

جامعة أبو بكر بلقايد - تلمسان

Université Abou Bakr Belkaïd de Tlemcen

Faculté de Technologie

Département de Génie Biomédical

Laboratoire de Recherche de Génie Biomédical

MEMOIRE DE PROJET DE FIN D'ETUDES

Pour l'obtention du Diplôme de

MASTER en GENIE BIOMEDICAL

Spécialité : Signaux et Images en Médecine

Présenté par : **BENNIHI Hasnia et CHEKIFI Meryem**

**SEGMENTATION NON SUPERVISEE DES
MASSES MAMMAIRES PAR LE MODELE DE
MARKOV CACHE**

Soutenu le 25 Mai 2016 devant le Jury

Mr.	ABDERRAHIM Mohammed El Amine	<i>MCA</i>	Université de Tlemcen	Président
Mme.	SETTOUTI Nesma	<i>MAA</i>	Université de Tlemcen	Examinatrice
Mr.	CHIKH Mohammed Amine	<i>PR</i>	Université de Tlemcen	Encadreur
Mr.	BECHAR Mohammed El Amine	<i>Doctorant</i>	Université de Tlemcen	Co-encadreur

Année universitaire 2015-2016



Dédicace

A mes très chers parents

Pour leurs sacrifices leur amour leurs prières et leurs soutiens.

A ma chère famille

Petit ou grand proche ou lointain.

A tous mes amis

*Sans qui la vie me semblerait bien fade exceptionnellement
Hasnia et Yasmin je vous souhaite la prospérité et le succès.*

A tous nos enseignants avec notre profonde considération.

Meryem



Dédicace

A mes très chers parents,

Que dieu les accueille dans son vaste Paradis

A ma chère famille,

Frères & Sœurs, Petit & grand, proche & lointaine.

A mes accompagnantes au cours de mes études :

CHERIFI Meriem & BELKACEM Chahira

A tous mes amis,

Sans qui la vie me semblerait bien fade, je vous souhaite la prospérité et le succès.

A tous nos enseignants avec notre profonde considération.

✍️ Hasmia

Remerciements :

Nous remercions Dieu de nous avoir accordé des connaissances de la science et de nous avoir aidés à réaliser ce travail.

Au terme de ce modeste travail nous remercions chaleureusement et respectivement tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste projet de fin d'étude.

Nos vifs remerciements vont tous d'abord à notre encadreur Mr. CHIKH Mohammed Amine.

Nous voulons également exprimer notre profonde gratitude pour notre Co-Encadreur Mr. BECHAR Mohamed Amine et BOUCHIKHI Sara qui nous ont encadrés tout le long de ce projet et qui nous ont apporté beaucoup de connaissances dans ce vaste domaine de « traitements & classification des images médicales ».

Tout notre respect et nos remerciements vont vers les membres du jury qui vont pleinement consacrer leur temps et leur attention afin d'évaluer notre travail, qui espérons le sera à la hauteur de leur attente.

Enfin, nos remerciements les plus sincères sont adressés à tous les professeurs, l'administration et le personnel de l'informatique qui ont contribué à forger nos connaissances et à assister notre formation.

Abstrait :

Le cancer du sein, reste la première cause de mortalité chez les femmes âgée plus de 40 ans à travers le monde. La mammographie est la technique de référence primordiale pour l'exploration du sein. Ce travail s'intéresse à la phase de classification non supervisée des images mammographiques de la base de donnée MIAS avec le model de Markov caché. Le procédé HMM est applicable à la segmentation des masses mammaires. Les résultats de la segmentation dans ce travail ont été comparés avec les résultats de la segmentation par Kmeans. L'évaluation de ces résultats a été validé par le classifieur SVM, ce dernier a permet une reconnaissance des masses malignes et bénignes avec un taux de classification de 83.33%, une sensibilité de 100%, et une spécificité de 66.67%.

Les résultats montrent que ces méthodes peuvent intégrer les SAD dans le domaine de l'oncologie.

Mots-clés: Model de Markov Caché (HMM), base de données MIAS, segmentation des masses mammaires, SAD (system d'aide au diagnostic), radiologie.

Abstract:

Breast cancer remains the leading cause of death of women aged over 40 years worldwide. Mammography is the primary reference technique for the exploration of the breast.

This work focuses on the clustering phase of mammographic images of MIAS database with the hidden Markov model. The HMM method is applicable to the segmentation of breast masses. The results of segmentation in this study were compared with the results of the segmentation by Kmeans. The evaluation of these results was made by the SVM classifier and has given us a classification rate of 83.33%, a sensitivity of 100% and a specificity of 66.67%.

The results show that these methods can integrate support ASD in the field of oncology.

Keywords: Hidden Markov Model (HMM), clustering, MIAS database, segmentation of mammary masses, ASD (aid systems for medical diagnosis), radiology.

الملخص:

يبقى سرطان الثدي السبب الرئيسي للوفاة عند النساء اللاتي تزيد أعمارهن عن 40 عاما في جميع أنحاء العالم. التصوير الاشعاعي هو التقنية الأساسية للكشف عن الثدي. يركز هذا العمل على مرحلة تصنيف الصور بالأشعة لقاعدة البيانات 'MIAS' باستعمال نموذج ماركوف المخفي. طريقة ماركوف تطبق لتقسيم الكتل الثديية. تمت مقارنة نتائج التجزئة في هذه

الدراسة مع نتائج "Kmeans". تم تقييم هذه النتائج وفقا للمصنف "SVM"، وأعطانا معدل تصنيف 83.33 %

حساسية 100 % وخصوصية 66.67 % . أظهرت النتائج أن هذه الطرق يمكن أن تدمج نظام المساعدة في التشخيص الطبي في مجال علم الأورام.

الكلمات المفتاحية: ماركوف المخفية النموذجي (HMM)، قاعدة بيانات MIAS، وتجزئة الكتل الثديية (نظام المساعدة في التشخيص الطبي) الطب الاشعاعي.

Table de matière :

Liste des figures	III
Liste des tableaux.....	IV
Liste des abréviations.....	V
Introduction générale.....	1

Chapitre I Etat de l'art

I. Introduction	2
II. Le cancer du sein	2
II.1 Les masses	4
III. Bases de données des mammogrammes numériques	5
III.1 Mammography Image Analysis Society (MIAS).....	5
IV. Etat de l'art	7
V. Conclusion.....	10

Chapitre II : Théorie des HMM

I. Introduction.....	11
II. Historique.....	11
III. Classification non supervisé	12
IV. Le prétraitement	12
IV.1 Le nettoyage	12
IV.2 Amélioration de contraste	13
IV.3 Suppression du muscle pectoral	13
IV.3.1 Le seuillage	13
IV.3.2 Etiquetage.....	13
V. Parcours de Hilbert-Peano	13
VI. Présentation et applications des modèles de Markov cachés	15
VI.1 Algorithme forward-backward	16
VI.2 Algorithme de Viterbi	18
VII. Apprentissage d'un HMM	19
VII.1 Apprentissage d'un HMM à structure connue	19
VII.2 Apprentissage d'un HMM à structure inconnue	23
VIII. Différents types de HMM	25

IX. Conclusion.....28

Chapitre III: Résultat et discussion

I. Introduction29

II. Rappelle.....29

III. Expérimentation31

IV. Résultat et discussion.....33

V. Conclusion.....36

Conclusion générale.....37

Bibliographie.....38

Liste des figures:

- Figure (I.1): Les types de masses.....4
- Figure (I.2): Différentes type de bords de masses5
- *Figure (I.3). Schéma explicative de la méthode.....10*
- Figure (II.1) — Construction du parcours de Peano pour une image 8*8.....16
- Figure (II.2) – modèle de chaine de Markov ergodique.....27
- Figure (II.3) – modèle de Bakis.....27
- Figure (III.1). Diagramme bloc représentatif de la méthode proposée.....33
- Figure (III.2).segmentation de la masse dans l'image mdb023 par HMM.....36
- Figure (III.3).segmentation de la masse dans l'image mdb028 par HMM.....36
- Figure (III.4).segmentation de la masse dans l'image mdb184 par HMM.....36
- Figure (III.5).segmentation de la masse dans l'image mdb023 par Kmeans....37
- Figure (III.6).segmentation de la masse dans l'image mdb028 par Kmeans....37
- Figure (III.7).segmentation de la masse dans l'image mdb184 par Kmeans....37

Liste des tableaux :

- Tableau (I.1): Nombre de mammogrammes dans la base de données MIAS.....6
- Tableau (I.2): Nombre de mammogrammes contenant des masses dans la base de données MIAS.....6
- Tableau (III.1) Base de donnée des images utilisée dans notre travaille.....32
- Tableau (III.2) Calcule de performance des deux méthodes38
- Tableau (III.3) Quelques caractéristiques géométrique et textural de la masse de l'image mdb023.....38

Liste des abréviations :

- HMM : Hidden Markov Model
- SVM : Support Vector Machine
- Kmeans : K_moyen
- MIAS : Mammography Image Analysis Society
- SFC : Space Filling Curves
- MMC : Modèle e Markov Caché
- CMC : Chaîne de Markov Caché

Introduction:

Le cancer du sein est la maladie le plus fréquent dans le monde ; il constitue la première cause de mortalité chez les femmes. L'Agence Internationale de l'Organisation Mondiale de la Santé pour la recherche sur le cancer a déclaré que chaque année, plus de 150.000 femmes meurent du cancer du sein et un million de nouveaux cas sont détectés à l'échelle mondiale.

Certain facteurs contribuent à l'augmentation de l'incidence de cancer de sein. Ce dernier est un ensemble de cellules malignes, qui se multiplient sans contrôle jusqu'à créer une tumeur, s'attaquent aux tissus sains avoisinants. Il peut exister sous forme de masses ou de dépôts de sel de calcium appelées micro ou macro-calcifications.

Comme toute forme de cancer, le dépistage et le diagnostic précoce du cancer du sein est l'un des facteurs les plus importants contribuant à la guérison de la maladie. En effet, la mammographie est l'outil le plus utilisé, car elle permet de produire des images de haute résolution du sein.

Malgré les évolutions des nouvelles technologies pour améliorer la qualité des images mammographiques numériques, l'interprétation et l'analyse de ces images restent une tâche difficile même pour un radiologue. Ceci est dû à la grande variabilité dans la forme des masses mammaires malignes et bénignes. Depuis quelques années, de nombreuses équipes de recherche tentent de développer un système d'analyse et d'interprétation automatique des images mammographiques numériques. Ces travaux de recherche s'orientent en deux grands axes: (1) la détection automatique qui vise à identifier les anomalies et à classer les régions d'un mammogramme en régions d'intérêts (suspectes) et (2) le diagnostic assisté par ordinateur qui vise à caractériser et à classifier l'anomalie détectée en bénigne ou maligne.

Le but de ce projet est de développer de nouvelle méthode de segmentation basé sur l'apprentissage machine (classification), ne nécessitant aucune intervention humaine et étant robuste à la faible qualité d'image médicale.

Ce projet décrit ces différentes étapes et illustre les résultats obtenus sur les mammogrammes de 161 femmes provenant de la base de données MIAS contenant 322 images. Ce mémoire est structuré en trois chapitres comme suivant :

Introduction générale :

Nous commençons par une présentation détaillée de la base de données MIAS, ainsi un résumé des différentes méthodes proposées dans la littérature pour l'extraction des masses appliqué sur cette base sous forme d'état de l'art et nous montrons le prétraitement utilisé pour améliorer la qualité des images avant d'appliquer notre méthode.

Le deuxième chapitre est consacré pour présenter la théorie du modèle de Markov caché (Hidden Markov Model HMM) qui est notre méthode de segmentation proposée.

Enfin, dans le troisième chapitre Nous présentons les techniques d'extraction des attributs de texture pour alimenter le classifieur retenu basé sur les SVM et nous illustrons les résultats obtenus par l'approche HMM sur les masses.

I. Introduction

Le cancer du sein, continue à être la première cause de mortalité chez les femmes âgée plus de 40 ans à travers le monde. Les études confirment que la détection précoce des cancers améliore le pronostic. La mammographie reste la technique de référence incontournable pour l'exploration du sein. Cependant tous les radiologues reconnaissent la difficulté de l'interprétation des mammographies.

C'est pour cela nous allons présenter dans ce chapitre quelques travaux qui ont facilité la tâche de la segmentation des masses mammaires en utilisant la base de donnée MIAS. En plus, nous allons proposer l'utilisation d'un prétraitement d'image qui permet une bonne amélioration de contraste d'image.

Avant tout il faut savoir qu'est-ce qu'un cancer de sein est qu'elles sont les types des masses qu'on va les détecter ?

II. Le cancer du sein

Le mot cancer du sein est un terme générique pour désigner la gamme entière des maladies malignes du sein. Les cellules malignes se multiplient sans contrôle jusqu'à créer une tumeur qui s'attaque aux tissus sains avoisinants. Cette tumeur peut propager des cellules cancéreuses dans tout l'organisme : on dit alors qu'elle devient métastatique [1].

Les facteurs tels que le vieillissement de la population, les nouveaux modes de vie, l'obésité due à l'alimentation riche en graisse, la consommation régulière d'alcool (plus courante chez les femmes depuis quelques décennies) ainsi que le traitement hormonal substitutif, contribuent à l'augmentation de l'incidence de ce cancer.

✓ La maladie du cancer du sein a six signes différents :

- ✚ Calcification
- ✚ Masse Spéculaire
- ✚ Masse circonscrite
- ✚ Masse bien définie
- ✚ Distorsion architecturel
- ✚ Asymétrie

II.1 Les masses :

Ce sont des lésions retrouvées en trois dimensions. Etant plus denses que le tissu mammaire adjacent, elles apparaissent blanches sur les mammogrammes. La taille ne prédit pas le caractère malin, sauf si les clichés successifs montrent une augmentation régulière de la taille. On distingue traditionnellement cinq formes différentes de masses : ronde, ovale, lobulaire, irrégulière ou une distorsion architecturale. Cette dernière forme inclut des rayons convergeant vers un point avec une rétraction focale. La figure (I.1) présente un ensemble de lésions.

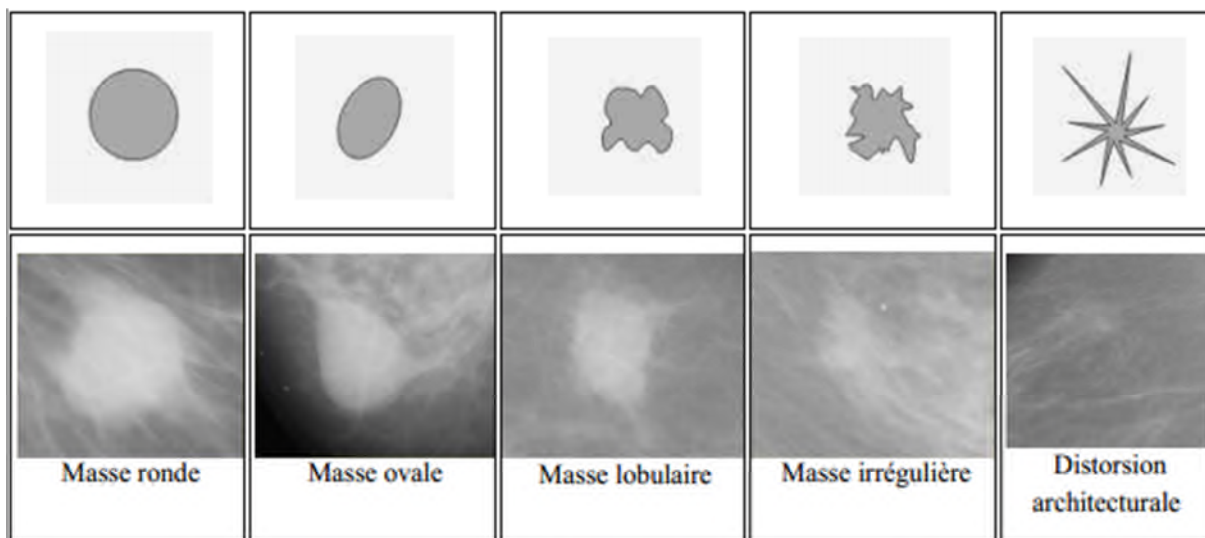


Figure (I.1): Les types de masses.

Les bords de la masse sont très importants à étudier lors de l'analyse sémiologique qui nécessite parfois l'utilisation de la loupe. On décrit cinq catégories de bords de type: circonscrits, effacés, micro-lobulés, mal définis et avec des spicules (voir figure I.2). Les bords effacés correspondent à une superposition de tissus adjacents. Les bords mal définis ou les spicules correspondent à une invasion dans le tissu sain.

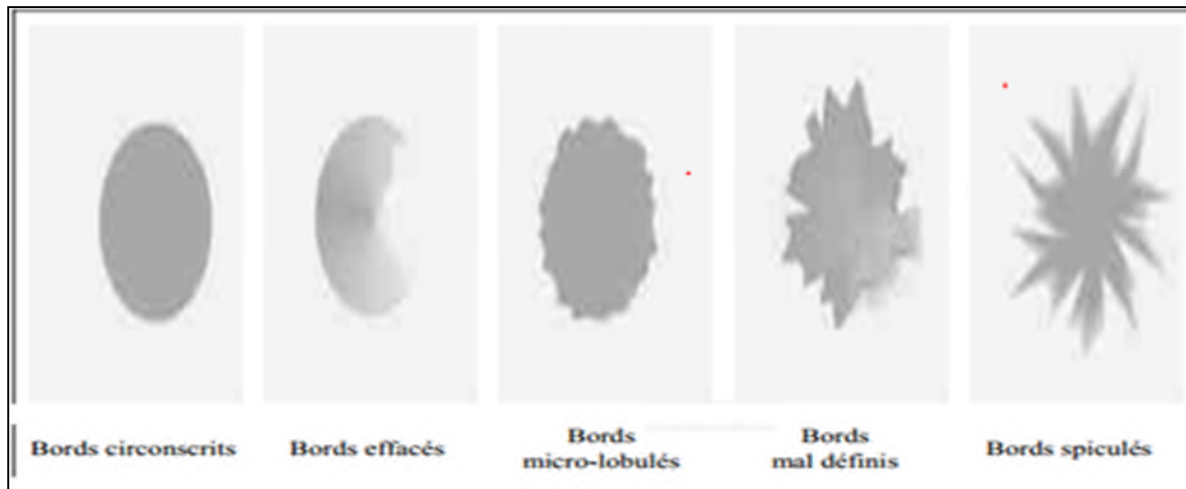


Figure (I.2) : Les différents types de bords de masses

III. Bases de données des mammogrammes numériques

Pour évaluer les performances d'une nouvelle approche d'analyse ou pour comparer de nouvelles méthodes [2], il est recommandé de les valider sur des bases de données de tests standards. En effet les chercheurs ont besoin de comparer leurs résultats sur les mêmes images. Malheureusement la plupart des bases de données d'images médicales réelles ne sont pas toujours accessibles au public. Actuellement parmi les bases de mammogrammes numérique les plus utilisée par les chercheurs est MIAS.

III.1 Mammography Image Analysis Society (MIAS)

La base de données MIAS [3] est le fruit de travail d'une organisation de groupes de recherche britanniques qui s'intéressent à la mammographie. Les images mammographiques sont obtenues après numérisation de films X-ray soigneusement sélectionnés à partir du programme national de dépistage fait à l'échelle du Royaume-Uni. La numérisation est faite avec un scanner Joyce-Lobel micro densitomètre qui est un dispositif linéaire dans la gamme de densité optique est de 0 à 3,2 et qui a une résolution de 50x50 micromètre. Chaque pixel est codé sur 8 bits. Toutes les images sont découpées de telle sorte qu'elles sont toutes de taille 1024x1024. Tous les noms des mammogrammes ont la forme suivante : mdbXXXBS, où:

- XXX : représentent le nombre de l'image, allant de 001 à 322 ;
- B est le côté du sein, il prend la valeur «L» pour le sein gauche ou «R» pour le sein droit (Left or Right);

- S est la taille de l'image qui peut être «s» pour la petite image (1600x4320 pixels), «m» pour des images de taille moyenne (2048x4320 pixels), «l» pour les grandes images (2600x4320 pixels) et «x» pour les très grandes images(4000x5200 pixels).

Chaque image est stockée au format brut: chaque numéro stocké dans le fichier correspond à la valeur du niveau de gris (de 0 à 255) du pixel correspondant dans l'image.

La base de données contient deux prises de vue du sein gauche et droit de 161 patients ; elle est constituée de 322 images classées en trois types :

- *Type normal* : il y a 208 cas;
- *Type bénin* : il y a 63 cas ;
- *Type malin* : il y a 51 cas

Pour chaque image, les radiologues ont décrit le type de la pathologie, sa localisation en donnant les coordonnées du centre de la pathologie, le diamètre d'une surface entourant la pathologie, sa taille, l'échelle, et d'autres informations utiles. Les tableaux (I.1) et (I.2) présentent les types de mammogrammes présents dans cette base de données. Sur ce tableau on peut trouver le nombre de mammogrammes contenant des calcifications, des masses, d'autres types d'anomalies, ainsi que le nombre de mammogrammes normaux, tous répartis selon le type du tissu mammaire (gras, glandulaire, ou dense).

<i>PATHOLOGIES</i>	<i>GRAS</i>	<i>GLANDULAIRE</i>	<i>DENSE</i>	<i>TOTAL</i>
<i>CALCIFICATION</i>	6	9	10	25
<i>MASSES</i>	17	15	10	42
<i>AUTRE</i>	17	15	16	48
<i>NORMAL</i>	66	65	76	207
<i>TOTAL</i>	106	104	112	322

Tableau (I.1): Nombre de mammogrammes dans la base de données MIAS

<i>MASSES</i>	<i>GRAS</i>	<i>GLANDULAIRE</i>	<i>DENSE</i>	<i>TOTAL</i>
<i>CIRCUMSCRIBED</i>	12	8	3	23
<i>SPICULATED</i>	5	7	7	19
<i>TOTAL</i>	17	15	10	42

Tableau (I.2): Nombre de mammogrammes contenant des masses dans la base de données

MIAS

Un fichier joint à la base de données décrit en détails les images. Il est organisé en sept colonnes de la manière suivante :

1^{ère} colonne : le numéro de référence de l'image dans la base de données MIAS

2^{ème} colonne: caractérise le type des tissus du sein. Il y a trois types :

- F : Graisseux (Fatty)
- G : Glandes Graisseux (Fatty-glandular)
- D : Dense (Dense-glandular)

3^{ème} colonne: Type de pathologie

- CALC: Calcification (Calcification)
- CIRC: Well-defined/circumscribed masses
- SPIC: Speculated masses
- MISC: Other, ill-defined masses
- ARCH: Architectural distortion
- ASYM: Asymmetry
- NORM: Normal

4^{ème} colonne: Sévérité de pathologie

- B- Benin
- M- Malin

5^{ème} et 6^{ème} colonne: les coordonnées (x, y) du centre de la pathologie

7^{ème} colonne: Rayon approximatif (en Pixel) d'un cercle enfermant l'anomalie.

IV. Etat de l'art

Comme toutes recherches, nous citons quelques travaux de la segmentation des masses mammaires par les différentes méthodologies de traitement d'images appliqué sur les images mammographiques.

Le travail de Djaroudib et al. [4] propose comme technique de prétraitement d'image d'appliquer le filtre de diffusion anisotropie SRAD (Speckle Reducing Anisotropic Diffusion)

et Contrast-limited Adaptive Histogram Equalization (CLAHE) pour l'amélioration de contraste de l'image. Gray Level Co-occurrence Matrix (GLCM) comme caractéristiques de texture avec une approche region-based sont utilisées. Ces caractéristiques viennent dans les précédentes phases pour l'étape de segmentation ou sont utilisées comme entrées pour l'étape de classification.

Pour évaluer la détection de bord, ils ont sélectionné les contours de masse identifiée, selon l'image d'expert et ils ont réalisé une évaluation quantitative.

Premièrement le Coefficient Dice :

- Pour le tissu dense : 93.39%
- Pour le tissu adipeux : 97.74%
- Pour le tissu Glandulaire : 96.67%

Deuxièmement, on calculant la Précision, sensibilité et F-mesure afin de calculer l'aire sous la courbe (area under the curve Az), $Az=0.81$.

Les avantages de la méthode de Shenghua Gu et al. [5] sont qu'ils peuvent être expliquées en trois aspects. Tout d'abord, en utilisant la méthode patch fusion comme la première étape pour fournir Région d'intérêt (ROI), peut améliorer l'efficacité. Deuxièmement, le processus de classification GHFCM dans ROI pour détecter la masse du sein est beaucoup plus rapide que dans l'ensemble de l'image. Enfin, en combinant la FCM générale (GFCM), qui est robuste au bruit avec les contraintes spatiales et hiérarchique FCM (HFCM) qui est robuste aux valeurs aberrantes avec une générale et flexible fonction distance pour introduire une FCM hiérarchique généralisée (GHFCM), les résultats de segmentation de masse du sein peuvent être obtenus de façon plus précise dans le ROI généré dans la première étape.

Ces résultats de segmentation indiquent que cette méthode qui intègre l'espace; la contrainte et la stratégie de classification hiérarchique est robuste au bruit et les valeurs aberrantes et est capable d'obtenir un contour très lisse de la masse du sein. Pour l'analyse quantitative, ils ont calculé le dice score de la méthode sur huit mammographies initiales, il est compris entre 0.7 et 0.9.

Somayeh Hosaini et al. [6] ont utilisé comme prétraitement le lissage le filtre médian. Pour bordure, l'opérateur Sobel a été utilisé pour calculer le gradient de pixels dans un

voisinage. Pour l'enlèvement des marges et la détection précise des principales zones, les bords ont été trouvés d'abord, puis, le bord de l'image a été profilé avec valeurs 0 et 1 à la fois verticalement et horizontalement.

Il est très important de disposer des modèles de différents tissus dans la segmentation des zones suspectes parce que l'accès à une répartition d'articulation parfaite des pixels est très difficile. Le HMM a été utilisé pour se rapprocher de la statistique marginale et conjointes des composants d'ondelettes. À travers le HMM, une variable d'état caché est considérée pour chacun des composants des ondelettes de contrôler leur grandeur ou petitesse.

Comme les résultats ont montré, la méthode proposée a été en mesure de détecter tous les cas de tumeurs dans des cas simples de la mammographie et en ce qui concerne les cas plus complexes, cette méthode pourrait détecter 92 % des cas. Le taux global de détection est 96%. Les résultats montrent également que la méthode proposée récemment fournit une détection plus précise des tumeurs du cancer du sein par rapport au procédé dans lequel la transformée en ondelettes est utilisée seule.

Dans ce travail, Anuj Kumar Singh et al. [7] introduisent une approche simple et efficace pour détecter la région du cancer dans l'image mammographique. Ainsi l'approche segmente la région du cancer dans l'image mammographique d'entrée.

Cette méthode consiste en deux étapes principales (1) la détection et (2) la segmentation. Dans la phase de détection, une opération de filtrage moyenne de taille de $s \times s$ et de seuillage est appliquée à l'image d'entrée originale qui délivre en sortie la zone de la région maligne.

Pour trouver les tissus malins, ils ont créé une fenêtre rectangulaire autour de la zone de la région en sortie, et appliqué la technique Max –moyen et les moins Variance. Dans la phase de segmentation, un patch de la tumeur se trouve en utilisant l'opération de fermeture morphologique et la technique de gradient d'image pour trouver les bords de la région. Ils ont mis en évidence le bord de la région résultante et les tissus malins détectés sur l'image originale d'entrée. Ils ont présenté les résultats de deux images mammographiques dont un temps approximatif de 4.2 secondes est pris pour exécuter l'expérience sur chaque image. Qui montre que la méthode est capable de donner des résultats à temps réel.

L'approche d'interprétation d'H.S. Rahli et al. [8] est basée sur trois outils qui sont : la matrice de cooccurrence qui est appliquée sur chaque région afin de calculer ses

caractéristiques : les caractéristiques d'intensité, les caractéristiques géométriques et les caractéristiques de texture, et une hybridation de la méthode de K-means et la méthode d'optimisation Search Harmony afin de détecter les tumeurs. Cette approche se base sur quatre étapes essentielles :

- Initialisation de la mémoire d'harmonie de manière aléatoire.
- Calcul de la probabilité d'appartenance des régions à chaque classe.
- Calcul de la fonction de fitness qui est la distance minimale des centroides des classes.
- Improvisation d'une nouvelle harmonie et la mise à jour de la HM (harmony memory)

Cette méthode a été appliquée sur des images mammographiques de taille 256*256 prétraitées.

Selon S.Julian Savari Antony et al. [10] Il y a deux étapes pour classer l'image. Tout d'abord, la qualité de l'image est augmentée en utilisant la méthode d'égalisation d'histogramme. La deuxième étape, les caractéristiques d'intensité sont calculées à partir de l'image de caractéristiques de forme. Les caractéristiques de la région tels que « Solidité », «Excentricité», « Zone Convexe », « Orientation », « Périmètre », «La longueur d'axes majeur », « Longueur d'axes mineur » qui sont extraites pour calculer les valeurs volumétriques. Par conséquent, la qualité de la classification de l'image est un objectif final de ce travail. Le résultat montre que la méthode proposée a augmenté la précision de classification à 99%. On utilise base de données des images MIAS.

V. Conclusion :

Dans ce chapitre, nous avons présenté des notions sur le cancer de sein, la base de donnée MIAS qui contient 322 images mammographiques, aussi un état de l'art des travaux réalisés sur la base de donnée MIAS ainsi le prétraitement appliqué sur l'image pour améliorer la qualité d'image.

Le chapitre suivant, on va entamer la théorie des chaînes de Markov caché qui est la méthode utilisée dans notre projet de fin d'étude.

I. Introduction

La classification des images est l'une des problèmes fondamentaux de la vision par ordinateur et la recherche dans le domaine du multimédia. Elle étant basée sur la notion d'apprentissage, soit dite non supervisée dans le cas où elle est basée sur un apprentissage non supervisé, soit dite supervisée si elle est basée sur un apprentissage supervisé.

L'objectif de ce chapitre est de présenter de façon détaillée le principe de la segmentation non supervisée d'images basée sur le modèle des chaînes de Markov cachées.

II. Historique

Les modèles de Markov cachés (Hidden Markov Model HMM) ont été introduits par Baum et ses collaborateurs dans les années 1960-70s [14]. Ce modèle est fortement apparenté aux automates probabilistes (PAs) [15]. Les HMM sont utilisées pour modéliser des séquences d'observations. Ces observations peuvent être de nature discrète ou continue. Le premier domaine auquel les HMM ont été appliqués est le traitement de la parole au début des années 1970 [16], [17]. Dans ce domaines; les HMM se sont rapidement imposés comme le modèle de référence et une majeure partie des techniques d'utilisation et d'implémentation des HMM ont été développés dans le cadre de ces applications. Ces techniques furent ensuite appliquées et adaptées avec succès au problème de la reconnaissance de textes manuscrites [18], [19] et d'analyse de séquences biologiques [20], [21]. Les HMM ont également été appliqués à d'autres domaines comme la reconnaissance d'image [22], la modélisation d'un signal musical [23] ou la génération de séquences de test pour circuits microélectronique [24].

Dans les années 90, une solution proposée par Benmiloud [25], a consisté à modéliser le processus caché X par une chaîne de Markov (processus 1D). Cette alternative nécessite alors l'utilisation d'un parcours (par exemple le parcours de Hilbert-Peano) pour transformer l'image 2D en un signal 1D. Dans ce contexte, l'estimation des paramètres du modèle avec différents algorithmes (EM, SEM, ICE) permet d'exprimer la loi de X , ainsi que les solutions bayésiennes, de façon explicite, sans mise en œuvre de méthodes d'approximation. En ce qui concerne la segmentation d'images, les CMC ont aussi été utilisées avec succès dans divers contextes [26], comme en segmentation d'images de télédétection [25, 27, 28, 29, 30], d'images médicales [31], d'images sonar, ou même d'images vidéo [25]. Dans le contexte de la segmentation d'images, les modélisations basées sur les chaînes de Markov constituent une alternative intéressante au modèle des champs de Markov cachés [32, 30], même si ces

derniers offrent une modélisation précise et plus intuitive des relations spatiales dans l'image. Le gain en temps de calcul est généralement très significatif, principalement en raison de la complexité de l'étape d'estimation de paramètres des HMRF, ce qui confère aux CMC un réel intérêt pour les problèmes de segmentation et d'analyse d'image.

III. Classification non supervisée

La classification non supervisée (aussi appelée clustering) est considérée comme une classification automatique qui consiste à utiliser des méthodes basées sur la notion d'apprentissage non supervisé. De ce fait, elle regroupe les méthodes ayant pour objectif de déterminer une typologie (ou partitionnement) décrivant un ensemble n d'individus ou d'objets à partir de p caractéristiques calculées pour chacun des n individus appartenant à K classes recherchées.

Contrairement à la classification supervisée dans laquelle l'appartenance des n individus est connue a priori, dans la classification non supervisée l'appartenance des n individus n'est pas connue. C'est justement cette appartenance qu'il s'agit de retrouver à partir des p caractéristiques [33].

IV. Le prétraitement

L'image obtenue de l'acquisition sera bruitée à cause de variation de luminosité, l'intensité et le contraste et qu'elle ne peut pas être traitée directement.

Les techniques de prétraitement d'image sont nécessaires afin de trouver l'orientation de la mammographie, filtrer le bruit et améliorer la qualité d'image. Pour cela on a utilisé les démarches suivantes :

1. Le nettoyage :

On a utilisé le filtre médian [11] qui est un filtre numérique non linéaire, souvent utilisé pour la réduction de bruit. Cela est une étape de prétraitement classique visant à améliorer les résultats de traitements futurs. La technique de filtre médian est largement utilisée en traitement d'images numériques car il permet sous certaines conditions de réduire le bruit tout en conservant les contours de l'image, et même il nous a donné des résultats qui nous ont aidés dans notre cas.

2. Amélioration de contraste :

L'égalisation d'histogramme : De façon générale, une image monochrome possède un histogramme de niveaux de gris dont les composantes non nulles sont groupées dans la partie basse, médiane ou haute de cet histogramme. Une transformation permettant un étalement de ces niveaux selon des proportions uniformes le long de la gamme est appelée égalisation. C'est une opération qui est appliquée dans deux buts :

- Etendre le contraste.
- Ramener le document dans une situation où il est aisé de la comparer à un autre.

Le principe est le suivant : on transforme l'histogramme de l'image en un histogramme plat par une table de transcodage que l'on veut strictement monotone afin de préserver l'ordre des contrastes dans l'image.

3. Suppression du muscle pectoral :

3.1 Le seuillage :

Le problème majeur du seuillage d'images est que l'on considère seulement les intensités des pixels et non les relations de voisinage entre eux. Pour un seuil choisi S , le compromis régions/pixels doit toujours être respecté. En d'autres termes, il faut trouver un seuil correct, souvent choisi par tâtonnement, minimisant les faux contours (pas trop bas) et conservant le maximum de régions complètes (seuil pas trop haut). Cependant, ce compromis reste un but difficile à réaliser [12].

3.2 Etiquetage (labellisation):

Une image d'étiquettes résulte d'une procédure d'étiquetage. Elle se fait par le choix d'une connexité qui permet de définir les propriétés de topologie des divers objets (4- ou 8-connexité sur les trames carrées, 6-connexité en trame hexagonale). L'étiquetage vient attribuer une étiquette semblable à tous les points connexes au sens de la segmentation [12].

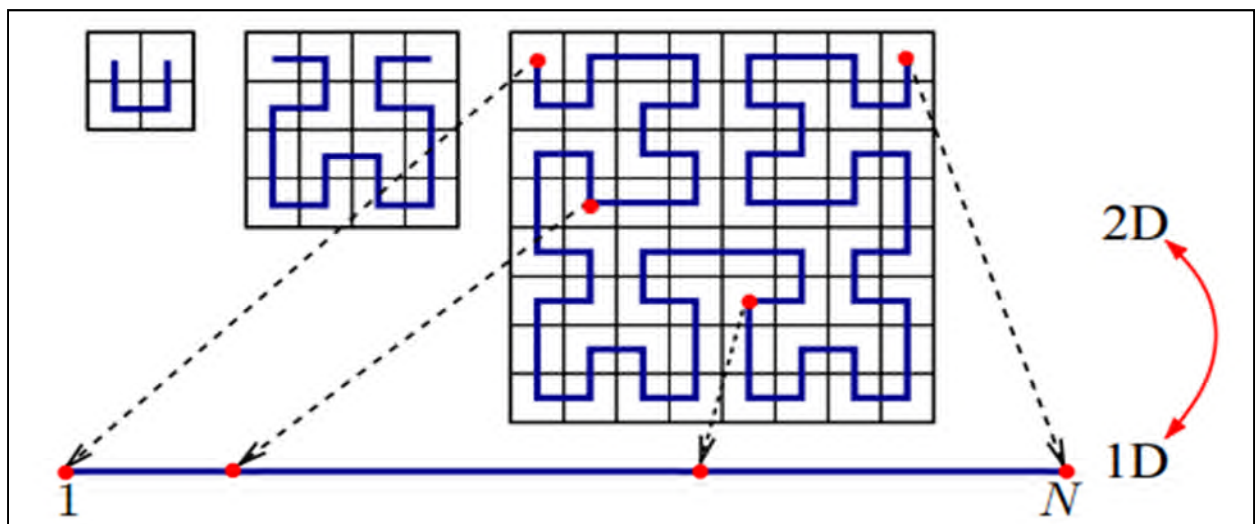
V. Parcours de Hilbert-Peano

Le parcours de Hilbert-Peano appartient à la famille des « Space Filling Curves » (SFCs), qui incluent les parcours en zigzag, le balayage en Z et les parcours standards (ligne par ligne, colonne par colonne). Les SFCs ont été employées dans une grande variété d'applications de traitement d'image telles que la compression, le tramage demi-tons (« half-

toning »), la reconnaissance de formes et l'analyse de texture (cf. [34, 35] et les références qu'elles contiennent). Tous ces parcours peuvent être employés pour convertir un ensemble 2D de pixels en séquence 1D.

Néanmoins, le parcours de Hilbert-Peano fournit une bonne exploitation du voisinage 2D, essentiellement grâce à la contrainte qu'il respecte : « tout pixel de la chaîne possède comme voisins, deux pixels voisins sur l'image ». Cette propriété, ainsi que l'aspect pseudo-aléatoire des changements de directions, implique que le parcours de Hilbert-Peano fonctionne bien (du point de vue statistique) pour de nombreuses familles d'images. La Fig 1.1 présente la construction du parcours de Hilbert-Peano pour une image $8 * 8$.

La première étape de tous les algorithmes de segmentation d'images basés sur le modèle des CMC consiste donc à transformer l'image bruitée en un vecteur. Tous les traitements s'effectuent sur ce vecteur, puis, une fois la segmentation de la chaîne achevée, l'image thématique est finalement reconstruite à l'aide du parcours de Hilbert-Peano inverse. Dans la suite, nous utilisons donc une structure 1D pour représenter une image. Quand nous parlerons d'image, il faudra donc comprendre la séquence issue de la transformation de l'image par le parcours de Hilbert-Peano. [L'application du parcours de Hilbert-Peano n'est possible que dans le cas d'images de taille $2^l * 2^q$ Nous utiliserons donc une version généralisée de ce parcours [36], dont la mise en œuvre est rendue possible sur des images dont les nombres de lignes et de colonnes sont paires].



Figure(II.1) — Construction du parcours de Peano pour une image $8*8$ (initialisation, étape intermédiaire et parcours final).

VI. Présentation et applications des modèles de Markov cachés :

Nous nous limitons ici à la modélisation de séquences de symboles appartenant à un alphabet fini.

Définition 1.1. Un HMM est défini par un quintuplé (S, Σ, T, G, π) , ou :

- S est un ensemble de N états,
- Σ est un alphabet de M symboles,
- $T = S \times S \rightarrow [0, 1]$ est la matrice de transition, indiquant les probabilités de transition d'un état à l'autre ; on note $P(s \rightarrow s')$ la probabilité de transition de l'état s vers l'état s' ,
- $G = S \times \Sigma \rightarrow [0, 1]$ est la matrice de génération, indiquant les probabilités de génération associées aux états ; on note $P(o | s)$ la probabilité de générer le symbole $o \in \Sigma$ à partir de l'état $s \in S$.
- $\pi : S \rightarrow [0, 1]$ est un vecteur de probabilités initiales de visite.

La procédure de génération d'une séquence $o_1 \dots o_T$ de symboles à l'aide d'un HMM consiste à partir d'un état s en suivant la distribution π , de se déplacer d'état en état suivant les probabilités de transition, et générer un symbole sur chaque état rencontré en utilisant la distribution de probabilité de génération associée à l'état. Lorsqu'un symbole a été généré, on choisit une transition sortante suivant la distribution de probabilité de transition associée à l'état courant, et la procédure est réitérée jusqu'à la $T^{\text{ème}}$ génération de symbole. Les HMMs définissent donc un processus stochastique non déterministe, ce qui explique le nom donné à ce modèle. On peut classer les principales applications des HMMs en deux catégories.

La première traite des problèmes de reconnaissance ou de classification, la seconde a trait aux problèmes de segmentation de séquences, c'est-à-dire au découpage d'une séquence en sous-séquences de différents types.

Trois problèmes de bases sont soulevés par l'utilisation des HMM, qui vont être abordés dans les parties suivantes :

1. étant donné une séquence d'observations $O = o_1 \cdot \dots \cdot o_T$ et un HMM H , comment évaluer efficacement $P(O|H)$, la probabilité d'observation de O étant donné H ?

2. étant donné une séquence d'observation $O = o_1 \cdot \dots \cdot o_T$ et un HMM H , comment choisir une séquence d'états $s_1 \cdot \dots \cdot s_T$ optimale dans un certain sens (i.e. expliquant au mieux O) ?

3. comment ajuster les paramètres d'un HMM H pour maximiser $P(O|H)$?

Le premier problème est un problème d'évaluation, qui peut également être vu comme un problème d'estimation de la capacité d'un modèle donné à reconnaître une séquence d'observations donnée. Le second problème se ramène à l'idée de dévoiler les états cachés S , sans y avoir accès directement. Dans la plupart des cas, le critère d'optimalité retenu influencera la séquence d'états calculée. Enfin, le troisième problème se ramène à l'entraînement d'un HMM par des séquences d'observations, en vue d'en optimiser les paramètres pour un problème spécifique donné.

VI.1 Algorithme forward-backward :

- Position du problème :

Si l'on veut calculer la probabilité de générer la séquence de symboles $O = o_1 \cdot \dots \cdot o_T$ à l'aide du HMM $H = (S, \Sigma, T, G, \pi)$, l'approche directe consiste à calculer la probabilité de génération pour chaque chemin possible et faire la somme de ces probabilités. La probabilité P_S de générer O suivant le chemin $S = s_1 \cdot \dots \cdot s_T$ est :

$$P_S = \pi(s_1)P(o_1|s_1) \cdot \dots \cdot P(s_{T-1} \rightarrow s_T)P(o_T|s_T) \quad (1.1)$$

La probabilité de générer O avec H obtenue en sommant sur l'ensemble des séquences d'états possibles est donc :

$$P(O | H) = \sum_{\{s_1 \dots s_T\}} P_S \quad (1.2)$$

Si N est le nombre d'états du HMM, alors le nombre de chemins possibles pour générer une séquence de symboles de longueur T est de l'ordre de N^T .

Comme, pour chaque chemin, le calcul de (1.1) demande de l'ordre de T opérations, le calcul de la probabilité de génération de O par H suivant (1.2) est en $O(TN^T)$. Ainsi, même pour des valeurs de T et N peu élevées, l'approche directe n'est clairement pas acceptable : avec $N = 5$ et $T = 100$ le calcul de $P(O | H)$ demande approximativement 10^{72} opérations. C'est pourquoi des alternatives doivent être mises en place.

Algorithme forward-backward

Soit $\alpha_t(s) = P(o_1 \cdots o_t, s_t = s | H)$ la probabilité d'avoir généré la séquence $O = o_1 \cdots o_t$ et d'être arrivé sur l'état s à l'instant t . Cette variable peut être calculée de manière inductive :

1. initialisation : $\alpha_1(s) = \pi(s) \cdot P(o_1 | s)$
2. induction : $\alpha_t(s) = [\sum_{s' \in S} \alpha_{t-1}(s') \cdot P(s' \rightarrow s) P(o_t | s)]$

Connaissant $\alpha_T(s)$ la probabilité d'avoir généré la séquence O et d'être arrivé sur s pour tout $s \in S$, le calcul de $P(O | H)$ est immédiat :

$$P(O | H) = \sum_{s \in S} \alpha_T(s) \quad (1.3)$$

La phase d'initialisation requiert une opération pour chaque état du HMM, donc au total $O(N)$ opérations. Pour la phase d'induction, pour chaque instant et chaque état, on réalise $O(N)$ opérations. Sommé sur l'ensemble des états et la totalité des instants, la phase d'induction requiert donc $O(N^2 T)$ opérations. Une fois les $\alpha_T(s)$ calculés, le calcul de $P(O | H)$ suivant (1.2) demande une opération pour chaque état, donc au total $O(N)$ opérations.

Le calcul de $P(O | H)$ à l'aide de cet algorithme, dit forward, ne requiert donc au total que $O(N^2 T)$ opérations. En reprenant la même application numérique que précédemment ($N = 5$ et $T = 100$), le calcul nécessite approximativement 3000 opérations.

Cet algorithme est appelé forward car l'induction est réalisée en avant : on calcule tout d'abord la probabilité de générer le premier symbole de la séquence, puis à chaque étape de l'induction on rajoute un symbole et on réitère la procédure jusqu'à ce que l'on ait calculé la probabilité de génération de la séquence entière.

Un algorithme similaire, l'algorithme backward, peut être utilisé pour réaliser ce calcul à l'envers. On utilise alors la variable backward $\beta_t(s) = P(o_{t+1} \cdots o_T | s_t = s, H)$ qui exprime la probabilité de générer la séquence $O = o_{t+1} \cdots o_T$ en partant de l'état s . L'induction suit alors le schéma :

1. initialisation : $\beta_T(s) = 1$
2. Induction: $\beta_t(s) = [\sum_{s' \in S} \beta_{t+1}(s') P(s \rightarrow s') P(o_{t+1} | s')]$

Connaissant la probabilité de générer la séquence O en partant de l'état s , le calcul de $P(O | H)$ peut alors être réalisé suivant la formule :

$$P(O | H) = \sum_{s \in S} \pi(s) \beta_1(s) \quad (1.4)$$

avec la même complexité.

VI.2 Algorithme de Viterbi

o Position du problème

Le problème posé ici est de trouver, étant donné une séquence de symboles $O = o_1 \cdot \dots \cdot o_T$ et un HMM $H = (S, \Sigma, T, G, \pi)$ la séquence d'états du HMM qui a la probabilité maximale de générer O . Ce qui nous préoccupe n'est pas la valeur de la probabilité maximale mais le chemin, appelé chemin de Viterbi, qui permet de générer la séquence O avec cette probabilité.

De manière similaire à l'approche utilisée pour le calcul de $P(O | H)$, l'approche directe pour résoudre ce problème consiste à calculer la probabilité de génération suivant tous les chemins possibles et de choisir parmi ces chemins celui qui a la probabilité la plus élevée. Cette approche a également une complexité en $O(TN^T)$ et est donc également inapplicable. L'algorithme de Viterbi est un algorithme de programmation dynamique très similaire à l'algorithme forward et qui permet de résoudre efficacement ce problème.

• Algorithme de Viterbi

Soit $\delta_t(s) = \max_{s_1 \dots s_{t-1}} P(s_1 \dots s_t = s, o_1 \dots o_t | H)$ la probabilité maximale de générer la séquence $O = o_1 \dots o_t$ suivant un unique chemin arrivant sur l'état s à l'instant t . De la même manière que pour $\alpha_t(s)$, cette variable peut être calculée de manière inductive :

1. Initialisation : $\delta_1(s) = \pi(s)P(o_1 | s)$
2. Induction: $\delta_t(s) = \max_{s' \in S} (\delta_{t-1}(s')P(s' \rightarrow s)) P(o_t | s)$

Connaissant $\delta_T(s)$ pour tous les états s , on peut calculer la probabilité maximale $P(O | H, V)$ de générer O avec H suivant un simple chemin V par :

$$P(O | H, V) = \max_{s \in S} (\delta_T(s)) \quad (1.5)$$

En fait, ce n'est pas la valeur de cette probabilité qui nous intéresse mais réellement le chemin qui permet de générer O avec cette probabilité. On doit donc, à chaque étape t de l'induction et pour chaque état s , mémoriser l'état $s' = \psi_t(s)$ qui maximise l'équation d'induction. Alors:

1. Initialisation: $\psi_1(s) = 0$
2. Induction: $\psi_t(s) = \operatorname{argmax}_{s' \in S} (\delta_{t-1}(s')P(s' \rightarrow s))$

Une fois les variables $\delta_t(s)$ et $\psi_t(s)$ calculées pour chaque étape de l'induction et pour chaque état, il ne reste plus qu'à lancer une procédure inductive de rétro-propagation pour "dérouler" le chemin de Viterbi $s_1^* \cdots s_T^*$:

1. initialisation : $s_T^* = \operatorname{argmax}_{s \in S} (\delta_T(s))$

2. induction : $s_t^* = \psi_{t+1}(s_{t+1}^*)$, $t \in \{T-1 \cdots 1\}$

Mise à part la phase de rétro-propagation, l'algorithme de Viterbi est très similaire à l'algorithme forward. La principale différence résulte de la maximisation des probabilités attachées aux états précédents au lieu du calcul de la somme de ces probabilités.

VII. Apprentissage d'un HMM

Les deux algorithmes précédents supposent disposer d'un HMM construit et paramétré de manière à modéliser de façon satisfaisante les séquences à traiter. La question au centre de ce paragraphe est celle de la construction d'un tel HMM. Dans le cas le plus favorable, le HMM recherché peut être construit directement à partir des connaissances a priori dont on dispose sur le domaine. Dans la plupart des applications, le HMM doit être construit à l'aide d'un algorithme d'apprentissage. Ces algorithmes sont appliqués sur un ensemble de séquences représentatives des séquences que l'on souhaite modéliser et appelées séquences d'apprentissage. On peut distinguer dans le problème de l'apprentissage d'un HMM deux cas de figure distincts, suivant que la structure (nombre d'états du HMM et transitions autorisées) est connue ou ne l'est pas. Lorsque la structure est connue, le problème se réduit à un problème d'entraînement consistant à estimer les paramètres numériques (distributions de probabilité de première visite, de transition et de génération) de manière à expliquer au mieux les séquences d'apprentissage. Pour certaines applications, on ne dispose pas de connaissances suffisantes pour inférer naturellement la structure du HMM. L'apprentissage devient alors encore plus difficile. Il ne suffit plus de paramétrer une structure mais il faut également déduire cette structure des exemples fournis.

VII.1 Apprentissage d'un HMM à structure connue

On souhaite ici apprendre, ou entraîner, un HMM à partir d'une structure connue. On dispose pour cela d'un ensemble d'apprentissage composé de séquences supposées

représentatives des séquences que l'on souhaite modéliser. Une approche possible est, suivant le principe du maximum de vraisemblance, de chercher les paramètres $\lambda = \langle T, G, \pi \rangle$ du HMM qui maximisent la probabilité de génération des séquences d'apprentissage. Soit $O = \{O_1 \cdot \dots \cdot O_K\}$ l'ensemble des séquences d'apprentissage. On suppose que ces séquences sont indépendantes et donc que la probabilité de générer l'ensemble d'apprentissage est simplement le produit des probabilités de génération de chacune des séquences. L'objectif est alors de trouver les paramètres $\lambda = \langle T, G, \pi \rangle$ maximisant :

$$P(O | H) = \prod_{k=1}^K P(O^k | H) \quad (1.6)$$

Une approche alternative au principe du maximum de vraisemblance est de chercher à maximiser la probabilité de génération des séquences d'apprentissage suivant leur chemin de Viterbi. Le but est alors de trouver les paramètres $\lambda = \langle T, G, \pi \rangle$ maximisant

$$P(O | H, v) = \prod_{k=1}^K P(O^k | H, V^k) \quad (1.7)$$

ou V^k est le chemin de Viterbi de la séquence O^k dans H , et $V = \{V^1 \cdot \dots \cdot V^k\}$. Les paramètres ne sont donc pas estimés en maximisant la vraie probabilité de génération des séquences d'apprentissage mais la probabilité de génération suivant les chemins les plus probables. Cette approche peut sembler moins rigoureuse que l'approche basée sur le principe du maximum de vraisemblance, mais elle est très utilisée en pratique et possède quelques arguments en sa faveur, en particulier celui selon lequel la probabilité de génération d'une séquence de symboles suivant son chemin de Viterbi est en général beaucoup plus élevée que suivant n'importe quel autre chemin. Cette observation a conduit à l'hypothèse, appelée hypothèse de Viterbi, que tous les chemins, excepté le chemin de Viterbi, ont une probabilité nulle ou négligeable d'engendrer la séquence.

Sous cette hypothèse la probabilité de génération d'une séquence O^k par un HMM H peut être approximée par la probabilité de génération de O^k suivant son chemin de Viterbi dans H . On a alors :

$$P(O | H) \approx P(O | H, V)$$

Suivant la méthode d'estimation des paramètres choisie, il existe deux heuristiques permettant l'entraînement des HMMs. Ces deux algorithmes, très similaires, sont tous deux issus d'une méthode générale servant à l'estimation des paramètres d'une grande famille de modèles probabilistes et appelée: algorithme d'*Expectation-Maximization* (EM).

Entraînement de Viterbi

Il existe un cas où il est possible de paramétrer la structure du HMM de manière à maximiser l'équation (1.7) de manière optimale. Ce cas est celui rencontré lorsque l'on connaît les chemins de Viterbi des séquences d'apprentissage dans la structure.

En effet, on peut alors associer à chaque état, chaque transition et chaque symbole attaché aux états du HMM le nombre de fois où ils sont utilisés pour générer l'ensemble d'apprentissage.

Soit n_s , $n_{s \rightarrow s'}$ et n_s^o respectivement le nombre de fois où l'état s est utilisé, le nombre de fois où la transition $s \rightarrow s'$ est utilisée et le nombre de fois où le symbole o est généré par l'état s dans les chemins de Viterbi. Alors (1.7) peut être réécrite :

$$P(O | H, V) = \left[\prod_{s \in S} \prod_{o \in \Sigma} P(o|s)^{n_s^o} \prod_{s' \in S} P(s \rightarrow s')^{n_{s \rightarrow s'}} \right] \quad (1.8)$$

Maximiser cette formule revient à maximiser indépendamment chacun de ses sous-produits. On peut donc estimer les paramètres T en maximisant l'expression $\prod_{o \in \Sigma} P(o|s)^{n_s^o}$ et les estimateurs de T et G recherchés se calculent donc par :

$$\hat{P}(s \rightarrow s') = \frac{n_{s \rightarrow s'}}{n_s} \quad (1.9)$$

$$\text{Et } \hat{P}(o | s) = \frac{n_s^o}{n_s} \quad (1.10)$$

Cette méthode d'estimation n'est possible que dans le cas favorable où les chemins de Viterbi sont connus. Lorsque ce n'est pas le cas, le problème est évidemment plus difficile et l'algorithme d'entraînement de Viterbi peut alors être une solution : c'est une méthode de réestimation itérative qui consiste, à partir d'un paramétrage initial du HMM, à calculer les chemins de Viterbi des séquences d'apprentissage à l'aide de l'algorithme de Viterbi. Les chemins de Viterbi sont utilisés pour calculer le nombre de fois où chaque transition, chaque état et chaque symbole attaché aux états est utilisé. On réestime alors à l'aide des formules (1.9) et (1.10) les paramètres du HMM, on reparamètre la structure à l'aide de ces estimations et on réitère la procédure jusqu'à stabilité. On peut montrer que $P(O | H, V)$ augmente à chaque itération et que l'algorithme converge vers un optimum local. Malheureusement, il existe généralement un grand nombre d'optima locaux et le paramétrage obtenu par ce processus dépend fortement du paramétrage initial choisi.

Entraînement de Baum-Welch

L'algorithme d'entraînement de Baum-Welch est un algorithme qui cherche à estimer les paramètres $\lambda = \langle T, G, \pi \rangle$ du HMM en maximisant (1.6). C'est un algorithme de réestimation itératif qui fonctionne sur le même principe que l'algorithme d'entraînement de Viterbi. Dans ce dernier, on compte le nombre de fois où chaque état, chaque transition et chaque symbole attaché aux états est utilisé dans les chemins de Viterbi. Ici on veut maximiser les probabilités de génération réelles, et non celles des chemins les plus probables.

On associe aux états, aux transitions et aux symboles le nombre de fois où ils sont utilisés pour toutes les séquences et tous les chemins susceptibles de générer les séquences, pondérées par la probabilité du chemin. Ces comptes pondérés sont alors utilisés pour réestimer les paramètres du modèle de la même manière que pour l'algorithme d'entraînement de Viterbi.

De manière plus formelle, considérons pour commencer une unique séquence d'observations $O = o_1 \cdot \dots \cdot o_T$. Il s'agit alors de trouver les paramètres $\lambda = \langle T, G, \pi \rangle$ maximisant la probabilité $P(O | H)$ de générer O avec H . Nous verrons ensuite comment étendre l'algorithme à un ensemble de séquences d'observations. Soit $\xi_t(s, s') = P(s_t = s, s_{t+1} = s' | O, H)$ la probabilité qu'en générant O avec H on passe par l'état s à l'instant t et par l'état s' à l'instant $t + 1$. On a alors:

$$\xi_t(s, s') = \frac{P(s_t = s, s_{t+1} = s', O | H)}{P(O | H)}$$

et en utilisant les variables forward et backward :

$$\xi_t(s, s') = \frac{\alpha_t(s)P(s \rightarrow s')P(o_{t+1}|s')\beta_{t+1}(s')}{\sum_{q \in S} \sum_{r \in S} \alpha_t(q)P(q \rightarrow r)P(o_{t+1}|r)\beta_{t+1}(r)}$$

Si $\gamma_t(s) = P(s_t = s | O, H)$ est la probabilité qu'en générant O avec H on se trouve sur l'état s à l'instant t , on a :

$$\gamma_t(s) = \sum_{s' \in S} \xi_t(s, s')$$

Si l'on somme $\gamma_t(s)$ sur l'ensemble des instants t , on obtient une quantité que l'on peut interpréter comme l'espérance du nombre de fois où l'état s est utilisé pour générer la séquence O . De même, si on somme $\xi_t(s, s')$ sur l'ensemble des instants t , on obtient une quantité que l'on peut interpréter comme l'espérance du nombre de fois où la transition $s \rightarrow s'$

est utilisée pour générer la séquence O . On a donc un estimateur \hat{H} du HMM défini par les expressions suivantes :

$$\hat{\pi}(s) = \gamma_1(s)$$

$$\hat{P}(s \rightarrow s') = \frac{\sum_{t=1}^{T-1} \xi_t(s, s')}{\sum_{t=1}^{T-1} \gamma_t(s')}$$

$$\hat{P}(O|s) = \frac{\sum_{t=1, o_t=0}^T \gamma_t(s)}{\sum_{t=1}^T \gamma_t(s)}$$

En utilisant ces estimateurs, Baum a montré que :

1. soit le modèle initial H définit un point critique de la fonction de vraisemblance, auquel cas $\hat{H} = H$
2. soit \hat{H} est tel que $P(O|\hat{H}) > P(O|H)$, c'est à dire que \hat{H} est plus à même de produire O que H

Dans le cas où l'on dispose d'un ensemble de séquences d'observations $O = \{O_1 \cdot \cdot \cdot O_k\}$, ces estimateurs se généralisent simplement en sommant numérateur et dénominateur sur l'ensemble des séquences d'apprentissage. On peut alors proposer une version complète de l'algorithme de Baum-Welch, en appliquant itérativement cette procédure d'estimation des paramètres de H sur O . Comme pour l'algorithme d'entraînement de Viterbi, on peut montrer que la probabilité de génération des séquences $P(O|H)$ augmente à chaque itération et que l'algorithme converge vers un optimum local. Néanmoins, dans ce cas là, l'optimum local n'est jamais atteint et il est nécessaire de définir un critère d'arrêt.

VII.2 Apprentissage d'un HMM à structure inconnue

Pour certaines applications, aucune connaissance ne permet d'inférer naturellement la structure du modèle de Markov caché. Le problème est alors, non seulement de paramétrer une structure, mais encore d'inférer cette structure à partir des exemples d'apprentissage. Plusieurs méthodes ont été proposées pour résoudre ce problème. Remarquons tout d'abord qu'une approche du type maximum de vraisemblance n'est clairement pas suffisante. En effet, il n'est pas difficile de construire un HMM qui a la probabilité maximale de générer les séquences d'apprentissage. Il suffit de construire un HMM équivalent à une disjonction de ces séquences : chaque chemin représente exactement une séquence d'apprentissage ; chaque état ne génère qu'un des symboles de l'alphabet et n'a qu'une transition sortante. Le HMM

construit de cette manière est spécifique aux séquences d'apprentissage (il ne peut générer que ces séquences), et est donc inutilisable en pratique puisque incapable de reconnaître d'autres séquences. On se trouve alors face à un problème classique en apprentissage, celui de la recherche d'un biais inductif, c'est-à-dire d'une règle de généralisation a priori permettant d'inférer des HMMs qui puissent être utilisés sur de nouvelles séquences.

Apprentissage par généralisation

L'apprentissage d'un HMM par généralisation procède de manière gloutonne et ascendante. Cette technique consiste à construire un HMM H spécifique comme exposé ci-dessus, et, par une opération élémentaire de fusion d'états, à transformer progressivement ce HMM en un modèle plus simple et plus général. Ce gain en généralité s'accompagne généralement d'une baisse de la vraisemblance des séquences d'apprentissage.

Cette méthode construit ainsi une suite de HMMs $(H_i)_{1 \leq i \leq n}$, où H_n est le HMM obtenu en fusionnant deux états de H_{n-1} . Après chaque fusion, les paramètres de la structure sont réestimés de manière à maximiser les chemins de Viterbi des séquences d'apprentissage. Plutôt que d'utiliser l'algorithme d'entraînement de Viterbi à chaque itération, les chemins de Viterbi sont supposés préservés par les fusions. En d'autres termes, on suppose que les chemins de Viterbi d'après la fusion peuvent être déduits des chemins de Viterbi d'avant la fusion en remplaçant simplement les occurrences des états fusionnés par l'état résultant de la fusion. Sous cette hypothèse, les paramètres du HMM se réestiment aisément grâce à (1.9) et (1.10). De plus, seuls les paramètres attachés au nouvel état et à ses transitions adjacentes ont à être réestimés puisque la fusion n'a pas d'effet sur les autres états et transitions.

Différentes techniques peuvent être utilisées pour guider la fusion au cours de l'apprentissage. Il faut d'abord choisir un couple d'états à fusionner à chaque itération, puis définir un critère d'arrêt pour l'algorithme. Par exemple, on peut utiliser un biais s'exprimant comme une probabilité a priori sur l'ensemble des HMMs possibles. Les critères de fusion est alors la maximisation d'une probabilité a posteriori construite suivant le théorème de Bayes :

$$P(H | O) = \frac{P(O | H)P(H)}{P(O)}$$

où $P(H)$ est la probabilité a priori du HMM H , $P(O | H)$ est la probabilité de générer les séquences d'apprentissage O avec H (cf. (1.6) ou (1.7)) et $P(O)$ est la probabilité d'observer les séquences O . De manière à favoriser les structures de tailles réduites, on peut

exprimer la probabilité a priori du HMM à l'aide d'une fonction du type $P(H) = e^{-l(H)}$, où $l(H)$ est la taille du HMM calculée en fonction du nombre d'états et de transitions. A chaque itération, on choisit alors le couple d'états dont la fusion engendre l'augmentation la plus importante de la probabilité a posteriori $P(H | O)$ ou plus simplement du produit $P(O | H)P(H)$, puisque $P(O)$ est une constante indépendante du HMM. La procédure est stoppée lorsque plus aucun couple ne permet d'augmenter cette probabilité. Le critère de fusion et d'arrêt définit donc un compromis entre généralisation et vraisemblance des séquences d'apprentissage.

Apprentissage par spécialisation :

A l'inverse de l'apprentissage par généralisation, l'apprentissage par spécialisation consiste à partir d'un HMM très général (par exemple composé d'un unique état qui boucle sur lui même), et par une opération de spécialisation (création d'un nouvel état, ajout d'une transition entre deux états existants), de construire progressivement un HMM plus spécifique de manière à augmenter la vraisemblance des séquences d'apprentissage.

VIII. Différents types de HMM

Quelques structures de modèles

Différentes topologies de HMM peuvent être envisagées, à commencer par le modèle ergodique, celui où chaque état peut être atteint à partir de n'importe quel autre en un nombre fini de transitions (fig. II.2). Le modèle est donc caractérisé par une matrice T aux coefficients strictement positifs.

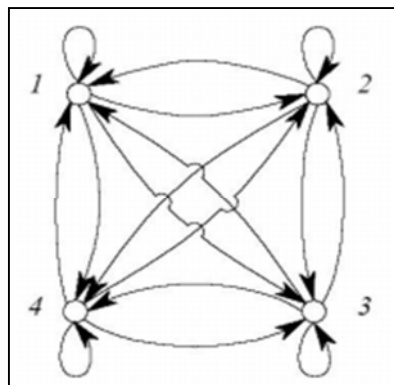


Fig. (II.2) – modèle de chaîne de Markov ergodique.

Une structure couramment utilisée est celle des modèles de Bakis, dits également modèles gauche-droite (figure II.3), ainsi appelés parce qu'ils n'autorisent aucune transition

d'un état vers un autre d'indice inférieur : les états qui se succèdent ont donc des indices égaux ou supérieurs aux précédents. Une fois dans le dernier état, le système est condamné à y rester : c'est pourquoi la probabilité initiale du premier état est posée égale à 1, les autres étant égales à 0.

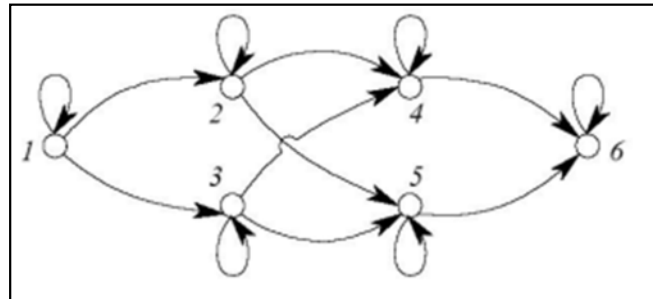


Fig. (II.3) – modèle de Bakis

Dans le même ordre d'idées, les sauts de nombreux états consécutifs sont souvent interdits ($T(i, j) = 0, j - i > \Delta$). Une variante du modèle gauche-droite est le modèle gauche-droite parallèle, utile, par exemple, en reconnaissance de la parole, où il permet de prendre en compte la possibilité de multiples prononciations d'un même mot.

Les procédures de réestimation des paramètres d'un modèle de Markov sont telles que, si certains de ces paramètres sont initialement mis à zéro, ils resteront nuls tout au long de la phase d'entraînement. Ceci permet donc de conserver les contraintes imposées à la structure du modèle.

Dans le cas spécifique d'un modèle de Bakis, l'entraînement ne peut toutefois pas être réalisé à partir d'une seule séquence d'observations car, étant donné la nature du modèle, trop peu d'observations dans une séquence sont associées à un état particulier. Afin de disposer de données suffisantes pour l'estimation de paramètres, la procédure de réestimation doit donc être modifiée de façon à prendre en compte des séquences d'observations multiples. Bien sûr, il existe évidemment encore bien d'autres topologies de Modèles de Markov Cachés.

Durée des états

On peut montrer que la probabilité qu'un état s_i produise d'observations élémentaires est une fonction exponentielle décroissante de d . Ceci constitue l'une des plus grandes faiblesses des modèles de Markov car cette modélisation s'avère inappropriée pour représenter la plupart des signaux physiques. Une amélioration possible du modèle est d'inclure explicitement les distributions $p_i(d)$ de la durée des états, au prix cependant d'une augmentation du nombre de

paramètres et, surtout, du volume de calcul. Une autre solution, beaucoup plus avantageuse, consiste à introduire une condition de durée minimale des états du modèle de Markov caché. Ceci s'avère particulièrement utile lorsque les séquences comportent davantage de boucles que de transitions d'un état vers un autre, les exponentielles prenant alors des valeurs beaucoup trop faibles.

Modèles de Markov à temps continu

Jusqu'à présent, seuls des Modèles de Markov Cachés modélisant des séquences d'observations discrètes ont été envisagés : ces observations prenaient des valeurs appartenant à un alphabet fini, et leurs probabilités étaient définies dans chaque état par un ensemble fini de valeurs. Le problème est que les observations sont souvent des vecteurs continus : l'utilisation de modèles à distributions discrètes implique donc une phase préalable de quantification de ces vecteurs, avec les dégradations qui en résultent. Il est dès lors intéressant d'inclure des densités d'observations continues dans les modèles de Markov. Néanmoins, afin de limiter le nombre de paramètres de ces distributions, il faut apporter quelques restrictions à la forme de ces densités de probabilité. La forme la plus commune de ces fonctions se présente comme une somme finie de gaussiennes, ou multi gaussiennes :

$$P(O|s) = \sum_{m=1}^M c_{sm} N_o(\mu_{sm}, \Sigma_{sm})$$

Les paramètres du modèle à ajuster sont alors, en plus de T , G et π , les coefficients de pondération c_{sm} , les vecteurs de moyenne μ_{sm} et les matrices de covariance Σ_{sm} . Les formules de réestimation nécessitent alors la définition de :

$$\gamma_t(s,m) = \left\{ \frac{\alpha_t(s)\beta_t(s)}{\sum_{s' \in S} \alpha_t(s')\beta_t(s')} \right\} \left\{ \frac{c_{sm} N_o(\mu_{sm}, \Sigma_{sm})}{\sum_{n=1}^M c_{sn} N_o(\mu_{sn}, \Sigma_{sn})} \right\}$$

qui généralise $\gamma_t(s)$ au cas continu. Ce coefficient s'interprète comme la probabilité de se trouver en s au temps t , en ne prenant en compte que la même composante de la multigaussienne de l'état considéré en o_t . Les formules de réestimation deviennent alors :

$$c_{sm} = \frac{\sum_{t=1}^T \gamma_t(s,m)}{\sum_{t=1}^T \sum_{n=1}^M \gamma_t(s,n)}$$

$$\mu_{sm} = \frac{\sum_{t=1}^T \gamma_t(s,m) o_t}{\sum_{t=1}^T \gamma_t(s,m)}$$

et la procédure demeure inchangée pour la réestimation des autres paramètres. La représentation continue par multigaussienne des probabilités d'émission de chaque état permet de bien mieux estimer les densités d'observations réelles qu'avec un modèle à temps discret. La grande quantité de paramètres qui doivent alors être déterminés représente cependant un lourd handicap du modèle : si d est la dimension des vecteurs de primitives, l'estimation des probabilités de génération $P(o|s)$ nécessite, pour chaque état, la connaissance de $(d^2 + d + 1)M$ paramètres, contre $|\Sigma|$ dans le cas d'un alphabet fini. L'entraînement d'un modèle de Markov à temps continu nécessite donc la disponibilité d'une très grande quantité de données.

Ces difficultés peuvent être limitées en supposant que les éléments non diagonaux des matrices de covariance sont nuls, c'est-à-dire en admettant que les composantes du vecteur caractéristique sont non corrélées. Ceci permet de réduire de d^2 à d le nombre de paramètres de chaque matrice de covariance. Une autre solution est aussi de limiter le nombre de gaussiennes : l'utilisation d'une seule gaussienne, même si cela restreint la qualité d'estimation des densités d'observations, peut conduire à de meilleurs résultats que l'utilisation d'un modèle de Markov à temps discret.

IX. Conclusion

Ce chapitre nous a permis d'avoir une idée générale sur la chaîne de Markov cachée, qui est une méthode de classification non supervisée utilisée pour la segmentation. Et comme tout classifieur elle est basée sur deux phases: celle d'apprentissage et celle de test.

Les chaînes de Markov cachées sont parmi les modèles de Markov cachés les plus simples et les plus anciens. Le prochain chapitre va présenter les résultats et ces interprétations qu'on a obtenues par notre méthode.

I. Introduction

La segmentation des structures dans les images biomédicales est une tâche essentielle à une multitude d'applications clés dont le diagnostic assisté, le suivi de pathologies et la recherche clinique. Bien qu'important, le processus de segmentation est cependant très complexe, dû notamment au faible contraste, et au bruit, typiquement présents dans les images médicales. Par ailleurs, alors que des approches efficaces de segmentation ont été proposées au fil des années, la plupart de ces approches requièrent l'intervention d'un humain pour étiqueter manuellement les régions d'intérêt dans l'image à segmenter, une tâche à la fois longue et complexe.

II. Rappel:

On a exploité des images de la base de donnée MIAS de taille 1024*1024 pixels dont le sein est infecté par des masses mammaires.

Les 15 images utilisées dans notre approche sont organisée dans le tableau comme suivant :

1^{ère} Colonne: numéro de référence de la base de données MIAS

2^{ème} Colonne : caractère du tissu d'arrière-plan

F - Fatty (adipeux)

G - Fatty-glandular (glandulair adipeux)

D - Dense-glandular (glandulair dense)

3^{ème} Type de pathologie

CIRC: Well-defined/circumscribed masses

SPIC: Spiculated masses

4^{ème} Colonne: sévérité de la pathologie :

B-bénin

M- malin

Image	Caractère du tissu	Type de masse	Sévérité d'anormalité
mdb025	F	CIRC	B
mdb028	F	CIRC	M
Mdb184	F	SPIC	M
Mdb290	D	CIRC	B
Mdb132	F	CIRC	B
Mdb148	F	SPIC	M
Mdb178	G	SPIC	M
Mdb195	F	SPIC	B
Mdb193	D	SPIC	B
Mdb019	G	CIRC	B
Mdb202	D	SPIC	M
Mdb188	G	SPIC	B
Mdb023	G	CIRC	M
Mdb021	G	CIRC	B
Mdb181	G	SPIC	M

Tableau (III.1) Base de donnée des images utilisée dans notre travaille

III. Expérimentation :

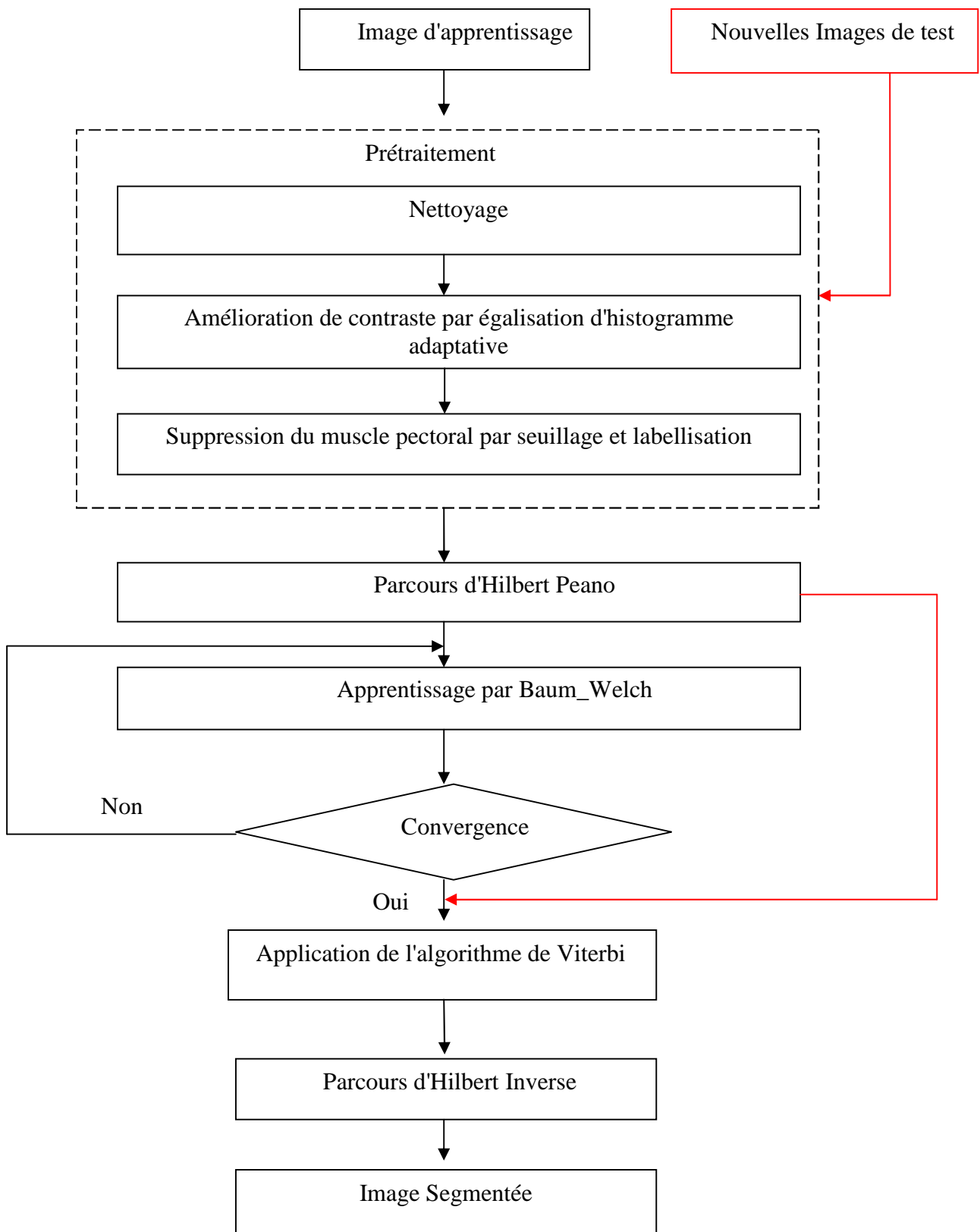


Figure (III.1). Diagramme bloc représentatif de la méthode proposée

Explication de la méthode proposée

1. Le prétraitement

On charge une image d'entrée, on effectue un prétraitement suivant ces étapes :

- Nettoyage

De bruit : par un filtre médian de taille 5*5

Des artefacts par seuillage manuel d'un seuil compris entre 20 et 250 pour tous les images, et labellisation pour extraire la région d'intérêt présentée par le sein.

- Amélioration de contraste par l'algorithme d'égalisation d'histogramme adaptative.
- Suppression du muscle pectoral par seuillage globale et labellisation

2. Parcours d'Hilbert Peano

Ce parcours est pour objectif de transformer l'image 2D à un vecteur 1D .ce dernier sera la séquence des données d'apprentissage

3. Apprentissage HMM par Baum_Welch

L'apprentissage du HMM est réalisé avec 3 états cachés, nous avons comme donnée le vecteur de la séquence des NG donc nous allons trouver les paramètres (A,B, Π) du model HMM qui correspond au mieux aux séquences observées après un nombre des itérations = 100.

4. Convergence

Pour chaque itération un changement fractionnel de la probabilité est calculé, et lorsqu'on atteint une valeur de probabilité maximale inférieure ou égale à 10^{-5} notre algorithme d'apprentissage va s'arrêter sinon nous allons recalculer les paramètres de notre model HMM.

5. Algorithme de Viterbi

Avec le model reconstruit à l'étape de l'apprentissage nous avons estimé les états cachés qui sont les classes dans notre cas en se basant sur la probabilité maximale.

6. Parcours d'Hilbert Inverse

Le vecteur des états cachés sera parcouru par le parcours d'Hilbert Peano inverse pour reconstruire la matrice des classes.

7. Test (segmentation)

Pour la phase de test nous avons utilisé 15 nouvelles images, l'algorithme de viterbi nous a permis de reconnaître la classe de chaque pixel de chaque image.

La dernière étape effectuée est la segmentation où nous allons segmenter les différentes régions de l'image et extraire la masse à partir de la région d'intérêt.

IV. Résultat et discussion:

Dans cette section, la segmentation par HMM est évaluée sur les images mammographiques en premier lieu (évaluation qualitative).et une validation est proposée pour prouver la robustesse et l'efficacité de notre méthode.

1. Evaluation qualitative

Pour cette partie nous avons évalué la segmentation par HMM, en comparant avec les résultats d'une segmentation par Kmeans.

Le principe de la méthode des "Kmeans" c'est que la classification se fait sur la base du critère des plus proches voisins. Celui-ci signifie que chaque pixel est affecté à une classe s'il est très proche de son centre de gravité.

Les figures suivantes représentent les résultats de la segmentation des masses de quelques mammogrammes par HMM et Kmeans respectivement, tel que :

(a) : représente l'image originale bien contrastée.

(b) : représente le masque binaire de la masse segmentée.

(c) : représente la masse segmentée.

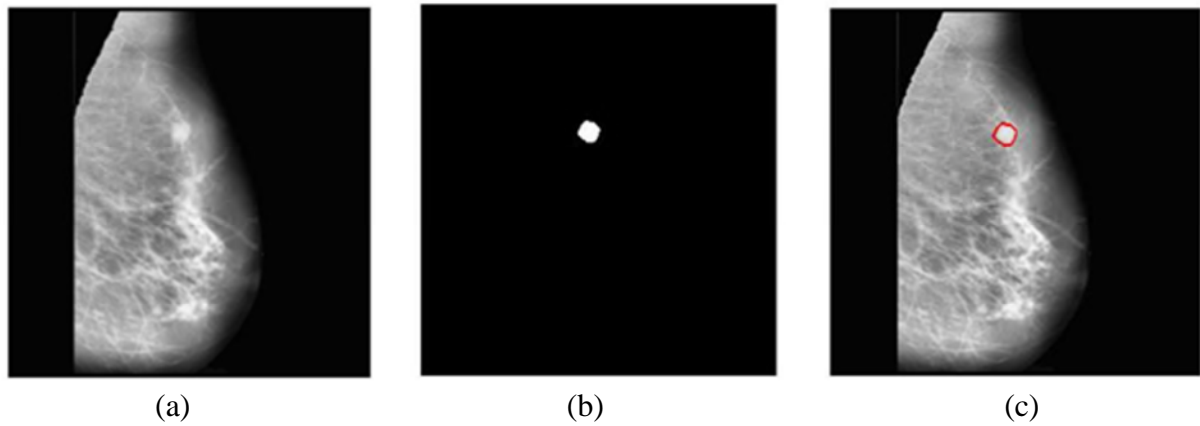


Figure (III.2).segmentation de la masse dans l'image mdb023 par HMM

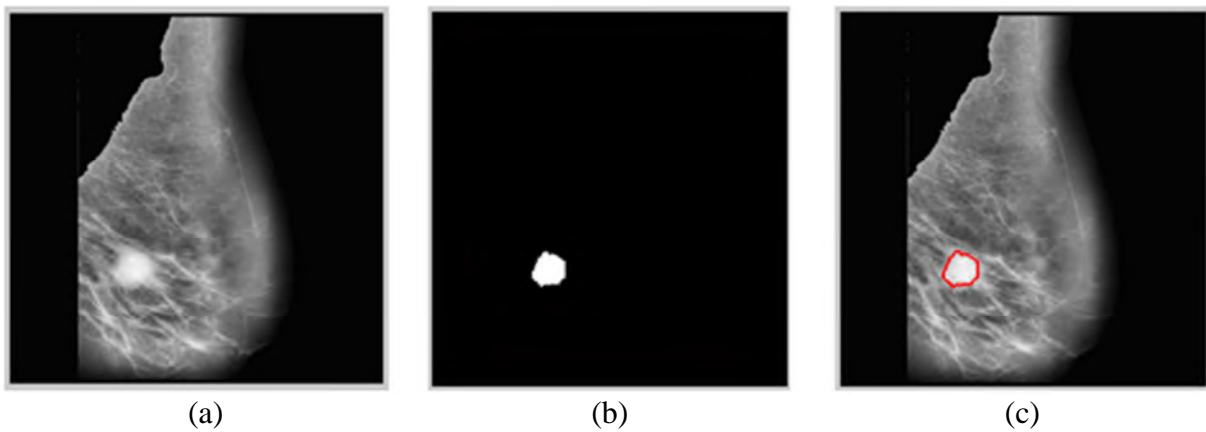


Figure (III.3). Segmentation de la masse dans l'image mdb028 par HMM

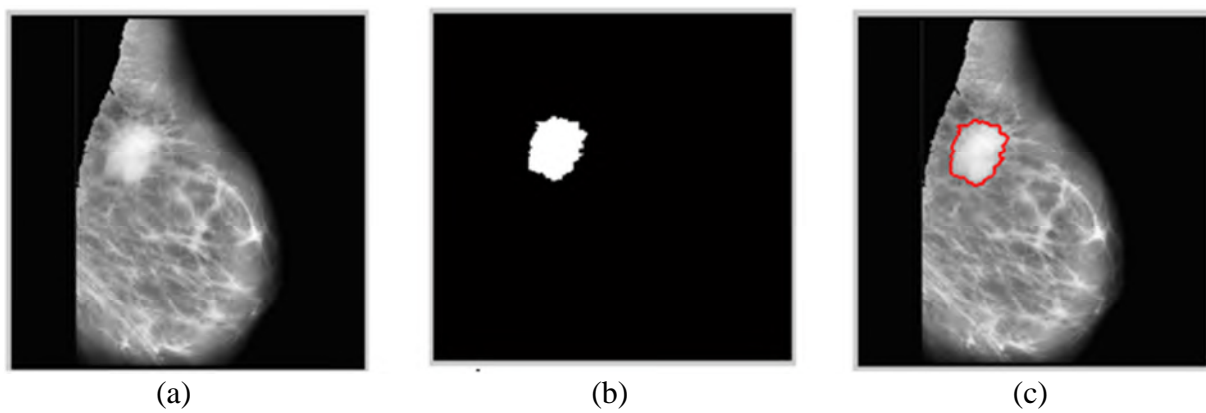


Figure (III.4). Segmentation de la masse dans l'image mdb184 par HMM

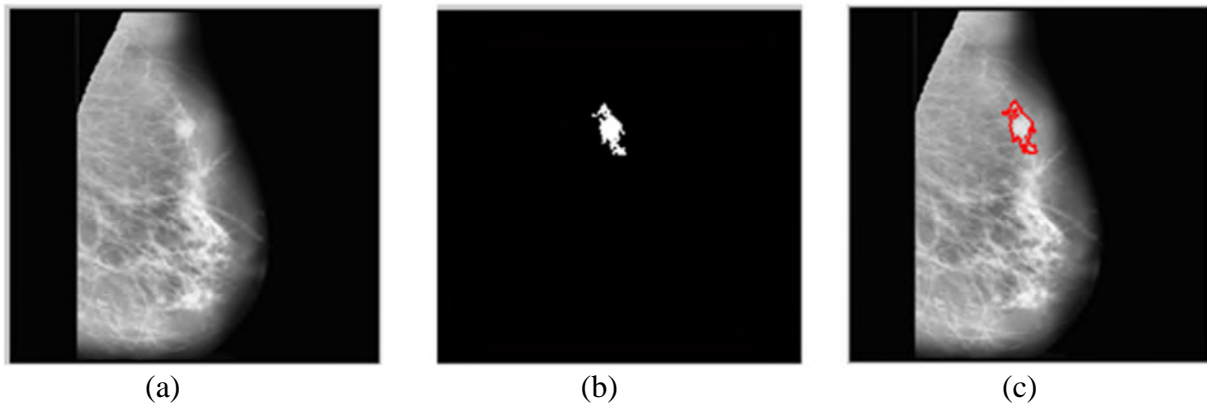


Figure (III.5). Segmentation de la masse dans l'image mdb023 par Kmeans

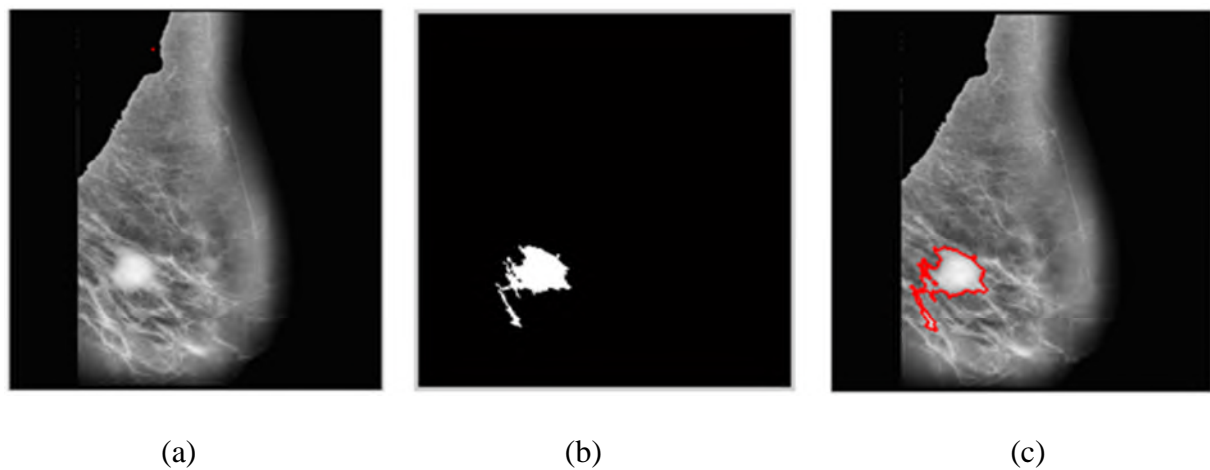


Figure (III.6). Segmentation de la masse dans l'image mdb028 par Kmeans

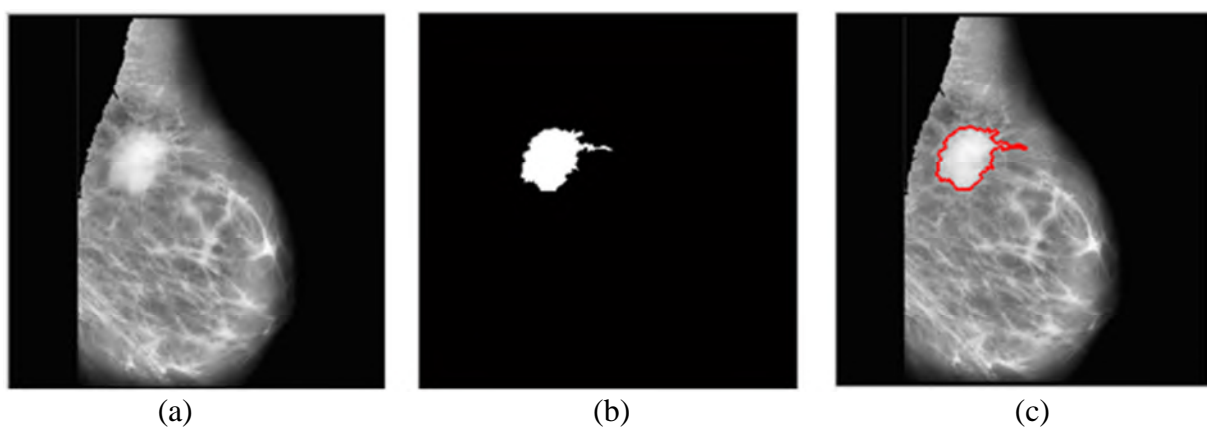


Figure (III.7). Segmentation de la masse dans l'image mdb184 par Kmeans

Il est montré clairement sur les figures que les masses ont été bien segmentées avec le model HMM proposé, par contre la segmentation par Kmeans, étant donné la similitude entre certaines parties du fond de l'image et des caractères de la masse.

2. Evaluation quantitative :

Nous avons calculé les 14 caractéristiques haralick et les moments géométriques du "gupta" de chaque image pour les deux méthodes de segmentation HMM et Kmeans et nous avons validé les résultats à l'aide d'un classifieur SVM (support vector machines) pour séparer les deux classes bénigne et maligne.

Le tableau ci-dessous représente le résultat de la classification par SVM :

La méthode	Taux de classification	sensibilité	spécificité
HMM	0.8333	1	0.6667
Kmeans	0.8333	0.6667	1

Tableau (III.2) Calcule de performance des deux méthodes

Bien que les deux méthodes ont donné le même taux de classification 83.33%, en terme de sensibilité notre méthode est sensible à la détection de la maladie à 100%.

Le classifieur SVM a reconnait toutes les classes de l'ensemble des 15 images de test sauf la classe de masse de l'image mdb023, elle est bien définit de la part du HMM Figure (III.2), mais elle est reconnait comme bénin tandis que le médecin a l'évalué comme étant malin, ce résultat du SVM est dû aux caractéristiques de la masse.

V. Conclusion :

Dans ce chapitre nous avons présenté les résultats de la segmentation des images mammographiques à l'aide de la technique de classification HMM, ainsi que son efficacité à la détection de la pathologie.

Conclusion et perspectives :

L'objectif général de ce modeste travail qui a été la détection des masses mammaires en utilisant le modèle de Markov caché comme méthode de segmentation, qui sert à l'aide au diagnostic, a été atteint.

Dans notre thèse, nous avons commencé par une définition du cancer de sein, ainsi, une indication sur la base de données MIAS, aussi un état de l'art des travaux réalisés sur cette base et le prétraitement appliqué sur laquelle pour améliorer sa qualité.

Ensuite, nous avons décrit la chaîne de Markov caché, qui est une méthode de classification non supervisée utilisée pour la segmentation. Et comme tout classifieur elle est basée sur deux phases: celle d'apprentissage et l'autre de test. Les chaînes de Markov cachées sont parmi les modèles de Markov cachés les plus simples et les plus anciens.

Enfin, nous avons présenté les résultats de la segmentation des images mammographiques à l'aide de la technique de classification HMM, ainsi que son efficacité à la détection de la pathologie.

Un futur travail visera à améliorer les moins bons résultats des patients atteints de masses mammaires par l'introduction de nombreuses images pour l'apprentissage et le test à fin de construire une base de données des caractéristiques plus riche que celle utilisée dans ce projet.

Bibliographie:

- [1] P. Kestener, Analyse multi fractale 2D et 3D à l'aide de la transformation en ondelettes : application en mammographie et en turbulence développée, PHD, l'université bordeaux I, Ecole Doctorale de Sciences Physiques et de l'Ingénieur, Novembre 2003.
- [2] Fatima EDDAOUDI, Aide à la détection des pathologies du sein par les techniques de segmentation et de codage de texture des images mammographiques, THÈSE DE DOCTORAT , Faculté des Sciences, 4 Avenue Ibn Battouta B.P. 1014 RP, Rabat –MAROC , 04 juillet 2012, <http://www.fsr.ac.ma>.
- [3] <http://peipa.essex.ac.uk/info/mias.html>.
- [4] Khamsa Djaroudib et al[], Textural Approach for Mass Abnormality Segmentation in Mammographic Images,
- [5] Shenghua Gu, et al[], A Novel Breast mass segmentation method based on patch merging and GHFCM, International Journal of Hybrid Information Technology Vol.8, No.10 (2015), pp.427-436 <http://dx.doi.org/10.14257/ijhit.2015.8.10.39> .
- [6] Sayedeh Somayeh Hosaini, Mehran Emadi, Fusion of Hidden Markov Model and Wavelet Transform to Recognise Breast Cancer, Ciência eNatura, v. 37 Part 2 jun. 2015, p. 139–145 ISSN impressa: 0100-8307 ISSN on-line: 2179-460X, Recebido: dia/mês/ano Aceito: dia/mês/ano.
- [7] Anuj Kumar Singh and Bhupendra Gupta, A Novel Approach for Breast Cancer Detection and Segmentation in a Mammogram, Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015), Department of Computer Science & Engineering, PDPM-IITDM, Jabalpur 482 005, India, Available online at www.sciencedirect.com.
- [8] H.S. Rahli, N. Benamrane, INTERPRETATION DES IMAGES DE MAMMOGRAPHIE PAR L'ALGORITHME SEARCH HARMONY, Laboratoire Imagerie et Vision Artificielle, Département d'Informatique, Université des Sciences et de la technologie Mohamed Boudiaf, Oran, Anale. Seria Informatică. Vol. X fasc. 2 – 2012, Annals. Computer Science Series. 10th Tome 2nd Fasc. – 2012.

[9] Leonardo de Oliveira Martins, *et al*[], Detection of Masses in Digital Mammograms using K-means and Support Vector Machine, *Electronic Letters on Computer Vision and Image Analysis* 8(2):39-50, 2009, Received 23th May 2008; accepted 16th July 2009, Correspondence to: ari@dee.ufma.br , Recommended for acceptance by Chulhee Lee ELCVIA ISSN:1577-5097, Published by Computer Vision Center / Universitat Autònoma de Barcelona.

[10] S.Julian Savari Antony, Dr.S.Ravi, A New Approach to Determine the Classification of Mammographic Image Using K-Means Clustering Algorithm, *International Journal of Advancements in Research & Technology*, Volume 4, Issue 2, February -2015 ISSN 2278-7763.

[11] https://fr.wikipedia.org/wiki/Filtre_m%C3%A9dian .

[12] <https://dspace.univtlemcen.dz/bitstream/112/1097/1/lakhdari-rachida.pdf>.

[13] Henri MAITRE, Chapitre 5, Description de contours et de formes

[14] L. E. Baum, T. Petrie, G. Soules, and N. weiss. A maximization technique occurring in statistical analysis of probabilistic function in Markov chains. *The annals of mathematical statistics*, 41(1): 164-171, 1970.

[15] F. Casacuberta, some relations among stochastic finite state networks used in automatic speech recognition. *IEEE Transaction on pattern Analysis and machine Intelligence*, 12(7): 691-695, July 1990.

[16] J. K Baker, The DRAGON system –an overview. *IEEE Transactions on acoustics, Speech and signal processing*. 23:24-29, 1975.

[17] L.R Rabiner. A tutorial on Hidden Markov Models and selected application in speech recognition. *Proceedings of the IEEE*, 77(2): 257-285, 1989.

[18] A. Kundu, Y. He and P. Bahl. Recognition of handwritten word: First and second order Hidden Markov Model based approach. In *CVPR'88 (IEEE computer society conference on computer vision and pattern recognition, Ann Arbor , MI, 5-9, 1988)*, pages 457-462, 1988

[19] M. Schenkel, I. Guyon, and D, Henderson. On-line cursive script recognition using time delay neural networks and Hidden Markov models. In *proc. ICASSP '94*, pages II-637-II-640, Adelaide, Australia, 1994.

- [20] David Haussler, Anders Krogh, Saira Mian, and Kimmen Sjolander. Protein modeling using Hidden Markov Models: analysis of globins. Technical Report UCSC-CRL-92-23, June 1992.
- [21] R. Durbin, S.Eddy, A. Krogh, and G.Mitchison. Biological sequence analysis probabilistic models of protein and nucleic acids. Cambridge University Press, 1998.
- [22] B.R Povlow and S.M Dunn. Texture classification using noncausal hidden Markov models. IEEE Transaction on pattern Analysis and machine intelligence, 17(10):1010-1014, October 1995.
- [23] C. Raphael. Automatic segmentation of acoustic musical signals using hidden Markov models. IEEE Transaction on pattern Analysis and Machine Intelligence, 21(4):360-370, 1998.
- [24] Laurent Brehelin, Oliver Gascuel, and Gilles Caraux. Hidden Markov models with Patterns and their application to integrated circuit testing. In 11th European conference on machine learning, 2000.
- [25] B. Benmiloud, Chaînes de Markov cachées et segmentation statistique non supervisée de séquences d'images, Thèse de doctorat, Université de Paris VII, Département CITI de l'INT, France, Décembre 1994.
- [26] K. Aas, L. Eikvil and R. Huseby, Applications of hidden Markov chains in image analysis, Pattern Recognition, vol. 32, no. 4, pp. 703–713, 1999.
- [27] N. Giordana and W. Pieczynski, Estimation of generalized multisensor hidden Markov chain and unsupervised image segmentation, IEEE Trans. Pattern Anal. Machine Intell., vol. 19, no. 5, pp. 465–475, 1997.
- [28] L. Fouque, A. Appriou and W. Pieczynski, Multiresolution hidden Markov chain model and unsupervised image segmentation, in IEEE Southwest Symp. Image Anal. Interpretation, (pp. 121–125), Austin, Texas, 2-4 April 2000.
- [29] S. Derrode, G. Mercier, J. Le Caillec and R. Garello, Estimation of sea-ice SAR clutter statistics from Pearson's system of distributions, in IEEE Int. Conf. Geosci. Remote Sensing, Sydney, Australia, July 9-13 2001

- [30] R. Fjørtoft, Y. Delignon, W. Pieczynski, M. Sigelle and F. Tupin, Unsupervised segmentation of radar images using hidden Markov chains and hidden Markov random fields , IEEE Trans. Geosci. Remote Sensing, vol. 41, no. 3, pp. 675–686, 2003
- [31] S. Faisan, L. Thoraval, J.-P. Armspach and F. Heitz, Hidden semi-Markov event sequence models : application to brain functional MRI sequence analysis , in IEEE Int. Conf. Image Processing , vol. 1, (pp. 880–883), Rochester, New York, USA, September 22-25 2002
- [32] F. Salzenstein, Modèles markoviens flous et segmentation non supervisée d’images , Thèse de doctorat, Université de Rennes I, Institut National des Télécommunications, Rennes, France, Décembre 1996.
- [33] MERATI Medjeded, Caractérisation de texture pour annotation, classification et recherche d’images médicales par le contenu, DOCTEUR en Science, Université Abou BEKR BELKAID Tlemcen, Faculté des sciences, Juin 2015,
- [34] R. Dafner, D. Cohen-Or and Y. Matias, Context-based space filling curves , Computer Graphics Forum, vol. 19, no. 3, 2000.
- [35] J. Hungersherffer and J. Wierum, On the quality of partitions based on space-filling curves , in Int. Conf. on Computational Science , (pp. 36–45), Amsterdam, The Netherlands, April 21-24 2002.
- [36] W. Skarbek, Generalized Hilbert scan in image printing , in Theoretical Foundations of Computer Vision (édité par R. Klette and W. G. Kropetsh), Akademik Verlag, Berlin, 1992.