUM 5.1. la figure 7.4 présente sa page écran principale.

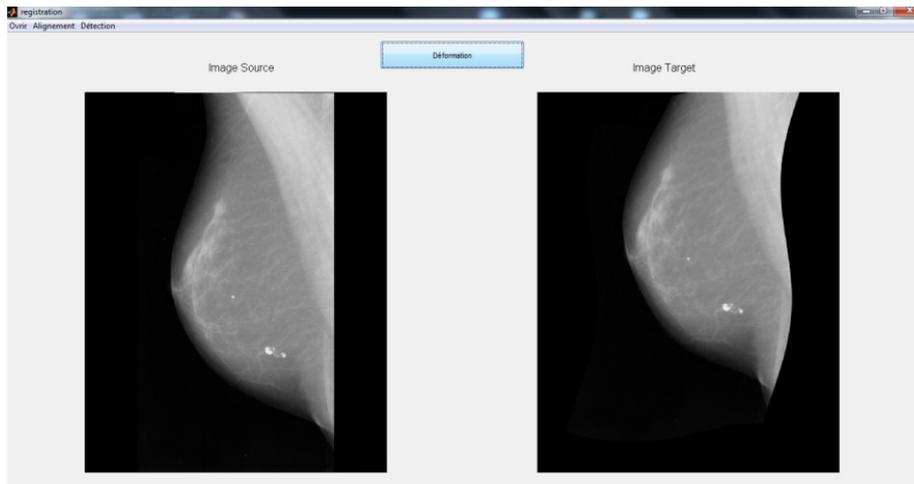


Fig.7. 4: Page principale de SASDUM 5.1

Le premier résultat de SASDUM 5.1 est une déformation rigide illustrée par la figure 7.5.

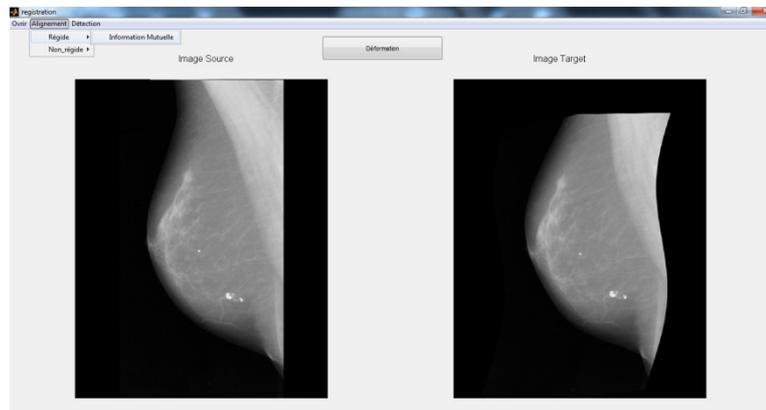


Fig. 7.5: Résultat de la déformation rigide de SASDUM 5.1

Le processus important de SASDUM 5.1 est la déformation non rigide. Il début par une extraction des points d'intérêt visualisé par la figure 7.6 (a) suivi d'une mise en correspondance figure 7.6 (b). SIFT est adopté pour ces deux opérations.

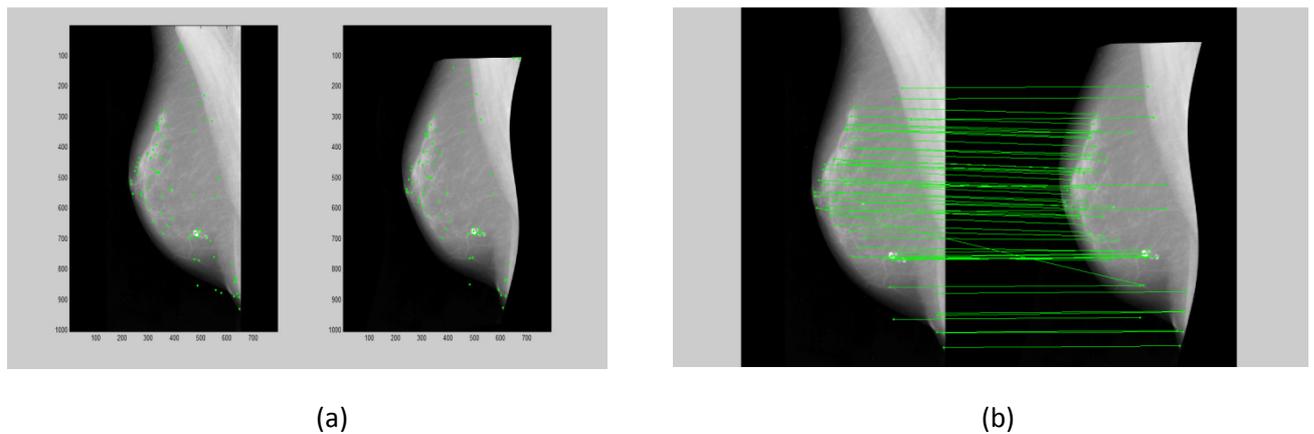


Fig. 7.6: Résultat de SASDUM 5.1 pour (a) la détection des points d'intérêts (b) la mise en correspondance des points détectés

GTM est par la suite appliqué pour éliminer les fausses correspondances. La figure 7.7 affiche les pages écran après application de GTM.

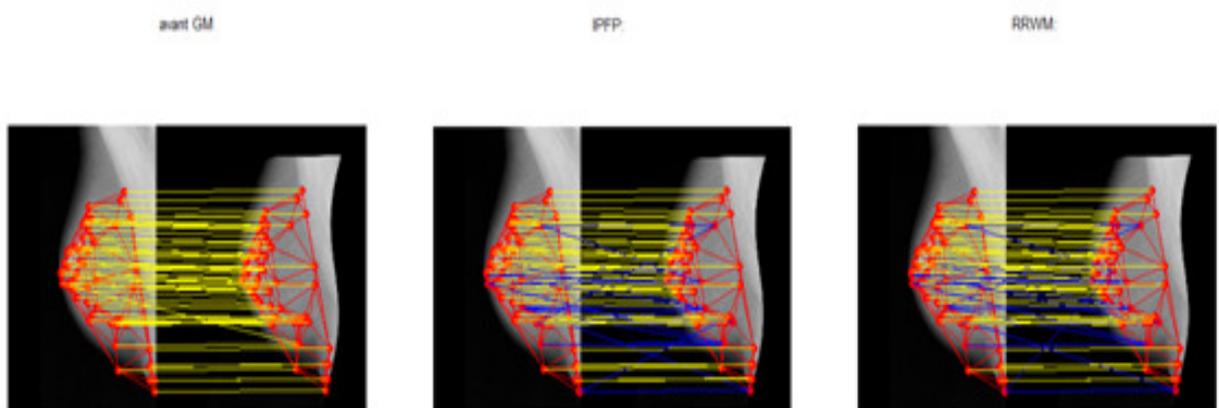


Fig. 7.7: Résultat de SASDUM 5.1 après application de GTM

La technique TPS utilise les points d'intérêts déterminés afin de déformer élastiquement l'image. Un tel résultat est présenté par la figure 7.8.

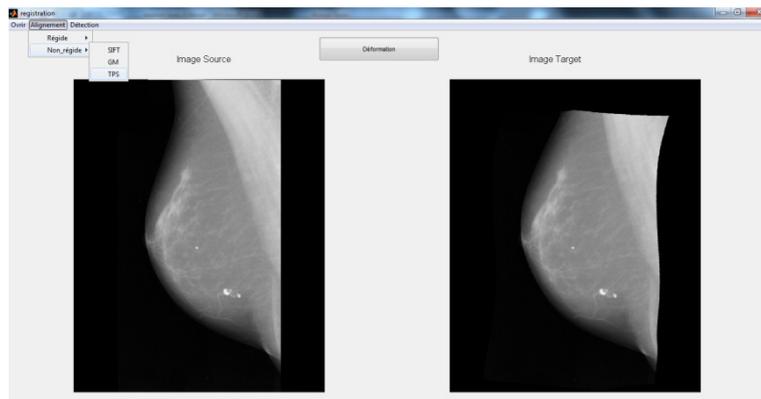


Fig. 7.8: Résultat de SASDUM 5.1 après application de la déformation non rigide

Enfin il est possible de cerner la tumeur sur les deux images par application de la méthode de segmentation par croissance de régions figure 7.9.



Fig.7.9 : Résultat de détection de la tumeur par SASDUM 5.1

7.4. Conclusion

Ce chapitre présente les différentes étapes de SASDUM 5.1 partie de SASDUM pour le suivi d'une tumeur mammaire par application d'un recalage non rigide. A l'aide de la méthode SIFT on obtient les point d'intérêts dans les deux image, suivi d'une élimination des fausse correspondance avec la méthode GTM, enfin on applique l'interpolation de TPS pour avoir notre image corrigée.

7.5 Références bibliographiques

- [1] P.I. Joseph V. Hajnal, Derek L.G. Hill. " Medical Image Registration", Edition Boca Raton : CRC Press, 2001.
- [2] A. ROCHE, "RECALAGE D'IMAGES MEDICALES PAR INFERENCE STATISTIQUE", PHD thesis, France, 2002.
- [3] H. Ngo, "RIGID TRANSFORMATIONS ON 2D DIGITAL IMAGES", PHD thesis, France, 2013.
- [4] C. Petitjean, "Recalage non rigide d'images par approche variationnelles statiques", PHD thesis, France, 2003.

- [5] M. RUBEAUX, "Approximation de l'Information Mutuelle basée sur le développement d'Edgeworth", PHD thesis, France, 2011.
- [6] Lowe D.G., "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol. 60, n0 2, 2004, p. 91-110.
- [7] Wendy Aguilar, Yann Frauel, Francisco Escolano "A robust Graph Transformation Matching for non-rigid registration". Image and Vision Computing, 27 :897–910, 2009.
- [8] L. Bookstein, "Principal warps: Thin plate splines and the decomposition of deformations". IEEE trans. PAMI, II(6):567-585, 1989.
- [9] F. Wahl, Digital Image Signal Processing. Artech House, 1987.
- [10] Bernard Colin, Marylène Troupé et. Jean Vaillant, " Information mutuelle et divergence : estimation, codage et classification".2005
- [11] Stamatakis EA, Ricketts IW, Cairns AY, Walker C, Preece PE, "Detecting Abnormalities on Mammograms by Bilateral Comparison, IEE Colloquium on Digital Mammography," London, UK, Mar. 27, 1996, pp.12/1-12/4.
- [12] Vujovic N, Brzakovic D, "Establishing the Correspondence between Control Points in Pairs of Mammographic Images," IEEE Trans. Image Processing, Vol.6, No. 10, pp.1388-1399, 1997.
- [13] Kostelec PJ, Weaver JB, Healy DM, "Multiresolution Elastic Image Registration," Medical Physics, Vol.25, No.9, pp.1593-1604, 1998.
- [14] Sallam MY, Bowyer KW, "Registration and Difference Analysis of Corresponding Mammogram Images," Medical Image Analysis, Vol.3, No.2, pp. 103-118, 1999.
- [15] Wendy Aguilar, Yann Frauel, Francisco Escolano, M. Elena Martinez-Perez, Arturo Espinosa-Romero, Miguel Angel Lozano, A robust Graph Transformation Matching for non-rigid registration, Image and Vision Computing, Vol. 27, pp. 897–910, 2009.
- [16] L. Shen, R. Rangayyan, and J. Desautels, "Detection and classification of mammographic calcifications," International Journal of Pattern Recognition and Artificial Intelligence, vol. 7, no. 6, pp. 1403–1416, 1993.

Sommaire du chapitre 8

1. Introduction
 2. Equipe de travail
 3. Démarche adoptée
 4. Réalisation
 5. Résultats obtenus
 6. Conclusion
 7. Références bibliographique
-

8.1. Introduction

Ce chapitre détaille une approche de recalage rigide de mammographies prises à des instants différées. L'approche proposée inclue trois étapes. La première consiste à découper l'image en imagerie. La deuxième utilise l'information mutuelle afin de déterminer quelle transformation est appropriée pour chaque imagerie. Enfin, la transformation majoritaire sera utilisée pour déformer l'image complète.

8.2. Equipe de travail

C'est Dr Melouah membre du projet PNR qui c'est chargé de la conception de la démarche de recalage et c'est sous ses directives que plusieurs étudiants en Master 2 ont travaillé. Bordjiba Abdelah, Arif Hichem, Djedaia Houria, Saoudi Maria, Atik Amina et Boumaza Rahma sont les étudiants en Master 2 qui ont participé dans l'implémentation de SASDUM 5.2.

8.3. Démarche adoptée

La figure 8.1 présente le schéma général de SASDUM 5.2. Il inclut les étapes suivantes :

1. Découpage de l'image source et de l'image cible en imagerie de mêmes tailles.
2. Transformation et interpolation des imagerie prisent deux à deux de l'image source et de l'image cible.
3. Calcul de l'information mutuelle de chaque paire d'imagerie et opter pour la transformation la plus majoritaire.
4. Optimisation par la méthode simplexe des paramètres de la transformation choisie.
5. Transformation et interpolation de l'image source en utilisant la transformation choisie avec les paramètres optimaux.

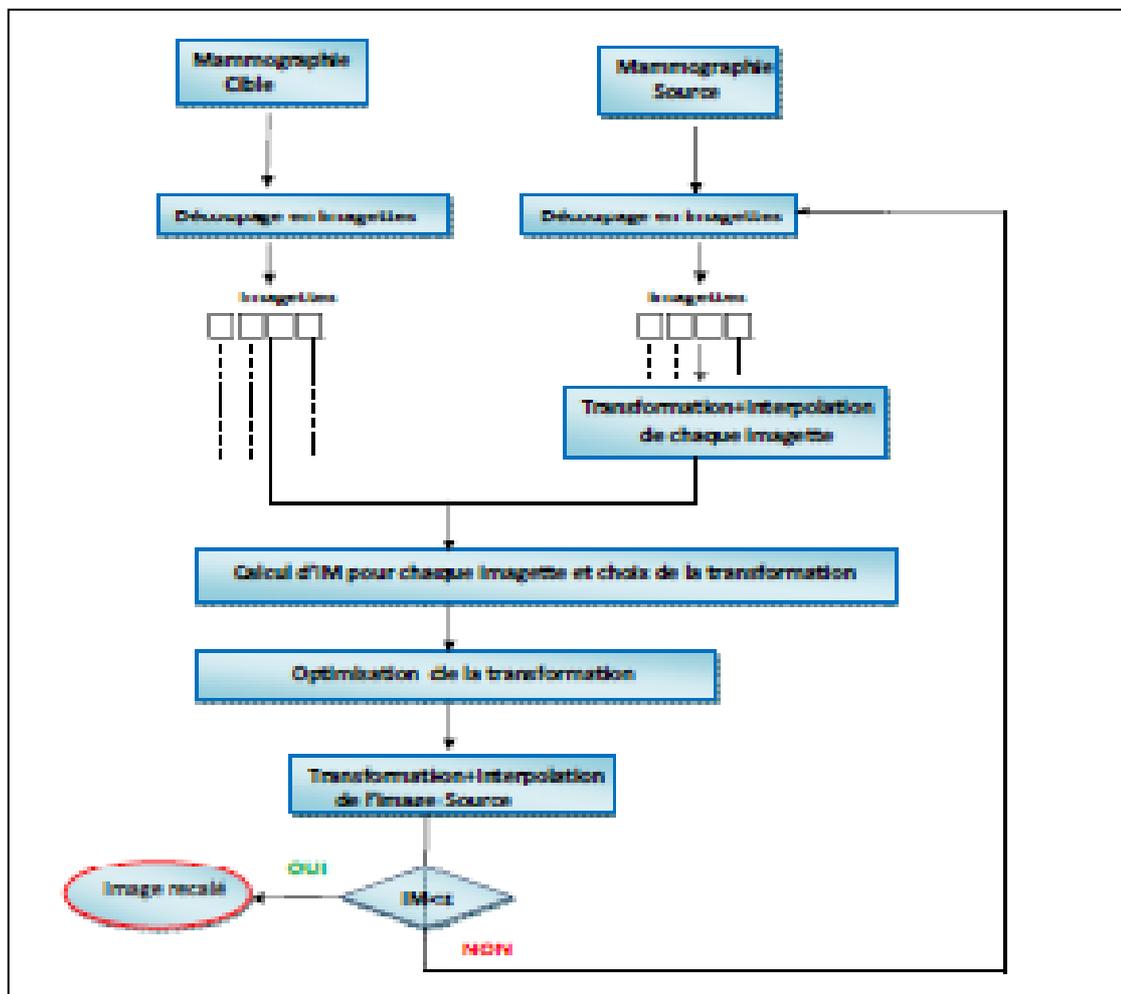


Fig 8.1. Principe Général de SASDUM 5.2

a. Découpage en imagerie:

Cette tâche consiste à découper chaque image (source et cible) en quatre imagerie. Ce qui revient à partitionner les pixels d'une image en quatre parties de même taille. Un résultat similaire à la figure 8.2 est obtenu.

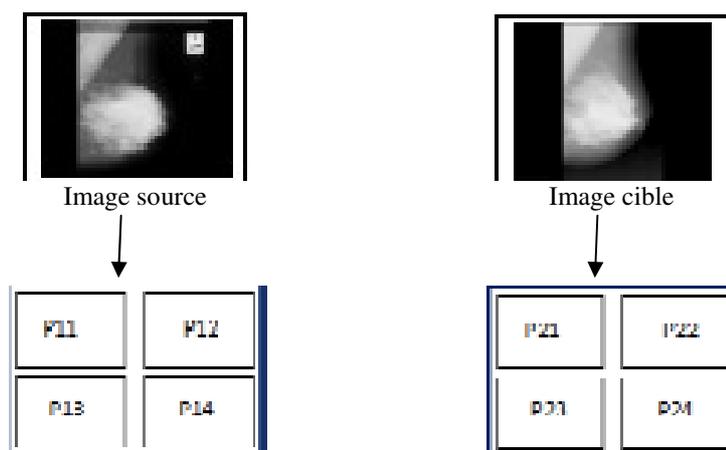


Fig 8.2. Principe de découpage de l'image source et de l'image cible

b. Transformation et interpolation des imagettes :

Chaque transformation rigide est appliquée sur chaque imagette. Par exemple la première imagette de l'image source (P21) subit une fois une translation et une autre fois une rotation. Ce qui revient à récupérer deux transformations de l'imagette P21 notant ces imagettes par P21(T) pour celle qui a subit une translation et P21(R) pour celle qui a subit une rotation. Ce même principe est appliqué sur les autres imagettes de l'image source à savoir P22, P23 et P24. Il faut noter que l'opération d'interpolation est exigée à chaque transformation d'imagette. L'interpolation NEAREST_NEIGHBOR [1] est celle adopté par cette partie de travail.

c. Calcul de l'information mutuelle et choix de la transformation :

L'information mutuelle se calcule directement à partir de l'histogramme conjoint des imagettes de l'image source et de l'image cible. Ainsi on dispose pour chaque paire d'imagettes deux valeurs de l'IM (figure 8.3):

- IM1(T) pour P11, P21(T) et IM1(R) pour P11, P21(R) ;
- IM2(T) pour P12, P22(T) et IM2(R) pour P12, P22(R) ;
- IM3(T) pour P13, P23(T) et IM3(R) pour P13, P23(R) ;
- IM4(T) pour P14, P24(T) et IM4(R) pour P14, P24(R).

IMi(.) (i=1,...,4) calcule la "planitude" de l'histogramme conjoint des imagettes P1i et P2i(.). En effet, plus la forme de l'histogramme conjoint est plane, moins l'information mutuelle est élevée. A l'opposé, moins les valeurs de l'histogramme conjoint seront identiquement réparties, plus la valeur de l'information mutuelle sera élevée. Le calcul de l'information mutuelle IMi(.) est comme suit [2,3]:

Soit $h_c(x, y)$ la valeur de l'histogramme conjoint au point $[x, y]$. On pose :

$$h_{CN}(x, y) = \frac{g(x,y)}{\sum_{a,b} h_c(a,b)} \dots\dots\dots 1$$

$$probaX(x) = \sum_b h_{CN}(x, b) \dots\dots\dots 2$$

$$probaY(y) = \sum_a h_{CN}(a, y) \dots\dots\dots 3$$

On a ainsi :

$$IM = \sum_{a,b} h_{CN}(a, b) \log \frac{h_{CN}(a,b)}{probaX(a)*probaY(b)} \dots\dots\dots 4$$

- $\sum_{a,b} h_c(a, b)$ représente le nombre de points utilisés pour créer l'histogramme conjoint.
- h_{CN} est l'histogramme conjoint normalisé. $h_{CN}(x, y)$ Peut donc se lire comme la probabilité qu'un point pris au hasard dans l'image soit la combinaison du niveau de gris x sur l'image 1 et du niveau de gris y sur l'image 2.
- $probaX(x)$ est aussi une distribution de probabilités. Pour un x donné c'est la probabilité que l'on trouve un point de niveau de gris x sur l'image 1.
- $probaY(y)$ est comme $probaX(x)$. Pour un y donné c'est la probabilité que l'on trouve un point de niveau de gris y sur l'image 2.

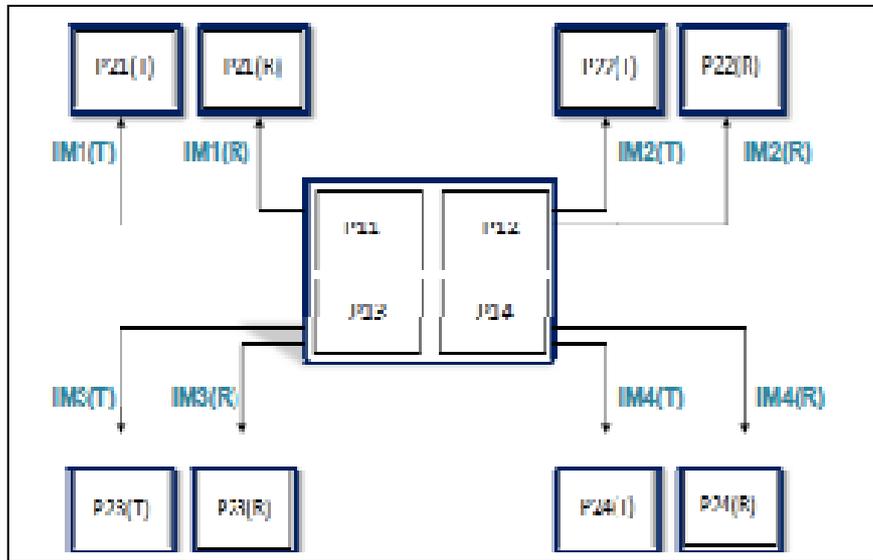


Fig 8.3. Information mutuelle de chaque imagette

Pour chaque paire d'imagettes la transformation qui donne la plus grande valeur de l'information mutuelle est sélectionnée.

Si $IM(P21(T), P11) > IM(P21(R), P11)$ alors choix-Translation++ sinon choix-Rotation++

Si $IM(P2n(T), P1n) > IM(P2n(R), P1n)$ alors choix-Translation++ sinon choix-Rotation++;

(Telle que n le nombre d'imagettes pour chaque image).

On récupère ainsi le nombre de choix de la transformation translation (choix-Translation) et le nombre de choix de la transformation rotation (choix-Rotation). C'est deux valeurs sont comparées et la transformation la plus majoritaire est la transformation choisie. C'est-à-dire : si choix-Translation > choix-Rotation alors choisir la translation sinon choisir la rotation.

La transformation la plus appropriée pour la majorité des imagettes est donc la transformation qui sera appliquée pour l'image source.

d. Optimisation :

Après le choix de la transformation une méthode d'optimisation est appliquée afin de converger vers les paramètres de transformation optimaux. C'est-à-dire les paramètres qui maximisent la valeur de l'information mutuelle [4]. En d'autre terme, si la translation est le résultat de la phase précédente alors on optimise les paramètres de la translation sinon on optimise les paramètres de la rotation. La méthode d'optimisation simplexe a été adoptée par cette partie du travail.

e. Transformation et interpolation de l'image source :

Ayant la transformation à appliquer avec ces paramètres optimaux il nous reste qu'à l'appliquer mais cette fois ci sur toute l'image source. Une itération du principe de recalage vient donc de s'achever. Un critère d'arrêt pour toutes les itérations est déterminé par la valeur de l'information mutuelle.

8.4. Réalisation

La figure ci-dessous est une vue écran de SASDUM 5.2.

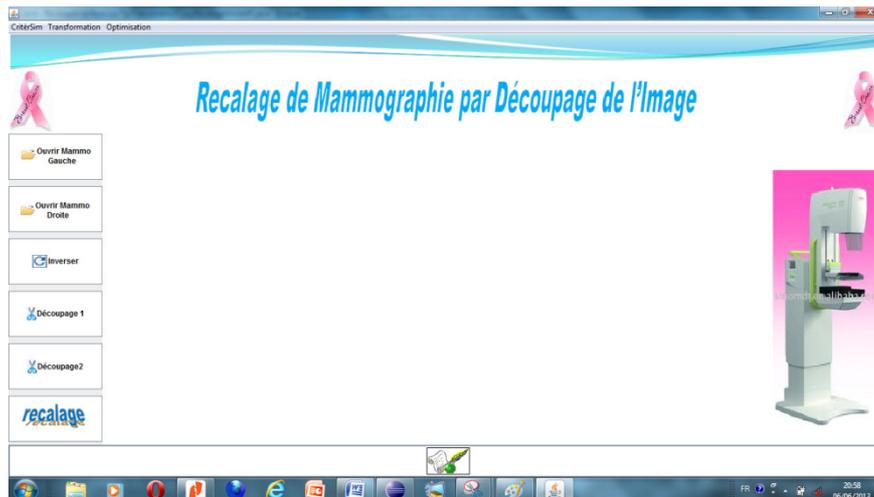
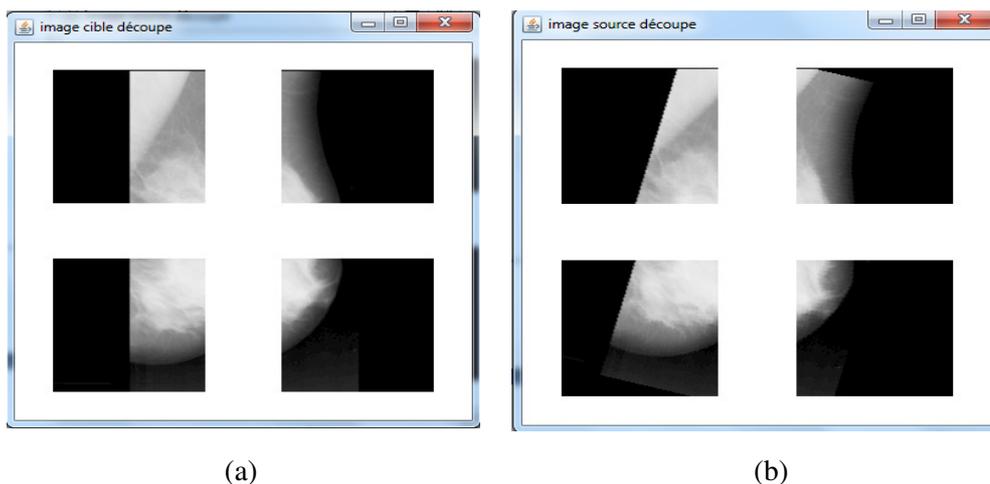


Fig 8.4. Vue de l'interface de SASDUM5.2
(Partie de SASDUM pour le suivi de l'évolution de la tumeur par approche rigide)

Par ailleurs, la figure 8.5 illustre le résultat d'un découpage de deux mammographies : source et cible.



(a) (b)
Fig 8.5. Résultat du découpage de (a) l'image source et de (b) l'image cible

Le processus de recalage consiste à transformer progressivement une image source pour la mettre en ligne avec une image cible. L'approche proposée assure un tel processus comme le montre la figure 8.6.

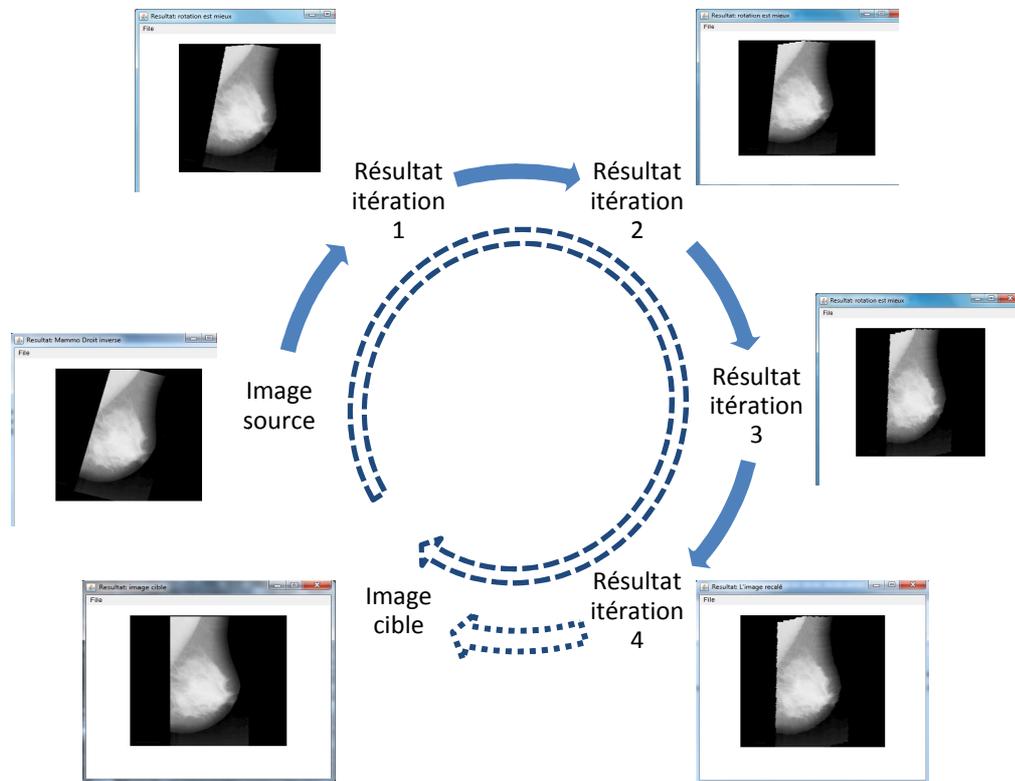


Fig 8.6. Exemple d'un cycle de recalage de SASDUM 5.2

8.5. Conclusion

Cette partie de SASDUM réalisée permet de recalibrer deux mammographies prises à des moments différents. Le processus de recalage passe par un découpage d'image en imagerie. L'information mutuelle nous permet de décider sur la meilleure transformation pour chaque imagerie. La transformation majoritaire sera adoptée pour déformer l'image source. Les résultats obtenus sont encourageants malgré qu'il y'a à dire sur l'opération d'optimisation. En effet cette dernière se piège, dans un grand nombre de cas, dans les minimaux locaux.

8.6. Références bibliographiques

- [1] Goshtasby, Ardeshir.; "2-D and 3-D image registration for medical, remote sensing, and industrial applications"; Wiley-Interscience publication, 2005.
- [2] Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan "Handbook of Biomedical Image Analysis: Volume III: Registration Models"; Edited by Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan, 2005.
- [3] M. RUBEAU, "Approximation de l'Information Mutuelle basée sur le développement d'Edgeworth", PHD thesis, France, 2011.
- [4] Mathieu Rubeaux, « Approximation de l'Information Mutuelle basée sur le développement d'Edgeworth : application au recalage d'images médicales », Thèse de Doctorat, Université De Rennes 1, 2011.

Sommaire du chapitre 9

1. Introduction
 2. Statistiques sur l'évolution du projet
 3. Statistiques sur les dépenses budgétaires
 4. Difficultés et obstacles rencontrés
 5. Retombé scientifique du projet
 6. Impact social du projet
 7. Conclusion
-

9.1. Introduction

Le traitement du cancer du sein peut se limiter à une simple chirurgie associée à une radiothérapie. Il peut toutefois nécessiter l'adjonction d'une chimiothérapie et/ou d'une radiothérapie.

Les malades qui poursuivent un traitement par chimiothérapie doivent joindre, périodiquement, les centres spécialisés pour prendre leurs traitements. Les protocoles de traitement par chimiothérapie sont établis cas par cas, conformément à des schémas de base validés par la communauté médicale internationale. Passer d'un protocole à un autre pour une patiente est envisageable si cette dernière répond négativement à son traitement. L'objectif de SASDUM est justement de répondre à une question clé « le protocole appliqué est-il efficace ou pas ? ».

La problématique de SASDUM a été posée par le partenaire socioéconomique et ce dernier à insister pour le respect de ces besoins. Ce chapitre est une récapitulation de toutes les tâches réalisées depuis le lancement de SASDUM jusqu'à ce jour.

9.2. Statistiques sur l'évolution du projet

SASDUM est un projet ambitieux qui a été décomposé en plusieurs phases (6 phases). Chaque phase a été attribuée à un groupe de travail. Le groupe est composé de doctorant, étudiants en master et de membres du projet. La coordination entre les différents groupes était assurée par le chef de projet qui a suivi en parallèle l'évolution de chaque groupe.

La collecte de données de SASDUM a été effectuée au sein de deux hôpitaux :

1. Hôpital CHU Ibn Rochd à Annaba sous la direction de Dr.Beldjehem Nedjmeddine (Service de radiologie et d'imagerie médicale CHU Annaba), et Dr.Aouras Hayette (Service de Gynécologie –Obstétrique CHU Ibn Rochd Annaba).
2. Hôpital de jour « Pont Blanc » à Annaba sous la direction de Dr.Djeddi et Dr.Amari des médecins spécialistes en oncologie.

SASDUM a vu une évolution pondérée, quelque'une de ces phases ont eu un rythme accéléré alors que d'autres ont avancé en ralenti. Une statistique sur chaque phase est donnée par le tableau suivant :

Tab.9.1 Statistique d'évolution du projet

Phase	Equipe de réalisation	Résultat de la phase	Taux de réalisation	Tâches restantes
1. Base mammaire	Dr Merouani F.H, Dr Layachi S., Ferkous (membre PNR) Benmazou, Hamoud (Doctrantes)	SASDUM.1	50%	<ul style="list-style-type: none"> • Poursuivre la collecte des données • Elaborer une version plus complète en prenant en considération les consignes du partenaire • Intégrer la partie dans le système
2. Détection de la tumeur	Dr Merouani H.F, Dr Melouah A., Ferkous (membre PNR) Ferhani Imen et Tameur Aicha (Master 2)	SASDUM.2	80%	<ul style="list-style-type: none"> • Intégrer la partie dans le système
3. Diagnostic de la tumeur	Dr Merouani Hayet Farida (membre PNR). Refai Ahlem, Djebbar Akila (doctorant) Gasmi Amina (master 2)	SASDUM.3	80%	<ul style="list-style-type: none"> • Intégrer la partie dans le système
4. Indexation	Dr.Merouani F.H. et Dr.Layachi Soumia (membre PNR) Benmazou Sarah (doctorante).	SASDUM.4.1 (Indexation basée sur les automates cellulaire)	80%	<ul style="list-style-type: none"> • Intégrer la partie dans le système
	Dr.H.F.Merouani, Dr.Y.Tlili (Membre PNR) S.Benmazou, M.Hamoud (Doctorat LMD)	SASDUM.4.2 (Indexation basée sur la loi Zipf et Zipf inverse)		<ul style="list-style-type: none"> • Intégrer la partie dans le système
5. Suivi de l'évolution de la tumeur	Dr Merouani F.H, Dr.Layachi Soumia (membre PNR) Salhi (Doctorante)	SASDUM.5.1 (Suivi par approche non rigide)	70%	<ul style="list-style-type: none"> • Intégrer la partie dans le système
	Dr Merouani H.F, Dr Melouah (membre PNR) Bordjiba Abdelah, Arif Hichem, Djedaia Houria et Saoudi Maria (Master 2).	SASDUM.5.2 (Suivi par approche rigide)		<ul style="list-style-type: none"> • Améliorer l'approche en appliquant un découpage selon le contenu de l'image • Tester sur des mammographies de la base • Intégrer la partie dans le système
6. Validation	Les experts : Dr.Beledjhem, Dr.Aoures	Validation des parties de SASDUM	40%	<ul style="list-style-type: none"> • Tester SASDUM • Valider SASDUM par les experts

9.3. Statistiques sur les dépenses budgétaires

Avec le budget octroyer nous avons put acquérir un certain nombre de besoins qui sont ventilés dans le tableau suivant :

Tab.9.2 Ventilation budgétaire

02	Matériels et mobiliers	
1	Matériels et instruments scientifiques et audiovisuels.	713624,7
2	Renouvellement du matériel informatique, achat accessoires, logiciels et consommables informatiques.	313994,868
3	Mobilier de laboratoire.	313994,868
4	Entretiens et réparations.	
	SOUS TOTAL	1341614,436
03	Fournitures	
1	Produits chimiques.	
2	Produits consommables.	
3	Composants électroniques, mécaniques et audio-visuels.	
4	Papeterie et fournitures de bureau.	42817,482
5	Périodiques.	
6	Documentation et ouvrages de recherche.	
7	Fournitures des besoins de laboratoires (animaux, plantes etc.)	
	SOUS TOTAL	42817,482
04	Charges Annexes	
1	Impression et édition.	28544,988
2	Affranchissements postaux.	14272,494
3	Communications téléphoniques, Fax, télex, télégramme, Internet.	
4	Autre frais (impôts et taxes, droits de douane, frais financiers, assurances, frais de stockage, et autre).	
5	Banque de données (acquisitions et abonnements).	
	SOUS TOTAL	42817,482
	TOTAL GENERAL	1427249,4

Ce que nous avons acquis jusqu'à ce jour : Papeterie, matériel informatique comme les ordinateurs, les imprimantes, disques durs, consommable et mobilier.

Nous regrettons malheureusement la lourdeur de la procédure d'achat, qui a fait que nous n'avons pas put bénéficier de micro-portables. Ceci dit, il reste encore des bons qui sont en attente de signature.

9.4. Difficultés et obstacles rencontrés

Tout le long du projet SASDUM plusieurs obstacles ont ralenti son évolution les plus importantes sont énumérés dans ce qui suit :

- les pannes successives et de longue durée des mammogrammes ;
- quelques patientes refusent de suivre leur traitement jusqu'au bout, soit elles se redirigent vers d'autres hôpitaux soit, elles ne donnent plus signe de vie malgré qu'elles ont suivi quelques séances de chimiothérapie ;
- les changements fréquents de la stratégie de dispatching des tâches dans l'hôpital, et les mutations de quelques médecins et administrateurs de leurs postes vers d'autres hôpitaux ;
- les nouveaux arrivants (médecins et radiologues) tardent à coopérer, dû à la charge intense qu'ils trouvent à la prise de leurs postes.
- les malades suivent leur traitement chimiothérapique au niveau de l'hôpital de jour Pont Blanc et, effectuent leurs examens de mammographies à l'hôpital CHU Ibn Rochd, ce transite alourdit

notre tâche qui exige une haute collaboration des malades d'un coté et des deux hôpitaux d'un autre coté.

- L'autre obstacle qu'on a rencontré c'est la lourdeur administrative pour l'acquisition du matériel.

9.5. Retombé scientifique du projet

La section suivante cite les différentes contributions scientifiques de toute l'équipe qui a participé dans le projet SASDUM

a. Publication(s)

1. **Melouah A** "A novel region growing segmentation algorithm for mass extraction in mammograms"; Modeling Approaches and Algorithms, SCI 488, pp. 95–104.;Springer International Publishing Switzerland 2013
2. Djebbar Akila, **Hayet Farida Merouani** "Optimizing retrieval phase in CBR through Pearl and JLO algorithms for medical diagnosis.", in International Journal of Advanced Intelligence Paradigms (IJAIP), Vol. 5, N°3.2013 pp.61-81 **ISSN** 1755-0386, indexed in **INDERSCIENCE**, ProQuest,ULRICH, DOAJ, OpenJ-Gate, Cabell, EI Compendex. <http://www.inderscience.com/ospeers/admin/author/articlelist>.
3. Djebbar Akila, **Hayet Farida Merouani** "Applying BN in CBR Adaptation –Guided Retrieval for Medical Diagnosis", in International Journal of Hybrid Information Technology, (IJHIT), Vol. 5, N°2, April 2012, pp.41-56. **ISSN 1738-9968**, indexed in **EBSCO**, ProQuest,ULRICH, DOAJ, OpenJ-Gate, Cabell, EI Compendex. <http://www.sersc.org/journals/IJHIT>
4. Djebbar Akila & **Hayet Farida Merouani** "Retrieval of cases by using a Bayesian Network applied to the medical diagnosis" The Mediterranean Journal of computers and networks, Vol.8, N°2,pp.69-74 April 2012, **ISSN 1744-2397**, indexed in **INSPEC**, **British Library**. www.medjcn.com
5. Benmazou Sara, **Layachi Soumia & Hayet Farida Merouani** "Segmentation in Images based Cellular Automata Reactive Agent Implemented in Netlogo Platform", in International Journal of Computer Application, Vol.45, N°24, May, 2012, pp.40-43. **ISBN: 978-93-80866-46-3**, **ISSN 09758887**, indexed in **Springer**, **INASP**, **DOAJ**, **LUND**, **BIOMed Central**, Editeur. Foundation of Computer Science <http://www.ijcaonline.org>

b. Conférences internationales

1. **A.Melouah**, A.Salhi, "Mass extraction in mammograms by an adaptive region growing technique" 2nd World Conference on Information Technology, 24-27 November 2011, Antalya – Turkey
2. **Melouh A.**, "Toward an adaptive extraction method of masses contours in digitized mammograms"; the 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA'13), 12-15 May 2013, Algiers, Algeria
3. **Melouah A.** "A novel region growing segmentation algorithm for mass extraction in mammograms", the fourth International Conference on Computer Science and its Applications CIIA'2013) 4-6 May 2013, Saida - Algeria.
4. Refai Ahlem, Djebbar Akila & **H.F.Merouani**, " RB-Maint: Un Modèle Probabiliste pour la Maintenance d'un Système RàPC" 7 éme Colloque sur l'Optimisation et les Systèmes d'Information, COSI'10, Avril 18-20, 2010, Ouargla, Algérie.
5. A.Djebbar **and H.F.Merouani**, "BN-CBR model for Diagnosis of Hepatic pathologies", Biomedical Engineering International Conference, biomeic'12, 10-11October 2012, Tlemcen, Algérie, **ISSN 2253-0886**.
6. Ahlam Refai & **H.F.Merouani**, " Application à l'aide de diagnostic médicale à base ce réseau bayésien dans le cadre de la maintenance" The Third International Conference on Image and Signal processing and their Applications, ISPA'12, 02-04 December 2012, Mostaganem, Algeria.
7. Refai Ahlem & **H.F.Merouani**, " Maintenance des Réseaux Bayésiens pour le Diagnostic des Pathologies des Seins" Second International Conference on Systems and Information Processing, ICSIP'11, May 15-17, 2011, Guelma, Algérie.
8. Benmazou Sarah, **Layachi Soumia & H.F.Merouani**, " Approche Automate Jeu de la Vie-SMA pour la Segmentation des Images IRM Cérébrales " Conférence Maghrébine sur l'Extraction et la Gestion des Connaissances, EGC'11,23-25 Novembre 2011, Tanger, Maroc.

9. Benmazou Sarah, **Layachi Soumia & H.F.Merouani** "Segmentation d'images sur un Couplage Automate-Agent", ICIST'11, 1st International Conférence on Information Systems and Technologies, 24-27 April 2011, Tebessa, Algeria, pp 525-532. **ISBN : 978-9931-9004-0-5**
10. Benmazou sarah, Layachi Soumia & **H.F.Merouani**, " Approche Automate Jeu de la Vie-SMA pour la Segmentation des Images IRM Cérébrales" The Second international conference on complex Systems, CISC'11,06-08 December 2011, Jijel, Algeria.
11. Benmazou Sarah, **Layachi Soumia & H.F.Merouani**, " Automates Cellulaire SMA pour la Segmentation des images " First International Conference on Signal, Image, Vision and their Applications, SIVA'11, at Guelma, November 21-24, 2011, Guelma, Algérie.
12. Hamoud Meriem & **H.F.Merouani**, "Computer Aided to medical diagnosis of Mammogram masses by the Biclustering and the power Laws" Workshop sur la Modelisation Mathématique et Controle, Université Badji Mokhtar Annaba, 20-22 Mai 2013, Annaba, Algeria
13. Hamoud Mériem, **Hayet Farida Merouani** "Detection of a region of interest in the Image Based on Zipf Laws", SITIS'11 at The 7th International Conference on Signal Image Technology & Internet Based Systems, Edited by Kokou Yelongnon, Richard Chbeir and Albert Dipanda, Dijon, France, 28 November-1 December 2011, pp.416-421, **ISBN:978-0-7695-4635-3/11,2011 IEEE DOI10.1109,Indexed in Scopus, IEEE explore and in ACM DL.**
14. Hamoud Meriem & **H.F.Merouani**, " Détection d'une Zone d'intérêt dans les images basée loi de Zipf" Conférence Maghrébine sur l'Extraction et la Gestion des Connaissances, EGC'11,23-25 Novembre 2011, Tanger, Maroc.

c. Conférences nationales

1. Sarah Benmazou, **Layachi Soumia & H.F.Merouani**, " Une approche AC-SMA pour la Segmentation IRM cérébrale" 1^{ère} Journées Doctorales du Laboratoire d'Informatique d'Oran, JDLIO'11, 31 MAI- 01 Juin 2011, Oran, Algérie.
2. Hamoud Meriem & **H.F.Merouani**, " Computer Aided Detection of Mammogram masses by the power Laws" Rencontre des Femmes Scientifiques Mediteraneennes, REFSCIME 2013, Université 209 Aout 1955 Skikda, 15-16 Avril 2013, Skikda, Algeria.
3. Hamoud Meriem & **H.F.Merouani**, Détection d'une Zone d'intérêt dans les images par la Loi de Zipf et la Loi de Zipf Inverse, 1^{ère} Journées Doctorales du Laboratoire d'Informatique d'Oran, JDLIO'11, 31 MAI- 01 Juin 2011, Oran, Algérie. www.univ-oran.dz/rencontres/jdllo/index.html

d. Autres

- Trois articles ont été acceptés dans des journaux de renommés et sont en cours de publication.
- Masters soutenus dans la thématique du projet:
 1. **Belleili Houssemeddine**" Détection des zones avec la fusion de la loi zipf et la loi Zipf inverse »
 2. **Saadi Asma** "Indexation des images par les Lois de ZIPf »
 3. **Khettabi Ryma**: « Evaluation de la segmentation Contour »
 4. **Merouani Naouel**:" Evaluation de la segmentation région »
 5. **Meghout Mohamed Salah** « Détection d'une zone d'intérêt dans les images basées fusion des lois zipf et zipf inverse »
 6. **Benhamadi Mostafa Islam** «Fusion des Lois Zipf et zipf inverse pour la détermination d'une zone d'intérêt dans les images »
 7. **Baaghana Kamel** « Indexation des images basées automates cellulaires »
 8. **Fertikh Meriem** « **Conception d'une** Base de données médicale de la ville de Annaba»

9. **Bendjama Tarek & Soltani Charif** « L'interprétation automatique des anomalies dans l'imagerie médicale » à l'Université de Guelma
10. **Ferhani Imen** ; " Détection des masses par une approche de croissance de région adaptative". 2012
11. **Tameur Aicha** ; " Détection des masses par une approche de croissance de région adaptative". 2012
12. **Bordjiba Abdelah**, « Recalage non rigide d'images IRM basé sur des points de repères » ; 2013.
13. **Arif Hichem**, « Recalage rigide d'images IRM par extraction des points d'intérêts » ; 2013.
14. **Saoudi Maria**, « Recalage de mammographie par transformation multiple » ; 2013.
15. **Djedaia Houria**, « Recalage de mammographies basé sur l'extraction du muscle pectoral » ; 2013.
16. **Atik Amina** ; « Recalage de mammographies basé sur l'information mutuelle » ; 2013.
17. **Boumaza Rahma**, « Recalage de mammographies par découpage de l'image » ; 2013.

9.6. Impact social du projet

Arrivé à la phase d'exploitation SASDUM a un impact important sur trois acteurs

- **Impact sur le malade** : Réduire la souffrance du malade et la durée de son traitement en lui évitant de poursuivre des traitements de chimiothérapie inutiles. Aller plus rapidement vers des traitements plus efficaces donc moins mutilants et avec moins de séquelles physiques, psychiques et sociales pour la malade.
- **Impact financier** : le rapport résultat/coût de la prise en charge d'une malade atteinte d'un cancer du sein reste non seulement décevant avec des risques de rechutes et de récurrence important mais a un coût très élevé, pouvant atteindre plusieurs millions de dinars par malade. Ce fait grève considérablement les budgets réservés à ce type de maladies. Le projet SASDUM permet d'éviter de continuer d'investir dans l'achat d'un traitement de chimiothérapie chère prescrit à une patiente, qui ne lui apportera aucune amélioration.
- **Impact sur le spécialiste** : SASDUM est un besoin formulé par les spécialistes eux-mêmes. Ces derniers sont en quête d'un outil qui les aide à :
 - ✓ Détecter une masse mammaire ;
 - ✓ Suivre l'évolution de la maladie.
 - ✓ Guider le processus de traitement en les informant sur l'efficacité de celui-ci les incitant ainsi à le maintenir ou à le modifier.

9.7. Conclusion

En récapitulant et faisant le point sur ce qui à été fait, nous pouvons avancer qu'une grande partie de SASDUM est réalisée.

Comme nous le voyons dans le tableau résumant les différentes phases de ce projet, et en donnant leurs taux d'avancement, nous pouvons dire que les objectifs tracés au départ sont atteints à 75%. Les autres 15% concernent principalement la base de données pour la ville d'Annaba qui doit impérativement être terminée (et c'est ce que nous sommes entrain de faire) ; l'autre phase restante inachevée est la validation du système dans sa totalité par les experts du domaine.

Dans ce chapitre de récapitulation, nous avons aussi présenté les différents travaux présentés dans des conférences Nationales et Internationales (17 articles). Certaines phases ont pu même être validées dans des publications journaux de renommés établis (5 publications), et d'autres sont en cours de publication (3 vont être publiés en 2014). Nous avons aussi cités les noms des étudiants encadrés en master 2 dans la thématique du projet.

« Plus de 45.000 nouveaux cas de cancer sont enregistrés annuellement en Algérie, indique les dernières statistiques de l'Organisation mondiale de la Santé (WHO). 20 types de cancer sont les plus répandus. Parmi les cinq cancers les plus fréquents chez la gent féminine, à savoir le cancer des ovaires, du col, de l'utérus et du côlon, le cancer du sein représente la première. »

Wafia ADOUANE

(tiré du Journal L'EXPRESSION du 22 Décembre 2012)

Le cancer du sein est devenu le cancer le plus fréquent chez la femme algérienne. Il touche le plus souvent des mères de famille, donc des femmes jeunes. Il est un réel problème de santé publique qui menace la société toute entière du fait qu'il affecte et bouleverse des familles entières en les privant encore trop souvent de leur noyau maternel. Le cancer du sein devient tellement fréquent que l'on a l'impression d'assister à une réelle épidémie. Il est devenu plus qu'urgent de se pencher sur ce problème. En Algérie, l'intérêt du dépistage et du suivi du cancer du sein n'est pas une simple vue de l'esprit. C'est bel et bien une nécessité qui ne s'exprime pas encore suffisamment mais qu'il est impératif de promouvoir et de mettre en application dans les meilleurs délais c'est justement l'objectif du projet SASDUM.

Le suivi de la maladie cancéreuse à était largement exposer dans la littérature mais aucun projet n'a réellement pris en charge complètement cette tâche. D'après nos recherches SASDUM est le seul projet de suivi du cancer du sein avant et après chimiothérapie. Si on arrive à finaliser SASDUM il saura le seul système de suivie réellement pratique avec une perspective de commercialisation.

La première innovation de SASDUM réside dans sa capacité de suivre l'évolution de la masse avant et après chimiothérapie. SASDUM apporte une aide précieuse à l'expert qui en fonction de ce suivi décide si le malade poursuit ou non son protocole de traitement.

La deuxième innovation de SASDUM est la conception d'une base de données propre à la ville d'Annaba, qui malheureusement n'a put être établie dus aux contraintes qui ne relèvent pas de nos prerogatifs.

Achever à 100% SASDUM peut s'installer et fonctionner sur un micro-ordinateur classique. Il ne demandera que des données comme les fichiers DICOM ou JPEG pour permettre le suivi d'une tumeur mammaire. De ce fait, le radiologue n'aura qu'à fournir les données à SASDUM pour pouvoir l'exploiter.

En fait, ce travail est loin d'être achevé ; comme la très bien souligné notre partenaire socio-économique (dans son attestation, annexe 1), le projet SASDUM est d'une telle envergure qu'il est pratiquement impossible de le terminer en deux ans. Tout fois notons, avec pertinence, l'importance, l'impact, le besoin ressenti par les spécialistes qui insistent pour obtenir un prototype exploitable.

Nous avons pris conscience de la difficulté à collecter et traiter des données réelles et brutes afin d'élaborer une base de données.

Aussi, en avançant dans ce travail d'autres perspectives et horizons sont ouverts à nous. Comme le prélèvement d'une biopsie dans la zone cancéreuse exacte. Un guidage automatique dans le prélèvement de la biopsie est plus que nécessaire et c'est aussi le ressenti du partenaire.

Annexe1

Dr. Beledjem NedjEddin
Maître assistant à l'hôpital Ibn Rochd
Annaba
Adresse e-mail : nadjmo@yahoo.fr

Le 17/10/13

A qui de droit

Objet : Attestation

Je soussigné Dr. Beledjem NedjEddin avoir supervisé une grande partie du projet SASDUM.

Cette supervision consiste en :

- ✓ Explications et discussions portant sur le diagnostic du sein et la répercussion de cette maladie en cas de diagnostic tardif.
- ✓ Octroi d'un certain nombre de mammographies.
- ✓ Discussions portant sur les différentes phases de ce projet.

Comme je soussigne avoir, à maintes reprises, rencontré les étudiants concernés par cette thématique en les orientant vers des besoins plus précis.

Les résultats obtenus dans les différentes phases de ce projet sont très satisfaisants ; certaines phases sont en voies d'achèvement.

Malheureusement, la base mammaire prévue pour la wilaya d'Annaba n'a pas pu se faire car les images n'étaient pas suffisantes, dues aux pannes fréquentes du mammogramme. Par ailleurs, il y a eu des difficultés pour l'obtention de mammographies avant et après chimiothérapies.

Le projet est d'une telle envergure qu'il est impossible de le terminer dans le délai prévu.

Je serai à votre disposition pour d'éventuelles informations.

Accepter, Monsieur mes respectueuses salutations.


Dr. BELDIEHEM N.
MD.RADIOLOGUE
Hospitalo-Universitaire
Radiologie Interventionnelle
Imagerie Médicale
ANNABA

Annexe 2

- **Publications et Communications**
 - Voir documents joins

- **Applications**
 - Voir CD