UNIVERSITY KASDI MERBAH OUARGLA
Faculty of New Information Technologies and Communication
Department of computer science and Information Technology

ACADEMIC MASTER Thesis

Domain :Computer science and Information Technology
Track : Computer science
Specialty: Industrial

Directed by: Chaima DEROUICHE

Topic

Cancer Computing: The early detection of Breast Cancer

Supported publicly
the : ../06/2018
Before the jury:

Pr.	Mohamed Amine KHARFI	Professor	President	UKM Ouargla
Dr.	Akram BOUKHAMLA	MC(B)	Supervisor	UKM Ouargla
Dr.	Djalila BELKEBIR	MC(B)	Examiner	UKM Ouargla

College year : 2017/2018

Acknowledgment

I would first like to thank ALLAH the Beneficent, the Merciful. Praise be to ALLAH, Lord of the Worlds. ALLAH who guides me along the way.

As well, I would like to express my deepest gratitude to my supervisor Dr. Akram BOUKHAMLA for providing valuable comments, wisdom, encouragements, guidance, giving me the possibility to conduct my thesis the maximum freedom, and helping me to achieve my goal, which is publishing paper in master degree.

I would also like to thank Pr. Mohamed Amine Kharfi and Dr. Djalila BELKEBIR for accepting to be part of my thesis' jury

My sincerest appreciation to the teacher who defined me the image processing filed , who was supporting me to do my best during 5 years of the studying, Dr. Meriem KORICHI.

Special thanks to teacher Leila AMRAN; who created enthusiasm of studying computer science field.

Also, I appreciate the great aid and support from all the teachers of computer science and information department, Ouargla University.

Finally, I am indebted to thank my family for their continuous support, encouragement and understanding. Special thanks go to my mother, my guardian angel, for her support in all my life.

Dedication

To the soul of my father; the inspiration of my life

To my mother; the one who creates challenges for me

To my grand mother; who praised me over my success.

To my aunt; who supports me to do better in my future

To my brothers; greatest gift my parents have ever gave me.

To my family, everyone helped me during my master degree from near and far.

Abstract

Nowadays, breast cancer still the most common human death, however women are the most exposed, thus, its cause remains unknown. In order to reduce the associated morbidity and mortality, it would be necessary to detect this illness on its early stage.
Several systems have been investigated to detect and diagnosis breast cancer in order to help radiologists in the screening step. CADe and CADx are the well known systems used to achieve this task. However, various steps are needed to obtain accurate results which consist in preprocessing, detection of abnormalities, classifying these latter as malignant or benign.
After mammographic acquisition, some irrelevant elements must be removed from the mammogram such as Pectoral muscle, radiopaque artifacts, etc. For that reason we proposed, a novel method based on the similarity between intensities to delineate the pectoral muscle boundary using measure features of semantic similarity between words in Natural Language Process (NLP) and Information Retrieval (IR) fields. Also, the proposed method was exploited to classify the mammogram, based on its intensity, as normal or abnormal.
The obtained results were promising since our approach gives an effective pectoral muscle edge and extracts region with high intensity which improve the accuracy of our proposed system IOBI.

Keywords: *Breast cancer, Mammogram, NLP, IR, Pectoral muscle boundary, Intensity, CADe, CADx,IOBI.*

ملخص

لا يزال سرطان الثدي السبب الأكثر شيوعا للوفاة البشرية، إلا أنّ النساء هم أكثر عرضةً لهذا الداء. فبما أنّ لم يتم التعرف عن سبب وجود هذا الداء، كان لزاما أن يتم الكشف عن هذا الداء في مراحله الأولى بغية الحد من نتائجه الوخيمة.

تم اقتراح العديد من الأنظمة للكشف عن سرطان الثدي وتشخصيه، وذلك لمساعدة أخصائي الأشعة في مرحلة الفحص، حيث تعتبر أنظمة الكاد (CADe،CADx) الأكثر استعمالا لهذا الكشف.

الا أنّ، هناك عدة خطوات ضرورية للحصول على النتائج ذات الصلة، والتي تتمثل في المعالجة المسبقة، والكشف عن التشوهات، وتصنيف هذه الأخيرة كخبيثة أو حميدة.

بعد الحصول على صور الماموجرام، يجب إزالة العناصر غير ضرورية المتمثلة في: العضلة الصدرية، القطع الأثرية وغيرها.

ولهذا السبب، اقترحنا طريقة مبنية على قياس التشابه لرسم حواف العضلة الصدرية، مستخدمين خصائص قياس التشابه الدلالي بين الكلمات في مجال معالجة اللغة الطبيعية(NLP) والبحث عن المعلومات(IR) . كما تم استخدام هذه الطريقة المقترحة في تصنيف الماموجرام على أساس الشدّة كطبيعية أو غير طبيعية.

النتائج التي تم الحصول عليها جيدة لأن الطريقة تعطي الحدود التي تمثل حواف العضلة الصدرية وتستخرج المناطق الأكثر كثافةً من ناحية الشدّة مما يحسن أداء النظام المقترح IOBI.

الكلمات الرئيسية : *سرطان الثدي، حواف العضلة الصدرية، الشدّة، NLP، IR ، CADe،CADx ، IOBI .*

Résumé

De nos jours, le cancer de sein reste la première cause de mort chez l'être humain, mais les femmes sont les plus exposées, car sa cause reste inconnue.
Plusieurs systèmes ont été investi pour détecter et diagnostiquer le cancer du sein dans le but d'aider les radiologues dans la phase du dépistage. CADe et CADx sont les systèmes les plus utilisé pour atteindre cette tâche. Cependant, plusieurs étapes sont nécessaire pour obtenir des résultats pertinentes, qui consistent à prétraitement, détection des anomalies, classifier ces dernier comme maligne ou bénigne.
Après l'acquisition de mammogram, quelques éléments qui sont pas essentiels doivent être enlevé , comme le muscle pectoral, les artefacts, etc. Pour cette raison, nous avons proposé une méthode basé sur la similarité entre les intensités pour tracer les limites du muscle pectoral, on utilisaient les caractéristiques du mesure du similarité sémantique entre les mots dans le traitement de langage naturelle NLP et la recherche d'information IR. Aussi, la méthode proposé a été exploitée a fin de classifier le mammogram, on se basant sur son intensité, comme étant normal ou anormal.
Les résultats obtenus sont prometteurs puisque notre approche donne des limites effective du muscle pectoral et extrait la région avec une intensité élevé ce qui améliore la performance de notre system proposé IOBI.
Mots clé : *cancer du sein, NLP, IR, les limites du muscle pectoral, Intensité, CADe, CADx, IOBI.*

Contents

List of Figures

List of Tables

List of Abbreviations

ANNs	Artificial Neural Networks.
ARCH	Architectural distortion .
ASYM	Asymmetry.
BS	Backward Selection.
C	Concept.
CADe	Computer Aided Detection.
CADi	Computer Aided Diagnosis.
CADx	Computer Aided Diagnosis.
CALC	Calcification.
CC	Cranio Caudal.
CIRC	Well-defined/circumscribed masses.
CLAHE	Contrast Limited Adaptive Histogram.
CTLM	Computed Tomography laser mammography.
DTMC	Discrete Time Markov Chain.
FS	Forward Selection.
HD	Hausdroff Distance.
IC	Information Content.
IOBI	Identification Of Breast Intensity.
IR	Information Retrieval.
GA	Genetic Algorithms.
KNN	K Nearest Neighbors.
LCE	Local Consistency Error.
MAE	Mean Absolute Error.
MC	MicroCalcification .
MIAS	Mammographic Image Analysis Society.
MISC	Other, ill-defined masses.
MLO	Medio Lateral Oblique.

MRI	Magnetic Resonance Imaging.
MST	Minimum Spanning Tree.
NORM	Normal.
NLP	Natural Language Process.
PDF	Probability Density Function.
PEM	Positron Emission Mammography.
PRI	Probabilistic Rand Index.
RANSAC	Random Sample consensus.
Sim	Similarity.
SGLD	Spatial Gray-Level Dependence.
SVM	Support Vector Machine.
SSEM	Single Sided Edge Marking .
SPIC	Spiculated masses.
TC	Tanimoto Coefficient.

General Introduction

1 General Introduction

Everyone of us has dream wants to live it, a goal wants to achieve it, but suddenly something comes out of your hands, affects your body and stop you to move, affects your mind and stop you to think, it is *breast cancer*. The most common cause of women death, and recently it affects men too.
Breast cancer is known as uncontrolled growth of cells inside the breast[14], As because the reason of its cause still remains unknown, prevent and protect human being from this killer illness seems as an impossible issue as consequence high morbidity and mortality. However, detecting and diagnosis this illness on its early stage can decrease the associated morbidity and mortality. Thus, several imaging modality techniques have been proposed to visualize inside the breast in order to detect abnormalities such as Mammography, which is considered as the best technique[95]. these techniques produce medical images which would be read by radiologists to diagnosis this illness. However; reading these medical images by radiologists is inaccurate thus, medical images have been taken from several patients, tiny abnormalities are not visible that means missing detection on the earliest stage of cancer, Moreover, they cannot distinguish between positive and negative anomalies in other words malignant and benign. Hence; a second reading for this kind of images is recommended indeed.
Various techniques have been investigated to assists radiologists in order to obtain accurate diagnosis such as CAD algorithms and image registration. But since medical images are affected during acquisition and contain some irrelevant elements which can bias the result of these techniques, preprocessing approaches have been proposed to restore images from disturbances and remove these elements.

2 Objective Of The Thesis

Diagnosis breast cancer automatically still remains big challenge thus, the information must be accurate which can be achieved by preprocessing methods, accurate removing of irrelevant elements, as well as the performance of CAD systems must be increased.
We have approached this problem by proposing a method that delineate the pectoral muscle boundary thus, the presence of this later affect the result of CAD algorithms, and developing a system IOBI that assists radiologists to identify the cardinality of suspicious regions whether it is normal or abnormal.

3 Organization Of The Thesis

Our thesis has been organized as follow :
Chapter 1: contains an explanation of various breast cancer image modalities techniques such as X-ray Mammography, breast ultrasound, Magnetic Resonance Imaging(MRI), Positron Emission Mammography(PEM) and Computed Tomography Laser Mammography(CTLM). The advantages and limitations of each technique are cited.
Chapter 2: This chapter reviews methods that have been proposed for preprocessing.
Chapter 3: describes the existing approaches that have been investigated to improve the performance of CAD systems and image registration technique.
Chapter 4: presents new method that we proposed to be used for preprocessing and IOBI system that we suggested to help radiologist to determine the cardinality of mammogram as normal or abnormal.

General conclusion which draws the conclusion of thesis, as it illustrates the main outcome of it and what more can be achieved in the future.

1

Medical Imaging Techniques For Breast Cancer

1 Introduction

There are Several diseases that cannot be diagnosis unless revealing internal structures hidden by the skin and bones. For that reason, medical imaging has been proposed as a technique to visualize inside body. Among the common diseases that have been diagnosed by this technique is Breast Cancer. It is known as an abnormal growth of gland(cells that are responsible of producing milk) or duct(cells that pass milk from gland to the nipple).
Variant techniques have been used to detect abnormal change inside the breast such as x-ray Mammography, Breast Ultrasound, Magnetic Resonance Imaging(MRI), Positron Emission Mammography(PEM) and Computed Tomography laser mammography(CTLM), these techniques produce medical image that would be read by radiologist in order to diagnosis this disease. In this chapter, we have explained characteristics of each medical imaging technique as well its limitations in detecting breast cancer.

2 X-ray Mammography

Mammography is an x-ray imaging technique that uses a low dose of X-ray system, which is designed specially for creating detailed image of breast called mammogram. This technique uses two plates to compress the breast between them horizontally then obliquely and x-ray is taken in each position[3](see Figure1.1).

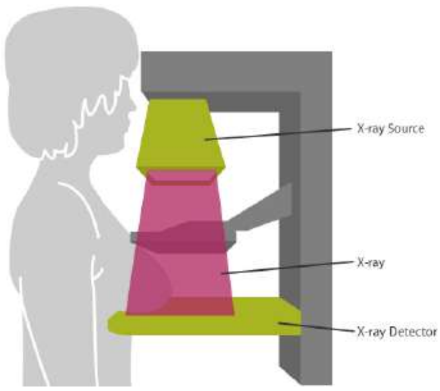


Figure 1.1: **Breast capture by mammography[3]**

There are two types of mammography; analog(that was in the past) and digital(advanced imaging technique).

2.1 Analog mammography

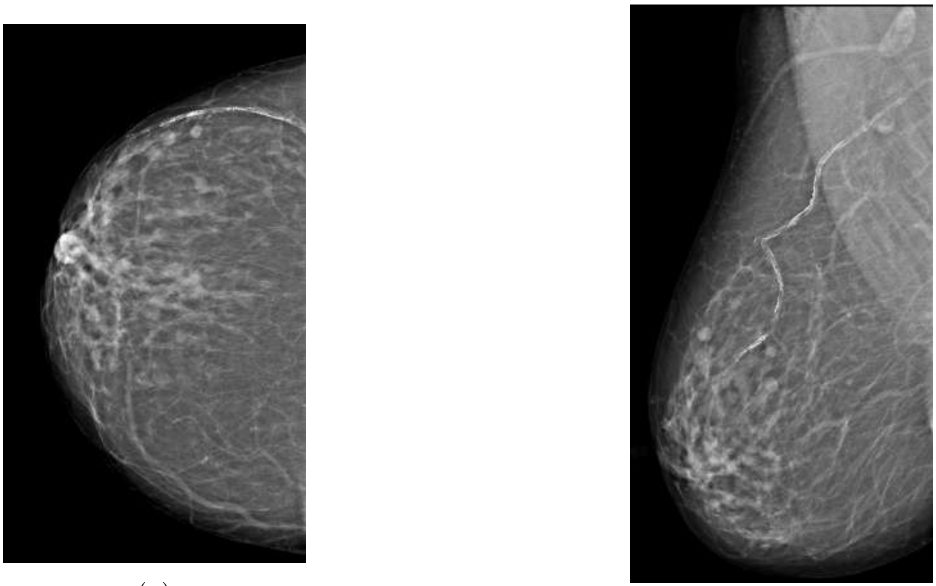
In analog mammography system(also called screening mammography); the image is captured by screen film, which is in cassette format, it is placed beneath the breast position. However; this technique perform poor contrast images because of image acquisition detector, storage and display device[2].

2.2 Digital mammography

In digital mammography system; the image is captured by chemical screen along a digital detector, which detect x-ray photons leaving the breast while chemical screen converts X-ray photons into light. Another device is used fiber optical reducer , it transfers the converted light to CCD (Charged Couple Device) where this latter converts it into digitized analogue signal; which is displayed on computer monitor[2].

2.3 Mammogram Views

Mammography technique offers two views to the breast,Cranio Caudal(CC); which is top-to-button view and Medio Lateral Oblique (MLO) view; which is a side view taken at angle[1](see Figure1.2).



(a) (b)

Figure 1.2: **Two basic views of mammographic image: (a) CranioCaudal (CC) view, (b) MedioLateralOblique (MLO) view [95]**

2.4 Advantages

- The test is a combination between sensitivity and specificity.
- The test is well tolerated[13].
- Mammography offers images with high-quality at a low radiation dose[80].
- Mammography is considered as an effective examination for detection of breast cancer before it becomes clinically palpable[4][80].

2.5 Limits

- Breast compression causes pain and anxiety.
- Radiation risk.
- False alarm[13].
- Recommend just for women that aged more than 40 years[80].

3 Breast Ultrasound

Breast Ultrasound is considered as a less harmful technology within other technologies, it uses sound waves to map the internal structure of the breast rather than using x-ray or other potentiality harmful types of radiation.As it is characterized by a quick visualization of the breast tissue(see figure1.3).

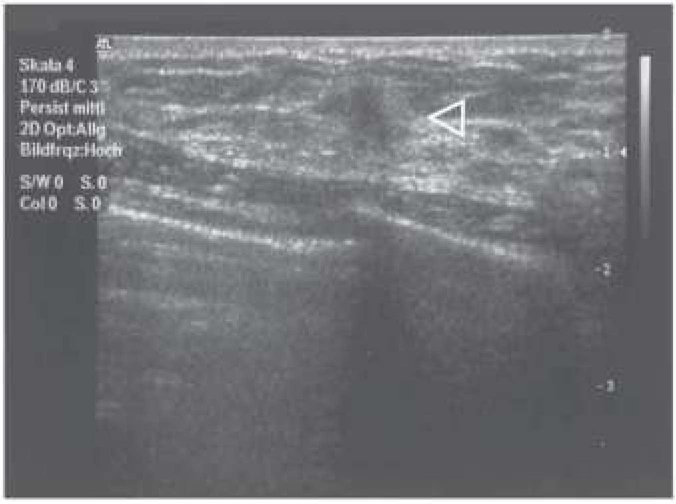


Figure 1.3: **Breast capture by Ultrasound[5]**

3.1 Advantages

- Cystic and solid are distinguishable.
- Patients under 30 years can take this examination.
- Solid masses are characterized in breast ultrasound examination.
- Painless

3.2 Limits

- Calcification(diffuse tumors) are not well captured (Low sensitivity).
- Low specificity result.
- Higher cost.
- Exploring the whole breast carefully takes long time.

4 Magnetic Resonance Imaging(MRI)

Magnetic Resonance Imaging is an invasive medical imaging test that produces detailed images of structure within body, including organ. This technique uses a combination of large magnet, radio wave and a computer to produce images.

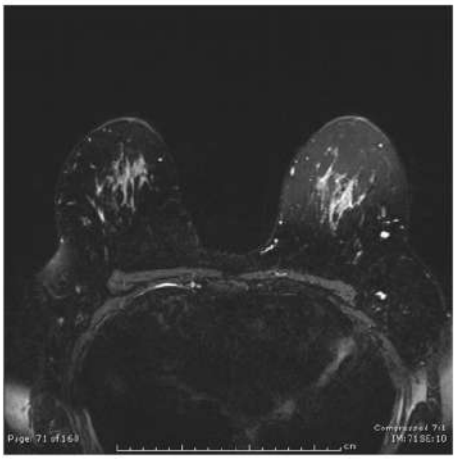


Figure 1.4: **Bilateral axial T2-weighed breast magnetic resonance image.[6]**

4.1 Advantages

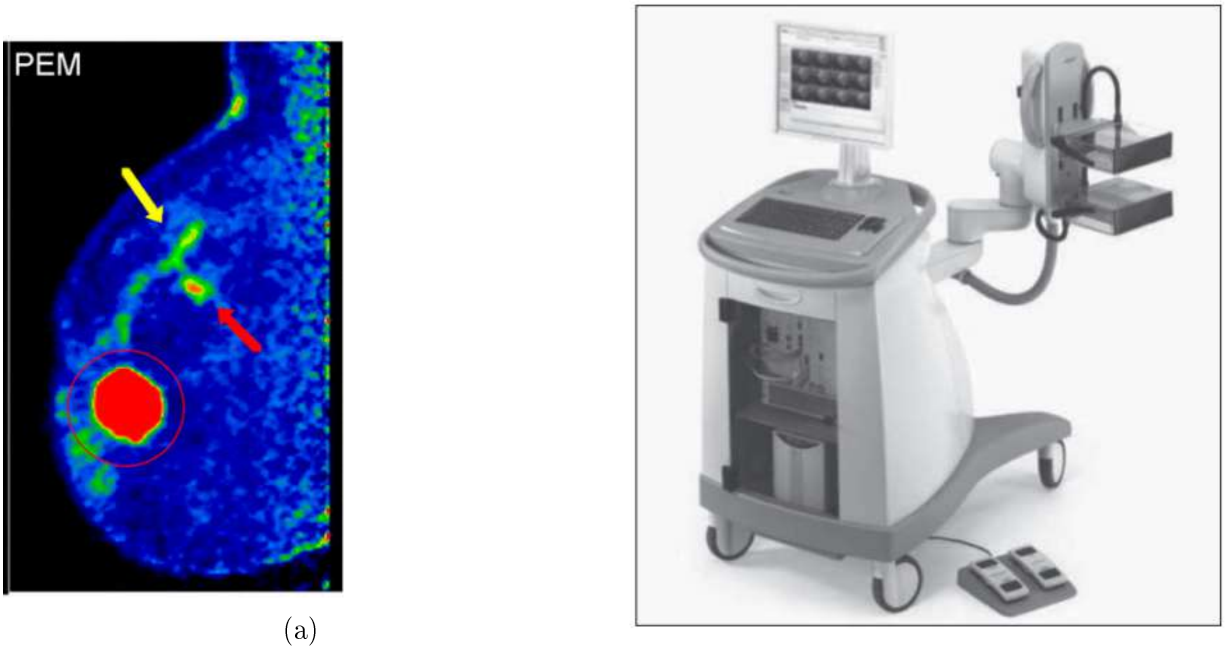
- Excellent tissue differentiation.
- High sensitivity of breast cancer .
- Capability of identifying breast masses as small as 1-2 mm.
- No limited by breast density.
- No ionizing radiation.

4.2 Limits

- Contraindicated in some patients
- Low specificity results in excess biopsies and additional follow-up.
- Indistinguishable intensity (caused by MRI acquisition or refers to the tissue of breast) in MRI images.
- Percutaneous biopsy under MRI guidance requires a stereotaxic device and needle that are nonmagnetic, and is a procedure that is under active development.

5 Positron Emission Mammography(PEM)

PEM is miniature to Positron Emission Tomography device that is dedicated to breast cancer diagnosis, it is used as alternative examination for people who cannot tolerate MRI[7]. PEM technique uses a pair of dedicated gamma radiation detectors placed above and below the breast and mild breast compression to detect coincident gamma rays after administration of fluorine-18 fluorodeoxyglucose(18F-FDG)[9].



(a) (b)

Figure 1.5: **(a) Breast capture[8] (b) by PEM [9].**

5.1 Advantages

- Higher sensitivity for breast cancer .
- Images with higher spatial resolution.
- It provides complementary information to conventional breast imaging modalities.
- Ability of detecting small hyper-metabolic lesions

5.2 Limits

- High radiation dose.
- Current limitations for biopsy.
- A single PEM study involving the use of a label-recommended radionuclide dose is associated with a 15-fold higher.

6 Computed Tomography laser mammography

Computed Tomography laser mammography(CTLM) is considered as an effective examination method for breast cancer thus , it is able to distinguish between malignant tumour and benign lesion.
CTLM uses laser beam of the wavelength equal to 808 nanometres, which is absorbed in blood pigments of physiological and pathological blood-vessels and is able to display their distribution[10].

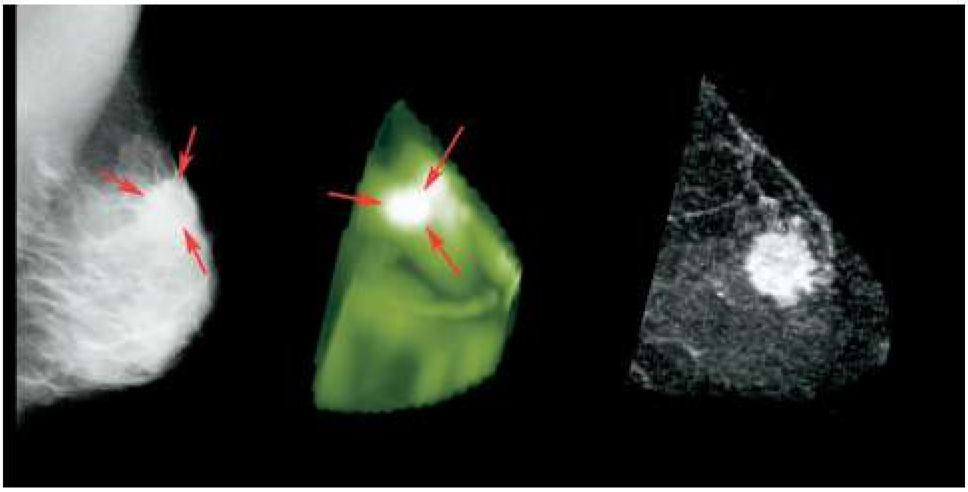


Figure 1.6: **Correlation between mammography, Computed Tomography Laser Mammography, and magnetic resonance imaging in a 28-year-old woman with an invasive ductal carcinoma and ductal carcinoma in situ (high grade) in her left breast[11]**

6.1 Advantages

- No ionizing radiation.
- No breast compression.
- No injection contrast agents.
- Empowering decision.
- Non-invasive/harmless.
- Easy and inexpensive to operate[12].

6.2 Limits

- Anxiety of patient before and during CT imaging.
- It takes much more time comparing to other imaging techniques.

7 Conclusion

Breast imaging modality techniques have been proposed to visualize inside the breast on purpose of diagnosis it, the results of these techniques consist in medical images that would be read by radiologist to detect abnormalities, however tiny abnormalities are not visible for radiologist as well they can not differentiate malignant from benign tumor. Thus; a second reading for medical images is recommended indeed.

2

Preprocessing and Breast Profile Segmentation

1 Introduction

Breast cancer is the most common cause of women death, for reason of its cause not yet fully known. However; detecting this illness in its early stage can reduce the associated morbidity and mortality rates.
Mammography is one of tools that used in the early detection of breast cancer[95]; it is used to visualize inside of the breast using a low dose of x-ray system to obtain medical image; where this latter would be reviewed by radiologist to detect abnormalities. However; radiologist cannot have an accurate result when abnormalities are not visible as well they cannot distinguish between malignant and benign; they are often classify abnormalities to malignant[14].
A sophistical wave of using computer in diagnosis of breast cancer has been emerged to avoid above problems; this diagnosis uses digital image that would be read by computer as well methods to detect abnormalities automatically which can be categorized into CADe (Computer aided detection) and CADi (computer aided diagnosis).
Before applying CAD algorithms, images need to be preprocessed. This chapter provides steps that have been used to prepare image for further process .
The first step ; preprocessing which is important to improve the quality of image ;second step is segmentation, which it is used to remove elements which are not considered in diagnosis.

2 Preprocessing

Quality of medical image is an indispensable factor, which helps to achieve accurate diagnosis and results. The presence of noise occurred during image acquisition in the image leads to suppress and blur important features in the image. For that reason, preprocessing is an important step used to improve the quality of images, as a result increasing the effectiveness of diagnosis[20].
Noise can be presented in digital image like as straight line, poor contrast and weak boundary[19].
Several approaches have been reported to restore images from disturbances such as:

2.1 Filter techniques:

Filter techniques are very effective in image restoration process which can be categorized into linear filter and non-linear filter; we mention methods that are widely used such as :

2.1.1 Mean filter

Mean filter is a linear filter technique ,it consists on replacing each pixel by the average value of intensities in neighborhood .
Example:

$$\begin{bmatrix} 2 & 2 & 2 \\ 2 & 5 & 2 \\ 2 & 1 & 1 \end{bmatrix} \xrightarrow{Meanfilter} \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2.11 & 2 \\ 2 & 1 & 1 \end{bmatrix}$$

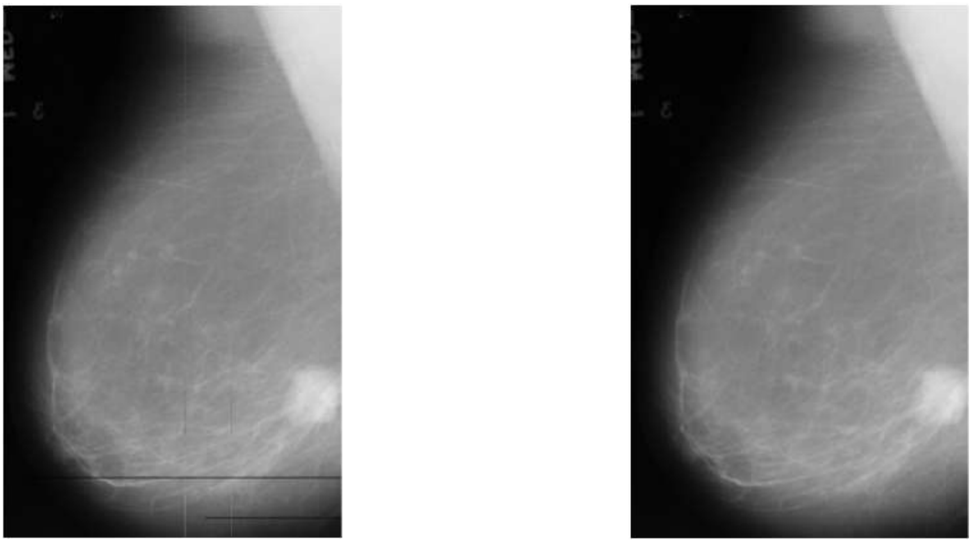
Beside mean filter can locally reduce variance and it is easy to apply;sometimes it does not give the desired result because [19]:

- Average operation leads to create another type of noise, which is blurring .

- In case of existing impulse noise (It is a short duration noise that may occur due image acquisition [17]) in image, applying mean filter will attenuate and diffuse without removing.
- The presence of a single pixel with a tiny value affects the mean value of all pixels in neighborhood significantly.

2.1.2 Median filter

It is non-linear filter; it used to save sharpness of image edge while removing noise [19] .Median filter works by sliding window that used to replace each pixel by the median value of its neighborhood, however; edges of the image are replaced by zero value that refers to black color in gray-scale. Median filter is an adequate method that it used to deal with straight line noise as it is illustrated in figure2.1 .



(a) **Corrupted image** (b) **Filtered image**

Figure 2.1: **Mammogram noise digitization removal using median filter** [22]

Hence; it may distribute noise in image if it is corrupted by a high degree of impulse noise[18]. Variant methods differ from median filter to palliate these limitations like as:

- **Max median filter**: where each pixel is replaced by the max value represented within its neighborhood.
- **Weighted median filter** : which refers to the higher weight value given under inspection to the filtered pixel[23].
- **Adaptive median filter** : several techniques that variant from median filter replace directly the filtered pixel by another value, however; in adaptive median filter,just the corrupted pixel is replaced by another value.
  The corrupted pixel is known by a different value from its neighborhood,moreover it is not structurally aligned within pixels to which it is similar[24].

2.2 Image Enhancement techniques:

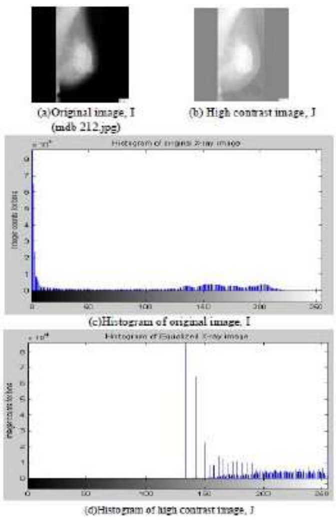
Image enhancement is considered as a crucial technique in medical image preprocessing, it increases the contrast and sharpness of images in the aim of providing an accurate input for further process[15]. Many techniques are used such as:

2.2.1 Histogram equalization:

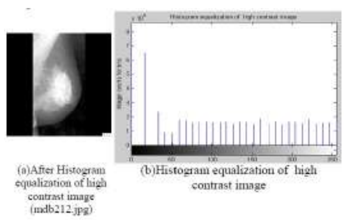
Since the weak contrast refers to the bad distribution for intensities values of an image in the histogram, it is necessary to map values back into their proper range.
Histogram equalization[16] is used to reach this goal by strewing the most frequent value intensities represented in histogram; as a result, the lowest intensities will appear by height frequent(i.e. see figure2.2). It is defined as following

$$\overline{T}(r_k) = int[\frac{T(r_k) - T_{min}}{1 - T_{min}} \times (L-1) + 0.5] \tag{2.1}$$

“Where T_{min} refers to the smallest value in the cumulative probability density function vector $T(r_k)$. $T(r_k)$ is the normalized value of each gray-level and L is 256 for eight bit gray scale image and $T(r_k)$is the transformed value for each gray level.”



(a)



(b)

Figure 2.2: **(a):Histogram equalization of mammogram,(b):contrast enhancement of mammogram** [25]

2.2.2 Contrast stretching :

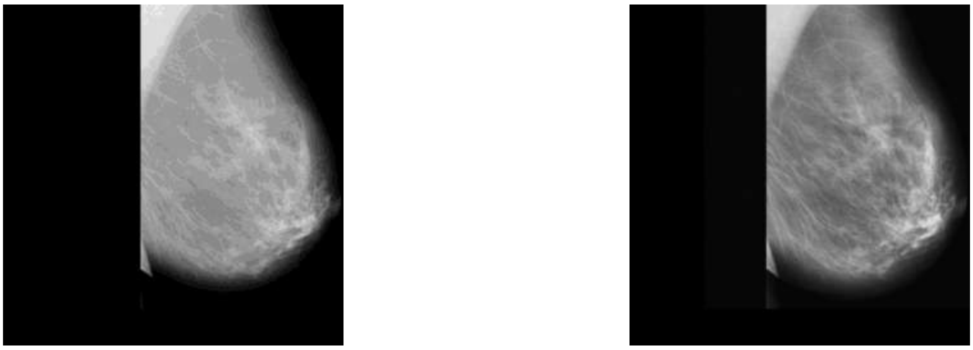
Contrast stretching(often called Normalization) differs from Histogram Equalization [15],it is a less enhancement technique that consists on stretching the range of intensity values of an image by a linear scaling function , which is defined as following:

$$P_{out} = (P_{in} - c)(\frac{b-a}{d-c}) + a \tag{2.2}$$

Where a, b refer to the lower , upper value limit concerned in this type of image (0-255), while c, d refer to the lowest , highest value represented in image respectively. Hence P_{out} may take values over value limits, in this case; values below 0 are set to 0 while the above 255 set to 255.

2.2.3 CLAHE :

CLAHE stands for Contrast Limited Adaptive Histogram Equalization, it has been approached special for medical images on purpose of reducing the effect of edge-shadowing produced in homogeneous regions(see figure2.3).
In CLAHE approach, the image is divided into small regions , so the histogram of each regions is calculated, then it is enhanced to limit contrast and to avoid amplifying any noise or edge-shadowing effect[26].



(a) (b)

Figure 2.3: **Original mammogram (a) after CLAHE(b)**[26].

3 Segmentation

Segmentation is the process of portioning image into regions, where each region groups pixels that go together. This process aims to separate the breast tissue from background of the mammogram furthermore; remove the pectoral muscle from the rest of breast tissue,thus abnormalities are located just inside the breast.
A lot of algorithms have been suggested for the two steps.

3.1 Separating the breast tissue from background

Background of mammogram is well known by radiopaque artifacts, low and high intensity as it is shown in figure2.4.

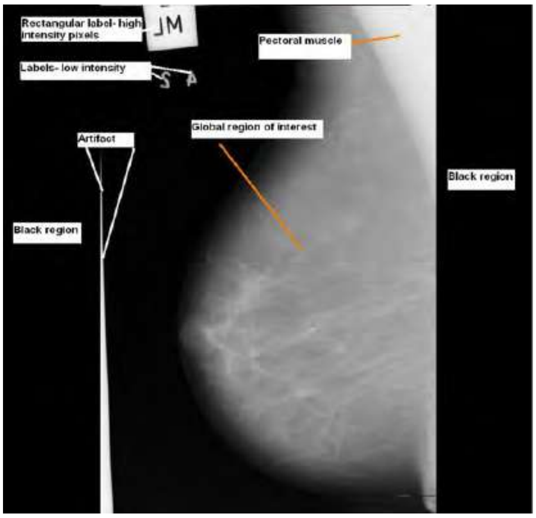


Figure 2.4: **Irrelevant elements in mammogram[30]**

There are several approaches that have been used to separate breast tissue from the mammogram's background such as:

- Morphological image processing
- connected components labeling

We explain these approaches in the following paragraphs.

3.1.1 Morphological image processing :

Morphological image processing is a collection of techniques that are used to adjust the shape of regions in image from distortion accrued during segmentation, however the image must be binary. Techniques that used are[28] :

3.1.1.1 Dilation :

Dilation is used to cover breaks and intrusions presented in the shape of region in an image, it grows and thickens the regions. Dilation is defined as following:

$$A \oplus B = \{s|(\hat{B})s \cap A\} \tag{2.3}$$

A is a set to be dilated , B is a structuring element that used to dilate the input (i.e. disk shaped), whereas; $(\hat{B})S$ is the reflection of B about its origin and followed by a shift by s. As well $A \oplus B$ is the set of all shifts that satisfy $\{s|(\hat{B})s \cap A\}$.

3.1.1.2 Erosion :

It is used to remove part joined objects as well extrusions.it is the opposite of dilation, it thins and shrinks object using the following equation .

$$A \ominus B = \{s|(B)s \subseteq A\} \tag{2.4}$$

Where A is input image , B is a structuring element.

3.1.1.3 opening & closing :

An object is defined by its contour, hence any distortion on contour leads to misunderstanding that object.
Imperfections that may exist on contour object are : break narrow and thin protrusions which are eliminated by opening operation while long thin gulfs,holes and gaps are eliminated by closing operation[28].

opening:

$$A \circ B = (A \ominus B) \oplus B \tag{2.5}$$

closing:

$$A \bullet B = (A \oplus B) \ominus B \tag{2.6}$$

3.1.2 connected components labeling

connected components labeling procedure is the process of grouping pixels that are similar in intensity as well connected to each other in one segment .The utility of this procedure is extracting object.
The classical approach for connected component labeling based on two subsequence of raster-scans.The first scan is used to give each pixel a temporary label depend on its neighbors values which are already assigned;at the end of this scan, different labels may contain the same component, cause of unconnected pixels with the same value, however different pixels are assigned to different label.Then; second scan is used to store equivalent label into one segment[29].

3.2 Pectoral muscle removal

The presence of pectoral muscle in mammogram can bias the result of CAD algorithms for many reasons: pectoral muscle occupies predominant region in mammogram , as well it has almost the similar pixel intensity of abnormalities.
Variant methods have been used to remove pectoral muscle accurately and efficiently over mammogram whether by detecting its edge or its region. Methods have been categorized as following[31]:

- Based on gray level
- Based on texture features
- Based on soft computing

3.2.1 Based on gray level

Extracting pectoral muscle based on gray level may be based on :

- Intensity
- Region
- gradient

3.2.1.1 Intensity based methods

Pectoral muscle is well defined by a dense and high intensity in digital mammogram, thus various methods use these characteristics to select pectoral muscle such as Histogram-based threshold, watershed transformation, Otsu thresholding.The utility of these methods in pectoral muscle removal have been proven in the following paragraphs.

(i) Histogram-based threshold:

Since The region of pectoral muscle is more noticeable in mammogram, histogram is used to extract this region.

Thangavel et al.[32] proposed an algorithm to extract pectoral muscle region.

First; histogram of image has been calculated, after; the global optimum intensity is used as threshold to select pixels that include pectoral muscle region as a result binary image is created, then; morphological operation such as dilation and erosion were applied on binary image to adjust pectoral muscle region, finally; this latter is removed from mammogram using the resulting binary image.

(ii) watershed transformation:

The idea of Watershed transformation came from the great divide which separates USA into two regions,Atlantic Ocean and Pacific Ocean. This divide constitutes a watershed line whereas Atlantic and Pacific are considered as catchment basins [38].Thus,this technique is used to detect pectoral edge since pectoral muscle region and ROI are separated by a line.

Camilus et al.[37] use watershed transformation to identify pectoral boundary,then a margin algorithm is proposed to combine several catchment basin produced during transformation as a result the accurate boundary is delineated. Figure2.5 illustrates result of the proposed algorithm .

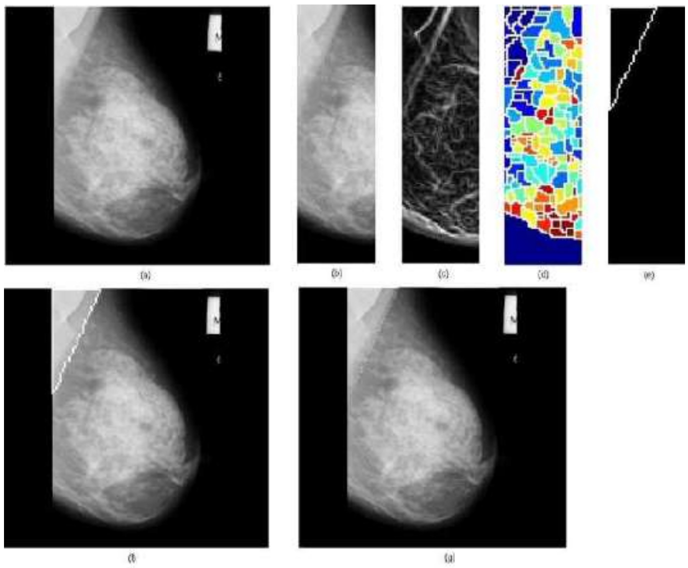


Figure 2.5: **(a) original mammogram, (b) ROI image, (c) ROI gradient, (d) watershed transformation of the ROI gradient, (e) identified pectoral muscle by the proposed merging algorithm, (f) identified pectoral muscle embedded over the original mammogram, (g) ground truth (pectoral muscle is marked white).**[37]

(iii) Otsu thresholding :

Thresholding technique is well known in image segmentation, it is used to extract target region from an image[39].

Lieu et al.[40] proposed an algorithm based on Otsu thresholding to identify the pectoral muscle border.
First, the image is binarized on a threshod which is selected by Otsu thresholding technique, after; morphological opening operation was applied to remove selected unconcerned regions, then;the pectoral muscle border is extracted, finally;regression analysis[41] were applied to adjust the selected pectoral border.

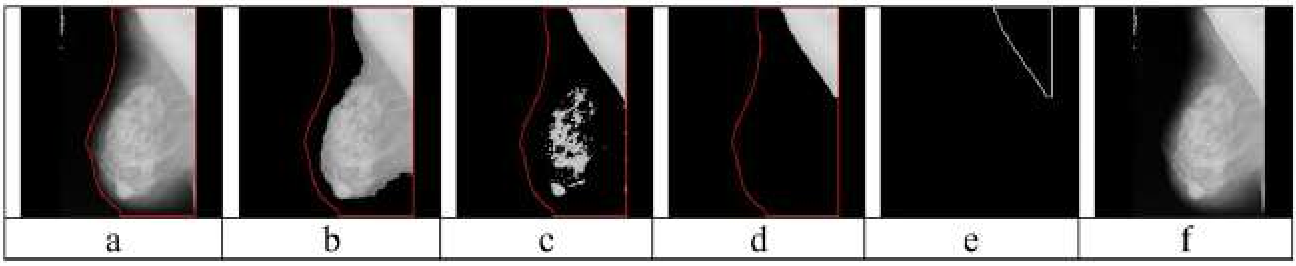


Figure 2.6: **The output of each steps of the presented algorithm, (a) breast region extracted by using MGVF snake, (b) after Otsu thresholding and morphological processing one time, (c) after Otsu thresholding and morphological processing two times, (d) after position detecting, (e) extracted pectoral muscle rough border, (f) final extracted pectoral muscle**[40]

3.2.1.2 Region based Methods

A region in image is defined by a group of connected pixels with similar characteristic.
Seed region growing[42] is a technique that is used to select region in image.The input of this algorithm is a pixel include the desired region to be selected,then; the process gathers all pixels that satisfy preset conditions like as distance.

Raba et al.[33] proposed an algorithm based on Seed Region Growing. Firstly; an adaptive histogram has been applied to separate the breast from background, secondly; the orientation of mammogram has been indicated in order to place a seed point in the pectoral muscle region, then; statistical region growing has been applied to extract the Region Of Interest (Breast tissue without irrelevant elements), finally; size restriction has been applied to avoid wrong growing.

3.2.1.3 gradient based methods

Pectoral muscle seems as a triangle at the edge of mammogram(MLO view), Thus; gradient based methods separate pectoral muscle from breast tissue by straight or curve line between them. A lot of techniques have been used to identify pectoral muscle region using gradient methods such as :

(i) tunable parametric edge detection :

Chandrasekhar et al.[34] improve the conventional edge detection paradigm to tunable parametric edge detection. The edge of pectoral muscle is detected in two steps:
The first step was an application of gradient techniques on the original image which consists on calculating the absolute values of *horizontal* and *vertical Sobel* gradient as well *range* and *standard deviation* using following equations respectively:

$$\varphi_1(\omega) = |(\omega_9 + 2\omega_8 + \omega_7) - (\omega_1 + 2\omega_2 + \omega_3)| \tag{2.7}$$

$$\varphi_v(\omega) = |(\omega_3 + 2\omega_6 + \omega_9) - (\omega_1 + 2\omega_4 + \omega_7)| \tag{2.8}$$

$$\varphi_r(\omega) = \max_{1\leq i\leq 9}[\omega_i] - \min_{1\leq i\leq 9}[\omega_i] \tag{2.9}$$

$$\varphi_s(\omega) = \sqrt{\left[\frac{1}{9}\sum_{i=1}^{9}\omega_i^2\right] - \left[\frac{1}{9}\sum_{i=1}^{9}\omega_i\right]^2} \tag{2.10}$$

ω refers to a vector that groups candidate pixel and its neighborhood in 3×3 window, pixels are indexed from 1 to 9 depends on raster fashion that means from top to bottom and left to right. Then; values of four gradient are normalized to ensure computability between them, as a result values can take [0,4] .
The second step was a detection of pectoral muscle edge, results of the four gradients were blended using the modified *logistic* function which is defined as following.

$$b_L(t, \varphi, \lambda, \beta) = \frac{1}{1 + \exp(-\lambda(\|\varphi\|_t - \beta))} \tag{2.11}$$

where λ and β are real positive constants; whereas the Minkowski t-norm$\|.\|$ is defined by :

$$\|\varphi\|_t - = \left[\sum_{i}^{p} |\varphi|^t\right]^{\frac{1}{t}} \tag{2.12}$$

Figure2.7 illustrates above steps.

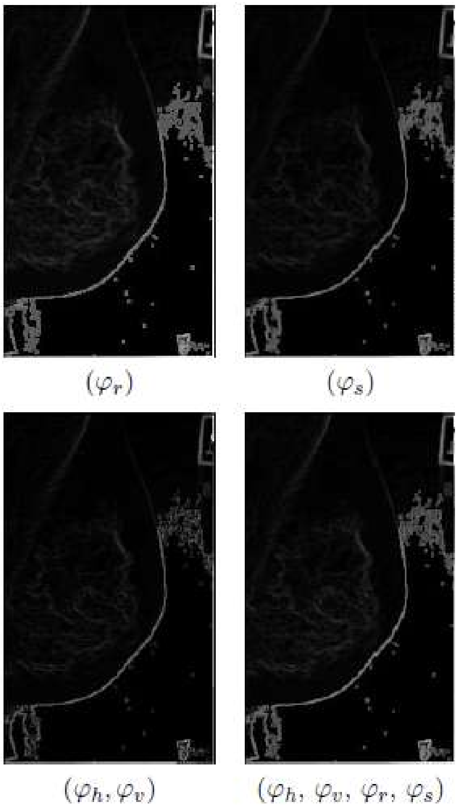


Figure 2.7: Pectoral muscle edge detection[34]

(ii) Hough Transform :

Hought transform is a procedure that is used to detect straight line in digitized images using angle-radius[43].This procedure is used to select pectoral muscle edge which is appeared approximately as a straight line in mammogram.

Ferrari et al.[36] suggested an algorithm based on Hough transform to delineate pectoral edge. First; the breast contour was detected using pair of Chain-code algorithm and contour model algorithm ,next;the processed region is determined, it is a part of breast region that was extracted where the pectoral muscle region is included in, thus to reduce the possibilities of the presence of linear structure.Then;the presented noise was removed using Gaussian filter[44],then; Sobel gradient methods[48] was applied on the processed region to detect the pectoral edge,finally; the straight line is specified using Hough transform.

(iii) Iterative Cliff detection :
Since straight line does not indicate the real edge of pectoral muscle,it is necessary to refine this straight line into a curve as consequence the edge of pectoral muscle is delineated accurately.Cliff detection technique is designed to curve straight lines.

Kowk et al.[45] used two steps to identify pectoral edge. First step was an estimation to the edge of pectoral by straight line using iterative threshold selection and gradient test, in the second step, the estimated edge was refined using iterative of cliff detection.

(iv) Random Sample consensus(RANSAC):
Gradient methods are used to detect pectoral edge, however the result of this process is accompanied by outliers lines which can make edge detection inaccurate.RANSAC algorithm[46] is used to interpret these outliers lines.
Molinara et al.[47] presented an algorithm based on RANSAC to detect pectoral edge.
The detection process starts by locating pectoral muscle at the top left corner in the image, after; the skin-line of the breast was identified the intensity of pixels that indicate skin-line are emphasized(normalization approach), then; Gaussian filter was applied to preserve large variation in image as consequence this latter would be used to detect pectoral edge, it is the characteristic that gradient methods based on to delineate edges. ,gradient method that has been used is Prewitt[48].Finally; RANSAC approach is used to delineate pectoral edge more accurately.

3.2.2 Based on texture features

Texture in an image is defined as some repeated pattern of small scale details[35].Hence; variation in textures identify the edge of pectoral muscle .
Techniques that based on this feature are:

- Transform Based Methods
- Probability / Polynomial
- Active Contour Based Methods
- Graph Theory Based Methods

3.2.2.1 Transform Based Methods
Transform technique consists on decomposing image to elementary, hence; texture features would be easy to analyze.
Various ideas suggested on transform based approach to extract pectoral muscle edge such as:

(i) Gabor wavelets :
An image can present different texture regions, where each texture is associated to a feature, hence; this latter is used for searching and retrieving data from image. In order to use this data, wavelet transform technique is used to extract features from image, such as Gabor Wavelet[49].

Ferrari et al.[36] proposed a method based on Gabor wavelets technique to identify pectoral muscle boundary.
The process stars by decomposing the images using *Gabor wavelets* technique,then *vector-summation* procedure was applied on the result to compute magnitude and phase of images, the magnitude value of each pixel is propagated in the direction of phase, the opposite phase orientation has been used to detect pixels that include the pectoral muscle boundary,finally a post-processing stage is used to eliminate unlucky pectoral boundary candidates.

(ii) Radon transform :
. Radon transforms is a suited approach used within image to detect features. Kinoshita et al.[50] used several approach to select pectoral edge.
First; wiener filter[51] was applied to remove radiopaque artifacts as well to restore original value of pixels that identify pectoral edge. After; pectoral edge is detected approximately by Canny filter[52] as a result pectoral muscle region is identified .Then; authors found all straight lines that may represent pectoral edge using radon transform, this latter was applied on the edge image , in the angle interval $[5°; 50°]$, $[-5°; -50°]$ for right and left breast respectively. Finally; the effective straight line was selected to delineate pectoral edge.Figure2.8 explains detection process of the proposed algorithm.

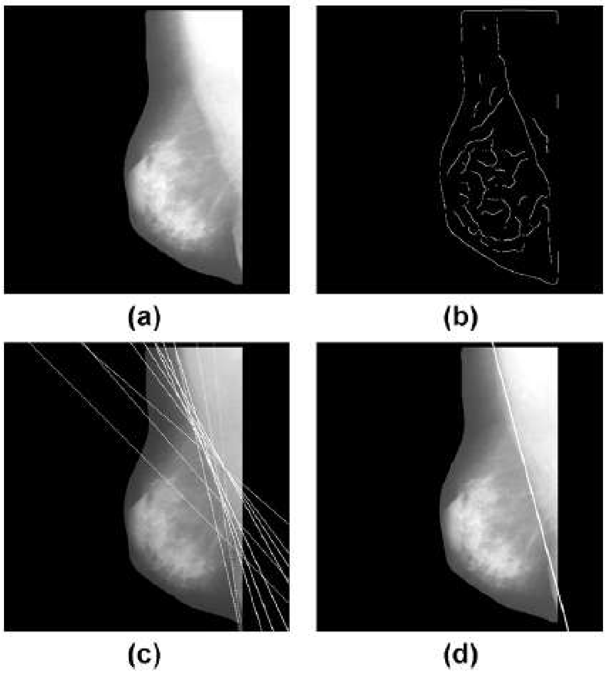


Figure 2.8: **(a) original image; (b) the edge image obtained using the Canny filter; (c) detection of straight-line candidates by the Radon transform; (d)selection of the pectoral muscle edge[50]**

(iii) Texture and intensity deviation :
Pectoral muscle region is represented in digital mammogram by an homogeneous texture and similar gray-level values approximately, while the border of this region is identified by high intensity deviation within its surrounding tissue.
Detection of pectoral edge has been approached by Li et al.[53] using these features.
Authors identify pectoral muscle region based on texture and intensity by different likelihood maps using spectral cluster in based segmentation and mean square deviation(MSD)matrix respectively.

Afterward; likelihood maps are combined to extract initial pectoral edge,finally; this latter was refined by Kalman filter to delineate pectoral edge efficiency.(see Figure2.9)

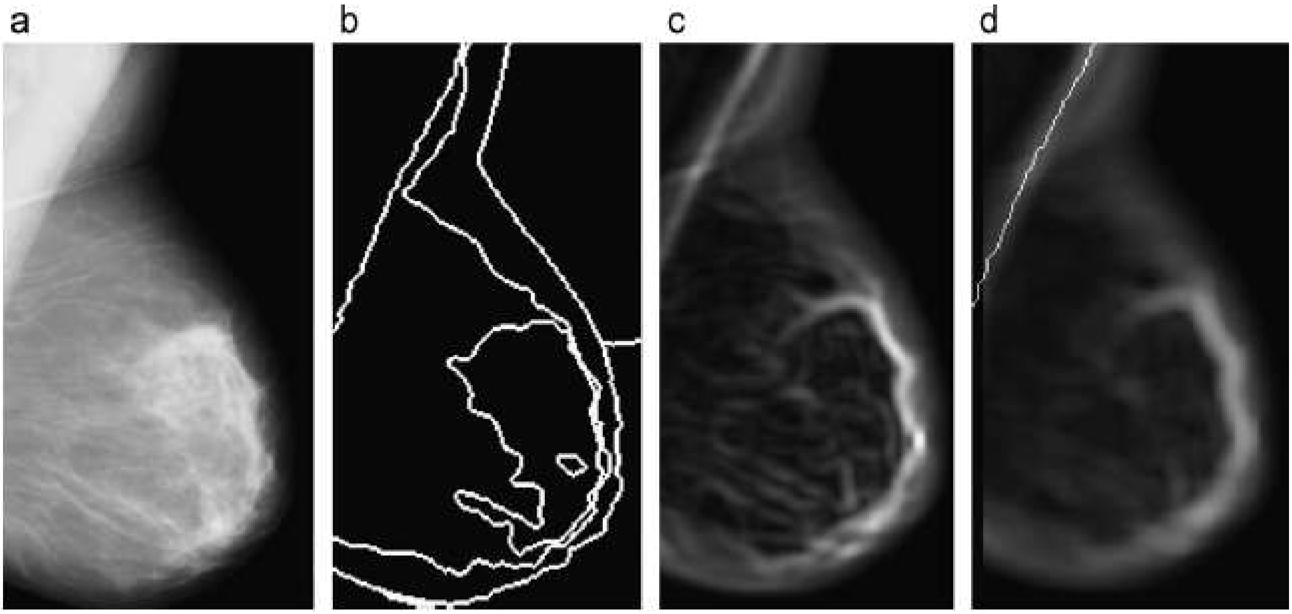


Figure 2.9: **The likelihood of the pectoral muscle of mdb095 in mini-MIAS;(a)is the original mammogram.(b)is the likelihood map in texture field.(c)is the likelihood map in intensity field.(d)is the likelihood map of the pectoral muscle edge[53].**

3.2.2.2 Probability / Polynomial based Methods

Edges in image can be detected by extracting pixel intensities candidate. Hence; probability and polynomial approaches have been used to classify pixel intensities using statistical parameters. These approaches are utilized to identify pectoral edge using texture, density and shape to deduce statistical parameters[31].
Few methods based on probability and polynomial have been suggested such as :

(i) Mean-shift clustering:

Clustering is the process of classifying pixels into different groups using measurement methods[57]. Mean-shift clustering[56][54] uses Probability Density Function(PDF)[55][54] as a measure to form clusters.
Sultana et al.[54] propose an algorithm based on Mean-shift clustering to select pectoral muscle region.
First; Breast tissue was separated from background using Histogram equalization and Labeling procedure. After; Mean-shift clustering was applied on resulting image to segment it, enough care for clustering result; authors applied Gaussian kernel[58] on the segmented image to associate discrete data to clusters.Finally; positioning and mean value intensity have been used as parameters to extract the target region.Figure 2.10 illustrates the result of the proposed algorithm.

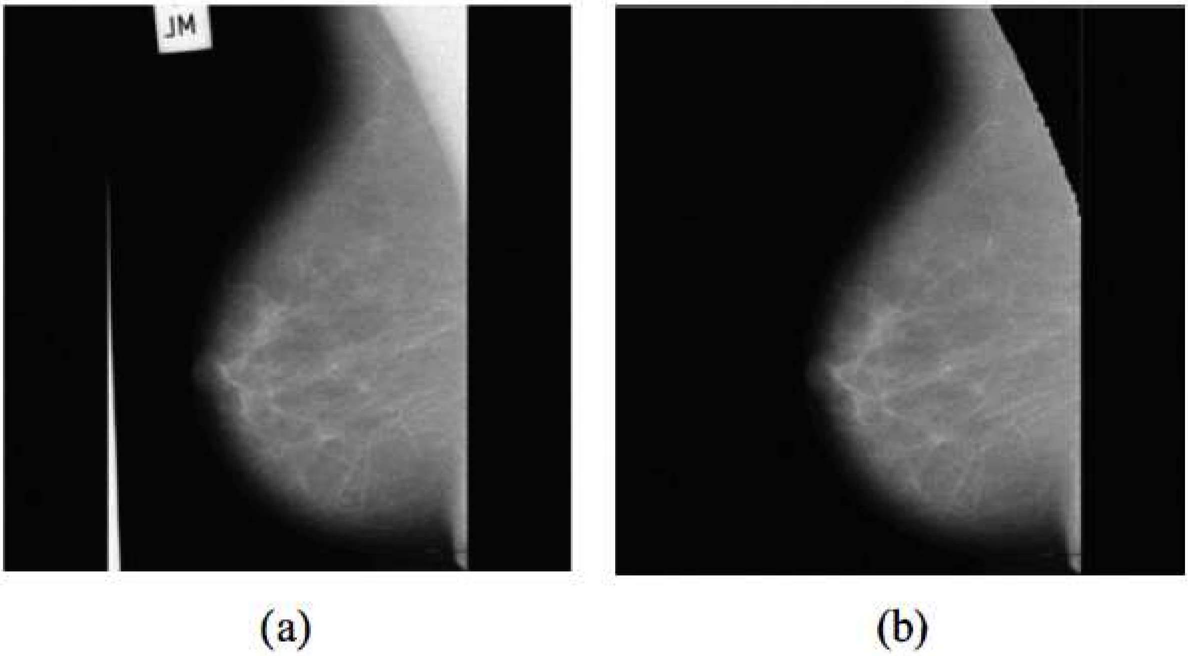


Figure 2.10: **(a)original image.(b)image after the pectoral muscle suppression was applied to the image[54].**

(ii) Position, intensity and texture information :

Pectoral muscle is observed by brightest regions in mammogram, from this characteristic Oliver et al.[65][31] extract characteristic of other regions to distinguish between regions as consequence pectoral region is removed. The other regions consist in background which is characterized by dark intensity while breast region is characterized by a variation of intensity between bright and dark.

Authors proposed a supervised single strategy to associate each pixel into its belonging region using several methods such are: probability density function which is calculated for each pixel (noticed A_R), Histogram; it is used to determine the intensity of each region(noticed I_R), Local Binary Patterns is used to characterize each pixel depending on its texture probability(noticed T_R).Hence; A_R, I_R and T_R are multiplied to calculate the likelihood of pixel belonging to each region, as a result each pixel is assigned to a region based on higher probability.

(iii) polynomial estimation :

Mustra et al.[59] proposed a robust automatic pectoral muscle segmentation from scanned mammograms, the process based on polynomial estimation to pectoral edge, it consists of the following steps. First; the position of pectoral muscle is indicated at 2/3 part of breast high in ROI. Second; CLAHE is used to enhance the contrast.Third; morphological operations have been applied to eliminate small objects and background noise.Fourth; threshold is calculated to create preliminary binary mask ,this latter would be used to choose 10 random points for polynomial fitting of pectoral edge. Finally;cubic linear fitting function and iterative linear fit function are used to smooth pectoral boundary. Functions are defined as following:

Cubic fitting function:

$$y = p_1x^3 + p_2x^2 + p_3 + P4 \tag{2.13}$$

Iterative linear fit function :

$$y = p_5x + p_6 \tag{2.14}$$

Where y is horizontal coordinate and x is vertical coordinate.

3.2.2.3 Active Contour Based Methods The idea of active contour(Snake) approach consists in indicating points that are belong the edge using smoothing filters(i.e. Gaussian), then minimizing all the external and internal potential energy of surrounding neighbor of the indicated points.This approach is widely applied in medical image segmentation,it is used to detect pectoral edge[31].
Several methods have been presented on active contour based method such as: Adaptive contour model,Discrete Time Markov Chain(DTMC) and Multiphase active contour method.The utility of these methods are discussed in the following paragraphs.

(i) Unimodal Thresholding Algorithm :
Unimodal distribution of points near edge of objects makes edge detection process difficult; for that reason, unimodal thresholding techniques are used to obtain a bi-level threshold of such image as a result an estimation of edge points is achieved[60][61].
To detect pectoral edge; Wirth et al.[60] use unimodal thresholding algorithm to identify an initial edge. After; this latter is enhanced by directed edge enhancement method. Then; morphological erosion was applied to remove to enlarge the enhance edge as well the presented noise is removed. Finally; authors improved the classical snake to identify pectoral edge, thus; classical snake cannot deal with topological variation in images[31].

(ii) Discrete Time Markov Chain(DTMC) :
DTMC is a random process that used to index set and state of discrete space using propriety of Markov(independence conditional of future evolution on the past)[62][31].Wang et al.[62] used this approach to model pectoral muscle boundary, because this latter is characterized by continuity and uncertainty.The result of this approach is Accompanied by a false points candidates, hence straight line has been used to validate rough boundary. The proposed process ends with a refinement to the detected edge using Snake algorithm with slight modification.Figure2.11 illustrates the result of the proposed algorithm.

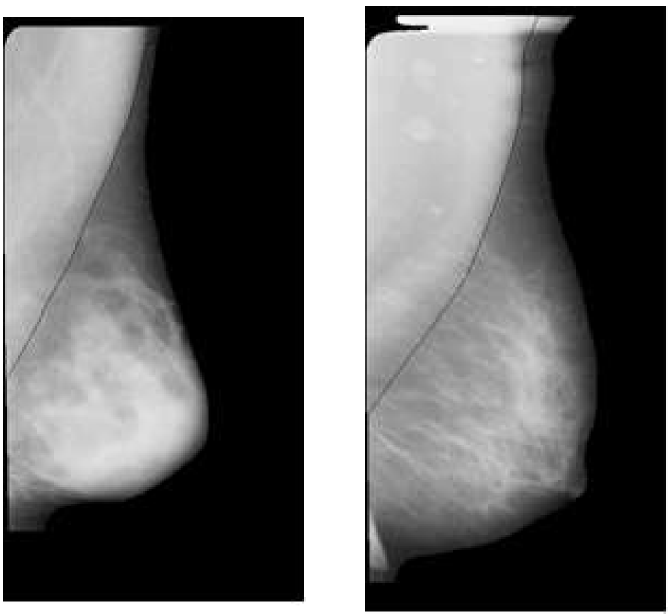


Figure 2.11: **Pectoral muscle boundary detection using DMTC and snake algorithm[62]**

(iii) Multiphase active contour method :
Beside active contour approach detects pectoral edge, it is also used to remove completely pectoral

muscle from mammogram.
Akram et al.[63] used modified active contour method proposed by Chan and Vise[64] to obtain pectoral edge, where this latter is detected in multiphase.As well authors defined a new term M^k to indicate the stopping point of the contour as a result the desired region is extracted as it is shown in Figure2.12.

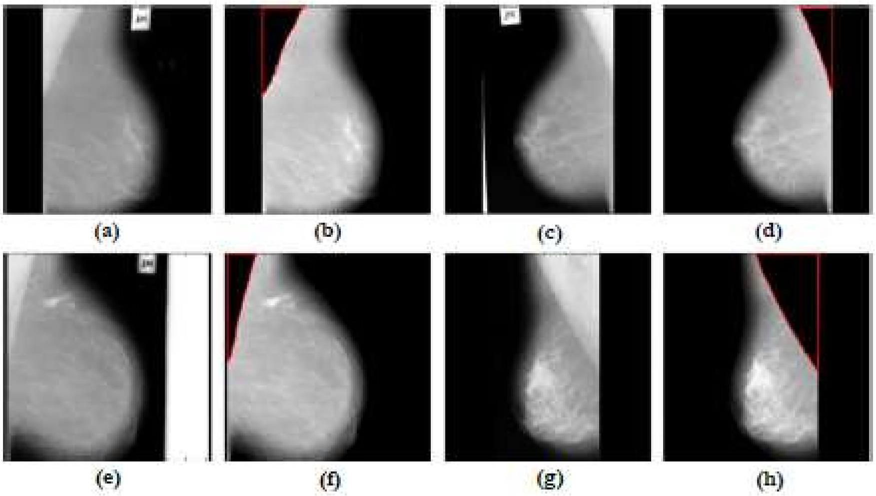


Figure 2.12: **Some of the successful preprocessing results obtained using the proposed algorithm. The ground truth contour of the actual pectoral muscle boundary is shown in red. (a), (c), (e) and (g) are images mdb006, mdb009, mdb132, and mdb249, respectively, from the mini-MIAS database[63].**

3.2.2.4 Graph Theory Based Methods

Partitioning image into regions has been proven as a difficult task in image segmentation for many reasons such as: regions are homogeneous as well variation in the scale of the image.Thus; several methods have been developed but no efficient result(method suits one situation).This latter problem has been solved by mapping image into graph[66].From the literature surveyed; several approaches have been presented on graph theory segmentation such as : Minimum Spanning Tree(MST) and Graph cut based method.We explained the usefulness of each method in digital mammogram segmentation in following paragraphs.

(i) minimum spanning tree(MST) :
MTS is an equivalent concept to shortest spanning tree which is explained in[67], MTS is used to cluster image based on graphs and that by producing connected groups of vertices(pixels)[66].
Ma et al[68] used MTS to identify the pectoral muscle in mammograms accompanied with Adaptive Pyramid Graph(AP is explained in[68]). The first is used to construct the graph of edge while the second is used to build pyramid graph of vertices.As a result the initial pectoral edge is indicated then it is refined using active contour approach.

(ii) Graph Cut based method :
Graph cut technique is used to segment region of interest, it consists of three major steps such are:formation of a graph, sorting of graph edges, and region merging[69][31].
Camilus et al.[69] presented an algorithm based on Graph cut approach to detect pectoral muscle. The process used features of pectoral muscle such as :

Distinguishable values of gray-scale and homogeneous features were considered as parameters to Graph cut based segmentation as a result ROI is segmented; as well pectoral muscle is identified.

Pectoral forms triangle shape at the top of image moreover; the edge may be approximated by straight line since the width of pectoral decreases gradually.these features were used to refine the detected edge using Bezier Curve approach.

3.2.3 Based on soft computing

Pectoral muscle suppression has been proven as a complicated task for many reasons, which poses the necessity to simplify this task. Soft computing is used to solve complicated problems, it consists of complementary elements fuzzy logic, genetic algorithm, neural computing and evolutionary computation[31]. A few researches have been presented on soft computing such as:

- Fuzzy Logic Based Method
- Genetic Algorithm Based Method
- Support Vector Machine Method

We discuss these methods in the following paragraphs.

3.2.3.1 Fuzzy Logic Based Method :

Since; the edge in image is abrupt variation in intensity within image, it is difficult to select pixels candidate in the edge. Fuzzy logic technique has been extended from the concept of partial truth, it is considered as a powerful approach to decision making[76][75].
Aroquiaraj et al.[75] used fuzzy logic approach and straight line technique to remove pectoral muscle region from mammogram. The proposed method was accompanied by several approaches to give more accuracy to the task such as: Mean Absolute Error (MAE), Local Consistency Error (LCE) , Tanimoto Coefficient (TC), Probabilistic Rand Index (PRI) and Hausdroff Distance (HD).

3.2.3.2 Genetic Algorithm Based Method :

Genetic Algorithms(GAs) use principles of evolution and natural genetics to solve complicated problems[70], that means problems would be solved in several stages as following :
First, possible solutions for a given problem are produced. After; each solution is evaluated provided as constraints imposed by the problem is respected. Then; best solutions are selected. Finally; new solutions are created by genetic manipulation operator[70][71].
Karnan et al.[72] exploited the benefit of GAs to detect pectoral edge. Digital mammogram is enhanced by using median filter approach, as well it is normalized. Pectoral muscle region is excluded to identify its edge, after; this latter was magnified by GA then; for more accuracy figure of merit was calculated to evaluate whether the detect edge is exact or not.

3.2.3.3 Support Vector Machine Method :

Support vector theory is designed for generalizing unseen data automatically by characterizing properties of learning machines[73].
Domingues et al.[74] used two steps to detect pectoral muscle contour automatically.
The first step was a prediction to endpoints of the contour, using pair of support vector regression models; while the second step was a delineation to the contour by computing the shortest path between endpoints of the contour.

3.2.4 Recent Researches

Pectoral muscle removal constructs big challenge till now, cause of characteristic of this region. recent researches have been presented to reduce imperfections of the previous methods which are an adaptive to variation in shape, density or tissue like as straight line and curved pectoral edge. We have explained the proposed recent algorithms in the following paragraphs such as based on Adaptive Gamma Corrections, Single Sided Edge Marking Method.

3.2.4.1 adaptive gamma corrections :

Gamma correction is considered as variation of histogram modification techniques obtained by several adaptive parameters γ [78].
Gardezi et al.[77] proposed an algorithm to remove pectoral muscle from mammogram by using two techniques. morphological operations were used to extract breast tissue from its background, while adaptive gamma correction technique was used to enhance mammogram according to tissues densities as well to provide a separation boundary between breast region and pectoral parenchyma.

3.2.4.2 Single Sided Edge Marking (SSEM) :

SSEM is new approach proposed by Toz et al.[79], based on the geometrical properties of pectoral muscle and its surrounding tissue. The result of this approach was refined by morphological operation to obtain the rough pectoral muscle region, then the boundary of pectoral was reconstructed by the linear interpolation method which has been redefined as a result the irrelevant region is well defined as it is illustrated in Figure2.13 .

(a) (b) (c)

Figure 2.13: **(a) Original mammogram. (b) Obtained mammogram image of rough pectoral muscle boundaries. (c) Mammogram image of the final borders obtained by linear interpolation.[79].**

4 Conclusion

Preprocessing and breast profile segmentation are crucial tasks. Thus; the obtained informations in mammogram must be accurate to detect anomalies.
We have described number of studies in preprocessing phase; which consists in removing noise, enhancement of mammograms and contrast manipulation, edge sharping, filtering. As we have undertaken studies to discover more about segmentation of breast profile which is divided into two steps; separating breast tissue from the background of mammogram and removing pectoral muscle. This latter remains a hard task till now for researchers because of the several characteristics of

this region ; thus we find in the literature several methods that have been proposed whether based on gray-level, texture features or soft computing.

3

Processing techniques

1 Introduction

Breast cancer is a serious and widespread disease that affects human body, but Women are more exposed to this type of cancer. protect and prevent the human from this disease has been proven as impossible issue since the cause of cancer disease remains unknown. Early diagnosis is crucial for effective treatment; as it needs accurate and more specific detection; which is currently possible by screening using only X-ray mammography, this latter is considered as a primary imaging technique for breast lesions detection[95].
However; reviewing mammograms by human observes is inaccurate. Thus; the enormous number of mammograms are generated in wide spread screening as consequence it is difficult to provide both accurate and uniform evaluation, some anomalies maybe missed as a result of fatigue[136], moreover misclassification to the detected anomalies.
Hence; there is a significant necessity of diagnosis cancer automatically to assist radiologist in improving the efficacy of screening programs and avoid misclassification. CAD systems are intended for diagnosis breast cancer automatically, it consists of CADe (Computer Aided Detection) which is used to detect abnormalities, and CADi(Computer Aided Diagnosis) which is used to classify abnormalities into benign or malignant.
We have discussed the utility of CAD systems in this chapter; which is divided into four sections. In the first section; we have reviewed different techniques that have been used in CADe system, while the second section illustrates CADi system, and in the third section we have explained image registration procedure which is used in diagnosis of cancer too, last section draws conclusion.

2 CADe System

Computer has become a crucial device for radiologist. It is used in the acquisition management, storage, and reporting of medical images beside that; computer programs have been developed to assist radiologist in detection potential abnormalities on diagnostic radiology examinations. This technique has been termed computer aided (or assisted) detection, commonly referred to as CADe[82].
Abnormalities of breast cancer maybe noticed as microcalcification, mass, architectural distortion or bilateral asymmetry[83]. There are several methods have been reported to detect these anomalies according to its characteristic, which have been explained in the following subsections.

2.1 MicroCalcification Detection

MicroCalcification (MC) is a tiny abnormal deposit of calcium scattered inside the breast, in the other hand; it is a small bright spots in mammograms[83]. There are several approaches have been proposed to detect MC; which are categorized into four categories; basic image enhancement methods, stochastic modeling methods, multiscale decomposition methods, and machine learning methods. Where each one uses a characteristic of MC to detect it[84][83].

2.1.1 Basic Image Enhancement

The common characteristic of MCs are brighter than its surrounding tissue in mammogram, hence; image enhancement techniques are used to improve the contrast of MCs as a result separating them from mammogram by an adequate threshold[83].
From the literature reviewed, the different ideas presented on basic image enhancement methods such as Nishikawa et al.[85]; whose used three steps to detect MCs.

The first step was a suppression of normal background of mammogram, by increasing the signal-ratio of MCs to noise-ratio as a result MCs are indicated in filtered image. In the second step; authors used variant methods to extract the indicated MCs from the filtered image. Third step is used to eliminate false positive signs (unconcerned regions).

2.1.2 Stochastic Modeling

The idea of stochastic modeling consists in using statistical difference of regions to detect MCs. For instance; Gurcan et al.[86] differentiate MCs from rest tissue in mammogram by computing the higher order statistics as they noticed that areas with MCs are characterized by non-Gaussian (nonzero skewness and kurtosis[86]) while areas with Mcs are noticed by a Gaussian-like distribution.

2.1.3 Multiscale Decomposition

Multiscale decomposition methods based on identifying MCs spot in several frequency content; one of the common methods that has been used in MCs detection is wavelet transform[87].
Strickland et al.[88] discussed undecimated biorthogonal wavelet transform to detect MC spots, whereas this latter are represented as a circular Gaussian shapes with variant width along different scales.

2.1.4 Machine Learning

Detecting MC spots in mammogram maybe performed by classifying pixel into its belong region. Hence; machine learning technique is used for automatic classification to pixels.
Yu et al.[89] presented an algorithm based on neural network approach to detect MC spots. Two stages were used; the first one was an identification to MCs by using mixed of methods consisting of wavelet components, gray-level statistic and labeling. While the second one was used to determine individual MC objects. Figure3.1 shows the result of their proposed algorithm.

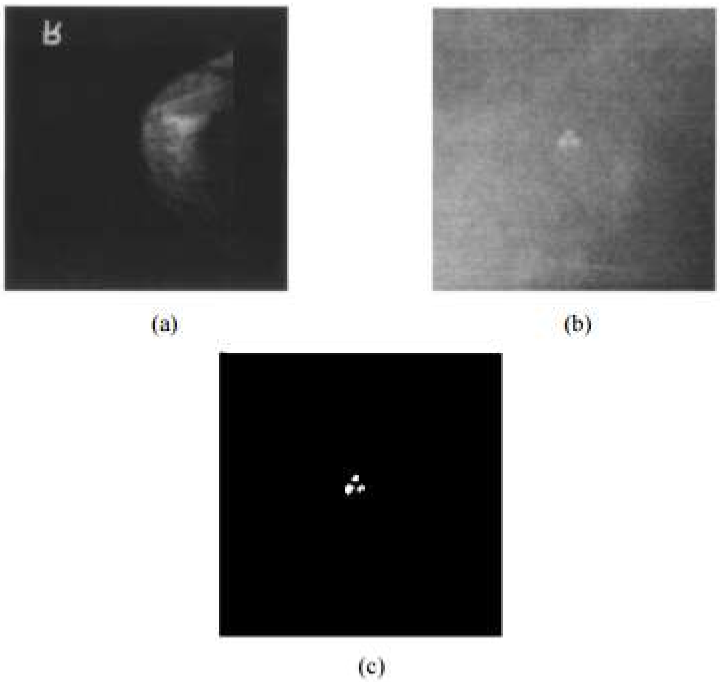


Figure 3.1: **. (a) Original mammogram . (b) A cluster of MCs in (a). (c) The ground truth image corresponding to the cluster of MCs in (b)[89].**

2.2 Mass Detection

A mass is defined as a space occupying lesion seen in different projections, as it is characterized by shape and margin. Nevertheless; mass detection procedure is based on two stages; suspicious region is detected firstly, then classifying it as a normal tissue or mass.
In the literature; generally the proposed methods that have been used to detect suspicious regions are based on pixel or region.

2.2.1 Pixel based

Kegelmeyer et al.[90][83] propose an algorithm based on pixel, where this latter is used to extract features which consist in Laws' texture and local oriented edge characteristic, then the extracted features were used to speculate a normal tissue from mass by using a binary decision tree classifier. Figure3.2 illustrates the result of this algorithm.

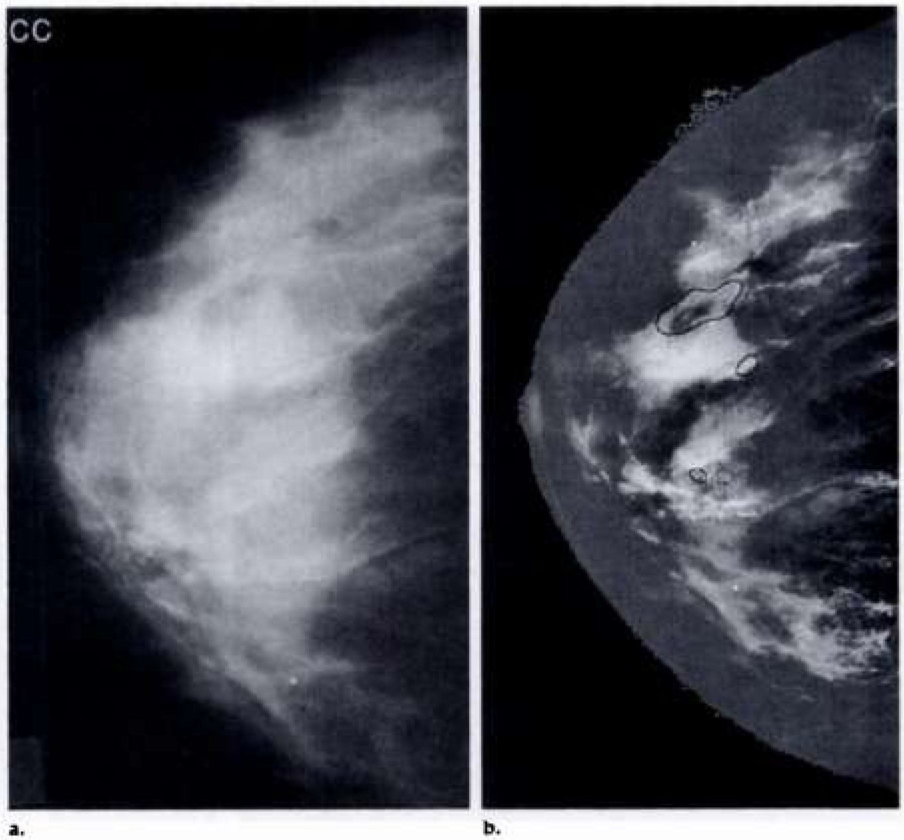


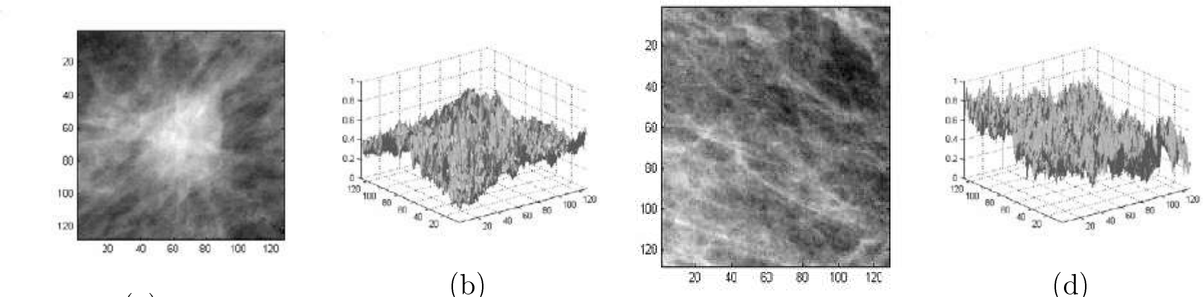
Figure 3.2: **Normal mammogram. (a) The original film version. (b) The mammogram in a with examples of typical false-positive findings (outlined areas) in the computer reports[90].**

2.2.2 Region based

Texture features play an important role in distinguishing between regions in image. Hence; Chan et al.[102] demonstrated that feasibility of using texture features is an effective way to discriminate regions containing spiculated or non-spiculated masses from those containing normal breast tissue. Authors used Spatial Gray-Level Dependence (SGLD) matrices to extract texture features. Then these latters were analyzed by linear discrimination approach to determine mass regions.

2.3 Detection of Architectural Distortion

Architecture Distortion is considered as the hardest breast cancer type to diagnosis by imaging modalities techniques, it is often missed during screening[83]. According to Bi-Rads[92] definition "The normal architecture (of the breast) is distorted with no definite mass visible. This includes speculations radiating from a point and focal retraction or distortion at the edge of the parenchyma."
Guo et al.[116] investigated the characteristics of architecture distortion by using Hausdorff fractal dimension approach; as for more accuracy authors used SVM classifier to distinguish between abnormal regions and normal ones. The result of their proposed method is illustrated in figure3.3
.



(a) (b) (c) (d)

Figure 3.3: **Sampled images and their surfaces. (a) "architectural distortion"; (b) surface of "architectural distortion" sample; (c) "normal tissue"; (d) surface of "normal tissue".**[116].

2.4 Detection of Bilateral Asymmetry

Asymmetry between the left and the right mammogram of one subject is an important radiological sign of breast cancer[94]. This asymmetry is indicated by the presence of a greater volume or density of breast tissue without a distinct mass, or more prominent ducts in one breast as compared to the corresponding area in the other breast[92][83].
Miller et al.[94] proposed an algorithm based on anatomical features to detect breast asymmetry. Where semiautomated texture based procedure were used to segment the glandular tissue, while measures of shape and registration cost between views were used to detect asymmetry.

3 CADi System

The idea of automatic studies on analysis of medical images has been emerged in 1960s. Where the computer was used to detect anomalies, thus; the computer output is more accurate than the result of human beings. However; in automatic diagnosis, computer output can be utilized by radiologists as an aid in making-decision but not replace them. This concept known as computer aided diagnosis, which is referred as CADi or CADx[96].
Diagnosis of breast cancer automatically based on three steps such are: **feature extraction**; which is used to extract characteristics of the detected anomalies, **feature selection**; which is used to select the effective characteristics that makes the differentiation between positive and negative anomalies, and **classification**; which is used to indicate the type of anomalies depending on the selected characteristics.
In this section; we have explained the process of CADi on the most common types of breast cancer as it is illustrated in the following subsections.

3.1 Feature Extraction

Discrimination between malignant and benign mammographic lesions acquires a recognition to several patterns that indicate both types. Hence; extraction of features that characterize each one of type is recommended indeed.
The extracted features from digital mammogram can be textural, structural or statistical[113]. We have reviewed these features in recognizing the two types MicroCalcification and Mass.

3.1.1 MicroCalcification

To identify the type of microcalcification; radiologists use several characteristics such as: the morphology and location of the cluster, the morphologies of the individual calcification particles, and the distribution of particles within the cluster. Hence; there are four major classes of features that have been deduced from radiologists' strategy, such are: Features that represent the morphology and location of the cluster, features that characterize the morphology and optical density of the individual calcification particles, features that describe the spatial distribution of the individual particles within the cluster, and finally features that represent the texture of the background tissue the calcifications are embedded in[98].
From the literature; variant methods have been proposed based on this taxonomy.

3.1.1.1 Morphology of the cluster :

The features that describes the morphology of microcalcification's cluster consist in: the area and the perimeter, the circularity, rectangularity, orientation, and eccentricity of the cluster. Where the basic one that commonly used by researches consists in the number of individual calcification particles in the cluster which can trivially be obtained from a segmentation of the individual calcifications(i.e. [105][98]). Beside that, other researches prefer to calculate the convex hull of the centroids or the contour pixels of the particles in a cluster to represent the cluster's shape[98].

3.1.1.2 Location of the cluster :

The location of the cluster may be taken as a feature to differentiate malignant lesions from benign ones. Thus; it has been proven that malignant lesions are more often located in the upper outer quadrant than other quadrants of the breast. For instance; Veldkamp et al.[106] extracted features of the clustered microcalcifications based on their location,and that by calculating the relative distance of a cluster to the pectoral muscle and the breast edge.

3.1.1.3 Morphology of individual calcification :

Beside using features of cluster's morphology; the characteristics of morphology of the individual calcification are obtained to distinguish between the two types. The nature of these features are statistical, which consist in mean, standard deviation, minimum, maximum, and medium. in the other hand; there are several approaches that have been used to describe the individual calcification such as: normalized Fourier descriptors[107] and Radon transform[108].

3.1.1.4 Optical density of individual calcifications :

Mammographic anomalies are noticed by high density. Hence; this latter is used as a measure to discriminate between anomalies. Statistical features are used in discrimination such as the mean and variance of the gray values ,and the contrast of individual particles[98]. Another features are considered to use like as thickness and volume of individual particles based on a mathematical model of image formation[109][98].

3.1.1.5 Distribution of individual calcifications :

The spatial distribution of individual calcifications within a cluster is described by statistical features, which consists in the distance between individual particles, the distances of the particle centroids to the cluster centroid, and the mean number of nearest neighbors of the particles in a cluster which is discovered by Leichter et al.[110].

3.1.1.6 Texture of the background tissue :

The presence of calcifications in breast, are noticed by a change in the texture of mammogram's background tissue. Hence; the type of features that used in classification of calcifications are textural. In the literature reviews, variant approaches have been used to extract texture features such as: gray-level co-occurrence matrices,Texture analysis based on wavelet packets[111] and multi-wavelet packet features[112].

3.1.2 Mass

Once the masses are separated from its surrounding normal tissue, a collection of features are extracted from its characteristics in order to be used in discrimination between benign and malignant masses[97].
Mammographic masses are characterized by shape, margin, texture, optical density and embedded calcification . Thus; Variant methods have been proposed based on these characteristics to extract features[98].

3.1.2.1 Shape:

Effective representation of the mass' shape by features lead to accurate result of classification. Thus; several groups used morphological features to represent the shape of mass such as: area and the perimeter as well as the circularity, rectangularity, orientation, and eccentricity[98]. An other kind of feature can be used to represent shape's mass like as pixel border. That by calculating the Euclidean distance of each pixel on the object contour to the object's centroid[99][98].

3.1.2.2 Margin characteristic

Margin characteristics are considered as an important measure. Hence, variant methods used pixels that are include the margin of mass to extract features. Rubber Band Straightening Transform (RBST)is an approach proposed by Sahiner et al.[100], it consists in transforming a band of margin pixels onto the Cartesian plane to extract features.

3.1.2.3 Optical density

Optical density of mass is represented by different features which consist in the mean gray level of the mass region, its local contrast, or by texture analysis features.

3.1.2.4 Texture features

Textures are a higher order features, thus; they are not used by radiologist. There are several approaches that have been proposed to extract features from texture of mass such as commonly used : gray-level co-occurrence matrices, gray-level run-length metrics and wavelet decompositions.

3.1.2.5 Embedded Calcification

It has been noticed that the region of segmented masses contain calcification. Hence; this noting is used as a feature to characterize mass. Shi et al.[114] proposed the number of microcalcifications presented in mass' region as a feature.

3.2 Feature Selection

An increase in the feature space's dimension leads to a decrease in the performance of discrimination task. Thus; the distribution of instances becomes increasingly sparse. In order to obtain accurate result in discrimination between malignant and benign lesions; it is necessary to select

the effective features that determine the cardinality of lesion. That, because features employed in CADx are various as consequence we find features that are highly significant for the discrimination of mammographic lesions and others are redundant.
Selecting ideal features in CADx system consists of two classes, which are: *filter approach*; which is used to select features based on a metric, and *wrapper approach*; which is employed to wrap the classifier in order to find optimal features subset and that by following a nonexhaustive searching strategy[98]. A lot of effort have been used to select the effective features, we mention the two techniques; which are well known:
forward selection (FS) and backward selection (BS), these two approaches works together but in different strategy in selecting features, where the first one starts its process by one variable then then adding other ones progressively incorporated into larger and larger subsets, while the second one takes all features as a set then deleting temporarily features which have largest value of selection criterion[127].

3.3 Classification

Classification of mammographic lesions into malignant or benign is based on a set of variant lesions which are already identified, this concept known as supervised classification[98].
There are several methods have been proposed to determine the type of mammographic lesions, we mention the ones that are commonly used in CADx system such as:

- Artificial Neural Networks (ANNs);
- K Nearest Neighbors (KNN);

We have explained the utility of each method in identifying the cardinality of mammographic lesion.

3.3.1 Artificial Neural Networks (ANNs) :

Numerous studies have been used to develop intelligent systems; in order to solve some problems like as pattern recognition. Some of these studies are inspired by biological neural network as; Artificial Neural Networks ANNs, which is used in the part of classification to assign an input identified by a feature vector to one of many prespecified classes[101].
For instance; Chan et al.[102] exploited the benefit of ANNs to predict whether the presence of microcalcifications is associated with malignant or benign pathology. First, Authors used Spatial Grey-level dependence (SGLD) matrices to extract texture features from mammograms, then; a stepwise feature selection is used to maximize the separation between malignant and benign classes, finally; ANNs classifier was trained and tested to recognize malignant and benign features as a result classifying the receiver pattern into one of the two classes.

3.3.2 K Nearest Neighbors (KNN)

The task of classification consists in predicting the cardinality of unlabeled class based on observation of discrete class labels. K Nearest Neighbors, also known as Nearest Neighbors Classifier; it assigns target pattern to one of the k nearest neighbor pattern, depends on distance measure like as Minkowski[103].
In the literature review, KNN was used for mammographic masses classification[104]; where ten texture and shape based features were used as a dataset to train nearest neighbors classifier.

Note

Feature extraction and classification are steps include both systems CADe and CADx , as they take the same definition(as it is defined above). However, the task is different, for CADe the two steps are used to differentiate normal tissue from abnormal tissue while CADx is used to distinguish between malignant and benign mammographic lesion.

4 Image Registration

Image Registration is known as spatial mapping [116]between two or more images, which are taken at different time, from different imaging modality techniques or from different viewpoints[115].On purpose of finding an ideal transformation or mapping approach, which will align the points in one image to their corresponding points in the other. Thus, search strategy is required to optimize similarity measure, which is used to determine the quality of the matching between images. As well feature space, which contains information concerning images like as the raw pixel values[116]. From the literature review, Proposed methods are classified into intra-image modality registration and inter-image modality registration.

4.1 inter-image modality registration

In this type of registration, the matching is applied on the same type of imaging modality. We have reviewed some methods that have been proposed for x-mammogram registration technique, which are classified into two basic categories feature-based and intensity-based, while other methods are an hybrid of the two categories.

4.1.1 Feature Based

The most common features that have been used in X-ray mammography registration techniques are *Control Points* which can be selected manually or located automatically on the nipple positions, breast boundary, or linear structures within the breast. However; selecting these features remains difficult task because of the compressibility of the breast that makes it a nonrigid structure. Moreover, it requires identification of anatomical features points in each mammogram to be registered.

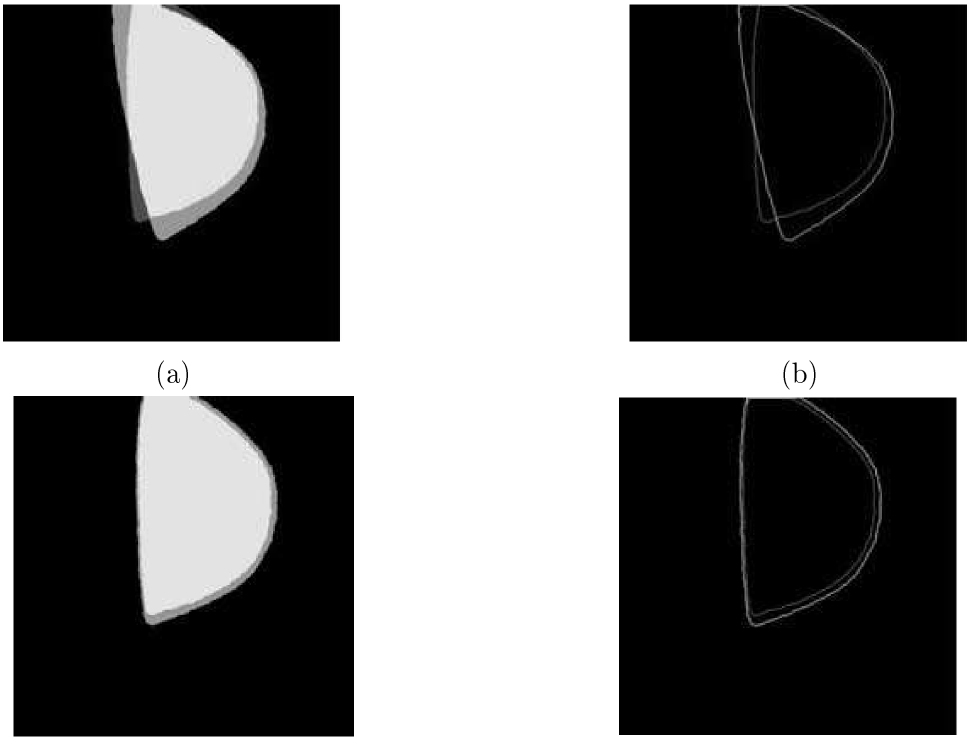
Sivaramakrishna et al.[117] approached this problem by mapping texture of mammogram by texture transformation as a result control points are selected. Then, to match the two images TPS-registration(Thin Plate Spline) were applied. TPS is a technique used "to interpolate data as a surface of an infinite plate under the imposition of point loads".

4.1.2 Intensity Based

Intensity-based registration methods use pixel values to match images. Brzakovic et al.[118] proposed an algorithm based on three steps. First; the principal axis methods have been used to register the two mammograms. Then, hierarchical region-growing technique was used to partition the mammograms. Finally; a various of methods such are rotation, translation and scaling were used for the result of partitioning registration.

4.1.3 Hybrid registration

In order to obtain an efficiency result, some researches preferred to use a combination of both registration technique feature-based and intensity-bases. Such as Suri et al.[119] who proposed an approach based on mask for registration strategy using adaptive segmentation technique. This mask were used to extract ROI as a result mutual information were used.



(a) (b) (c) (d)

Figure 3.4: **Overlaid images. First row: before registration; Second row: after registration[119].**

4.2 intra-image modality registration

The benefit of image registration technique between different imaging modality consists in taking accurate decision about the detected anomalies. Thus, this latter in a one type of imaging modality may not be clear unless combining its information with another type of imaging modality. In the literature review; X-ray mammography has been combined with MRI and Ultrasonography.

4.2.1 X-ray mammography and MRI registration

Since X-ray mammography has spatial whilst MRI provides a good tissue specificity, registration of both technique provides complementary pathological information for breast cancer diagnosis and surgical/treatment planning[120]. For instance; Martí et al.[121] presented an novel approach to correspond areas in X-ray and MR images of the same breast.

4.2.2 X-ray mammography and Ultrasonography registration

Detecting abnormalities in mammogram remains big challenge, specially in women with breast dense. Missing abnormalities by Using a combination of ultrasound and mammography has been proven that it is much smaller than using mammography alone, as its efficacy has been investigated by Malur et al.[122].

Note

Architectural distortion and bilateral asymmetry are considered as a sign of breast cancer[130][131]. Then; once it has been detected one of these two signs, another process of searching about abnormalities including this change, is launched to identify its cardinality. Few efforts have been used for diagnosis of these two sign of breast cancer. In the literature review; architectural distortion is also more difficult to diagnose because it can be subtle and variable in presentation[133], it is classified as malignant when the change integrates visible lesion such as mass and calcification, while it is benign in the presence of scan and soft tissue damage due to trauma[129]. As for bilateral asymmetry, it has been used to detect mass as Ericeira et al.[132] did.

5 Conclusion

Since the reviewing of mammograms by human observes may be inaccurate result, advanced researches have been investigated to make it accurate such as CADe system; which is used to detect abnormalities that may be missed by traditional detection, CADx system; which is used for automatic diagnosis , in other words classifying abnormalities into malignant or benign. More care about the accuracy of automatic diagnosis, image registration technique has been proposed to collect as much as possible information from several modality techniques to take decision whether the diagnosed mammogram contains cancer or no.
We have explained these advanced techniques for the early detection of breast cancer automatically, by referencing methods that have been used in each technique. As we noticed that the accuracy of the results of breast cancer prediction is still remains a big challenge for these techniques.

4

Detection of pectoral boundary & identification of intensity

1 Introduction

Detection diseases in their early stage can increase treatment options and survivability of patients, among these diseases; breast cancer. Mammography is considered as the best tool in identifying the subconscious lesions in breast[95]. However; the interpretation of this kind of medical imaging techniques remains a hard task as well it could lead to inaccurate detection. Thus, dense breast in mammogram may contain abnormal structures which are similar to the normal ones as consequence reading mammogram is a big challenge[136].
Before detecting abnormalities, the quality of mammogram must be improved as well irrelevant elements(radiopaque artifacts and pectoral muscle) must be segmented out from mammogram.
This chapter consists two main parts, the first one describes detection the boundary of pectoral muscle , while the second part represents extraction subconscious lesions in breast.

2 Material

To evaluate our methods, mini-MIAS[123] database is used. It is a reduced version of the original digital mammography database, this latter has been produced by Mammographic Image Analysis Society MIAS. It contains 322 mammograms and include the common abnormalities encountered in screening as well as the normal cases. All mammograms have been digitized to spatial resolution 200 μm and padded to obtain images with size of 1024×1024 pixels.
Detailed information[124] offered by mini-MIAS database consists in :

- referenced number of each mammogram
- Character of background tissue: Fatty F, Fatty-glandular G or Dense-glandular D (see 2nd column of Table4.1) .
- Type of abnormality represented in mammogram: Calcification CALC; Well-defined/circumscribed masses CIRC; Spiculated masses SPIC; Other, ill-defined masses MISC; Architectural distortion ARCH; Asymmetry ASYM and Normal NORM. The 3rd column of Table4.1 represent these information.
- Severity of abnormality Benign B or Malignant as it is illustrated in fourth Table4.1.
- Image-coordinates of center of abnormality x,y (5th ,6th column).
- The seventh column of table4.1 represent approximate radius (in pixels) of a circle enclosing the abnormality.

mdb001	G	CIRC	B	535	425	197
mdb009	F	NORM				
mdb058	D	MISC	M	318	359	27
mdb081	G	ASYM	B	492	473	131
mdb115	G	ARCH	M	461	532	117
mdb145	D	SPIC	B	669	543	49
mdb218	G	CALC	B	519	629	8

Table 4.1: Detailed information about Mammograms[124]

3 Pectoral Muscle Boundary Detection

Pectoral muscle region is appear as a triangle at the top of mammogram in MLO view, this appearances can bias the result of CAD algorithms. Hence; removing this region from mammogram is very important. We have proposed a novel approach based on the similarity between intensity to detect pectoral muscle boundary. The following subsections illustrate the contribution of our method.

3.1 Challenges

The region of pectoral muscle has several characteristics that make its removal from mammogram a big challenge and inaccurate sometimes. These characteristics consists in[31]:

- The absence of this region in some cases ;
- Pectoral muscle region and breast tissue are similar in textural information ;
- The boundary of pectoral muscle is variant from mammogram to another, it is concave in one and convex in another one;
- The mammograms have been taken from different patients
- Varying position, shape, size and texture from mammogram to another ;
- Overlapping features which lead to unclear and superimposed boundaries.

Hence; to find an approach which segment totally pectoral muscle region over a wide variety of mammogram , still remains big challenge.

3.2 Hypothesis

Various approaches have been proposed to extract pectoral muscle region, whether based on gray level thus; pectoral muscle is presented by a dense and high intensity in digital mammogram, based on texture feature since this latter identify its region; or based on soft computing because its removal considered as complicated task.
However; the majority of these methods have been employed on the region itself; whilst *accurate removing to the pectoral muscle region based on accurate detecting to its boundary.*

3.3 Similarity Between Intensity

In filed of Natural Language Process(NLP) and Information Retrieval (IR); words are characterize by polysemy and synonym, hence; choosing adequate word that refers to the intended meaning create big challenge to the human. In order to overcome this challenge, there was a necessity of finding out relationship between words. Thus ontologies are proposed to provide a formal specification of a shared conceptualization between words (i.e.taxonomy). As well methods to measure semantic similarity between words and concepts such as based on Paths (the length of the path that connect the concepts and their positions in the ontology), or based on Information Content(IC); which is defined as the probability P(c) of the presence of a concept (c) in ontology .The IC of concept can be quantified as follows .

$$IC(c) = \log^{-1} P(c). \tag{4.1}$$

Different approaches have been proposed for measuring semantic similarity between words such as Jiang et al.[125] who defined similarity between words as :

$$Sim(w_1, w_2) = IC(c_1) + IC(c_2) - 2 \times IC(LSuper(c_1, c_2)). \tag{4.2}$$

where c_1, c_2 refers to the set of possible senses for word in ontology w_1, w_2 respectively, while LSuper (lowest super-ordinate) refers to the shortest path from c_1 to c_2 in ontology.
We have proposed to use features of semantic similarity between words to quantify the similarity between pixels in intensity for gray-level images as it is illustrated in equation 4.3.

$$Sim(p_1, p_2) = IC(P_1) + IC(P_2) - 2 \times IC(LSuper(P_1, P_2)) \tag{4.3}$$

IC(P) refers to the probability of the presence of the pixel P, whilst in a selected region (i.e slid window) while LSuper refers to the the nearest intensity value to both pixels(P_1 and P_2) in the selected region.
Example : The similarity between $P_1 = 238$ and $P_2 = 240$ in the following part of gray-scale image is :
$Sim(P_1, P_2) = IC(P_1) + IC(P_2) + 2 * IC(LSuper(P_1, P_2))) = 0$
$IC(P_1) = -0.3$
$IC(P_2) = -0.3$
in this case $LSuper(P_1, P_2) = 238$, so $IC(LSuper(P_1, P_2)) = 0.6$.

238	240	241
238	240	241
239	241	241

the $IC(LSuper(P_1, P_2))$ is one of its neighborhood ,it is selected as follows: the matrix is reshaped to array which takes ascent order. then, the nearest intensity is the previous position of the smallest intensity between (P_1, P_2) in array.
The result "0" means that $P_1 and P_2$ are similar in intensity .

3.4 Proposed Method

Pectoral muscle removal is an important step before applying the CAD techniques; because this latter will be applied only inside the breast region. Accurate removing of pectoral muscle based on accurate detection of pectoral muscle boundary.
Our method consists of three steps, firstly; the quality of image is improved , secondly; the radiopaque artifacts are eliminated, finally; the boundary is identified. The following block diagram (figure 4.1) illustrates these steps.

3.4.1 Noise Removal

For the sake of improving the quality of image; we have used median filter approach, which is considered as an effective method for medical images noise.

3.4.2 Radiopaque Artifacts Suppression

On the part of removing radiopaque artifacts from mammogram; we have implemented the algorithm that has been proposed by Nagi et al.[22], it is based on morphological operations.

Input: Original mammogram

Noise removal using median filter

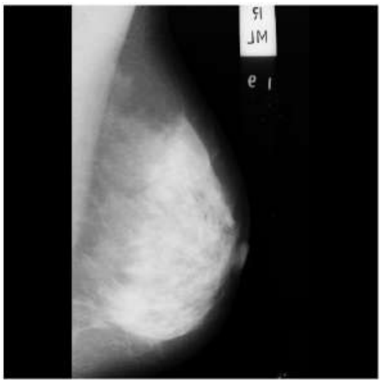
Artifacts suppression using Morphological operations

Pectoral muscle boundary detection

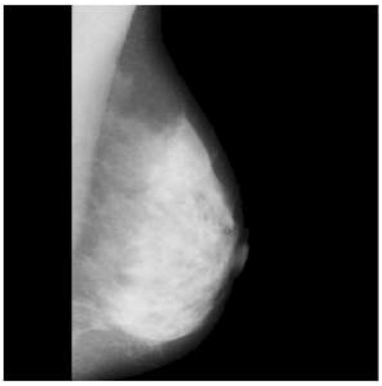
Output: pectoral muscle boundary

Figure 4.1: Block diagram for pectoral muscle boundary detection

First;the image is converted to binary image ,next; this latter is labeled using threshold T=18, then; the region which has big number of pixels has been extracted, after; morphological operations were applied to remove noise produced due previous steps, finally; the resulting image is multiplied by original image to obtain breast tissue. Figure4.2 illustrates all mentioned steps.

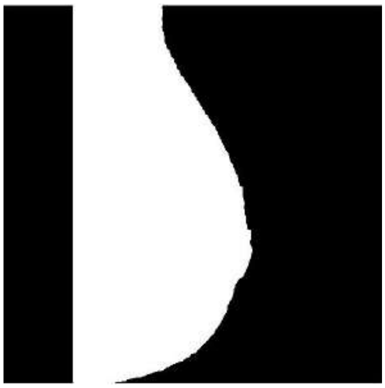


(a) **original mammogram**



(b) **mammogram without radiopaque artifact**

(c) **binary image**



(d) **large region**

Figure 4.2: **Artifacts Suppression using morphological operation**

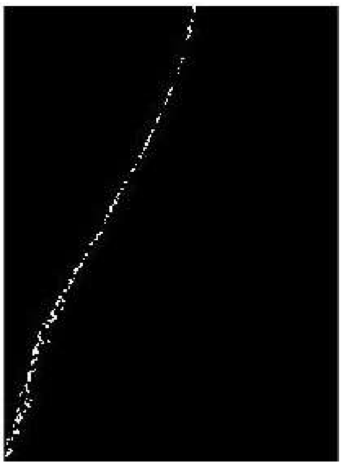
3.4.3 Pectoral Muscle Boundary Detection

In order to detect the boundary of pectoral muscle, we have proposed an algorithm based on similarity between pixels in intensity. Thus, a pixel include pectoral boundary has been selected, then the similarity between this latter and the rest of image pixels is computed after extracting the neighbors of each one,finally; the pixels which have similarity value less than threshold (0.2 which satisfy the majority of images) are selected, as a result pectoral muscle boundary is detected.

Algorithm
Begin
input: original image
output: pectoral boundary
if the breast orientation is left **then**
  fillip input
end if
Ps $\leftarrow$ Select a pixel that include pectoral boundary;
Enhance input ;
Foreach : pixel P in input
$region \leftarrow extractneighborsofPi$;
if $similarity(Ps, Pi, region) < 0,2$ **then**
  $Input(i) \leftarrow 1$;
else
  $Input(i) \leftarrow 0$;
end if;
Endfor.

The result of our method consists in binary image, where the white pixels refer to pectoral muscle boundary. Figure 4.3 illustrates the result.

(a)

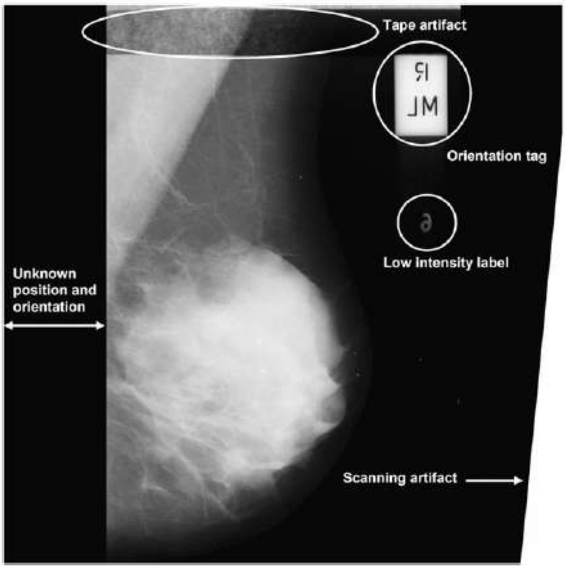


(b)

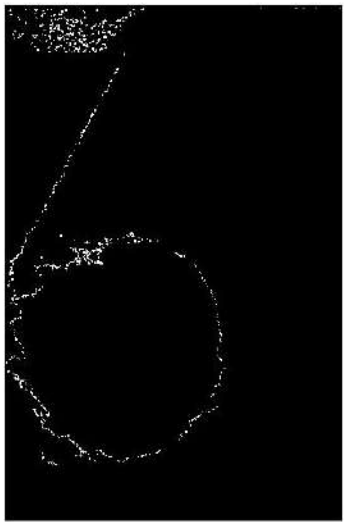
Figure 4.3: **Pectoral muscle boundary**

3.5 Experimental Result & Discussion

The proposed method has been applied on Mini-MIAS database, as it detects the edge of pectoral muscle even in the presence of some kind of artifact such as Top artifact (see figure 4.4) which still remains challenge.



(a)

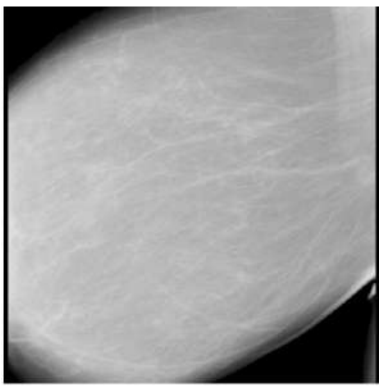


(b)

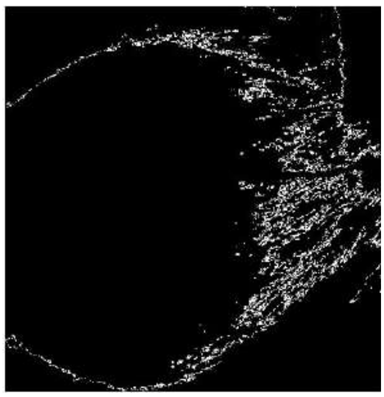
Figure 4.4: **(a) Common artifacts occurred in MIAS database images[126], (b) The result of our proposed method in the presence of top artifact**

However, Our approach faces some problems which consist in :

- Sometimes, the breast profile boundary is delineated, thus; the intensity of both boundaries are similar(see figure 4.5 & 4.4).



(a)



(b)

Figure 4.5: **breast profile boundary**

- The pectoral muscle boundary is not always detected because the selected pixel is not include the pectoral muscle boundary (see fig.4.6).

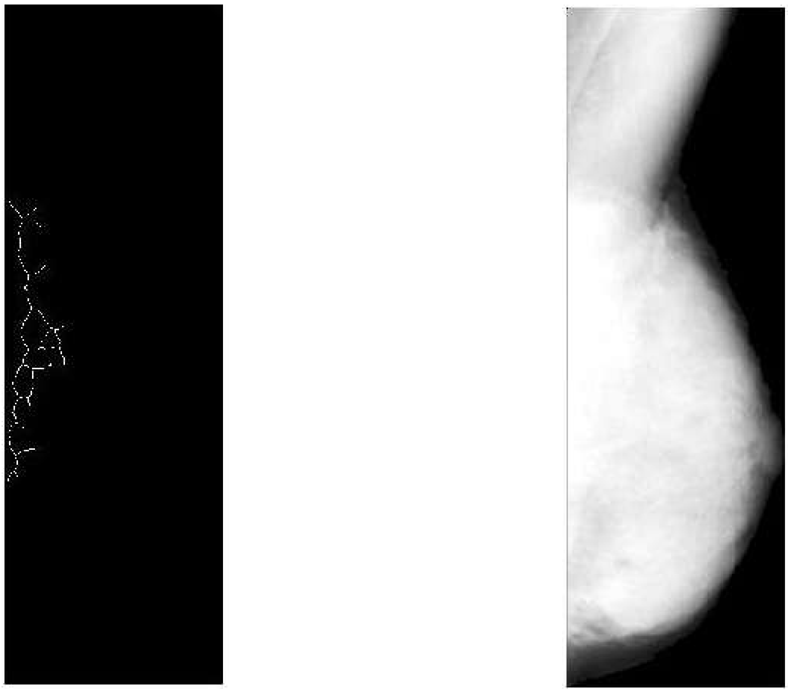


Figure 4.6: **invalid pectoral muscle boundary**

4 Identification Of Intensity

Abnormalities in mammogram are noticed by high intensity region within its surrounding regions. Thus, we have exploited the lack of our method which consists in delineating these regions to determine the cardinality of intensity.
The same algorithm has been used to select pixels which have similarity value higher than the threshold, then; the target region is isolated from resulting image, finally; features are extracted to classify it as a normal intensity or abnormal one.
we explain the process of features extraction and classification in the following subsections.

4.1 Feature Extraction

In order to distinguish between normal regions and abnormal ones, characteristics of each one are extracted. we have used statistical features thus, it has been previously proven that these kind of features give more information since they represent the gray level intensity measures. Moreover, they are useful in making decision for classification[135].

Feature	Expression	Measure of texture
Mean	$\mu = \frac{\sum_{ij} x_{ij}}{N}$	Mean pixel intensity
Standard Deviation	$\sigma = \sqrt{\frac{\sum_{ij}(x_{ij}-\mu)^2}{N}}$	The standard deviation of pixel intensity in the region of interest
Smoothness	$R = \frac{1}{(1+\sigma^2)}$	Measures the relative smoothness of intensity in a region
Skewness	$\frac{\sum_{ij}(x_{ij}-\mu)^3}{N\sigma^3}$	A measure of the asymmetry of the pixel values around the image mean
Uniformity	$\sum_{i=0}^{L-1} P(i)^2$	Measures the uniformity of intensity in the histogram
Kurtosis	$\frac{\sum_{ij}(x_{ij}-\mu)^4}{(N-1)\sigma^4}$	A measure of whether an image's intensity distribution is peaked or flat relative to the normal distribution
Average histogram	$AH_g = \frac{1}{L}\sum_{i=0}^{L-1} N(i)$	Estimation of the probability of occurrence of a gray level
Modified standard deviation	$\sigma_m = \sum_{ij}(x_{ij}-\mu)^2 P(x_{ij})$	A measure of average contrast
Modified skew	$MSK = \frac{1}{\sigma^3}\sum_{ij}(x_{ij}-\mu)^3 P(x_{ij})$	A measure of the asymmetry of the pixel values around the image mean

Table 4.2: Summary of statistical features[135]

4.2 Classification

Since we have used statistical features, it was necessary to choose a classifier that adapts with such type of features. Support vector machine(SVM) is a supervised machine-learning tool, it is based on statistical result learning theory. It is used to solve many problems such as classification; especially binary ones, and that by finding an hyperplane that segment the classes as consequences classification errors are minimized[127][135].
As the human thought, SVM learns by examples. Each example is represented by m numbers of data points($x_1, x_2, ..., x_m$), followed by label, which is in our case 1 or 0; 1 representing *normal* case while 0 representing *abnormal* one. The two cases are separated by generating the training examples[135].
SVM Classify an input into the nearest one of classes, thus; each example is represented by a vector(its data points) in dimensional space.

4.3 Training Data

The training data includes the class attribute and 9 attributes which consist in the statistical features, which are extracted from 62 mammograms from mini-MIAS database, including 15 normal mammograms and 47 abnormals ones, which include all abnormalities introduced by the database(i.e. Spiculated masses).
we present instances from dataset.

Mammogram	Mean	St-dev	Smoothness	Skewness	Kurtosis	Mod-St-dev	Mod-Skew	Avg.Histogram	Uniformity	Class	
mdb004	51.713	99.045	0.9999	1.3992	2.972	6.949e+05	69800	1947.6	0.6186	1	
mdb320	39.102	89.204	0.99987	1.8474	4.4249	1.1051e+6	68992	2345.8	0.70439	1	
mdb002	24.562	72.906	0.99981	2.6423	8.0165	1.7072e06	58602	2533.7	0.80653	0	
mdb184	8.722	45.497	0.99952	5.0357	26.415	3.0475e+06	35394	2373.3	0.93039	0	

Table 4.3: Extracted statistical features

4.4 Testing Data

Our testing data contains 44 mammograms, where 22 are normal and the rest are abnormal. we trained SVM classifier to give 0 to the abnormal intensity and 1 to the normal one. The result is illustrated in the following table4.4.

Type	Normal	Abnormal	Total
Class	22	22	44
Recognizes results	14	16	30
Classification rate	63.64%	72.73%	68.20%

Table 4.4: Result for breast intensity classification using SVM classifier

4.5 Discussion

The methodology that we suggested focus just on the regions with high intensity, thus we extracted statistic features that describes intensity which consists in mean, standard deviation, skew, kurtosis, modified standard deviation, modified skew, uniformity and average histogram. As consequences the classifier accuracy obtained is 68.20%, as we expected that the accuracy can be increased more, because the majority of abnormal mammograms which are classified as normal, they contains MISC which refers to other ill-defined masses. As for the wrong classification to normal one remains to the absence of a clear intensity regions in the tissue of breast.

5 IOBI System

The intensity of normal and abnormal region are indistinguishable for radiologist. In order to solve this problem; we propose IOBI system(Identification Of Breast Intensity). It helps radiologists to recognize the cardinality of intensity , whether it is normal or abnormal one.
IOBI system enables radiologist to upload the mammogram , that by pressing "*choose img* " button, then pressing "*High Int Regions*" button to extract regions with high intensity. As because pectoral muscle region characterized by high intensity, it is extracted by the one presented in the breast tissue, hence; it is necessary to isolate the breast intensity region, and that is done by "*Target Region*" button.
To identify the selected intensity region, IOBI presents two button, one to extract features corresponding to the region "*Extract Region*", while the other one gives the adaptable class to region based on the result of the extracted features "*Class*".
Figure4.7 illustrates what we have described in the previous paragraph.

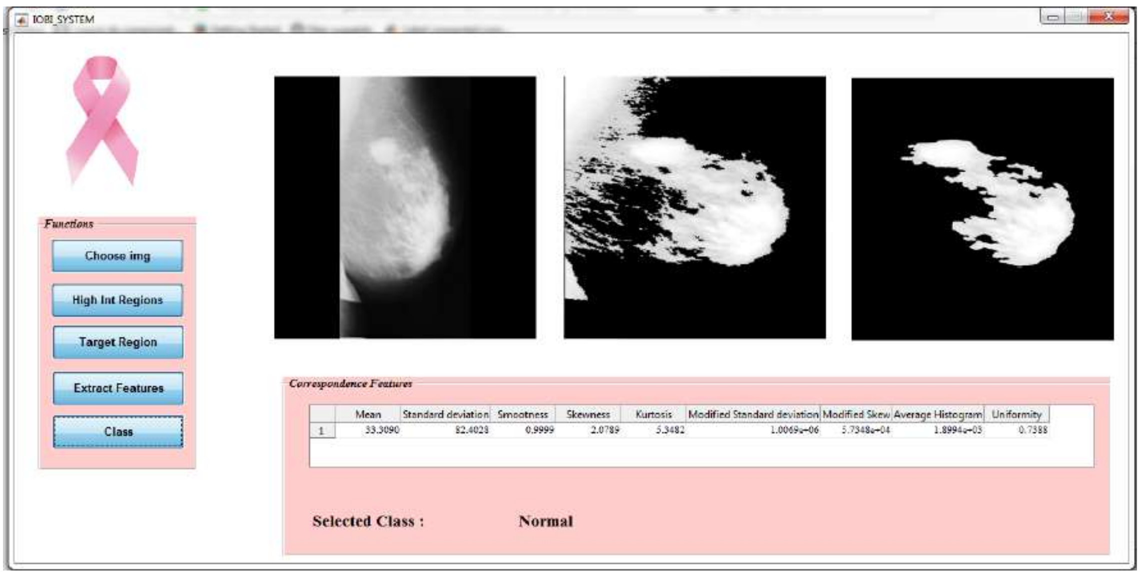


Figure 4.7: **Result of IOBI System**

6 Conclusion

In this chapter, we presented a method which measures similarity between pixels in intensity as consequences we could detect the boundary of pectoral muscle as we used this method to extract regions with high intensity in mammogram. As we have investigated that accurate removing of the pectoral muscle based on accurate detection of its boundary.

General Conclusion

General Conclusion

This thesis has focused basically on developing an approach for detecting the edge of pectoral muscle, which is an irrelevant element that must be removed from mammogram thus, its presence in mammogram can bias the result of CAD algorithms. As because abnormalities some times seem as a normal breast tissue, we proposed a IOBI system to distinguish between normal and abnormal tissue.

But before suggesting the proposed solutions, we reviewed the characteristics of some of image modality techniques that have been used to visualize inside breast which consist in X-ray Mammography, breast ultrasound, Magnetic Resonance Imaging(MRI), Positron Emission Mammography(PEM) and Computed Tomography Laser Mammography(CTLM) as well presented the advantage of each technique, its limitation in detecting breast cancer and its effect on human being physically and mentally. As a result we concluded that X-ray mammography is best examiner one for women thus its effect are less than the others except ultrasound, while this latter cannot be used as an effective one because it is used for women aged less than 40 years.

Then, we discussed the proposed methods that have been used to prepare mammogram for further process(CAD algorithms). The preparation consists in improving the quality of mammogram which is affected during acquisition, Among the methods that we cited, median filter which is considered as an effective one that restore mammogram from disturbances since it remove straight line which the mammograms are well known by this kind of noise. Removing irrelevant elements from mammogram is crucial task thus, cancer is located inside the breast, these elements consist in radiopaque artifacts which are removed by morphological operations, pectoral muscle region which its removal from mammogram remains big challenge because of its various characteristics such as varying position, shape, size and texture from mammogram to another. Hence, we found different methods based on different characteristics of this region such as based on gray-level, based on texture features, or on soft computing. The majority of these methods were used to remove totally pectoral region where we have noticed that *accurate removal of this region based on accurate detection of its boundary*, hence we have proposed to detect this boundary by computing the similarity between pixels in intensity.

Finally, we reviewed techniques that have been suggested to detect and diagnosis breast cancer automatically which consist in CADe, it is used to detected anomalies, CADx, it is used to identify the type of anomalies(benign or malignant) and image registration technique which is proposed to combine information from several image modalities in order to gather more information about anomalies. As because breast cancer has different views in mammogram such as mass, microcalcification, architecture distortion and bilateral asymmetry, the literature review contains various methods and process but till now the performance of these techniques are in need to be increased. Our proposition in this part consists in identifying the presented intensity in mammogram, thus anomalies are noticed by high intensity regions in mammogram while normal ones can take high intensity regions too. IOBI system is proposed to distinguish between the two high intensity regions.

The obtained results were promising thus, the proposed method detects the effective boundary of pectoral muscle as well extracts regions with highest intensity as a result using this latter by IOBI system to identify whether it is normal or abnormal one. IOBI achieved 68.20% which is acceptable since the majority of abnormal mammograms which are classified as normal contains masses that are not identified till now, as the normal one which are classified as abnormal does not contain a clear intensity and that can be solved future work.

Perspective

From the obtained results, we have noticed that several ideas can be realized in order to increase the accuracy of early detection of breast cancer which consist in:

- Improving the quality of resulting image of pectoral muscle boundary as consequence using it to remove its region from mammogram with high accuracy.
- Improving the accuracy of IOBI by finding solution for ill-defined region and intensity which is not clear .
- Using the extracted region by IOBI system to extract anomalies and classifying it as benign or malignant.

Achieved scientific work

Publication of a paper “Pectoral muscle boundary detection using digital mammograms”, 3-4 December 2017, The fifth international conference for image and signal processing and their applications(ISPA’17), as the paper has been selected as the best one of ISPA’17.

Bibliography

[1] Sampat, Mehul P., Mia K. Markey, and Alan C. Bovik. "Computer-aided detection and diagnosis in mammography." Handbook of image and video processing 2.1 (2005): 1195-1217.

[2] Breast Cancer-A Study (Early Detection); https://www.slideshare.net/mathupuji/medical-imaging-techniques-for-breast-cancer-a-study; 26/03/2018.

[3] Types of Medical Imaging-Understanding Medical Radiation; https://www.medicalradiation.com/types-of-medical-imaging/imaging-using-x-rays/mammography/; 26/03/2018.

[4] Ng, K. H., and M. Muttarak. "Advances in mammography have improved early detection of breast cancer." Journal-Hong Kong College of Radiologists 6 (2003): 126-131.

[5] Kuhl, Christiane K., et al. "Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer." Journal of clinical oncology 23.33 (2005): 8469-8476.

[6] Lin, Chen, Clark David Rogers, and Shadie Majidi. "Fat suppression techniques in breast magnetic resonance imaging: a critical comparison and state of the art." (2015).

[7] Reznik, A., B. J. M. Lui, and J. A. Rowlands. "An amorphous selenium based positron emission mammography camera with avalanche gain." Technology in cancer research & treatment 4.1 (2005): 61-67.

[8] MacDonald, Lawrence, et al. "Clinical imaging characteristics of the positron emission mammography camera: PEM Flex Solo II." Journal of Nuclear Medicine 50.10 (2009): 1666-1675.

[9] Glass, Shannon B., and Zeeshan A. Shah. "Clinical utility of positron emission mammography." Baylor University Medical Center Proceedings. Vol. 26. No. 3. Taylor & Francis, 2013.

[10] Bílková, A., V. Janík, and B. Svoboda. "Computed tomography laser mammography." Casopis lekaru ceskych 149.2 (2010): 61-65.

[11] Poellinger, Alexander, et al. "Near-infrared laser computed tomography of the breast: first clinical experience." Academic radiology 15.12 (2008): 1545-1553.

[12] CTLMâ (CT Laser Mammography) - Imaging Diagnostic Systems:imds.com/wp-content/uploads/2016/02/CTLM_DataSheet.pdf; 30/03/2018.

[13] National Research Council. Saving women's lives: strategies for improving breast cancer detection and diagnosis. National Academies Press, 2005.

[14] Ahmed, Eng Wael Abdel-Rahman Mohamed. COMPUTER AIDED DIAGNOSIS OF DIGITAL MAMMOGRAMS. Diss. FACULTY OF ENGINEERING, CAIRO UNIVERSITY GIZA, 2009.

[15] Mina, Luqman Mahmood, and Nor Ashidi Mat Isa. "Preprocessing technique for mammographic images." International Journal of Computer Science and Information Technology Research 2 (2014): 226-231.

[16] Elsawy, Nabila, et al. "Band-limited histogram equalization for mammograms contrast enhancement." Biomedical Engineering Conference (CIBEC), 2012 Cairo International. IEEE, 2012.

[17] Harikiran, J., B. Saichandana, and B. Divakar. "Impulse noise removal in digital images." International Journal of Computer Applications (0975–8887) Volume (2010).

[18] Teoh, Sin Hoong, Boon Tatt Koik, and Haidi Ibrahim. "Exploration of current trend on median filtering methods utilized in digital grayscale image processing." International Journal of Materials, Mechanics and Manufacturing 1.1 (2013).

[19] Ramani, R., N. Suthanthira Vanitha, and S. Valarmathy. "The pre-processing techniques for breast cancer detection in mammography images." International Journal of Image, Graphics and Signal Processing 5.5 (2013): 47.

[20] Sasikumar, Madhavan. "Medical Image Denoising Using Three Dimensional Discrete Wavelet Transform and Bilateral Filter." (2015).

[21] Teoh, Sin Hoong, Boon Tatt Koik, and Haidi Ibrahim. "Exploration of Current Trend on Median Filtering Methods Utilized in Dig- ital Grayscale Image Processing." Interna- tional Jour- nal of Materials, Mechanics and Manufactur- ing1.1 (2013).

[22] Nagi, Jawad, et al. "Automated breast profile segmentation for ROI detection using digital mammograms." Biomedical Engineering and Sciences(IECBES), 2010 IEEE EMBS Conference on. IEEE, 2010.

[23] Leavline, E. Jebamalar, and D. Asir Antony Gnana Singh. "Salt and pepper noise detection and removal in gray scale images: an experimental analysis." International Journal of Signal Processing, Image Processing and Pattern Recognition 6.5 (2013): 343-352.

[24] Leavline, E. Jebamalar, and D. Asir Antony Gnana Singh. "Salt and pepper noise detection and removal in gray scale images: an experimental analysis." International Journal of Signal Processing, Image Processing and Pattern Recognition 6.5 (2013): 343-352.

[25] Alam, Nashid, and Mohammed J. Islam. "Pectoral muscle elimination on mammogram using K-means clustering approach." International Journal of Computer Vision & Signal Processing 4.1 (2014): 11-21.

[26] Maitra, Indra Kanta, Sanjay Nag, and Samir Kumar Bandyopadhyay. "Technique for preprocessing of digital mammogram." Computer methods and programs in biomedicine 107.2 (2012): 175-188.

[27] Sapate, Suhas, and Sanjay Talbar. "An Overview of Pectoral Muscle Extraction Algorithms Applied to Digital Mammograms." Medical Imaging in Clinical Applications. Springer, Cham, 2016. 19-54.

[28] Goyal, Megha. "Morphological image processing." IJCST 2.4 (2011).

[29] Di Stefano, Luigi, and Andrea Bulgarelli. "A simple and efficient connected components labeling algorithm." Image analysis and processing, 1999. proceedings. international conference on. IEEE, 1999.

[30] Czaplicka, Kamila, and Helene Włodarczyk. "Automatic breast-line and pectoral muscle segmentation." Schedae Informaticae 2011.Volume 20 (2012): 195-209.

[31] Sapate, Suhas, and Sanjay Talbar. "An Overview of Pectoral Muscle Extraction Algorithms Applied to Digital Mammograms." Medical Imaging in Clinical Applications. Springer, Cham, 2016. 19-54.

[32] Thangavel, K., and M. Karnan. "Computer aided diagnosis in digital mammograms: detection of microcalcifications by meta heuristic algorithms." GVIP Journal 5.7 (2005): 41-55.

[33] Raba, David, et al. "Breast segmentation with pectoral muscle suppression on digital mammograms." Pattern Recognition and Image Analysis(2005): 153-158.

[34] Chandrasekhar, R., and Y. Attikiouzel. "Segmentation of the pectoral muscle edge on mammograms by tunable parametric edge detection." Australian Research Centre for Medical Engineering (ARCME) (2000): 573-579.

[35] Vese, Luminita A., and Stanley J. Osher. "Modeling textures with total variation minimization and oscillating patterns in image processing." Journal of scientific computing 19.1-3 (2003): 553-572.

[36] Ferrari, Ricardo J., et al. "Automatic identification of the pectoral muscle in mammograms." IEEE transactions on medical imaging 23.2 (2004): 232-245.

[37] Camilus, K. Santle, V. K. Govindan, and P. S. Sathidevi. "Pectoral muscle identification in mammograms." Journal of applied clinical medical physics 12.3 (2011): 215-230.

[38] Vincent, Luc, and Pierre Soille. "Watersheds in digital spaces: an efficient algorithm based on immersion simulations." IEEE Transactions on Pattern Analysis & Machine Intelligence 6 (1991): 583-598.

[39] Liao, Ping-Sung, Tse-Sheng Chen, and Pau-Choo Chung. "A fast algorithm for multilevel thresholding." J. Inf. Sci. Eng. 17.5 (2001): 713-727.

[40] Liu, Chen-Chung et al. “A pectoral muscle segmentation algorithm for digital mammograms using Otsu thresholding and multiple regression analysis.” Computers & Mathematics with Applications 64 (2012): 1100-1107.

[41] C.C. Liu, W.Y. Chen, Screw pitch precision measurement using simple linear regression and image analysis, Applied Mathematics and Computation 178 (2006) 390–404.

[42] Adams, Rolf, and Leanne Bischof. "Seeded region growing." IEEE Transactions on pattern analysis and machine intelligence 16.6 (1994): 641-647.

[43] Duda, Richard O., and Peter E. Hart. "Use of the Hough transformation to detect lines and curves in pictures." Communications of the ACM 15.1 (1972): 11-15.

[44] Aurich, Volker, and Jörg Weule. "Non-linear gaussian filters performing edge preserving diffusion." Mustererkennung 1995. Springer, Berlin, Heidelberg, 1995. 538-545.

[45] Kwok, Sze Man et al. Automatic pectoral muscle segmentation on mediolateral oblique view mammograms. IEEE transactions on medical imaging 23 9 (2004): 1129-40.

[46] Fischler, Martin A., and Robert C. Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." Readings in computer vision. 1987. 726-740.

[47] Molinara, Mario, Claudio Marrocco, and Francesco Tortorella. "Automatic segmentation of the pectoral muscle in mediolateral oblique mammograms." Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on. IEEE, 2013.

[48] Senthilkumaran, N., and R. Rajesh. "Edge detection techniques for image segmentation–a survey of soft computing approaches." International journal of recent trends in engineering 1.2 (2009): 250-254.

[49] Ma, Wei-Ying, and B. S. Manjunath. "A comparison of wavelet transform features for texture image annotation." Image Processing, 1995. Proceedings., International Conference on. Vol. 2. IEEE, 1995.

[50] Kinoshita, Sérgio Koodi, et al. "Radon-domain detection of the nipple and the pectoral muscle in mammograms." Journal of Digital Imaging 21.1 (2008): 37-49.

[51] Kumar, Suresh, et al. "Performance comparison of median and wiener filter in image denoising." International Journal of Computer Applications (0975–8887) Volume 12 (2010).

[52] Shrivakshan, G. T., and C. Chandrasekar. "A comparison of various edge detection techniques used in image processing." IJCSI International Journal of Computer Science Issues 9.5 (2012): 272-276.

[53] Li, Yanfeng, et al. "Pectoral muscle segmentation in mammograms based on homogenous texture and intensity deviation." Pattern Recognition 46.3 (2013): 681-691.

[54] Sultana, Alina et al. “Detection of pectoral muscle in mammograms using a mean-shift segmentation approach.” 2010 8th International Conference on Communications (2010): 165-168.

[55] Comaniciu, Dorin, and Peter Meer. "Mean shift analysis and applications." Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on. Vol. 2. IEEE, 1999.

[56] Comaniciu, Dorin, and Peter Meer. "Mean shift: A robust approach toward feature space analysis." IEEE Transactions on pattern analysis and machine intelligence 24.5 (2002): 603-619.

[57] Haralick, Robert M., and Linda G. Shapiro. "Image segmentation techniques." Computer vision, graphics, and image processing 29.1 (1985): 100-132.

[58] Cheng, Yizong. "Mean shift, mode seeking, and clustering." IEEE transactions on pattern analysis and machine intelligence 17.8 (1995): 790-799.

[59] Mustra, Mario, and Mislav Grgic. "Robust automatic breast and pectoral muscle segmentation from scanned mammograms." Signal processing 93.10 (2013): 2817-2827.

[60] Wirth, Michael A., and Alexei Stapinski. "Segmentation of the breast region in mammograms using active contours." Visual Communications and Image Processing 2003. Vol. 5150. International Society for Optics and Photonics, 2003.

[61] Rosin, Paul L. "Unimodal thresholding." Pattern recognition 34.11 (2001): 2083-2096.

[62] Wang, Lei, et al. "Automatic pectoral muscle boundary detection in mammograms based on Markov chain and active contour model." Journal of Zhejiang University SCIENCE C 11.2 (2010): 111-118

[63] Akram, Farhan, et al. "A preprocessing algorithm for the CAD system of mammograms using the active contour method." Applied Medical Informatics 32.2 (2013): 1.

[64] Chan, Tony F., and Luminita A. Vese. "Active contours without edges." IEEE Transactions on image processing 10.2 (2001): 266-277.

[65] Oliver, Arnau, et al. "One-shot segmentation of breast, pectoral muscle, and background in digitised mammograms." Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014.

[66] Morris, O. J., M. de J. Lee, and A. G. Constantinides. "Graph theory for image analysis: an approach based on the shortest spanning tree." IEE Proceedings F (Communications, Radar and Signal Processing). Vol. 133. No. 2. IET Digital Library, 1986.

[67] Suk, Minsoo, and Ohyoung Song. "Curvilinear feature extraction using minimum spanning trees." Computer vision, graphics, and image processing 26.3 (1984): 400-411.

[68] Ma, Fei, et al. "Two graph theory based methods for identifying the pectoral muscle in mammograms." Pattern Recognition 40.9 (2007): 2592-2602.

[69] Camilus, K. Santle, V. K. Govindan, and P. S. Sathidevi. "Computer-aided identification of the pectoral muscle in digitized mammograms." Journal of digital imaging 23.5 (2010): 562-580.

[70] Maulik, Ujjwal, and Sanghamitra Bandyopadhyay. "Genetic algorithm-based clustering technique." Pattern recognition 33.9 (2000): 1455-1465.

[71] Ribeiro Filho, José L., Philip C. Treleaven, and Cesare Alippi. "Genetic-algorithm programming environments." Computer 27.6 (1994): 28-43.

[72] Karnan, M., and K. Thangavel. "Automatic detection of the breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications." Computer methods and programs in biomedicine 87.1 (2007): 12-20.

[73] Smola, Alex J., and Bernhard Schölkopf. "A tutorial on support vector regression." Statistics and computing 14.3 (2004): 199-222.

[74] Domingues, Inês, et al. "Pectoral muscle detection in mammograms based on the shortest path with endpoints learnt by SVMs." Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE, 2010.

[75] Aroquiaraj, I. Laurence, and K. Thangavel. "Pectoral muscles suppression in digital mammograms using hybridization of soft computing methods." arXiv preprint arXiv:1401.0870 (2014).

[76] Aborisade, D. O. "Fuzzy logic based digital image edge detection." Global Journal of Computer Science and Technology (2010).

[77] Gardezi, Syed Jamal Safdar, et al. "Segmentation of pectoral muscle using the adaptive gamma corrections." Multimedia Tools and Applications 77.3 (2018): 3919-3940.

[78] Huang, Shih-Chia, Fan-Chieh Cheng, and Yi-Sheng Chiu. "Efficient contrast enhancement using adaptive gamma correction with weighting distribution." IEEE Transactions on Image Processing 22.3 (2013): 1032-1041.

[79] Toz, Guliz, and Pakize Erdogmus. "A Single Sided Edge Marking Method for Detecting Pectoral Muscle in Digital Mammograms." Engineering, Technology & Applied Science Research 8.1 (2018): 2367-2373.

[80] Tang, Jinshan, et al. "Computer-aided detection and diagnosis of breast cancer with mammography: recent advances." IEEE Transactions on Information Technology in Biomedicine 13.2 (2009): 236-251.

[81] Eltoukhy, Mohamed Meselhy, Ibrahima Faye, and Brahim Belhaouari Samir. "A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation." Computers in biology and medicine 42.1 (2012): 123-128.

[82] Castellino, Ronald A. "Computer aided detection (CAD): an overview." Cancer Imaging 5.1 (2005): 17.

[83] Tang, Jinshan, et al. "Computer-aided detection and diagnosis of breast cancer with mammography: recent advances." IEEE Transactions on Information Technology in Biomedicine 13.2 (2009): 236-251.

[84] EL NAQA, I. S. S. A. M., and YONGYI YANG. "Techniques in the detection of microcalcification clusters in digital mammograms." Medical Imaging Systems Technology: Volume 4: Methods in Diagnosis Optimization. 2005. 45-66.

[85] Nishikawa, Robert M., et al. "Computer-aided detection of clustered microcalcifications on digital mammograms." Medical and Biological Engineering and Computing 33.2 (1995): 174-178.

[86] Gurcan, M. Nafi, et al. "Detection of microcalcifications in mammograms using higher order statistics." IEEE Signal Processing Letters 4.8 (1997): 213-216.

[87] Strickland, Robin N., and He Il Hahn. "Wavelet transform methods for object detection and recovery." IEEE Transactions on image processing 6.5 (1997): 724-735.

[88] Strickland, Robin N., and Hee Il Hahn. "Wavelet transforms for detecting microcalcifications in mammograms." IEEE Transactions on Medical Imaging 15.2 (1996): 218-229.

[89] Yu, Songyang, and Ling Guan. "A CAD system for the automatic detection of clustered microcalcifications in digitized mammogram films." IEEE transactions on medical imaging 19.2 (2000): 115-126.

[90] Kegelmeyer Jr, W. Philip, et al. "Computer-aided mammographic screening for spiculated lesions." Radiology 191.2 (1994): 331-337.

[91] Chan, Heang-Ping, et al. "Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space." Physics in Medicine & Biology 40.5 (1995): 857.

[92] American College of Radiology. BI-RADS Committee. Breast imaging reporting and data system. American College of Radiology, 1998.

[93] Guo, Q., J. Shao, and V. Ruiz. "Investigation of support vector machine for the detection of architectural distortion in mammographic images." Journal of Physics: Conference Series. Vol. 15. No. 1. IOP Publishing, 2005.

[94] Miller, Peter, and Susan M. Astley. "Detection of breast asymmetry using anatomical features." Biomedical Image Processing and Biomedical Visualization. Vol. 1905. International Society for Optics and Photonics, 1993.

[95] Bozek, Jelena, et al. "A survey of image processing algorithms in digital mammography." Recent advances in multimedia signal processing and communications. Springer, Berlin, Heidelberg, 2009. 631-657.

[96] Doi, Kunio. "Computer-aided diagnosis in medical imaging: historical review, current status and future potential." Computerized medical imaging and graphics 31.4-5 (2007): 198-211.

[97] Retico, A., et al. "An automatic system to discriminate malignant from benign massive lesions on mammograms." Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 569.2 (2006): 596-600.

[98] Elter, Matthias, and Alexander Horsch. "CADx of mammographic masses and clustered microcalcifications: a review." Medical physics 36.6Part1 (2009): 2052-2068.

[99] Kilday, Judy, Francesco Palmieri, and Martin D. Fox. "Classifying mammographic lesions using computerized image analysis." IEEE transactions on medical imaging 12.4 (1993): 664-669.

[100] Sahiner, Berkman, et al. "Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis." Medical Physics 25.4 (1998): 516-526.

[101] Jain, Anil K., Jianchang Mao, and K. Moidin Mohiuddin. "Artificial neural networks: A tutorial." Computer 29.3 (1996): 31-44.

[102] Chan, Heang-Ping, et al. "Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network." Physics in Medicine & Biology 42.3 (1997): 549.

[103] Kramer, Oliver. "K-nearest neighbors." Dimensionality Reduction with Unsupervised Nearest Neighbors. Springer, Berlin, Heidelberg, 2013. 13-23.

[104] Arbach, Lina, et al. "Mammographic masses classification: comparison between backpropagation neural network (BNN), K nearest neighbors (KNN), and human readers." Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on. Vol. 3. IEEE, 2003.

[105] Wei, Liyang, et al. "A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications." IEEE transactions on medical imaging 24.3 (2005): 371-380.

[106] Veldkamp, Wouter JH, et al. "Automated classification of clustered microcalcifications into malignant and benign types." Medical Physics 27.11 (2000): 2600-2608.

[107] Shen, Liang, Rangaraj M. Rangayyan, and JE Leo Desautels. "Application of shape analysis to mammographic calcifications." IEEE transactions on medical imaging 13.2 (1994): 263-274.

[108] Bocchi, L., and J. Nori. "Shape analysis of microcalcifications using Radon transform." Medical Engineering and Physics 29.6 (2007): 691-698.

[109] Jiang, Yulei, et al. "Malignant and benign clustered microcalcifications: automated feature analysis and classification." Radiology 198.3 (1996): 671-678.

[110] Leichter, Isaac, et al. "The use of an interactive software program for quantitative characterization of microcalcifications on digitized film-screen mammograms." Investigative radiology 34.6 (1999): 394.

[111] Laine, Andrew, and Jian Fan. "Texture classification by wavelet packet signatures." IEEE Transactions on pattern analysis and machine intelligence 15.11 (1993): 1186-1191.

[112] Soltanian-Zadeh, Hamid, and Farshid Rafiee-Rad. "Comparison of multiwavelet, wavelet, Haralick, and shape features for microcalcification classification in mammograms." Pattern recognition 37.10 (2004): 1973-1986.

[113] Pradeep, N., et al. "Feature extraction of mammograms." International journal of Bioinformatics research 4.1 (2012): 241.

[114] Shi, Jiazheng, et al. "Characterization of mammographic masses based on level set segmentation with new image features and patient information." Medical physics 35.1 (2008): 280-290.

[115] Brown, Lisa Gottesfeld. "A survey of image registration techniques." ACM computing surveys (CSUR) 24.4 (1992): 325-376.

[116] Guo, Yujun, et al. "Breast image registration techniques: a survey." Medical and Biological Engineering and Computing 44.1-2 (2006): 15-26.

[117] Sivaramakrishna, Radhika. "Breast image registration using a textural transformation." (1997).

[118] Brzakovic, D., et al. "Mammogram analysis by comparison with previous screenings." Proc. 2nd Int. Workshop on Digital Mammography. 1994.

[119] Suri, J. S., et al. "Combining adaptive segmentation approach for improving multi-resolution image registration on X-ray mammograms acquired using Fischer's fused full field digital mammography and ultrasound system (FFDMUS)." IASTED BioMED (2005).

[120] Behrenbruch, Christian P., et al. "The use of magnetic resonance imaging to model breast compression in X-ray mammography for MR/X-ray data fusion." (2000).

[121] Martí, Robert, et al. "Mammographic X-ray and MR correspondence." Digital Mammography. Springer, Berlin, Heidelberg, 2003. 527-529.

[122] Malur, Sabine, et al. "Comparison of written reports of mammography, sonography and magnetic resonance mammography for preoperative evaluation of breast lesions, with special emphasis on magnetic resonance mammography." Breast Cancer Research 3.1 (2000): 55.

[123] Suckling, John, et al. "The mammographic image analysis society digital mammogram database." Exerpta Medica. International Congress Series. Vol. 1069. 1994.

[124] The mini-MIAS database of mammograms: http://peipa.essex.ac.uk/info/mias.html; 16/05/2018.

[125] Jiang, Jay J., and David W. Conrath. "Semantic similarity based on corpus statistics and lexical taxonomy." arXiv preprint cmp-lg/9709008 (1997).

[126] Mustra, Mario, and Mislav Grgic. "Robust automatic breast and pectoral muscle segmentation from scanned mammograms." Signal processing 93.10 (2013): 2817-2827.

[127] Luo, Shu-Ting, and Bor-Wen Cheng. "Diagnosing breast masses in digital mammography using feature selection and ensemble methods." Journal of medical systems 36.2 (2012): 569-577.

[128] Platt, John. "Sequential minimal optimization: A fast algorithm for training support vector machines." (1998).

[129] Bandyopadhyay, Samir Kumar. "Diagnosis of Breast Abnormalities in Mammographic Image." IJCST 2.1 (2011).

[130] Diniz, João Otávio Bandeira, et al. "Detection of mass regions in mammograms by bilateral analysis adapted to breast density using similarity indexes and convolutional neural networks." Computer Methods and Programs in Biomedicine (2018).

[131] Lakshmanan, Rekha, et al. "Detection of architectural distortion in mammograms using geometrical properties of thinned edge structures." Intelligent Automation & Soft Computing 23.1 (2017): 183-197.

[132] Ericeira, Daniel Rodrigues, et al. "Detection of masses based on asymmetric regions of digital bilateral mammograms using spatial description with variogram and cross-variogram functions." Computers in biology and medicine 43.8 (2013): 987-999.

[133] Gaur, Shantanu, et al. "Architectural distortion of the breast." American Journal of Roentgenology 201.5 (2013): W662-W670.

[134] https://www.irfanview.com.

[135] Subashini, T. S., Vennila Ramalingam, and S. Palanivel. "Automated assessment of breast tissue density in digital mammograms." Computer Vision and Image Understanding 114.1 (2010): 33-43.

[136] Soulami, Khaoula Belhaj, et al. "Detection of breast abnormalities in digital mammograms using the electromagnetism-like algorithm." Multimedia Tools and Applications (2018): 1-29.