

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université Abdelhamid Ibn Badis de Mostaganem l'UABM. Algérie.



Faculté des Sciences et de la technologie (FST)

Thèse présentée en vue de l'obtention
du diplôme de
Doctorat en sciences
Spécialité : Génie Électrique
Option : Signaux et Applications

Présentée Par

BOUKHOBZA MOHAMED EL HABIB

THÈME

**Application des réseaux de neurones artificiels pour
la classification automatique des tissus mammaires**

Soutenue le ... / ... / 2018

Devant le jury composé de :

| | | |
|------------|----------------------------|---|
| Président | BERRADJA KHADIDJA | Maître de conférences A, à l'Université de Mostaganem |
| Rapporteur | MIMI MALIKA | Professeur, à l'Université de Mostaganem |
| Examineur | AHMED FOITIH ZOUBIR | Professeur, à l'Université de Sciences et Technologies d'Oran |
| Examineur | CHAMA ZOUAOU | Professeur, à l'université de Sidi bel-Abbes. |

Dédicaces

A mes chers parents,

A ma femme et mes enfants

A toute ma famille,

A mes amis.

A tous je dis merci.

REMERCIEMENTS

Je remercie tout d'abord ma directrice de thèse M^{me}. MIMI MALIKA pour avoir mis à ma disposition tous les moyens nécessaires pour mener à bien ce projet de recherche, pour son soutien, les conseils précieux qu'elle m'a donnés au cours de ce travail et pour ses nombreuses qualités humaines.

Je tiens à exprimer mes remerciements à monsieur ABDELMALIK. TALEB AHMED, professeur, à l'IUT de Valenciennes et chercheur au LAMIH, pour son aide, sa disponibilité. Je remercie également monsieur MOHAMMED BENZOHRA, professeur, à l'université de ROUEN

*Je voudrais adresser mes sincères remerciements à M^{me} BERRADJA KHADIDJA, d'avoir accepté de présider l'honorable jury de cette thèse.
Je remercie vivement Messieurs : AHMED FOITIH ZOUBIR, Professeur A l'université de Sciences et Technologies d'Oran (USTO), CHAMA ZOUAOUI, Professeur à l'université de Sidi bel- Abbes.*

Je remercie également tous les membres du personnel du département de Génie Electrique de l'université de Mostaganem.

Mes remerciements vont également à tous les membres du laboratoire de recherche signaux et systèmes LSS.

Table des matières

| | |
|---|----|
| Résumé | 01 |
| Introduction | 04 |
| Chapitre 1- Etat de l'art sur L'interprétation de la mammographie | 07 |
| 1. L'interprétation de la mammographie | 07 |
| 1.1 Structure du sein | 07 |
| 1.2 Le cancer du sein | 08 |
| 1.3 La mammographie | 09 |
| 1.3.1 Les incidences en mammographie | 13 |
| 1.3.2 Mammographie numérique | 14 |
| 1.4 Le dépistage du cancer du sein | 15 |
| 1.5 Les systèmes de Détection/Diagnostic Assistées par Ordinateur (CAD) | 16 |
| 1.5.1 Détection assistée par ordinateur | 16 |
| 1.5.2 Diagnostic Assisté par Ordinateur | 17 |
| 1.6 Système d'aide au diagnostic /détection pour mammographie | 18 |
| 1.6.1 Prétraitement de la mammographie | 19 |
| 1.6.2 Classification des tissus mammaires | 22 |
| 1.6.3 Détection d'anomalie dans le tissu mammaire | 24 |
| 1.7 Base de données des mammographies | 26 |
| | |
| Chapitre 2- Les réseaux de neurones artificiels | 28 |
| 1 Réseaux de neurones artificiels | 28 |
| 2 Eléments biologiques du neurone | 28 |
| 3 Du neurone biologique au neurone formel | 30 |
| 3.1 Fonctions de transfert | 32 |
| 4 Le perceptron | 34 |
| 5 Perceptron multicouches | 34 |
| 6 Architecture des réseaux de neurones | 35 |
| 6.1 Les réseaux de neurones non bouclés | 35 |
| 6.2 Les réseaux de neurones bouclés | 36 |
| 7 L'apprentissage des réseaux de neurones | 37 |
| 7.1 Les types d'apprentissage | 37 |
| 7.1.1 L'apprentissage supervisé | 37 |
| 7.1.2 L'apprentissage non supervisé | 38 |

| | |
|--|----|
| 7.2 L'algorithme d'apprentissage | 39 |
| 7.2.1 L'algorithme de rétro-propagation | 39 |
| 8 Propriétés fondamentales des réseaux de neurones | 42 |
| 8.1 L'approximation fonctionnelle | 42 |
| 8.2 L'approximation parcimonieuse des réseaux de neurones | 43 |
| 9 Avantage et inconvénients des réseaux de neurones | 43 |
| 9.1 Avantages | 43 |
| 9.2 Inconvénients | 43 |
| | |
| Chapitre 3- Application des Techniques de Prétraitement sur une Mammographie | 44 |
| 1. Prétraitement de la mammographie | 44 |
| 2. Description de la méthode de prétraitement | 45 |
| 2.1 Orientation de la mammographie | 47 |
| 2.2 Ajustement de la mammographie | 47 |
| 2.2.1 Transformation linéaire de l'histogramme | 47 |
| 2.2.2 Correction gamma | 48 |
| 2.2.3 Application de L'ajustement sur une mammographie | 49 |
| 2.3 Méthode Otsu | 49 |
| 2.4 La binarisation | 51 |
| 2.5 Le filtre médian | 52 |
| 2.6 L'étiquetage par composantes connexes | 53 |
| 2.7 Suppression des artéfacts de l'image | 56 |
| 3. Application de la méthode | 56 |
| | |
| Chapitre 4- Classification automatique de la densité des tissus mammaires | 58 |
| 1. Système de classification automatique de la densité des tissus mammaires | 58 |
| 1.1. Classification automatique de la densité des tissus mammaires de la base mini-MIAS avec prétraitement d'image | 59 |
| 1.1.1 Les réseaux de neurones sélectionnés pour la classification de la base mini- MIAS | 68 |
| 1.1.2 Résultats et comparaison pour la base mini-MIAS | 70 |

| | |
|--|----|
| Chapitre 5- Détection automatique de la présence d'anomalie dans le tissu mammaire | 72 |
| 1. Détection d'anomalie dans le tissu mammaire | 72 |
| 2. Conception du système de détection | 73 |
| 2. 1 détection de la présence d'anomalie dans un tissu mammaire de type gras | 78 |
| 2. 2 détection de la présence d'anomalie dans un tissu mammaire de type gras-glandulaire | 81 |
| 2. 3 détections de la présence d'anomalie dans un tissu mammaire de type dense | 84 |
| 3. Avantages du <i>système de détection</i> | 86 |
| Conclusion | 88 |
| Annexe A | 90 |
| Annexe B | 91 |
| Références | 92 |

LISTE DES TABLEAUX

| | |
|---|----|
| Tableau. 1.1 Atténuation radiologique des composants mammaires | 11 |
| Tableau. 1.2 Les méthodes de prétraitement de la mammographie | 20 |
| Tableau. 1.3 Les méthodes de classification de la densité des tissus mammaires | 22 |
| Tableau. 1.4 Les méthodes de détection d'anomalie des tissus mammaires | 24 |
| Tableau. 2.1 Fonction de transfert $y_k = \varphi(sp)$ | 33 |
| Tableau 2.2 Algorithme de retro-propagation | 41 |
| Tableau. 3.1 Application de la méthode pour les différents types de tissus mammaires | 57 |
| Tableau 4 .1 trois mammographies de la base mini MIAS et leurs histogrammes après prétraitement | 60 |
| Tableau 4 .2 Des exemples de taux de reconnaissance des tissus mammaires conformément à la classification de la base mini-MIAS réalisé pour plusieurs configurations des réseaux de neurones | 67 |
| Tableau 4 .3 . Les paramètres des réseaux de neurones sélectionnés constituant le système de classification de la base mini-MIAS | 68 |
| Tableau 4 .4 : comparaison des méthodes conformément à la classification de la base mini-MIAS. | 70 |
| Tableau 5 .1 Nombre de paramètres associé à chaque taille de sous-image | 75 |
| Tableau 5 .2 Des exemples de taux de détection d'anomalies dans les tissus mammaires graisseux (F) réalisé pour plusieurs configurations des réseaux de neurones avec 256 entrées | 78 |
| Tableau 5 .3 Les paramètres du réseau de neurones sélectionné pour la détection de présence d'anomalie dans un tissu mammaire de type graisseux (F) avec 256 entrées | 79 |
| Tableau 5 .4 Taux détection de présence d'anomalie dans les tissus graisseux F associé à chaque taille de sous-image | 80 |
| Tableau 5 .5 Des exemples de taux de détection d'anomalies dans les tissus mammaires graisseux - glandulaire (G) réalisé pour plusieurs configurations des réseaux de neurones avec 256 entrées | 81 |
| Tableau 5 .6 Les paramètres du réseau de neurones sélectionné pour la détection de présence d'anomalie dans un tissu mammaire de type graisseux-glandulaire (G) avec 256 entrées | 82 |

| | | |
|---------------|--|----|
| Tableau 5 .7 | Taux détection de présence d'anomalie dans les tissus graisseux-glandulaire (G) associé à chaque taille de sous-image. | 83 |
| Tableau 5 .8 | Des exemples de taux de détection d'anomalies dans les tissus mammaires dense (D) réalisé pour plusieurs configurations des réseaux de neurones avec 256 entrées | 84 |
| Tableau 5 .9 | Les paramètres du réseau de neurones sélectionné pour la détection de présence d'anomalie dans un tissu mammaire de type dense (Dense D).avec 256 entrées | 85 |
| Tableau 5 .10 | Taux détection de présence d'anomalie dans les tissus graisseux-dense (Dense D) associé à chaque taille de sous-image | 86 |
| Tableau B.1 | Algorithme de Levenberg--Marquardt | 91 |

LISTE DES FIGURES

| | | |
|------------|---|----|
| Figure 1.1 | Structure du sein | 08 |
| Figure 1.2 | Les composants d'un appareil de mammographie | 10 |
| Figure 1.3 | Les éléments d'une mammographie | 12 |
| Figure 1.4 | Exemples d'incidences en mammographie : a) Cranio Caudale (CC), b) Médio Latérale Oblique (MLO) et c) Profil interne | 13 |
| Figure 1.5 | les étapes du système d'aide au diagnostic / détection d'une mammographie | 19 |
| Figure 2.1 | Deux neurones connectés | 29 |
| Figure 2.2 | Représente un neurone biologique d'après un dessin de Ramon Y Cajal, pionnier des neurosciences -1900; adroite le schéma d'un potentiel d'action, le signal élémentaire de communication entre neurones (Wul 06). | 30 |
| Figure 2.3 | Mise en correspondance neurone biologique / neurone artificiel | 31 |
| Figure 2.4 | Modèle d'un neurone artificiel | 32 |
| Figure 2.5 | Perceptron multicouches (MLP) à une couche cachée (φ_c fonction de la couche cachée φ_s fonction de la couche de sortie) | 34 |
| Figure 2.6 | Un réseau de neurones non bouclé - Perceptron multicouche | 35 |
| Figure 2.7 | réseau de neurones bouclé | 36 |
| Figure 3.1 | Deux mammographies de la base Mini-MIAS | 45 |
| Figure 3.2 | Organigramme de la méthode de prétraitement | 46 |

| | | |
|-------------|---|----|
| Figure 3.3 | Orientation de la mammographie | 47 |
| Figure 3.4 | Transformation linéaire de l'histogramme | 48 |
| Figure 3.5 | Ajustement de la mammographie | 49 |
| Figure 3.6 | Histogramme divisée en deux classes | 50 |
| Figure 3.7 | Mammographie en forme binaire | 52 |
| Figure 3.8 | exemple d'application du filtre médian | 53 |
| Figure 3.9 | application du filtre médian sur une mammographie en forme binaire | 53 |
| Figure 3.10 | Prédécesseurs d'ordre 4 (a), prédécesseurs d'ordre 8 (b) | 54 |
| Figure 3.11 | L'étiquetage (ou labellisation) d'une mammographie | 55 |
| Figure 3.12 | application de la suppression des artéfacts | 56 |
| Figure 4.1 | Organigramme du système de classification de la densité des tissus mammaires conformément à la classification de la base mini-MIAS avec prétraitement d'image | 62 |
| Figure 4.2 | . RNA à 3 couches, nc neurones cachés, 1 neurone de sortie et 30 d'entrées | 64 |
| Figure 4.3 | Les trois fonctions de transfert | 65 |
| Figure 4.4 | RNA spécialisé dans l'identification de tissus (F) | 68 |
| Figure 4.5 | RNA spécialisé dans l'identification des tissus (G) et (D) | 69 |
| Figure 5.1 | Mammographie subdivisé en sous-image | 74 |
| Figure 5.2 | Conception du réseau de neurones spécialisé dans la détection d'anomalie sur une mammographie | 76 |
| Figure 5.3 | RNA spécialisé dans la détection de présence d'anomalie dans les tissus gras (F). | 79 |
| Figure 5.4 | RNA spécialisé dans la détection de présence d'anomalie dans les tissus gras-glandulaire (G). | 82 |
| Figure 5.5 | RNA spécialisé dans la détection de présence d'anomalie dans les tissus dense(D). | 85 |
| Figure A.1 | Organigramme du système de classification de la densité des tissus mammaires conformément à la classification BI-RADS. | 90 |

Résumé.

Le cancer du sein est un problème de santé publique. La prolifération de cette maladie à pousser les chercheurs à développer des techniques d'aide au diagnostic. L'imagerie médicale constitue une technique importante pour assurer une qualité meilleure du diagnostic clinique. Reste le problème de la qualité de l'interprétation de l'image médicale. Les programmes de dépistage organisés du cancer du sein, conduit à une augmentation du nombre d'images médicales a interprété et les besoins de qualité en matière de santé publique rendent obligatoire le recours aux systèmes d'aide au diagnostic (SAD) par ordinateur. On propose un système à base de réseau de neurones artificiels capables de détecter la présence ou non d'anomalie sur une mammographie. Cependant, la qualité de l'interprétation d'une mammographie reste variable. Une des caractéristiques importantes de l'anatomie et de la physiologie du sein est la densité du tissu mammaire. La densité est importante pour deux raisons principales : premièrement, la densité mammaire accrue est associée à une diminution de la sensibilité de la mammographie pour la détection du cancer du sein (Sch 76). Deuxièmement, la densité du sein est l'un des plus importants facteurs de risque connus pour le cancer du sein (Pre 70) ; (Byd 14) .Pour toutes ces raisons, le classement automatique de la densité des tissus est un processus important dans le diagnostic. Dans cette étude, nous décrivons une méthode pour la classification de la densité globale du sein en utilisant les réseaux de neurones artificiels, cette approche nécessite une étape de prétraitement. Ensuite on utilise un système à base de réseau de neurones artificiels capable de détecter la présence ou non d'anomalie dans un tissu mammaire. La validité de notre méthode est démontrée en utilisant 180 mammographies de la base de données mini-MIAS.

Mots-clés : Mammographie, Classification automatique, Détection, Réseau de neurones artificiel, Systèmes d'aide au diagnostic, histogramme.

Abstract.

Breast cancer is an international public health concern. Proliferations of this disease push the researchers to develop techniques of assistance of diagnosis. The medical imagery constitutes an important technique for assuring a better quality the clinical diagnosis. Remain the problem of the quality of the interpretation of the medical image. The tracing routines organized of the breast cancer led to an increase amongst medical image interpreted and the needs for quality as regards public health make compulsory the recourse to the Assistance Systems Diagnosis (ASD) by computer. We propose a system containing a network of artificial neurons able to detect the presence or not anomaly on a mammography.

However, the quality of the interpretation of mammograms remains variable. One of the important characteristics in breast anatomy and physiology is breast tissue density. Density is important for two main reasons: First, increased breast density is associated with decreased mammographic sensitivity for the detection of breast cancer (Sch 76). Second, breast density is one of the strongest known risk factors for breast cancer (Pre 70); (Byd 14) . For these reasons, automatic tissue density classification is an important process in diagnosis.

In this study, we describe a method for overall breast density classification using artificial neural networks. This approach requires a preprocessing step.

Then we use system containing a network of artificial neurons able to detect the presence of anomalies on a breast tissue.

The validation of our method is demonstrated using 180 mammograms from mini-MIAS database.

Keywords: Mammography, Automatic Classification, Detection, artificial Network of neurons, Assistance Systems Diagnosis, histogram.

ملخص.

سرطان الثدي مشكلة صحية عمومية، انتشار هذا المرض دفع الباحثين لتطوير التقنيات المساعدة للتشخيص. التصوير الطبي هو تقنية هامة لضمان جودة أفضل للتشخيص طبي. لكن هناك مشكلة نوعية قراءة الصورة الطبية.

و هناك عدة حملات منظمة لإكتشاف سرطان الثدي، مما أدى إلى زيادة في عدد الصور الطبية، وينظر للإحتياجات النوعية في الصحة العمومية نكون بحاجة إلزامية إلى إستخدام أنظمة للمساعدة في التشخيص بإستعمال الحاسوب. ونقترح طريقة تعتمد على الشبكة العصبية الاصطناعية بغرض الكشف عن وجود أو عدم وجود سرطان بإستعمال صور الماموغرافي.

ومع ذلك، لا تزال نوعية قراءة الماموغرافي تتأثر بمتغيرات. واحدة من السمات الهامة لنسيج الثدي هو كثافة أنسجة الثدي.

الكثافة مهمة لسببين رئيسيين: أولاً، يرتبط زيادة الكثافة في نسيج الثدي بإنخفاض في حساسية التصوير الماموغرافي للكشف عن سرطان الثدي (Sch 76). ثانياً، كثافة نسيج الثدي هي واحدة من أهم عوامل الخطر المعروفة المتسببة في سرطان الثدي. (14 Byd) ; (70 Pre). لهذه الأسباب، تصنيف كثافة نسيج الثدي الألي هي عملية هامة في التشخيص.

هذه الدراسة، تقدم طريقة لتصنيف نسيج الثدي باستخدام الشبكات العصبية الاصطناعية، وهذه الطريقة تتطلب خطوة تحضيرية لمعالجة الصورة الماموغرافية. كما أنجزنا آلية تستعمل الشبكات العصبية الاصطناعية تسمح بالتعرف هل النسيج الثدي طبيعي أو غير طبيعي.

وقد تم التحقق من فعالية هذه الطريقة باستخدام 180 صورة ماموغرافية من قاعدة بيانات مياس.

الكلمات الإستدلالية: الماموغرافي، التصنيف الألي، الكشف، الشبكة العصبية الاصطناعية، أنظمة المساعدة للتشخيص، مدرج تكراري.

Introduction

Dans cette étude, nous allons étudier la problématique de la détection de la présence ou non d'anomalie dans un tissu mammaire en utilisant un modèle de calcul à base de réseau de neurones artificiels. La propriété d'approximation universelle parcimonieuse justifie, dans une large mesure, l'utilisation des réseaux de neurones dans le domaine des sciences de l'ingénieur (Riv 73).

Le cancer du sein est l'un des types de cancer les plus fréquents chez les femmes, la mammographie est la méthode la plus fiable pour la détection précoce des maladies du sein (Liy 48).

Les radiologues sont toujours obligés d'examiner très finement l'image pour faire un meilleur diagnostic. D'où l'importance des systèmes d'aide au diagnostic (SAD) par ordinateur développés durant ces vingt dernières années. Plusieurs techniques de traitement d'images ont été proposées.

Un système automatique de Diagnostic Assisté par Ordinateur sert à donner un second avis au radiologue. Ce système se résume dans la plupart des cas en 3 étapes (Nbe 60).

Un tel système commence généralement par un prétraitement afin d'obtenir une image plus lisible et moins bruitée. L'étape de classification de la densité des tissus mammaires est une étape importante. Puisque la densité mammaire a un impact majeur sur la fiabilité de l'interprétation des mammographies (Sti 80). Ensuite, la détection de la présence d'anomalie dans le tissu mammaire.

La densité mammaire est un terme radiologique, définie par l'atténuation du faisceau de rayons X au travers des constituants du parenchyme mammaire. La graisse radio transparente apparaît noire sur les clichés de mammographie. Le tissu glandulaire et fibro-conjonctif radio opaque apparaît blanc (Ser 81).

En 1976, Wolfe a décrit une association entre la densité mammaire et le risque de cancer du sein (Krl 47), (Mcc 52).

Les tissus mammaires à densité élevée sont beaucoup plus susceptibles de présenter un cancer du sein (Byd 15). Les femmes dont le tissu mammaire est plus dense doivent bénéficier d'examens plus avancés. Différentes méthodes de classification ont été appliquées pour classer les tissus mammaires, parmi lesquelles l'analyse topographique de la mammographie (Che 19), celle basée sur l'algorithme des k plus proches voisins (Oli 62),(Oli 63) , (Mus 57) et celle basée sur l'arbre de décision (Oli 62),(Oli 63).

Par ailleurs, une technique de classification utilisant un histogramme multi-résolution a été proposée (Muh 56), de même qu'une méthode utilisant une approche statique (Ptr 69).

La classification est un des problèmes récurrents en vision par ordinateur, problème sur lequel les réseaux de neurones donnent de bons résultats. Grâce aux deux propriétés suivantes :

La propriété d' « Approximateurs universels » : un Perceptron Multicouche doté d'une unique couche cachée, avec un nombre suffisant de neurones, peut faire l'approximation de n'importe quelle fonction avec la précision voulue (Hor 41).

La propriété d' « Approximateurs parcimonieux » : les réseaux à une couche cachée forment une famille d'approximateurs parcimonieux, ce qui signifie qu'à nombre égal de paramètres, il est possible d'approximer correctement davantage de fonctions qu'avec des polynômes (Bar 07).

Notre objectif est de développer un système capable de faire la classification des tissus mammaires et la détection automatique de la présence d'anomalie mammaire à base des réseaux de neurones artificiels.

Le vecteur d'entrée de chaque réseau de neurones est calculé à partir d'information issues de la mammographie.

La structure de l'étude est la suivante :

Chapitre 1 : Etat de l'art sur L'interprétation de la mammographie

Chapitre 2 : Les réseaux de neurones artificiels.

Chapitre 3 : Application des techniques de prétraitement sur une mammographie.

Chapitre 4 : Classification automatique de la densité des tissus mammaires.

Chapitre 5 : Détection automatique de la présence d'anomalie dans le tissu mammaire.

Chapitre 1

Etat de l'art sur l'interprétation de la mammographie

1. L'interprétation de la mammographie

En Algérie, le cancer du sein vient en tête des cas de cancer recensés avec 6625 nouveaux cas diagnostiqués en 2012 (Aou 04), (Ter 87). C'est le premier cancer de la femme et prend des proportions épidémiques renseignant sur les obligations en matière de prise en charge tant sur le plan préventif que curatif.

Son incidence connaît une progression exponentielle alarmante depuis environ 25 ans, à partir des années 1990 (Dgs 25).

La mammographie est non seulement pratiquée dans les campagnes de dépistage du cancer du sein, mais aussi pour le diagnostic et la localisation lors d'intervention chirurgicale (Chk 20).

La lecture des mammographies est un travail très exigeant pour les radiologues. Leurs jugements dépendent de la formation, de l'expérience et des critères subjectifs.

Les systèmes de diagnostic assistés par ordinateur (CAD) peuvent aider les radiologues à interpréter les mammographies. La combinaison du système CAD et des connaissances des experts améliorerait grandement la précision de détection.

1.1 Structure du sein

Le sein est une masse de tissu glandulaire, grasseux et conjonctif. Il est constitué de :

- lobules – glandes qui produisent le lait ;
- canaux – tubes qui transportent le lait des lobules au mamelon ;
- tissus grasseux et conjonctif – ils entourent et protègent les canaux et les lobules et donnent sa forme au sein ;
- aréole – surface ronde, rosée ou brunâtre qui entoure le mamelon ;
- mamelon – région située au centre de l'aréole d'où le lait est expulsé.

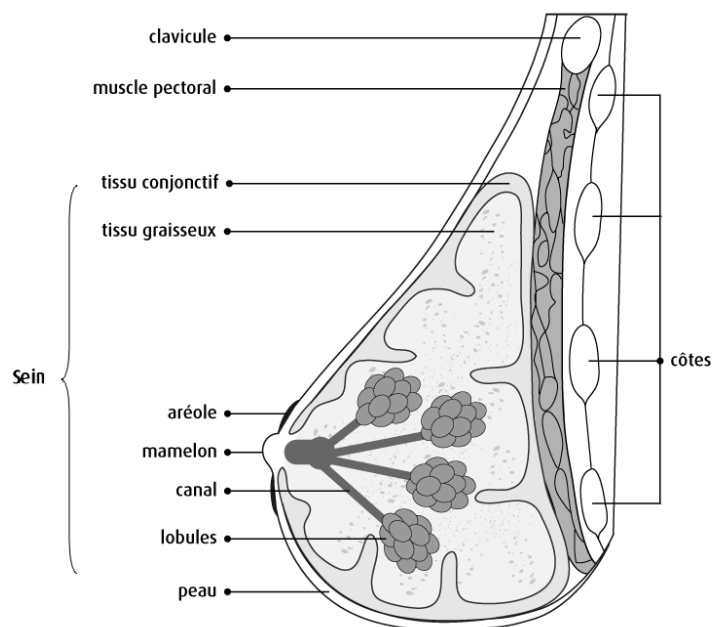


Figure 1.1 Structure du sein

1.2 Le cancer du sein

Le cancer du sein représente un véritable problème de santé publique. Dans le Monde, 540 000 cas de cancer du sein apparaissent chaque année et près de 300 000 femmes en meurent, généralement à cause d'un dépistage tardif (Aol 05).

Le cancer du sein est une tumeur maligne qui touche la glande mammaire. Les cellules malignes à l'origine de la tumeur se multiplient jusqu'à créer une tumeur qui s'attaque aux tissus sains. La tumeur peut aussi métastaser, c'est-à-dire, propager des cellules cancéreuses ailleurs dans l'organisme. Une tumeur est une prolifération anormale de cellules (Sar 74).

Une tumeur est considérée comme bénigne lorsque les cellules anormales restent à l'endroit où elles se sont développées. Une tumeur bénigne a un contour régulier. Ce ne sont pas des cancers. Elles n'ont souvent pas de conséquences graves pour la santé, mais dans le cas où la tumeur devient très grosse, elle peut engendrer une compression de l'organe où elle se situe (Sar 74).

Les cellules qui constituent les tumeurs malignes sont celles qui se multiplient le plus vite. Elles ont la particularité de détruire les cellules des tissus et organes dans lesquels elles se trouvent. Ces cellules sont capables de créer de nouveaux vaisseaux sanguins qui permettent à la tumeur de se nourrir et de se propager dans l'organisme. Ce sont les métastases. Une tumeur maligne a souvent un contour irrégulier (Sar 74).

1.3 La mammographie

La mammographie est une technique de radiographie, particulièrement adaptée aux seins de la femme.

L'appareil dédié à la réalisation d'une mammographie est le mammographe (figure 1.2) (Chk 20). Cet appareil se compose d'un tube radio-gène générateur de rayons X de faible énergie et d'un système de compression du sein.

Dans un premier temps, les deux seins sont comprimés à tour de rôle. Cette compression permet l'étalement des tissus mammaires ce qui facilite la visualisation des structures du sein et la réduction de la dose de rayons X délivrée. En deuxième temps, les deux seins sont exposés à une faible dose de rayons X. On obtient, alors, une projection du sein sur un détecteur plan. La radiographie est réalisée sur des films argentiques ou sur des systèmes de radiologie digitale de haute qualité (Chk 20).

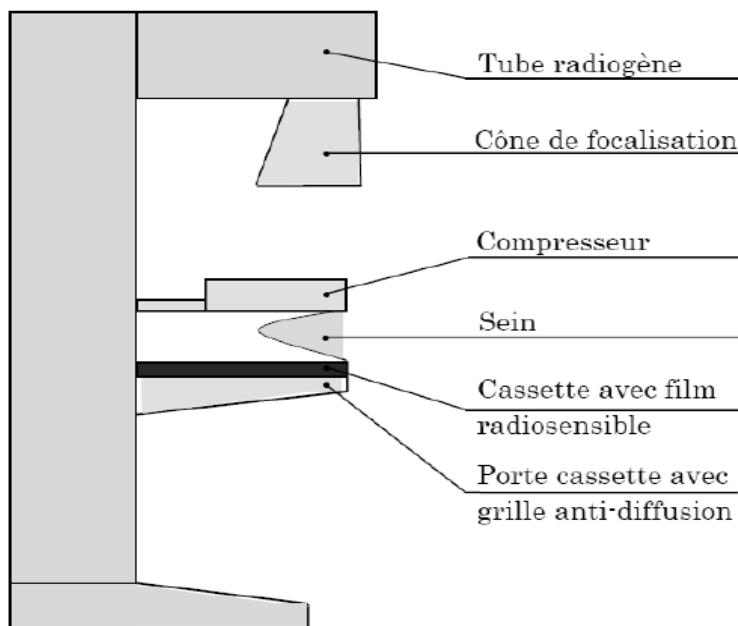


Figure 1.2 Les composants d'un appareil de mammographie (Chk 20)

La mammographie est le résultat d'atténuation d'un faisceau de rayons X traversant les différents tissus mammaires.

L'atténuation de ce faisceau dépend essentiellement de la composition des tissus traversés. En effet, la graisse est considérée comme une zone radio transparente vu qu'elle a une densité physique très légère. De ce fait, elle apparaît très sombre sur une mammographie (Chk 20).

En revanche, les zones radio opaques apparaissent claires et correspondent au tissu fibro-glandulaire et au calcium qui est le composant essentiel des lésions mammaires. Pour les matières prédominantes dans le sein, nous obtenons le tableau 1.1 (Chk 20).

Tableau. 1.1 Atténuation radiologique des composants mammaires (Chk 20).

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

Les mammographies sont souvent marquées avec une étiquette en forme rectangulaire. Contenant des informations concernant l'examen effectué et des labels opaques présentant la latéralité de la mammographie (R/L, Right/Left) ainsi que des indicateurs de projection, par exemple l'incidence oblique. Elle est également intitulée *médio-latérale oblique* (MLO) par référence à la direction du faisceau à rayons X depuis sa source jusqu'au détecteur. Une mammographie peut contenir des artefacts Figure 1.3.

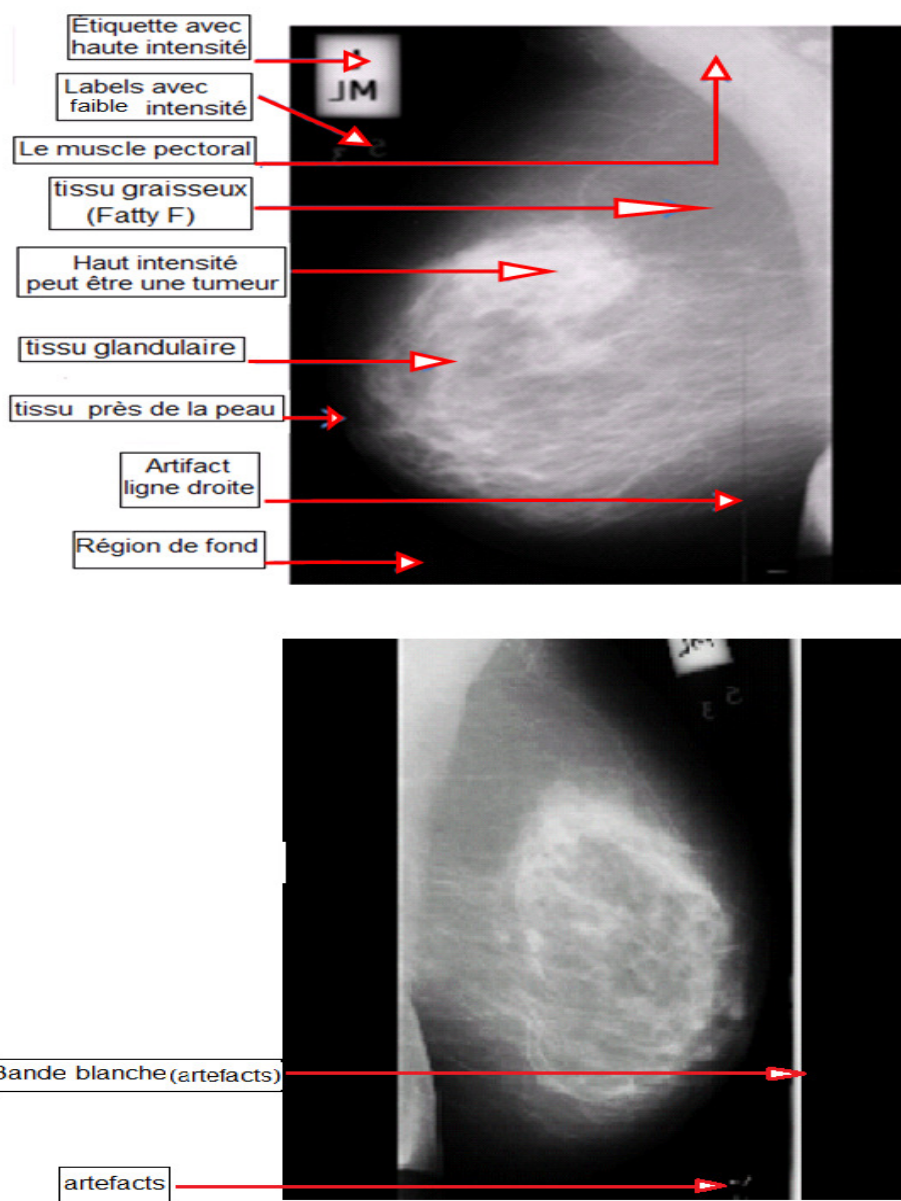


Figure 1.3 Les éléments d'une mammographie

1.3.1 Les incidences en mammographie

La mammographie est généralement prise sous différentes directions appelées incidences. Une bonne incidence a pour but de visualiser le maximum de tissu mammaire en l'étalant le plus possible sur la plaque radiographique.

Selon la partie du sein à laquelle s'intéresse l'examen, différentes incidences sont utilisées. Les incidences les plus fréquemment utilisées sont l'incidence de face appelée aussi Cranio Caudale (CC), l'incidence oblique externe nommée Médio Latérale Oblique (MLO) et l'incidence de profil.

Dans le cas de l'incidence de face (CC), le support de la cassette (le détecteur) ainsi que le tube radio-gène sont horizontaux. Cette incidence explore bien la région centrale et la région interne du sein.

Dans le cas de l'incidence oblique externe, le détecteur est incliné de 40° à 60° en fonction de la morphologie de la patiente de telle façon que l'orientation des fibres du muscle pectoral soit parallèle au détecteur.

Dans le cas de l'incidence de profil, le porte-cassette est vertical et placé entre les deux seins (profil interne). Cette incidence explore la région centrale et interne du sein. La figure 1.3 explique le positionnement du tube radio gène et du détecteur pour les différentes incidences définies ci-dessus (Chk 20).

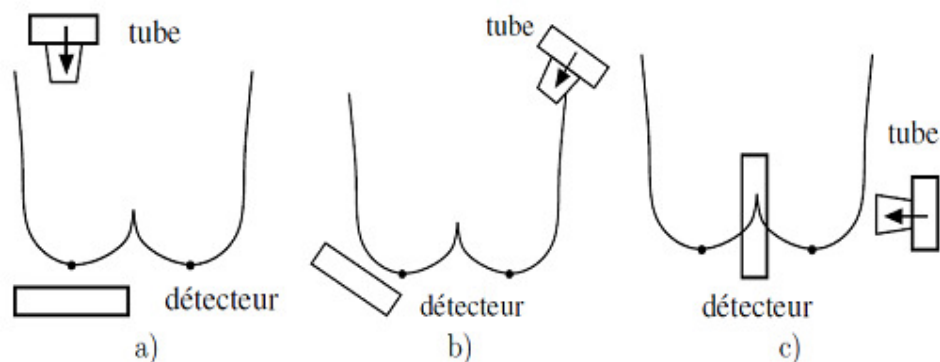


Figure 1.4 Exemples d'incidences en mammographie : a) Cranio Caudale (CC), b) Médio Latérale Oblique (MLO) et c) Profil interne

1.3.2 Mammographie numérique

La technique pour avoir une forme numérique d'un film radiographique consiste à utiliser une caméra ou bien un numériseur à balayage (scanner). Dans le premier cas, le signal vidéo acquis par la caméra est numérisé. Dans le deuxième cas, un faisceau laser de haute densité mesure la densité optique de chaque point radiographique (Ckh 21).

Le plus grand avantage de la mammographie numérique par rapport à la mammographie traditionnelle est le temps requis pour l'obtention d'une mammographie.

La mammographie numérique évitera le temps mis pour développer le film et permet de s'assurer si l'exposition du cliché est correcte ou si le sein est parfaitement positionné.

La mammographie numérique rend l'image numérique transportable ; les clichés de l'examen d'une patiente peuvent être envoyés, via réseau pour un diagnostic à distance par d'autres spécialistes. En plus, elle offre également la possibilité d'organiser une chaîne complète de gestion de l'information des données médicales. Quant à l'archivage des images, les grands locaux réservés à l'archivage des films dans les sous-sols des hôpitaux sont remplacés par des systèmes de stockage d'images numériques (disques durs). Pour un film, une fois révélée chimiquement, il sera impossible de modifier les contrastes locaux autour d'une zone d'intérêt par exemple. Il est possible de numériser le film, mais l'information resterait limitée. En numérique, on peut accéder aux données brutes affichées sur l'image numérique (Eda 32).

1.4 Le dépistage du cancer du sein

Selon l'Organisation Mondiale de la Santé (OMS): « le dépistage consiste à identifier de manière présomptive, à l'aide de tests appliqués de façon systématique et standardisée, les sujets atteints d'une maladie ou d'une anomalie passée jusque là inaperçue.

Les tests de dépistage doivent permettre de faire le partage entre les personnes apparemment en bonne santé mais qui sont probablement atteintes d'une maladie ou d'une anomalie donnée et celles qui en sont probablement exemptes» (Cor 22).

Le dépistage précoce du cancer du sein accroît la probabilité de guérison complète et permet dans certains cas un traitement moins intrusif.

Le test de dépistage du cancer du sein peut être un examen clinique ou radiologique appliqué à une population de femmes cible (Elk 33).

Le programme de dépistage organisé vise à détecter des anomalies, à un stade précoce. La réduction du risque de décès par cancer du sein est d'environ 30 % dans la population ayant effectivement participé au dépistage (Bra 13).

1.5 Les systèmes de Détection / Diagnostic Assistées par Ordinateur (CAD)

Un CAD (Computer Aided Detection/Diagnosis) est un système matériel et logiciel qui analyse des clichés médicaux et qui aide le spécialiste dans son travail de détection. Plus concrètement, il s'agit souvent d'un détecteur et d'un classifieur (KNN, arbres de décision, ...)(Hdc 39).

Les premiers travaux concernant les systèmes automatiques de traitement des mammographies avaient pour but d'offrir une seconde interprétation aux radiologues afin de les aider à détecter/diagnostiquer à un stade précoce.

La performance des radiologues dans le dépistage du cancer du sein serait améliorée par une aide des systèmes CAD.

1.5 .1 Détection assistée par ordinateur

Le taux de fausses interprétations négatives produites sur la mammographie classique est relativement élevé (entre 10 % et 30 %) (Paj 66). Pour cette raison, la détection assistée par ordinateur est devenue de plus en plus utilisée dans le dépistage par mammographie, elle améliore la sensibilité des mammographies. L'image de la mammographie sur film est convertie en image numérique par un numériseur. Cette image numérique est alors analysé par un ordinateur et affiché à l'écran. La détection assistée par ordinateur a recours à un logiciel pour parcourir l'image et fournir des messages (Ici 43).

Le détecteur réalise la détection d'anomalie en utilisant des informations déduite de la mammographie. Le classifieur a pour rôle de faire un tri.

Le CAD (Computer-aided Détection) peut-être considéré comme un complément précieux du médecin radiologue. Les médecins radiologues experts doivent porter leur choix sur des CAD composés d'algorithmes les plus performants notamment dans la détection des masses (Sym 85). Vu les avantages potentiels de ces technologies, depuis 1998, quatre systèmes de

détection assistée par ordinateur ont obtenu l'approbation de la FDA (Food and Drug Administration) aux Etats Unis. Le premier système conçu est le R2 Image Checker . En janvier 2002, deux ont été approuvés : le CADx Second Look et le iCAD MammoReader. Ces deux compagnies ont fini par fusionner et ils ont gardé l'appellation de Second Look. Ensuite le système Kodak DirectView CR Mammography Feature à été commercialisé (Chk 20).

1.5.2 Diagnostic Assisté par Ordinateur

Diagnostic vient du grec "diagnosis": connaissance. A l'origine utilisé dans le domaine médical, ce terme signifie: identification d'une maladie par ses symptômes. Plus généralement, on peut le définir comme: un jugement porté sur une situation, sur un état. Mais l'interprétation du mot diagnostic possède de nombreuses significations suivant les interlocuteurs auxquels on s'adresse: le diagnostic financier d'une société concerne la vérification de son état financier, le diagnostic médical s'intéresse à déterminer une maladie pour identifier les causes, le diagnostic industriel, quand à lui, a pour objet de trouver la cause d'une défaillance (Zwi 95).

Le domaine médical a rapidement adopté des outils d'aide au diagnostic. La mise en application d'un système d'aide au diagnostic en utilisant l'image médicale passe par plusieurs étapes: une étape de prétraitement, une étape de classification / identification, et une dernière étape pour la prise de décision finale par le personnel médical.

La classification des lésions mammaires par un radiologue est une classification humaine subjective qui peut facilement classer la même lésion de deux manières différentes. En revanche, un système automatique de Diagnostic Assisté par Ordinateur, basé sur les descripteurs (caractéristiques) de la lésion donne toujours les mêmes résultats pour le même cliché et peut donc servir de référence. Un système automatique de Diagnostic Assisté par Ordinateur sert à donner un second avis au radiologue (Chk 20).

Un tel système débute par un prétraitement afin d'obtenir une image plus lisible et moins bruitée. Ce prétraitement facilite la tâche des étapes suivante.

Une étape d'extraction des descripteurs mathématiques permet la classification de type de tissu et la détection de la présence d'anomalie. Ensuite l'étape finale consiste à prendre une décision.

La classification de la densité des tissus mammaire est une étape importante.

Cette information sera exploitée pour le diagnostic des tissus mammaires. Le concept de systèmes dédiés à l'aide au diagnostic (CAD = Computed Assisted Diagnosis) émerge depuis une quinzaine d'années. Plusieurs équipes se sont intéressées à des systèmes informatiques automatisés (Bal 06) (Doi 27) (Gig 36) pour l'aide à la détection et au diagnostic des lésions radiologiques en mammographie.

Ils sont destinés à une lecture informatique qui vise à aider mais pas à se substituer au radiologue (Bal 06).

L'analyse numérique des clichés de mammographie peut alors servir de deuxième lecture informatique pour le radiologue (Doi 26).

Plusieurs études récentes ont montré que ces systèmes CAD pouvaient augmenter de 15 à 20 % le taux de cancers détectés (Fre 35).

1.6 Système d'aide au diagnostic /détection pour mammographie

Le système d'aide au diagnostic /détection pour mammographie proposé est composé de plusieurs étapes décrites dans la figure 1.5 :

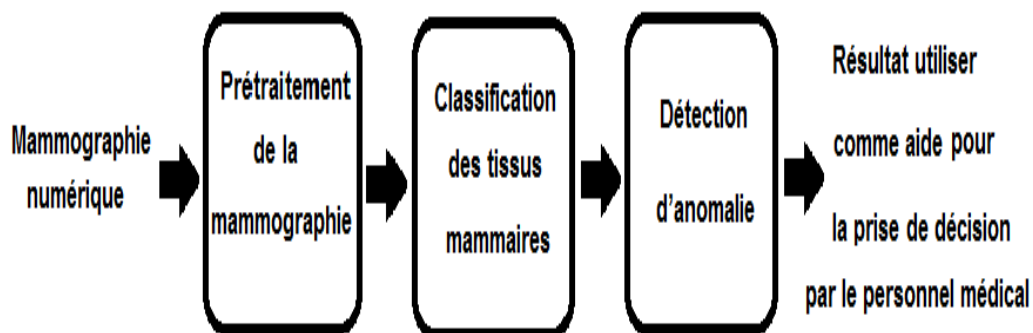


Figure 1.5 Les étapes du système d'aide au diagnostic / détection pour mammographie

L'ensemble des étapes représenté dans la figure 1.5 constitue un système d'aide au diagnostic pour mammographie, de plus la dernière étape permet la détection de la présence d'anomalie mammaire.

1.6.1 Prétraitement de la mammographie

La mammographie comporte des objets non-utiles au diagnostic médical. Notre objectif est l'extraction de la région d'intérêt d'une mammographie, la segmentation de cette région s'effectue en se basant sur des techniques de traitement d'images. On doit isoler cette zone du reste de l'image afin d'écartier l'influence de ces éléments sur les différentes étapes d'analyse ultérieures. Plusieurs études ont été menées pour effectuer l'extraction de cette zone d'image le tableau suivant est une représentation non exhaustive des travaux de prétraitement de mammographie.

Tableau. 1.2 Les méthodes de prétraitement de la mammographie

| Année | Description de la méthode | Base de données utilisé |
|-------|--|-------------------------|
| 2010 | Cet article présente une Amélioration de l'algorithme d'égalisation d'histogramme adaptative à contraste limité (CLAHE) pour supprimer le de bruit et les inhomogénéités d'intensité en arrière-plan tout en conservant les variations naturelles du niveau de gris des mammographies (Rah 71). | DDSM |
| 2011 | Le travail effectué dans cet article est présenté en trois phases. La première phase est de supprimer les parties d'image excessive qui sont dans les deux côtés de l'image. La deuxième phase est l'identification de la direction du sein et mettre toutes les images dans une seule direction; par l'utilisation du niveau de gris des deux moitiés de l'image. La troisième phase est la segmentation de la région du sein (Deh 24). | MIAS |
| 2012 | Ils proposent trois étapes distinctes. L'étape initiale implique l'amélioration du contraste. Ensuite, définir le rectangle pour isoler le muscle pectoral et finalement supprimer le muscle pectoral (Mai 49). | MIAS |
| 2013 | Dans cet article, ils ont proposé l'application de la normalisation, le filtrage et une technique d'étiquetage en composantes connexes pour l'extraction des régions mammaires et l'enlèvement du muscle pectoral (Bos 11). | MIAS |

| | | |
|------------------|--|---------------|
| 2014 | <p>L'algorithme proposer pour le prétraitement des mammographies comporte deux étapes: La première est l'amélioration de l'image. Huit techniques pour l'amélioration des mammographies ont été mises en œuvre : égalisation d'histogramme adaptative, filtre moyennneur, modification du contraste par étirement de l'histogramme, l'égalisation d'histogramme, transformation logarithemique, filtre médian, amélioration d'image basée sur la logique floue, Filtre Adaptatif 2-D</p> <p>- La deuxième étape consiste à extraire la région d'intérêt du reste de l'image (arrière-plan). A cet effet, une méthode basée sur le seuillage automatique a été utilisée (Min 54).</p> | Mini– MIAS |
| 2015 | L'article présente une technique d'amélioration du contraste par l'égalisation de l'histogramme (Aki 02). | MIAS |
| 2016 | <p>L'essai proposé montre que le filtre de Wiener et la technique d'égalisation d'histogramme adaptatif avec limitation de contraste (CLAHE), aident efficacement à améliorer la qualité de l'image.</p> <p>Le fond indésirable et le muscle pectoral sont éliminés en utilisant le seuillage et la segmentation d'image par croissance de région (Mak 50).</p> | mini- MIAS |
| Méthode proposée | L'approche proposée consiste à l'identification de la direction du sein et mettre toutes les images dans une seule direction et à appliquer la binairisation et la labellisation pour supprimer les d'artéfacts de la mammographie. | mini- MIAS |

1.6.2 Classification des tissus mammaires

Une des caractéristiques importantes de l'anatomie et de la physiologie du sein est la densité du tissu mammaire. La densité est importante pour deux raisons principales : premièrement, la densité mammaire accrue est associée à une diminution de la sensibilité de la mammographie pour la détection du cancer du sein (Sch 76). Deuxièmement, la densité du sein est l'un des plus importants facteurs de risque connus pour le cancer du sein (Pre 70), (Byd 14). Pour toutes ces raisons, le classement automatique de la densité des tissus est un processus important dans le diagnostic. Différentes méthodes de classification ont été appliquées, pour classer les tissus mammaires, le tableau suivant est une représentation non exhaustive de ces travaux :

Tableau. 1.3 Les méthodes de classification de la densité des tissus mammaires

| Année | La méthode utilisée | Base de données Utilisé |
|-------|--|-------------------------|
| 2005 | Cette étude présente deux méthodes : l'algorithme des k plus proches voisins et l'arbre de décision (Oli 62). | MIAS |
| 2006 | une technique de classification de la densité mammaire en utilisant un histogramme multi-résolution a été proposée (Muh 56). | MIAS |

| | | |
|------------------|--|---|
| 2011 | <p>Plusieurs méthodes ont été employées pour la classification de la densité mammaire :</p> <ul style="list-style-type: none"> - Les arbres de classification et de régression - L'algorithme des k plus proches voisins - Machine à vecteurs supports –SVM-(Tzi 93). | Mini MIAS |
| 2012 | <p>Dans cet article, ils proposent l'utilisation du classifieur bayésien naïf et l'algorithme des k plus proches voisins (Mus 57).</p> | Mini-MIAS et KBD-ER digital mammography |
| 2013 | <p>Cet article présente une méthode fondée sur l'estimation de la densité mammaire. Le profil global de la densité du tissu mammaire est représenté à l'aide d'une carte topographique (Che 19).</p> | MIAS et DDSM |
| Méthode proposée | <p>Dans l'approche proposée, les valeurs déduites de l'histogramme seront traitées par des réseaux de neurones artificiels (RNA) pour la classification des tissus mammaires.</p> | Mini –MIAS |

1.6.3 Détection d'anomalie dans le tissu mammaire

Nous allons étudier la problématique de la détection de la présence ou non d'anomalie dans une mammographie, plusieurs études ont été menées pour la détection des anomalies (tableau 1.4).

Tableau. 1.4 Les méthodes de détection d'anomalie des tissus mammaires

| Année | Description de la méthode | Base de données Utilisé |
|-------|--|--|
| 2010 | Un système automatique pour l'aide au diagnostic est appliqué à la détection et la localisation des zones suspectes dans les mammographies. L'algorithme proposé contient trois étapes essentielles; dans la première étape une méthode fondée sur les contours actifs et la carte de chanfrein est utilisée pour la détection des contours de la glande-mammaire. Dans la deuxième étape une méthode de segmentation permet de segmenter le muscle pectoral, le but des deux étapes précédentes est de limiter la région d'intérêt, dans la troisième étape un contour actif génétique est utilisé pour localiser les régions de masses suspectes (Fer 34). | MIAS |
| 2012 | Cet article présente une approche basée sur une hybridation de la méthode de K-means et la méthode de la recherche harmonique (<i>Harmony Search</i> – HS) afin de détecter les tumeurs (Ral 72). | Des images mammographiques de taille 256*256 |

| | | |
|-----------------|--|------------|
| 2015 | <p>L'étude propose une technique de diagnostic assistée par ordinateur en utilisant l'hybridation de l'algorithme génétique par l'algorithme d'optimisation par essaim de particules GA-PSO, pour la sélection des caractéristiques de texture.</p> <p>Le classificateur KNN est utilisé pour classer les tissus mammaires normaux et anormaux (Sun 78).</p> | Mini-MIAS |
| 2016 | <p>Dans cet article un filtre de Gabor sert à calculer des caractéristiques qui seront utilisés pour sélectionner les pixels les plus pertinents. Finalement, une classification des tissus mammaires en deux classes anormale / normale est réaliser en utilisant une Machine à vecteurs de support (SVM) (Tor 90).</p> | Mini-MIAS |
| Méthode proposé | <p>L'étude présente propose une interprétation des mammographies par un système à base de réseau de neurones artificiels capables de détecter la présence ou non d'anomalie mammaire.</p> | Mini -MIAS |

1.7 Base de données des mammographies

Plusieurs études ont utilisé la base mini-MIAS ceci nous offre une large possibilité de comparaison et les images issue de la base min-MIAS occupe moins d'espace mémoire vue leur taille en pixel ce qui représente un avantage.

La base mini-MIAS (Mammographic Image Analysis Society) (Suc 82) réalisée par un groupe de chercheurs britanniques qui s'intéressent à l'interprétation de la mammographie.

La base mini-MIAS nous offre des mammographies en niveaux de gris codées sur 8 bits (noir \equiv 0, blanc \equiv 255) avec une résolution de 200 μ m par pixel, la taille de chaque image est de 1024x1024 pixels. Dans cette base d'images, il y a seulement trois catégories de tissus mammaires (Suc 82):

- grassex (Fatty F) ;
- grassex-glandulaire (Fatty glandular G) ;
- dense (Dense D).

Cependant il existe d'autre base de données tel que : La base DDSM (l'université de la Floride du Sud).

La classification actuellement utilisée dans la base DDSM provient du lexique BI-RADS de l'American College of Radiology (ACR) et consiste à apprécier visuellement le pourcentage de tissu radio-opaque. Quatre types sont ainsi décrits :

- BI-RADS I : Sein presque entièrement graisseux (< 25 % de glande mammaire) ;
- BI-RADS II : présence d'opacités fibro-glandulaires éparses (approximativement 25 à 50 % de glande mammaire) ;
- BI-RADS III : tissu mammaire dense et hétérogène, ce qui peut compliquer la détection de petites masses (approximativement 51 à 75 % de glande mammaire) ;
- BI-RADS IV : tissu mammaire très dense, ce qui peut diminuer la sensibilité de la mammographie (> 75 % de glande mammaire) (Tar 86).

Chapitre 2

Les réseaux de neurones artificiels

1 Réseaux de neurones artificiels

Depuis une dizaine d'années, l'utilisation des réseaux de neurones artificiels (RNA) s'est développée dans de nombreuses disciplines (sciences économiques, écologie et médecine...). Ils sont notamment appliqués pour résoudre des problèmes de classification, de prédiction, de reconnaissance des formes (Dre 28).

Dans le cadre du traitement des données, les RNA constituent une méthode d'approximation de systèmes complexes, particulièrement utile lorsque ces systèmes sont difficiles à modéliser (Scm 77).

Les réseaux de neurones artificiels sont des réseaux fortement connectés de processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit. Toute structure hiérarchique de réseaux est évidemment un réseau (Tou 91).

En général le réseau de neurones possède :

- Les neurones d'entrée qui reçoivent des données.
- Les neurones de sortie qui envoient des données par la sortie du système.
- Les neurones cachés dont les signaux d'entrée et de sortie demeurent dans le système.

Les RNA sont inspirés de la biologie et représentent un modèle mathématique du fonctionnement du neurone biologique (Jod 45).

2 Eléments biologiques du neurone

Les neurones constituent les éléments de base du système nerveux. Il en existe une grande variété remplissant diverses fonctions. Nous nous intéresserons à la cellule nerveuse.

La partie principale d'une cellule nerveuse, ou neurone, est appelé soma. La membrane cellulaire qui la recouvre forme des ramifications appelées dendrites. Pour communiquer, un neurone envoie des signaux vers d'autres

neurones à travers une fibre appelée l'axone. Enfin, c'est à travers les synapses, que le signal provenant du soma d'un neurone arrive aux dendrites, ou au soma d'autres neurones (Mej 53). La figure suivante montre une vue de deux neurones biologiques connectés.

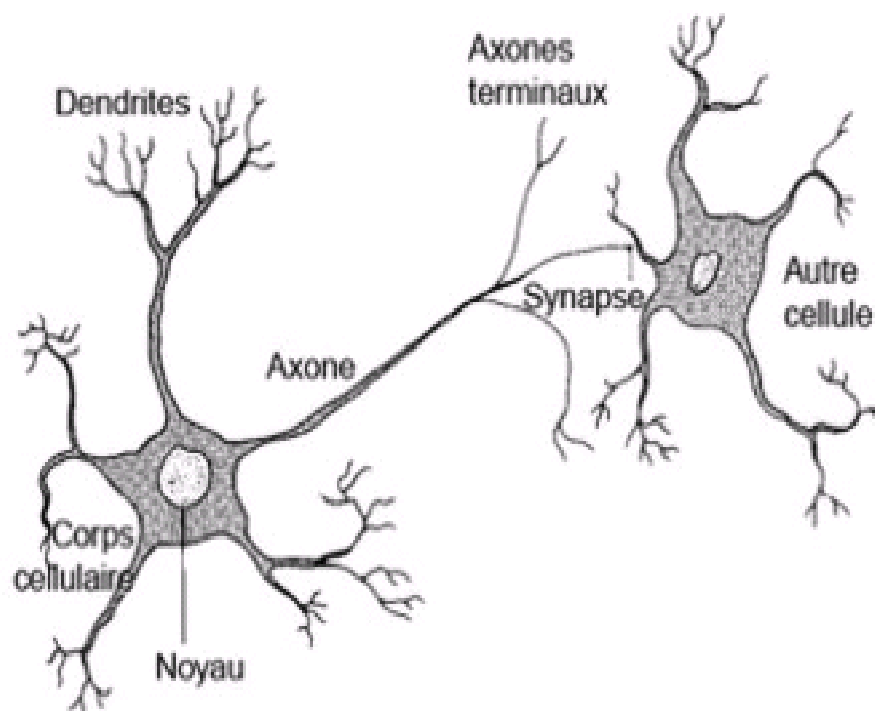


Figure 2.1 Deux neurones connectés

Un processus électrochimique complexe permet la transmission de signaux dans le système nerveux. Le signal standard est un potentiel d'activation électrique d'une durée de 1 à 5 ms et d'une amplitude d'environ 100 mV, qui se propage le long de l'axone. Arrivé à l'extrémité de l'axone, le potentiel d'action engendre une excitation de nature chimique sous forme d'un neurotransmetteur diffusé dans la fente synaptique.

Lorsque le neurotransmetteur atteint le neurone suivant, il génère un potentiel synaptique. Ce dernier provoque une excitation ou une inhibition du neurone. Le neurone reçoit de tels potentiels synaptiques de toutes les

cellules nerveuses avec les quelles il est en liaison. Il effectue une sommation si le résultat obtenu est supérieur à un seuil, le soma génère un potentiel d'action. Dans les détails il y a beaucoup de variations de ce concept. En fait, la description des processus à l'intérieur d'un seul neurone est un domaine de recherches très actives. Pour la suite, le simple concept d'une sommation d'entrées suivie d'un effet seuil est suffisant (Wul 94).

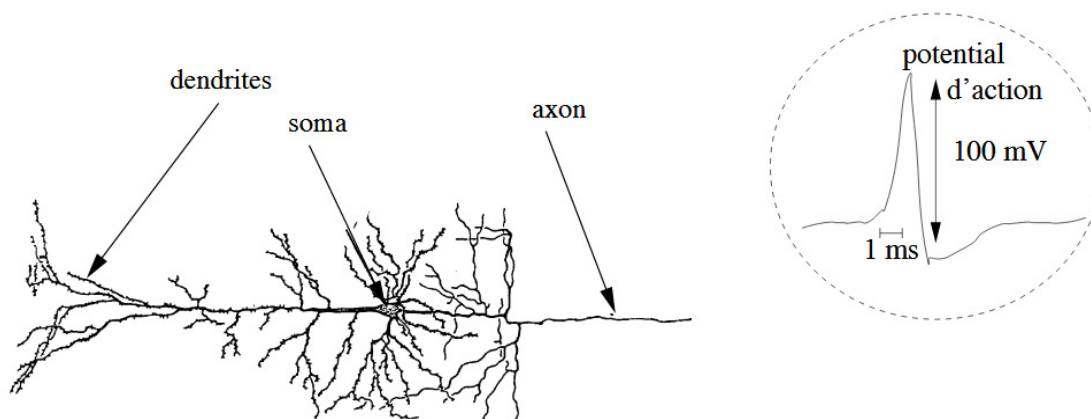


Figure 2.2 Représente un neurone biologique d'après un dessin de Ramon Y Cajal, pionnier des neurosciences -1900; adroite le schéma d'un potentiel d'action, le signal élémentaire de communication entre neurones (Wul 94).

3 Du neurone biologique au neurone formel

Le neurone formel est une modélisation mathématique qui reprend les principes de fonctionnement du neurone biologique. Le premier neurone formel est proposé en 1943 par McCulloch et Pitts.

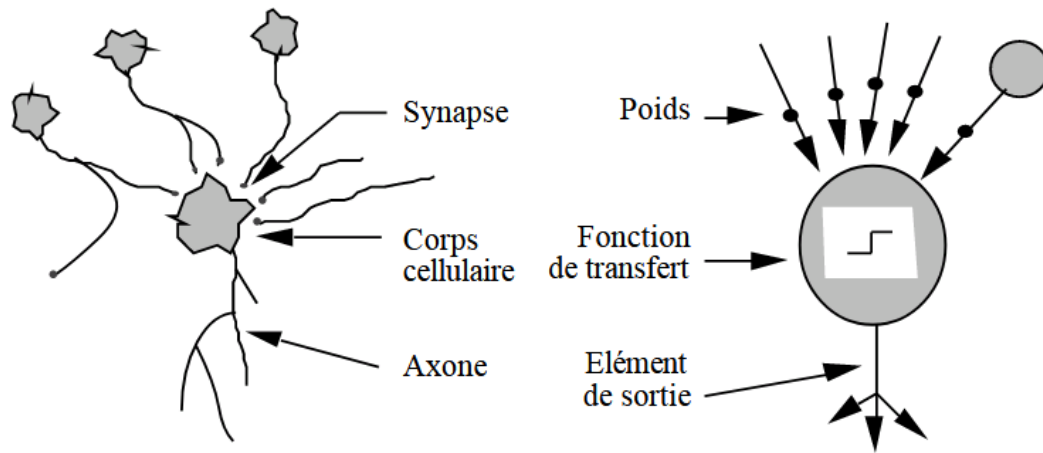


Figure 2.3 Mise en correspondance neurone biologique / neurone artificiel (Tou 91)

Un neurone formel est généralement constitué de plusieurs entrées et d'une sortie par analogie aux dendrites et à l'axone. Les grands principes y sont repris, notamment la sommation des entrées. A chacune de ces entrées, est associé un poids qui correspond aux actions excitatrices et inhibitrices des synapses.

Une phase d'apprentissage est alors utilisée pour ajuster ces coefficients. Soit un neurone k , on lui présente en entrée les signaux $x_1, x_2 \dots x_p$.

Ils sont respectivement multipliés par des coefficients, appelés poids synaptiques, $w_{k1}, w_{k2} \dots w_{kp}$. On définit alors la somme pondérée v_k .

$$v_k = \sum_{j=1}^p w_{kj} \cdot x_j \quad (2,1)$$

Le neurone commence à devenir actif lorsque la sommation pondérée dépasse un seuil (ou biais) noté b_k . Il est additionné à cette grandeur, pour ensuite être transformé par la fonction d'activation notée « φ »:

$$y_k = \varphi \left(b_k + \sum_{j=1}^R w_{k,j} \cdot x_j \right) \quad (2,2)$$

Pour $sp = \left(b_k + \sum_{j=1}^R w_{k,j} \cdot x_j \right)$, On à : $y_k = \varphi(sp)$ (2,3)

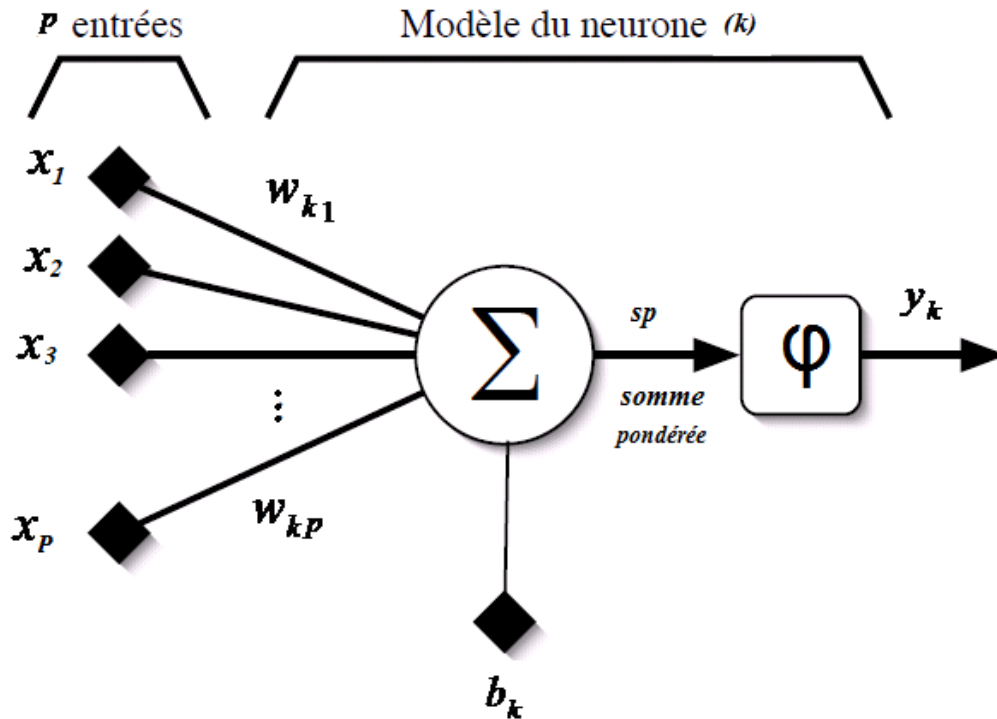


Figure 2.4 Modèle d'un neurone artificiel (Par 67).









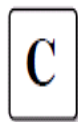
On peut faire un parallèle entre ce modèle mathématique et le neurone biologique. Il faut bien comprendre que notre neurone artificiel est un modèle pragmatique qui, comme nous le verrons plus loin, nous permettra d'accomplir des tâches intéressantes (Par 67).

3.1 Fonctions de transfert

Le choix de la fonction d'activation est important pour obtenir un modèle utile en pratique, car il influe sur les propriétés du neurone formel (Sav 75). Différentes fonctions de transfert pouvant être utilisées comme fonction

d'activation du neurone sont énumérés au tableau 2.1 (Par 67). Les plus utilisées sont les fonctions «linéaire » et «sigmoïde».

Tableau. 2.1 Fonction de transfert $y_k = \varphi(sp)$ (Par 67).

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

4 Le perceptron

Le perceptron est la forme la plus simple du réseau de neurones, et permet de classifier correctement des objets appartenant à deux classes linéairement séparables. Le perceptron est formé d'un seul neurone qui possède un seuil ainsi qu'un vecteur de poids synaptiques ajustables, tout comme le modèle de neurone de McCulloch & Pitts (Moh 58).

5 Perceptron multicouches

La mise en cascade de perceptrons conduit à ce qu'on appelle le perceptron multicouches. Lorsque le vecteur de caractéristiques d'un objet est présenté à l'entrée du réseau, il est communiqué à tous les neurones de la première couche. Les sorties des neurones de cette couche sont alors communiquées aux neurones de la couche suivante, et ainsi de suite. La dernière couche du réseau est appelée couche de sortie, les autres étant désignées sous le terme de couches cachées car les valeurs de sortie de leurs neurones ne sont pas accessibles de l'extérieur (Moh 58).

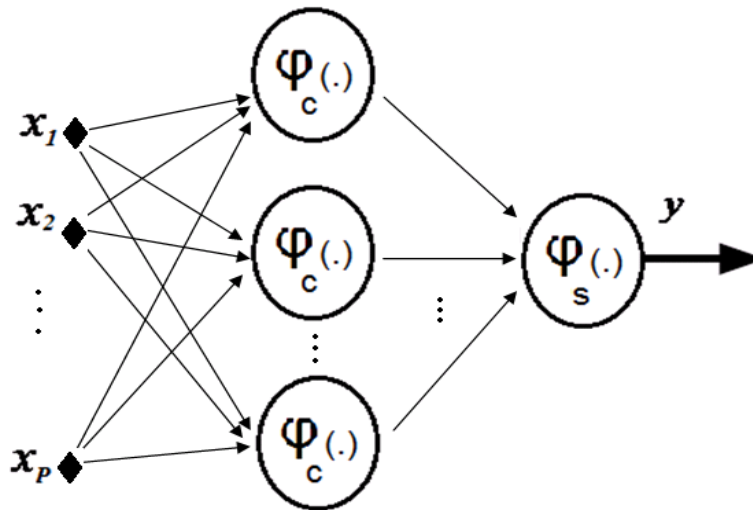


Figure 2.5 Perceptron multicouches (MLP) à une couche cachée (φ_c fonction de la couche cachée

φ_s fonction de la couche de sortie)

6 Architecture des réseaux de neurones

On distingue deux grands types d'architectures de réseaux de neurones : Les réseaux de neurones non bouclés et les réseaux de neurones bouclés (Drf 30).

6.1 Les réseaux de neurones non bouclés

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonctions algébriques de ses entrées, par composition des fonctions réalisées par chacun de ses neurones (Drf 30). Un réseau de neurones non bouclé est représenté graphiquement par un ensemble de neurones "connectés" entre eux, l'information circulant des entrées vers les sorties sans "retour en arrière". Dans la très grande majorité des applications, les réseaux de neurones sont des formules algébriques dont les valeurs numériques sont calculées par des programmes d'ordinateurs, non des objets physiques (circuits électroniques spécialisés) (Drf 30). La figure suivante représente un réseau de neurones non bouclé.

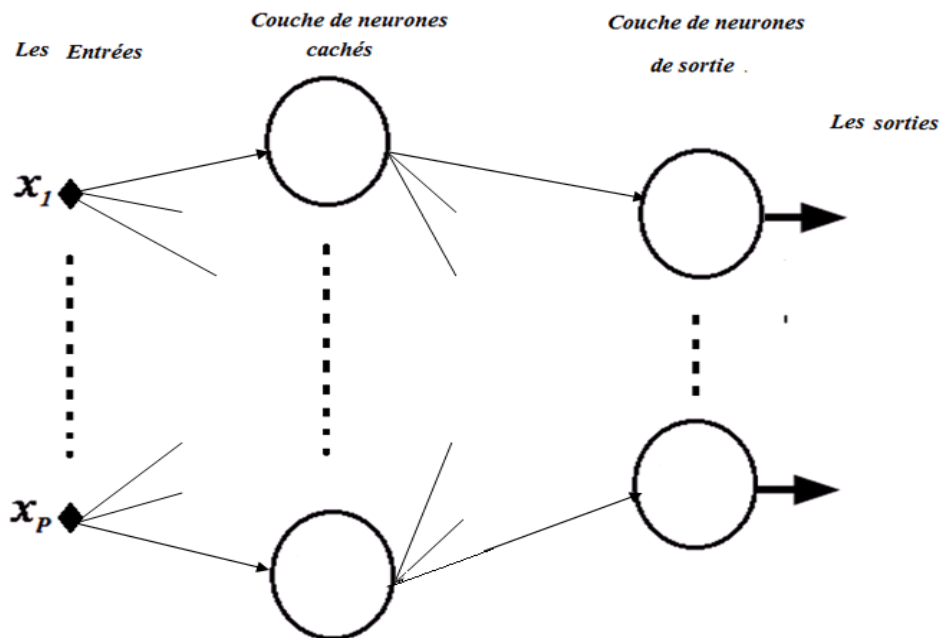


Figure 2.6 Un réseau de neurones non bouclé - Perceptron multicouche - (Drf 30).

Les réseaux de neurones non bouclés sont des objets statiques : si les entrées sont indépendantes du temps, les sorties le sont également. Ils sont utilisés principalement pour effectuer des tâches d'approximation de fonction non linéaire, de classification ou de modélisation de processus statiques non linéaires (Drf 30).

6.2 Les réseaux de neurones bouclés

Les réseaux de neurones bouclés peuvent avoir une topologie de connexions quelconque, comprenant notamment des boucles qui ramènent aux entrées la valeur d'une ou plusieurs sorties. Il faut évidemment qu'à toute boucle soit associé un retard : un réseau de neurones bouclé est donc un système dynamique (Drf 30).

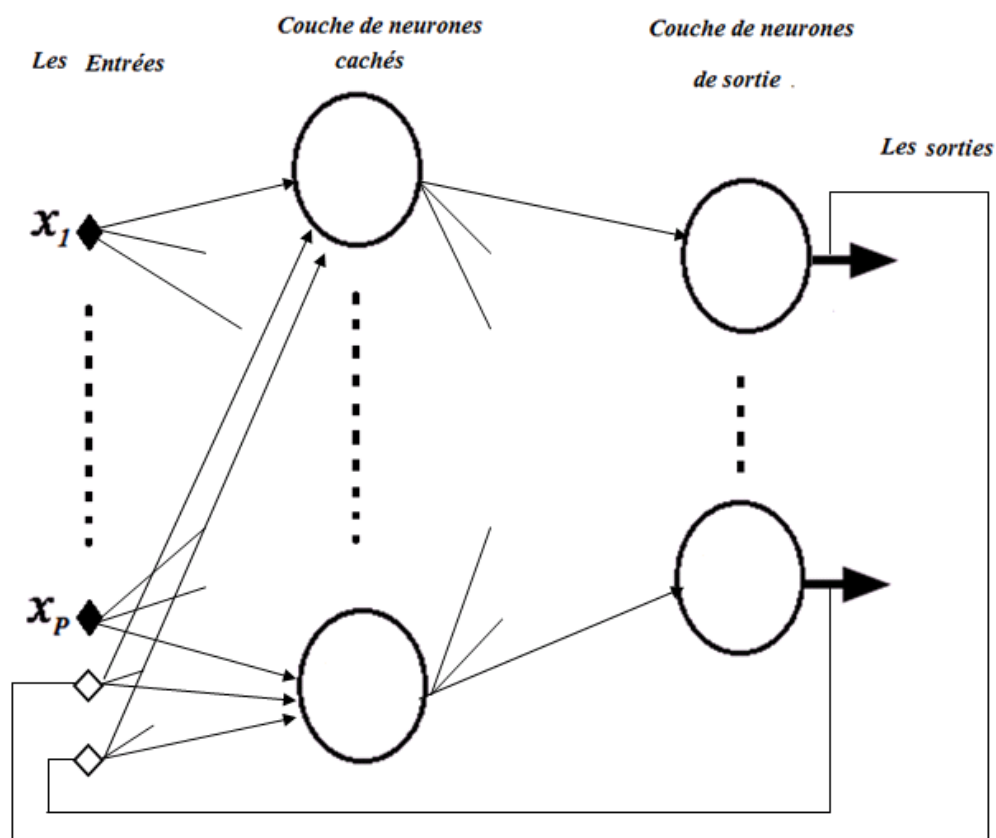


Figure 2.7 Réseau de neurones bouclé

Les réseaux de neurones bouclés sont utilisés pour effectuer des tâches de modélisation de systèmes dynamiques, de commande de processus, ou de filtrage (Drf 30).

7 L'apprentissage des réseaux de neurones

L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. "C'est-à-dire un changement dans la valeur des poids qui relient les neurones d'une couche à l'autre" (Bov 12). Soit le poids w_{kj} reliant le neurone k à son entrée j . Au temps t , un changement de poids peut s'exprimer simplement de la façon suivante (Otm 64) :

$$\Delta w_{kj}^{(t)} = w_{kj}^{(t+1)} - w_{kj}^{(t)} \quad (2,4)$$

Avec $w_{kj}^{(t+1)}$ et $w_{kj}^{(t)}$ représentent respectivement la nouvelle valeur et l'ancienne du poids w_{kj} .

7.1 Les types d'apprentissage

On distingue principalement deux types d'apprentissages : l'apprentissage « supervisé » et l'apprentissage « non supervisé ».

7.1.1 L'apprentissage supervisé

On peut affecter à un réseau non bouclé une tâche qui consiste à réaliser une fonction algébrique non linéaire ; celle-ci peut être :

- connue analytiquement : le réseau réalise alors une tâche d'approximation de fonction.
- inconnue analytiquement, mais pour laquelle on dispose de valeurs, en nombre fini, qui sont entachées de bruit si ces valeurs résultent de mesures effectuées sur un processus physique, chimique, économique, biologique, etc... : le réseau réalise alors une tâche de modélisation. C'est essentiellement pour ce dernier type d'applications que sont mis en œuvre les réseaux de neurones à apprentissage supervisé. On connaît donc, en tous points ou seulement en certains points, les valeurs que doit avoir la sortie du réseau en fonction des entrées correspondantes : c'est en ce sens que l'apprentissage est «supervisé», métaphore qui signifie qu'un «professeur » peut fournir au réseau des «exemples» de ce que celui-ci doit faire (Drf 29).

7.1.2 L'apprentissage non supervisé

Un réseau de neurones non bouclé peut également être utilisé dans un but de visualisation ou d'analyse de données : on dispose d'un ensemble de données, représentées par des vecteurs de grande dimension, et l'on cherche à les regrouper, selon des critères de ressemblance qui sont inconnus a priori. Ce type de tâches est connu en statistique sous le nom de méthodes «d'agrégation» (en anglais clustering).

On peut utiliser les réseaux de neurones non bouclés pour réaliser une tâche assez voisine : à partir de données, décrites par des vecteurs de grande dimension, trouver une représentation de ces données dans un espace de dimension beaucoup plus faible (typiquement de dimension 2) tout en conservant les «proximités» ou « ressemblances» entre ces données. Il n'y a donc pas là de « professeur », puisque c'est au réseau de découvrir les ressemblances entre les éléments de la base de données, et de les traduire par une proximité dans la « carte » de dimension 2 qu'il doit produire. Les réseaux à apprentissage non supervisé les plus étudiés et utilisés sont les « cartes auto-organisatrices » ou « cartes de Kohonen» (Drf 29).

7.2 L'algorithme d'apprentissage

L'apprentissage consiste en un entraînement du réseau ; on présente au réseau des entrées et on lui demande de modifier sa pondération de telle sorte que l'on retrouve la sortie correspondante. L'algorithme consiste dans un premier temps à propager vers l'avant les entrées jusqu'à obtenir une entrée calculée par le réseau. La seconde étape compare la sortie calculée à la sortie réelle connue. On modifie alors les poids de telle sorte qu'à la prochaine itération, l'erreur commise entre la sortie calculée et connue soit minimisée. Malgré tout, il ne faut pas oublier que l'on a des couches cachées. On rétro-propage alors l'erreur commise vers l'arrière jusqu'à la couche d'entrée tout en modifiant la pondération. On répète ce processus sur tous les exemples jusqu'à ce qu'on obtienne une erreur de sortie considérée comme négligeable. Dans un cadre plus général, l'apprentissage d'un réseau de neurones est basé sur la minimisation de l'erreur entre la sortie calculée de celui-ci et une réponse désirée, qu'elle soit donnée dans le cas supervisé ou générée dans le cas non supervisé. Cette minimisation est obtenue grâce à l'application d'un algorithme, le plus répandu étant l'algorithme de rétro-propagation (Moh 58).

7.2.1 L'algorithme de rétro-propagation

Cet algorithme repose sur la minimisation de l'erreur quadratique entre les sorties calculées et celles souhaitées (désirée). Le terme rétro-propagation du gradient provient du fait que l'erreur calculée en sortie est transmise en sens inverse vers l'entrée. Le signal d'erreur à la sortie du neurone k à l'itération n est (Moh 58):

$$e_k(n) = d_k(n) - y_k(n) \quad (2,5)$$

Où y_k est la sortie du neurone k et d_k la réponse désirée. On définit alors l'énergie instantanée de l'erreur à l'instant d'itération n par :

$$\mathcal{E}(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n) \quad (2,6)$$

Où C est l'ensemble des neurones de la couche de sortie du réseau. L'énergie d'erreur moyenne peut alors être calculée :

$$\mathcal{E}_{moy} = \frac{1}{N} \sum_{n=1}^N \mathcal{E}(n) \quad (2,7)$$

Pour minimiser \mathcal{E}_{moy} , on calcule son gradient par rapport à chaque poids w_{kj} , puis on modifie les poids dans le sens inverse du gradient. Celle-ci peut être interprétée comme une fonction de coût qui traduit la performance de l'apprentissage. En effet l'objectif de l'apprentissage est d'ajuster les paramètres libres du réseau, à savoir poids synaptiques et biais, afin de minimiser \mathcal{E}_{moy} (Moh 58). L'algorithme de rétro-propagation est basé sur une approximation similaire à celle de l'algorithme des moindres carrés. En effet, on estime que la correction à appliquer aux poids synaptiques w_{kj} de la couche i est proportionnelle à la dérivée partielle de l'énergie instantanée par rapport au poids synaptique (Moh 58):

$$\Delta w_{kj}^i = -\eta \frac{\partial \mathcal{E}(n)}{\partial w_{kj}^i} \quad (2,8)$$

η , est le taux d'apprentissage de l'algorithme de rétro-propagation. L'associativité des dérivées partielles permet d'obtenir la loi suivante (Moh 58) :

$$\Delta w_{kj}^i = \eta \delta_k^i(n) y_k^i(n) \quad (2,9)$$

où δ_k^i est le gradient local de la couche i défini par :

$$\delta_k^i(n) = -\frac{\partial \mathcal{E}(n)}{\partial v_{kj}^i} = e_k^i(n) \varphi_k^i(v_k^i(n)) \quad (2,10)$$

Si i est une couche de sortie, l'erreur se calcule directement d'après la relation (2.5)

$$\text{Sinon on a : } e_k^i(n) = \sum_j \delta_j^{i+1}(n) w_{jk}^{i+1}(n) \quad (2,11)$$

L'application de l'algorithme de rétro-propagation pour l'entraînement d'un réseau de neurones est résumée dans le tableau 2.2 (Moh 58).

Tableau 2.2 Algorithme de retro-propagation

1. Initiation du réseau

Choisir aléatoirement les poids synaptiques et les biais dans une distribution uniforme de moyenne nulle et de variance telle que les champs induits des neurones se trouvent dans la partie non saturée des fonctions d'activation sigmoïdales.

2. Présenter les entrées $\vec{x}(n)$ au réseau, calculer les sorties correspondantes et $\vec{e}(n)$ le vecteur d'erreur de la couche de sortie.

3. La propagation :

Calculer le champ induit local de chaque neurone k de chaque couche i :

$$v_k^i(n) = \sum_{j=0}^{m_k} w_{kj}^i(n) y_j^{i-1}(n)$$

4. La rétropropagation :

Calculer les gradients locaux de chaque neurone $\delta_k^i(n)$:

$$\delta_k^i(n) = -\frac{\partial \mathcal{E}(n)}{\partial v_{kj}^i} = e_k^i(n) \varphi_k^i(v_k^i(n)) \text{ si } i \text{ est une couche de sortie}$$

$$\delta_k^i(n) = \varphi_k^i(v_k^i(n)) \sum_{j=0}^{m_k} w_{kj}^i(n) y_j^{i-1}(n) \text{ sinon}$$

Puis mise à jour des poids synaptiques :

$$w_{kj}^i(n+1) = w_{kj}^i(n) + \eta \delta_k^i(n) y_j^{i-1}(n)$$

où η est le taux d'apprentissage

5. Itération des étapes 2 à 4 jusqu'à avoir rencontré un critère d'arrêt.

L'amélioration de l'algorithme de rétro-propagation est basée sur le développement en série de Taylor de la fonction coût, aussi l'approximation de Newton permet de calculer l'optimum de la fonction cout. Ceci, implique l'existence de l'inverse du Hessien. D'autre part, même s'il existe, le calcul de celui-ci nécessite beaucoup de mémoire. Une alternative consiste à approximer le Hessien plutôt que de le calculer. C'est ce qui est fait dans l'algorithme de Levenberg-Marquardt (Moh 58). - Annexe B -

8 Propriétés fondamentales des réseaux de neurones

8.1 L'approximation fonctionnelle

Les Réseaux de Neurones Artificiels appliquent, entre autre, le principe de l'approximation fonctionnelle. Ainsi, ils apprennent une fonction en utilisant des exemples de la dite fonction. Un des exemples le plus simple est celui d'un Réseau de Neurones artificiels apprenant la fonction XOR (OU-exclusif), mais il pourrait tout aussi bien apprendre à déterminer la langue d'un texte, ou la présence d'une tumeur dans une image passée au rayon-X (Ste 83).

Les variables d'entrée d'un problème consistant à trouver une tumeur sur une image au rayon-X pourraient être les valeurs de pixels de l'image, mais elles pourraient également prendre d'autres valeurs extraites de l'image en question. Les données de sortie pourraient être soit une valeur binaire soit une valeur à virgule flottante représentant la probabilité de la présence d'une tumeur dans l'image (Ste 83).

La propriété d'approximation peut être énoncée de la manière suivante : toute fonction bornée suffisamment régulière peut être approchée avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini (Hor 40).

8.2 L'approximation parcimonieuse des réseaux de neurones

La spécificité des réseaux de neurones réside dans le caractère « parcimonieux » de l'approximation : à précision égale, les réseaux de neurones nécessitent moins de paramètres ajustables (les poids des connexions) que les approximateurs universels couramment utilisés ; plus précisément, le nombre de poids varie linéairement avec le nombre de variables de la fonction à approcher, alors qu'il varie exponentiellement pour la plupart des autres approximateurs (Hor 42). En pratique dès qu'un problème fait intervenir plus de deux variables, les réseaux de neurones sont, en général, préférables aux autres méthodes (Drf 30).

9 Avantage et inconvénients des réseaux de neurones

9.1 Avantages (Mou 59)

- Approximateur universel parcimonieux et classifieur universel.
- Rapidité d'exécution
- Robustesse des solutions, résistance au bruit des données
- Facilité de développement

9.2 Inconvénients (59)

- Le choix de l'architecture est critique
- Le temps d'apprentissage peut être long
- Présence de minima locaux de la fonction de coût
- Difficultés d'interprétation des résultats.

Chapitre 3

Application des techniques de prétraitement sur une mammographie

1. Prétraitement de la mammographie

La mammographie est une radiographie qui permet le dépistage du cancer du sein, mais seulement une région de cette image est importante, d'où l'intérêt d'appliquer un prétraitement pour focaliser le traitement numérique sur la région d'intérêt de la mammographie.

Plusieurs études ont été menées pour effectuer l'extraction de cette zone d'image (Bos 11),(Deh 24),(Mai 49), (Sri 79), (Gok 37),(Rah 71).

Dans cette étude, nous présentons une méthode pour isoler la région d'intérêt d'une mammographie basée sur plusieurs techniques de traitement d'images, cette méthode est appliquée sur une base d'images.

Le prétraitement de la mammographie pour extraire la région d'intérêt est une étape importante pour assurer la fiabilité des Systèmes d'Aide au Diagnostic (CAD) par ordinateur.

La mammographie comporte des objets non-utiles au diagnostic médical. Notre objectif est l'extraction de la région d'intérêt d'une mammographie, on doit isoler cette zone du reste de l'image à savoir, les artefacts et le bruit afin d'écartier l'influence de ces éléments sur les différentes étapes d'analyse ultérieures.

Notre méthode est appliquée sur des mammographies provenant de la base mini-MIAS : Mammographic Image Analysis Society.

Par observation de la figure (3.1) on constate la possibilité d'avoir deux orientations différentes des mammographies pour cette raison la première étape de prétraitement consiste à reconnaître l'orientation de la mammographie et d'orienter toutes les mammographies vers la droite.

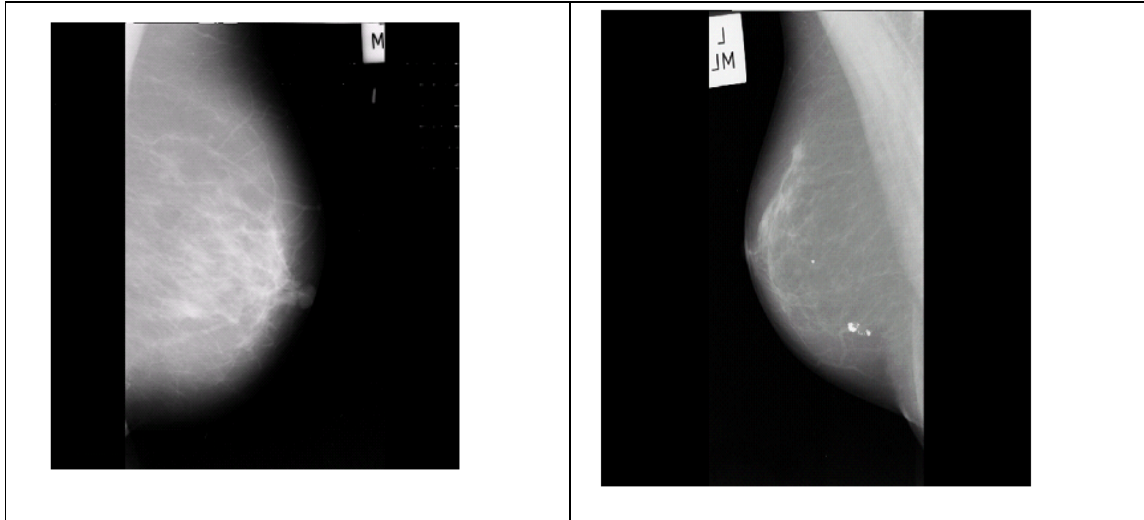


Figure 3.1 Deux mammographies de la base Mini-MIAS

2. Description de la méthode de prétraitement

Pour extraire la région d'intérêt on doit réaliser une segmentation basée sur plusieurs techniques.

En commence par une orientation du sens de la mammographie vers la droite et un ajustement de l'image, cette image ajustée est utilisée pour déterminer un seuil par méthode de Otsu afin d'obtenir une image en forme binaire. On applique sur cette dernière une labellisation afin de supprimer les artefacts de l'image, en phase finale on obtient la région d'intérêt de la mammographie. Cette approche sera appliquée aux différents types de tissus mammaires de la base de données mini-MIAS, pour observer sa fiabilité, figure (3.2).

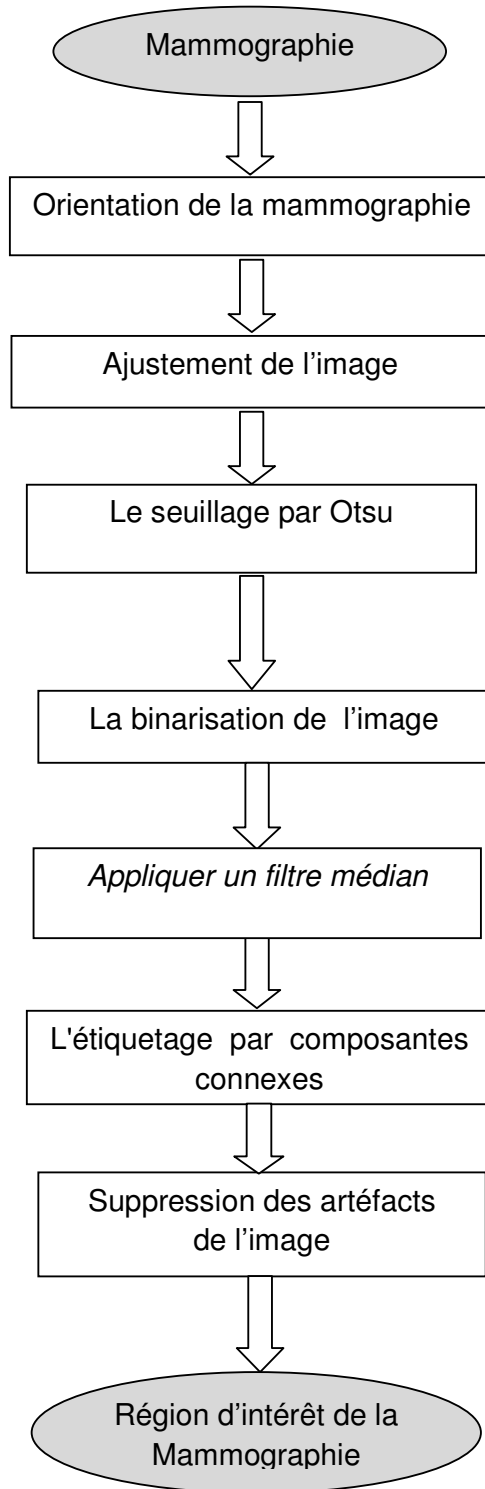


Figure 3.2 Organigramme de la méthode de prétraitement

2.1 Orientation de la mammographie

Pour reconnaître le sens d'orientation de la mammographie et mettre toutes les images dans une seule direction; par l'utilisation du niveau de gris des deux moitiés de l'image (Deh 24).

on compare la somme des niveaux de gris de la moitié droite et la moitié gauche de l'image pour identifier le sens d'orientation.

La mammographie est mémoriser sous forme de matrice et l'interversion des colonnes gauche vers la droite permet de changer le sens d'orientation de la mammographie.

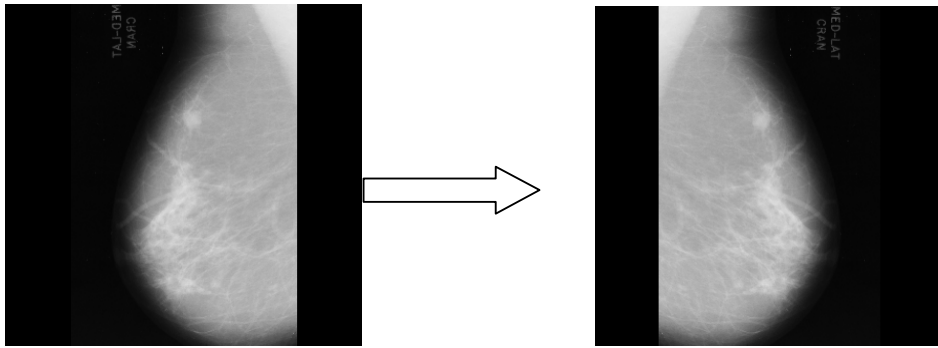


Figure 3.3 Orientation de la mammographie

2.2 Ajustement de la mammographie

L'ajustement des niveaux de gris permet de modifier le contraste et/ou la luminosité d'une image monochrome. L'ajustement de l'image sera réaliser en deux phases :

- transformation linéaire de l'histogramme ;

- Correction gamma.

2.2.1 Transformation linéaire de l'histogramme

On appelle dynamique l'intervalle $[a,b]$, où a et b sont respectivement les niveaux de gris minimal et maximal présents dans l'image. La dynamique maximale est $[0,255]$ pour une image 8 bits (Isi 44).

L'amélioration peut être effectuée par manipulation d'histogramme, une transformation linéaire à pour but de redistribuer les niveaux de gris de l'image pour qu'ils occupent toute la bande de nuances possible, c.-à-d. la dynamique maximale.

Une image numérique étant essentiellement discrète. Dans ce qui suit x et y sont des entiers qui représentent les coordonnées des pixels et f représente les valeurs du niveau de gris pour chaque pixel de l'image originale.

soit $f(x,y)$ qui représente l'image originale et $g(x,y)$ l'image améliorée, pour l'amélioration de l'image on applique la transformation suivante (Isi 44):

$$g = \alpha f + \beta$$

$$\alpha = \frac{g_{\max} - g_{\min}}{b - a}$$

$$\beta = \frac{g_{\min} \cdot b - g_{\max} \cdot a}{b - a} \quad (3.1)$$

Pour $g_{\max} = 255$ et $g_{\min} = 0$ on obtient la figure suivante :

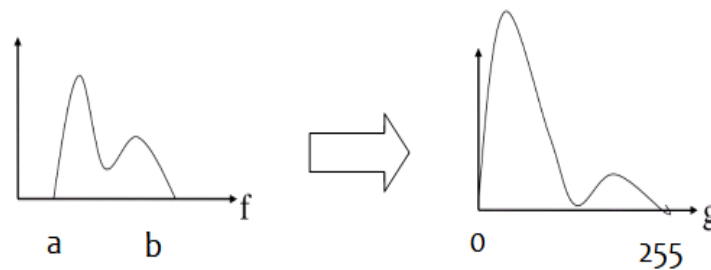


Figure 3.4 Transformation linéaire de l'histogramme (Isi 44).

2.2.2 Correction gamma

Afin de rendre les régions et les objets composants l'image bien distincte et bien séparées les uns des autres on applique une Correction gamma par la formule suivante (Isi 44):

$$g(x, y) = \frac{g_{\max}}{g_{\max}^{\gamma}} f(x, y)^{\gamma} \quad (3.2)$$

- Lorsque $\gamma > 1$: Cette transformation augmente la plage dynamique des intensités élevées : dilatation des zones claires.
- Lorsque $\gamma < 1$: elle augmente la plage dynamique des intensités faibles : dilatation de zones sombres (Isi 44).

2.2.3 Application de l'ajustement sur une mammographie

L'Ajustement de l'image permet de diminuer l'effet du faible contraste entre les régions du fond de l'image et le sein et permet également la réduction de l'invisibilité des contours vis-à-vis du fond, figure (3.5).

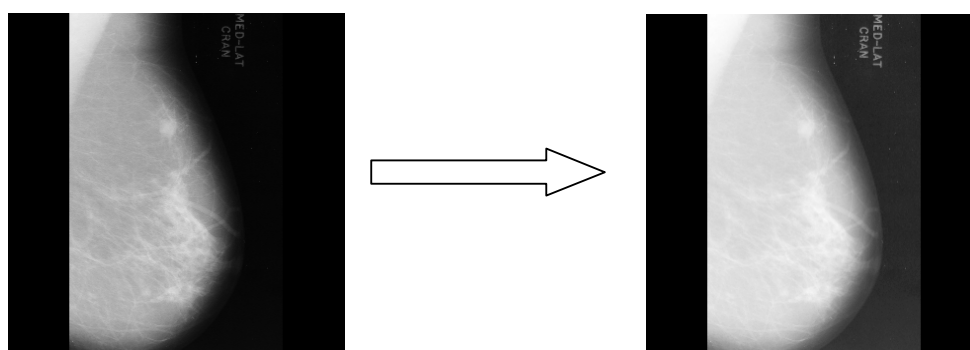


Figure 3.5 Ajustement de la mammographie

2.3 Méthode Otsu

La méthode d'Otsu (Ots 65) est une des méthodes de calcul de seuil automatique pour la mise sous forme binaire d'une image à niveau de gris. Otsu assimile la difficulté de détermination d'un seuil adapté s pour la binarisation de l'image à un problème de meilleure classification des pixels en deux classes. Une de ces classes contient toujours le fond et l'autre les objets de l'image, figure (3.6).

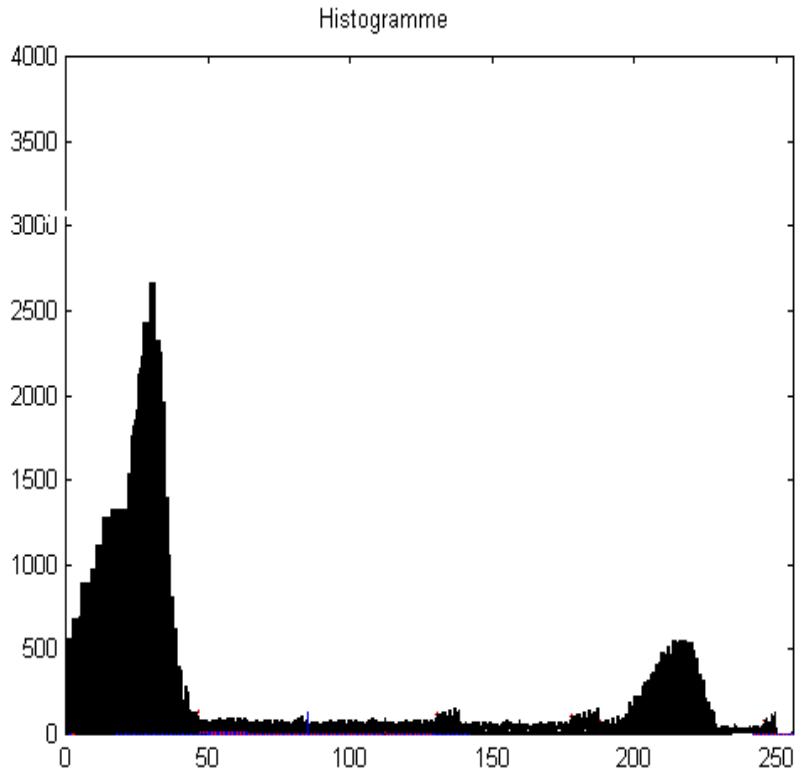


Figure 3.6 Histogramme divisée en deux classes

La méthode d'Otsu (Ots 65) fait partie des méthodes de recherche de seuil les plus utilisées en traitement d'image. Cette technique permet de choisir un seuil qui minimise la variance des intensités des pixels entre deux classes de pixels et qui maximise la différence des moyennes d'intensité de ces deux classes (Tol 88).

On considère que les mammographies traitées ont 256 niveaux de gris. Soit h l'histogramme d'une mammographie. On pose p_k la probabilité de distribution du niveau k , définie par (Tol 88) :

$$p_k = \frac{h(k)}{N} \quad (3.3)$$

Avec N le nombre total de pixels de l'image. La variance interclasse pondérée est définie par :

$$\sigma_{\omega}^2(s) = \omega_1(s)\sigma_1^2(s) + \omega_2(s)\sigma_2^2(s) \quad (3.4)$$

Avec s le seuil de séparation des deux classes et :

$$\omega_1(s) = \sum_{k=1}^s p_k \quad ; \quad \omega_2(s) = \sum_{k=s+1}^{256} p_k = 1 - \omega_1(s) \quad (3.5)$$

et

$$\sigma_1^2(s) = \sum_{k=1}^s (k - \mu_1(s))^2 \frac{p_k}{\omega_1(s)} \quad (3.6)$$

$$\sigma_2^2(s) = \sum_{k=s+1}^{256} (k - \mu_2(s))^2 \frac{p_k}{\omega_2(s)} \quad (3.7)$$

Avec les moyennes de classe définies ainsi :

$$\mu_1(s) = \sum_{k=1}^s \frac{kp_k}{\omega_1(s)}; \quad \mu_2(s) = \sum_{k=s+1}^{256} \frac{kp_k}{\omega_2(s)} \quad (3.8)$$

L'algorithme d'Otsu consiste à trouver le seuil s qui minimise $\sigma_{\omega}^2(s)$ (Tol 88).

2.4 La binarisation

Les pixels de l'image sont partagés par un seul seuil s en deux classes : ceux qui appartiennent au fond et ceux qui appartiennent aux objets de l'image, le seuil est déterminée au préalable.

La binarisation consiste à faire un seuillage. Si la valeur du niveau de gris du pixel est en dessous du seuil, nous lui associons la valeur 0. Si la valeur du niveau de gris du pixel est égale ou supérieur au seuil nous lui donnons la valeur 1.

$$g(x, y) = \begin{cases} 0, si : f(x, y) < s \\ 1, si : f(x, y) \geq s \end{cases} \quad (3.9)$$

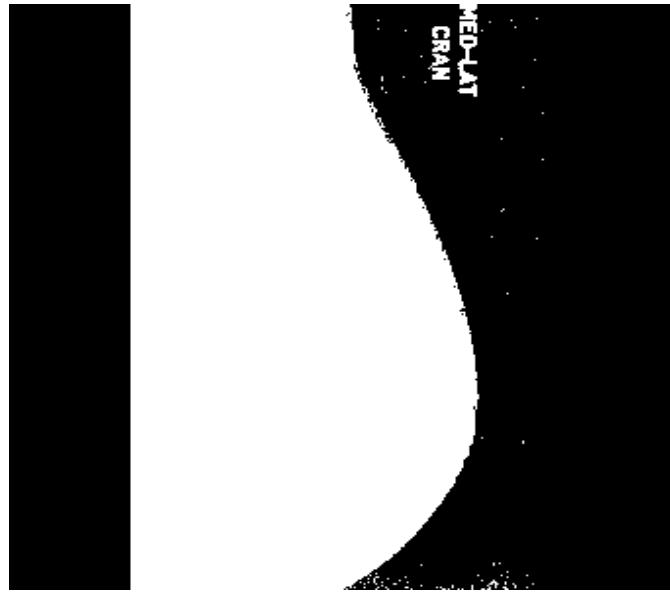


Figure 3.7 Mammographie en forme binaire

2.5 Le filtre médian

Déplacer une fenêtre de taille impaire sur le support image et remplacer le pixel central (sur lequel est positionnée la fenêtre) par la valeur médiane des pixels inclus dans la fenêtre (Dub 31). La médiane d'un ensemble de valeur, est la valeur "du milieu" (Mat 51). Le principe de l'algorithme du filtre médian est le suivant (Mat 51) :

1. Chercher les valeurs des voisins du pixel courant,
2. Ordonner ces valeurs (tri),
3. Sélectionner celle du milieu.

Ce filtre est utile pour contrer l'effet Poivre et Sel (P&S) c'est-à-dire des faux « 0 » et « 255 » dans l'image à 8bits (Ber 08) et permet de préserver les contours (Mat 51). La figure suivante présente un exemple d'application du filtre médian.

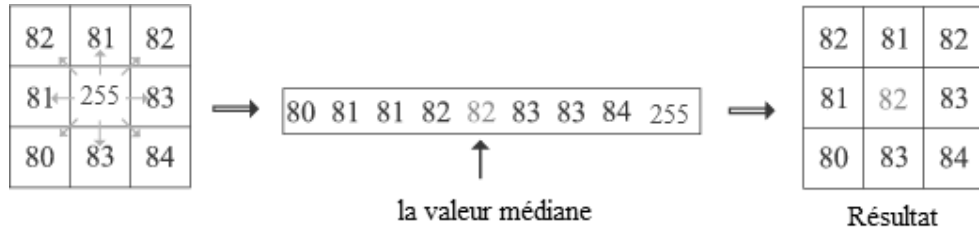


Figure 3.8 exemple d'application du filtre médian.

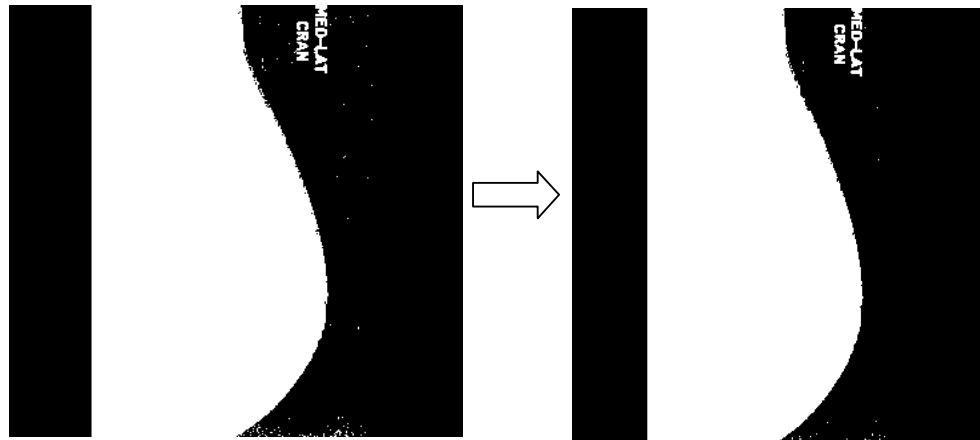


Figure 3.9 application du filtre médian sur une mammographie en forme binaire

2.6 L'étiquetage par composantes connexes

Sur une image binaire, on peut considérer deux choses: le fond de l'image, en noir et les objets ou régions en blanc sur l'image, l'étiquetage par composantes connexes (ou labellisation) de l'image repose sur cette considération.

Le concept de l'étiquetage en composantes connexes s'applique généralement sur les images binaires, et consiste à trouver les différentes composantes connexes de l'image et à les étiqueter. L'étiquetage consiste à

affecter une étiquette identique à tous les pixels d'une même composante connexe (Toh 89).

L'étiquetage en composantes connexes d'une image binaire attribue à tous les pixels d'une composante connexe la même valeur entière. Chaque composante connexe est alors identifiée par son étiquette. Les algorithmes d'étiquetage détectent les adjacences entre les pixels et définissent l'étiquette du point courant en fonction de celle de ses points voisins. Ils opèrent en balayant l'intégralité de l'image. L'idée de l'algorithme repose sur un balayage séquentiel de l'image (Toh 89). On considère pour chaque point P_x ses prédécesseurs déjà traités. Les prédécesseurs d'un point dépendent du type de connexité considéré (Figure 3.10).

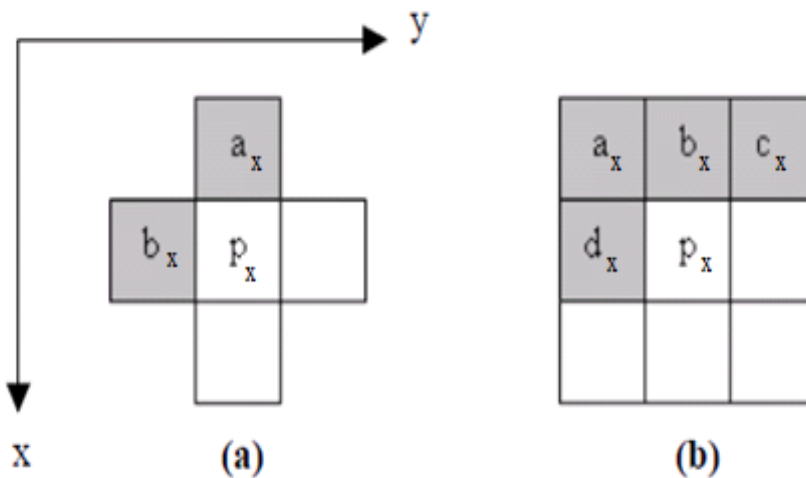


Figure 3.10 Prédécesseurs d'ordre 4 (a), prédécesseurs d'ordre 8 (b) (Toh 88).

En voisinage d'ordre 4, P_x a deux prédécesseurs de coordonnées $(x-1,y)$ et $(x,y-1)$. En voisinage d'ordre 8 il en a quatre de coordonnées $(x,y-1)$, $(x-1,y-1)$, $(x-1,y)$ et $(x-1,y+1)$, le principe est le suivant (en voisinage 4 sur lequel est basé l'algorithme ci-dessous) (Toh 89) :

1. balayer l'image ligne par ligne de haut en bas et de gauche à droite pour détecter les pixels qui sont à 1.
2. Si P_x appartient au fond, passer à la position d'après sinon;
 - Si a_x et b_x appartiennent au fond affecter une nouvelle étiquette à P_x .
 - Si a_x ou b_x n'appartient pas au fond, affecter son étiquette à P_x .
 - Si a_x et b_x n'appartiennent pas au fond, mais qu'ils ont des étiquettes différentes, affecter la plus petite étiquette à P_x et considérer les deux étiquettes comme équivalentes.
3. Après ce balayage, les pixels n'appartenant pas au fond sont étiquetés. Certaines étiquettes sont équivalentes, on leur affecte la même étiquette.
4. Faire un 2^{ème} balayage pour affecter aux pixels leur étiquette finale.

Le principe est de détecter les régions blanches de l'image en leur attribuant à chacune un label (numéro d'identification) différent, figure (3.11).

Pour une image binaire, on produit une image dite de label qui comportera, pour chaque pixel de l'image binaire, le label qu'il lui est associé (Aub 03).

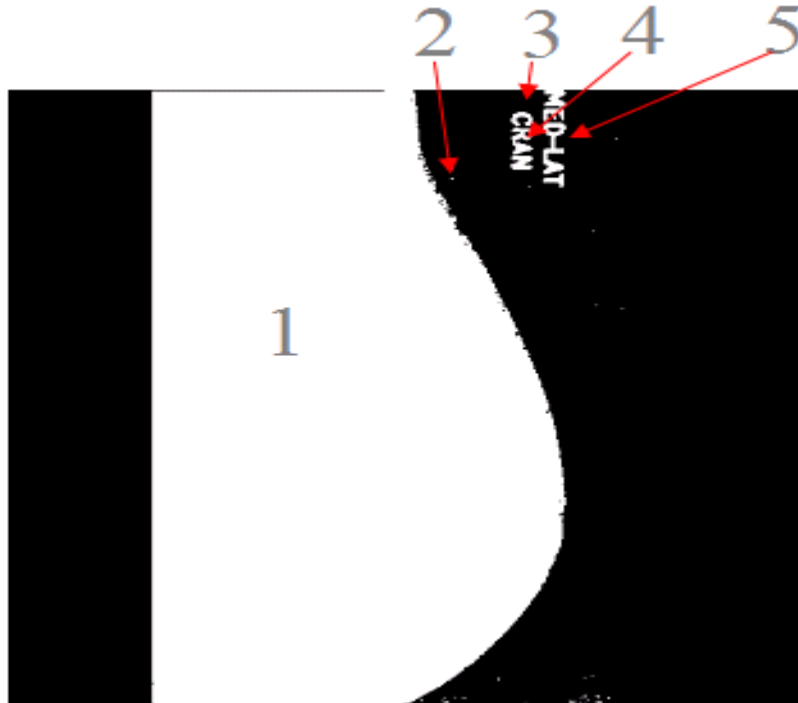


Figure 3.11 L'étiquetage (ou labellisation) d'une mammographie

2.7 Suppression des artefacts de l'image

Cette étape consiste à supprimer les artefacts de l'image en utilisant le résultat de la labellisation, figure (3.12). $gt(x,y)$ représente l'image avec étiquetage, $f(x,y)$ l'image originale et $gs(x,y)$ l'image avec suppression des artefacts, il suffit d'appliquer les relations suivante :

$$\begin{cases} \text{Si } gt(x,y)=1 \text{ Alors } gs(x,y) = f(x,y) \\ \text{Si non } gs(x,y) = 0 \end{cases} \quad (3.10)$$

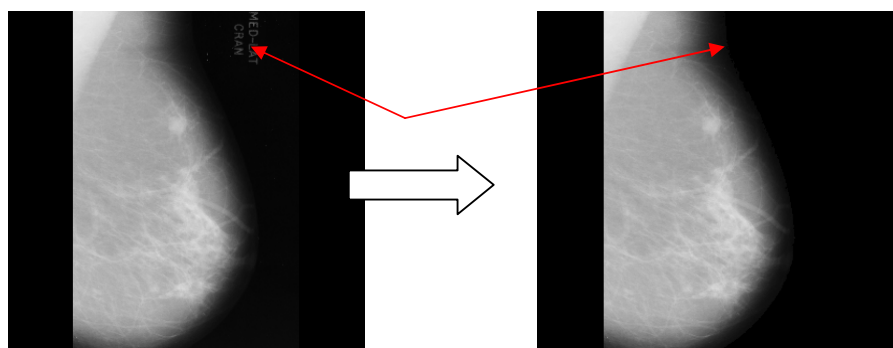
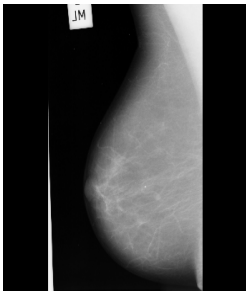

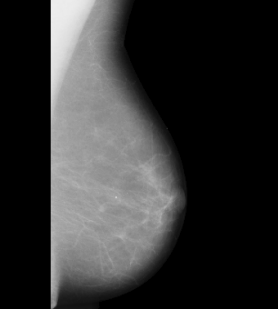
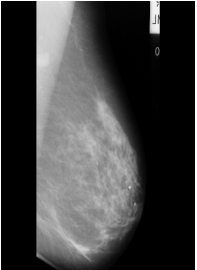

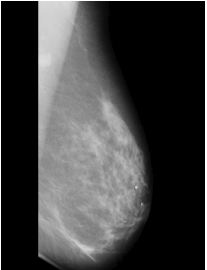
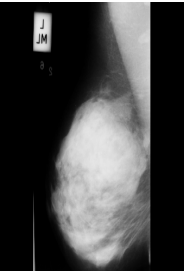

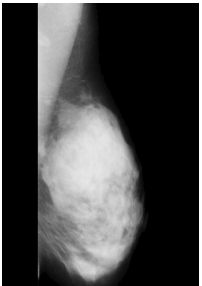


Figure 3.12 application de la suppression des artefacts

3. Application de la méthode

Le tableau suivant présente l'application de la méthode pour les différents types de tissus de la base mini-MIAS, pour chaque cas on expose la mammographie d'origine, les éléments enlevés et l'image de la région d'intérêt de la mammographie.

Tableau. 3.1 Application de la méthode pour les différents types de tissus mammaires -mini-MIAS-

| Type de tissu | Image d'origine | Les éléments supprimés | Région d'intérêt |
|---------------|---|--|---|
| F |  |  |  |
| G |  |  |  |
| D |  |  |  |

L'application de la méthode de prétraitement décrite précédemment sur les mammographies de la base DDSM, a posé des problèmes d'espace mémoire pendant la phase d'exécution des calculs.

Chapitre 4

Classification automatique de la densité des tissus mammaires

1. Système de classification automatique de la densité des tissus mammaires

La densité mammaire a une influence sur la fiabilité de l'interprétation des mammographies (Sti 80) et plusieurs chercheurs décrivent une association entre la densité mammaire et le risque de cancer du sein (Krl 47), (McC 52).

Notre objectif est de développer un système de classification automatique de la densité des tissus mammaires à base de réseaux de neurones artificiels. Le vecteur d'entrée de chaque réseau de neurones est calculé à partir de l'histogramme de la mammographie.

La densité mammaire peut être quantifiée à partir de l'histogramme d'une mammographie (Byn 16). Dans l'approche proposée, les valeurs prélevées de l'histogramme seront traitées par des réseaux de neurones artificiels (RNA).

Il n'existe pas de règle théorique permettant de choisir la meilleure configuration d'un réseau de neurones artificiels, c'est la raison pour laquelle la conception d'un réseau de neurones se fait de manière expérimentale.

Les approches utilisant les réseaux de neurones posent souvent des problèmes de convergence, d'initialisation des poids synaptiques et de choix de la topologie du réseau de neurones. Pour résoudre ces problèmes, on doit appliquer les considérations pratiques suivantes :

- Initialiser tous les poids du réseau à des petites valeurs aléatoires (Che 18);
- Un taux d'apprentissage compris entre 0 et 1, ce taux détermine l'importance de la modification des poids durant l'apprentissage (Sté 84) ;
- L'utilisation d'une base d'apprentissage comprenant les exemples d'entrée ainsi que les sorties désirées associées, permet de réaliser un apprentissage supervisé. Les paramètres du modèle vont ainsi s'adapter en comparant à chaque fois les sorties obtenues et les sorties désirées (Cha 17) ; (Mit 55).

Un problème de classification multi-classe peut-être résolu par un réseau de neurones à plusieurs sorties ou bien un système composé de plusieurs réseaux de neurones, chaque réseau possédant une seule sortie binaire.

Notre choix d'utiliser un système de classification composé de plusieurs réseaux de neurones est justifié par le fait que plusieurs chercheurs (Ahm 01); (Ngu 61) ;(Pat 68) affirment que le système multi-réseaux de neurones est plus performant que le réseau de neurones unique a plusieurs sorties.

L'élaboration de notre système de classification s'effectue, selon les étapes suivantes (Bou 09) :

- établissement de l'histogramme de la mammographie ;
- prélèvement des valeurs d'entrée du système de l'histogramme;
- normalisation des valeurs ;
- envoyer les valeurs normalisées à l'entrée des réseaux de neurones.

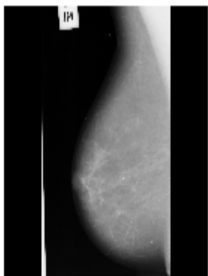
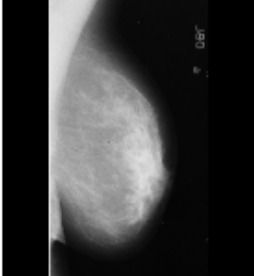
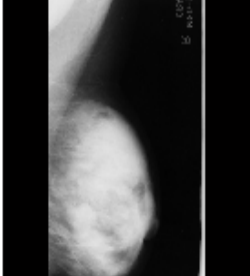
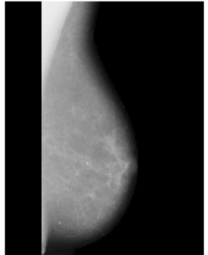
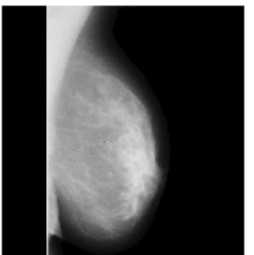
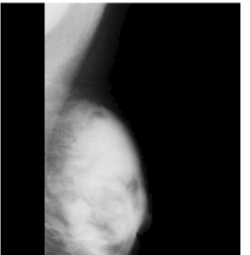
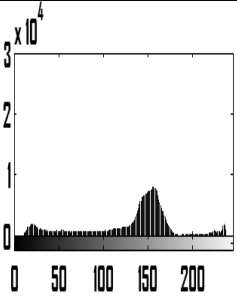
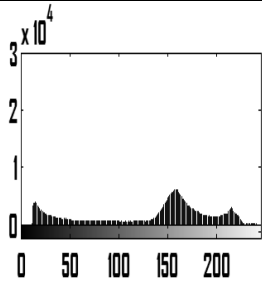
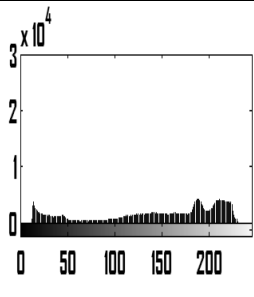
1.1. Classification automatique de la densité des tissus mammaires de la base mini-MIAS avec prétraitement d'image

La méthode est appliquée pour la classification des tissus de la base mini-Mias avec prétraitement des images.

La classification de la base mini-MIAS s'effectuera par l'utilisation de deux réseaux de neurones (Bou 09), puisque dans cette base d'images, il y a seulement trois catégories de tissus mammaires (Suc 82).

Le tableau suivant présente les histogrammes des différents types de tissus de la base mini- MIAS avec prétraitement d'image :

Tableau 4.1 trois mammographies de la base mini MIAS et leurs histogrammes après prétraitement

| Type de tissu | Tissu grassex (F) | Tissu grassex-glandulaire (G) | Tissu dense (D) |
|---|---|--|---|
| Image d'origine |  |  |  |
| Image après prétraitement |  |  |  |
| Histogramme de chaque Image après prétraitement |  |  |  |

Chaque histogramme du tableau 4.1 présente un pic, sur la droite, cette partie représente la zone d'intérêt et doit être traitée puisqu'il correspond aux tissus mammaires.

L'utilisation d'un nombre limité des valeurs fournies par l'histogramme permet de réduire le nombre d'entrées, ceci afin de diminuer le temps de calcul et accélérer l'apprentissage du système de classification de tissus mammaires (Bou 09).

Le nombre de valeurs à utiliser est choisi de manière arbitraire. La zone d'intérêt est limitée par les niveaux de gris 100 et 245. À partir de cet intervalle, on prend 30 valeurs régulièrement réparties pour former un vecteur d'entrée « E ». Chacune de ces valeurs représente le nombre de pixels pour chaque niveau de gris (Bou 09).

En pratique, il est recommandé d'effectuer une normalisation, pour chaque vecteur d'entrée « E » (Drf 29).

On calcule la valeur moyenne « m » et l'écart-type « σ » par rapport aux composantes du vecteur d'entrée « E », ensuite une normalisation du vecteur « E » est réalisée pour obtenir le vecteur normalisé « E' » suivant la formule ci-dessous :

$$E' = \frac{E - m}{\sigma} \quad (4,1) \text{ (Drf 29)}$$

Chacune des valeurs du vecteur E' représente le nombre de pixels normalisés pour chaque niveau de gris, ces valeurs servent à former le vecteur d'entrée : E1'.....E30'. De sorte que le vecteur normalisé « E' » a une moyenne voisine de zéro et un écart-type de l'ordre de 1 (Drf 29). Les valeurs du vecteur « E' » sont introduites dans la couche d'entrée des réseaux de neurones.

Le système de classification de la densité des tissus mammaires de la base mini - MIAS est composé de deux étapes successives, chaque étape utilise un réseau de neurones, le premier réseau permet d'identifier les tissus mammaires gras (Fatty F), ce même réseau de neurones peut indiquer que le tissu traité et non gras dans ce cas, le deuxième réseau a pour

rôle d'identifier les tissus graisseux-glandulaires (G) et les tissus denses (D)(Bou 09) tel qu'illustré à la figure (4.1).

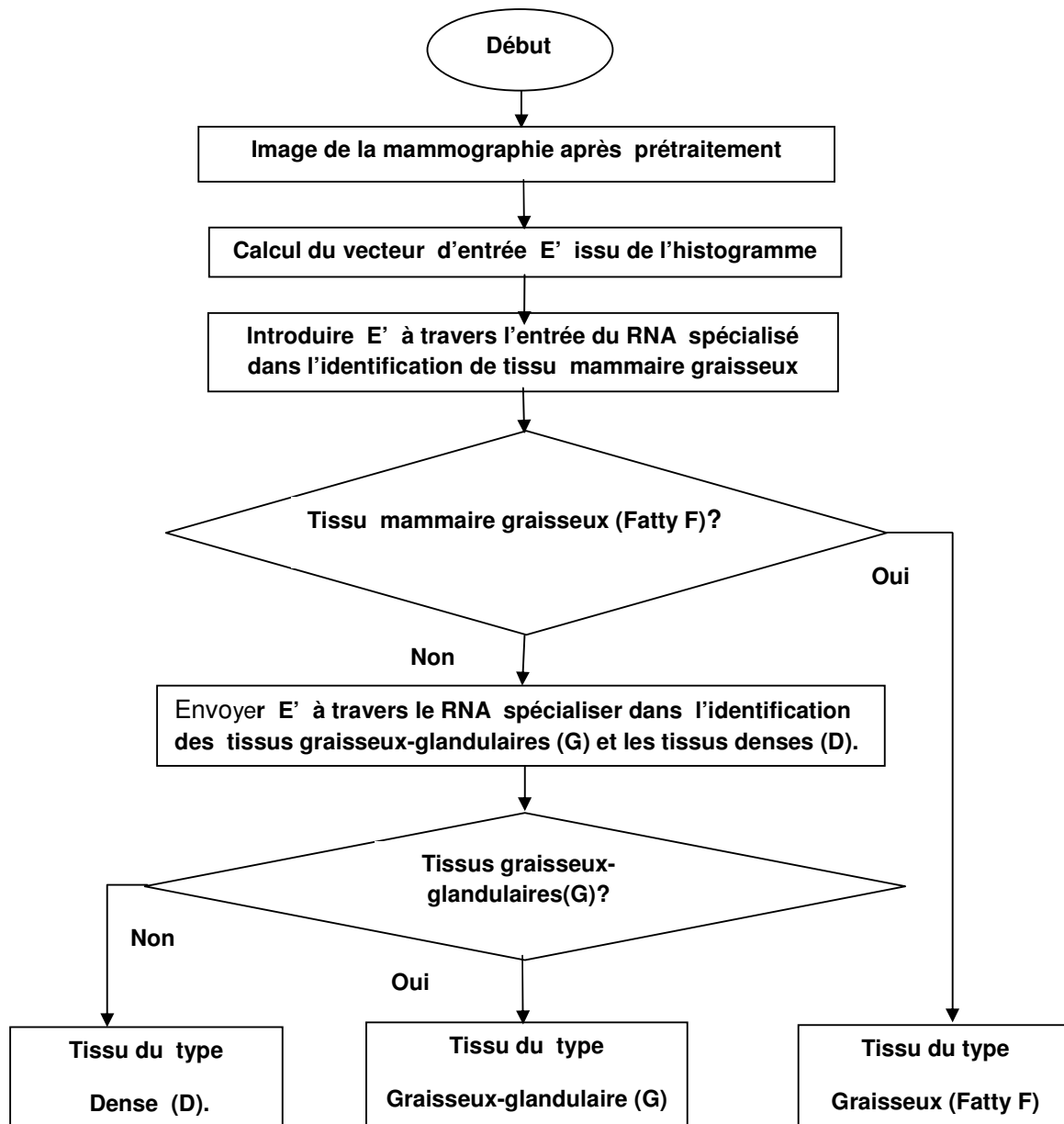


Figure 4.1 Organigramme du système de classification de la densité des tissus mammaires conformément à la classification de la base mini-MIAS avec prétraitement d'image

Quatre vingt pour cents des mammographies traitées seront utilisées pour l'apprentissage et 20 % pour le test, pour chaque réseau de neurones (Bou 09). Pour la conception d'un système de classification de tissu mammaire de la base mini-MIAS, on a choisi 180 mammographies réparties en trois sous-groupes de 60 images pour chaque type de tissu (Bou 09).

La première étape consiste à calculer des indicateurs à partir de l'histogramme de l'image en niveaux de gris après prétraitement.

Le système de classification est formé de deux réseaux de neurones. Les réseaux de neurones de notre application sont composés de trois couches (Bou 09) figure 4.2 :

- une couche d'entrée, elle contient autant de neurones que le nombre d'éléments du vecteur d'entrée $E' : E1' \dots E30'$;
- pour ne pas augmenter la complexité du réseau, nous avons opté pour une seule couche cachée (Bou 09) ;
- une couche de sortie avec un neurone "S", cette sortie permet d'identifier le type du tissu, "S" peut prendre la valeur "0" ou "1" (Bou 09).

La fonction d'activation (ou fonction de transfert) sert à convertir le résultat de la somme pondérée des entrées d'un neurone en une valeur de sortie, cette conversion s'effectue par un calcul de l'état du neurone et permet d'introduire une non-linéarité dans le fonctionnement du neurone (Cyb 23).

Les fonctions d'activation et le nombre de neurones dans la couche cachée ont été choisis expérimentalement ; le modèle le plus performant sera sélectionné (Bou 09).

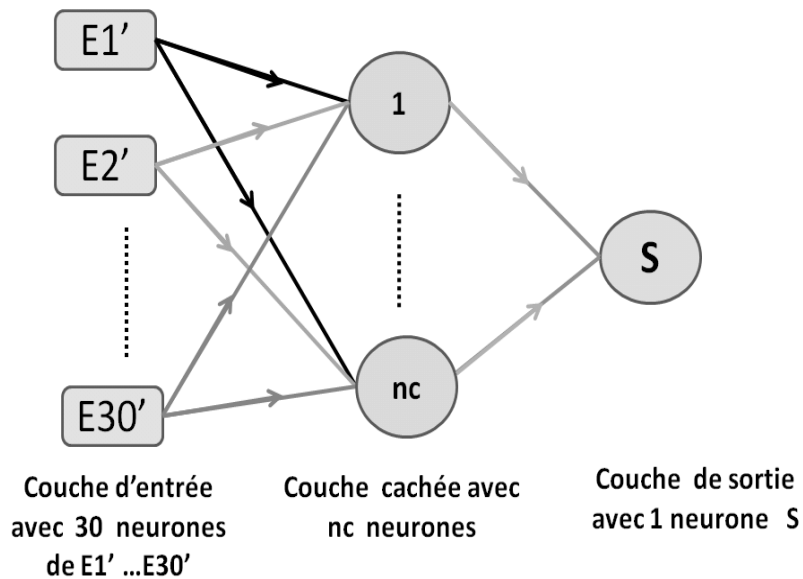


Figure 4.2 . RNA à 3 couches, nc neurones cachés, 1 neurone de sortie et 30 d'entrées.

L'expérimentation du réseau de neurones, nécessite l'application des tâches suivantes (Bou 09):

- faire varier nc de 1 à 50, nc étant le nombre de neurones en couche cachée ;
- changer fc, fc désignant la fonction de transfert de la couche cachée ;
- changer fs, fs désignant la fonction de transfert de la couche de sortie.

Pour simplifier, on considère un seul type de fonction par couches de neurones, trois fonctions de transfert seront envisagées (Bou 09):

- la fonction linéaire : 'purelin', figure (4 .3-a) ;
- la tangente hyperbolique : 'tansig', figure (4 .3-b) ;
- la fonction sigmoïde : 'logsig', figure (4 .3-c).

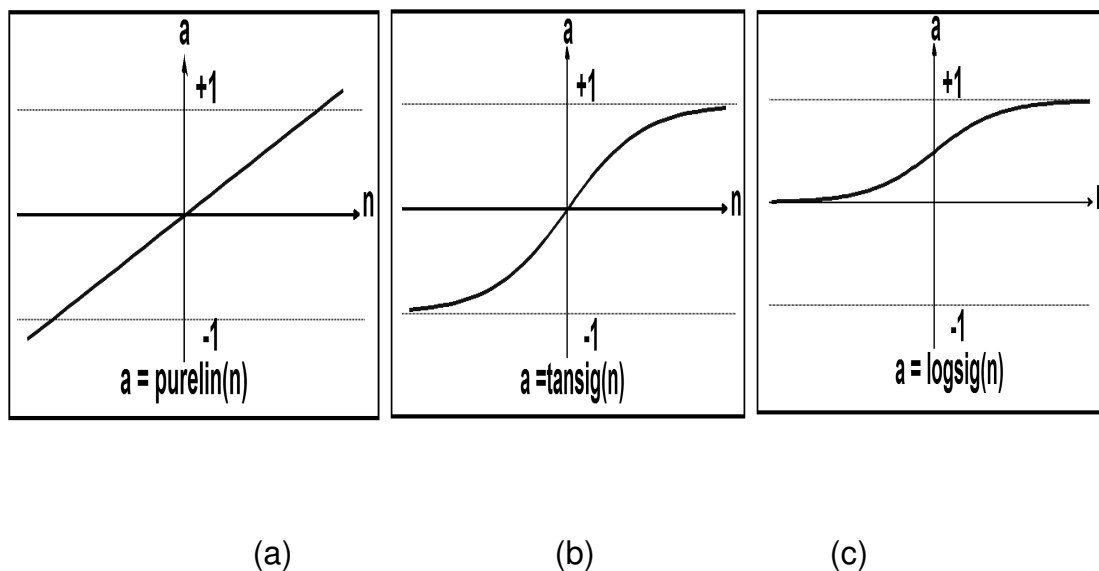


Figure 4.3 Les trois fonctions de transfert

L'utilisation des fonctions 'purelin', 'tansig' et 'logsig' au niveau de la couche cachée et de la couche de sortie, permet d'obtenir 9 couples de fonctions d'activation (Bou 09) :

tansig-tansig ; tansig-logsig ; tansig-purelin ;

logsig-logsig ; logsig-tansig ; logsig-purelin ;

purelin-purelin ; purelin-logsig ; purelin-tansig.

L'apprentissage est effectué pour chaque configuration, avec les paramètres suivants (Bou 09):

- nombre maximum d'itérations (Epochs) = 1000 ;
- gradient minimum = 1 e-10.

L'apprentissage est arrêté si l'une de ces conditions est satisfaite (Bou 09).

Critères de performance :

Les performances de chaque réseau de neurones sont observées par comparaison des résultats estimés par le réseau avec les états réels. Le taux de reconnaissance est défini comme étant le rapport entre le nombre des cas reconnus correctement et le nombre des cas présentés pour le test du système (Tre 92).

$$\text{Taux de reconnaissance} = \frac{\text{Bonnes décisions}}{\text{Nombres total d'exemples}} \times 100$$

(4,2)

On choisie le réseau le moins complexe en cas d'égalité des performances (Bou 09).

Le tableau 4.2 est une présentation non exhaustive des taux de reconnaissance obtenue pour la base mini-MIAS, en variant le nombre de neurones en couche cachée et les couples des fonctions de transfert.

Tableau 4 .2 Des exemples de taux de reconnaissance des tissus mammaires conformément à la classification de la base mini-MIAS réalisé pour plusieurs configurations des réseaux de neurones.

| Type de tissus identifié par le réseau de neurones | | Nombre de neurones cachés nc | fc = logsig fs = purelin | fc = purelin fs = logsig | fc = logsig fs = tansig |
|--|-------------------------|------------------------------|-----------------------------|-----------------------------|----------------------------|
| RNA spécialisé dans l'identification de tissu mammaire du type grassex (F). | | 5 | 33,33 % | 91,67 % | 25,00 % |
| | | 15 | 58,33 % | 83,33 % | 8,33 % |
| | | 25 | 25,00 % | 58,33 % | 8,33 % |
| | | 35 | 25,00 % | 83,33 % | 0,00 % |
| | | 45 | 16,67 % | 83,33 % | 0,00 % |
| | | 50 | 8,33 % | 91,67 % | 8,33 % |
| RNA spécialisé dans l'identification de tissu mammaire du type grassex-glandulaire (G) et dense (D). | Grassex-glandulaire (G) | 2 | 00,00 % | 83,33% | 0,00 % |
| | | 12 | 8,33 % | 33,33% | 8,33 % |
| | | 22 | 16,67 % | 58,33 % | 8,33 % |
| | | 32 | 8,33 % | 91,67 % | 0,00 % |
| | | 42 | 16,67% | 83,33 % | 0,00 % |
| | | 50 | 8,33% | 50,00 % | 0,00 % |
| | Dense (D) | 2 | 33,33 % | 58,33 % | 8,33% |
| | | 12 | 16,67 % | 41,67 % | 83,33 % |
| | | 22 | 8,33 % | 41,67 % | 83,33 % |
| | | 32 | 0,00 % | 83,33 % | 75,00 % |
| | | 42 | 16,67% | 50,00 % | 66,67 % |
| | | 50 | 8,33% | 66,67 % | 66,67 % |

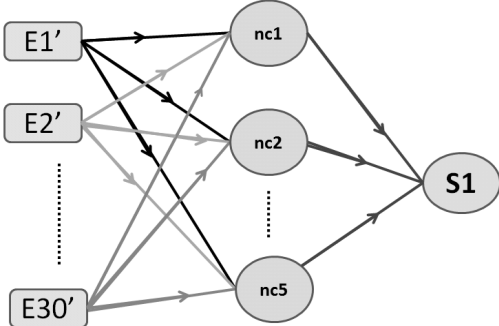
* fc désigne la fonction de transfert de la couche cachée.

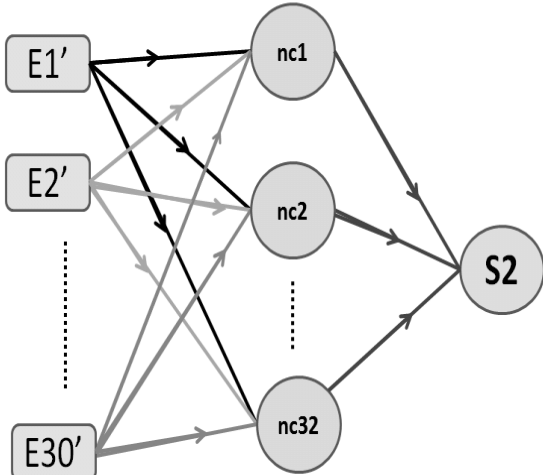
* fs désigne la fonction de transfert de la couche de sortie.

1.1.1 Les réseaux de neurones sélectionnés pour la classification de la base mini- MIAS

Le tableau suivant indique les paramètres de chaque réseau de neurones sélectionné constituant le système de classification de la base mini-MIAS.

Tableau 4 .3 . Les paramètres des réseaux de neurones sélectionnés constituant le système de classification de la base mini-MIAS.

| N° d'étape | Type de tissu identifié par le réseau de neurones | Les paramètres du réseau de neurones sélectionné |
|------------|--|---|
| Étape 1 | Réseau de neurones spécialisé dans l'identification de tissu mammaire graisseux (F). | <ul style="list-style-type: none"> - nombre d'images utilisées NI = 60*3 types de tissus - nombre d'images d'apprentissage NA = 144 - nombre d'images testées NT = 36 - nombre de neurones d'entrée NE = 30 - nombre de neurones cachés NC = 5 - fonction de transfert de la couche cachée : Purelin - fonction de transfert de la couche de sortie : Logsig - un neurone de sortie : S1, telle que : <ul style="list-style-type: none"> * si $S1 = 0 \Rightarrow$ tissu graisseux (F) ; * si $S1 = 1 \Rightarrow$ autre type de tissu. <div style="text-align: center; margin-top: 20px;">  <p style="text-align: center;"> Couche d'entrée avec 30 neurones de E1' ...E30' Couche cachée avec 5 neurones de nc1....nc5 Fonction de transfert : Purelin Couche de sortie avec 1 neurone S1 Fonction de transfert : Logsig </p> </div> <p style="text-align: center; margin-top: 10px;">Figure 4.4 RNA spécialisé dans l'identification de tissu (F)</p> |

| | | |
|----------------|--|---|
| <p>Étape 2</p> | <p>Réseau de neurones spécialisé dans l'identification des tissus graisseux glandulaires (G) et les tissus denses (D).</p> | <ul style="list-style-type: none"> - nombre d'images utilisées NI = 60*2 types de tissus - nombre d'images d'apprentissage NA = 96 - nombre d'images testées NT = 24 - nombre de neurones d'entrée NE = 30 - nombre de neurones cachés NC = 32 - fonction de transfert de la couche cachée : Purelin - fonction de transfert de la couche de sortie : Logsig - un neurone de sortie : S2, telle que : <ul style="list-style-type: none"> * si $S2 = 0 \Rightarrow$ tissu graisseux glandulaire ; * si $S2 = 1 \Rightarrow$ tissu dense. <div style="text-align: center; margin: 20px 0;">  </div> <div style="display: flex; justify-content: space-around; text-align: center;"> <div data-bbox="735 1186 938 1297"> <p>Couche d'entrée avec 30 neurones de E1' ...E30'</p> </div> <div data-bbox="964 1186 1198 1325"> <p>Couche cachée avec 32 neurones de nc1....nc32 Fonction de transfert : Purelin</p> </div> <div data-bbox="1222 1186 1446 1325"> <p>Couche de sortie avec 1 neurone S2 Fonction de transfert : Logsig</p> </div> </div> <p>Figure 4.5 RNA spécialisé dans l'identification des tissus (G) et (D).</p> |
|----------------|--|---|

1.1.2 Résultats et comparaison pour la base mini-MIAS

Les mammographies utilisées pour évaluer notre méthode et les autres méthodes proviennent de la base mini-MIAS.

Tableau 4 .4 : comparaison des méthodes conformément à la classification de la base mini-MIAS.

| Réf – année Type de tissu | (Oli 62) | (Muh 56) | (Mus 57) | Méthode proposée |
|--|-----------------|-----------------|-----------------|-------------------------|
| F | 63,33 % | 80,16 % | 87,73 % | 91,67 % |
| G | 58,33 % | 74,71 % | 75,00 % | 91,67 % |
| D | 66,66 % | 77,00 % | 83,50 % | 83,33 % |
| TR global | 62,77 % | 77,57 % | 82.50 % | 88,89 % |

Le taux de reconnaissance (TR) global n'est pas toujours la moyenne des taux de chaque type de tissu, puisque le taux global dépend du nombre de tissus de chaque type utilisé dans le test (Bou 09).

Le système de classification de la densité des tissus mammaires est simple à élaborer, puisqu'il utilise des réseaux de neurones à une seule couche cachée. L'entrée de chaque réseau utilise seulement les données issus des histogrammes (Bou 09).

De plus, la classification de la densité des tissus mammaires s'applique sur des bases engendrées par des campagnes de dépistage du cancer du sein,

ce qui entraîne la production d'une masse importante d'informations à traiter. Ceci inique que la simplicité de la méthode est un facteur important puisqu'il permet de réduire le temps de calcul (Bou 09).

Notre système de classification est composé de plusieurs réseaux de neurones et chaque réseau est spécialisé dans l'identification d'un seul type de tissu mammaire ou deux conformément à la classification mini-MIAS. Cette spécialisation offre une possibilité d'adaptation de notre approche aux différents types de tissus mammaires (Bou 09).

Cette spécialisation explique la différence des résultats entre la méthode proposée et les autres méthodes (Bou 09).

L'observation du tableau (4 .4), nous permet de constater la performance de la méthode proposée pour chaque type de tissu mammaire.

L'application du système de classification de tissus mammaire de la Figure (4.1) sur la base DDSM est possible, mais il faut adapter le nombre de réseau de neurone au nombre de type de tissu conformément à la classification DDSM (Bou 09)-Annexe A.

Chapitre 5

**Détection automatique de la
présence d'anomalie dans le tissu
mammaire**

1. Détection d'anomalie dans le tissu mammaire

Dans cette partie, nous allons étudier la problématique de la détection de la présence ou non d'anomalie dans le tissu mammaire en utilisant un modèle de calcul à base de réseau de neurone artificiel. Les radiologues spécialistes détectent autour de 70 % des cas de cancers (Ral 72). La massification du dépistage, la diminution du nombre de radiologues spécialistes et les exigences de qualité en matière de santé publique rendent indispensable le recours à des techniques capables d'aider les radiologues à traiter plus de cas tout en améliorant les performances en matière de détection. Les Systèmes d'Aide au Diagnostic (SAD) par ordinateur répondent à ce besoin, car ils permettent à un radiologue même débutant d'améliorer ces performances.

Un système de détection de présence d'anomalie mammaire représente, un Système d'Aide au Diagnostic (SAD) par ordinateur, ce system s'avérerait un outil précieux, surtout dans un contexte de dépistage. Le nombre d'images à traiter, en particulier dans les cas de dépistage systématique rend particulièrement intéressant le recours à des techniques automatisées.

Pour atteindre cet objectif on a utilisé une base d'images, sur ces images on a effectué un prétraitement pour l'extraction des paramètres descriptifs de chaque mammographie. Ces paramètres seront injectés dans un réseau de neurones artificiel spécialisé pour chaque type de tissu. Ces réseaux ont pour fonction la classification d'une mammographie en deux catégories :

- Mammographie normale ;
- Mammographie avec anomalie.

Tous les radiologues reconnaissent la difficulté de l'examen d'une mammographie qui s'accroît encore par le type de tissus du sein examiné, pour cette raison notre système de détection d'anomalie sera composé de plusieurs réseaux de neurones chaque réseau a pour objectif de détecter la présence d'anomalie pour un seul type de tissu mammaire, ces réseaux fonctionnent de manière séparée. Le système de détection bénéficie des résultats de prétraitement exposés dans le chapitre 3 et du système de classification des tissus mammaires étalés dans le chapitre 4.

2. Conception du système de détection

Dans ce travail on applique des techniques capables de détecter la présence ou non d'anomalie dans le tissu mammaire, la base mini MIAS regroupe trois catégories de tissus mammaires, 60 images sont utilisées pour chaque type de tissu de la base mini MIAS, donc le système de détection d'anomalie est réalisé par l'utilisation de 180 images.

Pour chaque type de tissu on répartit les mammographies en deux ensembles le premier ensemble constitué de 48 images est utilisé pour l'apprentissage du réseau le deuxième ensemble de 12 images sert au test du réseau. Voici les étapes à suivre pour la conception d'un système de détection de présence d'anomalie mammaire à base de réseau de neurones:

Etape 1. Division de l'image en blocs

Au début chaque image est subdivisée en blocs de $n \times n$ pixels. Chacun de ces blocs représente une partie de l'image. Cette subdivision nous permet de parcourir l'ensemble des pixels de chaque image (Bou 10). Cette méthode qui consiste à diviser l'image en sous-images par un nombre de rectangles égaux a été appliquée par plusieurs chercheurs (Kha 46) (Gon 38).

3 tailles différentes des blocs sont testé, afin d'observer l'influence de la taille des sous-images sur la performance du système d'aide au diagnostic à base de réseau de neurones (Bou 10) figure (5.1).

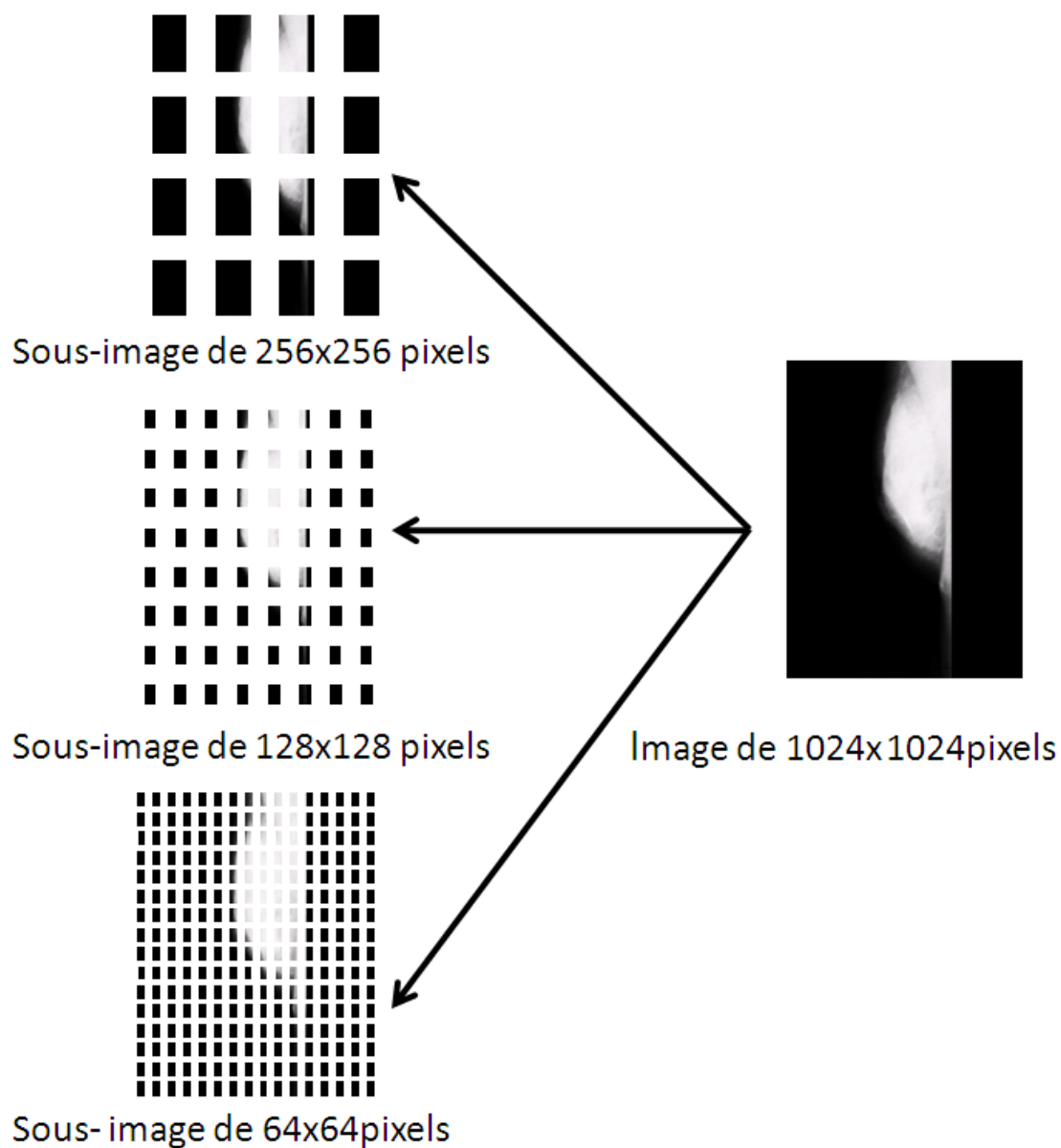


Figure 5.1 Mammographie subdivisé en sous-image (Bou 10)

Etape 2. Calcule de la moyenne

Ensuite on effectue le calcul des paramètres associés à cet ensemble de sous-images. À partir de chaque sous-image on calcule la valeur moyenne du niveau de gris \bar{I} .

$$\bar{I} = \frac{1}{l * c} \sum_{ij} I(i, j) \quad (5,1)$$

l et c représente respectivement le nombre de ligne et de colonne de chaque sous image, $I(i, j)$ représente la valeur du niveau de gris du pixel (i, j) tel que i varie de '1' jusqu'à ' l ' et j varie de '1' jusqu'à ' c '.

A ce niveau on obtient pour chaque image un nombre k de paramètres descriptifs, ce nombre varie suivant la taille de sous-images utilisé, k représente aussi le nombre de sous-images (Bou 10).

Tableau 5.1 Nombre de paramètres associé à chaque taille de sous-image (Bou 10)

| <i>Taille des sous-images</i> | <i>K nombre de paramètres</i> |
|-------------------------------|-------------------------------|
| <i>256x256 pixels</i> | <i>16</i> |
| <i>128x128 pixels</i> | <i>64</i> |
| <i>64x64pixels</i> | <i>256</i> |

Etape 3. Configuration du réseau de neurones

Le réseau de neurones de notre application est composé de trois couches. La couche d'entrée du réseau sert à intégrer des paramètres descriptifs des mammographies, la couche de sortie permet d'identifier la présence ou non d'anomalie (Bou 10).

Le rôle de la couche cachée est d'effectuer des traitements sur les données d'entrée en utilisant des fonctions de transfert et un nombre de neurone caché

nécessaire afin d'atteindre ou approcher l'objectif voulu, ce traitement nous permet d'extraire des données qui seront par la suite utilisées par la couche de sortie. Le nombre de neurones en couche d'entrée dépend du nombre de paramètres descriptifs utilisés, puisqu'on a calculé la valeur moyenne des niveaux de gris de chaque sous-image donc on doit utiliser k neurones d'entrées : $E_1 \dots E_k$ (Bou 10).

Le nombre de neurones de sortie est fixé à une seule sortie "S", puisqu'on cherche à détecter la présence ou non d'anomalie sur chaque image (Bou 10), tel que :

- $S=1 \rightarrow$ Présence d'anomalie.
- $S=0 \rightarrow$ Absence d'anomalie.

Le nombre de neurones en couche cachée et les fonctions d'activation ont été choisies expérimentalement ; on a opté pour le modèle le plus performant.

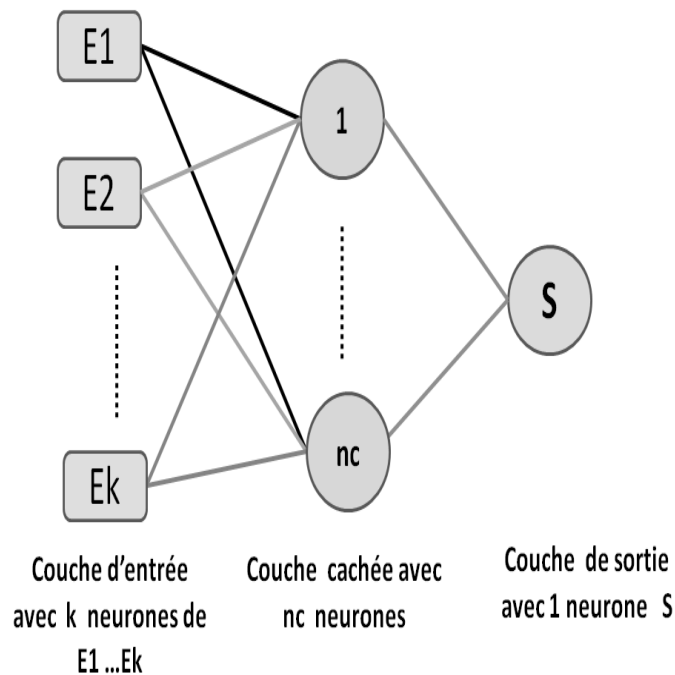


Figure 5.2 Conception du réseau de neurones spécialisé dans la détection d'anomalie sur une mammographie (Bou 10)

"nc" Indique le nombre de neurones en couche cachée. Le réseau a été expérimenté en variant "nc" de 1 à 15, figure (5.2).

Trois fonctions de transfert seront envisagées : deux fonctions sigmoïdes, l'une à sortie positive et négative 'tansig'. L'autre à sortie uniquement positive 'logsig', et une fonction linéaire 'purelin'.

Etape 4. Critères de performance du réseau de neurones

Afin d'observer les performances de chaque réseau de neurones, on compare les états calculés par le réseau de neurones avec les états réel. La performance est calculée en pourcentage. Les tests sont faits sur 12 images, ces images ne doivent pas être les mêmes utilisées dans l'apprentissage.

Le taux de reconnaissance est défini par la relation (4,2) dans le chapitre 4. On choisie le réseau le moins complexe en cas d'égalité des performances.

Etape 5. Apprentissage du réseau

On effectue un apprentissage supervisé pour chaque configuration, avec les paramètres suivants :

- * Nombre maximum d'itérations (Epochs) = 100.
- * Gradient minimum = $1 \text{ e-}10$.

L'apprentissage est arrêté si l'une de ces conditions est satisfaite. On a utilisé 48 images pour l'apprentissage du réseau et 12 images pour le test du réseau pour chaque type de tissu.

Etape 6 . Expérimentation des différents réseaux de neurones :

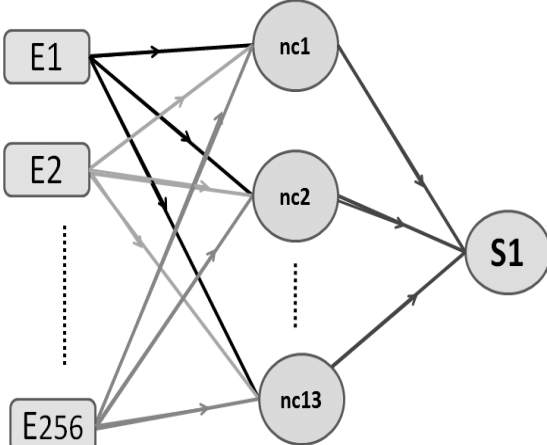
2.1 Détection de la présence d'anomalie dans un tissu mammaire de type grasseux

Pour une taille des sous images de 64x64pixels et 256 valeurs d'entrées on obtient les tableaux suivants :

Tableau 5 .2 Des exemples de taux de détection d'anomalies dans les tissus mammaires *grasseux* (F) réalisé pour plusieurs configurations des réseaux de neurones avec 256 entrées

| | Nombre de neurones cachés nc | fc = logsig fs = purelin | fc = purelin fs = logsig | fc = logsig fs = tansig |
|--|-------------------------------------|-------------------------------------|-------------------------------------|------------------------------------|
| RNA spécialisé dans la détection d'anomalies sur les tissus mammaires du type grasseux (F). | 3 | 16,67 % | 50,00 % | 25,00 % |
| | 5 | 00,00 % | 41,67 % | 25,00 % |
| | 7 | 25,00 % | 50,00 % | 25,00 % |
| | 10 | 41,67 % | 50,00 % | 08,33 % |
| | 13 | 16,67 % | 66,67 % | 16,67 % |
| | 15 | 16,67% | 50,00% | 16,67 % |

Tableau 5.3 Les paramètres du réseau de neurones sélectionné pour la détection de présence d'anomalie dans un tissu mammaire de type gras (F) avec 256 entrées

| Fonction du réseau de neurones | Les paramètres du réseau de neurones sélectionné | | | |
|---|--|--|--|--|
| <p>Réseau de neurones spécialisé dans la détection de présence d'anomalie dans un tissu mammaire de type gras (F)</p> | <ul style="list-style-type: none"> - nombre d'images utilisées NI = 60 mammographies pour tissu mammaire de type de <i>gras</i> (F) - nombre d'images d'apprentissage NA = 48 - nombre d'images testées NT = 12 - nombre de neurones d'entrée NE = 256 - nombre de neurones cachés NC = 13 - fonction de transfert de la couche cachée : Purelin - fonction de transfert de la couche de sortie : Logsig - un neurone de sortie : S1, telle que : <ul style="list-style-type: none"> * si $S1 = 0 \Rightarrow$ <i>Absence d'anomalie</i>; * si $S1 = 1 \Rightarrow$ <i>Présence d'anomalie</i>. <div style="text-align: center; margin: 20px 0;">  </div> <table style="width: 100%; border: none; margin: 0 auto;"> <tr> <td style="text-align: center; width: 33%;">Couche d'entrée avec 256 neurones de E1 ...E256</td> <td style="text-align: center; width: 33%;">Couche cachée avec 13 neurones de nc1....nc13 Fonction de transfert : Purelin</td> <td style="text-align: center; width: 33%;">Couche de sortie avec 1 neurone S1 Fonction de transfert : Logsig</td> </tr> </table> <p style="text-align: center; margin-top: 20px;">Figure 5.3 RNA spécialisé dans la détection de présence d'anomalie dans les tissus gras (F).</p> | Couche d'entrée avec 256 neurones de E1 ...E256 | Couche cachée avec 13 neurones de nc1....nc13 Fonction de transfert : Purelin | Couche de sortie avec 1 neurone S1 Fonction de transfert : Logsig |
| Couche d'entrée avec 256 neurones de E1 ...E256 | Couche cachée avec 13 neurones de nc1....nc13 Fonction de transfert : Purelin | Couche de sortie avec 1 neurone S1 Fonction de transfert : Logsig | | |

Le tableau suivant nous permet d'observer l'influence de la taille des sous-images sur la performance du réseau de neurone sélectionné pour la détection de la présence d'anomalie *dans les tissus gras* F.

Tableau 5.4 Taux de détection de présence d'anomalie dans les tissus gras F associé à chaque taille de sous-image

| Taille des sous-images | K nombre de paramètres d'entrées | Taux de reconnaissance |
|-------------------------------|---|-------------------------------|
| <i>256x256 pixels</i> | <i>16</i> | <i>50 %</i> |
| <i>128x128 pixels</i> | <i>64</i> | <i>50 %</i> |
| <i>64x64 pixels</i> | <i>256</i> | 66,67% |

On observe du tableau 5.4 que la diminution de la taille des sous-images engendre une augmentation du taux de reconnaissance. Le meilleur taux de reconnaissance obtenu est de 66.67 % réalisé avec une taille de sous image de 64x64 pixels.

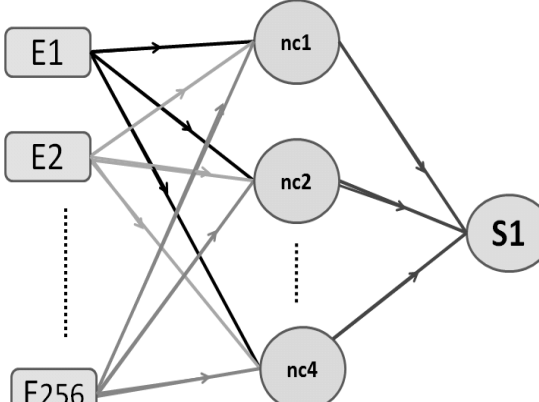
2. 2 Détection de la présence d'anomalie dans un tissu mammaire de type grasseux-glandulaire

Pour une taille des sous images de 64x64 pixels et 256 valeurs d'entrées on obtient les tableaux suivants :

Tableau 5.5 Des exemples de taux de détection d'anomalies dans les tissus mammaires *grasseux-glandulaire (G)* réalisé pour plusieurs configurations des réseaux de neurones avec 256 entrées

| | Nombre de neurones cachés nc | fc = logsig fs = purelin | fc = purelin fs = logsig | fc = logsig fs = tansig |
|---|---|-------------------------------------|-------------------------------------|------------------------------------|
| RNA spécialisé dans la détection d'anomalies Sur les tissus mammaires <i>grasseux- glandulaire (G)</i> . | 1 | 58,33 % | 50,00 % | 50,00 % |
| | 4 | 41,67 % | 83,33 % | 50,00 % |
| | 7 | 16,67 % | 50,00 % | 41,67 % |
| | 10 | 00,00 % | 50,00 % | 33,33 % |
| | 13 | 16,67% | 41,67 % | 25,00 % |
| | 15 | 00,00 % | 41,67 % | 16,67 % |

Tableau 5.6 Les paramètres du réseau de neurones sélectionné pour la détection de présence d'anomalie dans un tissu mammaire de type graineux-glandulaire (G) avec 256 entrées

| Fonction du réseau de neurones | Les paramètres du réseau de neurones sélectionné |
|--|--|
| <p>Réseau de neurones spécialisé dans la détection de présence d'anomalie dans un tissu mammaire de type <i>graineux-glandulaire (G)</i></p> | <ul style="list-style-type: none"> - nombre d'images utilisées NI = 60 mammographies pour tissu mammaire de type <i>graineux-glandulaire (G)</i> - nombre d'images d'apprentissage NA = 48 - nombre d'images testées NT = 12 - nombre de neurones d'entrée NE = 256 - nombre de neurones cachés NC = 4 - fonction de transfert de la couche cachée : Purelin - fonction de transfert de la couche de sortie : Logsig - un neurone de sortie : S1, telle que : <ul style="list-style-type: none"> * si $S1 = 0 \Rightarrow$ Absence d'anomalie; * si $S1 = 1 \Rightarrow$ Présence d'anomalie. <div style="text-align: center; margin-top: 20px;">  <p style="display: flex; justify-content: space-around; margin-top: 10px;"> Couche d'entrée avec 256 neurones de E1 ...E256 Couche cachée avec 4 neurones de nc1...nc4 Fonction de transfert : Purelin Couche de sortie avec 1 neurone S1 Fonction de transfert : Logsig </p> </div> <p style="text-align: center; margin-top: 20px;">Figure 5.4 RNA spécialisé dans la détection de présence d'anomalie dans les tissus graineux-glandulaire (G).</p> |

Le tableau suivant nous permet d'observer l'influence de la taille des sous-images sur la performance du réseau de neurone sélectionné pour la détection de la présence d'anomalie *dans les tissus* gras-seux-glandulaire

Tableau 5.7 Taux de détection de présence d'anomalie *dans les tissus* gras-seux-glandulaire (G) associé à chaque taille de sous-image

| <i>Taille des sous-images</i> | <i>K nombre de paramètres d'entrées</i> | <i>Taux de reconnaissance</i> |
|-------------------------------|---|-------------------------------|
| <i>256x256 pixels</i> | <i>16</i> | <i>41,67 %</i> |
| <i>128x128 pixels</i> | <i>64</i> | <i>50,00 %</i> |
| <i>64x64 pixels</i> | <i>256</i> | <i>83,33 %</i> |

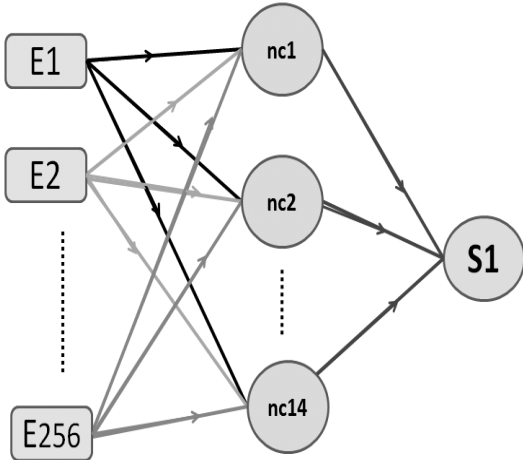
2. 3 Détection de la présence d'anomalie dans un tissu mammaire de type dense

Pour une taille des sous images de 64x64 pixels et 256 valeurs d'entrées on obtient les tableaux suivants :

Tableau 5 .8 Des exemples de taux de détection d'anomalies dans les tissus mammaires dense (D) réalisé pour plusieurs configurations des réseaux de neurones avec 256 entrées

| | Nombre de neurones cachés nc | fc = logsig fs = purelin | fc = purelin fs = logsig | fc = logsig fs = tansig |
|---|-------------------------------------|-------------------------------------|-------------------------------------|------------------------------------|
| RNA spécialisé dans la détection d'anomalies sur les tissus mammaires dense (<i>Dense D</i>). | 1 | 25,00 % | 50,00 % | 33,33 % |
| | 4 | 00,00 % | 50,00 % | 00,00 % |
| | 7 | 00,00 % | 50,00 % | 33,33 % |
| | 10 | 00,00 % | 50,00 % | 08,33 % |
| | 14 | 00,00 % | 75,00 % | 16,67 % |
| | 15 | 16,67 % | 50,00 % | 25,00 % |

Tableau 5.9 Les paramètres du réseau de neurones sélectionné pour la *détection de présence d'anomalie dans un tissu mammaire de type dense (Dense D)*.avec 256 entrées

| <p>Fonction du réseau de neurones</p> | <p>Les paramètres du réseau de neurones sélectionné</p> |
|--|---|
| <p>Réseau de neurones spécialisé dans la détection de présence d'anomalie dans un tissu mammaire de type <i>dense (D)</i>.</p> | <ul style="list-style-type: none"> - nombre d'images utilisées NI = 60 mammographies pour tissu mammaire de type <i>dense (D)</i> - nombre d'images d'apprentissage NA = 48 - nombre d'images testées NT = 12 - nombre de neurones d'entrée NE = 256 - nombre de neurones cachés NC = 14 - fonction de transfert de la couche cachée : Purelin - fonction de transfert de la couche de sortie : Logsig - un neurone de sortie : S1, telle que : <ul style="list-style-type: none"> * si $S1 = 0 \Rightarrow$ <i>Absence d'anomalie;</i> * si $S1 = 1 \Rightarrow$ <i>Présence d'anomalie.</i> <div style="text-align: center; margin: 20px 0;">  </div> <div style="display: flex; justify-content: space-around; font-size: small;"> <div style="text-align: center;"> <p>Couche d'entrée avec 256 neurones de E1 ...E256</p> </div> <div style="text-align: center;"> <p>Couche cachée avec 14 neurones de nc1....nc14 Fonction de transfert : Purelin</p> </div> <div style="text-align: center;"> <p>Couche de sortie avec 1 neurone S1 Fonction de transfert : Logsig</p> </div> </div> <p>Figure 5.5 RNA spécialisé dans la détection de présence d'anomalie dans les tissus dense (D).</p> |

Le tableau suivant nous permet d'observer l'influence de la taille des sous-images sur la performance du réseau de neurone sélectionné pour la détection de la présence d'anomalie *dans les tissus de type dense* (Dense D).

Tableau 5.10 Taux de détection de présence d'anomalie *dans les tissus dense* (Dense D) associé à chaque taille de sous-image

| Taille des sous-images | K nombre de paramètres d'entrées | Taux de reconnaissance |
|------------------------|----------------------------------|------------------------|
| <i>256x256 pixels</i> | <i>16</i> | <i>50,00 %</i> |
| <i>128x128 pixels</i> | <i>64</i> | <i>58,33 %</i> |
| <i>64x64 pixels</i> | <i>256</i> | <i>75,00%</i> |

3. avantages du système de détection

Notre système à base de réseau de neurone artificiel possède les avantages suivants:

- Simplicité du calcul des paramètres descriptifs de chaque mammographie.
- Détection automatique et rapide de la présence ou non dans un tissu mammaire par le réseau de neurone artificiel.
- Possibilité d'adapter le système à d'autres images provenant de différentes bases médicales avec différentes résolutions.

L'essai du système avec une taille de sous-image inférieure à 64x64 pixels pénalise le délai d'apprentissage du réseau de neurones. La base de données doit être également enrichie pour couvrir une vaste gamme de cas.

Notre étude nous a permis de développer un réseau de neurone capable de détecter la présence ou non d'anomalie pour chaque type de tissu avec un taux de réussite important, ce modèle représente un intérêt, en particulier dans les cas de dépistage systématique ce qui engendre un nombre important d'image à interpréter.

Ce système permet au radiologue de concentrer son attention sur les cas les plus difficiles.

De plus, cette étude a démontré l'influence du choix de la taille des sous images et le nombre de neurones, les fonctions de transferts utilisées sur l'efficacité du réseau de neurones artificiel.

Conclusion

Cette étude a exposé la possibilité de l'extraction de la zone d'intérêt d'une mammographie par l'utilisation de plusieurs techniques de traitement d'image numérique. On a utilisé la méthode du seuil d'Otsu, la binarisation et la labellisation, pour supprimer les artefacts.

La deuxième phase a été consacrée à l'application des réseaux de neurones artificiels, dans la classification automatique de la densité des tissus mammaires.

Le système de classification est développé grâce à l'exploitation des indicateurs issus des histogrammes des images en niveaux de gris par des réseaux de neurones artificiels. Ces indicateurs ont l'avantage d'être d'une grande simplicité d'emploi.

Notre système de classification est composé de plusieurs réseaux de neurones artificiels. Chaque réseau du système de classification est spécialisé dans la reconnaissance, soit d'un seul type de tissu mammaire ou deux.

Les réseaux de neurones artificiels peuvent réaliser la détection de la présence ou non d'anomalie sur une mammographie pour chaque type de tissu avec un taux de reconnaissance important. De plus, cette étude a démontré l'influence du choix de la taille des sous images et le nombre de neurones, les fonctions de transferts utilisées sur l'efficacité du réseau de neurones artificiels pour la détection d'anomalies.

Les réseaux de neurones artificiels peuvent être utilisés pour la conception des Systèmes d'Aide au Diagnostic (SAD) par ordinateur.

Le principal avantage des réseaux de neurones artificiels réside dans:

- Leur capacité d'apprentissage automatique, ce qui permet de résoudre des problèmes sans nécessiter d'écriture de règles complexes.

Les résultats expérimentaux montrent l'efficacité de notre système proposé. En perspective d'autres bases de données de mammographie seront testées.

Annexe A

Le système de classification de la densité des tissus mammaires de la base DDSM est composé de plusieurs étapes successives, chaque étape utilise un réseau de neurones (Bou 09) :

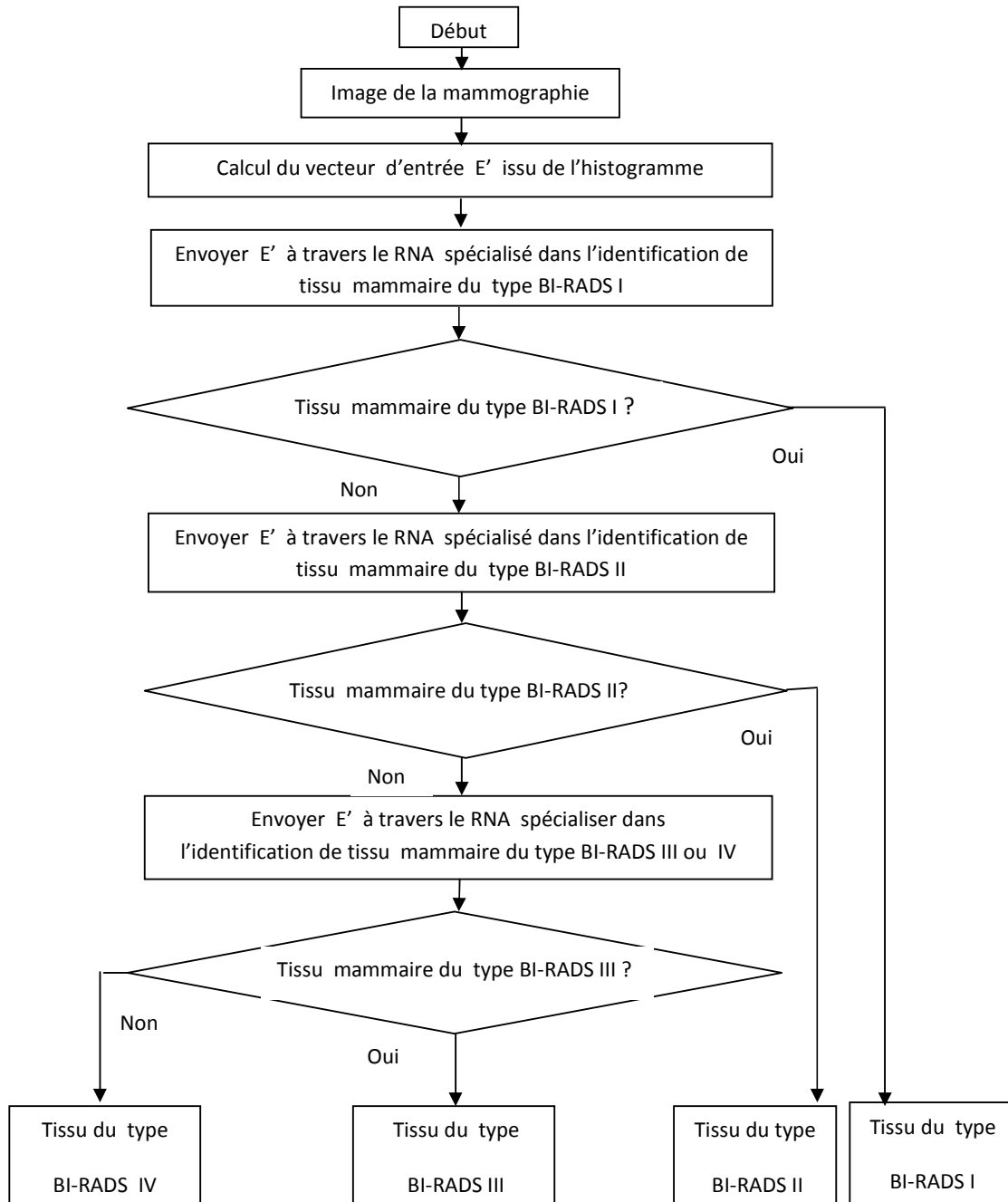


Figure A .1 Organigramme du système de classification de la densité des tissus mammaires conformément à la classification BI-RADS (Bou 09).

Annexe B

Tableau B.1 Algorithme de Levenberg-Marquardt (Moh 58)

-
1. Présenter les entrées $\vec{x}(n)$ au réseau, calculer les sorties correspondantes, le vecteur d'erreur $\vec{e}(n)$ et le calculer la fonction de coût

$$\zeta_n(\vec{w}) = \sum_{k=1}^N e_k^{(n)}(\vec{w})$$

-
2. Calculer la matrice jacobienne $J_n(\vec{w})$

-
3. Mettre à jour les poids :

$$\vec{w}_{n+1} = \vec{w}_n - (J_n^T(\vec{w}_n)J_n(\vec{w}_n) + \mu_n I)J_n^T(\vec{w}_n)\zeta_n(\vec{w}_n)$$

-
4. Calculer : $\zeta_n(\vec{w}_{n+1}) = \sum_{k=1}^N e_k^{(n)}(\vec{w}_{n+1})$:

Si $\zeta_n(\vec{w}_{n+1}) < \zeta_n(\vec{w}_n)$ alors $\mu_{n+1} = \mu_{n+1} - \nu$, où ν est une constante, puis retourner à l'étape 1

Si $\zeta_n(\vec{w}_{n+1}) > \zeta_n(\vec{w}_n)$ alors $\mu_{n+1} = \mu_{n+1} + \nu$, puis retourner à l'étape 3 pour mettre à jour μ_{n+1} .

-
5. Itération des étapes 2 à 4 jusqu'à avoir rencontré un critère d'arrêt.
-

Références

- (Ahm 01) AHMAD, Zainal et ROSLIN, Fairuoze. Modeling of real pH Neutralization Process using Multiple Neural Networks (MNN) Combination Technique. *structure*, 2007, vol. 19, p. 20.
- (Aki 02) AKILA, K., JAYASHREE, L. S., et VASUKI, A. Mammographic image enhancement using indirect contrast enhancement techniques—a comparative study. *Procedia Computer Science*, 2015, vol. 47, p. 255-261.
- (Aub 03) AUBREE Loïs, "Optimisation d'algorithmes de traitement d'images sur une architecture utilisant le GPGPU,"Rapport de stage ST40, L'université de technologie Belfort-Montbéliard, 2013.
- (Aou 04) AOUCHICHE, Salima, BAKLI, Lila, et MEBAREK, Bouchaala Fatheh Encadreur. L'anxiété chez les femmes atteintes de cancer du sein: étude de six cas à CHU de Bejaia. 2017. Thèse de doctorat. Université de Bejaia.
- (Aol 05) ALOULOU, Sofia, MAHFOUDI, A. E., OMRANI, A. E., et al. Facteurs liés au diagnostic tardif du cancer du sein: expérience du CHU Mohammed VI Marrakech. *Pan African Medical Journal*, 2015, vol. 21, n°1.
- (Bal 06) BALLEYGUIER, C., BOYER, B., ATHANASIOU, A., et al. Comprendre et utiliser le CAD (Aide informatisée au diagnostic) en mammographie. *Journal de Radiologie*, 2005, vol. 86, no 1, p. 29-35.
- (Bar 07) BARRON, Andrew R. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Transactions on Information theory*, 1993, vol. 39, no 3, p. 930-945.
- (Ber 08) BERGOUNIOUX, Maitine. Quelques Méthodes de Filtrage en Traitement d'image. 2011.
- (Bou 09) BOUKHOBZA, Mohamed El Habib et MIMI, Malika. Automatic classification of breast tissue density. *TRAITEMENT DU SIGNAL*, 2016, vol. 33, no 4, p. 441-459.
- (Bou 10) BOUKHOBZA, M. et MIMI, M. Detection automatique de la presence d'anomalie sur une mammographie par réseau de neurones artificiels. *courier du savoir*, 2012, no 3, p. 103-108.

- (Bos 11) BOSS, R., THANGAVEL, K., et DANIEL, D. Automatic mammogram image breast region extraction and removal of pectoral muscle. arXiv preprint arXiv:1307.7474, 2013.
- (Bov 12) BOUYEYRON, Charles. Modélisation et classification des données de grande dimension: application à l'analyse d'images. 2006. Thèse de doctorat. Université Joseph-Fourier-Grenoble I.
- (Bra 13) BRAILLON, Alain. Correspondance à propos de l'article «Dépistage du cancer du sein». Des conclusions... aux actes!. La Presse médicale, 2014, vol. 43, no 11.
- (Byd 14) BOYD, N. F., BYNG, J. W., JONG, R. A., et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. JNCI: Journal of the National Cancer Institute, 1995, vol. 87, no 9, p. 670-675.
- (Byd 15) BOYD, Norman F., GUO, Helen, MARTIN, Lisa J., et al. Mammographic density and the risk and detection of breast cancer. New England Journal of Medicine, 2007, vol. 356, no 3, p. 227-236.
- (Byn 16) BYNG, Jeffrey William, BOYD, N. F., FISHELL, E., et al. The quantitative analysis of mammographic densities. Physics in Medicine & Biology, 1994, vol. 39, no 10, p. 1629.
- (Cha 17) CHERIF, Aymen. Réseaux de neurones, SVM et approches locales pour la prévision de séries temporelles. 2013. Thèse de doctorat. Tours.
- (Che 18) CHERIF, Mounia. Capacité d'une mémoire associative à fonction de sortie chaotique. 2010.
- (Che 19) CHEN, Zhili, OLIVER, Arnau, DENTON, Erika, et al. Automated mammographic risk classification based on breast density estimation. In : Iberian Conference on Pattern Recognition and Image Analysis. Springer, Berlin, Heidelberg, 2013. p. 237-244.
- (Chk 20) CHEIKHROUHOU, Imen. Description et classification des masses mammaires pour le diagnostic du cancer du sein. 2012. Thèse de doctorat. Université d'Evry-Val d'Essonne.
- (Ckh 21) C.Khampheng, S.Taupiac, L'avenir de la radiologie analogique 2000, Projet DESS "TBH.
- (Cor 22) CORBILLON, E., POUILLIÉ, A., BLONDET, E., & MISSOUR, S. Guide méthodologique: comment évaluer a priori un programme de dépistage 2004. . Agence nationale d'accréditation et d'évaluation en santé (ANAES), 1-68.

- (Cyb 23) CYBENKO, George. Approximation by superpositions of a sigmoidal function. Mathematics of control, signals and systems, 1989, vol. 2, no 4, p. 303-314.
- (Deh 24) DEGHANI, Sara et DEZFOOLI, Mashallah Abbasi. A method for improve preprocessing images mammography. International Journal of Information and Education Technology, 2011, vol. 1, no 1, p. 90.
- (Dgs 25) Direction Générale des Structures de Santé., Manuel de Prise en Charge Du Cancer du Sein Février 2016, Algérie.
- (Doi 26) DOI, Kunio, MACMAHON, Heber, KATSURAGAWA, Shigehiko, et al. Computer-aided diagnosis in radiology: potential and pitfalls. European journal of Radiology, 1999, vol. 31, no 2, p. 97-109.
- (Doi 27) DOI, Kunio, GIGER, Maryellen L., NISHIKAWA, Robert M., et al. Computer aided diagnosis of breast cancer on mammograms. Breast Cancer, 1997, vol. 4, no 4, p. 228-233.
- (Dre 28) DREW, Philip J. et MONSON, John RT. Artificial neural networks. Surgery, 2000, vol. 127, no 1, p. 3-11.
- (Drf 29) DREYFUS, Gérard, MARTINEZ, Jean-Marc, SAMUELIDES, Manuel, et al. Réseaux de neurones-Méthodologie et applications. 2002.
- (Drf 30) DREYFUS, Gérard, et al. Les réseaux de neurones. Mécanique industrielle et matériaux, 1998, vol. 51.
- (Dub 31) Dubuisson, J Séverine. Bases du traitement des images Filtrage d'images. 22 octobre 2010.
- (Eda 32) EDDAOUDI, Fatima. Aide à la détection des pathologies du sein par les techniques de segmentation et codage de texture des images mammographiques. 2012.
- (Elk 33) ELKAOU Mina Evaluation du programme national de détection précoce du cancer du sein cas de la région du grand Casablanca.2014.
- (Fer 34) FERKOUS Chokri, et FARIDA, Merouani Hayet. Détection des Lésions dans les Mammographies avec les Contours Actifs.2010. Conference: Seminaire National d'Informatique SNIB'10, At Biskra Algérie, Volume: 7.
- (Fre 35) FREER, Timothy W. et ULISSEY, Michael J. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. Radiology, 2001, vol. 220, no 3, p. 781-786.

- (Gig 36) GIGER, Maryellen L., DOI, Kunio, MACMAHON, H., et al. An" intelligent" workstation for computer-aided diagnosis. *Radiographics*, 1993, vol. 13, no 3, p. 647-656.
- (Gok 37) Gokila. Deepa. G , "Mammogram Image Segmentation Using Fuzzy Hybrid with Particle Swarm Optimization (PSO) ," *International Journal of Engineering and Innovative Technology (IJEIT)* Volume 2, Issue 6, December 2012.
- (Gon 38) GONG, Yihong, CHUAN, Chua Hock, et XIAOYI, Guo. Image indexing and retrieval based on color histograms. *Multimedia Tools and Applications*, 1996, vol. 2, no 2, p. 133-156.
- (Hdc 39) HENG-Da, Cheng , CAI, Xiaopeng, CHEN, Xiaowei, et al. Computer-aided detection and classification of microcalcifications in mammograms: a survey. *Pattern recognition*, 2003, vol. 36, no 12, p. 2967-2991.
- (Hor 40) HORNIK, Kurt, STINCHCOMBE, Maxwell, et WHITE, Halbert. Multilayer feedforward networks are universal approximators. *Neural networks*, 1989, vol. 2, no 5, p. 359-366.
- (Hor 41) HORNIK, Kurt. Some new results on neural network approximation. *Neural networks*, 1993, vol. 6, no 8, p. 1069-1072.
- (Hor 42) HORNIK, Kurt, STINCHCOMBE, Maxwell, WHITE, Halbert, et al. Degree of approximation results for feedforward networks approximating unknown mappings and their derivatives. *Neural Computation*, 1994, vol. 6, no 6, p. 1262-1275.
- (Ici 43) icis 2005 Institut canadien d'information sur la santé. L'imagerie médicale au Canada 2005 , https://secure.cihi.ca/free_products/MedImag05_f.pdf.
- (Isi 44) http://www.isimg.rnu.tn/useruploads/cours/11352918809_traitement-des-images-chapitre1.pdf
- (Jod 45) JODOUIN, Jean-François. *Les Réseaux de neurones: Principes et définitions*. Hermes, 1994.
- (Kha 46) KHARBOUCHE, Said, VANNOORENBERGHE, Patrick, LECOMTE, Christèle, et al. Histogrammes spatiaux couleur optimisés pour l'indexation d'images par le contenu. In : 20° Colloque sur le traitement du signal et des images, FRA, 2005. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 2005.
- (Krl 47) KERLIKOWSKE, Karla. *The mammogram that cried Wolfe*. 2007.

- (Liy 48) LIYAKATHUNISA, C. N. et KUMAR, Ravi. A Novel and Efficient Lifting Scheme based Super Resolution Reconstruction for Early Detection of Cancer in Low resolution Mammogram Images. International Journal of Biometrics and Bioinformatics, 2011, vol. 5, n° 2.
- (Mai 49) MAITRA, Indra Kanta, NAG, Sanjay, et BANDYOPADHYAY, Samir Kumar. Technique for preprocessing of digital mammogram. Computer methods and programs in biomedicine, 2012, vol. 107, no 2, p. 175-188.
- (Mak 50) MAKANDAR, Aziz et HALALLI, Bhagirathi. Pre-processing of mammography image for early detection of breast cancer. Int J Comput Appl, 2016, p. 0975-8887.
- (Mat 51) Algorithmes pour le traitement de l'image Mathias Ortner 7 mai 2004/ ftp://ftp-sop.inria.fr/ariana/mortner/linfo/AI_trim_final.pdf
- (Mcc 52) MCCORMACK, Valerie A. et DOS SANTOS SILVA, Isabel. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiology and Prevention Biomarkers, 2006, vol. 15, no 6, p. 1159-1169.
- (Mej 53) MEJIA MANTILLA, Carlos. Architectures neuronales pour l'approximation des fonctions de transfert: application a la teledetection. 1992. Thèse de doctorat. Paris 11.
- (Min 54) MINA, Luqman Mahmood et ISA, Nor Ashidi Mat. Preprocessing technique for mammographic images. International Journal of Computer Science and Information Technology Research, 2014, vol. 2, p. 226-231.
- (Mit 55) MITCHELL, Tom M., et al. Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 1997, vol. 45, no 37, p. 870-877.
- (Muh 56) MUHIMMAH, Izzati et ZWIGGELAAR, Reyer. Mammographic density classification using multiresolution histogram information. In : Proceedings of the International Special Topic Conference on Information Technology in Biomedicine, ITAB. 2006.
- (Mus 57) MUŠTRA, Mario, GRGIĆ, Mislav, et DELAČ, Krešimir. Breast density classification using multiple feature selection. automatika, 2012, vol. 53, no 4, p. 362-372.
- (Moh 58) Mohamed Assaad HAMIDA. Introduction aux Méthodes de Contrôle Intelligent .2014 .https://elearn.univ-ouargla.dz/2013-2014/courses/INTRODUCTIONAUXMETHO/document/Cours_HAMIDA.pdf?cidReq=INTRODUCTIONAUXMETHO-

- (Mou 59) MOUTARDE, Fabien. Introduction aux réseaux de neurones. Ecole des Mines de Paris, 2007.
- (Nbe 60) N. Ben, HAMAD, BENROMDHANE, N., TAOUIL, Khaled, et al. Réduction des Faux Positifs dans les Systèmes d'Aide au Diagnostic du Cancer du Sein. In : SETIT 2007 4th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications, Tunisia. 2007.
- (Ngu 61) NGUYEN, Hanh H. et CHAN, Christine W. Multiple neural networks for a long term time series forecast. Neural Computing & Applications, 2004, vol. 13, no 1, p. 90-98.
- (Oli 62) OLIVER, Arnau, FREIXENET, Jordi, BOSCH, Anna, et al. Automatic classification of breast tissue. In : Iberian Conference on Pattern Recognition and Image Analysis. Springer, Berlin, Heidelberg, 2005. p. 431-438.
- (Oli 63) OLIVER, Arnau, FREIXENET, Jordi, et ZWIGGELAAR, Reyer. Automatic classification of breast density. In : Image Processing, 2005. ICIP 2005. IEEE International Conference on. IEEE, 2005. p. II-1258.
- (Otm 64) OTMANI, Imene. L'analyse discriminante et le perceptron multicouche.
- (Ots 65) OTSU, Nobuyuki. A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics, 1979, vol. 9, no 1, p. 62-66.
- (Paj 66) PARIKH, Jay. Digital mammography: current capabilities and obstacles. Journal of the American College of Radiology, 2005, vol. 2, no 9, p. 759-767.
- (Par 67) PARIZEAU, Marc. Réseaux de neurones. GIF-21140 et GIF-64326, 2004, vol. 124.
- (Pat 68) PATAN, Krzysztof. Artificial neural networks for the modelling and fault diagnosis of technical processes. Springer, 2008.
- (Ptr 69) PETROUDI, Styliani, KADIR, Timor, et BRADY, Michael. Automatic classification of mammographic parenchymal patterns: A statistical approach. In : Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE. IEEE, 2003. p. 798-801.
- (Pre 70) PREVRHAL, Sven, SHEPHERD, John A., SMITH-BINDMAN, Rebecca, et al. Accuracy of mammographic breast density analysis: results of formal operator training. Cancer Epidemiology and Prevention Biomarkers, 2002, vol. 11, no 11, p. 1389-1393.

- (Rah 71) RAHMATI, Peyman, HAMARNEH, Ghassan, NUSSBAUM, Doron, et al. A new preprocessing filter for digital mammograms. In : International Conference on Image and Signal Processing. Springer, Berlin, Heidelberg, 2010. p. 585-592.
- (Ral 72) RAHLI, H. S. et BENAMRANE, N. INTERPRETATION DES IMAGES DE MAMMOGRAPHIE PAR L'ALGORITHME SEARCH HARMONY. Annals. Computer Science Series, 2012, vol. 10, no 2.
- (Riv 73) RIVALS, Isabelle, PERSONNAZ, L., DREYFUS, G., et al. Modélisation, classification et commande par réseaux de neurones: principes fondamentaux, méthodologie de conception et illustrations industrielles. Les réseaux de neurones pour la modélisation et la commande de procédés, JP Corriou, ed.(Lavoisier Tec & Doc, 1995), 1995.
- (Sar 74) SARAH Hibo, Qu'est-ce que le cancer du sein ? Service Etudes du Secrétariat National des FPS 2010 02/515.17.67. http://doc.hubsante.org/doc_num.php?explnum_id=10349
- (Sav 75) SAVOYE, Gras. ALTERNATIVE NEURONALE EN TARIFICATION SANTE.
- (Sch 76) SCHETTER, Susann E. Breast density as an independent risk factor for cancer. J Am Osteopath Coll Radiol, 2014, vol. 3, p. 10-19.
- (Scm 77) SCHMITT, Aurore, LE BLANC, Benoit, CORSINI, M.-M., et al. Les réseaux de neurones artificiels. Un outil de traitement de données prometteur pour l'anthropologie. Bulletins et mémoires de la Société d'Anthropologie de Paris, 2001, no 13 (1-2).
- (Sun 78) SUNDARAVINAYAGAM, Subashini et SANKARI, S. Bhavani. Detection and classification of masses in mammograms using a hybrid GA-PSO-KNN Approach. IJARTET, 2015, vol. 2, no 5, p. 36-40.
- (Sri 79) SRIKARTHIK, A. et SIVAKUMAR, R. Preprocessing and Enhancement for Mammogram Images Using Unified Approach. International Journal of Computer & Organization Trends, 2013, vol. 3, no 3, p. 56-61.
- (Sti 80) STINÈS, Joseph. La densité mammaire: un concept radiologique., 26 journées de la SFSPM 2004, Nancy, p. 150 – 158.
- (Ser 81) Seror J Y. Quelles explorations chez une patiente à densité mammaire élevée ?.2005. Centre de Radiologie Duroc.

- (Suc 82) SUCKLING, John, PARKER, J., DANCE, D., et al. The mammographic image analysis society digital mammogram database. In : Exerpta Medica. International Congress Series. 1994. p. 375-378.
- (Ste 83) Steffen Nissen. Création d'un réseau de neurones – c'est facile, www.software20.org Software 2.0 No 2/2005
- (Sté 84) STÉPHANE, Tufféry. Data Mining et statistique décisionnelle: L'intelligence des données. Éditions Technip, 2012.
- (Sym 85) Symposium franco-algérien sur les nouvelles technologies du numérique pour l'aide au diagnostic et le dépistage des cancers du sein en Algérie - 17 et 18 janvier 2009 – Alger -http://www.santemaghreb.com/algerie/symposium_fenics2009.htm
- (Tar 86) Tardivon A. (2008). Quels risques, pour quelles femmes ? Densité mammaire et cancer du sein, 30 journées de la SFSPM 2008, La Baule, p. 93 – 101.
- (Ter 87) Terki. N et col. 2015. Incidence annuelle du cancer du sein dans la population Algérienne diagnostiquée au niveau des centres de référence d'anatomopathologie en Algérie. Le fascicule de la santé n°19.
- (Tol 88) TOULOUSE, Tom. Estimation par stéréovision multimodale de caractéristiques géométriques d'un feu de végétation en propagation. 2015. Thèse de doctorat. Université Pascal Paoli.
- (Toh 89) TOUHAMI, Sami. Classification non supervisée de pixels d'images couleur par analyse d'histogrammes tridimensionnels.
- (Tor 90) TORRENTS-BARRENA, Jordina, PUIG, Domenec, MELENDEZ, Jaime, et al. Computer-aided diagnosis of breast cancer via Gabor wavelet bank and binary-class SVM in mammographic images. Journal of Experimental & Theoretical Artificial Intelligence, 2016, vol. 28, no 1-2, p. 295-311.
- (Tou 91) TOUZET, Claude. Les réseaux de neurones artificiels, introduction au connexionnisme. EC2, 1992.
- (Tre 92) TREMBLAY, Guillaume. Optimisation d'ensembles de classifieurs non paramétriques avec apprentissage par représentation partielle de l'information. 2004. Thèse de doctorat. École de technologie supérieure.

- (Tzi 93) TZIKOPOULOS, Stylianos D., MAVROFORAKIS, Michael E., GEORGIU, Harris V., et al. A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry. computer methods and programs in biomedicine, 2011, vol. 102, no 1, p. 47-63.
- (Wul 94) Wulfram Gerstner, Benjamin Leroy-Beaulieu Sing Souksavanh Jean-Luc Beuchat . RESEAUX DE NEURONES ARTIFICIELS Une introduction a l'apprentissage automatique ,Laboratory of Computational Neuroscience (LCN) School of Computer and Communication Sciences 2006 (4eme edition), <https://epfl.wikispaces.com/file/view/ch1ch2.pdf>
- (Zwi 95) ZWINGELSTEIN, Gilles. Diagnostic des défaillances: théorie et pratique pour les systèmes industriels. Hermès, 1995.