

IMAGE PROCESSING

**Development of Algorithms for detecting Architectural
Distortion and Enhancing Microcalcification features from
Pectoral Muscle delineated Mammograms**

Thesis submitted to the Cochin University of Science and
Technology for the award of the degree of Doctor of Philosophy
under the Faculty of Engineering

By
REKHA LAKSHMANAN

Under Supervising Guide
Dr.VINU THOMAS

**Department of Electronics
Govt. Model Engineering College, Thrikakkara
Kochi-682022, Kerala, India.**

February 2016

Department of Electronics
Govt. Model Engineering College, Thrikakkara
Kochi-682022, Kerala, India



Certified that this thesis entitled " Development of Algorithms for detecting Architectural Distortion and Enhancing Microcalcification features from Pectoral Muscle delineated Mammograms " is a bonafide record of the research work done by Rekha Lakshmanan under my supervision in the Department of Electronics, Govt. Model Engineering College, Thrikakkara, Kerala, India. The contents of this thesis have not been submitted and will not be submitted to any other University or Institute for the award of any degree/diploma.

Kochi-21
February 2016

Dr. Vinu Thomas
(Supervising Guide)
Department of Electronics
Govt. Model Engineering College
Kochi-21

Department of Electronics
Govt. Model Engineering College, Thrikakkara
Kochi-682022, Kerala, India



Certificate

Certified that all the relevant suggestions and modifications suggested by the audience during the pre-synopsis seminar and recommended by the Doctoral Committee of the candidate have been incorporated in this thesis entitled “Development of Algorithms for detecting Architectural Distortion and Enhancing Microcalcification features from Pectoral Muscle delineated Mammograms”

Kochi-21
February 2016

Dr. Vinu Thomas
(Supervising Guide)
Department of Electronics
Govt. Model Engineering College
Kochi-21

DECLARATION

I hereby declare that the work presented in this thesis entitled "Development of Algorithms for detecting Architectural Distortion and Enhancing Microcalcification features from Pectoral Muscle delineated Mammograms" is a bonafide record of the research work done by me under the supervision of Dr. Vinu Thomas, Associate Professor, Department of Electronics, Govt. Model Engineering College, Thrikakkara, Kerala, India.

Kochi-21
February 2016

Rekha Lakshmanan

To my husband.....

Acknowledgement

The completion of this research would not have been possible if not for the immense help rendered by my teachers, family, friends and well-wishers. I am grateful to God Almighty for his blessings and for giving me this opportunity to address an issue of concern for women in the society.

I am indebted to my supervising guide Dr. Vinu Thomas, for his able guidance and support at every stage of my research. He was always there with knowledge, encouragement and proper direction. The valuable advice that I received from him and the fruitful discussions I had with him have inspired me in my research.

I extend my sincere gratitude to the former and current Principals of Model Engineering College, Dr. P Sureshkumar and Dr. V. P. Devassia and former heads of the Department of Electronics Engineering at Model Engineering College, Dr. Mini M.G and Dr. Jayasree.V.K, for providing all facilities for this research.

I express my sincere gratitude to Prof. Shiji T. P, faculty member of Govt. Model Engineering College for her selfless support, constant motivation and guidance as an elder sister. I am thankful to Prof. Sumitha Mathew, faculty member of Govt. Model Engineering College for her personal attention, guidance and grammatical correction of the manuscripts. I am grateful to Dr. Binesh T, faculty member of Govt. Model Engineering College for his generous care, guidance and encouragement.

I am grateful to Dr. G Madhu, Principal, Dean, Faculty of Engineering, CUSAT and Dr. Vinod Kumar M N, Professor, School of Engineering, CUSAT for their excellent lectures and guidance for the timely completion of course work for this research.

I express my profoundest gratitude to the radiologists of Lakeshore Hospital Kochi, Dr. Suma M Jacob and Dr. Thara Pratab for spending precious time with me for providing database, fruitful discussions and

verification of the results. I am also thankful to Mr. Biju, technical staff of Lakeshore Hospital for helping out with the data collection.

I gratefully acknowledge Dr. Rangaraj M Rangayyan, Professor with the Department of Electrical and Computer Engineering, and Adjunct Professor of Surgery and Radiology, at University of Calgary, Alberta, Canada for his advice and support. I also acknowledge A. Mencattini, P. Casti, and M. Salmeri from the Department of Electronic Engineering, University of Rome Tor Vergata (Italy) for providing the pre-processed images for comparison.

I express my sincere thanks to the management of KMEA Engineering College, Aluva for granting me leave and financial support to complete my research.

I have no words to express my sincere gratitude to my husband, Vibin S for his sacrifice, compromise, patience, love and affection throughout. His constant support, advices and valuable discussions always motivated me in facing all difficult situations. I am thankful to my kids, Swathy V R and Sathvik Vibin for their understanding. I always feel sorry for not having spent adequate time with them during my research. I am obliged to my parents, in-laws, brothers and sisters who are my all-time supporters. I also would like to thank Mrs. Sumathy, who was ever ready to take care of us during my research period.

I would like to express my thanks to my co- research fellows Arya Devi P.S, Rashid M E, Jibi John, Aparna Devi P S, Jagadeesh Kumar, Joseph George K N, Swapna Viju, Vineetha George, Neethu M S, Anoop T R, Asha R S, Simi Z. Sreeba and Lisa Bento for their support and cooperation. I am also thankful to all the teaching, non-teaching and technical staff of Govt. Model Engineering College and KMEA Engineering College for their support.

This research was supported by financial assistance from the Kerala State Council for Science, Technology and Environment (KSCSTE) and Technical Education Quality Improvement Programme (TEQIP-II) of the

Ministry of Human Resources Development, Govt. of India. Their support is gratefully acknowledged.

Rekha Lakshmanan

Abstract

Breast cancer detection is an important social requisite as it is the leading cause of death due to cancer among women. The mortality rate of breast cancer is second among all cancers. The cause for breast cancer is not known to date and early detection & treatment are the only means to reduce breast cancer related deaths. Mammography is the main radiological tool that is employed for identifying breast cancer at the earliest stage. Computer aided techniques have great relevance in detection of abnormalities from mammographic images, as often the features associated with various abnormalities are difficult to detect and might be missed by even trained radiologists. In addition, when screening mammography is employed, a large number of mammographic images need to be checked for signs of abnormality, justifying the use of computer aided diagnosis.

Three problems are addressed in this thesis: delineation of the pectoral muscle region by properly identifying the pectoral muscle boundary, detection of architectural distortion and enhancement of microcalcification features in the mammographic images. Two novel methods were developed for identifying the pectoral muscle boundary from mediolateral oblique view mammograms that employed multiscale decomposition and local segmentation. The breast area is extracted after this step following the removal of the Pectoral muscle region. The breast abnormalities are searched for in this region. Architectural distortion is the most commonly missed abnormality in mammograms. A novel method for detecting architectural distortion is proposed in this thesis that employs geometrical features obtained from selected edge structures in the mammographic image. These features are used to train a feedforward neural network classifier initialized using metaheuristic algorithms for better classification. Microcalcification is another breast cancer symptom which is

said to be the most commonly occurring. However the visibility of the microcalcification structures is often poor, especially when they are located in dense parenchymal tissues. Therefore an algorithm is proposed to enhance such features, employing the singularities, viz. zero-crossings and modulus maxima of coefficients obtained after computing the contourlet transform of the mammographic image. Contourlet transform is employed for the directional information it provides.

All algorithms are evaluated against similar works from current literature and the results are promising. The standard databases employed in literature, the mammographic image analysis society database (MIAS) of the University of Sussex, UK, and the digital database for screening mammography of the University of South Florida, USA were employed to evaluate the algorithms. Ground truth information is provided in both databases for all images. Since in India, breast cancer is often found at a younger age, a study on a cross section of the Indian populace, with quite a good number of dense mammograms, was also undertaken, with the help of expert radiologists of Lakeshore hospital, Kochi, Kerala.

The results obtained are highly promising.

Table of Contents

Chapter 1: Introduction	1
1.1 Introduction	2
1.2 Anatomy of breast	3
1.3 Breast Imaging Modalities	5
1.3.1 Magnetic resonance imaging	5
1.3.2 Ultrasound imaging	6
1.3.3 Ductogram (Galactogram)	7
1.3.4 Positron emission tomography	8
1.3.5 Thermal imaging	9
1.3.6 Mammography	10
1.3.6.1 Major views of mammograms	14
1.3.6.1.1 Medio lateral oblique view	14
1.3.6.1.2 Cranio caudal view	15
1.3.6.2 Computer aided analysis of mammograms	16
1.3.6.2.1 Evaluation of CAD Techniques	18

1.4	Comparative study in terms of detection features for various imaging modalities	20
1.5	Organization of the thesis	21
1.6	Bibliography	23
Chapter 2: Scope and formulation of the problem		
2.1	Breast cancer	29
2.2	Major types of breast cancer	30
2.2.1	Ductal carcinoma	30
2.2.2	Lobular carcinoma	33
2.2.3	Rarely occurring breast cancers	33
2.3	Reasons for breast cancer	34
2.4	Major breast cancer symptoms	35
2.4.1	Microcalcification	35
2.4.2	Mass	36
2.4.3	Architectural distortion	37
2.4.4	Bilateral asymmetry	38
2.5	Bi-Rads categories	39
2.6	Scope and relevance	40

2.7	Database used	44
2.7.1	MIAS database	44
2.7.2	DDSM database	47
2.7.3	Lakeshore database	47
2.8	Bibliography	48
Chapter 3:	Pectoral muscle delineation	53
3.1	Introduction	54
3.2	Literature review	55
3.3	Proposed methods for PM boundary detection	58
3.3.1	Pre-processing operation	59
3.3.1.1	Removal of unwanted artifacts and label	59
3.3.1.2	Detection of orientation view of mammogram	62
3.3.1.3	Triangular ROI selection	63
3.3.1.4	Contrast enhancement using gray level grouping	64
3.3.2	Detection of major PM boundary segment	65
3.3.2.1	Method 1: using contours on homogeneous ROI	65

	Homogeneous region	
3.3.2.1.1	extraction using SUSAN filtering	65
3.3.2.1.2	Contour Extraction	69
3.3.2.1.3	Extraction of maximally stretched contour segments	69
3.3.2.1.4	Selection of intersecting segments	70
3.3.2.1.5	Selection of major PM boundary segment using orientation and eccentricity criteria	71
3.3.2.1.6	Extension of major PM boundary segment	73
3.3.2.2	Method 2: Using edge detector on coarse ROI	75
3.3.2.2.1	Gaussian pyramid decomposition	75
3.3.2.2.2	Canny edge detection at coarse level	75
3.3.2.2.3	Removal of unwanted pixels	77

	3.3.2.2.4	Selection of major PM boundary segment using orientation and eccentricity criteria	78
	3.3.2.2.5	Gaussian pyramid reconstruction	80
	3.3.2.2.6	Extension of major PM boundary segment	80
3.4		Experimental setup and performance measures	80
3.5		Results and discussion	83
3.6		Summary	93
3.7		Bibliography	95
Chapter 4:		Detection and classification of architectural distortion	99
4.1		Introduction	100
4.2		Literature review	102
4.3		Detection and classification of architectural distortion	106
	4.3.1	Detection of suspicious regions	106
	4.3.1.1	Spatial mask modification	107

4.3.1.2	Contour extraction	109
4.3.1.3	Detection of contour centroid	111
4.3.1.4	Detection of suspicious regions of architectural distortion	112
4.3.2	Edge feature extraction	112
4.3.3	Thinning	117
4.3.4	Geometrical analysis of edge structures	118
4.3.5	Classification	125
4.3.5.1	Back propagation neural network for classification	127
4.3.5.2	BPNN with metaheuristic optimization	129
4.3.5.2.1	Cuckoo search algorithm	130
4.3.5.2.2	Bat algorithm	132
4.3.5.2.3	Cuckoo search/ Bat BPNN classification	133
4.3.6	Feature selection	134
4.4	Experimental setup	137
4.5	Results and Discussion	138
4.6	Summary	152

4.7	Bibliography	155
Chapter 5:	Enhancement of Microcalcification	165
5.1	Introduction	166
5.2	Literature review	167
5.3	Proposed methods	170
5.3.1	Extraction of contourlet coefficients	172
5.3.2	Singularity detection of contourlet coefficients using zero crossings/modulus maxima	182
5.3.3	Suspicious microcalcification detection using parent child relationship in contourlet transform	183
5.4	Performance measures	185
5.5	Results and discussion	187
5.6	Summary	198
5.7	Bibliography	199
Chapter 6:	Concluding remarks and future work	207
6.1	Concluding remarks	208
6.2	Future work	210
6.3	Bibliography	212

Appendix 1:	Database specifications	213
Appendix 2:	Pseudocode for Cuckoo search/Bat algorithm and Cuckoo search/Bat initialized BPNN classification	217
List of publications		219
Index		223

List of Figures

1.1	Sagittal section of female breast	4
1.2	Breast MRI	6
1.3	Breast ultrasound	7
1.4	Ductography	8
1.5	PET image of breast	9
1.6	Thermal image of breast	9
1.7	Mammography unit	11
1.8	MLO view of a mammogram collected from Lakeshore hospital.	15
1.9	CC view of a mammogram collected from Lakeshore hospital	16
1.10	Receiver operating characteristics (ROC) curve	20
2.1	Ductal carcinoma.	31
2.2	Breast cancer development from normal, in situ, invasive, and metastatic carcinoma.	32
2.3	Lobular carcinoma.	33
2.4	Microcalcification	36

2.5	Mass in a mammogram, mdb028 of MIAS database, within blue boundary of ground truth information.	37
2.6	Architectural distortion in mdb115 with ground truth information in blue circle.	38
2.7	Bilateral asymmetry in mdb081 and mdb082 with ground truth information marked as blue circular boundary.	39
2.8	Graphical representation of incidence rate (IR) and mortality rate (MR) of breast cancer	41
3.1	Pectoral muscle, wedge, label, marker and noise in mammographic images from MIAS database	56
3.2	Flowchart of the methods for the detection of Pectoral muscle boundary.	60
3.3	Noise removal	61
3.4	Mammographic image after thresholding operation, detection of largest area component	62
3.5	Boundary defect in a mammogram, mdb085 from MIAS database.	63
3.6	Reduced ROI detection.	64
3.7	Single Univalued Segment Assimilating Nucleus (SUSAN) principle	66
3.8	Single Univalued Segment Assimilating Nucleus (SUSAN) spatial filter mask.	67
3.9	Intensity similarity functions.	67
3.10	Homogeneous region extraction in mammogram using SUSAN filtering technique.	68
3.11	Contours of extracted homogeneous region in reduced ROI	69

3.12	Maximally stretched contour segment extraction.	70
3.13	Maximally stretched contour segments satisfying intersection criteria	71
3.14	Orientation angle of image region	71
3.15	Maximally stretched contour segments of mdb004 satisfying orientation property of PM boundary	72
3.16	Major PM boundary segment satisfying eccentricity for mdb004.	73
3.17	Extension of major PM boundary component	74
3.18	Extended PM boundary of mdb004 obtained using method 1	74
3.19	Three level Gaussian Pyramid decomposition on mdb004	76
3.20	Canny edge detection on images obtained for a 3 level Gaussian Pyramid decomposition on mdb004	76
3.21	Edge structures of coarse representation of mdb004	77
3.22	Pixels involved in connected edge component (scanning from top to bottom)	78
3.23	Edge structures obtained after removing unwanted pixels on coarse image of reduced ROI.	78
3.24	Edge components of mdb004 satisfying orientation property of PM boundary	79
3.25	Edge component of mdb004 satisfying eccentricity property	79
3.26	Major PM boundary segment after gaussian pyramid reconstruction	80

3.27	PM boundary of mdb004 drawn by radiologist and by method2	81
3.28	False positive and false negative region.	82
3.29	Histogram showing RNE	85
3.30	PM boundary of mdb125 drawn by radiologist, method 1 and method 2.	86
3.31	PM boundary of mdb031 with small PM region using ground truth information, method 1 and method2.	87
3.32	PM boundary of mdb130 (dense breast) drawn by radiologist, method 1 and method 2.	89
3.33	PM boundary of mdb013 drawn by method 1 and method 2.	90
3.34	PM boundary of mdb028 drawn by method 1 and method 2.	91
3.35	PM boundary marked on dense mammographic image, mdb112 drawn by radiologist, method 1 and method 2.	92
3.36	PM boundary marked on a mammographic image obtained from Lakeshore hospital: original and boundary marked by proposed method 2.	93
3.37	PM boundary of mammogram obtained from Lakeshore hospital: original, PM boundary marked using method 2 and extracted breast region using detected PM boundary.	94
4.1	Flowchart of the proposed method for architectural distortion detection.	107
4.2	Images obtained after applying modified SUSAN filtering operation on mammograms from MIAS (mdb115), Lakeshore hospital (LSH-RMLO-264412) and DDSM databases (1078-lcc).	108

4.3	Breast region and contours of breast region with ground truth information.	108
4.4	Different types of spatial mask for reducing the number of suspicious regions in breast region	109
4.5	Spatial mask and resultant contours for mammograms	104-105
4.6	Contours and centroids (red spots) of suspicious homogeneous region using radiating filter mask.	112
4.7	Suspicious regions around each centroid of respective contours	113
4.8	Polar diagram of the components of a fourier series at a point in a signal and Fourier components plotted head to tail.	115
4.9	Edge features of mammograms from MIAS, Lakeshore hospital and DDSM database obtained using phase congruency.	118
4.10	Strong edge structures of breast region on images from MIAS, DDSM and Lakeshore hospital databases	119
4.11	Linear structures retained after thinning operation	120
4.12	Converging lines retained from the linear structures of a malignant and normal regions	121
4.13	Structures satisfying quadrant criterion 1 for malignant and normal regions	123
4.14	Structures satisfying quadrant criterion 2 for malignant and normal regions	124
4.15	Retained abnormal structures in malignant and normal regions of mammogram collected from MIAS database.	125

4.16	Retained abnormal structures in malignant and normal regions of mammogram collected from DDSM database.	126
4.17	Architecture of a feed forward back propagation neural network	127
4.18	Levy flight distribution	131
4.19	Training error for seven hidden neurons	137
4.20	Malignant region of dense mammogram from Lakeshore database, indicating ground truth information, suspicious ROIs, and retained structures that satisfy all criteria	140
4.21	Normal region of dense mammogram (LSH-LCC-262437) from Lakeshore database, indicating ground truth information, suspicious ROIs, and retained structures that satisfy all criteria	141
4.22	Mammogram (LSH-RCC-237263) with malignancy at sub areola region. Ground truth information, suspicious ROIs and structures satisfying the various criteria are shown.	142
4.23	Normal region of the mammogram (LSH-RCC-237263). Suspicious ROIs and structures satisfying various criteria are shown.	143
4.24	Mammogram (LSH-RMLO-237263) with malignancy at sub areola region. Ground truth information, suspicious ROIs and structures satisfying the various criteria are shown.	144

4.25	Normal region of the mammogram (LSH-RMLO-237263). Suspicious ROIs and structures satisfying various criteria are shown.	145
4.26	Mammogram (LSH-RMLO-182539) with focal retraction. Ground truth information, suspicious ROIs and structures satisfying the various criteria are shown.	146
4.27	Mammogram (LSH-RMLO-1209) with focally retracted malignancy. Ground truth information, suspicious ROIs and structures satisfying the various criteria are shown.	147
4.28	Mammogram (LSH-RCC-1209) with RCC view. Ground truth information, suspicious ROIs and structures satisfying the various criteria are shown.	148
4.29	ROC curves for Cuckoo search backpropagation classification of mammographic images collected from MIAS database, DDSM database and Lakeshore hospital.	149-150
4.30	ROC curves for Bat Backpropagation classification of mammographic images collected from MIAS database, DDSM database and Lakeshore hospital.	151-152
4.31	Malignant region of dense mammogram (LSH-LMLO-262437) Ground truth information, suspicious ROIs and structures satisfying the various criteria are shown.	153
5.1	Le Gal's classification of microcalcification	167
5.2	Representation of a curve using basis elements of wavelet and contourlet transforms	173
5.3	Block diagram of contourlet transform	174

5.4	Laplacian pyramid decomposition: analysis and synthesis	174
5.5	Wedge shape frequency partition for a 3-level DFB decomposition that leads to 8 (2^3) subbands	175
5.6	DFB decomposition blocks in two level DFB of contourlet transform	175
5.7	Directional subbands: Impulse responses at first two level of DFB and directional subbands of two level DFB.	176
5.8	Swapping of filtering and downsampling operation using multirate identity	177
5.9	Quincunx filter bank with resampling operation from third level of DFB in Contourlet Transform	177
5.10	Resampling operation on the mammographic region of mdb241 and resultant image with resampling matrices R_0 , R_1 , R_2 and R_3 .	178
5.11	Block diagram for a three level DFB of Contourlet Transform	179
5.12	Resampled responses of quincunx filter bank on third level of DFB.	179
5.13	Combined impulse response of bandpass arm of LP and the 8 th directional filter in the DFB	181
5.14	Impulse response of the subbands after 4 level directional decomposition of contourlet transform	181
5.15	Parent-child relationships among the contourlet transform coefficients	185
5.16	Enhancement of mdb 249: original, enhanced results using zero-crossings coefficients of wavelet transform and contourlet transform and modulus maxima coefficients of wavelet transform and contourlet transform.	188

5.17	Enhancement of mdb 241: original, enhanced results using zero-crossings coefficients of wavelet transform and contourlet transform and modulus maxima coefficients of wavelet transform and contourlet transform.	190
5.18	Comparison of performance measures to evaluate contrast, visual quality and sharpness of enhanced mammographic images using zero-crossings and modulus maxima on wavelet and contourlet coefficients.	191-192
5.19	Segmented MCCs obtained for mdb241 using the proposed methods compared against wavelet based methods	193
5.20	Segmented MCCs obtained for mdb249 using the proposed methods compared against wavelet based methods	194
5.21	Segmented MCCs. obtained for mdb245 using the proposed methods compared against wavelet based methods.	195
5.22	Enhanced MCCs obtained for mammograms from the DDSM database with the proposed methods.	197

List of Tables

2.1	Global comparison of incidence and mortality rate statistics with Indian scenario	40
3.1	Performance analysis: comparison of values reported for the proposed method with existing methods	84
4.1	Classification features of architectural distortion	136
4.2	Comparison of classification results for DDSM database.	153
4.3	Comparison of classification results for MIAS database.	154
5.1	Comparison of contrast improvement index(CII) in various methods	196

List of Abbreviations

ACS	American Cancer Society
AD	Architectural Distortion
ANE	Area Normalized Error
ANN	Artificial Neural Network
AUC	Area Under Curve
BI-RADS	Breast Imaging Reporting and Data System
BPNN	Back Propagation Neural Network
CADe	Computer Aided Detection
CADx	Computer Aided Diagnosis
CC	Cranio-Caudal
CII	Contrast Improvement Index
CS	Cuckoo Search
CT	Contourlet Transform
DCIS	Ductal Carcinoma InSitu
DDSM	Database for Screening Mammography
DFB	Directional Filter Bank
DITI	Digital Infrared Thermal Imaging
DM	Digital Mammography
DP	Diagonal Pixel
FN	False Negative
FNF	False Negative Fraction

FP	False Positives
FPF	False Positive Fraction
GLG	Gray Level Grouping
HD	Hausdorff distance
HP	Horizontal Pixel
ICMR	Indian Council of Medical Research
IDC	Invasive Ductal Carcinoma
ILC	Invasive Lobular Carcinoma
IR	Incidence Rate
LCIS	Lobular Carcinoma In-Situ
L-MLO	Left Medio-Lateral Oblique
LP	Laplacian Pyramid
MCC	Microcalcification
MIAS	Mammographic Image Analysis Society
MLO	Medio-Lateral Oblique
MR	Mortality Rate
PDFB	Pyramidal Directional Filter Bank
PET	Positron Emission Tomography
PM	Pectoral Muscle
PNL	Posterior Nipple Line
R-MLO	Right Medio-Lateral Oblique
RNE	Row Normalized Error
ROC	Receiver Operating Characteristics

ROI	Region of Interest
RP	Reference Pixel
SFM	Screen Film Mammography
SUSAN	Single Univalue Segment Assimilating Nucleus
TB _c	Target to Background Contrast Ratio
TEN	Tenengrad Criterion
TNF	Specificity
TN	True Negative
TP	True Positives
TPF	Sensitivity
US	Ultra Sound
VP	Vertical Pixel
WHO	World Health Organization
WT	Wavelet Transform

Chapter 1

Introduction

Computer aided techniques assists radiologists in providing a second opinion in mammographic analysis for early breast cancer detection. This chapter presents a brief introduction on breast cancer. The fundamental knowledge on breast anatomy is important in understanding the pathology of breast cancer. The two major views of mammogram which are significant in analyzing the areas of abnormalities are also discussed. Descriptions are given on various breast imaging modalities which are vital in early breast cancer detection. The relevance of screening and computer aided algorithms for the detection of abnormalities in its early stage is also highlighted. A brief sketch of organization of chapters included in the research work is also provided.

1.1. Introduction

Breast cancer is one of the most alarming cancers commonly diagnosed among women. Breast cancer stands first in developed countries and second in developing countries like India [Imran, 2011]. Even though there is no substantiating evidence about the exact reasons for the occurrence of breast cancer, some of the major risk factors include aging, late birth to first child, nulliparity, family history of breast cancer, lack of breast feeding and genetic mutation [Augustine, 2014].

The incidence rate of breast cancer is more in developed countries than in developing countries but the mortality rate is less in developed countries [Ferlay, 2014]. The reason for the reduction in mortality rate is mainly due to the awareness about screening techniques and treatment facilities. The analysis of breast image using various modalities such as mammogram along with computer aided methods assists radiologist in identifying the presence of abnormalities in its initial stage. The detection of abnormalities in the early stage avoids the spread of malignancy that leads to cancer death.

A radiologist observes the presence of breast abnormality by analyzing different views of both left and right breast images. Even though double reading of mammogram provides better detection of malignancy; it is more time consuming and expensive. Hence computer aided methods in association with detailed examination of breast images yields a more accurate and assured identification of cancerous areas.

In this chapter, the anatomy of breast, various breast imaging modalities such as x-ray imaging, magnetic resonance imaging, ultrasound imaging, thermal imaging etc., major views of mammogram, computer aided detection/ diagnosis and the performance metrics of computer aided detection methods are discussed. The organization of research work in the following chapters is also summarized.

1.2. Anatomy of breast

The study of anatomy of female breast is indispensable to understand breast cancer. Figure 1.1 shows the anatomy of a women's breast. Female breast, a mammary gland is a heterogeneous structure consisting of different types of tissues in which the major ones are glandular and stromal. Glandular tissue houses ducts, lobules etc. which involves in generation and transportation of milk. Alveolus, the milk secreting unit is the basic unit of glandular tissue. Stromal tissue consisting of fat tissue and fibrous connective tissue provides support and shape to the breast. Stroma, the non-parenchyma tissues consisting of fatty and connective tissues renders blood supply to parenchyma tissues of breast.

The two important system of breast for immunity and purification are lymph and vascular system. Lymph system composing of lymph nodes and lymphatic vessels involves in resisting diseases and removal of waste materials. Vascular system consists of a network of blood vessels to transport blood between breast tissues and the rest of the body. Women's breast, the mammary gland is located on top of muscle layers of chest.

Breast parenchyma includes 15-20 lobes of glandular tissue. Lobes are the milk generating glands which are isolated by fibrous tissue septa. Each lobe consists of numerous smaller lobules which in turn consist of tiny cavities called alveoli where milk is produced. The characteristic shape of breast is provided by these lobes. Lactiferous ducts, a minute milk carrier tubes carries milk from lobes to nipples. Nipple is a small projection on the breast surface with 15- 20 duct opening from secretory glands inside the breast tissue [Kenneth, 2013]. These ducts form a radial pattern from the nipple at the center of a dark area of skin called the areola. Lactiferous sinus is a dilated portion of duct under areola for storing milk [Dixon, 2011], [Valerie, 2011], [Moinfar, 2007]. Fat lobules are formed by a bunch of fat cells encompassed in fine layer of single membrane. The fat cells in a fat lobule use the same terminal of vascular supply [Jeffrey, 2000]. Spaces

Chapter 1 : Introduction

around lobules and ducts are filled with fat, ligaments and connective tissues.

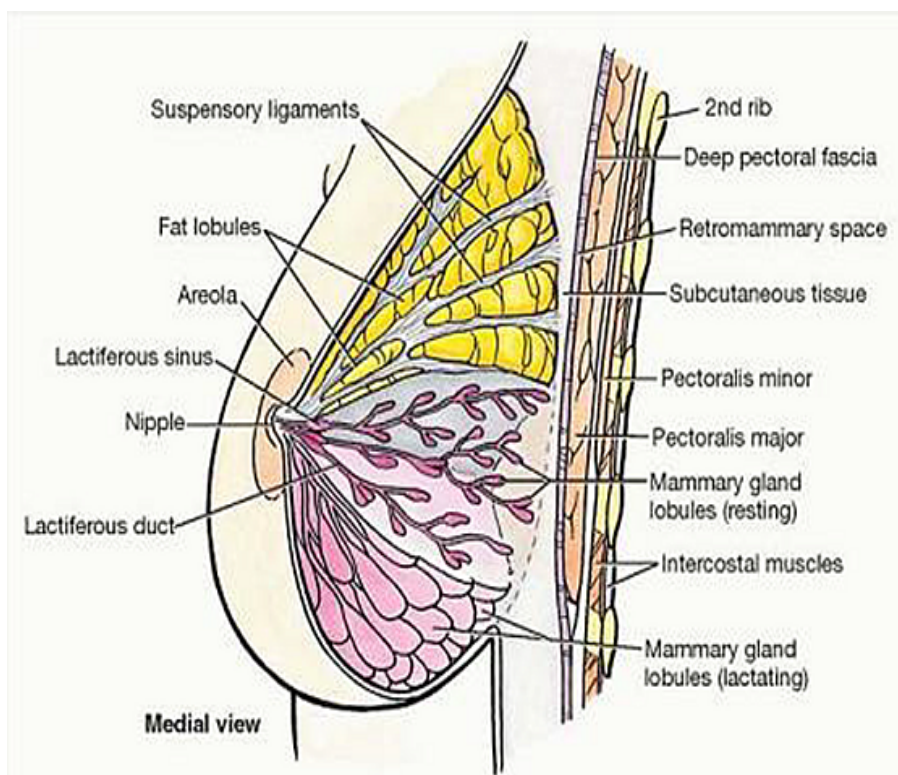


Figure 1.1: Sagittal section of female breast (from (MOORE, 2004)) [Moore, 2004]

Breast covers most of the chest area from second rib level to sixth rib level in front of the human rib cage [Kenneth, 2013]. The breasts are located directly over the pectoralis major (chest) muscle. Pectoralis major is the largest chest muscle which controls shoulder movements and fits hand to the body. Pectoralis minor, the smallest chest muscle that lies under pectoralis major occupies the third to fifth ribs [Moore, 2004]. Suspensory ligaments consisting of fibrous tissue septa extend from the deep fascia to the skin in order to provide support to the breast on the chest wall. These ligaments which are observed throughout the breast are also known as ‘Cooper’s ligaments’. Retromammary space consists of a layer of loose

connective tissue that differentiates the breast from the deep fascia and contributes some extent of movement over underlying structures [Valerie, 2011]. Subcutaneous tissues consists mainly of fat tissues, connective tissues, blood vessels and nerves which has a role as heat provider, skin binder, shape provider, energy reservoir and shock absorber [Jay, 2009]. Intercoastals are the muscles between the ribs. During lactation, the lobules in mammary glands enlarge and the glandular tissues become more prominent than the connective. The generated milk flows from alveoli to nipple through ducts by contraction of muscle like cells called myoepithelial cells.

1.3. Breast imaging modalities

The World Health Organization (WHO) [Ferlay, 2014] statistics emphasize the need of early breast cancer detection in order to reduce the mortality rate among women. The abnormalities of breast are identified by analyzing images of breast. Various imaging modalities that are used for screening breast cancer comprises of magnetic resonance imaging (MRI), ultrasound (US), mammogram, ductogram, positron emission tomography (PET), digital infrared thermal imaging (DITI). Currently, although other modalities like MRI, DITI and US are popular, mammogram is considered as the gold standard generally used for screening of breast with average risk. Even though there is advancement in imaging modalities, a combined approach is more efficient than single modality in detecting abnormalities due to breast cancer [Gheonea, 2011]. Analysis using various imaging technique reduce the possibility of missing the cancerous region.

1.3.1. Magnetic resonance imaging

MRI is a non-invasive imaging technique which uses magnets and radio waves to generate high quality images with good resolution [Stephan, 2010]. A very detailed image of breast is obtained by capturing the pattern

released by the energy of absorbed radio waves. The visibility of abnormal breast tissues are improved by injecting a contrast substance called galladium into the vein of arm [ACS, 2015]. The major advantages of MRI images are the non-exposure to radiation, possibility in analyzing simultaneously, dense breasts and imaging of breasts with inverted nipples. The major disadvantages of MRI include its expensiveness, consumption of time, increased false positives, noneligibility for women with internal metal object, difficulties in identifying microcalcification and insitu carcinomas. MRI images are usually taken as an additional tool for confirming the presence of breast tumours. Figure 1.2 illustrates breast MRI technique.

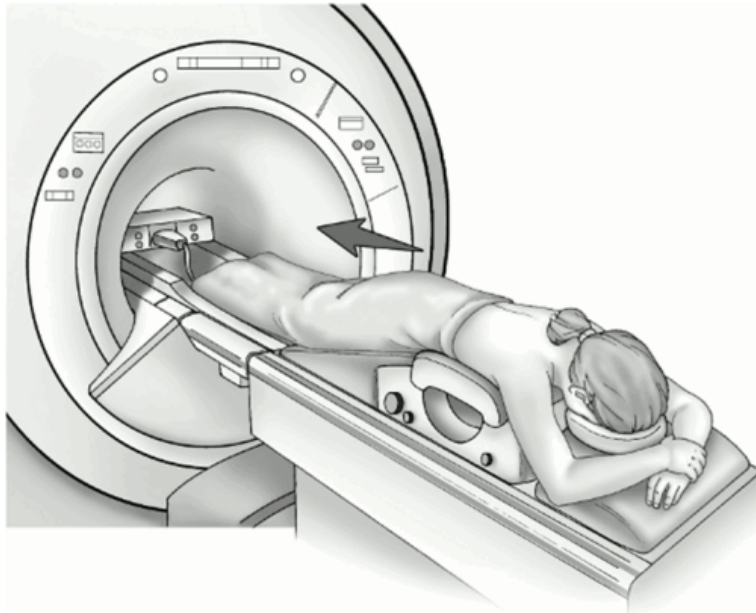


Figure 1.2: Breast MRI [ACS, 2013].

1.3.2. Ultrasound imaging

US is mainly used for detecting breast lesions. In US imaging technique, the high frequency sound waves generated by the transducer are directed to the breast tissues. The reflected sound waves obtained from the breast tissues are used to form two dimensional ultrasound images. The

ultrasound images are not considered as an effective technique in breast cancer screening due to the presence of false positive and false negative results [Teh, 1998]. US technique is considered as a supplementary tool to evaluate suspicious regions obtained using mammogram. The quality of US image is less compared to that of an MRI image. US imaging is a non-invasive technique which is not painful, is safer and is getting popular day by day as it is not using radiations. It is less expensive than MRI but more than that of mammogram. Even though US imaging technique is not good for identifying microcalcifications as in mammogram, it is found superior in imaging lesions in dense breasts and soft tissues. Figure 1.3 illustrates the breast US operation.



Figure 1.3: Breast Ultrasound [ACR, 2003]

1.3.3. Ductogram (Galactogram)

The reason for breast cancer with nipple discharge is identified using the technique known as ductogram or galactogram. A thin metal tube is placed through the opening of the duct where the nipple discharge occurs. In order to identify the presence of tumors in the passage of duct, a small amount of contrast substance is passed through the metal tube. The color

variation is useful in analyzing the size and shape of tumors [Valerie, 2011], [Sharmin, 2013].



Figure 1.4: Ductography [Sharmin, 2013]

1.3.4. Positron emission tomography

In PET, a special camera is used to take breast images by injecting radioactive substances into the blood stream. Cancerous tissues are active cells which absorb radioactive substances. The PET scanner forms an image by detecting the γ rays generated by radioactive materials. The advantage of PET is that it is not adversely affected by the presence of breast density, earlier surgery or radiotherapy [Gheonea, 2011]. The PET images are very expensive with less resolution. The exposure to radiation also generates side effects. Figure 1.5 shows a PET image of the breast. The parameters of the body including the blood flow, use of oxygen, metabolism of glucose etc. can be measured by the doctor through PET scan.



Figure 1.5: PET image of breast.
(Courtesy: <http://breast-cancer.ca/pets-bcanc/>)

1.3.5. Thermal imaging

In thermal imaging technique, a breast image is taken by a heat sensing camera. Figure 1.6 illustrates a thermal image of breast. The metabolic rate of malignant breast tissue is high compared to that of normal breast tissues.

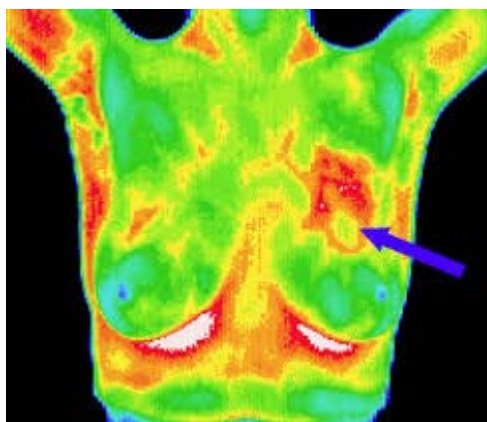


Figure 1.6: Thermal image of breast
(Courtesy:<http://southwestmedicalthermalimaging.com/breast-screening-thermography.html>)

As these rapidly multiplying cancerous cells are in need of new blood vessels for providing nutrients, the temperature surrounding these areas are high [Ng, 2009]. Studies on thermal breast images show that it is an effective screening tool for breast cancer. A study on thermal imaging technique and mammography shows that the latter is superior. Thermal imaging method is capable to detect only a quarter of the number of cases of carcinoma detected using mammogram [ACS, 2013].

1.3.6. Mammography

Mammography is currently considered as the golden standard among other imaging modalities in detecting breast cancer at its early stage. It is the cost effective and globally acceptable technique for early breast cancer detection. The image acquired through the exposure of x-rays on a breast is called mammogram. Modern mammography uses very low levels of radiation, usually about a 0.1 to 0.2 rad dose per x-ray. In order to get an x-ray image of breast, it is placed between two 2 plates to flatten and spread the breast tissue.

Mammograms appear as a black and white image of the breast tissue on a film or as a digital computer image that is read, or interpreted, by a radiologist [ACS, 2014]. As breast consists of various tissues such as fatty tissue, fibroglandular tissue, tumour tissues etc., rate of absorption of x-ray varies from tissue to tissue. The two dimensional image consisting of pixel intensities gives the characteristic features of various breast tissues through which the x-ray passes. As the rate of absorption of x-ray photons increases, the color is changed from black to white. Fatty tissues absorb and scatter x-rays less than fibroglandular tissues and hence appears black whereas calcium absorbs and scatters more x-rays [Giovanni, 2008]. Breast cancer is more easily detected in the case of fatty breast compared to dense breast [Alan, 2005]. Malignant calcifications and masses appear bright in mammograms.

The major components of a mammography unit comprises of an x-ray tube, compression paddle, grid, receptors etc., as shown in figure 1.7. The x-ray tube is used to produce characteristic x-ray energy. The desired x-ray energy for mammography is around 17- 24 kV [K Thayalan, 2014]. Anode material or the target material is usually made up of molybdenum (Mo) or rhodium (Rh) or tungsten (W) in rare cases [Lancaster].

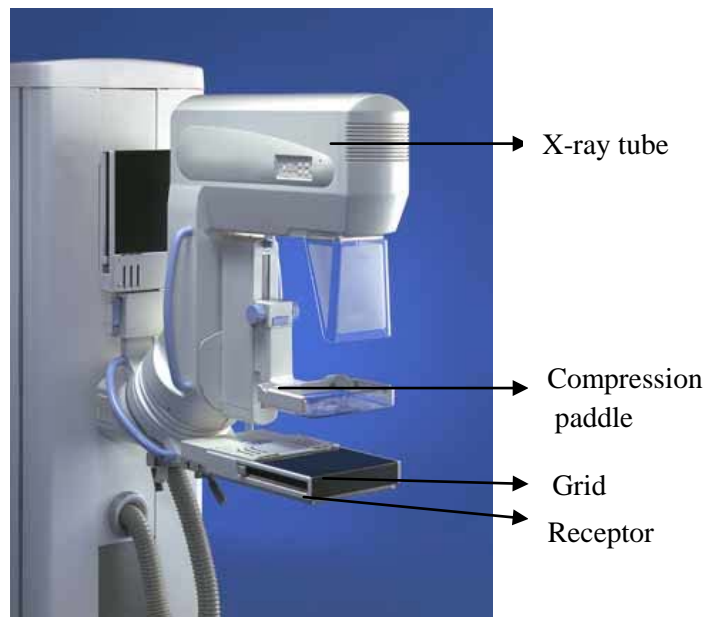


Figure 1.7: Mammography Unit, Siemens MAMMMOMAT 3000 Nova.

(https://www.cee.siemens.com/web/ua/ru/medecine/detection_dia_gnosis/mam/mammomat_nova/Documents/mammomat-3000-nova-mammography_applications-00009756.pdf)

Contrast among various tissues is useful in identifying the malignant structures present. Molybdenum and tungsten are the major x-ray tube anode material used in Siemens mammomat 3000 nova. Filters in x-ray tube kept on the x-ray path absorb the low and high x-ray energies in order to produce the desired energy. The usual range of x-ray tube voltage is 25-35 kV. A better quality of image is possible with operating voltage higher than the atomic number of the image structures. Since the most commonly visible microcalcification with atomic number, 20 is higher than that of the

three primary breast tissues (adipose, fibrous and glandular) ranging from 6 to 8 kV value around 25 adequate for penetration [Richard, 2012]. In Siemens mammomat 3000 Nova shown in figure 1.7, the current in x-ray tube with 25 kV, anode material as Mo is 150 mA and W is 188 mA. The advantage of high current in x-ray reduces the motion artifacts as well as the exposure time of radiation. The photon energy used in mammogram unit is comparatively less than that of a normal x-ray unit [Peter, 2015]. The photon energy must not have a low or high value. A high photon energy value results in reduced image contrast whereas a low value necessitates large patient dose due to inadequate penetration of x-ray. The x-rays are emitted from a small area of anode known as focal spot. The size of focal spot in the mammogram device is usually smaller than other radiographic devices as small focal spot size yields a sharper and a detailed high resolution mammographic image. The small size of focal spot within 0.1-0.3 mm, small distance between breast and image receptor, large distance between breast and focal spot are essential in reducing the geometrical blurriness [Ellen, 2007]. The compression paddle is used for compressing the breast in order to reduce the thickness of breast for uniform penetration of x-ray energy. Compression of breast improves the visibility of malignant lesion by spreading the overlying tissues of parenchyma. The radiation dose can be reduced with proper compression. Mammographic image blurring is reduced by holding the breast still using the compression paddle. Grid in the mammography unit which is thinner than other commonly used grid is helpful in improving the quality of image. It is efficient especially for the case of thick and dense breasts by reducing the spreading out of x-rays with more contrast [Robson, 2010]. Around 80% - 90% spreading of x-rays are reduced by placing grid in a mammography device. Even though the presence of grid increases the image quality, the exposure to x-ray radiation is two to three times more than that of nongrid mammographic unit [Richard, 2012]. Automatic exposure control is an essential part of mammography unit to provide consistent image receptor exposure for

various thickness and density compositions of breast tissues for the given kV. It is crucial as there is a difficulty for the radiologist in identifying the exact density composition of breast tissues [Richard, 2012]. The major advantages in using automatic exposure control include the reduction of repeated exposure and the exact time of exposure of x-rays.

x-ray breast images can be captured either on film known as screen film mammography (SFM) or directly to computer called digital mammography (DM). SFM was a very common image capturing technique around two decades ago. Currently DM is preferred over SFM due to its advantages like high resolution, wide dynamic range, zooming, magnifying and enhancing the image [Faridah, 2008] of especially dense breasts. The digital mammogram can be exchanged very easily among experts in reaching a conclusion compared to SFM. Over and above, a 45% processing time can also be saved in DM [Ranganathan, 2007]. Even though the cost of DM is comparatively higher than SFM, the technological advancement of DM outperforms the SFM in other aspects.

The database used in the proposed research work includes digital mammograms from standard databases such as mammographic image analysis society (MIAS) [Suckling, 1997], database for screening mammography (DDSM) [Paola, 2013] and images collected from Lakeshore hospital, India [Lakeshore]. The details of databases are provided in the second chapter.

The detection of abnormalities in the beginning stage avoids growing and spreading of cancerous cells throughout the body and thereby improving the survival rate. The mammograms are commonly used for screening as well as diagnostic purpose. Screening mammograms are the x-ray images of breast for women having no previous symptoms of breast cancer [ACS, 2014] whereas diagnostic mammograms are for women with previous cancerous symptoms. Screening mammograms are usually taken as part of routine medical check-up. Screening of breast using

mammograms is scientifically suitable for identifying anomalies in its beginning [Kopans, 2000]. It reduces the mortality rate due to breast cancer by 28% [Weedon, 2014]. The purpose of diagnostic mammogram is different from that of screening mammogram. Women with suspicious breast (having lumps, discharge etc.) are advised to take diagnostic mammogram. A thorough inspection of the suspicious area is performed using diagnostic mammogram. In order to evaluate the doubtful area, a magnification view or spot view is recommended. Other supplementary imaging modalities are also recommended for further analysis. Depending on the severity of breast cancer, the verification of diagnostic mammogram leads to a routine yearly or half-yearly checkup and biopsy.

1.3.6.1. Major views of mammogram

The two major standard views of mammogram involved in screening are the cranio caudal (CC) view and medio lateral oblique (MLO) view. Both these two views are essential in analyzing breast image for early breast cancer detection. Radiologists usually consider these two views in different angles for analyzing the mammographic images to reach a conclusion. In both views, x-rays are used for image acquisition. Compression of breast for x-ray imaging is proportional to the spreading of tissues which in turn maximizes the quality of resultant breast image. The reason for compressing breast is to reduce the exposure of x-rays on the breast [Harjit, 2011].

1.3.6.1.1. Medio lateral oblique view

Medio lateral oblique (MLO) view is considered as the primal view of mammogram because of the visibility of most of the breast tissues [Lawrence, 2004]. The whole breast image appears in the MLO view. An MLO view of mammogram is obtained by fixing an x-ray tube and a film holder in a direction parallel to pectoralis major muscle. X-rays are passed

to a compressed breast image from upper inner quadrant to lower outer quadrant [Michael, 2001] in order to acquire greater amount of breast tissues by minimizing the overlapping breast tissues. The visibility of pectoral muscle in MLO view of mammogram ensures the proper positioning of breast over the detector [William, 2007]. MLO view of right and left breast is distinguished by checking the label or position of pectoral muscle on the mammographic image. Right-MLO (RMLO) and left- MLO (LMLO) labels indicate the MLO view of the right and left breasts respectively. The pectoral muscle at the top right corner indicates a right breast image whereas the left breast has a pectoral muscle at top left corner. Figure 1.8 shows the MLO view of mammogram collected from the Lakeshore hospital, India [Lakeshore Hospital].

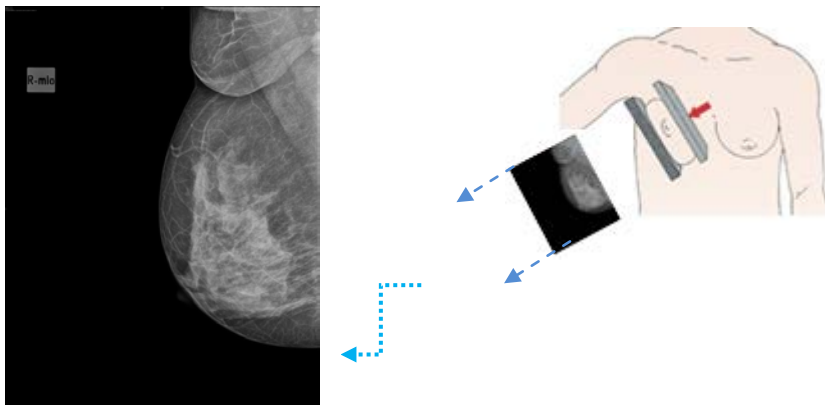


Figure 1.8: MLO view of a mammogram collected from Lakeshore hospital.

1.3.6.1.2. Cranio caudal view

Cranio caudal (CC) view images the breast by passing the x-rays from top to bottom. In CC view the x-ray device is positioned perpendicular to the floor and the breast is kept horizontally on the film holder [22]. The CC view complements the MLO view in the visual representation of medial, lateral, central and subareolar breast tissues. In CC view, nipple is clearly visible but pectoral muscle is minimally displayed. In CC view breast tissues in the medial and lateral portion of the breast should be

visible. Figure 1.9 shows the CC view of mammogram collected from the Lakeshore hospital.

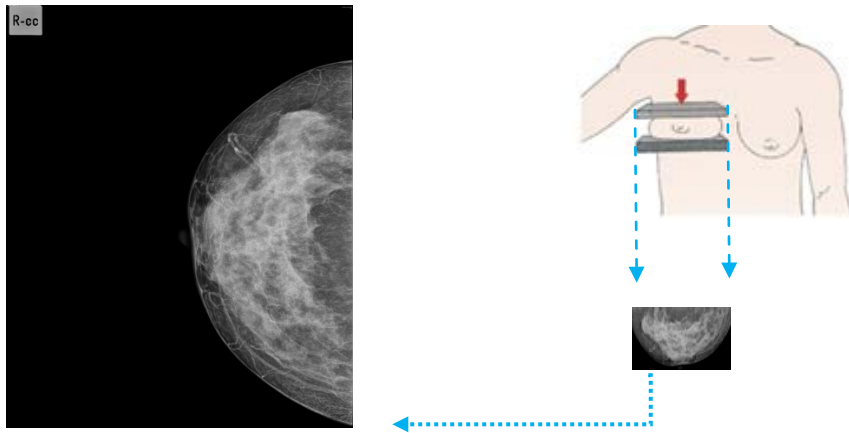


Figure 1.9: CC view of a mammogram collected from Lakeshore hospital

In MLO view, the posterior nipple line (PNL) is formed as the perpendicular line from the anterior pectoral muscle boundary to the nipple. For CC view, PNL is the line between the nipple and back edge of breast. The breast tissues in MLO and CC views are assured as maximally visible by checking the PNL distance between both views. The PNL distance of both views within 1 cm is considered as proper view with inclusion of all breast tissues.

1.3.6.2. Computer aided analysis of mammograms

Screening of mammogram by expertise radiologists classifies the mammogram as negative or positive. But the more the number of mammograms, various image acquisition artifacts may mislead to the interpretation of mammogram as being normal or abnormal. The examination of a single mammogram by different radiologists referred to as ‘Double reading of mammogram’ improves the interpretation thereby reducing misclassification [Brown,1996] and increasing high sensitivity and screening efficiency [Warren, 1995]. ‘Double reading of mammogram’ is

more expensive and time consuming along with heavy work load compared to computer aided techniques. Assessment by radiologists along with computer aided techniques is a good choice of early breast cancer detection. The literature [Morton, 2006], [Brem, 2003] shows that there is 7.62% increase in the number of detected breast cancers with 21.2% improvement in sensitivity.

A computer aided analysis of mammogram requires knowledge about the information on features of mammographic images and computer programming techniques. A database with proper number of mammographic images consisting of normal and abnormal cases is essential in computer aided analysis. Detection of various malignancies of the breast in their early stage improves the treatment thereby reducing the mortality rate among women [Freer, 2001].

The two major models of computer aided systems are computer aided detection (CADe) and computer aided diagnosis (CADx). Computer aided detection or diagnosis (CAD) systems have now emerged as a challenging and remarkable topic of interest among researchers as well as radiologists [Tang, 2009]. The need for detecting malignancies such as calcifications, masses, architectural distortions, bilateral asymmetry etc., and the incorrect classification of normal and abnormal mammograms are some of the major reasons for improving the computer aided systems among researchers whereas the results assists radiologists in providing a reliable analysis. CADe systems help in improving the accuracy of breast cancer detection by providing a second opinion whereas the CADx systems aids in making decisions between follow-up and biopsy [Sampat, 2005]. The interpretation of the abnormality in CADe system is left to radiologists [Warren, 1995]. In CADx, radiologists take the decision about the assessment, type and stage of the disease by considering the interpretations of radiologists or by the results of CADe system on mammograms [Vyborny, 2000], [Petrick, 2013].

1.3.6.2.1. Evaluation of CAD techniques

The four major categories of the detection results are true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN).

True positive:

Correct detection in which malignant case appears as malignant.

True negative:

Correct detection in which normal case appears as normal.

False positive:

Incorrect detection in which normal case appears as malignant.

False negative:

Incorrect detection in which malignant case appears as normal.

As FP and FN are the errors in detection, the results with minimum number of false positives and false negatives make the system more efficient. FP detection may generate anxiety in the patient. The FP findings may lead to unnecessary biopsies and treatment. The FN result is a more serious problem as the CAD system failed to identify the lesion on a patient. The FN detection may lead to an irrecoverable state of malignancy. The evaluation of a CAD system is reported through various performance metrics such as Sensitivity and Specificity which are calculated on the basis of number of FP, FN, TP and TN.

Sensitivity is referred as the capability of CAD system for identifying malignant breast. It is also known as true positive fraction. A high value of sensitivity is used for characterizing the detection of malignant cases. Sensitivity is inversely proportional to FN and is defined as

$$\begin{aligned}
 \text{Sensitivity/ True positive fraction, } TPF &= \frac{TP}{TP + FN} \\
 &= \frac{\text{Number of identified malignant cases}}{\text{Total number of malignant cases}} \quad (1.1)
 \end{aligned}$$

Specificity is referred as the capability of CAD system for identifying normal breast. Specificity is also known as true negative fraction. A high value of specificity represents the detection of normal cases. Specificity is inversely proportional to FP and is defined as

$$\begin{aligned}
 \text{Specificity/ True negative fraction, } TNF &= \frac{TN}{TN + FP} \\
 &= \frac{\text{Number of identified normal cases}}{\text{Total number of normal cases}} \quad (1.2)
 \end{aligned}$$

$$\begin{aligned}
 \text{False Positive Fraction, } FPF &= \frac{FP}{FP + TN} = 1 - \text{specificity} \\
 &= \frac{\text{FP case}}{\text{Total number of normal cases}} \quad (1.3)
 \end{aligned}$$

$$\begin{aligned}
 \text{False negative fraction, } FNF &= \frac{FN}{FN + TP} = 1 - \text{sensitivity} \\
 &= \frac{\text{FN case}}{\text{Total number of malignant cases}} \quad (1.4)
 \end{aligned}$$

$$\begin{aligned}
 \text{Accuracy, } A &= \frac{TP + TN}{TP + FP + FN + TN} \\
 &= \frac{\text{Number of correct assessments}}{\text{Total number of malignant cases}} \quad (1.5)
 \end{aligned}$$

The sensitivity and specificity values are used for plotting receiver operating characteristics (ROC) curve [Hanley, 1982]. ROC curve is a graph plotted as shown in figure 1.10 with sensitivity along x axis and FPF along y axis. The efficiency of CAD system is visible in ROC curve. An ideal ROC curve with high accuracy (100% sensitivity and 100% specificity) passes through the upper left corner [Zweig, 1993]. The area

under curve (AUC) represents the accuracy of the CAD system. A high value of AUC indicates a high classification performance.

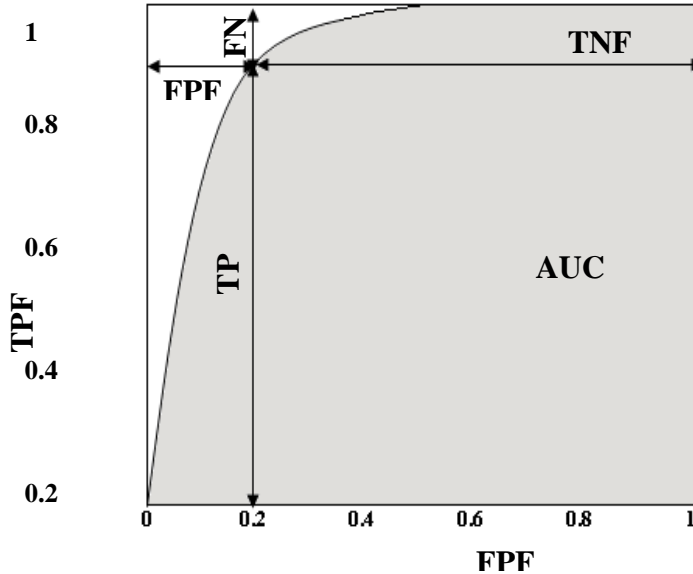


Figure 1.10: Receiver Operating Characteristics (ROC) curve

1.4. Comparative study in terms of detection features of breast imaging modalities

In addition to its ability to provide adequate visualization of soft tissue abnormalities, the particular strength of x-ray mammography is the ability to depict subtle calcifications. However in the case of dense breasts, mammography is seen to miss many cancers. In such cases, US imaging is used as an adjunct to mammography [Berg, 2016]. However ultrasonography is not good at detecting microcalcification. Thermography is a promising screening tool because of its ability to diagnose breast cancer at least ten years in advance. The disadvantage with thermography is that less depth information is obtained with thermography as attenuation of infrared rays in tissue is very high [Berg, 2016]. MRI performs better compared to other modalities in detecting abnormalities in dense images. In

addition it is painless and non-invasive. The disadvantages include more time consumption, inferior in detecting in situ cancer and inability to image calcifications. It is also expensive compared to other imaging modalities [Berg, 2016]. Many studies have been conducted on the performance of PET in the evaluation of suspicious breast lesions. Although PET can be a useful adjunct to mammography in characterizing breast tumors, this technique is limited by a low sensitivity in detecting small tumors and lobular carcinomas. Radiation exposure and the high cost of PET imaging has limited the use of this tool in the routine diagnosis of primary breast cancer [Sree, 2011].

1.5. Organization of the thesis

The addressing problem of the proposed work is to develop automated algorithms for

- extracting breast region by delineating pectoral muscle,
- detection and classification of Architectural Distortion and
- enhancement of microcalcification features

in mammographic images for assisting radiologists in early breast cancer detection.

The accuracy of computer aided systems for breast cancer detection can be improved by analyzing the breast area excluding pectoral muscle and other unwanted artifacts such as noise, labels, markers, wedges etc. Various computer aided detection schemes have been proposed in the literature for delineating the pectoral muscle from the mammogram. In this research work, two novel methods are proposed to identify pectoral muscle boundary using a local segmentation and multiscale decomposition technique. The research work implements a novel method for the detection and classification of architectural distortion. Two different metaheuristic algorithms along with feature selection are applied to improve the performance of classification technique. The proposed method of research

Chapter 1 : Introduction

make use of the directionality properties of contourlet transform for enhancing features of microcalcification in order to improve the visibility of anomaly in a mammogram. The proposed research work is structured in six chapters as mentioned below.

Chapter 1: The objective, scope and relevance of the thesis, brief summary of the thesis and its organization in various chapters are explained.

Chapter 2: An introduction, the current statistics of breast cancer, anatomy of the breast, various imaging modalities, views of mammogram, major breast cancer symptoms, Bi-rads categories, computer aided analysis and performance evaluation are discussed.

Chapter 3: Two novel methods for pectoral muscle boundary detection are discussed in this chapter. A literature review on pectoral muscle methods is done. A performance evaluation of these two methods along with those in the literature was compared.

Chapter 4: A novel method for the detection and classification of the most commonly missed abnormality, architectural distortion is proposed. An optimized neural network classification algorithm was used for optimizing the performance of the prescribed method.

Chapter 5: An enhancement operation using contourlet transform for the easy identification of the most commonly occurring breast cancer system, microcalcification is done. Performance metrics of enhancement operations are analyzed to check the efficiency of the proposed method.

Chapter 6: In chapter 6, the conclusions obtained on the proposed methods are discussed. The major outcomes along with the future work are described.

1.6. Bibliography

Alan, Stevens., & James, S. Lowe. (2005). Human Histology, (3rd edition), published by Elsevier.

American Cancer Society. Breast Cancer Facts & Figures 2013-2014. Atlanta: American Cancer Society, Inc. 2013.

American Cancer Society. Cancer Facts & Figures. (2015). Atlanta: American Cancer Society.

American College of Radiology, ACR BI-RADS—Mammography, Ultrasound & Magnetic Resonance Imaging, 4th ed. Reston, VA: Amer. Coll. Radiol., 2003.

Augustine, P., Jose, R., Peter, A., Lal, A, A., Prabhakar, J., & Sreedharan, J. (2014). Risk Factors of Breast Cancer in Kerala, India - A Case Control Study, Academic Medical Journal of India. 2(1), 7–13.

Carl, J. Vyborny., Maryellen, L. Giger., Robert, M. Nishikawa. (2000). Computer- Aided Detection and Diagnosis of Breast Cancer, Radiologic Clinics, 38(4), 725 – 740.

Dixon, M., & Thomas, J. (2006). Symptoms, assessment, and guidelines for referral. ABC of Breast Diseases, (3rd Edition), Blackwell Publishing Ltd, Massachusetts.

Ellen, Shaw, De, Paredes. (2007). Atlas of Mammography, Lippincott Williams & Wilkins.

Gheonea, Ioana, Andreea., Raluca, Pegza., Luana, Lascu., Simona, Bondari., Zoia, Stoica., & A. Bondari. (2011). The Role of Imaging Techniques in Diagnosis of Breast Cancer, Current Health Sciences Journal,

Chapter 1 : Introduction

37(2).

Giovanni, Di, Giacomo., Nicole, Pouliart., Alberto, Costantini., & Andrea, de, Vita. (2008). Atlas of Functional Shoulder Anatomy, Springer Science & Business Media.

Hanley, J, A., McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.

Harjit, Singh., & Janet, Neutze. (2011). *Radiology Fundamentals: Introduction to Imaging & Technology*, Springer Science & Business Media.

Imran, Ali., Waseem, A. Wani., & Kishwar, Saleem. (2011). Cancer Scenario in India with Future Perspectives, *Cancer Therapy*, 8, 56-70.

Jackie, Brown., Stirling, Bryan., & Ruth, Warren. (1996). Mammography screening: an incremental cost effectiveness analysis of double versus single reading of mammograms, *BMJ*, 312(7034), 809-12.

Jay, R. Harris., Marc, E. Lippman., C. Kent, Osborne, & Monica, Morrow. (2009). *Diseases of the Breast*, (4th edition), Lippincott Williams And Wilkins.

Jeffrey, A. Klein. (2000). *Tumescent Technique: Tumescent Anesthesia & Microcannular Liposuction*, 1e, Mosby, St. Louis.

J. Ferlay., I. Soerjomataram., R. Dikshit., S. Eser, C. Mathers., M. Rebelo., D. M. Parkin., D. Forman., & F. Bray. (2014). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN-2012. *International Journal of Cancer*, 136(5), E359-E386.

Jinshan, Tang., Rangayyan, R. M., El, Naqa, I. (2009). Computer-Aided Detection and Diagnosis of Breast Cancer with Mammography: Recent Advances, IEEE Transactions on Information Technology in Biomedicine, 13(2), 236 – 251.

Kenneth, L., & Bontrager, John, Lampignano. (2013). Textbook of Radiographic Positioning and Related Anatomy, Elsevier Health Sciences. (8th Edition).

Kopans, D. B. (2000). Double reading. Radiol Clin North Am. 38(4), 719-24.

K. Thayalan. (2014). The Physics of Radiology and Imaging, illustrated Edition, JP Medical Ltd.

Lakeshore Hospital and Research Centre, Cochin, Website: <http://www.lakeshorehospital.com>.

Lancaster, Jack, L. Ch. 4: Physical determinants of contrast, Physics of Medical X-ray Imaging (PDF). University of Texas Health Science Center.

Lawrence, W. Bassett., & R. Edward, Hendrick., Quality Determinants of Mammography: Clinical Practice Guideline, DIANE Publishing.

Marilyn, J. Morton., Dana, H. Whaley., Kathleen, R. Brandt., & Kimberly, K. Amrami. (2006). Breast Imaging: CAD Interpretation of Screening Mammograms, Radiology, 239(2), 375-383.

M. H. Zweig., & G. Campbell (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry. 39(4), 561-577.

Michael, S. Sabel. (2009). Essentials of Breast Surgery: A Volume in the

Chapter 1 : Introduction

Surgical Foundations Series Surgical Foundations, Elsevier Health Sciences.

Moinfar, F. (2007). Essentials of diagnostic breast pathology. Berlin Heidelberg: Springer-Verlag.

Moore, K. L., Agur, A. M., & Dalley, A. F. (2004). Essential Clinical Anatomy, (4th Edition), (W. Kluwer, Ed.).

Ng, EYK. (2009). A review of thermography as promising noninvasive detection modality for breast tumour. Int. J Therm. Sci., 48, 849-859.

Paola, Casti., Arianna, Mencattini., Marcello, Salmeri., Antonietta, Ancona., Fabio, Felice, Mangieri., Maria, Luisa, Pepe., & Rangaraj, Mandayam, Rangayyan. (2013). Automatic Detection of the Nipple in Screen-Film and Full-Field Digital Mammograms Using a Novel Hessian-Based Method. Journal of Digital Imaging, 26(5), 948-957.

Peter, Hogg., Judith, Kelly., & Claire, Mercer. (2015). Digital Mammography: A Holistic Approach, Springer.

Petrick, N., Sahiner, B., Armato, S. G., Bert, A., Correale, L., Delsanto, S. Chan, H. P. (2013). Evaluation of computer-aided detection and diagnosis systems. Medical Physics, 40(8), 087001:1-17.

Ranganathan, S., & Faridah, Y. (2007). Moving into the digital era: a novel experience with the first full-field digital mammography system in Malaysia, Ng KH Singapore Med J., 48(9), 804-807.

Richard, Carlton. Arlene, Adler. (2012). Principles of Radiographic Imaging: An Art and A Science, (5th edition), published by Cengage Learning.

R. M. L. Warren., & W. Duffy. (1995). Comparison of single reading with double reading of mammograms, and change in effectiveness with experience, *The British Journal of Radiology*, 68(813), 958-962.

R. Brem, J. Baum, M. Lechner, S. Kaplan, S. Souders, L. Naul, & J. Hoffmeister. (2003). Improvement in sensitivity of screening mammography with computer-aided detection: A multiinstitutional trial, *Amer. J. Roentgenol*, 181(3), 687–693.

Robson, K. J. (2010). Advances in mammographic imaging. *The British Journal of Radiology*, 83(988), 273-275.

Sampat, M. P., & Markey, M. K. (2005). In: Computer-aided detection and diagnosis in mammography *Handbook of Image and Video Processing*, (Bovik AC, Editor), published by Academic Press, 1195–217.

Sharmin, Akhtar, Rupa., Mahbooba, Ishrat., & Sonia, Shahid. (2011). Ductography: The Diagnostic Procedure of Choice in Patients with Abnormal Nipple Discharge, *J Enam Med. Col.*, 1(2), 97.

Sree, S. V., Ng, E. Yin- Kwee., Acharya, R. U., & Faust, O. (2011). Breast imaging: A survey. *World Journal of Clinical Oncology*, 2(4), 171–178.

Stephan, P. (2010). Lung and breast carcinomas, *Mammography and Breast MRIs*. Elseiveir, (Hayat MA, Ed.).

Suckling, J. (1994). The Mammographic Image Analysis Society Digital Mammogram Database *Exerpta Medica. International Congress Series*, 1069, 375-378.

The, W., Wilson, A. R. (1998). The role of ultrasound in breast cancer screening: A consensus statement by the European Group for Breast Cancer

Chapter 1 : Introduction

Screening, *Eur. J Cancer*, 34, 449 - 450.

Weedon Fekjær, Harald., Romundstad, Pål., & Vatten- Lars. (2014). Modern mammography screening and breast cancer mortality: population study, 348, g3701.

Timothy, W. Freer., & Michael, J. Ulissey. (2001). Screening Mammography with Computer-aided Detection: Prospective Study of 12,860 Patients in a Community Breast Center, *Radiology*, 220(3).781–786.

Wendie, A. Berg. (2016). Current Status of Supplemental Screening in Dense Breasts, *Journal Of Clinical Oncology*, 34(23).

William, E. Brant., & Clyde, A. Helms. (2007). *Fundamentals of Diagnostic Radiology*, Lippincott Williams & Wilkins.

Valerie, Andolina., & Shelly, Lillé. (2011). *Mammographic Imaging: A Practical Guide*, Lippincott Williams & Wilkins.

Y. Faridah. (2008). Digital versus screen film mammography: a clinical comparison, *Biomed Imaging and Intervention Journal*. 4(4).

Chapter 2

Scope and formulation of the problem

Computer aided techniques assists radiologists in providing a second opinion in mammographic analysis for early breast cancer detection. This chapter covers breast cancer and its types, breast cancer statistics and the various reasons for breast cancer. The major symptoms of breast cancer and the Bi-Rads categories are also discussed. The relevance of screening and computer aided algorithms for the detection of abnormalities in its early stage is also highlighted. The details about the mammographic images used for the proposed research is also explained.

2.1. Breast cancer

Cancerous cells that develop within the body harm the healthy cells and tissues leading to death [Pat, 1999]. Breast cancer, an unregulated cell growth in breast is 90% curable, once it is detected in its early stage [Jerrold, 2004]. These uncontrolled groups of cells can be either benign or malignant. Benign cases with unregulated growth or tumor is generally associated with pain [Daniel, 2013]. The uncontrolled growth is usually not visible in its initial stage. Early identification of breast cancer is possible through screening. The breast cancer symptoms in later stage includes breast lump, thickening, swelling, distortion, tenderness, skin irritation, redness, nipple retraction, nipple discharge etc. [ACS, 2014]. The lobes and ducts are usually considered as the origin of breast cancer. In some cases, the breast cancer starts from the intermediate tissues.

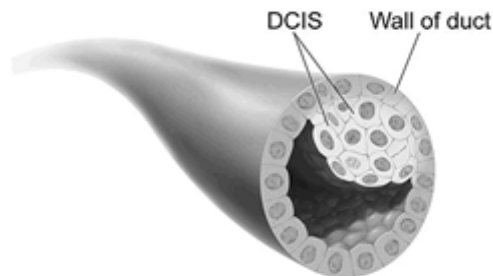
2.2. Major types of breast cancer

Breast cancers are mainly categorized as ductal and lobular as the milk generating lobules and transporting ducts are observed as the major regions of malignancy [ACS, 2015]. Some types are also observed in the in-between tissues containing glandular, fibrous and fatty tissues. The less frequently occurring breast cancers are inflammatory, paget, phyllodes and angiosarcoma. The complicated breast cancers among woman can be of more than one type [Judith, 2001].

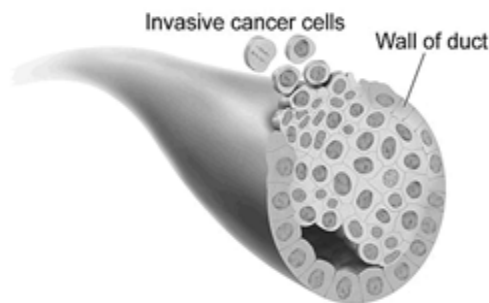
2.2.1. Ductal carcinoma

Ductal carcinoma originates in the epithelial cells that lines the milk ducts. Around 65% of breast cancer originates in duct [Katherine, 2011]. The two categories of ductal carcinoma are ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC). Ductal carcinoma in situ is a non-invasive breast cancer whereas invasive ductal carcinoma belongs to

invasive breast cancer. Figure 2.1 shows ductal carcinoma in situ and invasive ductal carcinoma, the two types of ductal carcinoma.



(a) Ductal carcinoma insitu



(b) Invasive ductal carcinoma

Figure 2.1: Ductal carcinoma. (a) Ductal carcinoma in situ, (b) invasive ductal carcinoma (courtesy: <http://www.oncolink.org/types>).

IDC spreads outside the ductal epithelium and surrounding myoepithelial cells to nearby normal breast tissues whereas DCIS is enclosed inside the ductal epithelium cell. Figure 2.2 represents the developmental stages of cancer growth in a duct. The normal duct consists of an inner epithelial cell, surrounding myoepithelial cell and an outer basement membrane. The supporting breast tissue consists of white blood cells (leukocytes), cell in connective tissue generating fibrous protein (fibroblasts), a fibroblast with characteristics of soft muscle

Chapter 2 : Scope and formulation of the problem

cells (myofibroblasts) and cells which line the interior of blood vessels and lymph vessels (endothelial). Even though it is not included in aggressive category, there is always a risk of around 20-25% chance of it developing into invasive breast cancer in a later stage [Seth, 2008], [Katherine, 2011]. Ductal carcinoma in situ constitutes around 2.5% of breast cancer [Katherine, 2011]. DCIS breaks the basement membrane reducing the number of myoepithelial cells with an increase in epithelial, fibroblasts, myofibroblasts and lymphocytes cells. The complete disappearance of myoepithelial cell as well as basement membrane leads to IDC. The spreading of cancerous cells to the neighbouring tissues, lymph vessels, blood cells and to other parts of the body results in the generation of cancerous growth in other organs, leading to metastases.

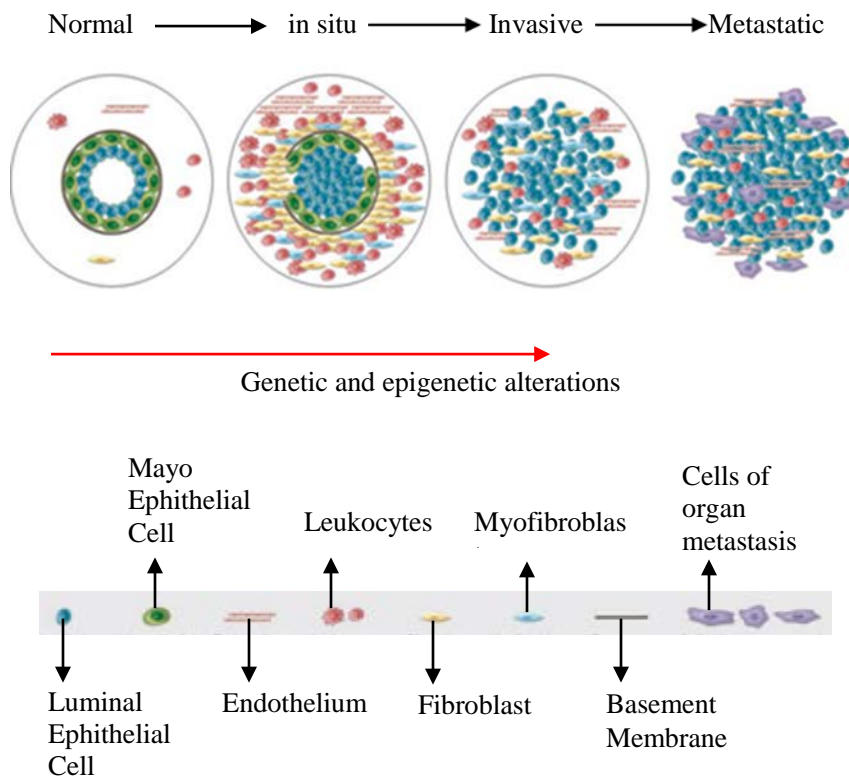


Figure 2.2: Breast cancer development from normal, in situ, invasive, and metastatic carcinoma ^[Kornelia, 2007].

2.2.2. Lobular carcinoma

Lobular carcinoma originates in the breast lobules. 10-15% of breast cancers originate in lobules [Katherine, 2011]. The two categories of lobular carcinoma are lobular carcinoma in situ (LCIS) and invasive lobular carcinoma (ILC). Lobular carcinoma in situ, also known as "lobular neoplasia" belongs to non-invasive category. The cancer like cells in lobular carcinoma in situ spreads only in lobules of milk generating glands. There is a probability for LCIS to develop into ILC or DCIS in its later stage [Bridget, 2011]. ILC is invasive category of lobular carcinoma with spreading of cancer cells to different parts of the body by breaking lobules. Figure 2.3 shows the normal lobule, LCIS and ILC. According to ACS, the symptom of ILC are thickening or hardening of breast instead of a lump formation [ACS, 2014]. The most commonly found invasive carcinoma is IDC and ILC stands second contributing to 5-15% of invasive breast cancer [Hacer 2012].

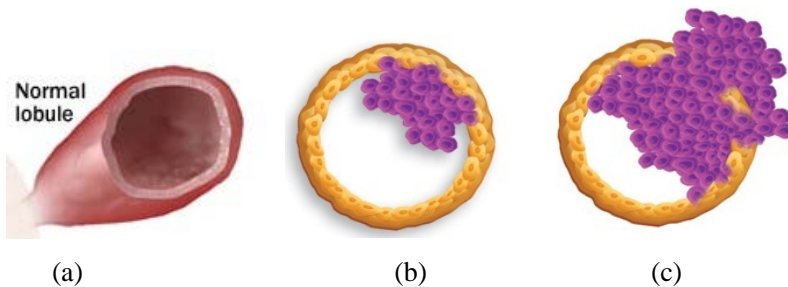


Figure 2.3: Lobular carcinoma. (a) Normal lobule, (b) lobular carcinoma in situ, (c) invasive lobular carcinoma. (Courtesy: <http://beasurvivoravera.com/breast-anatomy-and-function/>)

2.2.3. Rarely occurring breast cancers

Inflammatory, paget, phyllodes and angiosarcoma are considered as rarely occurring abnormalities since they constitute not more than 3% of breast cancers. Even though inflammatory breast cancer is occurring less frequently, the spreading of disease is very fast compared to other commonly occurring breast cancers. Some of the symptoms of

inflammatory breast cancer are hard breast, thickened and red colored breast skin. Paget is a breast cancer that originates in duct, extends to nipple skin and then to areola. In most cases it is associated with either DCIS or IDC. Phyllodes breast cancer originates in connective tissues of breast rather than ducts or lobules. Angiosarcoma emanates in cells that line blood or lymph vessels [ACS, 2014].

2.3. Reasons for breast cancer

Even though the exact reasons of breast cancer are not available, the major possible causes of breast cancer are categorized as modifiable or preventable and non-modifiable or non-preventable [Lakshmi, 2013]. The modifiable factors include overweight, physical inactivity, alcohol consumption etc. whereas the non-modifiable factors are age, family history, medical history etc.

Estrogen hormone, generated by ovaries is essential for the growth and development of breast, and reproductive system and organs until menopause. But the presence of high level of estrogen hormone after menopause, generated by fat tissues increases the risk of breast cancer [ACS, 2014]. Physical inactivity increases the fat deposits which in turn increases the estrogen level. Similarly alcohol consumption and smoking among women increases the risk of breast cancer [ACS, 2014]. Hormonal replacement therapy (HRT) raises the risk of breast cancer with adverse prognostic characteristics. Among the two major categories of HRT, combination HRT consisting of estrogen and progesterone is more dangerous than HRT with estrogen [Ronald, 2000].

The mutation of human genes, BRCA1 and BRCA2 which produces tumor suppressor proteins generates a high risk of breast cancer [Campeau, 2008]. Breast cancer risk increases with age due to the possibility in mutation of genes. Breast cancer among women in young age is more destructive [Nelson, 2006]. A long time exposure of estrogen hormone in the body due to menarche at earlier age and menopause at later

age results in an increased risk of breast cancer [Key, 2001]. Pregnancy before the age of 30 and breast feeding reduce the risk of breast cancer [Huiyan, 2007]. Women having some hereditary factors as well as previously diagnosed cancer have a high risk of breast cancer [Dixon, 2008]. Women with 75% dense breast tissues constituted by epithelium and stroma are more prone to breast cancer compared to the one with less dense tissue [Norman, 2007].

A study on risk factors of female breast cancer conducted in Regional cancer centre, Thiruvananthapuram, Kerala (RCC), shows that the factors found to be consistent with the well known standards are women in latter age, previous history of breast cancer, nulliparity, pregnancy after age 30, lack of breast feeding and nonuniform cycle of menstruation. Women in higher socio economic classes with better living conditions are found to be more susceptible to breast cancer [RCC, 2014]

2.4. Major breast cancer symptoms

Mammogram, the x-ray image of breast portraits the various tissue structures of breast. The mammographic images are analyzed by the radiologists to identify breast abnormalities. These anomalies on mammographic images are mainly characterized into four classes [Rangayyan, 2007]. The major symptoms of breast cancer are microcalcification, mass, architectural distortion and bilateral asymmetry.

2.4.1. Microcalcification

Microcalcification, one of the major anomalies of breast cancer is considered as the most commonly occurring breast cancer symptom. Microcalcifications are tiny deposits of calcium which appears as white speckles in mammogram. The size of microcalcification ranges from 0.1 to 1 mm with an average diameter of 0.3 mm [Lanyi, 1988]. The low contrast and tiny size of microcalcification obstructs its visibility from superimposed

breast tissues and noise. Benign calcification generally of larger size than malignant microcalcification are coarser, round, easily visible with smooth margin whereas malignant microcalcification are usually visible with magnifying glass [ACR, 1998]. Figure 2.4 represents a mammogram, mdb241 with microcalcification in MIAS database [Suckling, 1994]. Usually benign calcifications are calcium oxalate dehydrate and malignant calcifications are calcium hydroxy apatite [Haka, 2002].

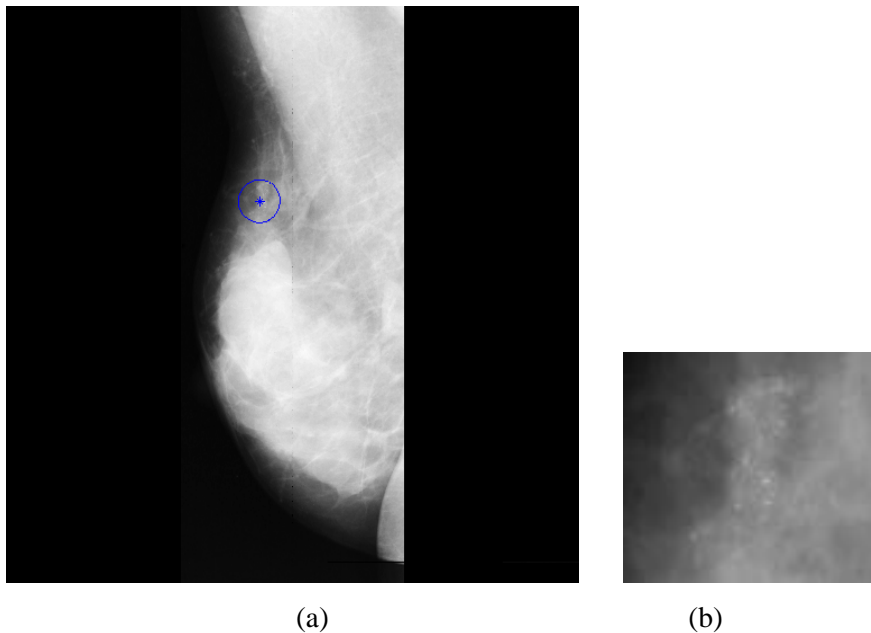


Figure 2.4: Microcalcification. (a) Malignant area with microcalcification in mdb241 is within the blue border representing ground truth information. (b) Enlarged view of malignant region.

2.4.2. Mass

Mass, a space occupying lesion and one of the major anomalies of breast cancer is visible in at least two different projections [ACR, 2003]. Masses differ in shape, margin and density. Masses are often hard and painless. A benign mass is generally characterized as low dense or isodense lesion with round or oval shape and smooth or circumscribed margins. A

lobule shaped mass can be benign or suspicious. A mass with more density consisting of more fatty tissues are found to be highly suspicious. A malignant mass is usually irregular in shape with speculated or rough margins [Mislav, 2009]. Figure 2.5 depicts a mammogram, mdb028 from MIAS database with malignant region within blue border of ground truth information.

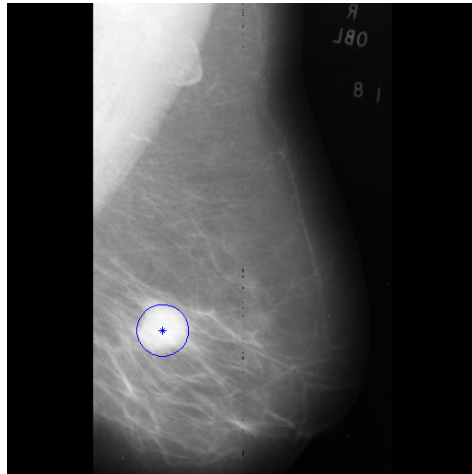


Figure 2.5: Mass in a mammogram, mdb028 of MIAS database, within blue boundary of ground truth information.

2.4.3. Architectural distortion

Architectural distortion is defined as alteration of normal architecture of a breast. Radiating patterns appear to emanate from a central region with no definite mass visible. A focal retraction at the edge of parenchyma is also considered as a symptom of Architectural Distortion is usually associated with other abnormalities such as mass and microcalcification [ACR, 2003]. Even though there are enormous works related to the most commonly missed abnormalities such as mass and microcalcification, moderately less studies have been reported on architectural distortion detection. The subtle nature and lesser contrast of architecturally distorted breast tissue structures obstructs the identification

of architectural distortion [Banik, 2013]. Figure 2.6 represents an architecturally distorted mammogram, mdb115 from MIAS database. The ground truth information provided by the radiologist is represented as the area with blue border.

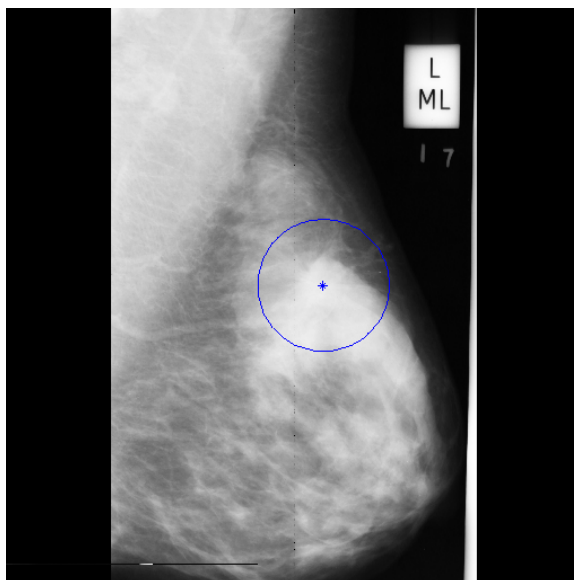


Figure 2.6: Architectural distortion in mdb115 with ground truth information in blue circle.

2.4.4. Bilateral asymmetry

Bilateral asymmetry is a breast cancer symptom obtained by verifying corresponding right and left mammographic image. Bilateral asymmetry is the lack of evenness among right and left mammogram. The difference in the right and left images are mainly due to its density or fibroglandular tissue structural dissimilarity. The asymmetry among right and left breast may be an indication of development of other breast cancers [Rangayyan, 2005]. Global and focal asymmetry are the two categories of bilateral asymmetry [ACR, 2003]. The difference in volume of fibroglandular tissues of corresponding areas of both breasts defines global asymmetry whereas the circumscribed area of asymmetry is associated with

lack of margins and suspicious mass. Figure 2.7 depicts mammogram with bilateral asymmetry.

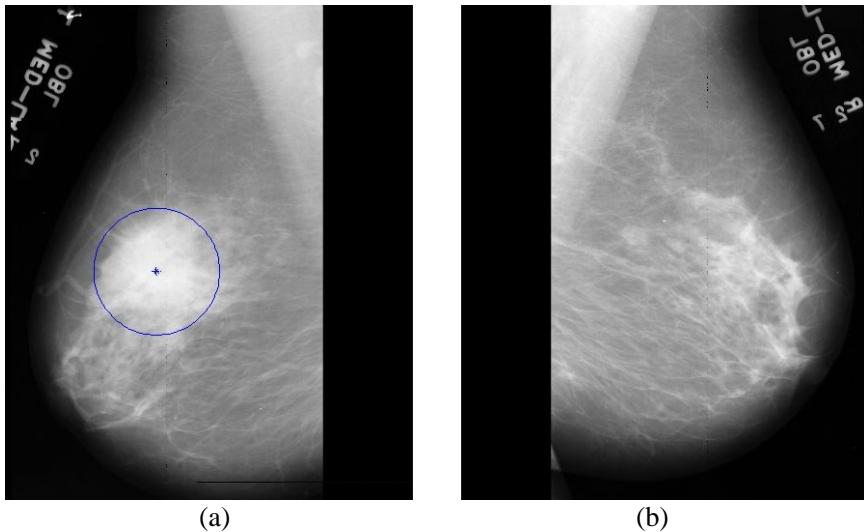


Figure 2.7: Bilateral asymmetry in mdb081 and mdb082 with ground truth information marked as blue circular boundary

2.5. Bi-Rads categories

Breast imaging reporting and data system (BI-RADS) provide a standardized method for categorizing the mammographic interpretation using numbers from 1-6 [ACS, 2014]. The various categories are listed below.

- Category 1:** An incomplete assessment of mammogram suggesting additional images of breast for a clear view of abnormality
- Category 2:** A complete assessment of mammogram with no malignancy.
- Category 3:** A complete assessment of mammogram with 98% of probability being benign. Repeat checkup during six months till the result become stable.
- Category 4:** A biopsy is preferred as the mammogram looks suspicious.
- Category 5:** A biopsy is compulsory as there is 95% chance of being cancer.
- Category 6:** Assessment of mammogram in response to the treatment of a proven case of malignancy.

2.6. Scope and relevance

According to the global statistics, the incidence rate of breast cancer stands in the second position [Ferlay, 2014] among all cancers. WHO reports states that breast cancer incidence rate in India is 1, 44,937 among new cases [Ferlay, 2014]. According to the statistics of american cancer society (ACS) [ACS, 2015], around 40,000 women died in 2015 due to breast cancer, 232,670 and 62,570 women were affected by invasive breast cancer and ductal carcinoma respectively. Breast cancer, the second cause of cancer fatality among women in India contributes 7% in the global scenario.

An analysis in India [Agarwal, 2008] reports annual breast cancer diagnosis of about 10,000 new breast cancer cases. Women above 40 are more susceptible to breast cancer in Asia. A study conducted by Govt. Medical College, Thiruvananthapuram (under Pain and Palliative Clinic) shows that the most commonly occurring cancer in women in India is breast cancer [Asha, 2014]. The population based cancer registry [2009 - 2011] reports that around 25-31% of cancers in Indian cities are due to breast cancer. Various reports and fact sheets in the last two decades indicate an increase in breast cancer incidence rate as well as death rate in India [Curado, 2011]. As per the reports of GLOOCAN (WHO) [Ferlay, 2014] depicted in table 2.1, women mortality rate in India due to breast cancer for the year 2012 is 70,218 ranking the highest position in world. The graphical plot of the statistics is shown in figure 2.8.

Table 2.1: Global comparison of incidence and mortality statistics with Indian scenario [Ferlay, 2014].

	World	India	World	India
year	Incidence rate	Incidence rate	Mortality rate	Mortality rate
2012	16,71,149	1,44,937	5,21,907	70,218
2015	17,90,861	1,55,863	5,60,407	75,957
2020	19,79,022	1,74,706	6,22,676	85,869

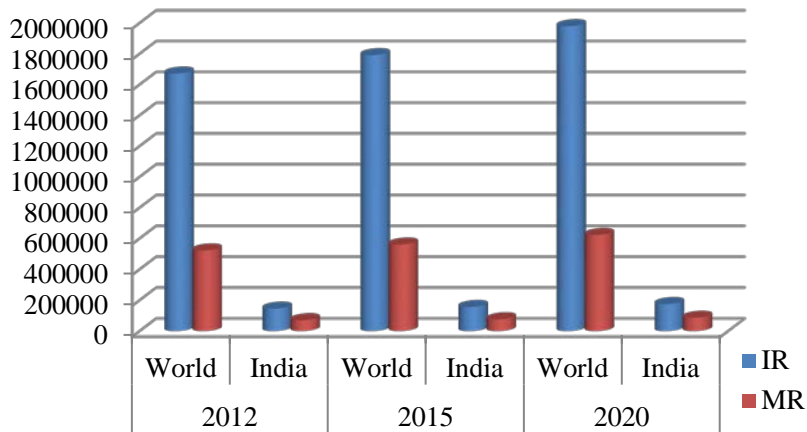


Figure 2.8: Graphical representation of incidence rate (IR) and mortality rate (MR) of breast cancer

In India, among the two newly identified breast cancer cases, one woman dies due to breast cancer [Ashutosh, 2015]. According to the Indian council of medical research (ICMR) study reports during the period from 1982 to 2005, breast cancer incidence rate in metropolitan cities has doubled. Statistics reports in Kerala shows that the increased breast cancer incidence rate contributes to one-third of all types of cancers among women [RCC, 2014]. A study conducted in 2012, warns that breast cancer among women in India leads to cervix cancer [Priya, 2012] by 2020.

According to Meshram [Meshram 2009], 90% of breast cancer stages are identified in its advanced stages. Early detection of breast cancer is very effective in reducing the mortality rate of breast cancer patients [Unger-Saldaña, 2014]. Routine checkup for breast cancer increases the chance of survival rate among women to about 10-25%. Mammography is a commonly used modality to detect breast cancer in its beginning stage as it is localized in a small area of breast without much spread. It helps to reduce the false negative results so as to avoid unnecessary biopsies and treatments. Reports based on various studies suggest that breast cancer screening dramatically reduces mortality rate resulting in 30% fewer death

Chapter 2 : Scope and formulation of the problem

annually [Kopans, 2004]. The possibility of observational oversights can be reduced by double reading of mammogram. There is a probability for missing 10 - 30% of cancers in mammogram depending on an increased density [Kolb, 2002]. As the interpretation and judgment of mammogram is a difficult task, a double reading is preferred. A double reading of mammogram always improves the detection rate by 3% - 15% [Kopans, 2000], [Harvey, 2003]. The difficulty in double reading includes high cost, more time consumption and less availability of an experienced radiologist. The prognosis of mammogram by an experienced radiologist assisted with adequate computer aided techniques reduces the false positive results thereby avoiding unnecessary biopsies [Ayres, 2005]. The computer aided techniques assists radiologists in identifying suspicious regions of mammogram for further analysis [Kolb, 2002]. These techniques are useful in highlighting abnormal regions which may be missed out by radiologists [Metz, 2001]. Computer aided techniques improves the breast cancer detection rates by 20-21% [Calas, 2012].

Automated computer aided systems for early breast cancer detection using mammograms; a low cost and low dose x-ray image is still an area of research due to various reasons.

1. Interpretation of low contrast images can be increased by differentiating the minor difference in intensity between the normal tissue and the malignant disease.
2. Presence of noises in mammographic images leading to false negatives can be avoided.
3. Performance efficiency of radiologists in analyzing mammograms can be improved.
4. Mortality rate can be reduced by detecting the symptoms of cancer lesions such as microcalcification, minute and low contrast anomaly as well as architectural distortion, most commonly missed abnormality in its early stage.

5. Efficient feature selection improves the accuracy of the system.
6. Removal of pectoral muscle from breast region reduces false positives.
7. Unnecessary biopsies can be avoided by reducing the false positives and false negatives.

Computer aided techniques using mammogram for early breast cancer detection is a demanding research area. The addressing problem of the proposed work is to develop automated algorithms for enhancing the most commonly occurring abnormality, microcalcification and the most commonly missed malignancy, and ie; architectural distortion from pectoral muscle delineated mammograms.

Radiologists usually compare right and left views of mammograms using landmarks such as nipple and pectoral muscle for the detection of breast cancer abnormalities. The omission of pectoral muscle from the mammographic image using computer aided techniques improves the accuracy of breast cancer screening. The pectoral muscle region may adversely affect anomaly detection due to its analogy with breast region. The presence of artifacts such as noise, wedges, opaque markers, labels etc. also adversely affects the precise detection of breast cancer. Extraction of the region of interest from the various perceptible and structurally distinct regions helps us to confine the search for anomalies only on the breast region.

Microcalcification, the most commonly occurring abnormality is a non-palpable symptom of breast cancer. Due to the smaller size, subtle nature, low contrast, presence of noise and dense tissues reduce the visibility of microcalcification in mammogram. The microcalcification enhancement techniques improve the visibility of microcalcification so that the radiologist can efficiently distinguish the presence of abnormality in mammogram.

Architectural distortion is a major breast cancer symptom which is considered as the most commonly missed abnormality [Rangayyan, 2007]. The diversity in appearance, subtle nature and the presence of dense tissues makes the detection of architectural distortion a challenging task for the radiologists. Architectural distortion, a common finding of false negative mammography reduces the detection of breast cancer in its early stage thereby increasing the mortality rate. Architectural distortion is usually found with other abnormalities such as microcalcification and mass. Compared to other malignancies such as mass and microcalcification, a less number of developed methods [Rangayyan, thesis, 2013] in the field of Architectural Distortion Detection makes it a promising research area.

2.7. Databases employed

The proposed works are experimented on mammograms from three databases (1) MIAS database, (2) DDSM database and (3) mammograms collected from Lakeshore hospital, India. The details of the databases employed are given as Appendix 1.

2.7.1. MIAS database

The mammographic image analysis Society (MIAS) [Suckling, 1994] database from Royal Marsden Hospital, U.K is a collection of 322 mammographic images. The database consists of normal and malignant mammograms in MLO view. The malignant images include symptoms such as microcalcification, circumscribed masses, spiculated lesions, ill-defined masses, architectural distortion and asymmetric density. The images are of size 1024 x 1024 with 200 micron pixel edge. The ground truth information is provided in the database with reference number of image, type of mammographic tissues (dense, glandular etc), category of mammogram (normal, microcalcification, mass, architectural distortion, bilateral

asymmetry), type of abnormality (benign or malignant) and circular region of abnormality with center coordinates and radius.

2.7.2. DDSM database

Database for screening mammography (DDSM) databases have been created by scanning images on x-ray films from Massachusetts General hospital, Wake Forest University School of Medicine, Sacred Heart Hospital and Washington University of St. Louis School of Medicine. The four different scanners, DBA M2100 Image-Clear, Howtek 960, the Lumisys 200 Laser, and the Howtek MultiRad850 provide a sampling rate of 42, 43.5, 50, 43.5 $\mu\text{m}/\text{pixel}$ respectively. The database includes 2,620 images containing both MLO and CC view of mammograms. The sizes of the mammogram images in the DDSM database are different.

2.7.3. Lakeshore database

The mammogram images in Lakeshore database are obtained from the hospital itself, India [Lakeshore]. The Lakeshore database is the only database with digital mammograms acquired using the Siemens MAMMOMAT 3000 NOVA digital mammography. Around 100 images were collected for experimental analysis.

2.8. Bibliography

Agarwal, G., & Ramakant, P. (2008). Breast Cancer Care in India: The Current Scenario and the Challenges for the Future. *Breast Care*, 3(1), 21-27.

American Cancer Society. *Cancer Facts & Figures*. (2015). Atlanta: American Cancer Society.

American College of Radiology (ACR). (1998). *Illustrated Breast Imaging Reporting and Data System (BI-RADS®)*, American College of Radiology, (3rd edition), (Reston,VA:Editor), 1998.

American College of Radiology. (2003). *BI-RADS—Mammography, Ultrasound & Magnetic Resonance Imaging*, (4th edition), (Reston, VA: Editor), Amer. Coll. Radiol.

Asha,K. P., D. Indu., M. S. Roopasree., Roshini, R. Lal., & V. G. Roshni.(2014). Clinical Symptoms and Psychosocial Problems of Cancer Patients Attending Pain and Palliative Care Clinic, *Journal of Academia and Industrial Research*, 2(10).

Ayres,F. J. & R. M. Rangayyan. (2005). Characterization of architectural distortion in mammograms. *IEEE Engineering in Medicine and Biology Magazine*, 24(1), 59–67.

Bridget, A. Oppong., & Tari A. King. (2011). Recommendations for Women With Lobular Carcinoma in situ (LCIS), *Oncology Journal*, *Breast Cancer*, 25(11), 1051-6, 1058.

Calas, Maria, Julia, Gregorio., Gutfilen, Bianca., Pereira, Wagner, Coelho, de, Albuquerque. (2012). CAD and mammography: why use this

Chapter 2 : Scope and Formulation of the problem

tool?. Radiol Bras, São Paulo, 45(1), 46-52.

Campeau, P. M., Foulkes, W. D., & Tischkowitz, M. D. (2008). Hereditary breast cancer: New genetic developments, new therapeutic avenues. *Human Genetics*, 124(1), 31–42.

Curado, M. P. (2011). Breast cancer in the world: incidence and mortality, *Salud Publica Mex.* 53(5), 372-84.

Daniel, D. Chiras, (2013). *Human Biology*, published by Jones & Bartlett Publishers.

Dixon, J. M. (2006). *ABC of breast diseases*. (3rd ed.) B. Publishing, Ed.

Ferlay,J., I. Soerjomataram., R. Dikshit., S. Eser, C. Mathers., M. Rebelo., D. M. Parkin., D. Forman., & F. Bray. (2014). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN-2012. *International Journal of Cancer*, 136(5), E359-E386.

Hacer, Haltas., Reyhan, Bayrak. Sibel, Yenidunya., Dilek, Kosehan., Meral, Sen., & Kayihan, Akin.(2012). Invasive lobular carcinoma with extracellular mucin as a distinct variant of lobular carcinoma: a case report, *Diagnostic Pathology*, 7(1), 91.

Haka, A., Shater-Peltier, K., Fitzmaurice, M., Crowe, J., Dasari, R., & Feld, M., Identifying Microcalcifications in Benign and Malignant Breast Lesions by Probing Differences in Their Chemical Composition Using Raman Spectroscopy. *Cancer Research*, 62, 5375-5380.

Chapter 2 : Scope and formulation of the problem

Harvey, S .C., Geller, B., Oppenheimer, R. G., Pinet, M., Riddell, L., & Garra, B. (2003). Increase in cancer detection and recall rates with independent double interpretation of screening. *Am. J. Roentgenol*, 180(5), 1461–1467.

Huiyan, Ma., Leslie, Bernstein., Malcolm, C. Pike., & Giske, Ursin. (2006). Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res.* 8(4).

Jerrold, S. Greenberg., George, B. Dintiman., & Barbee, Myers, Oakes. (2004). *Physical Fitness and Wellness: Changing the Way You Look, Feel, and Perform*, (3rd edition), published by Human Kinetics.

Judith, Peacock. (2001). *Breast Cancer Perspectives on disease and illness*, published by Capstone.

Katherine, A. Schneider. (2011). *Counseling About Cancer: Strategies for Genetic Counseling*, (3rd edition), published by John Wiley & Sons.

Key, T, J., & Verkasalo, P. K. (2001). Banks E. Epidemiology of Breast cancer. *Lancet Onco.*, 12(13), 133 – 40.

Kolb, T. M. , J. Lichy., J. H. Newhouse. (2002). Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations, *Radiology*, 225 (1) 165–175.

Kopans, D. B. (2004). Sonography should not be used for breast cancer screening until its efficacy has been proven scientifically, *Am. J.*

Roentgenol, 182(2), 489-491.

Kornelia, Polyak. (2007). Breast cancer: origins and evolution, Science in medicine, The Journal of Clinical Investigation, 117 (11), 3155–3163.

Lakeshore Hospital and Research Centre, Cochin, Website: <http://www.lakeshorehospital.com>.

Lakshmi, R., Vijayalakshmi, S., Athira, Raju., & Teena Mary Joy. (2013). Assessment of various risk factors of Breast cancer, International Journal of Pharmacy and Pharmaceutical Sciences, 5(4), 675-678.

Lanyi,M., (1988). Diagnosis and Differential Diagnosis of Breast Calcifications. (Berlin, Heidelberg : Editor), published by Springer-Verlag,.

Meshram, I. I., Hiwarkar, P. A., & Kulkarni, P. N. (2009). Reproductive Risk Factors for Breast Cancer: A Case Control Study, Online Journal of Health and Allied Sciences, 8(3).

Metz, C. E., (2001). ROC methodology in radiologic imaging, Invest Radiol, 21 (9), 720–733.

Mislav, Grgic., Kresimir, Delac., & Mohammed, Ghanbari. (2009). Recent Advances in Multimedia Signal Processing and Communications, Volume 231 of Studies in Computational Intelligence, Springer Science & Business Media.

Nelson N. (2006). Migrant Studies Aid Search for Factors Linked to Breast Cancer Risk, JAMA, 8(7), 436-438.

Chapter 2 : Scope and formulation of the problem

Norman, F. Boyd., Helen, Guo., Lisa, J. Martin., Limei, Sun., Jennifer, Stone., Eve, Fishell., Roberta, A. Jong., Greg, Hislop., Anna, Chiarelli., Salomon, Minkin., & Martin, J. Yaffe. (2007). Mammographic Density and the Risk and Detection of Breast Cancer, *The new england journal of medicine*, 356(3), 227-36.

Paola, Casti., Arianna, Mencattini., Marcello, Salmeri., Antonietta, Ancona., Fabio, Felice, Mangieri., Maria, Luisa, Pepe., & Rangaraj, Mandayam, Rangayyan. (2013) Automatic Detection of the Nipple in Screen-Film and Full-Field Digital Mammograms Using a Novel Hessian-Based Method. *Journal of Digital Imaging*, 26(5), 948–957.

Pat, Kelly., & Mark, Levine. (1999). *What You Need to Know about Breast Cancer: A Book for Women with Breast Cancer and Those who Care about Them* Empowering Press Series Pmph USA Ltd Series, PMPH-USA.

Priya, Shetty. (2012). India faces growing breast cancer epidemic, *The Lancet* , 379(9820) , 992 – 993.

Rangaraj, M. Rangayyan., Fábio, J. Ayres., J. E. Leo, Desautels, (2007). A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs, *Journal of the Franklin Institute*, 344(312–348).

Rangaraj, M. Rangayyan., Shantanu, Banik., Jayasree, Chakraborty., Sudipta, Mukhopadhyay., & J. E. Leo Desautels. (2013). Measures of divergence of oriented patterns for the detection of architectural distortion in prior mammograms, *International Journal of Computer Assisted Radiology and Surgery*, 8(4), 527-545.

Chapter 2 : Scope and Formulation of the problem

Rangayyan,R. M., (2005). Biomedical Image Analysis. Published by CRC Press, Boca Raton, FL.

RCC Hospital Cancer Registry. (2014). Regional Cancer Centre, Thiruvananthapuram [Internet]. Trivandrum: Regional Cancer Center, Trivandrum, Available from: www.rcctvm.org.

Ronald, K. Ross., Annlia, Paganini-Hill., Peggy, C. Wan & Malcolm, C. Pike. Effect of Hormone Replacement Therapy on Breast Cancer Risk: Estrogen Versus Estrogen Plus Progestin. (2000). Oxford Journals Medicine & Health, J. Natl. Cancer Inst., 92(4), 328-332.

Seth, Daniel, Shumate. (2008). Computational Model of Ductal Carcinoma in situ of the Breast, University of Arkansas.

Shantanu, Banik., Rangaraj, M. Rangayyan. (2013). Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer, 47th volume of Synthesis lectures on biomedical engineering, published by Morgan & Claypool.

Suckling, J. (1994). The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. International Congress Series, 1069, 375-378.

Suckling, J. (1994). The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. International Congress Series, 1069, 375-378.

Unger-Saldaña, Karla. (2014). Challenges to the Early Diagnosis and Treatment of Breast Cancer in Developing Countries, World Journal of

Chapter 2 : Scope and formulation of the problem

Clinical Oncology 5(3), 465–477.

Chapter 3

Pectoral Muscle delineation

Pectoral muscle (PM) occupies a significant portion of the MLO view of mammogram. The PM region may adversely affect anomaly detection due to its close similarity to the breast region. The unwanted artifacts during image acquisition such as noise, wedges, opaque markers etc. along with labels are also removed from mammographic images. This chapter discusses two novel methods for PM boundary detection. The performances of the methods were analyzed by comparing them with ground truth information provided by radiologists and the results were compared against those given by along with four algorithms in current literature. The proposed methods utilize the geometrical properties of edge structures for extracting PM boundary. In order to obtain the suspicious segments of PM boundary, the first method uses contours of homogeneous regions in breast area and the second method utilizes canny edge detector on a coarse breast region. An intensity similarity approach extends the detected major PM boundary segment to the two boundaries of mammogram. Experimental analyses were carried out on mammograms obtained from the mammographic image analysis database (MIAS) and on mammographic images obtained from Lakeshore hospital, Kochi. The proposed methods yield low values for average false positive, average false negative and hausdorff distance. It is observed from the performance analysis that 97% of images have an average error less than 4 mm and 3 mm for method 1 and method 2 respectively. Among the two novel methods, PM boundary extracted by method 2 is closer to radiologist drawn PM border.

3.1. Introduction

The removal of PM from the mammographic image using computer aided techniques improves the accuracy of breast cancer screening along with a reduction in computation time for searching the whole mammogram. The PM in MLO views are unique and it appears visually continuous enough to be identified easily even by the untrained observer. Even though the PM boundary seems to be easily detected, it is difficult to accurately delineate them in mammographic images. Evidence of abnormal axillary lymph nodes can be obtained by the internal analysis of pectoral region [Ferrari, 2004]. The similarity in densities and texture among PM and dense breast tissues results in inefficacy of detection methods. Inaccurate detection may in turn lead to an increase in false negative and false positive rates. The similarity in density and texture among PM region and breast parenchyma could bias the detection of various abnormalities. Intensity gradient at the PM boundary may generate false alarms. Around 52% of error is due to the misinterpretation of breast cancer signs and 43% of missed abnormalities results from overlooking signs [Ferrari, 2004]. Radiologists usually compare right and left views of mammograms using landmarks such as nipple and PM for the detection of breast cancer abnormalities.

In a mammographic image, breast region is the area that contains the region of interest for breast cancer analysis whereas the non-breast region contains PM, labels as well as artifacts such as markers, wedges etc. [Masek, 2004]. Figure 3.1 shows various artifacts as well as PMs of mammographic images from the MIAS database. Extraction of region of interest from the various perceptible and structurally distinct regions helps us to confine the search for anomalies only to the breast region. Thus removal of non breast region considerably reduces the false positives along with reduction in processing time. The accuracy in breast region interpretation reduces unnecessary biopsies.

This chapter describes two methods for extracting breast region, obtained after the removal of PM region, label, opaque markers and noise. For removing noise, labels, and other unwanted artifacts, weiner filtering and a global threshold operation are performed. The PM for the LMLO view is visible on top left corner of mammographic image. An optimal threshold value is applied on the selected region of interest (ROI) to determine the connected region at the top left corner as an approximate PM region. The geometrical properties of edge structures on a contrast enhanced coarse version of the mammographic image are used to identify the major PM boundary component.

The images from MIAS database [Suckling, 1997] and the databases created from images supplied by the Lakeshore hospital, Kochi [Lakshmanan, 2014] were considered for experimental analysis. Automatically detected PM boundaries were compared with ground truth information provided by the radiologist. Qualitative as well as quantitative analyses were performed based on these results. The mentioned results were analyzed and compared with four well known methods in current literature to benchmark the proposed method.

3.2. Literature review

Some of the major works proposed for the automatic detection of PM boundary includes those based on straight line, curve, statistics, Gabor wavelet, and watershed transformation. The high density regions of PM are clearly visible in the upper part of most MLO views of mammograms [Homer, 1997]. Karssemeijer [Karssemeijer, 1998] developed a pectoral removal method based on the application of a threshold value on the peak values obtained using hough transform. As breast tumors are frequently projected in the lower areas of the pectoral and the presence of the intensity gradient at the PM edge may easily generate false alarms or miss tumors, the position of PM has great significance. Yam et al, [Yam, 2001] detected

curved pectoral boundary using a dynamic programming method on the straight line boundary obtained by Hough transform.

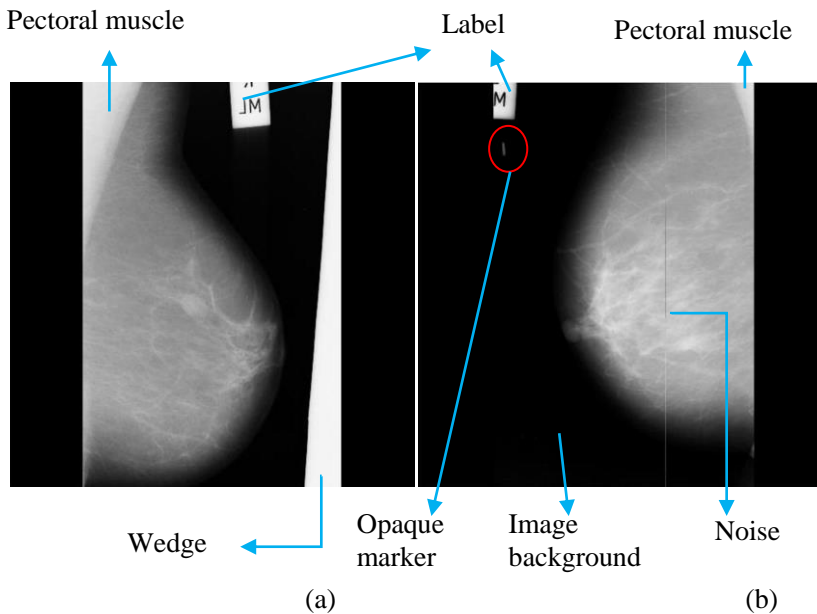


Figure 3.1: Pectoral muscle, wedge, label, marker and noise in mammographic images from MIAS database, mdb012, (b) mdb013.

Kwok, et al, [Kwok, 2004] proposed an automatic method for PM extraction by estimating a straight line with correct location and orientation where an iterative cliff detection method is applied for an accurate segmentation. Ferrari et al, [Ferrari, 2004] enhanced the PM edge by employing a group of Gabor filters. The phase of each pixel in the filtered image is propagated until it reaches a location where two opposite directions of flow encounter each other. Bajger [Bajger, 2005] determined an estimate of PM boundary by initializing an active contour using minimum spanning tree. Ma utilized a graph based approach utilizing adaptive pyramid [Ma, 2007]. The quality of mammograms is assessed by the position, shape and visibility of PM [Christina, 2008]. Even though the boundary of PM seems visible, it is very difficult to demarcate the PM from the mammographic images due to the large fluctuation in edge potency.

Chakraborty et al, [Chakraborty, 2011] proposed a method to automatically segment the PM from mammographic image. The various properties of regions of PMs exploited in the proposed method are its high intensity nature, sharp intensity variation at edge, approximate triangular shape at upper corner, and straight line edge with certain orientation. Chakraborty realized a straight line boundary approximation using weighted average gradient, adaptive band division and selection. A precise muscle boundary is obtained by taking a small region around the line and finding out more accurate points employing local gradient search. Cardoso et al, [Cardoso, 2010] proposed a method to separate the PM in a mammogram by finding the shortest path between two regions in a graph, after conveniently modeling the image as a weighted graph. Camilus et al, developed an automatic PM segmentation method using watershed transformation. Correction of PM region avoiding over segmentation is achieved from watershed line indicating the PM edge [Camilus, 2011].

Even though the approximate PM boundary can be detected using line based techniques the PM boundaries having curved nature makes such methods inefficient. The primary focus of the proposed method is to detect principal boundary component of PM in a reduced region of interest where the PM lies. In most methods, it is difficult to separate the PM boundary accurately from dense tissues. The reason for this is mainly due to the similarity and low contrast nature of breast and PM tissues. A histogram based contrast enhancement method, gray level grouping (GLG) is used to improve the contrast between PM and the dense mammary tissues. The boundary coordinates of PM is utilized to extract the breast region in mammographic image for anomaly detection. In mammograms, factors such as intensity, linearity, edge strength and visual texture demand an accurate and automatic detection of edge [Sauders, 2007]. The three regions including breast, background and PM are modeled using a supervised single strategy based on position, intensity and texture information is utilized for extracting breast region in mammogram [Oliver, 2014]. An accurate breast

MRI segmentation in CAD systems and breast density assessment considers PM delineation as an important preprocessing step [Lianghao, 2014], Fooladivanda, 2014].

The peculiar characteristics of PMs utilized for finding its boundary are

1. PM region usually occupies a predominant portion in MLO view of mammograms.
2. PM boundary is roughly a straight line [Ferrari, 2004].
3. Sharp intensity discontinuities exist at the PM boundary.
4. PM lies at the top left corner of LMLO view of mammographic image.
5. PM boundary line orientation of left oriented image lies within 30° to 80° [Ferrari, 2004].
6. PM boundary is visually separable but difficult to segment.
7. PM region can be considered roughly as a triangular region [Chandrasekhar, 2001].
8. Width of PM region gradually decreases from top to bottom.
10. The two edges of PM boundary are part of breast image boundary.
11. PM boundary is clearly visible in its approximation image.

3.3. Proposed methods for PM boundary detection

The detection of PM boundary coordinates helps to extract breast region in a mammogram. The proposed PM removal technique consists of four major steps.

1. Preprocessing operation to generate a triangular ROI including the PM.
2. Detection of suspicious segments of PM boundary.
3. Selection of major segment of PM boundary.
4. Extension of PM segment to the mammographic image border.

A preprocessing operation on mammogram removes unwanted artifacts and labels. The visibility of the PM boundary is improved using a contrast enhancement technique. Two methods are proposed for the

detection of suspicious segments including PM boundary. The first method generates contour segments of homogeneous regions whereas the second method extracts the edges of coarse image as suspicious PM boundary segments. In both methods, a segment which satisfies the property of PM boundary is selected as the major PM boundary segment. The end points of the identified segment are extended to the two borders of mammogram to obtain complete PM boundary coordinates. The boundary coordinates so obtained are utilized to generate the breast region. The flow charts of both the proposed methods are shown in figure 3.2.

3.3.1. Pre-processing operation

The preprocessing operation in the PM boundary detection reduces the search area of mammogram for effective breast cancer analysis. The major steps of preprocessing operation consists of

1. Unwanted artifact removal
2. Keeping the orientation of mammogram towards left
3. Enhancing the contrast
4. Selection of a triangular ROI containing PM region.

3.3.1.1 Removal of unwanted artifacts and label

Removal of unwanted artifacts from the breast region is a prerequisite operation. Here a significant breast area is extracted from various artifacts, labels and background region. The unwanted artifacts of a mammographic image include wedges, opaque markers, noise etc. [Masek, 2004]. In the LMLO view of mammogram, the PM region occupies the top left corner.

The first step of preprocessing operation consists of three steps including

1. Noise reduction
2. Global thresholding
3. Large area detection.

Sauders et al, concluded that the quantum noise due to low-counts of x-ray

photons in acquisition system is very prominent in mammographic images.

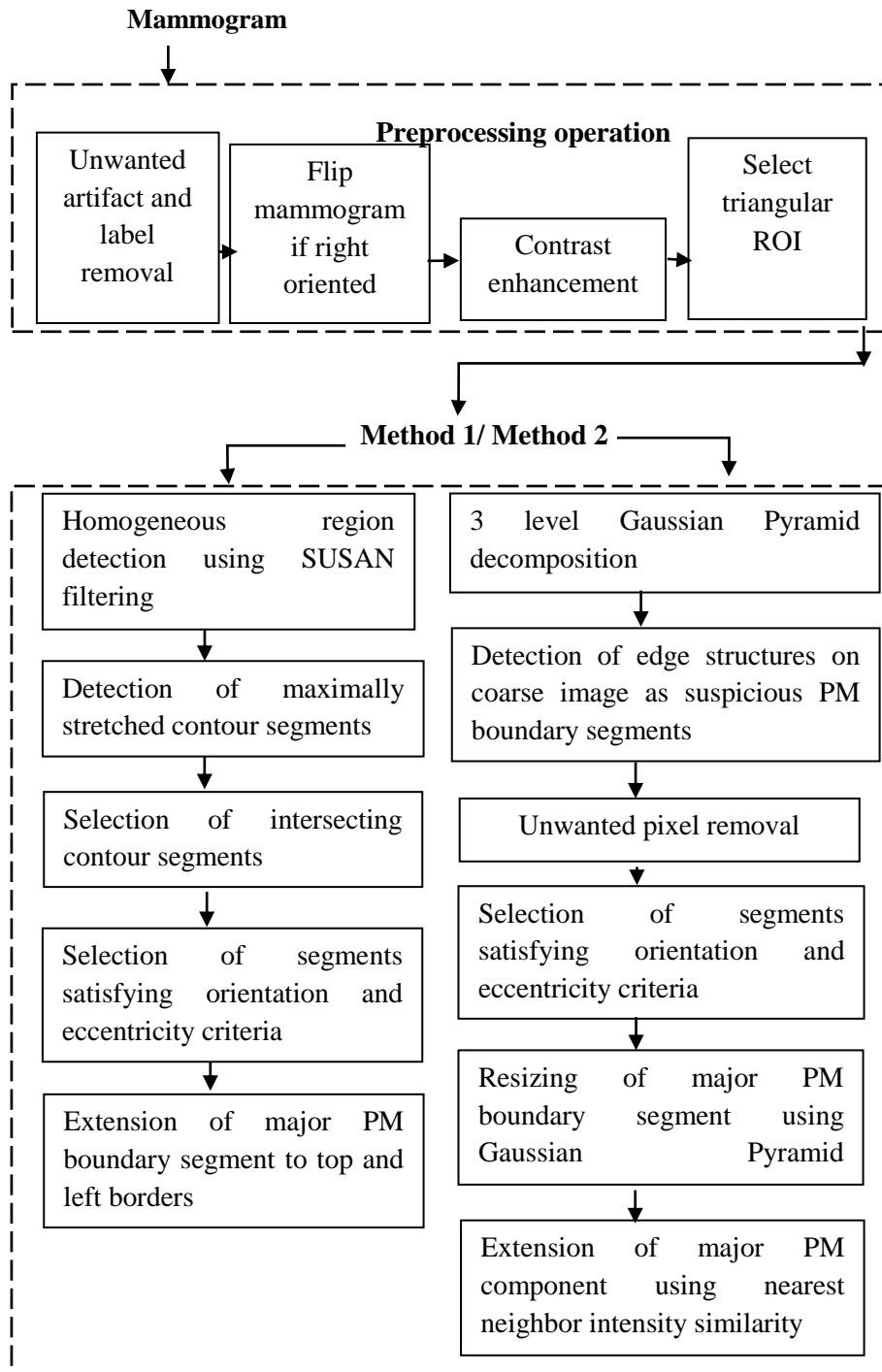


Figure 3.2: Flowchart of the methods for the detection of pectoral muscle boundary

During breast cancer analysis, the presence of noise may bias the judgments of radiologist [Sauders, 2007]. Therefore an edge preserving noise removal using wiener filtering is performed on the mammographic image [Mayo, 2004]. An enhanced smoothing is performed in regions with less variance. Figure 3.3 shows the resultant image obtained after wiener filtering of mdb013 from MIAS database.

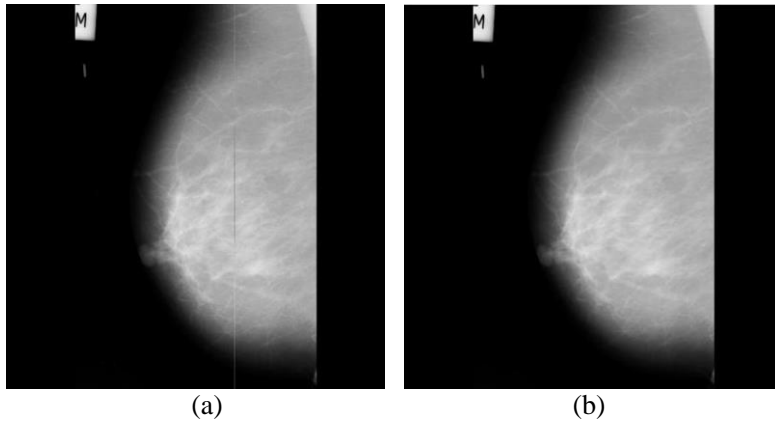


Figure 3.3: Noise removal (a) original, (b) after Wiener filtering of mammogram, mdb013

A global threshold is employed to obtain a binary image. The threshold, TH used for this operation is selected as

$$TH = (I_{\min} + I_{\max}) \times 0.1 \quad (3.1)$$

where I_{\min} and I_{\max} are the minimum and maximum intensity value of image, I respectively. This value for TH was obtained empirically.

The binary image with largest area connected component represents the breast region along with PM. The result obtained after thresholding operation is depicted in figure 3.4(a). The resultant image consists of various regions representing labels, opaque markers, wedges and breast region with PM. The connected components belonging to various regions are grouped together using eight connectivity [Mayo, 2004].

The largest connected component representing breast area including PM region is set to 1 and the rest is assigned as 0 as shown in figure 3.4(b). This component will act as a mask to extract the breast region from the mammographic image. The noise suppressed image after wiener filtering is multiplied by this mask to obtain the ROI, which includes the PM region and the breast region. This is illustrated in figure 3.4(c).

3.3.1.2. Detection of orientation view of mammogram

An MLO view of mammographic image clearly shows the PM as a triangular region [Ma, 2007]. The orientation of the MLO view is found by checking the total number of nonzero pixels in a column at the boundary of ROI. For an LMLO view, the number of pixels in the column corresponding to the left boundary of ROI is more than that of right boundary.

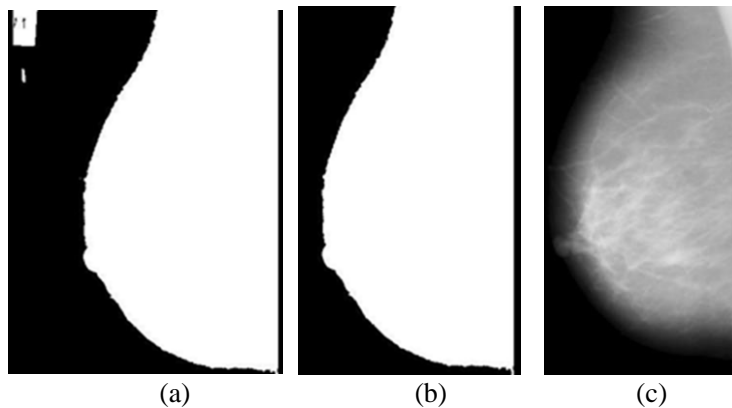


Figure 3.4: Mammographic image, mdb013 after (a) thresholding operation (b) largest area component, (c) retained large area component in noise.

PM appears at the top left corner in an LMLO view and top right corner in an RMLO view of mammogram. In order to simplify the geometrical analysis of PM boundary, the algorithm is designed in such a way as to keep the position of PM region at left side of ROI. An ROI with PM at the right side is flipped to left so that the MLO view is left oriented and the PM

is always at the top-left corner.

In the proposed method, the position of PM is determined by finding the nonzero pixels in a column lying ten pixels away from the left and right boundary of ROI. The reason for not choosing the left/ right boundary for pixel count is to avoid the defects or scanning imperfection at the boundary as found in the case of mdb085 [Suckling, 1994] shown in figure 3.5.

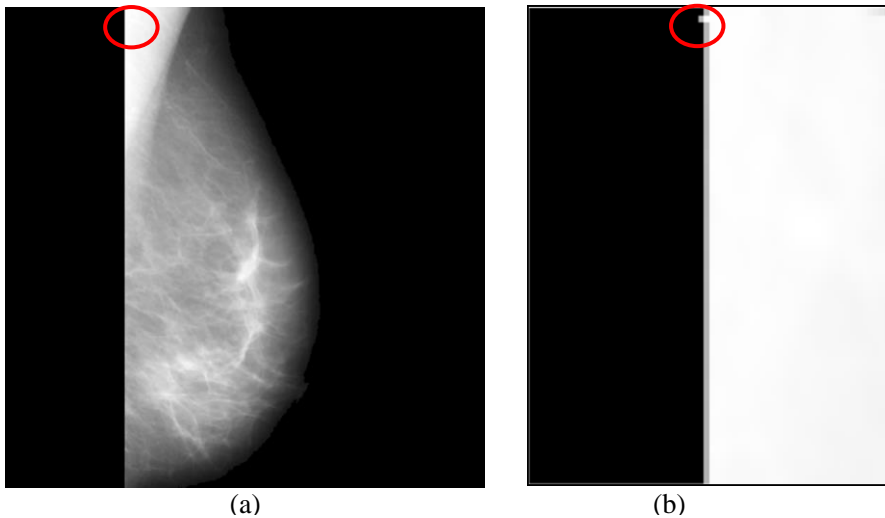


Figure 3.5: Boundary defect in a mammogram, mdb085 from MIAS database. (a) ROI including breast region and PM, (b) Enlarged portion of defect on the boundary of ROI. Red circle encloses the defect on the boundary

3.3.1.3. Triangular ROI selection

The PM region occupies the top left corner of LMLO view of mammogram. Figure 3.6(b) depicts a mammogram after removing unwanted artifacts, labels, noise and background region of figure 3.6(a) from MIAS database. Hence a triangular ROI is selected with the top leftmost pixel as one vertex and the bottom left pixel as the second vertex. A pixel located 50 pixels to the right of the first non-zero pixel on the first row, is selected as the third vertex of this triangle to account for imperfections if any that occurred during the thresholding operation in the

preprocessing step. Figure 3.6(c) shows a triangular ROI of mdb004 from MIAS database after removing unwanted artifacts, labels, noise and background region.

3.3.1.4. Contrast enhancement using gray level grouping (GLG)

Enhancement of contrast improves the visual quality of an image. gray level grouping (GLG) proposed by Abidi [Chen, 2006], [Lakshmanan, 2008] is an efficient and effective automatic technique to enhance the contrast of medical images. GLG, a histogram based technique excels various conventional contrast enhancement techniques such as histogram equalization, Hanmandhu's fuzzy based method [Hanmandhu, 2006] etc. GLG process tends to spread the histogram components uniformly over the grayscale, preventing the histogram components from concentrating on particular locations of the grayscale. The GLG algorithm outperformed other contrast stretching methods in increasing the contrast between the PM region and the breast region in the triangular ROI. The result of GLG contrast stretching for the image mdb004 from the MIAS mammographic database is shown in figure 3.6(d), where the boundary of PM is identifiable.

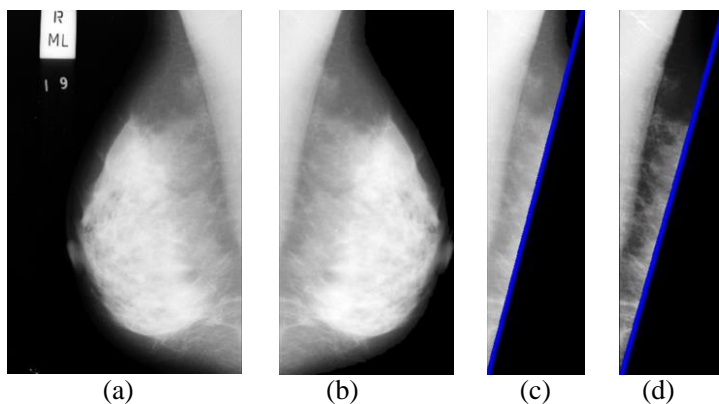


Figure 3.6: Reduced ROI detection. (a) Original image, mdb004 (b) flipped image after unwanted artifact removal, (c) triangular ROI (d) triangular ROI after contrast enhancement (reduced ROI).

3.3.2. Detection of major PM boundary segment

The enhanced triangular ROI consisting of PM and breast regions is used for finding the PM boundary. The PM region is nearly homogeneous with a sharp discontinuity at the border between PM and breast region. Hence the strong intensity variations in the enhanced triangular ROI are utilized for finding PM boundary segments. The following two methods are proposed for obtaining the suspected PM boundary segments: The first method extracts suspected PM boundary segments using contour segments of homogeneous regions obtained using SUSAN [Smith, 1997] filtering technique. In the second method, a multiscale decomposition is used for getting an approximation image. An edge detector on the coarse image yields strong edge structures. A technique for removing unwanted pixels from the edge structures generates suspected PM boundary segments. A major component of the PM boundary is selected from these segments using geometrical properties of the PM boundary.

3.3.2.1. METHOD 1: Using contours on homogeneous ROI

The various steps involved in the first method for selecting major segment of PM boundary include

1. Homogeneous region extraction
2. Contour detection
3. Detection of maximally stretched contour segments
4. Selection of intersecting contour segments
5. Selection of major PM boundary segment using orientation and eccentricity criteria
6. Extension of major PM boundary segment to top and left borders

3.3.2.1.1. Homogeneous region extraction using SUSAN filtering

The PM region with relatively high intensity values is roughly homogeneous [Kwok, 2004], [Ferrari, 2004]. The reduced ROI as shown in Figure 3.6(d) consists of regions of PM as well as breast. A homogeneous

region extraction method on reduced ROI groups similar intensity areas such as PM region. In the proposed work, single univalue segment assimilating nucleus (SUSAN) [Smith, 1997] is used for extracting similar regions. It is a low level non-linear filtering technique for local segmentation. The local regions are formed on the basis of pixel intensity similarity. A circular mask is moved over the entire breast region for grouping regions with similar intensity. Figure 3.7 shows a dark image on a white background. A circular mask shown in figure 3.7 at four different positions explains the SUSAN principle for grouping local regions. The center pixel of image region marked as red spot in figure 3.7, enclosed in the circular mask for processing is called nucleus. The locally grouped pixels with similar intensity value as that of nucleus are called univalue segment assimilating nucleus (USAN). Figure 3.7(b) shows USAN area of the circular mask in green shade. This technique isolates similar local structures of the image. The number of pixels assimilated in USAN area decides whether the nucleus belongs to a homogeneous, edge or corner region. The number of pixels in USAN area of a homogeneous region is very high compared to that of an edge which in turn is more than that of a corner region. The homogenous regions in an image are obtained using a circular mask of size 37 pixels.

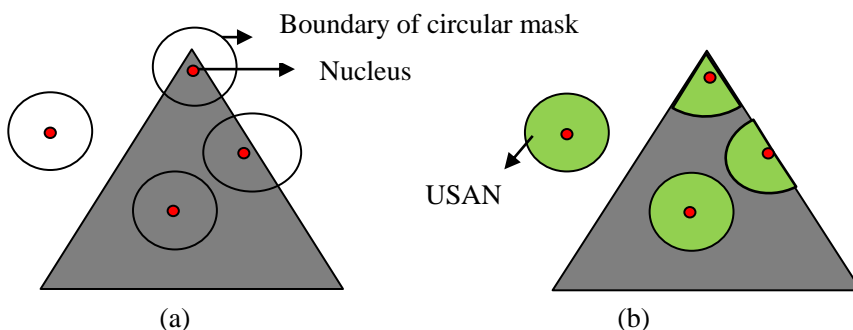


Figure 3.7: Single univalue segment assimilating nucleus (SUSAN) principle. (a) Four circular mask with nucleus (red) placed a different position on an image for explaining SUSAN principle, (b) USAN area (green shade)

0	0	1	1	1	0	0
0	1	1	1	1	1	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	0
0	0	1	1	1	0	0

Figure 3.8: Single univalue segment assimilating nucleus (SUSAN) spatial filter mask

The original function, $C(\vec{r}, r_0)$ for checking the similarity is given in equation 3.2.

$$C(\vec{r}, r_0) \begin{cases} 1 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| \leq t \\ 0 & \text{if } |I(\vec{r}) - I(\vec{r}_0)| > t \end{cases} \quad (3.2)$$

where I is the image and \vec{r} and \vec{r}_0 are the nucleus and a pixel enclosed in the USAN area respectively and t is a user defined brightness difference threshold. The value of t is obtained through empirical analysis. A smoother and stable form of equation 3.2 proposed by Smith [Smith, 1997] is used in the proposed method 1 and is given in equation 3.3.

$$C(\vec{r}, r_0) = e^{-\left(\frac{|I(\vec{r}) - I(\vec{r}_0)|}{t}\right)^6} \quad (3.3)$$

The representation of intensity similarity functions in equation 3.3 and 3.4 are represented in figure 3.9 as 'a' and 'b' respectively. A smooth brightness comparison function is better than a sharp cut off function as the pixels brightness vary slightly without having much effect on C , even near the threshold. Noise suppression is also possible through the use of a smoothed profile than a sharp cut off spatial profile.

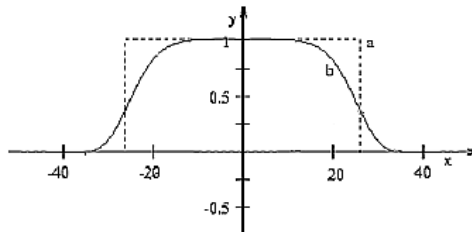


Figure 3.9: Intensity similarity functions. Dotted curve (a) shows a sharp cut off function and continuous curve (b) represents a Gaussian function.

Chapter 3: Pectoral Muscle Delineation

For each nucleus, the aggregate of C for each enclosed pixel is calculated as given in equation 3.4.

$$n(\vec{r}_0) = \sum_r C(\vec{r}, \vec{r}_0) \quad (3.4)$$

A modification in SUSAN filtering technique is employed here to obtain the homogeneous region rather than edge features in the triangular ROI. In the proposed method, for extracting a homogeneous region representing PM, the accumulated sum, $n(\vec{r}_0)$ is compared with a global threshold $g = 3n_{max}/4$, where n_{max} is the maximum value of n , to obtain the response, R as given in equation 3.5. In the original SUSAN filtering method, a nucleus is included as an edge if the accumulated sum is less than g .

$$R(\vec{r}_0) = \begin{cases} g - n(\vec{r}_0) & \text{if } n(\vec{r}_0) > g \\ 0 & \text{otherwise} \end{cases} \quad (3.5)$$

The value of R in equation 3.5 decides whether the pixel should be included in the local homogeneous region or not. Figure 3.10 shows the resultant image, mdb004 after applying homogeneous region extraction technique on reduced ROI.



Figure 3.10: Homogeneous region extraction in mammogram, mdb004 using SUSAN filtering technique.

3.3.2.1.2. Contour extraction

Contours are closed paths enclosing regions with uniform intensity. Contours represent the path of discontinuities. The locally grouped homogeneous regions in reduced ROI are extracted using contours. Figure 3.11 shows the contours of homogeneous region in mdb004. One of the regions inside the contours of reduced ROI represents the PM region.

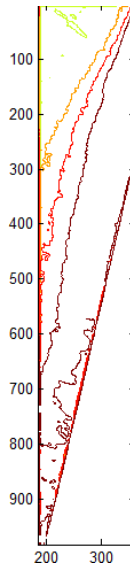


Figure 3.11: Contours of extracted homogeneous region in reduced ROI for mdb004.

3.3.2.1.3. Extraction of maximally stretched contour segments

The contours corresponding to each homogeneous region are closed paths enclosing the corresponding area. In an LMLO view of mammogram, PM region is roughly considered as a triangular area where the top and left side are borders of the mammographic image [Ferrari, 2004], [Chandrasekhar, 2001], [Ma 2007]. The third side of PM region is in touch with breast area. In the proposed method, the contour of third side of the PM region in touch with breast area is selected for analyzing the properties of PM boundary.

The PM boundary segment is selected as a segment of a contour that has the maximum lengths in the direction of x and y axes, and labeled as the maximally stretched segment of that contour. One of the maximally stretched contour segments for the ROI of the image mdb004 is shown in thick red colour in Figure 3.12.

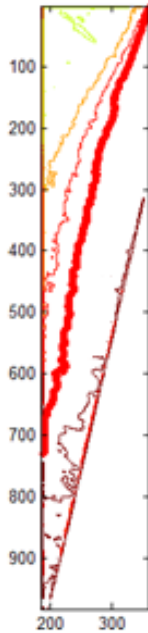


Figure 3.12: Maximally stretched contour segment extraction. One of the maximally stretched contour segment of mdb004 representing PM boundary.

3.3.2.1.4. Selection of intersecting segments

The major PM boundary component is one of the maximally stretched contour segments in the ROI. As the PM region in LMLO view of mammogram is always situated in the triangular ROI selected as discussed earlier, a line connecting the top-left corner to the midpoint of the hypotenuse of this triangle will always intersect the major PM boundary segment. Figure 3.13 shows such maximally stretched contour segments satisfying the criteria of intersection.

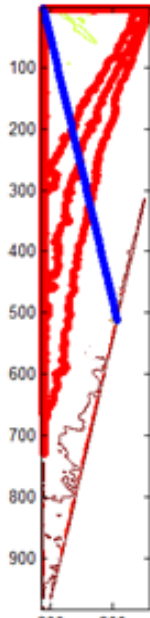


Figure 3.13: Maximally stretched contour segments satisfying intersection criteria. Suspected PM boundary contours (red colour). Blue line represents intersecting line for finding suspected PM boundary.

3.3.2.1.5. Selection of major PM boundary segment using orientation and eccentricity criteria

The major segment corresponding to PM boundary is selected from the maximally stretched contour segments by analyzing geometrical properties such as orientation and eccentricity. The property of each maximally stretched contour segment is obtained by fitting it in an ellipsoid as shown in figure 3.14.

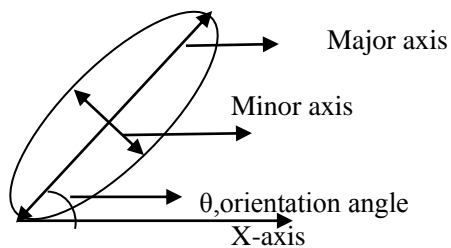


Figure 3.14: Orientation angle of image region

Chapter 3: Pectoral Muscle Delineation

The angle between the major axis and x-axis provides the orientation of a component. Here orientation angle of the PM boundary component (represented as major axis in figure 3.14) from the x-axis is supposed to lie between 30 and 80 degree [Ferrari, 2004]. Figure 3.15 shows the suspected PM boundary components satisfying this orientation property. It is obvious that the number of suspicious PM boundary segments satisfying the orientation property can be more than one. Therefore the orientation property by itself is not enough to select the major PM boundary component from the suspicious contour segments. As the PM boundary is approximately a straight line [Ferrari, 2004], the eccentricity property is chosen as the next criteria for selecting major PM boundary component.



Figure 3.15: Maximally stretched contour segments of mdb004 satisfying orientation property of PM boundary

Eccentricity of an elliptical region can be calculated either as the ratio of the distance between the foci of the ellipse and its major axis length or as the ratio of distance from the center to a focus and distance from that focus to a vertex. The range of eccentricity value is between 0 and 1 where

0 indicates a circular region while 1 represents a line segment [Edward, 2011]. Segments satisfying orientation property with an eccentricity value near to 1 are therefore better candidates to be selected as PM boundary component. In the proposed method, a major PM boundary segment is selected as the longest suspected segment having an eccentricity value greater than 0.99. Figure 3.16 depicts the major PM boundary segment of mdb004 obtained using method 1.

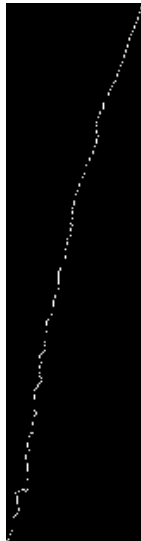


Figure 3.16 : Major PM boundary segment satisfying eccentricity for mdb004.

The major PM boundary segment is selected as the PM boundary if the endpoints are at the top and left boundary of mammogram. Otherwise an extension process based on neighbouring pixel intensity similarity is applied for extending this segment to the top boundary and the left boundary

3.3.2.1.6. Extension of major PM boundary segment

The principal PM boundary component can be extended to the top and left boundary of mammographic image by selecting pixels based on the assumption that the PM region is roughly of triangular shape. Figure 3.17 shows the selection criteria to be considered while extending the major

Chapter 3: Pectoral Muscle Delineation

PM boundary component. RP, VP, HP and DP in figure 3.17 represent reference pixel, vertical pixel, horizontal pixel, diagonal pixel respectively. From the reference pixel, the next neighbouring pixel to be included for extension can be vertical, horizontal or diagonal. For extending upwards, the neighbouring pixel in the outward direction is considered while an inward extension direction is utilized in the case of extensions towards left boundary. On comparison of intensity value of RP with that of VP, HP and DP, the pixel with nearest intensity value to that of RP can be identified and selected as next pixel for extension. Figure 3.18 depicts the extended major PM boundary component of mdb004 satisfying intersection, orientation and eccentricity criteria.

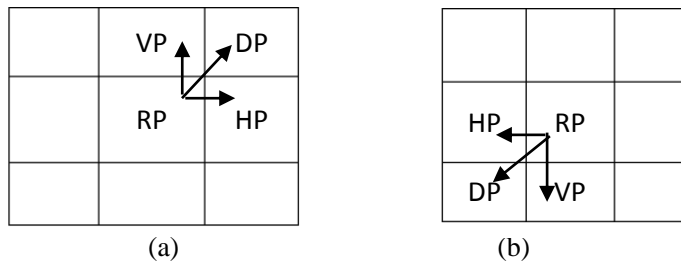


Figure 3.17 Extension of major PM boundary component (a) to top boundary (b) to left boundary of mammographic image.

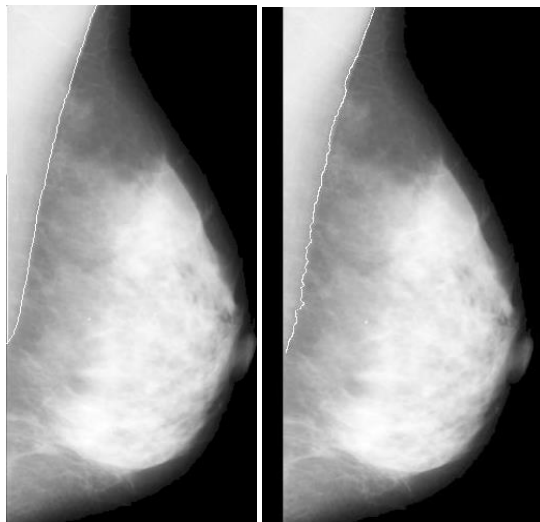


Figure 3.18: Extended PM boundary of mdb004 obtained using method 1.

3.3.2.2. METHOD 2: Using edge detector on coarse ROI

In method 2, the PM boundary is extracted using the following steps

1. Gaussian pyramid decomposition
2. Canny edge detection
3. Removal of unwanted pixels
4. Selection of major PM boundary segment using orientation and eccentricity criteria
5. Gaussian pyramid reconstruction
6. Extension of major PM boundary segment

3.3.2.2.1. Gaussian pyramid decomposition

A multiscale representation of images at various resolutions is useful in analyzing prominent features such as edge [Lopez, 2011]. A hierarchical representation is obtained by recursive low-pass filtering and downsampling. In the proposed method, a multiscale representation of mammographic region is obtained using a three level Gaussian decomposition with 9/7 filter [Burt, 1983]. Figure 3.19 illustrates the various stages of three level decomposition using Gaussian pyramid on enhanced triangular ROI. Here I_0 is the original image, I_1 , I_2 and I_{coarse} are the images at next levels with reduced resolution. I_{coarse} is the approximation or coarse image. Reduced resolution approximation is hierarchically obtained at each level.

The strong edges such as the PM boundary are visible in the coarse resolution whereas weak edges are suppressed. Hence edge detection on coarse image generates prominent edges including PM border.

3.3.2.2.2. Canny edge detection at coarse level

Canny edge detection [Canny, 1986] with good localization features is used to detect and locate sharp discontinuities representing the PM boundary from the coarse image. Figure 3.20 illustrates the need for

performing a three scale smoothing operation for the detection of edges. The disruption of strong edges obtained using approximation image at the third level are less compared to that in high resolution images.

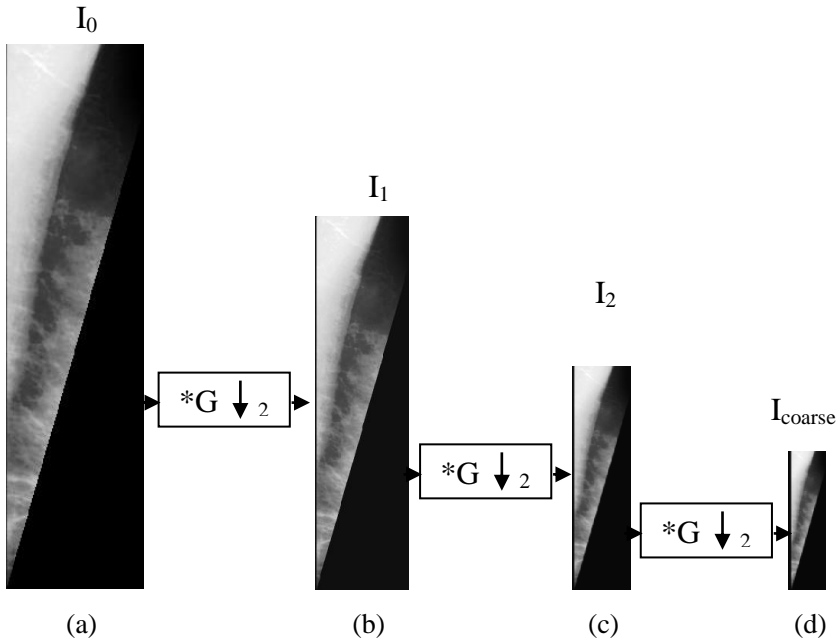


Figure 3.19: Three level gaussian pyramid decomposition on mdb004

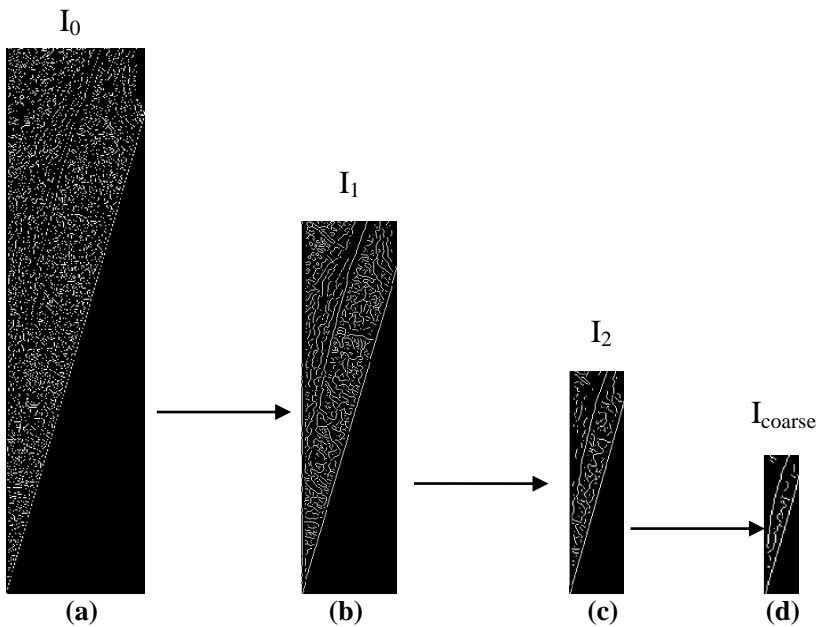


Figure 3.20: Canny edge detection on images obtained for a 3 level Gaussian pyramid decomposition on mdb004

In canny edge detection, a gradient operation is performed on a smoothed image. A double threshold is applied to detect the PM boundary. Figure 3.21 illustrates the resultant edge structures obtained for mdb004 from MIAS database. In figure 3.21(a), a reduced ROI is utilized to get the edge image, whereas in figure 3.21(b), the entire image is used. It is obvious that the reduction to the ROI results in less number of false edges.



Figure 3.21: Edge structures of coarse representation of mdb004 (a) with reduced ROI, (b) on entire mammographic image.

3.3.2.2.3. Removal of unwanted pixels

PM region is roughly triangular in shape [Ferrari, 2004]. Hence the number of pixels in PM region narrows down while scanning from top to bottom. The possible pixel patterns of the edge components obtained for a PM boundary are shown in figure 3.22. For each edge component, analyzed from the top to bottom, RP denotes the reference pixel, HL the pixel horizontally left of RP, VD the pixel vertically down to RP and DDL the pixel diagonally left of RP. While scanning from top to bottom, the next RP may move in any of the three directions viz., DDL, HL or VD. Pixels in

the direction of HR_2 , i.e., two pixels to the right of RP, are removed as the PM with an approximate triangular shape will not extend to the right.

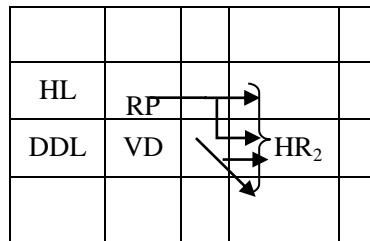


Figure 3.22: Pixels involved in connected edge component (scanning from top to bottom)

Removal of such deviating pixels from the suspicious segments helps for an efficient analysis of properties of edge structures in later stages. Figure 3.23 shows the resultant image after removing these unwanted branches from the ROI for mdb004.



Figure 3.23: Edge structures obtained after removing unwanted pixels on coarse image of reduced ROI.

3.3.2.2.4. Selection of major PM boundary segment using orientation and eccentricity criteria

The connected components obtained after removing the unwanted pixel branches are utilized for selecting the major component representing PM boundary. The orientation and eccentricity properties mentioned in

section 3.3.2.1.5 are used for selecting a major PM boundary component from the suspicious segments. Figure 3.24 shows the retained suspicious segments of mdb004 within an orientation of 30° to 80° . Among the retained structures, the component with a maximum eccentricity value is selected as the major PM boundary component. The retained structure is shown in figure 3.25.



Figure 3.24: Edge components of mdb004 satisfying orientation property of PM boundary



Figure 3.25: Edge component of mdb004 satisfying eccentricity property.

3.3.2.2.5. Gaussian pyramid reconstruction

A gaussian pyramid reconstruction [Burt, 1983] is performed with the major PM boundary component. The size of reconstructed image remains same size as that of original image used for decomposition. Figure 3.26 illustrates the reconstructed image of major PM boundary segment.



Figure 3.26: Major PM boundary segment after Gaussian pyramid reconstruction

3.3.2.2.6. Extension of major PM boundary segment

The major PM boundary component satisfying orientation and eccentricity criteria is extended to the top and left borders as per the method described in section 3.3.2.1.6. Figure 3.27 depicts the result of PM boundary component of mdb004 obtained using method 2.

3.4. Experimental setup and performance measures

A total of 83 images from MIAS database [Suckling, 1994] taken for the experimental analysis are of size 1024x1024 with 200 μm sampling interval and 8-bit gray-level quantization. The 83 mammographic images of MIAS database that were used by Ferrari et al, [Ferrari, 2004] were

employed in the proposed method for PM boundary extraction.



Figure 3.27: PM boundary of mdb004 drawn by (a) radiologist, (b) method 2

The PM boundaries obtained were compared with ground-truth contours provided by Dr. Rangayyan, University of Calgary [Ferrari, 2004]. Performance evaluation and comparison of the proposed method with four methods from current literature [Ferrari, 2004], [Camilus, 2011], [Chakraborty, 2012] were conducted using three performance measures including the false positive average, false negative average and hausdorff distance [Ferrari, 2004], [Camilus, 2011], [Bajger, 2005], [Ma, 2007], [Huttenlocher, 1993]. The row normalized error [Bajger, 2005] of the proposed method was also computed to find out the average error per row.

Figure 3.28 represents the false positive and false negative region for performance analysis. The PM boundary drawn by radiologists is selected as the reference for finding false positive and false negative region. False positive region is formed when the PM boundary generated by the proposed method is outside the ground truth boundary that is when the

detected PM boundary moves into the breast region. False positive region is calculated as the region outside the ground truth into the breast region, but inside the PM boundary formed by the proposed method. False negative region is obtained when the PM boundary drawn by the proposed method is inside the ground truth boundary, i.e., inside the PM region. False negative region is computed as the region outside the proposed algorithm detected PM boundary but inside the reference boundary.

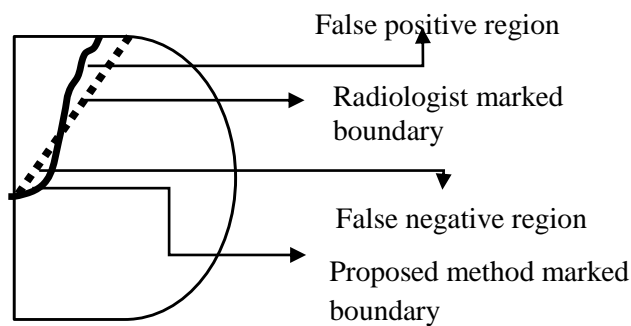


Figure 3.28: False positive and false negative region. Dotted line indicates the pectoral boundary drawn by radiologists whereas the solid line represents the PM boundary obtained by the proposed method.

False positive average and false negative average are respectively the mean of false positive and false negative rates of all N images used for the evaluation, which are 83 in this case. These parameters, generally known as area normalized error (ANE) are dependent on the included area of PM region. The inclusion of PM region during image acquisition may vary according to the positioning of breast [Ma, 2007]. Hence row normalized error (RNE), which is independent of area of PM region, is also included in the proposed work as a performance measure. For RNE , the area between boundaries of PM drawn by radiologist and proposed method were normalized using total rows occupied by PM. The mismatch between the relative positions of boundaries obtained by proposed method and that drawn by radiologists is obtained using hausdorff distance. The various

performance measures such as false positive average (FP_{ave}), false negative average (FN_{ave}) and row normalized error (RNE) are defined below.

$$FP_I = \frac{1}{AR} \sum_{i=1}^{nr} \max[0, HCP(i) - HCR(i)] \quad (3.6)$$

$$FP_{ave} = \frac{1}{N} \left[\sum_{k=1}^N FP_{I_k} \right] \quad (3.7)$$

$$FN_I = \frac{1}{AR} \sum_{i=1}^{nr} \max[0, HCR(i) - HCP(i)] \quad (3.8)$$

$$FN_{ave} = \frac{1}{N} \left[\sum_{k=1}^N FN_{I_k} \right] \quad (3.9)$$

$$RNE = \frac{1}{nr} \left[\sum_{k=1}^N HCR(i) - HCP(i) \right] \quad (3.10)$$

where value of m indicates the row size of the image, AR represents area of PM region marked by the radiologist, nr represents the maximum number of rows among radiologist and proposed method identified PM border, $HCP(i)$ represents the horizontal coordinate in i^{th} row of PM boundary obtained using the proposed method, $HCR(i)$ implies horizontal coordinate in i^{th} row of PM boundary drawn by radiologists, N represents the total number of images considered for performance analysis (here 83). The range of k is from 1 to N to find the average false positive and the average false negative. Hausdorff distance (HD) is defined as

$$HD(R,P) = \max(h(R,P), h(P,R)) \quad (3.11)$$

$$h(R,P) = \max_{x \in R} \min_{y \in P} D(x,y) \quad (3.12)$$

where $D(x,y)$ is the Euclidean distance between two points x and y .

3.5. Results and discussion

The results obtained from the proposed method were compared with the ground-truth contours detected using hough and gabor based methods described in Ferrari et al, [Ferrari, 2004], watershed method by Camillus [Camillus, 2011] and average gradient method by Chakraborty [Chakraborty, 2012]. The average performance analysis for 83

mammographic images from MIAS database is indicated in table 3.1.

Computationally efficient codes were developed in MATLAB 8.2 computing environment, and the CPU had as the processor an AMD Athlon II X2 250- 3 GHz processor with 2 GB RAM. The average time taken for the first method was 41.8 second whereas the time taken by the second method was 3.72 second. It included 0.971 seconds for extracting the triangular ROI.

Table 3.1. Performance analysis: Comparison of values reported for the proposed method with existing methods.

Performance Measures	Methods					
	Hough (2004)	Gabor (2004)	Watershed (2011)	Average gradient (2012)	Method 1 (contour)	Method 2 (multiscale)
Hausdorff Distance	33.16	22.69	19.25	19.28	17.09	15.31
FP avg	0.0198	0.0058	0.0085	0.00422	0.0030	0.0028
FN avg	0.2519	0.0577	0.0488	0.0393	0.0421	0.0367

From table 3.1, it is evident that the performance measures calculated for the proposed method 2 is better compared to other techniques indicating that the results are closer to the ground truth PM boundary.

The methods 1 and 2 were also evaluated for the RNE, and the results are illustrated in figure 3.29. In the proposed work, 5 pixels represent 1 mm. 97% of images considered in method 1 have an RNE error less than 4 mm per row whereas for the method 2 this figure is less than 3mm. The proposed methods outperform the approach suggested by Ma [Ma, 2007], where RNE is considered as the parameter for evaluation for which the RNE similarly calculated, is less than 5 mm.

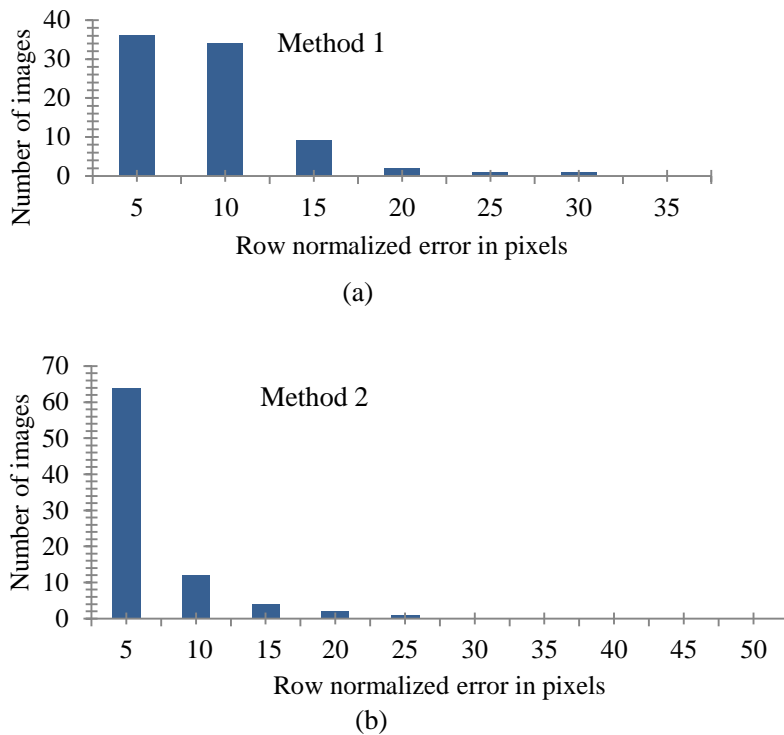


Figure 3.29: Histogram showing RNE of (a) method 1 and (b) method 2. The y axis represents the total number of images and x axis represents the pixel range. The first bar shows the number of images with RNE in the range (0-5] pixels, whereas the second bar represents RNE in the range (5-10] ^[29].

The effectiveness of the algorithms to detect PM boundary were inspected by applying them on various types of mammographic images. Figures 3.30(b) and 3.30(c) demonstrate the boundary detection on mdb125 which has a multilayered PM. The results obtained for the PM boundary are very close to the ground truth PM boundary as indicated in figure 3.30(a).

The presence of dense tissue structures, small PM region, strong edges which are parallel to the PM boundary, misleading false edges are some of the major factors that could adversely affect the accurate delineation of PM boundary. The proposed methods were found good in extracting the PM boundary from mammogram. The performance

Chapter 3: Pectoral Muscle Delineation

evaluation metrics show that the method 2 is the best method for PM boundary delineation.

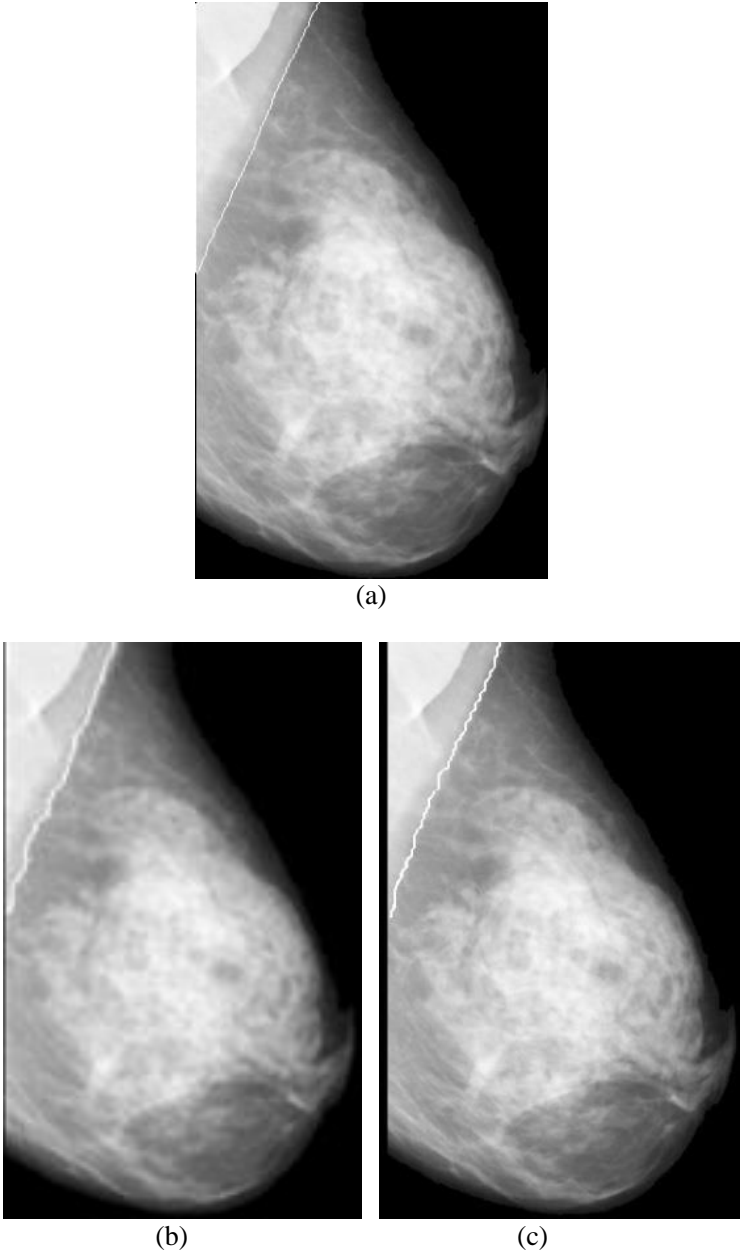
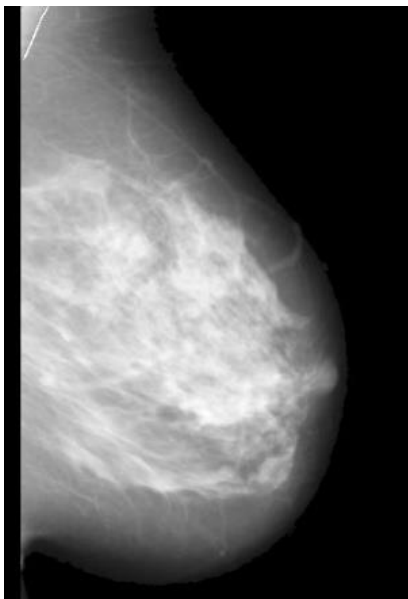
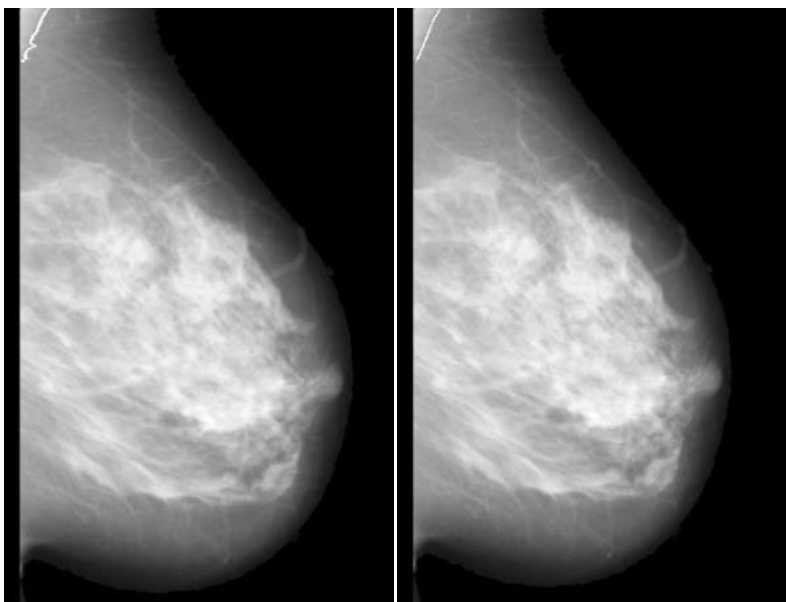


Figure 3.30: PM boundary of mdb125 drawn by (a) radiologist, (b) method 1 and (c) method 2.

Figure 3.31 presents the results for a mammographic image, mdb031 with a minuscule PM region. The small PM region is efficiently identified by the methods 1 and 2.



(a)



(b)

(c)

Figure 3.31: PM boundary of mdb031 with small PM region using (a) ground truth information, (b) method 1 and (c) method 2.

Chapter 3: Pectoral Muscle Delineation

The presence of dense tissues in the PM region can obscure the accurate detection of PM boundary. Figure 3.32(a) shows a dense glandular mammogram, mdb130 with the ground truth for the PM boundary marked by the radiologist. The PM boundary detected by methods 1 and 2 respectively are shown as figure 3.32 (b) and figure 3.32 (c) respectively. It is observed that the proposed methods have succeeded in identifying the correct PM boundary for this dense mammogram.

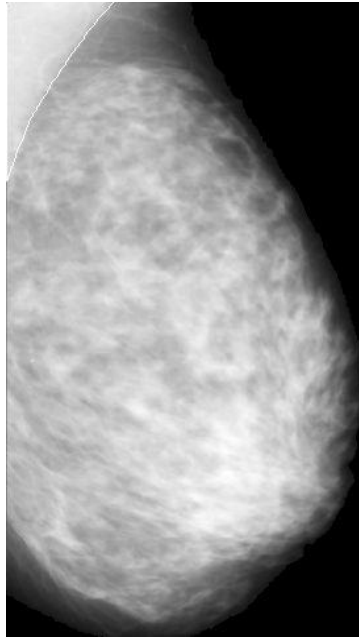
Figure 3.33 shows the results for a glandular noisy right oriented mammogram, mdb013 with small PM region. Both the proposed methods have identified the PM boundary.

Figure 3.34 show the results for mdb028 from the MIAS database, which has a false edge that could obscure the true PM boundary. The methods 1 and 2 have succeeded in identifying the true boundary even in the presence of the false edge.

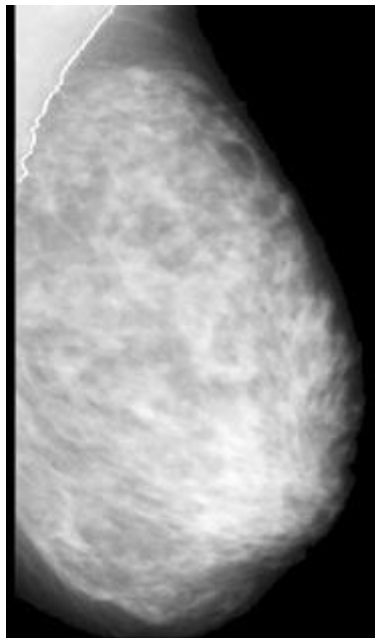
Inaccurate detection is found in some dense images where the PM boundary is obscured by the presence of dense tissues. Figure 3.35 shows the results for a dense mammogram, mdb112, where the algorithms have not properly identified the PM boundary.

In the later chapters, the second method is used due to its efficiency in PM boundary detection. Figure 3.36 depicts the result of the proposed method 2 for the mammographic image collected from Lakeshore Hospital, India. PM boundary coordinates thus obtained can be used to delineate PM region from a mammogram to generate breast region for further analysis of cancer.

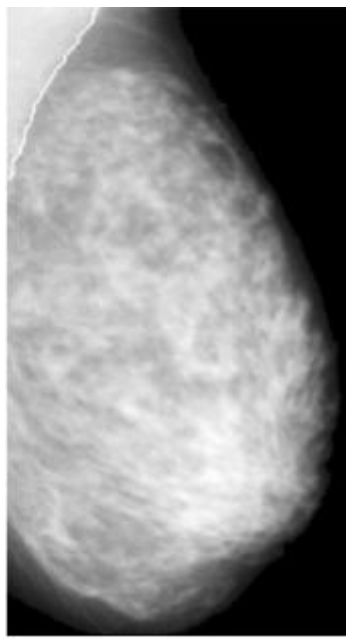
Figure 3.37 depicts the extracted breast region using method 2 for a mammogram collected from Lakeshore hospital, Kochi, India. The proposed method 2 is successful in delineating pectoral muscle boundary in cases where there are drastic intensity variations in pectoral muscle boundary.



(a)

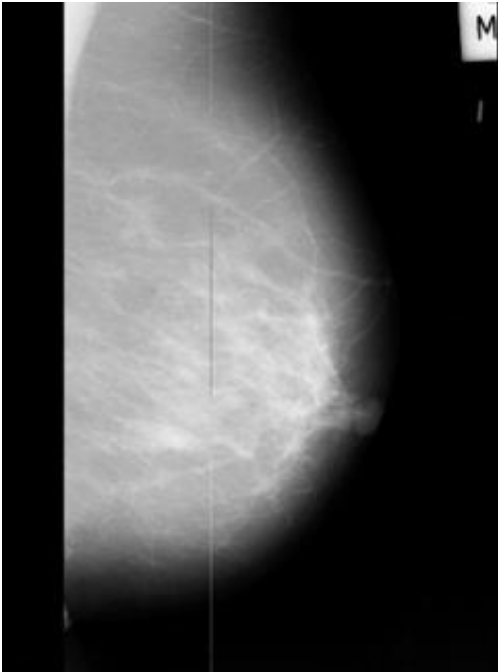


(b)

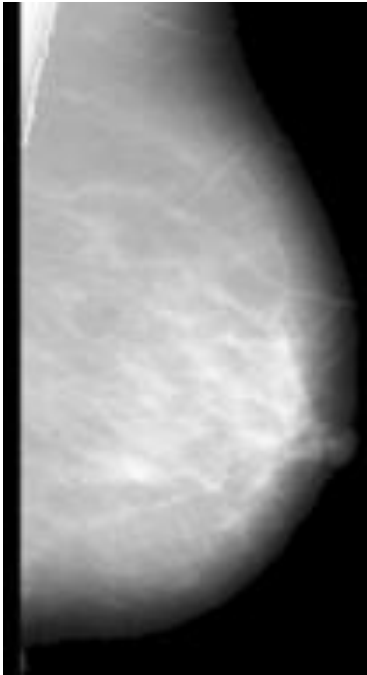


(c)

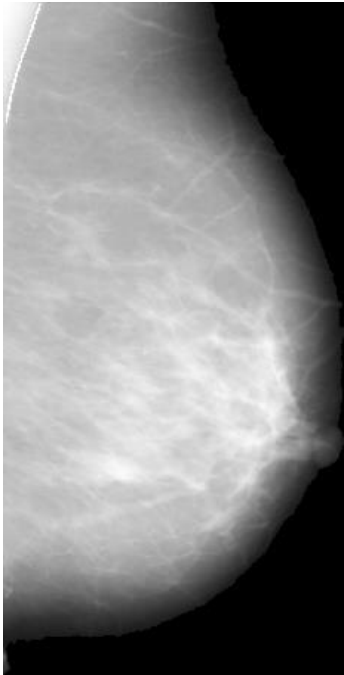
Figure 3.32: PM boundary of mdb130 (dense breast) drawn by (a) radiologist, (b) method 1, (c) method 2.



(a)

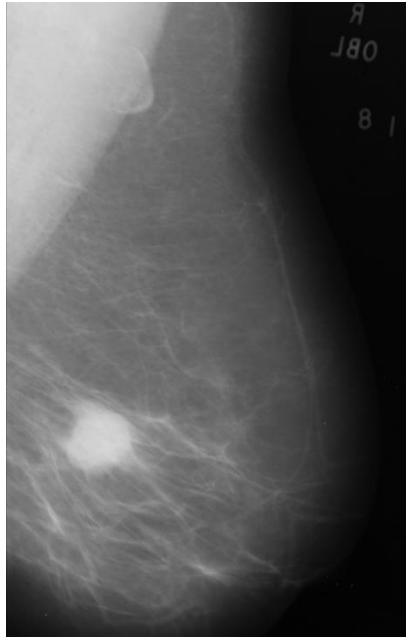


(b)

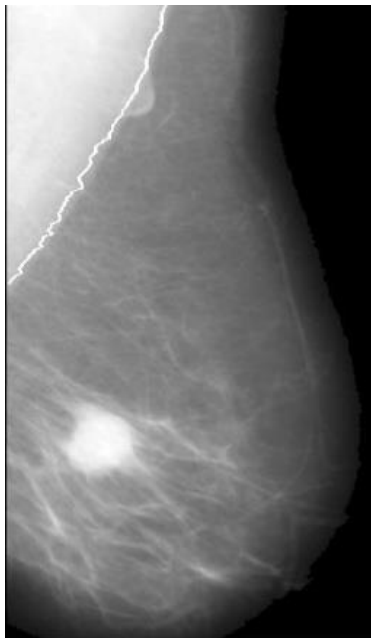


(c)

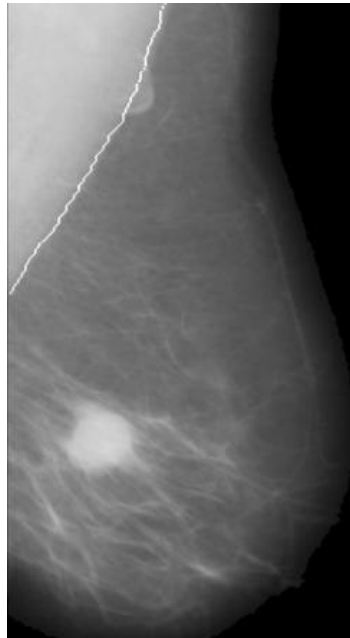
Figure 3.33: PM boundary of mdb013 (a) flipped image. PM boundary drawn



(a)

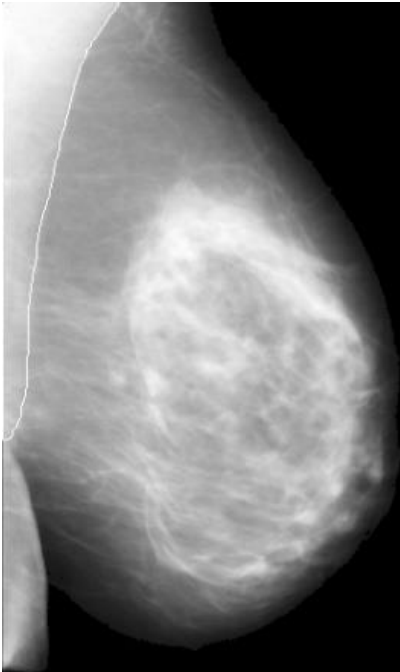


(b)

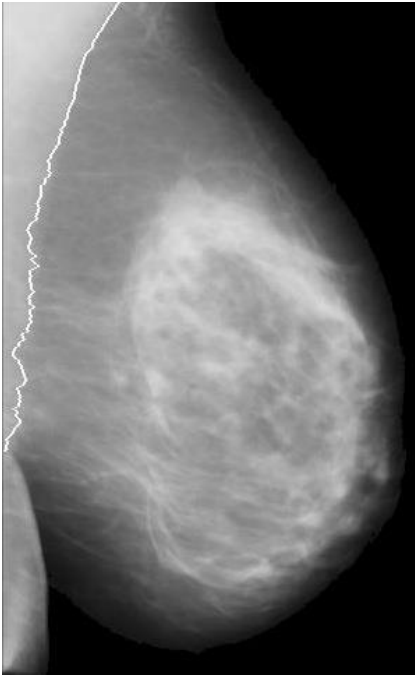


(c)

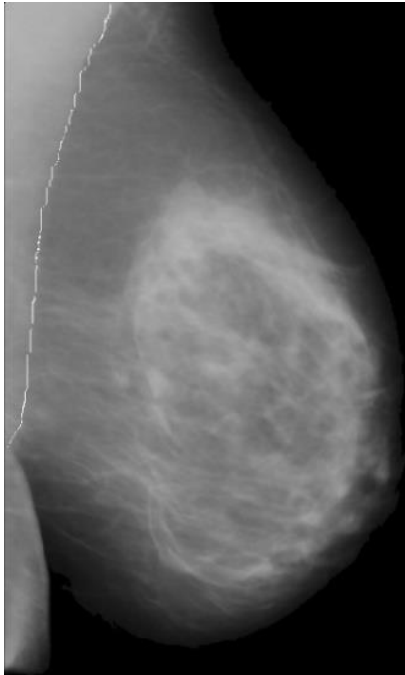
Figure 3.34: PM boundary of (a) mdb028 drawn by (b) method 1 and (c) method 2.



(a)



(b)



(c)

Figure 3.35: PM boundary marked on dense mammographic image, mdb112

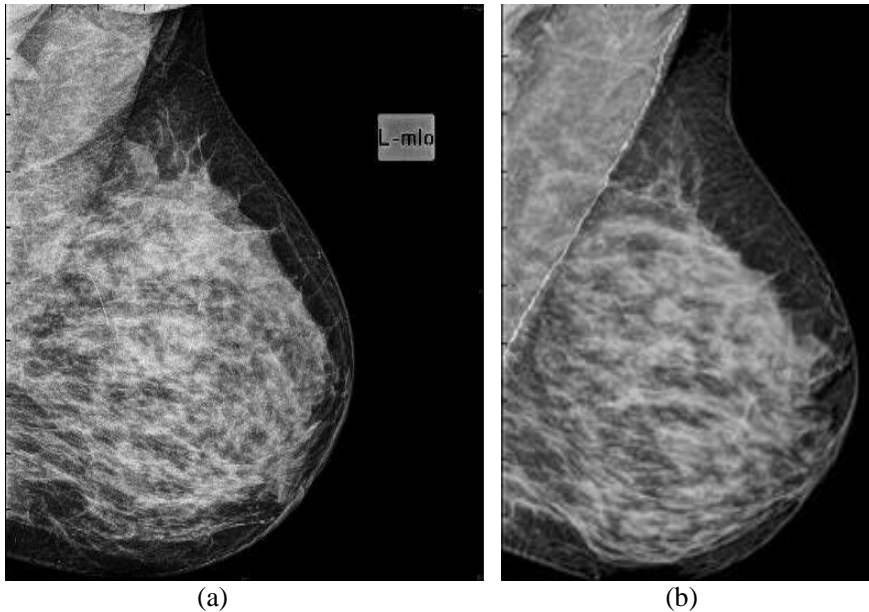


Figure 3.36: PM boundary marked on a mammographic image obtained from Lakeshore hospital. (a) Original, (b) boundary marked by proposed method 2.

3.6. Summary

In this chapter, two novel methods were developed for the extraction of PM boundary. The MIAS mammographic database and mammograms of a representative set of Indian population, provided by Lakeshore hospital, Kochi, India, were employed to validate the algorithm. In the first method, suspicious PM boundary segments were extracted using contours of homogeneous regions. The second method used edge detection on a multiscale decomposed mammogram for identifying the suspicious segments of PM. Efficiency of the proposed methods were validated using performance measures such as false positive average, false negative average, row normalized error and hausdorff distance. The performance measures of proposed methods were evaluated with four methods from current literature and were promising. The performance metrics obtained for the proposed method 2 are better than method 1. The results of method 2

indicate that the ground truth is closely met by the algorithm.

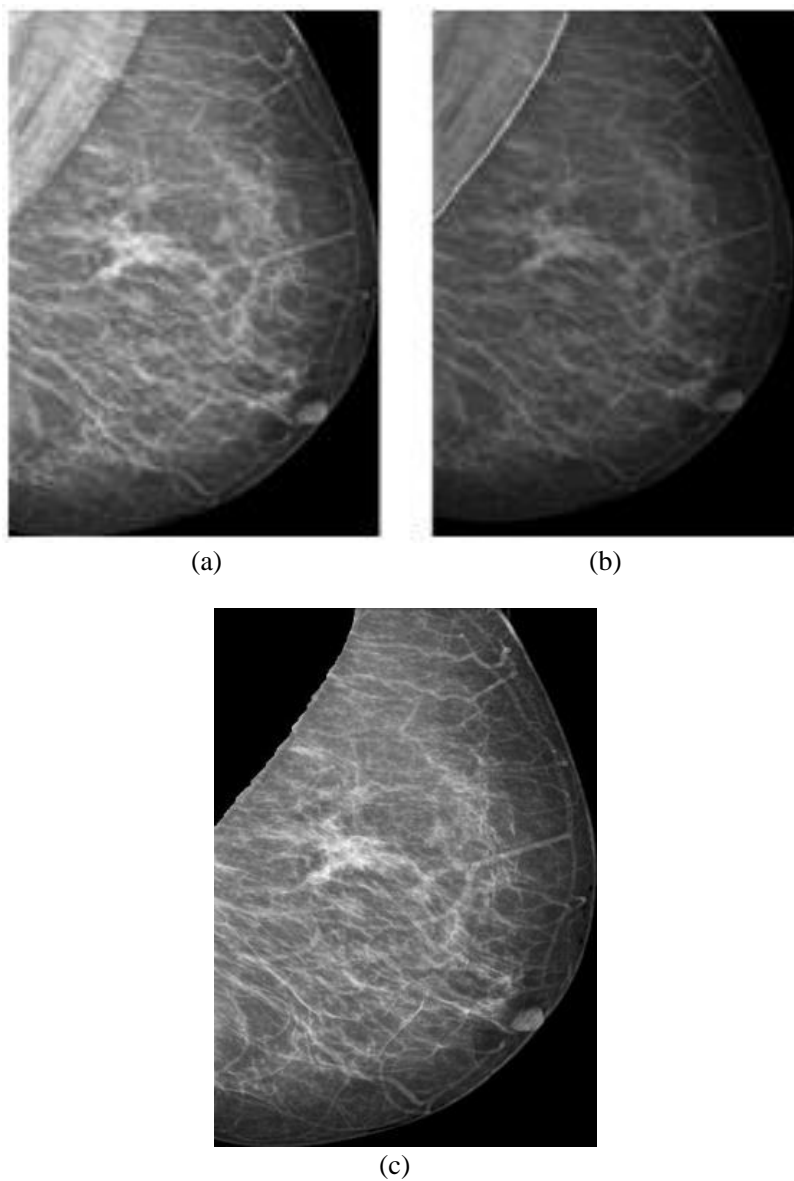


Figure 3.37: PM boundary of a mammogram obtained from Lakeshore hospital. (a) Original, (b) PM boundary marked using method 2 and (c) the extracted breast region using detected PM boundary.

3.7. Bibliography

Bajger, M., Ma, F., & Bottema, M. J. (2005). Minimum spanning trees and active contours for identification of the pectoral muscle in screening mammograms, In: *Digital Image Computing: Techniques and Applications conference*, 47–53.

Burt, P. J., Adelson, E. H. (1983). The Laplacian Pyramid as a Compact Image Code, *IEEE Transactions on Communications* , 31(4), 532-540.

Camillus, K. Santle., Govindan, V. K., & Sathidevi, P. S. (2011), Pectoral muscle identification in mammograms. *Journal of Applied Clinical Medical Physics*, 12(3).

Canny, J. F., (1986). A computational approach to edge detection, *IEEE Trans. On Pattern Anal. Machine Intell.*, PAMI-8(6), 679-698.

Chakraborty, J., S. Mukhopadhyay., V. Singla., N. Khandelwal., & P. Bhattacharyya. (2012). Automatic Detection of Pectoral Muscle Using Average Gradient and Shape Based Feature, *Journal of Digit Imaging*, 25(3), 387–399.

Chandrasekhar, R. & Y. Attikiouzel. (2001). Segmentation of the pectoral muscle edge on mammograms by tunable parametric edge detection, In *Advances in Signal Processing and Computer Technologies*, (G. Antoniou., N. Mastorakis., & O Panfilov, editors), 55-60, World Scientific and Engineering Society, WSES Press, Athens, Greece.

Chen, Z , B. R. Abidi., D. L. Page., & M. A. Abidi. (2006). Gray-level grouping (GLG): an automatic method for optimized image contrast enhancement--Part I: the basic method, *IEEE Trans Image Processing*, 15(8), 2290-302.

Chapter 3: Pectoral Muscle Delineation

Christina, Olsén. (2008). Towards Automatic Image Analysis for Computerised Mammography, PhD Thesis, Department of Computing Science Umeå University, Sweden.

Edward. J. Deneck. (2011). Barron's Regents Exams and Answers: Earth science-- the physical setting. Barron's Educational Series.

Ferrari, R. J., R. M. Rangayyan, J. E. L. Desautels, R. A. Borges., & A. F. Frère. (2004). Automatic Identification of the Pectoral Muscle in Mammograms, IEEE transactions on medical imaging, 23(2).

Fooladivanda, A., S. B. Shokouhi., M. R. Mosavi., & N. Ahmadinejad. (2014), Atlas-based automatic breast MRI segmentation using pectoral muscle and chest region model, Biomedical Engineering (ICBME), 21th Iranian Conference on, Tehran, 258-262.

Homer, M. J. (1997). Mammographic Interpretation: A Practical Approach, McGraw-Hill, Boston, MA.

Huttenlocher, D. P., G. A. Klanderman., & W. J. Rucklidge. (1993), Comparing images using the Hausdorff distance, IEEE Trans. Pattern Anal. Machine Intell., 15(9), 850–863.

Jaime, S. Cardoso., Ines, Domingues., Igor, Amaral., Ines, Moreira., Pedro, Passarinho., Joao, Santa, Comba., Ricardo, Correia., & Maria, J. Cardoso. (2010). Pectoral muscle detection in mammograms based on polar coordinates and the shortest path”, 32nd Annual International Conference of the IEEE EMBS Buenos Aires, Argentina.

Karssemeijer, N., (1998). Automated classification of parenchymal patterns in mammograms, Phys. Med. Biol., 43(2), 365–378.

Lakshmanan, R., M. S. Nair., M. Wilscy., & R. Tatavarti. (2008). Automatic Contrast Enhancement for Low Contrast Image: A comparison of Recent Histogram based Techniques, International Conference on Computer Science and Information Technology- ICCSIT, 269-276.

Lakshmanan, R., Shiji, T. P., Vinu, Thomas., Suma, M. J. & Thara, Pratab. (2014). Pectoral Muscle Boundary detection - A preprocessing method for early breast cancer detection, In: International Forum on Multimedia and Image Processing (IFMIP), World Automation Congress (WAC), published by IEEE, 258 – 263.

Lianghao, Han., John, H. Hipwell., Björn, Eiben., Dean, Barratt., Marc, Modat., Sebastien, Ourselin., & David, J. Hawkes. (2014). A Nonlinear Biomechanical Model Based Registration Method for Aligning Prone and Supine MR Breast Images, IEEE Transactions on Medical Imaging, 33(3), 682-694.

Lopez-Molinaa, C., B. De, Baetsb., H. Bustincea., J. Sanza., & E. Barrenecheaa. (2013). Multiscale edge detection based on Gaussian smoothing and edge tracking, 44, 101-111.

Ma, F., Bajger, M., Slavotinek, J. P., & Bottema, M. J. (2007). Two graph theory based methods for identifying the pectoral muscle in mammograms, Pattern Recognition, 40(9), 2592–2602.

Madasu, Hanmandlu., Devendra, Jha. (2006). An Optimal Fuzzy System for Color Image Enhancement, IEEE Transactions on Image Processing, 15(10), 2956-2966.

Masek, M., (2004). Hierarchical segmentation of mammograms based on pixel intensity, Ph.D. dissertation, Centre for Intelligent Information Processing Systems, School of Electrical, Electronic and Computer

Chapter 3: Pectoral Muscle Delineation

Engineering, The University of Western Australia.

Mayo, P., Rodenas, F., & Verdu, G. (2004). Comparing methods to denoise mammographic images, *Engineering in Medicine and Biology Society, IEMBS '04, 26th Annual International Conference of the IEEE*, 1, 247-250.

Oliver, A., X. Lladó., A. Torrent., & J. Martí. (2014). One-shot segmentation of breast, pectoral muscle, and background in digitised mammograms, *IEEE International Conference on Image Processing (ICIP), Paris*, 912-916.

Saunder, R. S. Jr., J. A. Baker., D. M. DeLong., J. P. Johnson., & E. Samei. (2007). Does image quality matter? Impact of resolution and noise on mammographic task performance, *Medical Physics*, 34(10), 3971-3981.

Smith, S., & J. Brady. (1997). SUSAN-a new approach to low level image processing, *International Journal of Computer Vision*, 23(1), 45-78.

Suckling, J., (1994). The Mammographic Image Analysis Society Digital Mammogram Database Excerpta Medica. *International Congress Series*, 1069, 375-378.

Sze, Man, Kwok., Ramachandran, Chandrasekhar., Yianni, Attikiouzel., & Mary, T. Rickard. (2004). Automatic Pectoral Muscle Segmentation on Mediolateral Oblique View Mammograms, *IEEE transactions on medical imaging*, 23(9).

Yam. M., Brady, M., Highnam, R., Behrenbruch, C., English, R., & Kita, Y. (2001). Three-dimensional reconstruction of microcalcification clusters from two mammographic views, *IEEE Trans Med Imaging*, 20(6), 479-489.

Chapter 4

Detection and classification of architectural distortion

Architectural distortion is the most commonly missed breast abnormality during visual examination of the mammogram and therefore its computer aided detection is challenging. A novel method for detecting architectural distortion is proposed in this chapter. The basis of the proposed method lies in the analysis of geometrical properties of abnormal patterns that correspond to architectural distortion in mammograms. Preprocessing methods are employed for the elimination of PM region from the mammogram. Regions that are candidates to contain centroids of architectural distortion are identified using a modification of the isotropic SUSAN filter. Edge features computed using phase congruency is thinned using gradient magnitude maximization. Three geometric properties namely eccentricity to retain near linear structures, perpendicular distance from each such structure to the centroid of the edges and quadrant support membership of these edge structures are used for finding features for classification. A feed-forward neural network, trained using a combination of backpropagation and a metaheuristic algorithm based on cuckoo search (CS)/ bat, is employed for classifying the suspicious regions as normal or malignant. The bat algorithm initialized classification yields a better sensitivity and specificity for the MIAS, DDSM and Lakeshore databases than CS initialized BPNN.

4.1. Introduction

Architectural distortion (AD), a breast cancer symptom, is observed as a distortion of normal structure of breast tissues and is said to be the most commonly missed abnormality [Knutzen, 1993]. Normal breast tissue structures represented as curvilinear structures appear to converge towards the nipple whereas the abnormal tissues related to AD are straight lines that appear to emanate from a central region with no definite mass [Hashimoto, 2011]. The point from which the linear radiating lines appear to originate is possibly located at the center of cancer [ACR, 1998]. According to the research conducted by Johnston [Johnston, 1991] on 308 biopsies of breast lesions, 67% were found to be AD and 80% of them were found to be invasive. 12 – 45% of breast carcinomas missed in mammography are found to be AD [Jasionowska, 2010]. The presence of subtle and uneven appearance of the abnormal structures corresponding to AD, normal breast parenchyma and dense tissues in mammogram make it difficult to identify the malignancy during screening [Sampat, 2005].

The fine nature of abnormal structures corresponding to AD leads to false negative findings on screening mammograms [Matsubara, 2004]. The presence of AD symptom in a single view is dangerous and should be treated seriously [Hashimoto, 2011]. AD is usually associated with other abnormalities such as mass and microcalcification [ACR, 1998]. In the studies conducted by Knutzen and Gisvold for determining the likelihood of malignant diseases in 859 cases of mammogram, 14 out of 59 architecturally distorted lesions were associated with calcifications [Knutzen, 1993]. AD is an important sign of malignancy as half to two-third of the cases are often observed with other symptoms of breast cancer [Matsubara, 2004]. Radiologists detect AD by analyzing the breast structure for finding the presence of radiating structures from a central region or distortion at the edge of parenchyma. A comparison of various breast cancer detection in mammograms using computer aided techniques shows

that more research is required in the field of AD detection than detection of microcalcification and mass [Baker, 2003]. The true positive detection of architecturally distorted abnormalities of breast reduces the mortality rate among women.

AD, the indirect sign of malignancy, is a difficult to detect symptom in mammographic screening for early breast cancer detection. Among 300 cases of breast cancer affected mammographic images, 20% is due to indirect sign of malignancy such as AD, asymmetry, single dilated duct and increased density [Sickles, 1986]. A clear definition on characterization of normal mammograms is difficult [Kopans, 1989], [Bassett, 1997]. The different mammographic structures are visible due to the difference in attenuation coefficient. The normal breast structures consisting of ducts, blood vessels, ligaments, fibroglandular tissues etc. converge towards nipple whereas the architecturally distorted abnormal patterns concentrate towards the region of distortion [Rangayyan, 2013]. The presence of overlapping normal breast tissues may appear as architecturally distorted structures which lead to false negative cases [Burrell, 1996]. According to Homer, [Homer, 1997], the identification of radiating breast structures corresponding to AD is more difficult than focally retracted structures. In dense mammogram, the fat trapped within fine spiculating structures surrounding the lesion appear as bright [Ellen, 2007] making the detection of AD features difficult.

According to Hashimoto [Hashimoto, 2011], the three major visible locations of AD are peripheral, central and subareolar positions of the breast region. Peripheral distortion retracts the edge of parenchyma. Central distortion at the center portion of fibroglandular tissue consists of numerous intersecting straight lines deviated from the direction towards nipple. Subareolar distortion is usually associated with nipple retraction. Central and subareolar distortions are more challenging than peripheral distortion. In dense breasts, the presence of breast parenchyma obstructs the visibility of abnormal lines representing AD. The subareolar region contains several

normal lines converging towards nipple. An isolation of abnormal structures in the subareolar region becomes difficult.

Literature review of computer aided algorithms for early breast cancer detection shows that more volume of work has been undertaken for identification of microcalcification and mass symptoms on mammograms compared to AD [Baker, 2003].

4.2. Literature review

Baeg [Baeg, 2002] achieved 90% sensitivity for detection of Architectural distortion using features based on a weighted gradient of the mammographic image. Training and testing of the method by back propagation neural network was conducted on 80 and 324 mammographic region respectively. Matsubara developed methods for AD detection for 17 cases around skin line and 38 cases within mammary gland [Matsubara, 2003]. The technique for focal retraction detection utilizes features such as the size and position of depressed thick mammary gland regions by binarization and top-hat morphological operations. The technique for detecting AD within mammary glands exploits features such as isotropy index on linear structures of mammary gland with a high value of concentration index. These methods yielded sensitivity of 94% for focal retraction and 84% for malignancy within mammary glands.

Ichikawa [Ichikawa, 2004] developed an automated computer aided method for detecting regions of AD. The mean curvature with positive sign is used to find the mammary glands. The concentration indices of linear structures extracted using mean curvature detect suspicious area of AD. A classification is performed using features such as size, mean pixel value, mean concentration index, mean isotropic index, contrast and features from the power spectrum. Experimental analysis on 94 mammographic images consisting of 41 architecturally distorted cases provided 76% accuracy and 80% sensitivity.

Chapter 4: Detection and Classification of Architectural Distortion

Guo [Guo, 2005] classified 40 mammographic ROI consisting of 21 normal regions of interest and 19 abnormal ROI with AD using support vector machine and radial basis function neural network. The experimental results show that support vector machine and radial basis function neural network provides a classification accuracy of 72.5% and 65% respectively. Matsubara suggested a method [Matsubaraa, 2005] to detect AD by employing shape index and curvedness for extracting structures corresponding to mammary glands in density varying regions of mammographic images. The direction of mammary gland structures situated in various radial areas originating from nipple is compared with the expected normal structures converging towards the nipple. The structures which are not aligned in the expected direction towards nipple are retained and others are rejected. A region with high value of concentration index is selected as suspicious.

Sampat performed detection method with 80% sensitivity on 45 architecturally distorted images in digital database for screening mammography (DDSM) [Sampat, 2005]. Linear structures representing AD are extracted by applying a filter on radon transform of the mammographic image. A threshold is applied to local peaks along the columns of transformed image to obtain linear structures in the image. Converging spiculations in the spatial domain are detected by applying a pair of radial spiculation filters on an inverse transformed image.

Rangayyan [Rangayyan, 2006] detected potential sites of AD by analyzing phase portrait maps of selected core curvilinear structures using a Gabor filter bank. Nonmaximal suppression technique is used to identify core curvilinear structure pixels. The method applied on 19 mammograms from MIAS database yielded 84% sensitivity. In an AD method proposed by Nakayama [Nakayama, 2008], the linear structures of mammogram were identified by analyzing eigen values of hessian matrix on three scales. Six objective features consisting of convergence indices and distributions of

Chapter 4: Detection and Classification of Architectural Distortion

linear structures in the direction towards nipple in three scales distinguished the normal and abnormal region of interest with a sensitivity of 71.3%.

Banik [Banik, 2009] obtained 80% sensitivity for detecting the possible sites of AD using Gabor filters, phase portrait modeling and classification. For each region, fractal dimension and haralick features were extracted. Out of these features, sum average, node value, difference variance, contrast, entropy, correlation, energy and fractal dimension were selected using the area under receiver operating characteristic (ROC) curves. Sequential forward selection and logistic regression were employed for the classification. Banik detected [Banik, 2011] probable regions of AD with 90% sensitivity using Gabor filters, phase portrait [Rangayyan, 2006] and features such as the angular spread of power and laws texture energy measures. The method was experimentally analyzed using 106 mammograms, with 4224 regions of interest including 301 regions of AD. Rangayyan [Rangayyan, 2013] developed a computer aided algorithm for AD with 80% sensitivity based on the divergence of orientation patterns in a region from the expected pattern of orientation. As a preprocessing operation, the breast region is delineated from the mammographic image by otsu's thresholding and morphological opening operation. The oriented patterns of breast region are obtained using a Gabor filter bank. The upper and lower breast boundaries are approximated using a polynomial curve. The deviation of breast tissue orientation from the expected orientation is utilized for detecting novel features such as index of divergence of spicules, radially weighted difference measure, angle-weighted difference measure, angle-weighted difference in the entropy of spicules etc. AD method proposed by Banik [Banik, 2013] based on Gabor filter, phase portrait analysis and classification using entropy based features achieved a sensitivity of 80%. The method was applied on 4,224 regions of interest with 301 malignant regions and the incorporated features included shannon's entropy, tsallis entropy and rényi entropy for the classification.

Chapter 4: Detection and Classification of Architectural Distortion

Zhiqiang [Zhiqiang, 2011] employed Gabor filter bank for the orientation estimation. Phase portrait template matching is performed at various scale for finding match with lesions of different size. Feature map is generated from phase portrait template at a particular scale for which a match with lesion is obtained. The maximum value and the entropy in the feature map indicating the likelihood of suspicious patterns are selected as features of AD. High values for both the features indicate that the region is architecturally distorted. A sensitivity of 93% is achieved for 500 abnormal and 1000 normal cases.

A method for detecting AD proposed by Amit [Amit, 2012] on the MIAS database achieved a sensitivity of 89%. In this the orientation field obtained from an average orientation is used in the formation of a directional filter. Suspicious regions of AD are extracted as a directionally different region by employing the filtered output of the directional filter. The regions with low variations of spectral density and low density of closed contours represent an architecturally distorted region.

Abhijith [Abhijith, 2013] proposed a method with 89.4% classification accuracy for differentiating the architecturally distorted regions and non architecturally distorted regions using control charts. ROI with size 128 by 128 are filtered using entropy to differentiate the distorted region from the normal. The linear patterns of filtered regions are extracted using Gabor filter. Shewhart control chart generates statistical measures and parameters for different orientations to generate a graphical plot which distinguishes architecturally distorted regions from normal regions. Otega obtained 87.5% sensitivity for detecting suspicious regions of Architectural distortion for 40 mammograms from DDSM [O'tega, 2014]. Features of CLAHE enhanced mammographic image were extracted using Gabor filter bank. Seven templates characterizing AD were created for finding a high value of cross correlation coefficient among Gabor filtered mammographic region and template. Such regions are considered suspicious areas of AD.

4.3. Detection and classification of architectural distortion

The proposed method for AD focuses on finding the suspicious regions by analyzing the geometrical properties of edge structures within the ROI. The suspicious regions are extracted using a modified SUSAN filtering technique [Smith 1995], [Smith, 1996] which is explained in chapter 3. Thinned edges are extracted using phase congruency [Kovesi, 1999] and gradient magnitude maximization [Canny, 1986]. Geometrical properties of these edges are then analyzed to extract features associated with these regions. A CS/ bat algorithm initialized back propagation algorithm is used for classifying regions into normal and abnormal. The flow chart of the proposed method is illustrated in figure 4.1.

4.3.1. Detection of suspicious regions

Most of the existing methods for detecting AD require processing of the entire mammogram for identification of abnormal region. In the proposed method, suspicious ROIs are selected from the entire breast region using a modified SUSAN filtering technique explained in section 3.3.2.1.1 of chapter 3.

Figure 4.2 shows SUSAN filter outputs for mammograms from MIAS, Lakeshore hospital and DDSM databases respectively. The homogeneous regions of images are grouped together based on the intensity similarities. AD features are typically radiating spicules that emanate from small central homogeneous regions [Hashimoto 2011] that are captured by the isotropic SUSAN mask. The extracted regions include the abnormal as well as normal regions.

The contours of homogeneous breast regions for the images in figure 4.2 are shown in figure 4.3. These contours are employed for detecting the radiating patterns that correspond to AD. However it is observed that the number of contours to be processed after their extraction

from the SUSAN filter output is large, which calls for employing a mask more tuned to the detection of AD features.

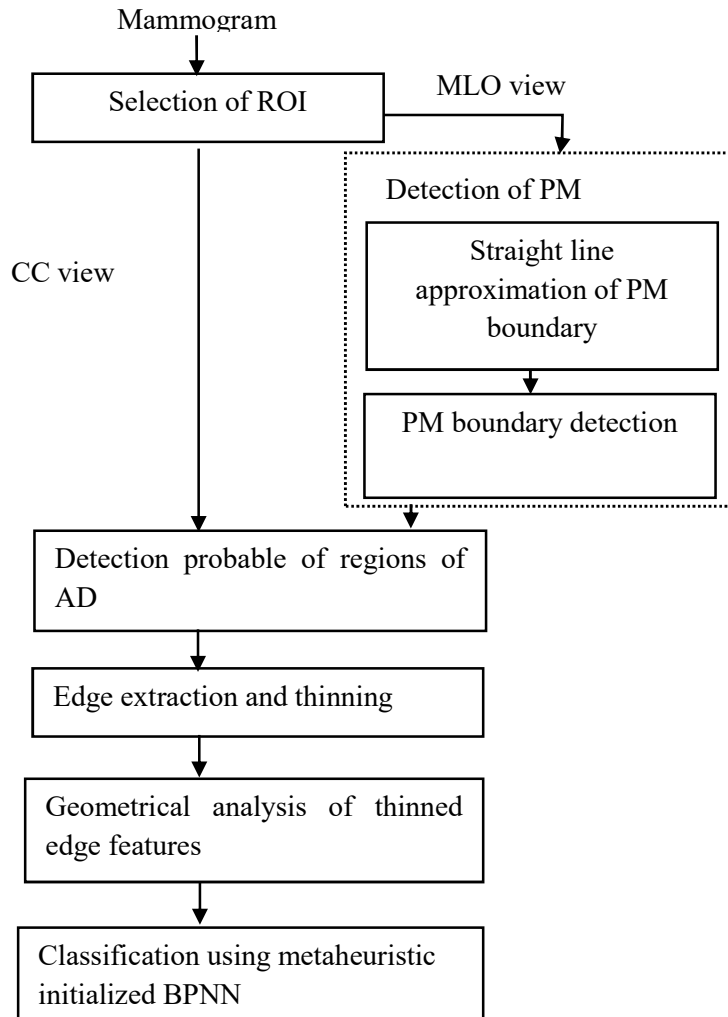


Figure 4.1: Flowchart of the proposed method for architectural distortion detection

4.3.1.1. Spatial mask modification

Various modifications of the spatial mask of SUSAN filter were performed to reduce the number of contours used for further processing, without missing out on any of the AD features. Figure 4.4 shows the different spatial masks that were tried out.

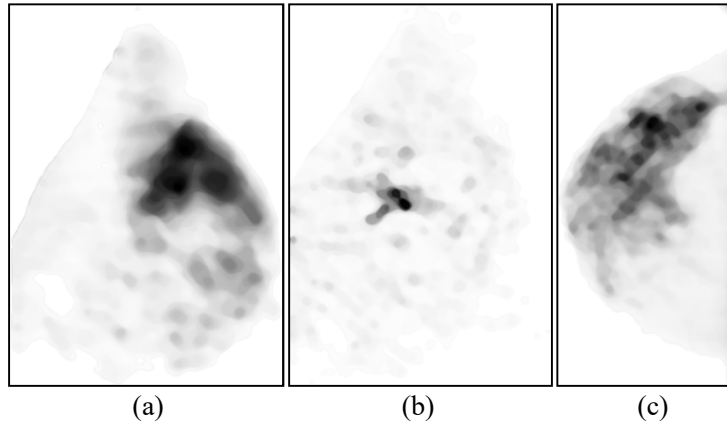


Figure 4.2: Images obtained after applying modified SUSAN filtering operation on mammograms from MIAS (mdb115), Lakeshore Hospital (LSH-RMLO-264412) and DDSM Database (1078-lcc).

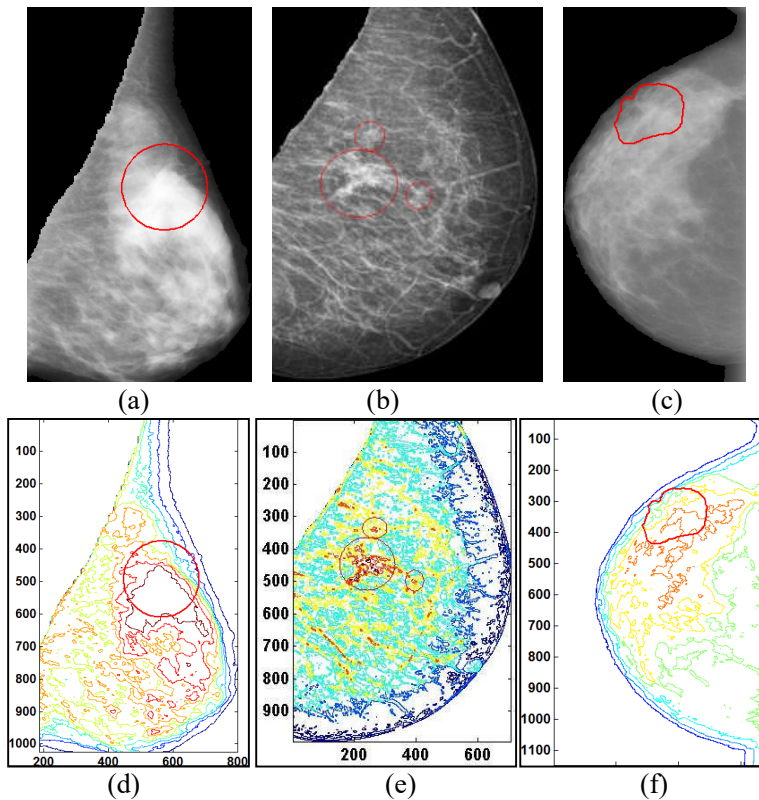


Figure 4.3: Breast region and contours of breast region with ground truth information. Breast region of (a) mdb115 from MIAS database, (b) mammogram from Lakeshore hospital, (c) 1078LCC from DDSM database. Contours of (d) mdb115, (e) mammogram from Lakeshore hospital, (f) 1078LCC

4.3.1.2. Contour extraction

Boundaries of region with iso-intensity are extracted by finding contours of breast image obtained using the modified SUSAN filters. Figure 4.5 shows the contours of mammograms from the three databases employed, on applying the spatial masks in figure 4.4.

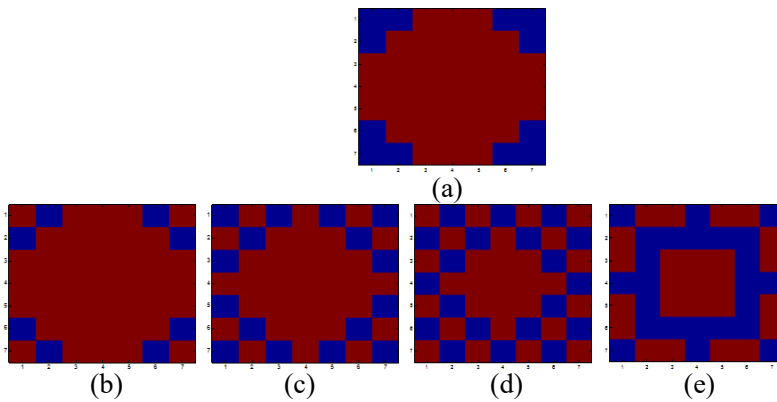
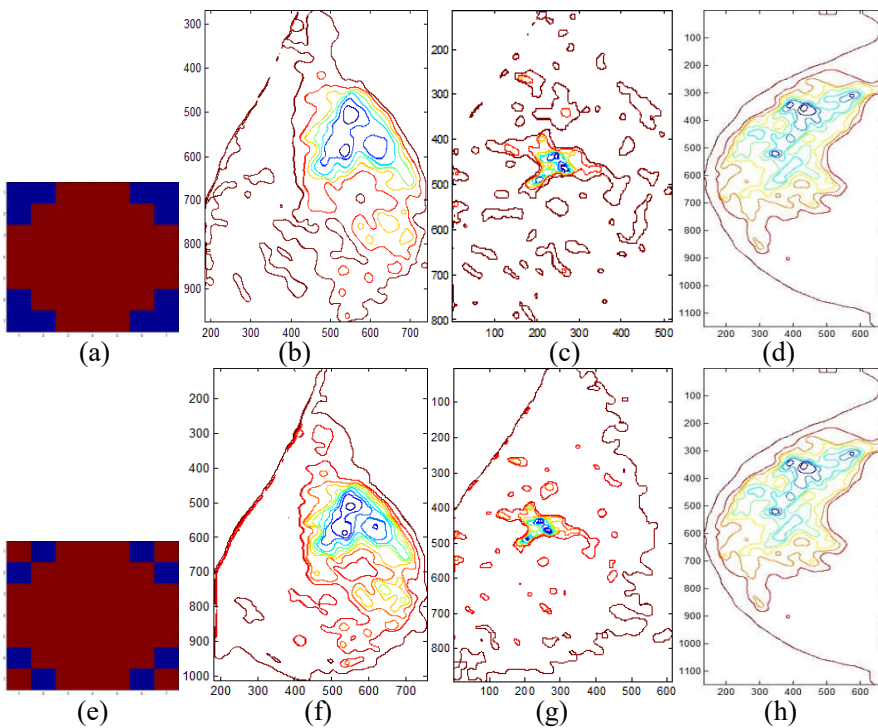


Figure 4.4: Different types of spatial mask for reducing the number of suspicious regions in breast region



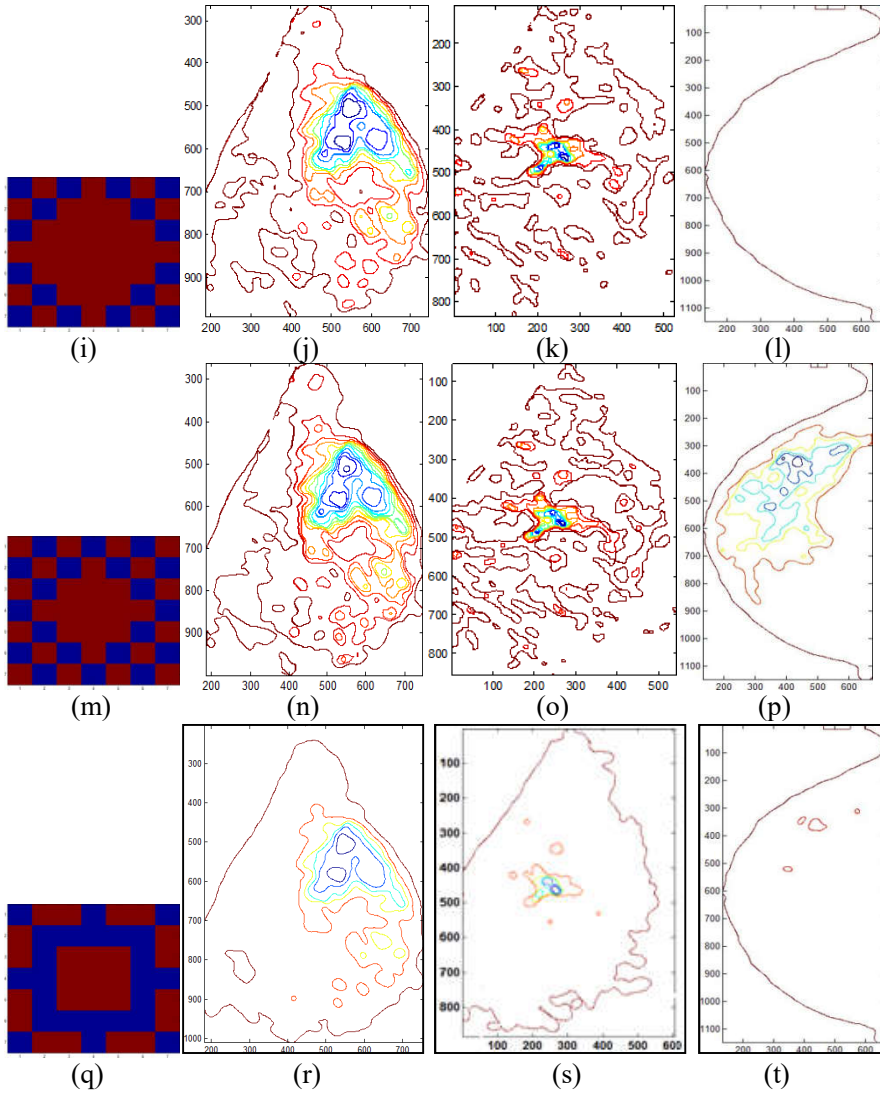


Figure 4.5: Spatial mask and resultant contours for mammograms as shown in 4.3(a), 4.3(b) and 4.3(c) from MIAS (mdb115), Lakeshore hospital (LSH-RMLO-264412) and DDSM Database (1078-lcc) respectively.

The spatial mask shown in figure 4.5(q) is found promising in finding regions with radiating patterns emanating from a central region. The radiating filter formed by modifying the SUSAN filter significantly reduces the number of processing regions without missing any of the origins of

abnormality. Figure 4.5(q) illustrates the effectiveness of the radiating mask in grouping the homogeneous regions of the mammogram, where the origins of the AD lie. When the radiating mask is employed, it is seen that the number of contours obtained (figure 4.5(r), 4.5(s), 4.5(t)) has been reduced significantly and the center of AD, indicated by the ground truth information is always seen to be included within one of the contours. It is observed that on an average, 32% computational savings is obtained with the radiating mask compared to the SUSAN mask.

4.3.1.3. Detection of contour centroid

Centroids of contours are computed from the contour coordinates [Bourke, 1998] as given below, to obtain the point of origin of AD.

$$C_x = \frac{1}{6A} \sum_{i=0}^{N-1} (x_i + x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \quad (4.1)$$

$$C_y = \frac{1}{6A} \sum_{i=0}^{N-1} (y_i + y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \quad (4.2)$$

$$A = \frac{1}{2} \sum_{i=0}^{N-1} (x_i y_{i+1} - x_{i+1} y_i) \quad (4.3)$$

Here x_i and y_i are the coordinates of the contour with i ranging from 0 to $N-1$, N being the number of vertices in the contour. C_x and C_y are the coordinates of contour centroid and A is the contour area. The red spots in figure 4.6(b) represent the contour centroids of figure 4.6(a).

4.3.1.4. Detection of suspicious regions of architectural distortion

The identified centroids obtained are used for selecting ROIs. According to Sampat et al, [Sampat, 2005] the average length of architecturally distorted spicules measured by the radiologists is 17.3 ± 9.7 mm. Hence ROIs centered on the identified centroid and large enough to include the largest spicule are considered. The size of this region depends on the resolution at which the images are digitized. Figure 4.7 shows breast regions with suspicious ROI of for mammograms from Lakeshore hospital,

MIAS and DDSM databases. Geometrical properties of these suspicious regions are analyzed further for abnormality identification.

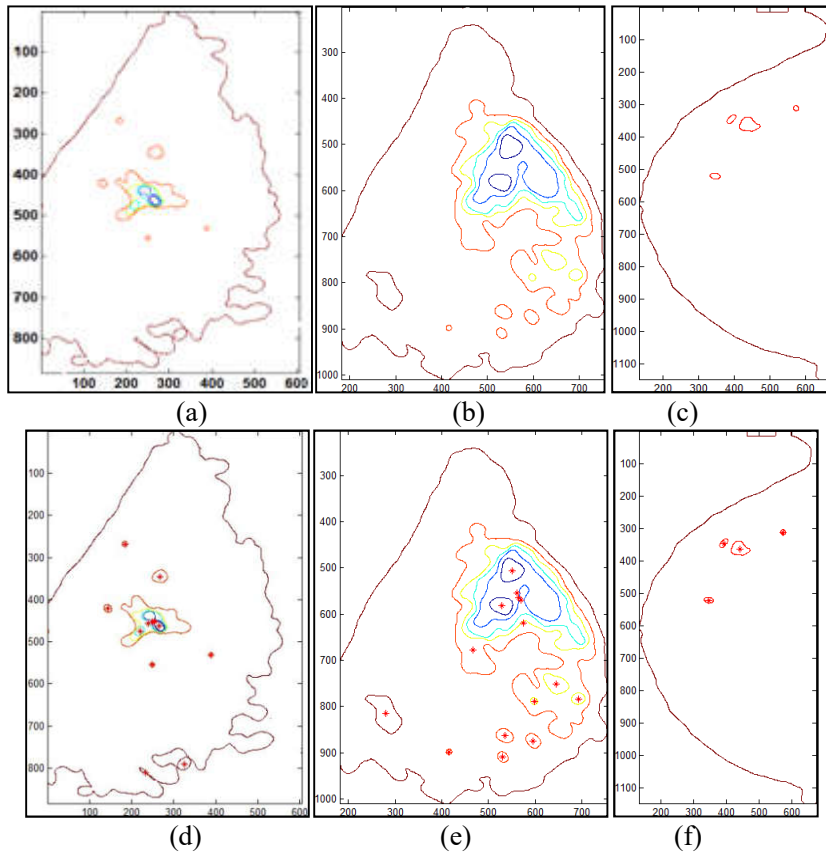


Figure 4.6: Contours and centroids (red spots) of suspicious homogeneous region using radiating filter mask. (a), (b) and (c) are the contours and (d), (e) and (f) are the corresponding centroids of mammograms of Lakeshore, MIAS and DDSM database as shown in 4.3(a), 4.3(b) and 4.3(c).

4.3.2 Edge feature extraction

Image features such as edge structures in a suspicious region of a mammogram can provide information regarding the presence of radiating patterns of AD. Though the gradient operators are good in edge feature extraction, there could be problems due to inaccurate localization and variations due to illumination and use of derivatives.

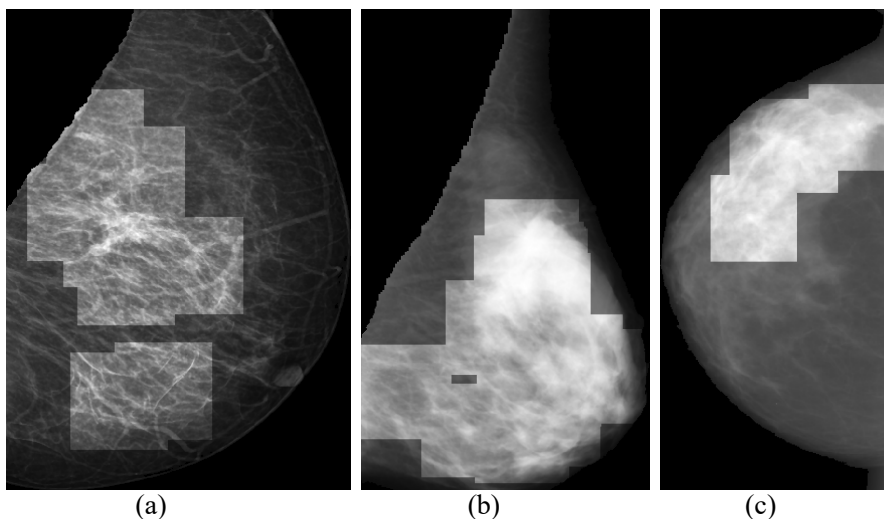


Figure 4.7: Suspicious regions around each centroid of respective contours shown in Figure 4.6.

In the proposed method, edge features are extracted using phase congruency [Kovesi, 1999], an energy based model employing log Gabor filter. Log Gabor functions are able to encode natural images more efficiently than ordinary Gabor functions, which would over represent the low frequency components and under-represent the high frequency components in any encoding. The advantage of log Gabor transform for radial filtering is that it is gaussian on a logarithmic scale and has zero DC components [Kovesi, 1999], [Zhiqiang, 2011], [David, 1987]. According to Morrone [Heeger, 1987], [Morrone, 1986], [Morrone, 1987], the features of an image are obtained at points where phase congruency is maximum. The phase congruency function is defined as

$$Phase\ congruency, PC(x) = \max_{\theta \in [0, 2\pi]} \frac{\sum_n A_n \cos(\phi_n(x) - \theta)}{\sum_n A_n} \quad (4.4)$$

where $\sum_n A_n \cos(\phi_n(x))$ represents the Fourier series expansion of a signal, A_n represents the amplitude of n^{th} Fourier component, $\phi_n(x)$ denotes the local phase of the Fourier component at position x and θ is the weighted phase angle of all the Fourier terms. The Taylors expansion of

$\cos(\theta) \approx 1 - \frac{x^2}{2}$ in equation (4.4) shows that the phase congruency is maximum when the weighted standard deviation of phase angles is minimum. In order to simplify the calculation of phase congruency, Venkatesh [Venkatesh, 1989] used peaks of local energy function for finding the points with maximum phase congruency, defined as

$$\text{Local energy function, } E(x) = \sqrt{F^2(x) + H^2(x)} \quad (4.5)$$

where $H(x)$ is the Hilbert transform of $F(x)$. The phase congruency function which is directly proportional to energy function is given in equation (4.6). From equation (4.6), it is clear that the peak value of phase congruency corresponds to the peak value of energy.

$$E(x) = PC(x) \sum_n A_n \quad (4.6)$$

Figure 4.8 shows the polar diagram of Fourier components at a point in a signal. The association among phase congruency, energy and the sum of Fourier amplitudes is illustrated in figure 4.8(b). Figure 4.8(b) shows the Fourier components plotted head to tail as complex vectors. $F(x)$ is the original signal with removed DC component and $H(x)$, its Hilbert transform representing the summed vectors to the real and imaginary axis respectively. The magnitude of the summed vector is depicted as $E(x)$. From figure 4.8, it is clear that

$$E(x) = \sum_n A_n \cos(\phi_n(x) - \theta) \quad (4.7)$$

From equations (4.6) and (4.7), phase congruency, PC is defined as

$$PC(x) = \max_{\theta \in [0, 2\pi]} \frac{\sum_n A_n \cos(\phi_n(x) - \theta)}{\sum_n A_n} = \frac{E(x)}{\sum_n A_n} \quad (4.8)$$

Instead of using Hilbert transform, local energy can be calculated by convolving the signal with a pair of filters in quadrature, i.e., the convolution is performed by odd and even symmetric filters with zero mean and same sum of squares value [Kovesi, 1999].

The transfer function of log Gabor filter is given in equation (4.9):

$$lg(x) = e^{-\frac{(\log(\omega/\omega_0))^2}{2} - \frac{(\log(k/\omega_0))^2}{2}} \quad (4.9)$$

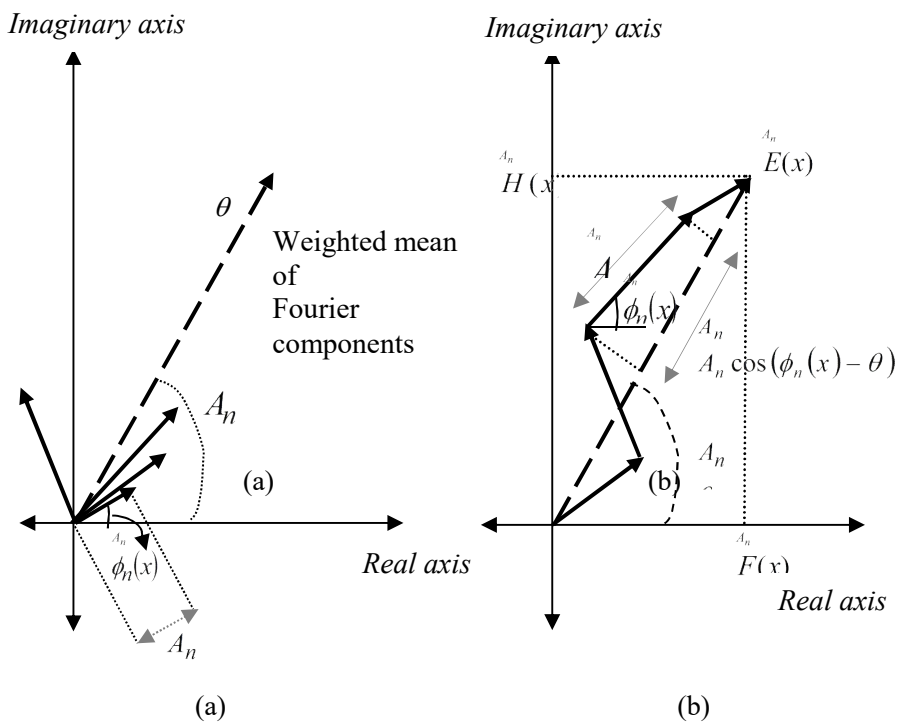


Figure 4.8: Polar diagram of the (a) components of a Fourier series at a point in a signal [33], (b) Fourier components plotted head to tail.

where ω_0 is the center frequency of filter. In order to maintain a constant shape ratio, (k/ω_0) is kept constant for different values of ω_0 . For a scale n , let M_n^e and M_n^o represent even symmetric (cosine) and odd symmetric (sine) wavelets for a signal I . The response vector obtained for this

quadrature filter pairs are given in equation (4.10).

$$[e_n(x), o_n(x)] = [I(x) * M_n^e, I(x) * M_n^o] \quad (4.10)$$

Amplitude of the transform at a wavelet scale, n is given in equation (4.11).

$$A_n(x) = e_n(x)^2 + o_n(x)^2 \quad (4.11)$$

Phase of the transform at a wavelet scale, n is given in equation (4.12).

$$\phi_n(x) = a \tan 2(e_n(x), o_n(x)) \quad (4.12)$$

The plotting of response vectors for each scale is similar to that of Fourier components represented in figure 4.8. Hence $F(x)$ is formed by projecting the summed response of the even filters to the axis corresponding to the output of even symmetric filter and $H(x)$ is obtained by projecting the summed response of odd filters to the axis corresponding to the output of odd symmetric filter

$$F(x) \cong \sum_n e_n(x) \quad (4.13)$$

$$H(x) \cong \sum_n o_n(x) \quad (4.14)$$

$$A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2} \quad (4.15)$$

$$E(x) = \sqrt{F(x)^2 + H(x)^2} \quad (4.16)$$

A group of log Gabor wavelets at N different scales and M orientations were used to compute phase congruency in two dimensions. A set of filters are generated by multiplying radial log Gabor filters and angular filters which filter the mammographic image. Log Gabor filters are calculated over N scales and angular filters are computed over M orientations. Thereafter the filtered image is converted back to spatial domain. For a particular orientation, a composite image is generated by

adding the filtered images obtained over the scales. A final edge image is formed combining all such composite images, by retaining pixel values from the composite image for orientations having the maximum magnitude. Figure 4.9 show the extracted edge features of mammographic images from Lakeshore hospital, MIAS and DDSM database respectively where a structural distortion is clearly visible in the abnormal regions.

4.3.3. Thinning

The suspicious regions of AD include normal structures such as blood vessels, ligaments, ducts, fibroglandular tissues etc. along with abnormal structures [Rangayyan, 2006]. The distributions of mammary gland are approximated to linear structures [Matsubara, 2004]. A subsequent thinning operation using gradient magnitude maximization [Canny, 1986] generates the strong edge structures. The obtained edge points are categorized into strong, weak and insignificant based on a double threshold value. The insignificant edge points are not included in the final set of edges. The strong edge points as well as weak edge points connected to strong edge points are included in the final edge set.

Figure 4.10 shows strong edges of mammograms from MIAS, DDSM and Lakeshore hospital database after applying the thinning operation on edge features. The geometrical analysis of extracted strong edge structures differentiates the abnormal patterns from normal.

4.3.4. Geometrical analysis of edge structures

A suspicious region contains strong thinned edges associated with normal and abnormal structures. Geometrical analyses of these strong edge structures reduce the number of these normal structures.

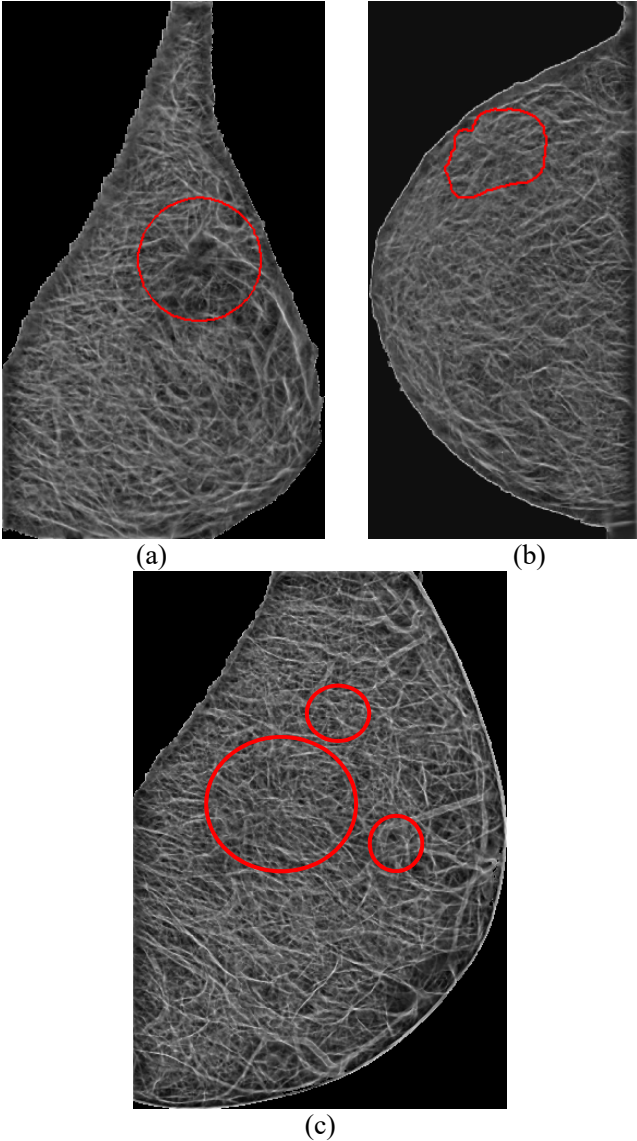


Figure 4.9: Edge features of mammograms from (a) MIAS, (b) Lakeshore hospital and (c) DDSM database obtained using phase congruency. Red circles indicates ground truth information

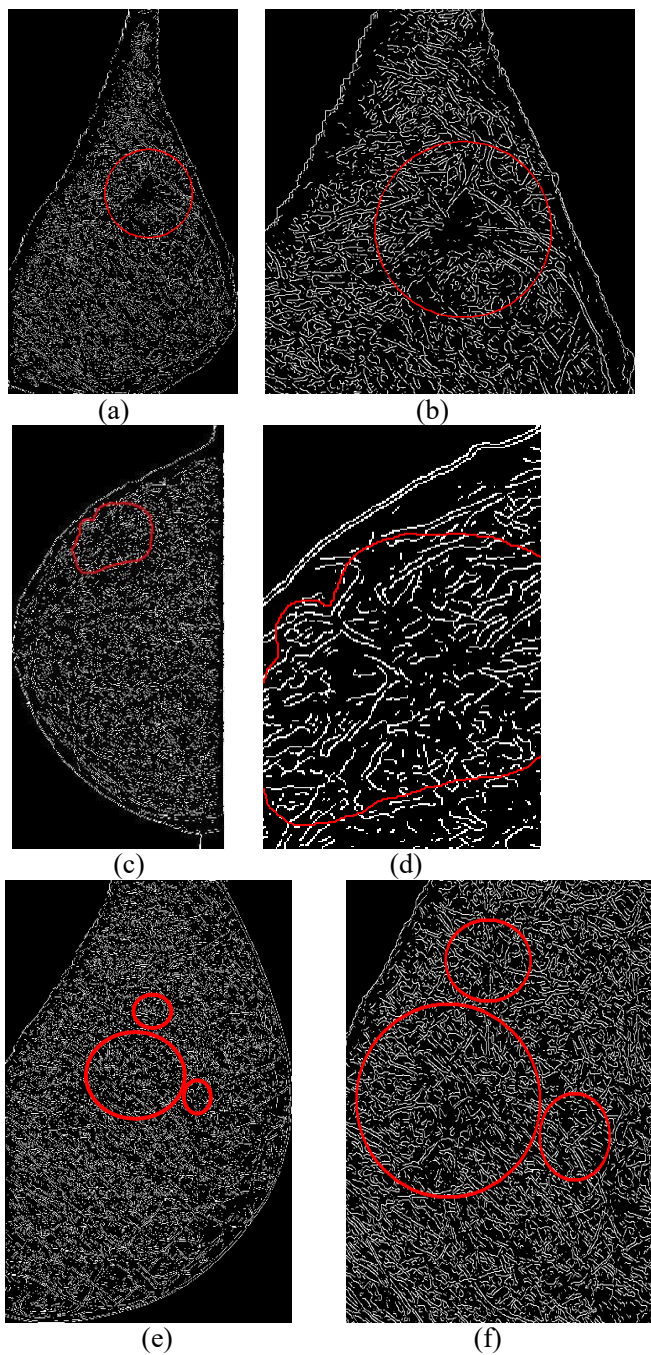


Figure 4.10: Strong edge structures: (a) Strong edge structures of breast region, (b) zoomed region of (a) in MIAS database. (c) Strong edge structures of breast region, (d) zoomed region of (c) in DDSM database. (e) Strong edge structures of breast region, (f) zoomed region of (e) in Lakeshore hospital. Red circles indicate ground truth information provided by the radiologists.

The geometrical properties incorporated in the proposed work are

1. Eccentricity, to retain near linear structures.
2. Perpendicular distance of thinned edge structures from the centroid.
3. Quadrant support membership of edge structures.

Features for classification are generated from these properties. Architecturally distorted structures are approximately linear [Rangayyan, 2006]. Only those structures having eccentricity near to one are retained [Math, 2009], [Deneck, 2011]. Figure 4.11 shows linear structures in a normal and abnormal region of mdb115 from MIAS database.

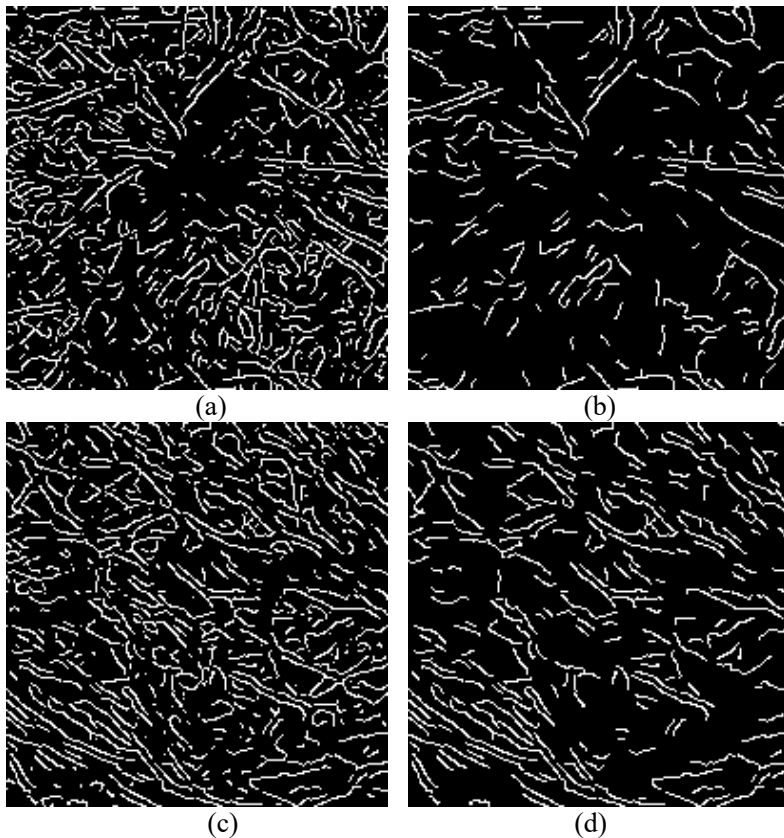


Figure 4.11: Linear structures: (a) and (c) shows the thinned structures of a malignant and normal region in mdb115 from MIAS database; (b) and (d) represents the corresponding linear structures.

The abnormal structures converge towards the central portion of AD [Hashimoto, 2011]. Hence in the method proposed, only the linear components converging towards the centroid of the region are retained. Each component is extended towards the boundary of the ROI. The components with shortest (perpendicular) distance from the centroid within a threshold (here taken as 4 mm) are retained.

Figure 4.12 shows the results of converging linear structures for a normal and abnormal suspicious region on mammogram, mdb115 from MIAS database.

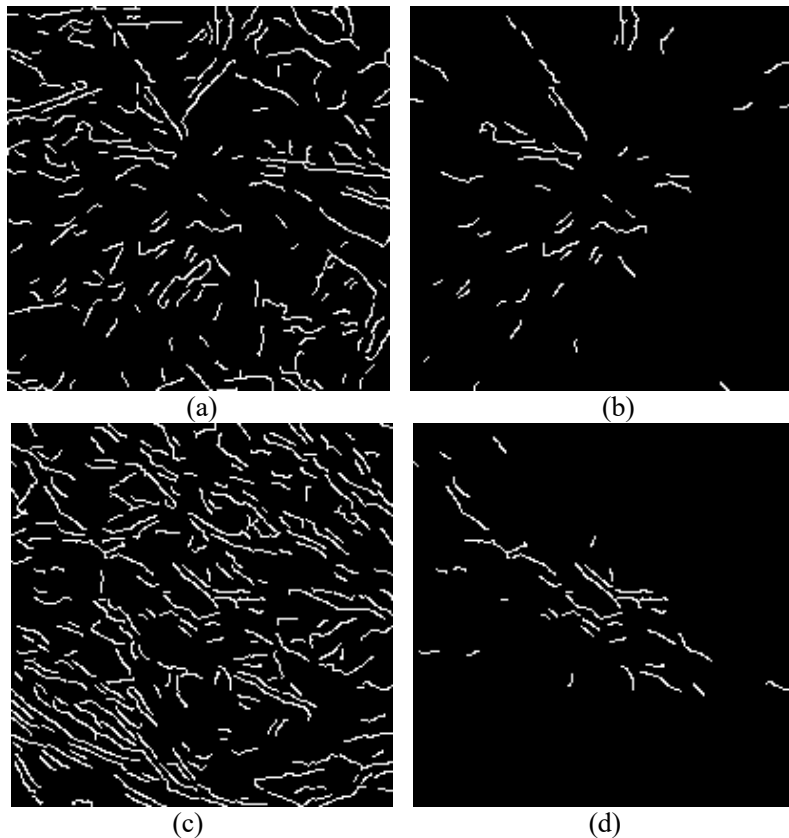


Figure 4.12: Converging lines. (a) Linear structures of a malignant region, (b) converging structures of (a), (c) linear structures of a normal region and (d) converging structures of (c)

Usually, the normal breast tissue structures converge towards nipple whereas abnormal structures appear to emanate from a central originating region [Hashimoto, 2011]. Here, the retained converging components may include linear structures oriented towards nipple. In order to remove normal lines extended through and beyond the identified centroid of suspicious region than ending there as in the case AD, two criteria based on quadrant division through the centroid is used.

1. Quadrant crossing
2. Quadrant slope

The quadrant division of the ROI employs the centroid as the origin. If an edge structure exists across two quadrants, the possibility of it emanating from a region centered on the centroid of the converging structures is minimal, and such edge structures are removed. Figure 4.13 shows the results of normal structure removal for normal and abnormal suspicious ROIs in mdb115.

In addition, only those structures satisfying the corresponding quadrant slopes are retained as abnormal structures. For example, all lines emanating from the centroid, in the first quadrant, will have orientations in the range $0-90^0$, in the second quadrant, orientations in the range $91-180^0$ and so on. Figure 4.14 shows the results after removing such structures.

Figure 4.15 shows the structures so obtained for abnormal and normal ROIs for images from the MIAS database. Figure 4.16 shows similar results for images from the DDSM database.

An initial set of thirteen features are extracted from the geometrical properties of these detected abnormal structures. The features employed are represented in table 4.1. This initial set is pruned in the next step and used to train a feed forward neural network [Rumelhart, 1986] to classify the suspected regions into sites of AD or normal regions.

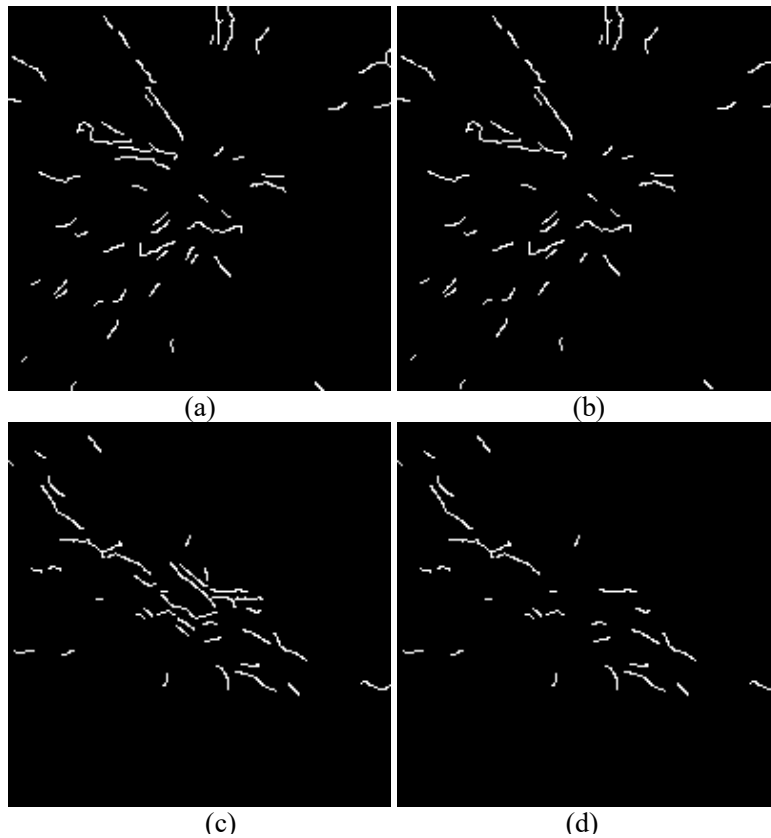


Figure 4.13: Quadrant membership criteria 1. (a) Concentrating structures of a malignant ROI, (b) structures satisfying quadrant membership criteria 1 of (a), (c) concentrating structures of a normal ROI, (d) structures satisfying quadrant membership criteria 1 of (c) for mdb115 of MIAS database.

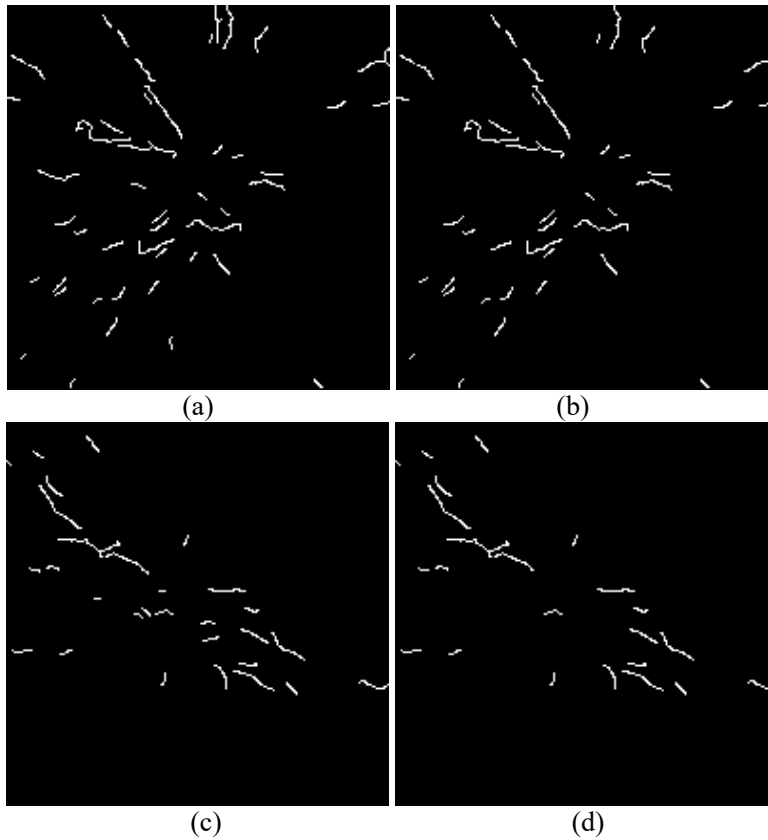


Figure 4.14: Quadrant membership criteria 2. (a) Concentrating structures of a malignant ROI, (b) structures satisfying quadrant membership criteria 2 of (a), (c) concentrating structures of a normal ROI, (d) structures satisfying quadrant membership criteria 2 of (c) for mdb115 of MIAS database.

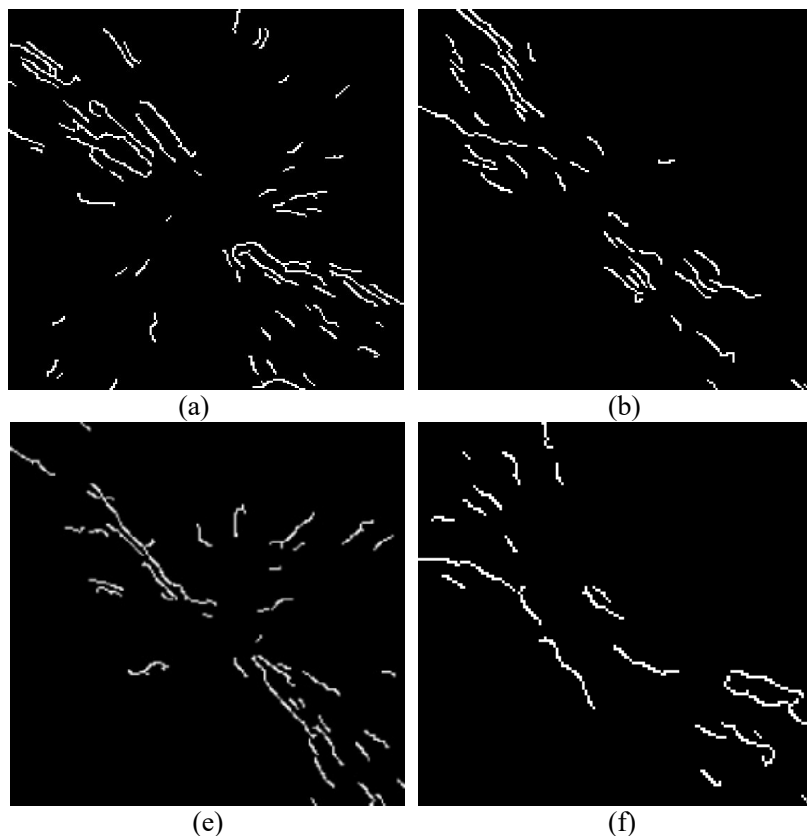


Figure 4.15: Abnormal structures in malignant and normal regions of mammogram collected from MIAS database. Retained normal patterns in (a) malignant region of mdb155, (b) normal region of mdb155, (c) malignant region of mdb158 and (d) normal region of mdb158

4.3.5. Classification

A classification technique using features is essential to classify various regions of mammograms as normal or abnormal. The final step of the proposed method is to perform classification of suspicious regions of AD using various features extracted from the results of geometrical analysis of edge structures and backpropagation neural network (BPNN) [Rumelhart, 1986]. A BPNN with initial weights optimized by the CS algorithm [Xin-She Yang, 2014], [Jiao-hong, 2014] or bat algorithm [Yang,

2013] is used to classify the suspicious regions as normal and abnormal based on the selected features. The initial set of features is cross validated to generate an optimized set of features. The number of neurons in the hidden layers are optimized using the area under curve (AUC) [Hajian-Tilaki, 2013].

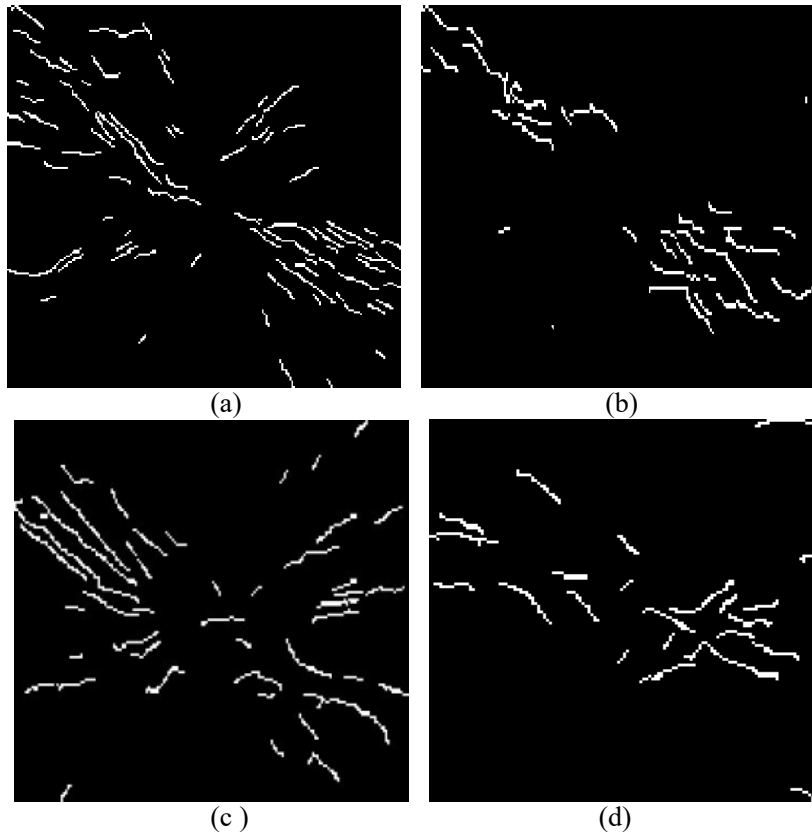


Figure 4.16: Abnormal Structures: (a) Abnormal structures in mammograms from DDSM database. Retained patterns in (a) malignant region of 3058_LCC, (b) normal region of 3058_LCC, (c) malignant region of 3375_RCC and normal region of 3375_RCC

4.3.5.1. Backpropagation neural network for classification

Artificial neural networks (ANN) are biologically inspired classification algorithm consisting of input, output and hidden layers of neurons [Rumelhart, 1986]. Feed forward neural networks with BPNN is one of the popular category of ANN. The general structure of a BPNN network as shown in Figure 4.17 consists of at least three layers of neuron with first layer as input layer, second as hidden layer and third as output layer.

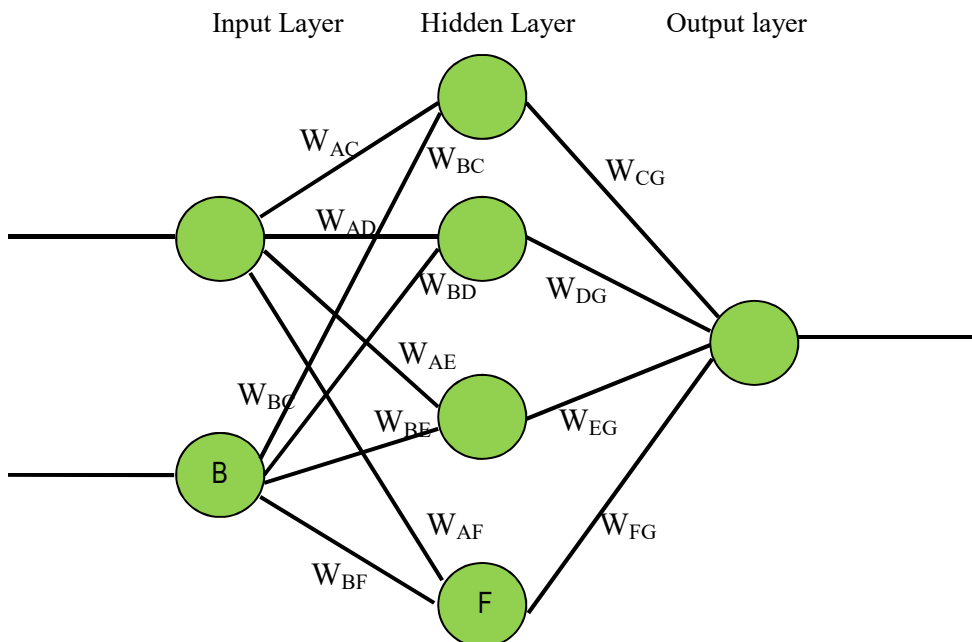


Figure 4.17: Architecture of a feed forward back propagation neural network

Each node in the network is fully connected to all nodes in the next layer from input to output layer. Input patterns are given at the input layer. The number of nodes in the input layer is selected as the number of input patterns. A three layer feed forward back propagation neural network [Rumelhart, 1986] with one hidden layer is used in the proposed work for classification [Rumelhart, 1986], [Zhen-Guo, 2011].

Training and testing are the two major phases of a neural network classification technique. The main goal of training is to find best sets of weights and bias values for BPNN using gradient-descent search. Testing is to classify the given unknown input to various output classes. Learning process in the training stage is an iterative step. A supervised error based learning technique is used in BPNN for training purpose. Random values are used as the initial connecting weights. The inputs to a neuron multiplied by the corresponding weights are summed up to evaluate the activation function [Bekir, 2011] to generate output of each node. The output generated is actual output. The error is calculated on the basis of difference in desired output, $DOut_G$ and actual output, $AOut_G$. The performance of the network model is improved by adjusting the weights on the basis of back propagated error. In the proposed work, sigmoid function is used as the activation function. Error at node, G at output layer is calculated as shown in equation (4.17),

$$E_G = AOut_G (1-AOut_G) (DOut_G - AOut_G) \quad (4.17)$$

where $AOut_G (1-AOut_G)$ is the derivative of the activation function. The error of the output node, G is used to update the weight connecting nodes of hidden layer and output node. Let W_{old} be the old weight and W_{new} be the new weight connecting node C of hidden layer and G of output layer. The updated weight, W_{new} is calculated as given in equation (4.18)

$$W_{new_CG} = W_{new_CG} + \alpha W_{old_CG} + (\eta E_G AOut_C) \quad (4.18)$$

As the desired output at hidden layer, $DOut_C$ is not available, the error of node, C in hidden layer is obtained by backpropagating the error, E_G at node G of output layer as given in equation (4.19).

$$E_C = AOut_C(1- AOut_C)(E_G W_{old_CG}) \quad (4.19)$$

Weight of the hidden layer node is updated as

$$W_{new_AC} = W_{old_AC} + \alpha W_{old_AC} + \eta E_C AOut_A \quad (4.20)$$

The learning process is repeated until the error is within a prescribed tolerance. In the proposed work, the iterative operation is performed until the training error decreases to 0.001. A large number of neurons in hidden layer results in slow convergence whereas less number reduces the reliability. Too fast or slow learning rate is not beneficial for the convergence of the neural network.

The proposed method employed a learning rate, η of 0.8 which is in between 0.1 and 1, suggested by Li Min et al., [LiMin, 1994]. Momentum coefficient to avoid the error fluctuations in convergence was selected as 0.7, which is advised to be in between 0.01 and 1 [Kumar, 2005]. Since large number of nodes in hidden layer reduces the converging speed and small number of nodes reduces reliability, we have initially selected the number of nodes in hidden layer as 13, same as that of the input.

4.5.2. BPNN with metaheuristic optimization

Metaheuristic approaches [Osman, 1996] were employed to initialize the BPNN with optimized weights to classify the suspicious ROIs as normal or abnormal. The major drawbacks of conventional BPNN algorithm are slow convergence speed, problems of local minima and the random choice of initial weights and bias. A proper initialization of weights and bias yields an ideal BPNN with high performance. If the selection of the weights and bias is improper, convergence of the algorithm will be very slow and the training process would take a long time. More importantly, the solutions could get trapped in local minima for problems with multiple solutions.

Metaheuristic algorithms are stochastic optimization algorithms that work on randomization [Christian, 2003]. The optimized solution can be either a maximum or minimum depending upon the type of problem.

Diversification/ exploration and intensification/ exploitation are the two major factors of metaheuristic algorithms [Christian, 2003]. Intensification and diversification generates the solution on a local and global scale respectively. Diversification avoids local minima problem. The metaheuristic algorithms employed in this work to optimize the initial weights and bias of the BPNN are the CS and bat algorithm. This approach improves the sensitivity and specificity of the classification of the suspicious regions by the BPNN.

4.3.5.2.1. Cuckoo search algorithm

CS algorithm is a bio-inspired metaheuristic algorithm based on the breeding of cunning cuckoo birds [Yang, 2013]. Flight movement of many animals and insects is recognized as a random walk. Levy Flight is a random walk that is characterized by a series of straight jumps chosen from a heavy-tailed probability density function which is used to model the flight movement of birds and insects. In statistical terms, it is a stochastic algorithm for global optimization that finds a global minimum. Levy flight concept is used in CS Algorithm [Nazri, 2013]. Figure 4.18 shows the Levy flight distribution. Levy flights used in CS algorithm utilize a balanced combination of local and global random walk controlled by a probability $p_a \in [0,1]$.

The implementation of CS algorithm is based on three major assumptions as follows

1. Each cuckoo lays single egg at a time in a randomly chosen host nest.
2. The high quality eggs that survive best are permitted to the next generation.
3. The number of host nest is fixed in advance. The host bird identifies the intruding egg with a probability, $p_a \in [0, 1]$. On recognizing intrusion, the host birds dispossess the parasitic egg or give up the nest for a new one.

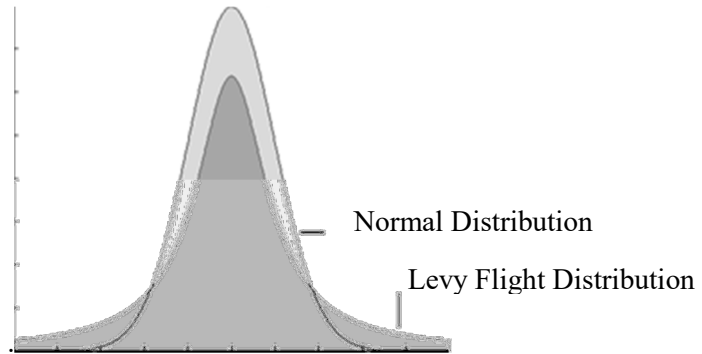


Figure 4.18: Levy flight distribution [Yan, 2009].

The pseudocode for CS algorithm is described in Appendix 2. CS algorithm starts with a set of possible solutions. Based on the fitness function, a randomly selected poor solution (egg in a nest) is replaced by a new and potentially better (cuckoo- egg) solution. Each egg in the nest represents a solution whereas the new solution is represented by cuckoo egg. Levy flight is performed for generating new solutions. In the proposed method, each nest is assumed to occupy a single egg. The new solution is obtained by implementing the Levy flight as in equation (4.21).

$$x_i^{t+1} = x_i^t \oplus \alpha Levy(s, \lambda) \quad (4.21)$$

where x_i^{t+1} is the new solution for i^{th} cuckoo, α is the step size scaling factor and s is the step length. The Levy flight provides a random walk whose random step length is drawn from a Levy distribution as per equation (4.22). The step length, s needed for Levy flight distribution is given in equation (4.23)

$$Levy \sim s^{-\lambda}, \quad (1 < \lambda \leq 3) \quad (4.22)$$

$$s = \frac{u}{|v|^{\frac{1}{\beta}}} \quad (4.23)$$

where u and v are obtained from the normal distribution as per equations (4.24), (4.25), (4.26) and (4.27).

$$u = N(0, \sigma_u^2) \quad (4.24)$$

$$v = N(0, \sigma_v^2) \quad (4.25)$$

$$\sigma_u = \left\{ \frac{\Gamma(1 + \beta) \sin(\pi\beta / 2)}{\Gamma[(1 + \beta) / 2] \beta 2^{(\beta-1)/2}} \right\}^{\frac{1}{\beta}} \quad (4.26)$$

$$\sigma_v = 1 \quad (4.27)$$

4.3.5.2.2. Bat algorithm

Bat algorithm is a bio-inspired metaheuristic algorithm proposed by Xin-She Yang [Yang, 2013]. Bat algorithm makes use of the echolocation behavior of microbats. Bats emit very large and loud sound pulse in order to search/detect prey even in darkness. The sound pulse helps them to avoid obstacles, identify the type, distance and speed of the prey. Bats increase the loudness of pulse with a rate of 10 - 20 pulse/ sec while searching the prey. The pulse remains for around 8 - 10 ms with 25 khz – 150 khz frequency and 2 mm - 14 mm wavelength. On reaching near to the prey, the loudness is decreased whereas the pulse rate is increased to a rate of 200 pulse/ sec. This algorithm is derived using echolocation behavior of bats with varying rate of pulse emission and loudness [Yang, 2013].

The three major assumptions used in bat algorithm are as below:

1. All bats use echolocation to sense distance, and they also 'know' the difference between food/ prey and background barriers in some manner.
2. Bats fly randomly with velocity v_i at position x_i with a frequency f_{\min} , varying wavelength λ and loudness A_0 to search for prey. They can automatically adjust the wavelength (or frequency) of their emitted pulses and adjust the rate of pulse emission $r \in [0, 1]$, depending on the proximity of their target.

3. Although the loudness can vary in many ways, we assume that the loudness varies from a large (positive) A_0 to a minimum constant value A_{\min} .

The pseudocode for bat algorithm is described in Appendix 2. For a random bat fly, an adjustment in frequency parameter followed by an update of velocity and position is performed according to the following equations (4.28), (4.29) and (4.30).

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta \quad (4.28)$$

$$v_i^{t+1} = v_i^t + (x_i^t - x^*)f_i \quad (4.29)$$

$$x_i^{t+1} = x_i^t + v_i^t \quad (4.30)$$

where $\beta \in [0,1]$ is the random vector drawn from a uniform distribution. f_i is the adjusted frequency. Here f_{\min} and f_{\max} are the minimum and maximum frequency, x^* is the current best global position. The new solution is generated as given in equation (4.31).

$$x_{new} = x_{old} + \epsilon A^t \quad (4.31)$$

where $\epsilon \in [-1,1]$ and A^t is the average loudness of the time step t .

The loudness and pulse emission rate are updated at each iteration as in equations (4.32) and (4.33)

$$A_i^{t+1} = \alpha A_i^t \quad (4.32)$$

$$r_i^{t+1} = r_i^0 [1 - \exp(-\gamma)] \quad (4.33)$$

where α and γ are constants. The initial values of A and r are chosen randomly depending on the problem.

4.3.5.2.3. Cuckoo search/ Bat BPNN classification

BPNN with slow speed of convergence and local minima problem are optimized using metaheuristic optimization algorithm such as CS and bat algorithm. The optimized weights obtained using metaheuristics

algorithm is utilized for improving the efficiency of BPNN. The pseudocode for a CS/ bat initialized BPNN classification is given in appendix 2.

4.3.6. Feature selection

Feature selection provides a smaller but more distinguishing subset compared to the starting data, selecting the distinguishing features from a set of features and eliminating the irrelevant ones. The main goal of feature selection is to reduce the dimension of the data by finding a small set of important features that can give good classification performance. This results in both reduced processing time and increased classification accuracy [Uzer, 2013]. Feature selection requires a search strategy, sequential backward selection (SBS), to select candidate subsets and an objective function to evaluate these candidates. Here, k -fold cross-validation method is utilized in order to increase the reliability in the phase of classification during the selection of features. In k -fold cross-validation, the original sample is partitioned into k sub-samples randomly. Of the k sub-samples, a single sub-sample is retained as the validation data for testing the model, and the remaining $k-1$ sub-samples are used as training data. The cross-validation process is then repeated k times (the folds), with each of the k sub-samples used exactly once as the validation data. The average of k results from the folds gives the test accuracy of the algorithm.

In this work, a set of thirteen features are extracted from the geometrical properties of the detected abnormal structures. These properties are represented in table 4.1. The performance of the method depends upon the features selected for classification [Rangayyan, 2013]. A set of optimized features are selected from the thirteen features using k -fold cross validation [Math, 2009]. Feature selection methods reduce the redundant set of features from the initial set thereby reducing the complexity of the neural network architecture [Girish, 2014]. A large number of features increases

the storage needs of the network. Unreliable and inaccurate classification and difficulty in generalization [Wang, 2008], [Nandi, 2006] are some of the disadvantages with increased number of features.

Sequential backward selection is employed as the search strategy in the proposed method [Fukunaga, 1990]. Initially all features are included to check the classification performance. Elimination of a single feature is done until there is performance degradation. In order to evaluate the performance of the classifier by estimating the error rate, a 10 fold cross validation is performed. In 10 fold cross validation; training samples are randomly separated into 10 groups. For each trial, a single group is considered as a test set and the remaining 9 groups are used for training. The procedure is repeated with different test sets and the average estimate of misclassification rate. The optimal features selected from Table 1 are 1, 2, 4, 8, 9, 10, 12 and 13.

The possibility of over fitting is reduced by also optimizing the number of hidden layer neurons in the feed forward NN used for classification [Chen, 2006]. The initial set of thirteen features was optimally reduced to eight as discussed earlier, thereby reducing the number of input neurons from thirteen to eight. In order to reduce the number of hidden layer neurons, a backward approach [Ghana, 2013] is selected.

Initially, the number of neurons in the hidden layer is selected by a rule of thumb [Saurabh, 2012]. According to this rule, the number of neurons in the hidden layer can be (1) in between the size of input and output layer or (2) the sum of size of output layer and two-third of the size of the input layer. Initially the number of hidden neurons is selected as six, as per the second rule of thumb. Now as per the first rule, the size of the hidden layer may go up to 8. Among the three possibilities for the number of hidden layer neurons (6, 7 or 8) as per the above rule of thumb, the optimal number is selected by finding the Area Under Curve (AUC) of the ROC (Receiver Operating Characteristics) curves [Hajian, 2013] which are computed for all the three cases.

Table 4.1. Classification features of Architectural Distortion

1. Total number of linear structures after applying eccentricity criteria in suspicious ROI ($TotL$)
2. Total number of converging lines : ($TotC$)
3. Total number of structures satisfying quadrant criteria1 ($TotQ1$)
4. Total number of structures satisfying quadrant criteria2 ($TotQ2$)
5. Normalized converging value, $NorC$ is obtained as the ratio of total number of converging lines to the total number of linear structures ($NorC = TotC / TotL$)
6. Normalized quadrant criteria1, $NorQ1$ is obtained as the ratio of total number of structures satisfying quadrant criteria1 to the total number of linear structures ($NorQ1 = TotQ1 / TotL$)
7. Normalized quadrant criteria2, $NorQ2$ is obtained as the ratio of total number of structures satisfying quadrant criteria2 to the total number of linear structures ($NorQ2 = TotQ2 / TotL$)
8. Normalized quadrant criteria2-converging, $NorQ2C$ obtained as the ratio of total number of structures satisfying quadrant criteria2 to the total number of converging lines ($NorQ2C = TotQ2 / TotC$)
9. Normalized stepwise quadrant criteria1, $StQ1$ is obtained as the total number of structures satisfying quadrant criteria1 to the total number of converging lines ($StQ1 = TotQ1 / TotC$)
10. Normalized stepwise quadrant criteria2, $StQ2$ is obtained as the total number of structures satisfying quadrant criteria2 to the total number of structures satisfying quadrant criteria1 ($StQ2 = TotQ2 / TotQ1$)
11. Standard deviation of total lines, $SDTot$ is obtained as the standard deviation among total number of linear structures, total number of converging lines, total number of structures satisfying quadrant criteria1 and total number of structures satisfying quadrant criteria2, ($SDTot = SD(TotL, TotC, TotQ1, TotQ2)$)
12. Stepwise Standard deviation, $SDSt$ is obtained as the standard deviation among Normalized stepwise quadrant criteria1 and Normalized stepwise quadrant criteria2 ($SDSt = SD(StQ1, StQ2)$)
13. Standard deviation of normalized value, $SDNor$ is obtained as the standard deviation among Normalized converging value, Normalized quadrant criteria1 and Normalized quadrant criteria2. ($SDNor = SD(NorC, NorQ1, NorQ2)$)

It is observed that the AUC is maximum for the case of seven neurons in hidden layer and thus the size of the feed forward NN was fixed as having eight input neurons, seven hidden layer neurons in a single hidden layer and

one neuron in the output layer. Figure 4.19 depicts the training error for seven neurons in hidden layer.

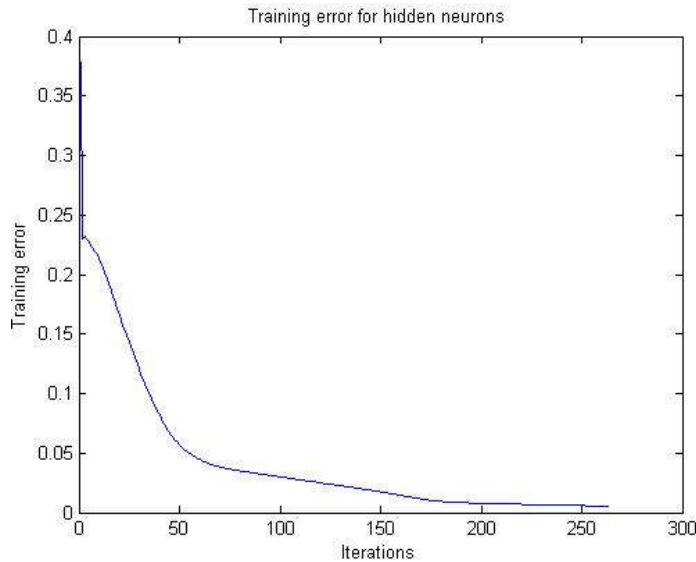


Figure 4.19: Training error for seven hidden neurons

4.4. Experimental setup

The experimental analyses were carried out on 60 images from MIAS database [Suckling, 1994], 100 images from DDSM databases [Paola, 2013], and 100 mammographic images from Lakeshore hospital, Kochi, India [Lakeshore hospital]. The results of mammograms obtained from Lakeshore hospital, India were verified by the senior radiologists, Dr. Suma M Jacob (M.D) and Dr. Thara Pratab (M.D) of Lakeshore hospital Kochi, India, who were associated with this research. For the MIAS and DDSM databases, the ground truth information provided was used for validation.

The Lakeshore database that was created for the purpose of this research had digital mammograms (acquired using the Siemens MAMMOMAT 3000 NOVA digital mammograph), whereas both the MIAS and DDSM databases have been created by scanning x-ray films.

4.5. Results and discussion

Detection of AD in a dense breast is a challenging task for the radiologists. Fat getting trapped in between radiating structures obscures the visibility of abnormalities [Ellen, 2007]. The proposed method focuses on distorted architecture and radiating structures for detecting AD patterns in dense breast. Figure 4.19 and figure 4.20 show the various steps involved in detecting the features from the ROIs (square regions within red and blue square border) that are detected by the radiating mask in a dense mammogram, (LSH_LCC_262437) collected from Lakeshore hospital, India. The ROI in red square border shown in figure 4.19 is positive for AD, whereas that in figure 4.20 is normal. The structures that are retained after the geometric analysis of the thinned edge features are seen to be distributed in different directions for the architecturally distorted region. The red circle shown in Figure 4.19(a) and 4.20(a) indicates the ground truth information provided by the radiologists.

Figure 4.21 and 4.22 represents various steps involved with malignant and normal ROI of a mammogram (LSH_RMLO_237263) with abnormality at subareolar region. A focal retraction indicating malignancy is visible near the nipple. The right cranio caudal (RCC) view of above mammogram in figure 4.21 and 4.22 is depicted in figure 4.23 and 4.24. Figure 4.25 and 4.26 represents suspicious ROIs of a mammogram with focal retraction. Figure 4.27 shows the RCC view of mammogram (LSH_RMLO_1209) with focal retraction.

Classification is performed for 60 images from MIAS database containing 101 ROIs (11 malignant and 90 normal ROIs), 100 images from DDSM database containing 599 ROIs (78 malignant and 521 normal ROIs) and 100 images collected from Lakeshore hospital database containing 454 ROIs (30 malignant and 424 normal ROIs). A total of 28 ROIs (11 malignant and 17 normal ROIs), 35 ROIs (13 malignant and 22 normal ROIs) and 30 ROIs (5 malignant and 25 normal ROIs) were used for

training Lakeshore, MIAS and DDSM database respectively. Specificity, sensitivity and accuracy were computed to evaluate the proposed algorithms and for comparison with other methods (these parameters are explained in chapter 2). When the conventional BPNN was employed for the classification with the initial weights randomly initialized, the specificity and sensitivity values computed were 72% & 70%, 69% & 68% and 76% & 75% respectively for the MIAS, DDSM and Lakeshore databases. These values are unacceptably low. However when the initial weights were found using the CS/ bat algorithms, there were significant improvements in the sensitivity and specificity values. The ROC curves for MIAS, DDSM and Lakeshore database using CS/ bat initialized BPNN are shown in figure 4.28 and 4.29 respectively.

Figure 4.31 shows the example of an erroneous classification for the case of a malignant mammogram (LSH_LMLO_262437) collected from Lakeshore hospital which was recognized as normal.

The ROI extracted using modified SUSAN filter includes the region of malignancy. The classification step failed in identifying this dense mammogram as abnormal. The highly dense tissues present in the breast region obstruct the visibility of abnormal patterns which leads to misclassification. Even though the abnormal patterns obtained after geometrical analysis is distributed in different direction, the number of retained abnormal patterns are found less. However it is also to be observed that for the corresponding CC view shown in figure 4.25, the proposed method was successful in identifying the abnormality.

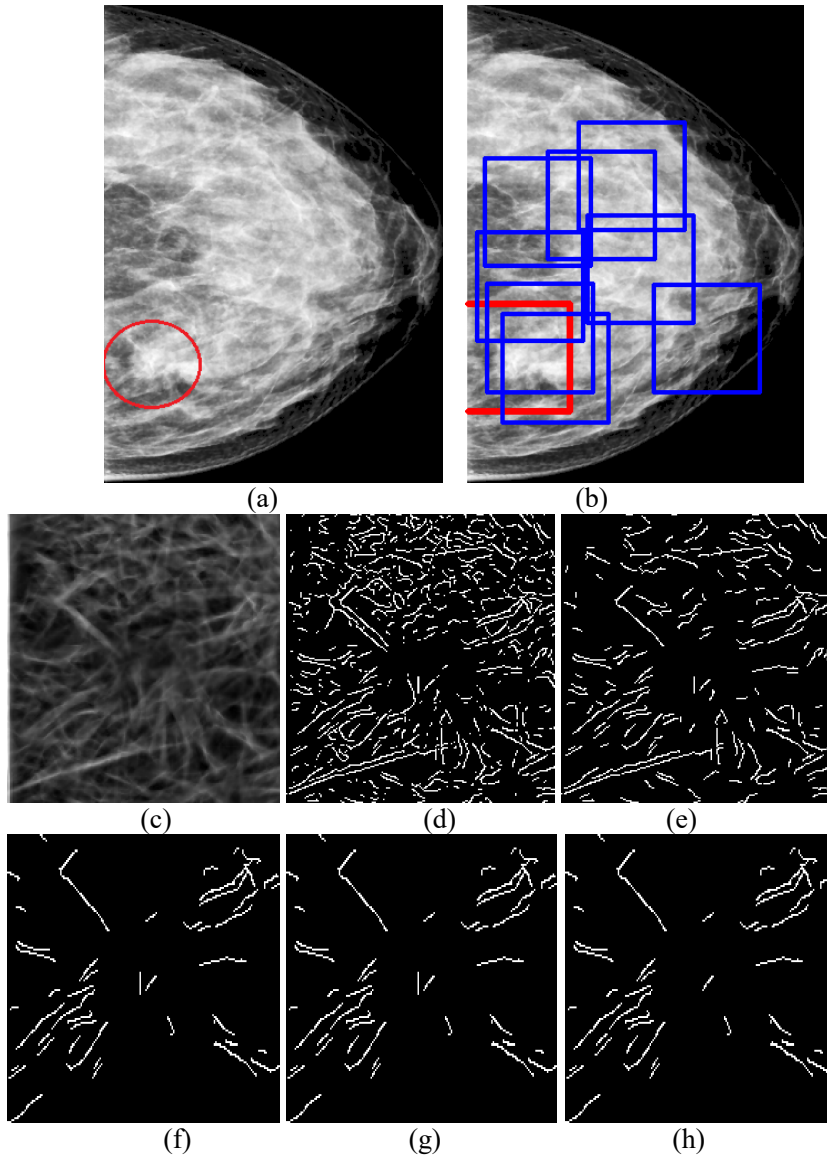


Figure 4.20: Malignant region of dense mammogram: LSH_LCC_262437). (a) Ground truth information in red circle, (b) suspicious ROIs (c) edge features, (d) thinned edge structures, (e) linear structures, (f) converging structures, (g) structures satisfying quadrant criteria1, (h) structures satisfying quadrant criteria2 in malignant ROI of (b) with red border.

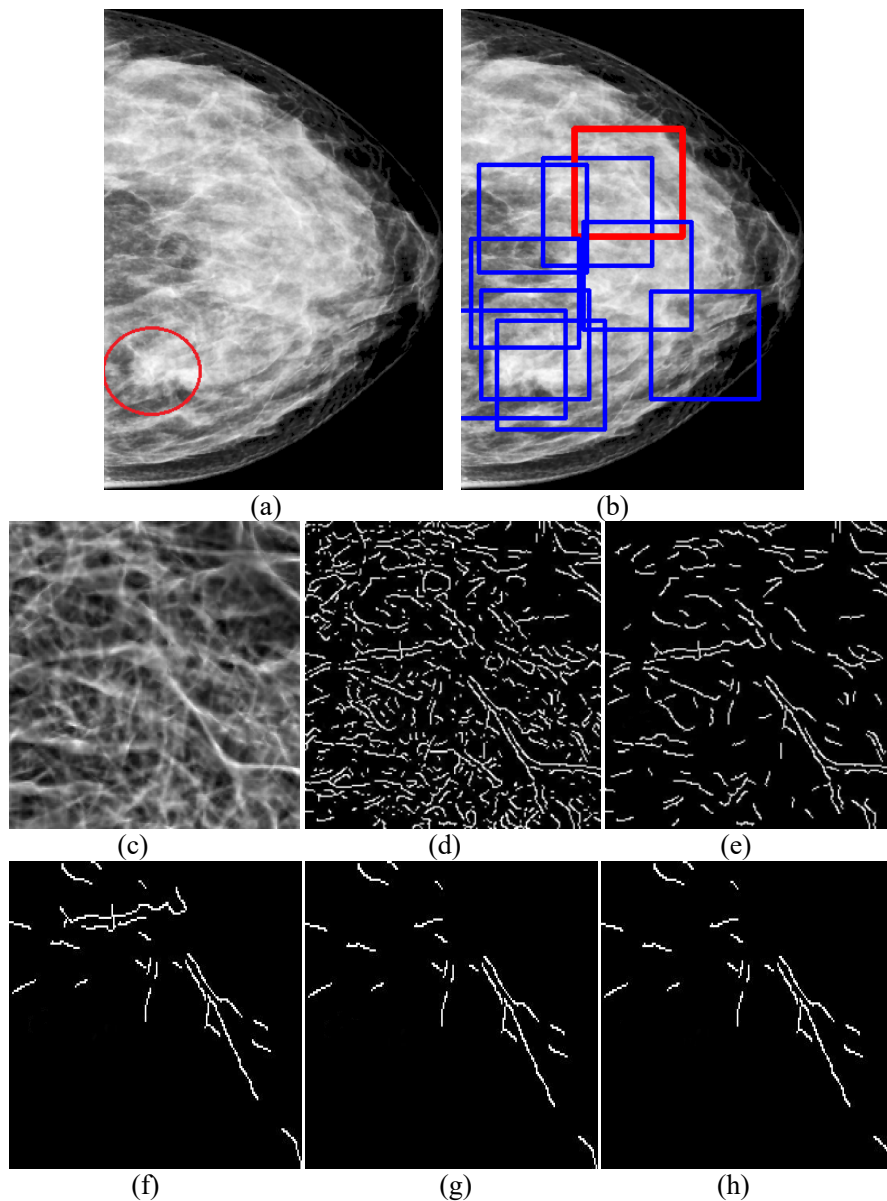


Figure 4.21: Normal region of dense mammogram (LSH_LCC_262437). (a) Ground truth information in red circle, (b) suspicious ROIs (c) edge features, (d) thinned edge structures, (e) linear structures, (f) converging structures, (g) structures satisfying quadrant criterial, (h) structures satisfying quadrant criteria 2 in normal ROI of (b) with red border

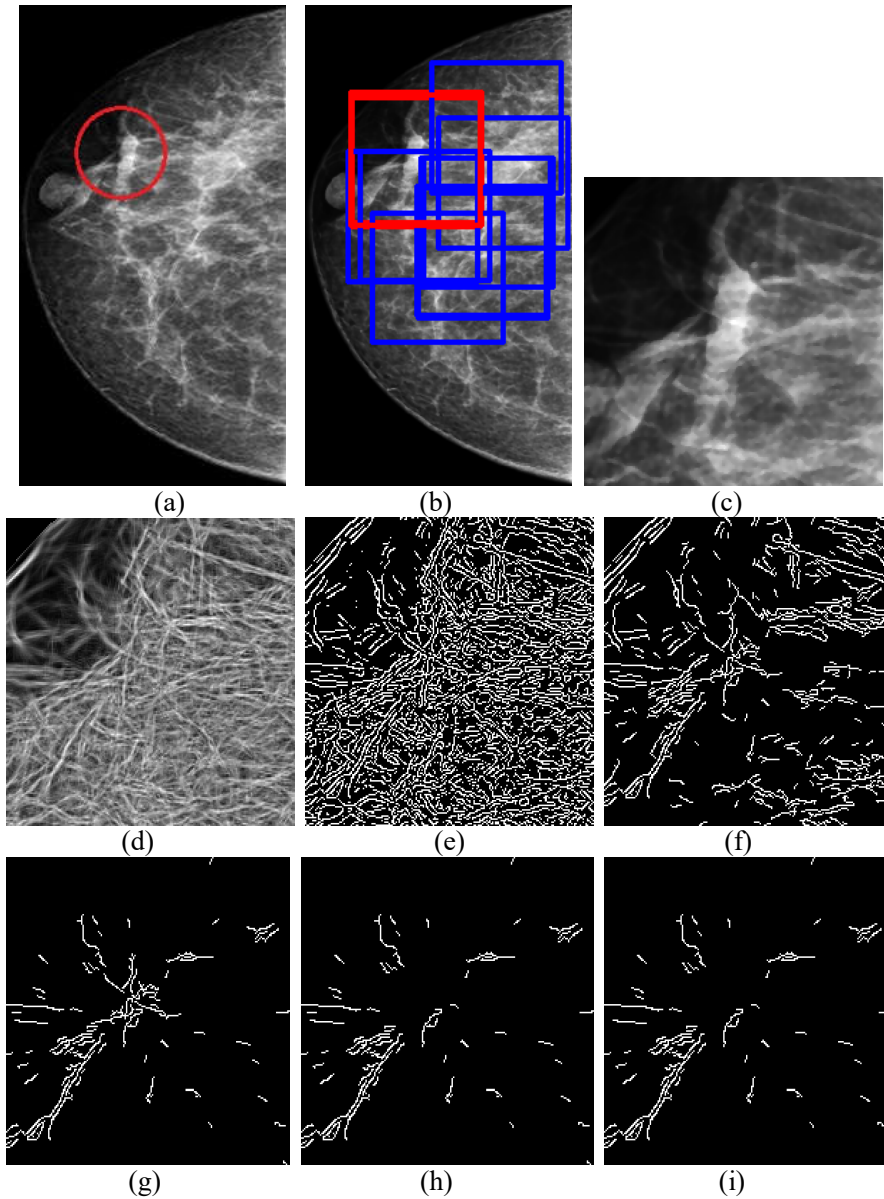


Figure 4.22: Mammogram (LSH_RCC_237263) with malignancy at sub areola region: (a) Ground truth information in red circle, (b) suspicious ROIs (c) malignant ROI, (d) edge features, (e) thinned edge structures, (f) linear structures, (g) converging structures, (h) structures satisfying quadrant criteria 1, (i) structures satisfying quadrant criteria 2 in malignant ROI of (b) with red border.

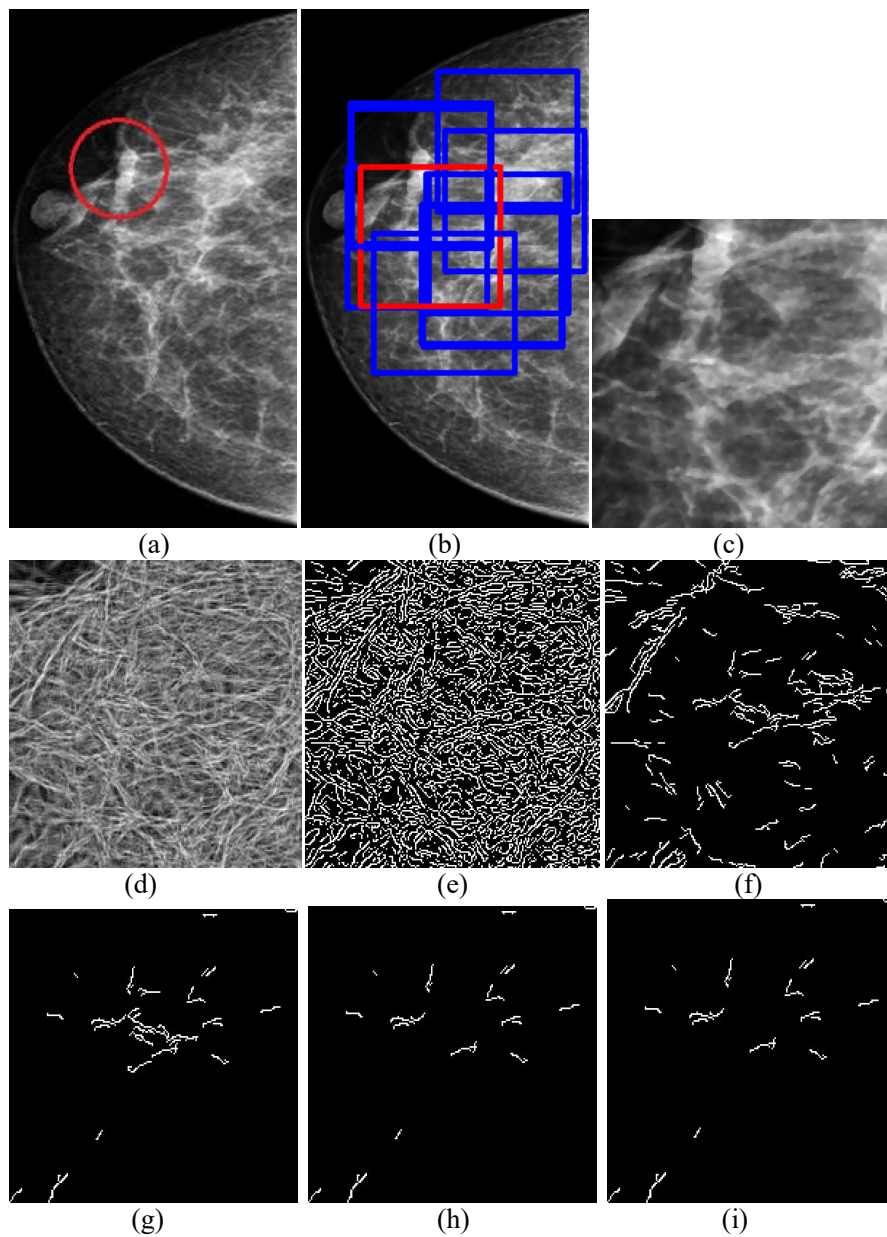


Figure 4.23: Normal region in a mammogram LSH_RCC_237263) with abnormality at sub areola region: (a) Ground truth information in red circle, (b) suspicious ROIs (c) normal ROI, (d) edge features, (e) thinned edge structures, (f) linear structures, (g) converging structures, (h) structures satisfying quadrant criteria 1, (i) structures satisfying quadrant criteria 2 in normal ROI of (b) with red border.

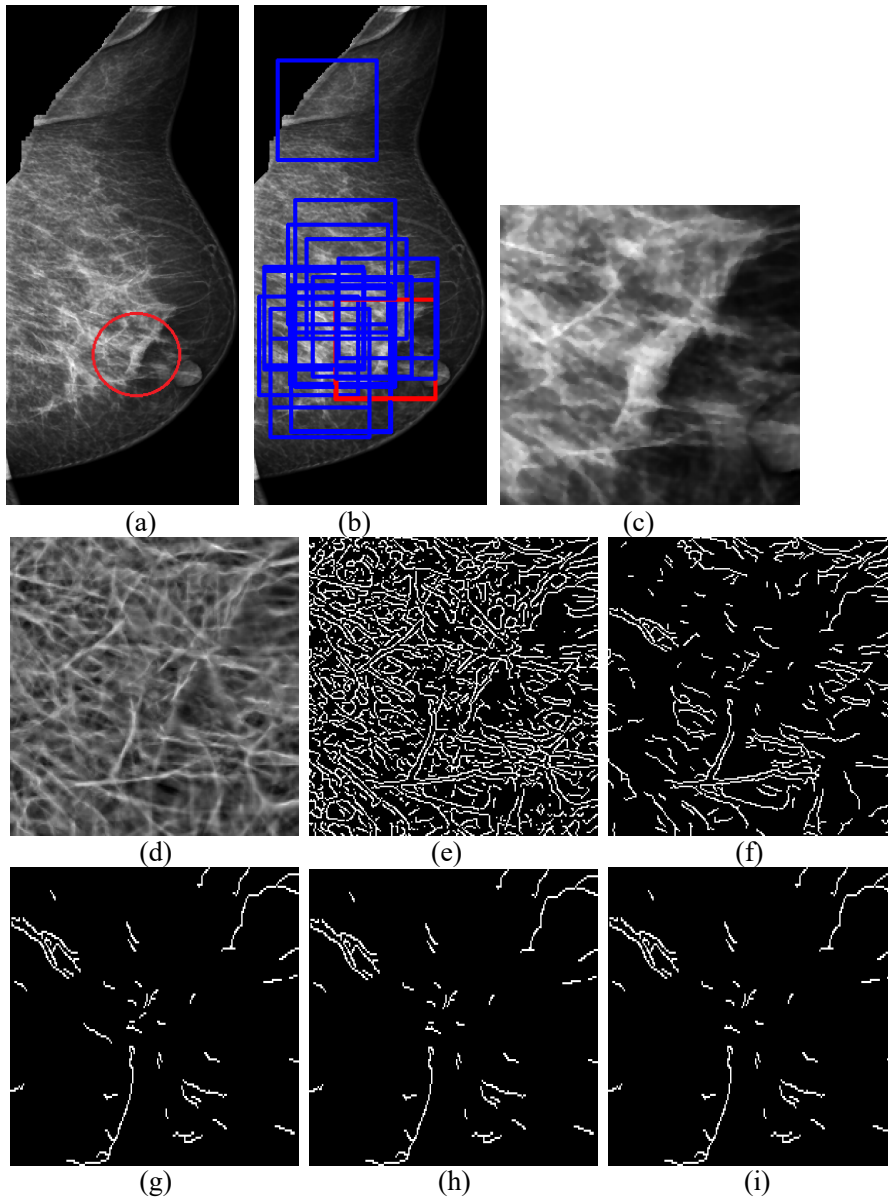


Figure 4.24: Mammogram (LSH_RMLO_237263) with malignancy at sub areola region: (a) Ground truth information in red circle, (b) suspicious ROIs (c) malignant ROI, (d) edge features, (e) thinned edge structures, (f) linear structures, (g) converging structures, (h) structures satisfying quadrant criteria 1, (i) structures satisfying quadrant criteria 2 in malignant ROI of (b) with red border

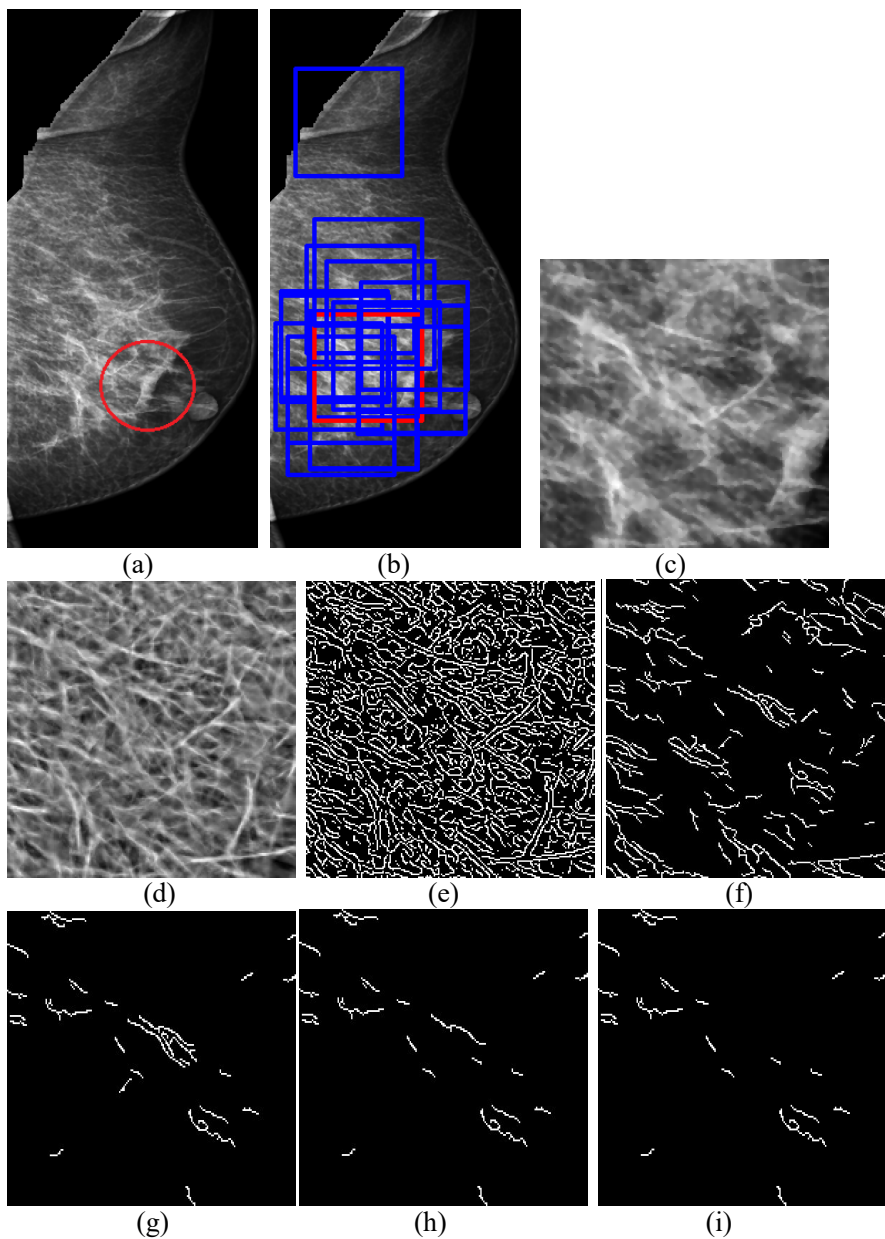


Figure 4.25: Normal region of a mammogram (LSH_RMLO_237263) with malignancy at sub areola region: (a) Ground truth information in red circle, (b) suspicious ROIs (c) normal ROI, (d) edge features, (e) thinned edge structures, (f) linear structures, (g) converging structures, (h) structures satisfying quadrant criteria 1, (i) structures satisfying quadrant criteria2 in normal ROI of (b) with red border.

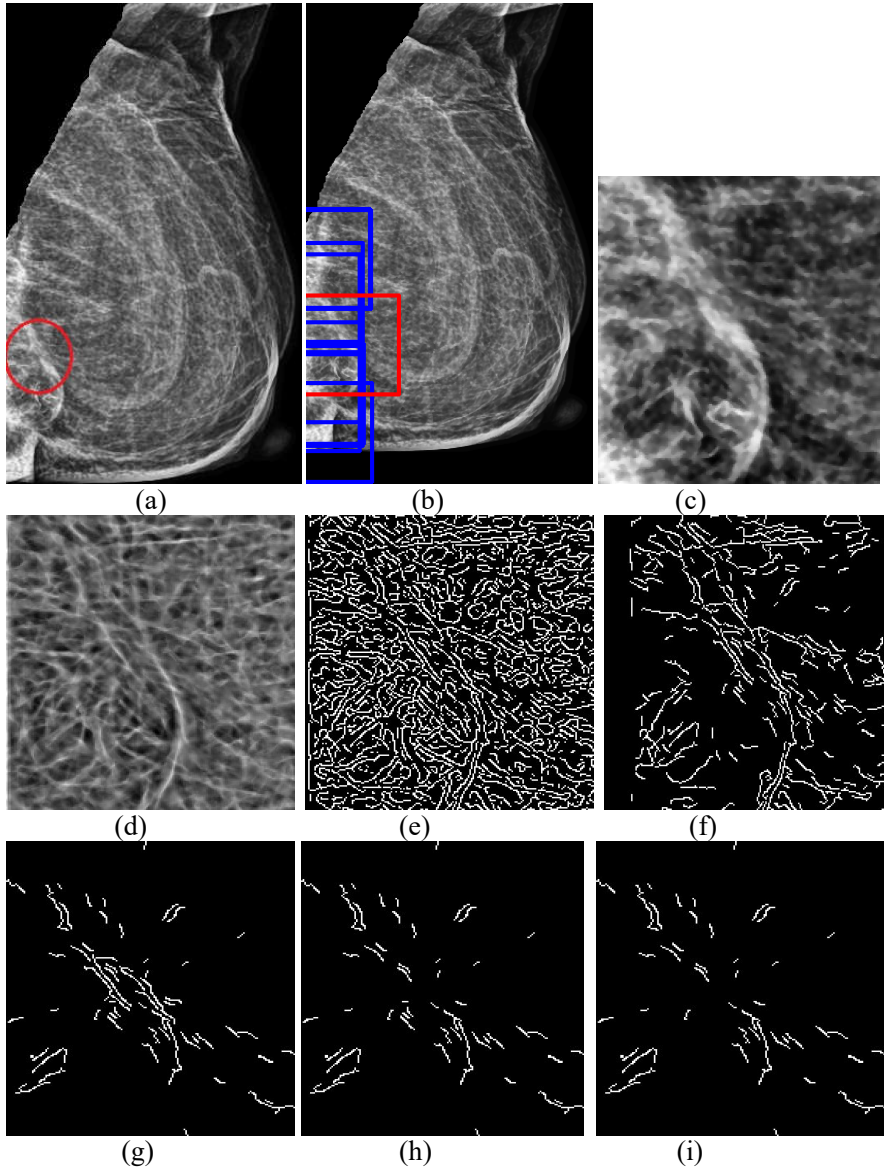


Figure 4.26: Mammogram (LSH_RMLO_182539) with focal retraction: (a) Ground truth information in red circle, (b) suspicious ROIs (c) malignant ROI, (d) edge features, (e) thinned edge structures, (f) linear structures, (g) converging structures, (h) structures satisfying quadrant criteria 1, (i) structures satisfying quadrant criteria 2 in malignant ROI of (b) with red border.

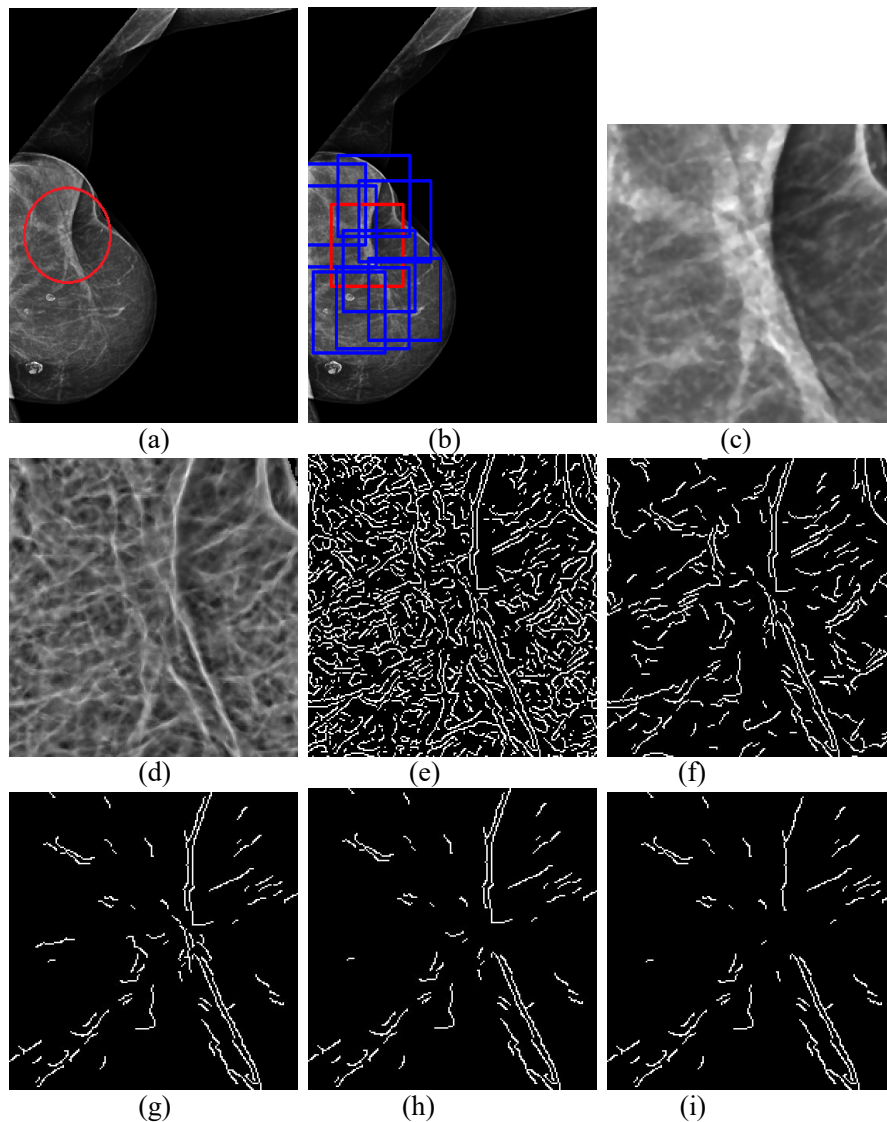


Figure 4.27: Mammogram (LSH_RMLO_1209) with focally retracted malignancy: (a) Ground truth information in red circle, (b) suspicious ROIs, (c) malignant ROI, (d) edge features, (e) thinned edge structures, (f) linear structures, (g) converging structures, (h) structures satisfying quadrant criteria 1, (i) structures satisfying quadrant criteria 2 in malignant ROI of (b) with red border.

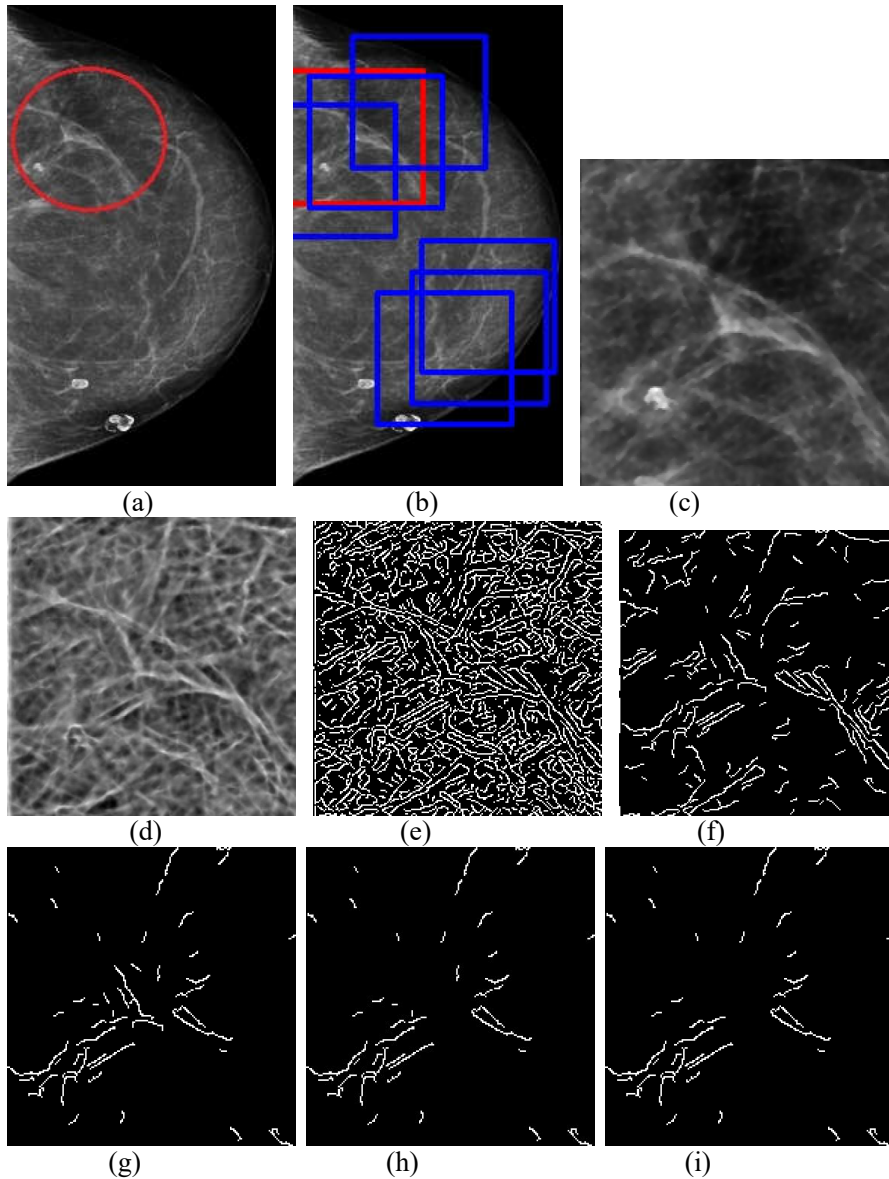
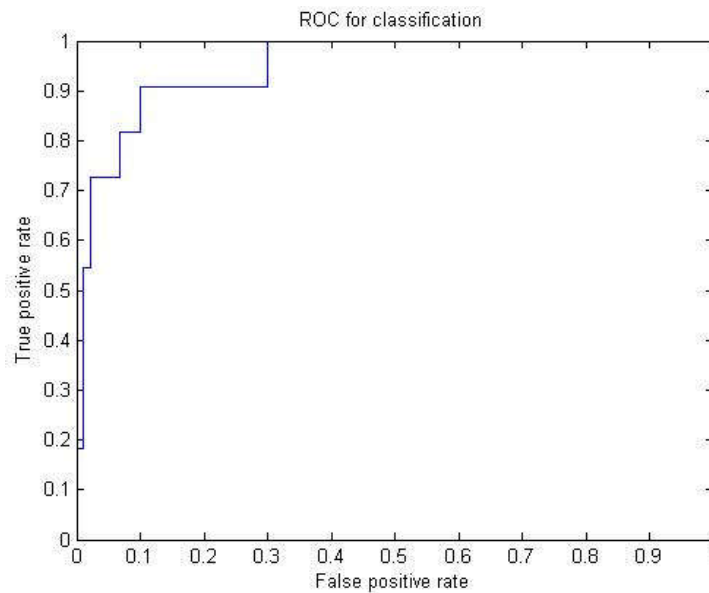


Figure 4.28: Mammogram (LSH_RCC_1209) with RCC view of figure 4.26 with focal retraction: (a) Ground truth information in red circle, (b) suspicious ROIs, (c) malignant ROI, (d) edge features, (e) thinned edge structures, (f) linear structures, (g) converging structures, (h) structures satisfying quadrant criteria 1, (i) structures satisfying quadrant criteria 2 in malignant ROI of (b) with red border.

Chapter 4: Detection and Classification of Architectural Distortion

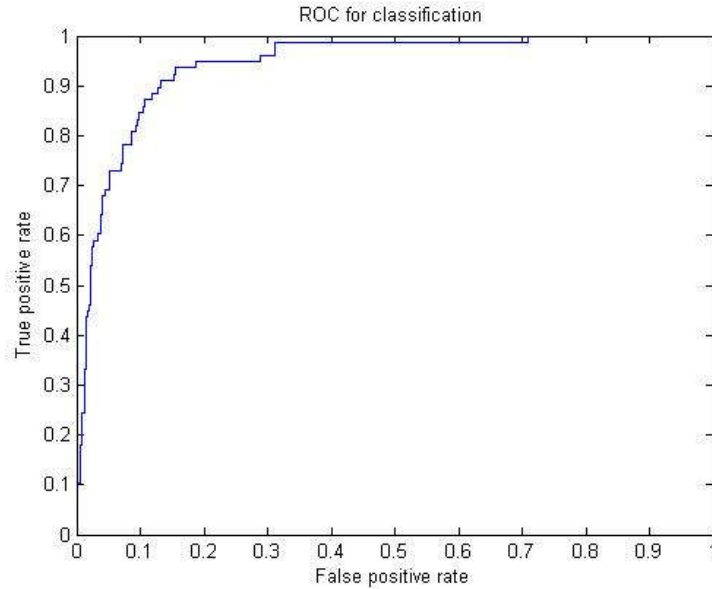
The CS initialized BPNN classification yields a sensitivity of 89%, 89.8% and 97.6% and specificity of 90.9%, 85% and 96.7% respectively for the MIAS, DDSM and Lakeshore databases which is highly promising. Accuracy of the proposed method for Lakeshore database is 97.6% with 0.2% FP/ image, MIAS database is 89.1% with 1.2% FP/ image and DDSM database is 89.1% with 2.5% FP/ image.

The bat initialized BPNN classification yields a sensitivity of 90%, 89.3% and 97.2% and specificity of 90.9%, 83.6% and 96.7% respectively for the MIAS, DDSM and Lakeshore databases which is highly promising. Accuracy of the proposed method for Lakeshore Database is 97.1% with 0.2% FP/ image, MIAS database is 90.1% with 1.2% FP/ image and DDSM database is 88.6% with 2.6% FP/ image. The performance of the proposed method is compared against some state of the art methods in literature for DDSM and MIAS database respectively in the table 4.3 and table 4.4 respectively.

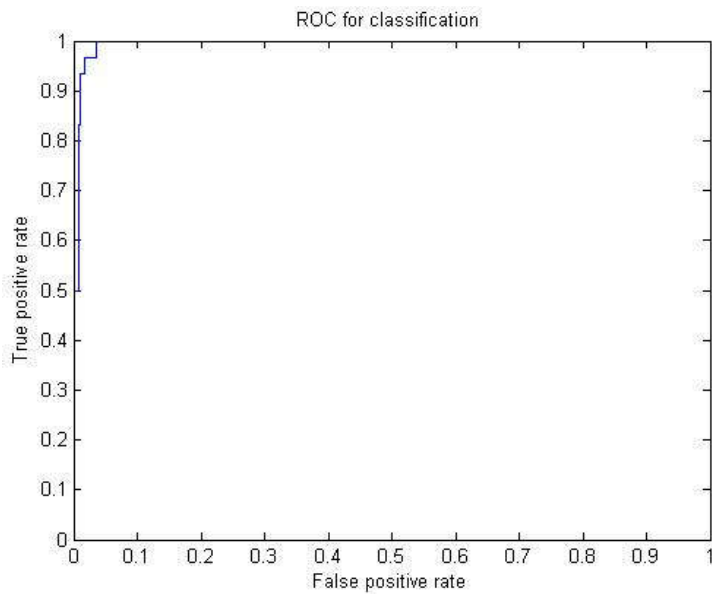


(a)

Chapter 4: Detection and Classification of Architectural Distortion



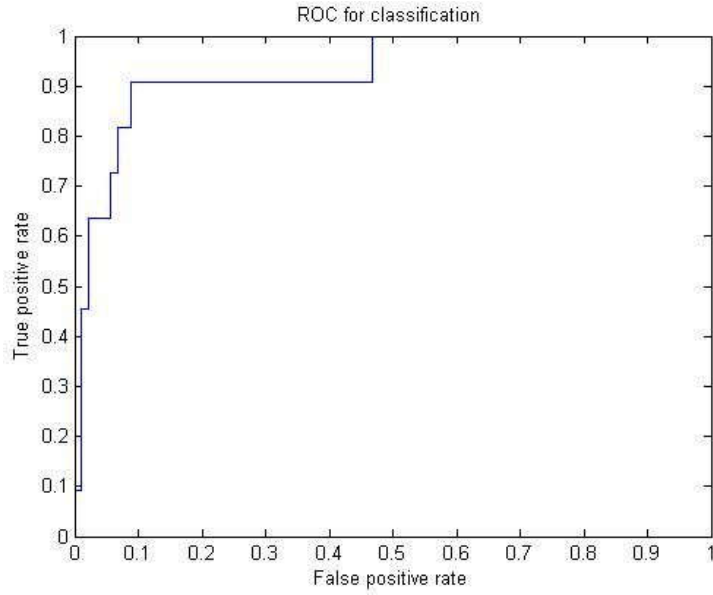
(b)



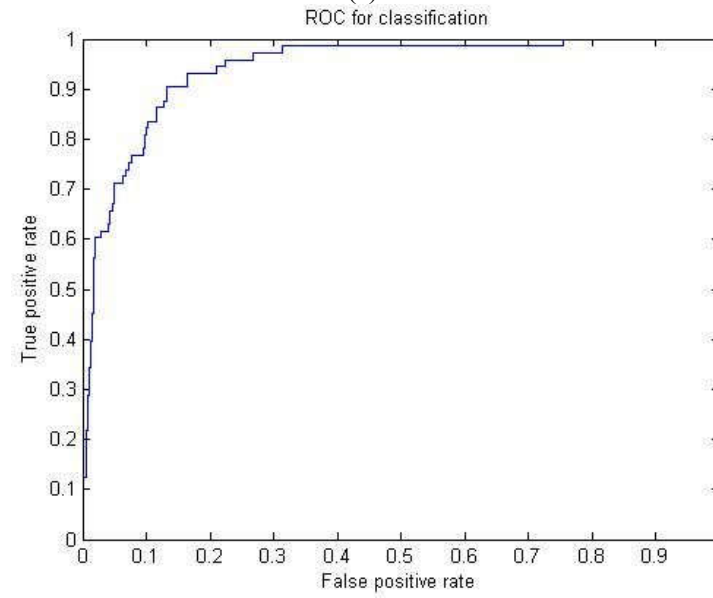
(c)

Figure 4.29: ROC curves for CS BPNN classification of mammographic images collected from (a) MIAS database (b) DDSM database and (c) Lakeshore hospital.

Chapter 4: Detection and Classification of Architectural Distortion



(a)



(b)

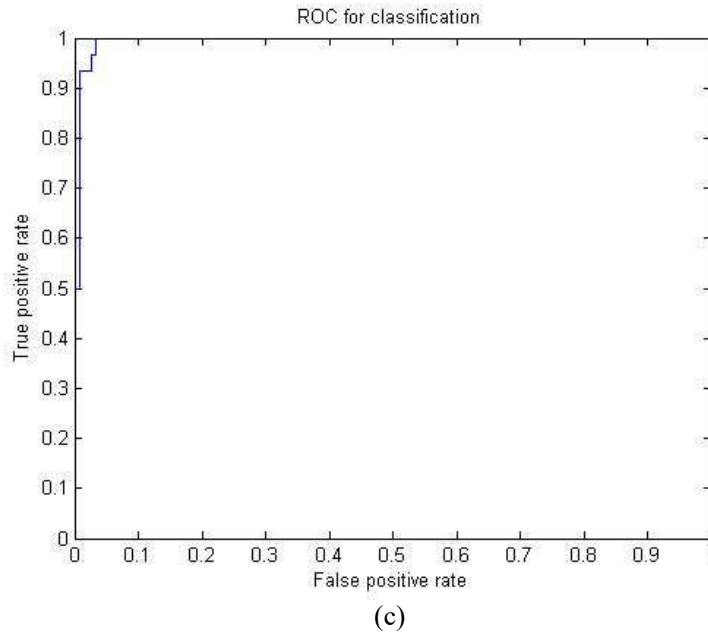


Figure 4.30: ROC curves for bat BPNN classification of mammographic images collected from (a)MIAS database (b) DDSM database and (c) Lakeshore hospital.

4.6. Summary

The proposed method for detection of AD in mammograms shows immense potential. The images from MIAS and DDSM databases as well as images from a representative set of Indian populace from the Lakeshore hospital, Kochi were employed to validate the results. The accuracy of the proposed method was validated using performance measures. A BPNN initialized with the CS/ bat metaheuristic algorithm was employed for the classification of the detected ROIs into normal or with AD features. Use of metaheuristic algorithms to optimize the initial weights of the BPNN significantly improved specificity and sensitivity values.

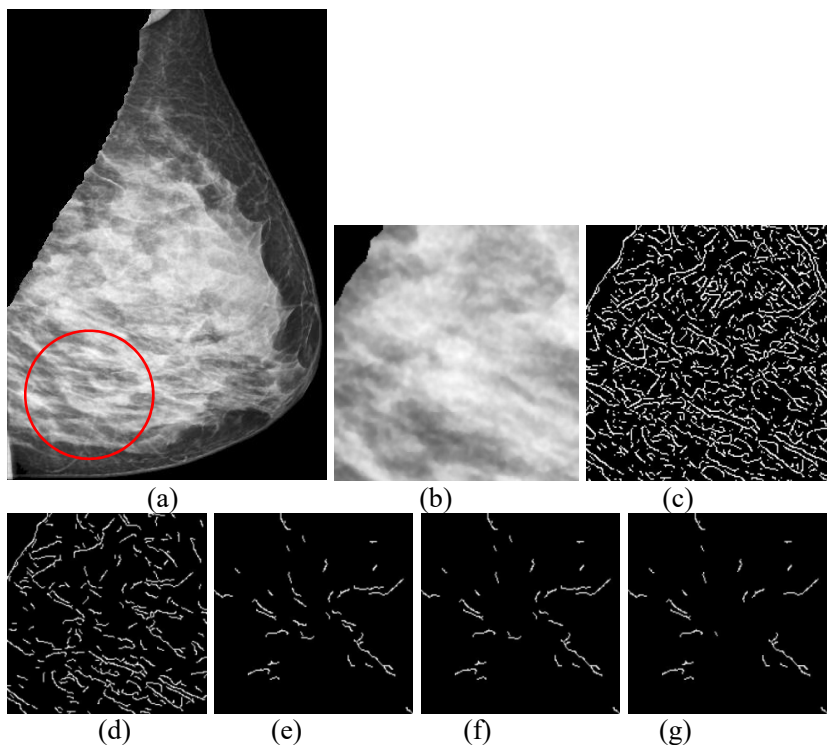


Figure 4.31: Malignant region of dense mammogram (LSH_LMLO_262437): (a) Ground truth information in red circle, (b) malignant ROI, (c) thinned edge structures, (d) linear structures, (e) converging structures, (f) structures satisfying quadrant criteria1, (g) structures satisfying quadrant criteria2 in malignant ROI

Table 4.2. Comparison of classification results for DDSM database.

Author	Performance		
Matsubara et al., 2004 ^[Matsubara, 2004]	80% sensitivity		
Sampat, 2005 ^[Sampat, 2005]	80% sensitivity with 14 FP/ image		
Magdalena Jasionowska, 2010 ^[Jasionowska, 2010]	68% sensitivity with 0.86 FP/ image.		
Proposed work (CS/Bat BPNN), 2015	CS	Bat	
	89.8%	89.3%	sensitivity
	85%	83.6%	specificity
	89.1% with 2.5% FP/ image	88.6% with 2.6% FP/ image	accuracy

Chapter 4: Detection and Classification of Architectural Distortion

Table 4.3. Comparison of classification results for MIAS database.

Author	Performance		
Ayers and Rangayyan, 2005 [Ayers, 2005]	76.5% sensitivity, 76.4% specificity		
Özekes et al, 2005 [Özekes, 2005]	89.02% accuracy		
Guo, 2005 [Guo, 2005]	72.5% accuracy		
Ayers and Rangayyan, 2007 [Ayers, 2007]	84% sensitivity with 4.5 FP/ image		
Amit Kamra et. al., 2012 [Amit, 2012]	89% sensitivity		
Anuradha et. al., 2013 [Anuradha, 2013]	90 % sensitivity, 80 % specificity, 82.86 % accuracy		
Otega, 2013 [otega, 2003]	sensitivity, 79%		
Proposed work(CS/Bat BPNN), 2015	CS	Bat	
	89%	90%	sensitivity
	90.9	89.3%	specificity
	89.1% with 1.2% FP/image	90.1% with 1.2% FP/image	accuracy

4.7. Bibliography

Abhijith, Bailur., Amit, Kumar, Sharma., Shona, Saseendran., Madhav, Kurup., Prithi, Rajagopalan., Abhinav. (2013). A Novel Approach to Clearly Distinguish Architectural distorted area from Mammogram Images, Proc. of Int. Conf. on Advances in Computer Science, published by Elsevier.

American College of Radiology (ACR). (1998). Illustrated Breast Imaging Reporting and Data System (BI-RADS®), American College of Radiology, (3rd edition), (Reston, VA: Editor), 1998.

Amit, Kamra., Sukhwinder, Singh., & V.K Jain. A Novel Method for Detection of Architectural distortion in Mammogram. (2012). ACEEE International Journal on Information Technology, 2(2), 5.

Anuradha, C. Phadke. & Preeti, P. Rege. (2013). Classification of Architectural distortion from Other Abnormalities in Mammograms, International Journal of Application or Innovation in Engineering & Management (IJAIEEM), 2(2).

Ayres, F. J. & R. M. Rangayyan, (2007). Reduction of false positives in the detection of Architectural distortion in mammograms by using a geometrically constrained phase portrait model, International Journal of Computer Assisted Radiology and Surgery, 1(6), 361–369.

Ayres, F. J. & R. M. Rangayyan. (2005). Characterization of Architectural distortion in mammograms, IEEE Engineering in Medicine and Biology Magazine, 24(1), 59–67.

Chapter 4: Detection and Classification of Architectural Distortion

Baeg, S. & N. Kehtarnavaz. (2002). Classification of breast mass abnormalities using denseness and Architectural distortion. *Electronic Letters on Computer Vision and Image Analysis*, 1(1):1–20.

Baker, J. A., E. L. Rosen., J. Y. Lo., E. I. Gimenez., R. Walsh. & M. S. Soo. (2003). Computer-aided detection (CAD) in screening mammography: Sensitivity of commercial CAD systems for detecting Architectural distortion. *American Journal of Roentgenology*, 181, 1083–1088.

Banik, S., R. M. Rangayyan. & J. E. L. Desautels. (2013). Measures of angular spread and entropy for the detection of Architectural distortion in prior mammograms. *International Journal of Computer Assisted Radiology and Surgery*, 8(1), 121–134.

Banik, S., R. M. Rangayyan. & J. E. L. Desautels. (2009). Detection of Architectural distortion in prior mammograms of interval-cancer cases with neural networks. In *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, 6667–6670.

Banik, S., R. M. Rangayyan. & J. E. L. Desautels. (2011). Detection of Architectural distortion in prior mammograms. *IEEE Transactions on Medical Imaging*, 30(2), 279–294.

Bassett, L. W., V. P. Jackson., R. Jahan., Y. S. Fu. & R. H. Gold. (1997). *Diagnosis of Diseases of the Breast*. W. B. Saunders Company, 501-518.

Bekir, Karlik. & A. Vehbi, Olgac, (2011). Performance Analysis of Various Activation Functions in Generalized MLP Architectures of Neural Networks, *International Journal of Artificial Intelligence And Expert*

Systems, 1(4), 111-122.

Bourke, P. (1998), Calculating the area and centroid of polygon, Online: <http://local.wasp.uwa.edu.au/~pbourke/geometry/polyarea/>

Burrell, H. C., D. M. Sibbering., A. R. M. Wilson., S. E. Pinder., A. J. Evans., L. J. Yeoman., C.W. Elston, I. O. Ellis., R. W. Blamey. & J. F. R. Robertson. (1996). Screening interval breast cancers: Mammographic features and prognostic factors. *Radiology*, 199(4), 811–817.

Canny, J. (1986). A computational approach to edge detection, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-8(6), 679-698.
Chen., K. & L. Wang. (2006), *Trends in Neural Computation*, Volume 35 of *Studies in Computational Intelligence*, published by Springer.

Christian, Blum. & Andrea, Roli. (2003)., *Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison*, *ACM Computing Surveys*, 35(3), 268-308.

David, J. Field. (1987). Relations between the statistics of natural images and the response properties of cortical cells, *Journal of Optical Society of America*, 4(12). 2379–2394.

Deneck, Edward. J. (2011). *Barron's Regents Exams and Answers: Earth science-- the physical setting*, published by Barron's Educational Series.

Ellen, Shaw. & De, Paredes. (2007). *Atlas of Mammography*, published by Lippincott Williams & Wilkins.

Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition*. (2nd edition), Academic Press Professional, Inc., San Diego, CA.

Chapter 4: Detection and Classification of Architectural Distortion

Ghana Sheila. K & S. N. Deepak. (2013). Review on Methods to Fix Number of Hidden Neurons in Neural networks, *Mathematical Problems in Engineering*, 1-11.

Girish, Chandrashekar. & Ferat, Sahin. (2014). A survey on feature selection methods, *Computers and Electrical Engineering*, Elsevier, 40, 16–28.

Guo, J. Shao & V. Ruiz. (2005), Investigation of support vector machine for the detection of Architectural distortion in mammographic images. *Journal of Physics: Conference Series*, 15, 88–94.

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation, *Caspian Journal of Internal Medicine*, 4(2), 627 – 635.

Hashimoto. Beverly. (2011). *Practical Digital Mammography*, published by Thieme.

Heeger, David, J. (1987). Model for the extraction of image flow, *Journal of Optical Society of America*, 4(8).

Homer. M. J. (1997). *Mammographic Interpretation: A Practical Approach*, (2nd edition), McGraw-Hill, New York, NY.

Ichikawa.,T., T. Matsubara., T. Hara., H. Fujita., T. Endo. & T. Iwase. (2004). Automated detection method for Architectural distortion areas on mammograms based on morphological processing and surface analysis, (J.

M. Fitzpatrick., & M. Sonka: editors), *Proceedings of SPIE Medical*

Imaging, Image Processing, San Diego, CA, 920–925.

Jasionowska. M., A. Przelaskowski., A. Ruczynska. & A.Wroblewska. (2010). A two-step method for detection of Architectural distortions in mammograms. In *Information Technologies in Biomedicine*, Published by Springer-Verlag, 69, 73–84.

Jiao-hong, Yi., Wei-hong, Xu., & Yuan-tao, Chen. (2014). Novel Back Propagation Optimization by Cuckoo Search Algorithm, *The Scientific World Journal*, 1-8.

Johnston.J.A. & C. Z. Clee. (1991). Analysis of 308 localisation breast biopsies in a New Zealand hospital, *Australas Radiol.*, 2(35), 148-51.

Knutzen. M & J. J. Gisvold. (1993). Likelihood of malignant disease for various categories of mammographically detected, nonpalpable breast lesions. *Mayo Clinic Proceedings*, 68, 454–460.

Kopans. D. B. (1989). Preoperative Imaging-guided Needle Placement and Localization of Clinically Occult Lesions, *Breast Imaging*, (C. Mc. Cormick: Editor), published by J B Lippincott Company, Philadelphia, 320–341. *Breast Imaging*. J. B. Lippincott Company.

Kovesi, P. (1999). Image features from phase congruency, *Videre: Journal of Computer Vision Research*, Massachusetts Institute of Technology, 1(3).

Kumar, S. (2005). *Neural Networks*, McGraw-Hill, New York.

Lakeshore Hospital and Research Centre, Cochin, Website:

Chapter 4: Detection and Classification of Architectural Distortion

<http://www.lakeshorehospital.com>

LiMin, Fu., Hui-hunag, Hsu. & Jose, C, Principe. (1994). A knowledge-Based Approach to Supervised Incremental Learning, Proceedings of IEEE international Conference on Neural Networks and World Congress on Computational Intelligence, 3, 1793-1798.

Math, (2009). Image Processing Toolbox: User's Guide (R2009a). available on www.mathworks.com

Matsubara. T, T. Ichikawa., T. Hara., H. Fujita., S. Kasai., T. Endo. & T. Iwase. (2003). Automated detection methods for Architectural distortions around skinline and within mammary gland on mammograms, International Congress Series: Proceedings of the 17th International Congress and Exhibition on Computer Assisted Radiology and Surgery, (H. U. Lemke., M. W. Vannier., K. Inamura., A. G. Farman., K. Doi., & J. H. C. Reiber : Editors), published by Elsevier, London, UK, 950–955.

Matsubara. T., T. Ichikawa., T. Hara., H. Fujita., S. Kasai., T. Endo. & T. Iwase. (2004). Novel method for detecting mammographic Architectural distortion based on concentration of mammary gland. International Congress Series, Elsevier B.V., 1268, 867–871, June 2004. Proceedings of the 18th International Congress and Exhibition on Computer Assisted Radiology and Surgery (CARS2004).

Matsubara.,T., T, D. Fukuokab., N. Yagic., T. Harac., H. Fujitac., Y. Inenagad., S. Kasaid., A. Kanod., T. Endoe. & T. Iwasef. (2005). Detection method for Architectural distortion based on analysis of structure of mammary gland on mammograms, International Congress Series, published by Elsevier, 1281, 1036-1040.

Chapter 4: Detection and Classification of Architectural Distortion

Morrone, M. C. & R. A. Owens. (1987). Feature detection from local energy. *Pattern Recognition Letters*, 6 (5), 303-313.

Morrone, M. C., J. R. Ross., D. C. Burr. & R. A. Owens. (1986). Mach bands are phase dependent. *Nature*, 324(6094), 250-253.

Nakayama, R., R. Watanabe., T. Kawamura., T. Takada., K. Yamamoto. & K. Takeda. (2008). Computer aided diagnosis scheme for detection of Architectural distortion on mammograms using multiresolution analysis. In *Proceedings of the 22nd International Congress and Exhibition on Computer Assisted Radiology and Surgery, Barcelona, Spain*, 3(1), S418–S419.

Nandi, R. J., A. K. Nandi., R. M. Rangayyan. & D. Scutt. (2006). Classification of breast masses in mammograms using genetic programming and feature selection. *Medical and Biological Engineering and Computing*, 44, 683–694.

Nazri, Mohd, Nawi., Abdullah, Khan. & Mohammad, Zubair, Rehman. (2013). A New Back-propagation Neural Network Optimized With Cuckoo Search Algorithm, *Computational Science and Its Applications, Lecture Notes in Computer Science*, published by Springer Berlin Heidelberg, 7971, 413-426.

Osman, Ibrahim, H. & Gilbert, Laporte. (1996) Metaheuristics: A bibliography *Annals of Operations Research*. 63(5), 511-623.

O'tega, A. Ejofodomi., Edikan, Nse, Gideon. Gbenga, Olalekan, Oladipo. & Etse, Rosemary, Oshomah. (2014). Automated detection of Architectural distortions in mammograms using template matching,

Chapter 4: Detection and Classification of Architectural Distortion

International Journal of Biomedical Science and Engineering, 2(1):1-6.

Özekes, S. , O. Osman. & A. Y. Çamurcu. (2005). Computerized detection of Architectural distortions in digital mammograms. In Proceedings of the 19th International Conference on Computer Assisted Radiology and Surgery (CARS2005), 1281, 1396, Berlin, Germany.

Paola, Casti., Arianna, Mencattini., Marcello Salmeri., Antonietta, Ancona, Fabio, Felice, Mangieri., Maria, Luisa, Pepe. & Rangaraj Mandayam Rangayyan. (2013). Automatic Detection of the Nipple in Screen-Film and Full-Field Digital Mammograms Using a Novel Hessian-Based Method, *Journal of Digital Imaging*, 26(5), 948–957.

Rangayyan. R. M, S. Banik., J. Chakraborty., S. Mukhopadhyay. & J. E. L. Desautels. (2013). Measures of divergence of oriented patterns for the detection of Architectural distortion in prior mammograms. *International Journal of Computer Assisted Radiology and Surgery*, 8, 527–545.

Rangayyan. R. M. & F. J. Ayres. (2006). Gabor filters and phase portraits for the detection of Architectural distortion in mammograms. *Medical and Biological Engineering and Computing*, 44, 883–894.

Rangayyan. R. M., S. Banik., J. Chakraborty., S. Mukhopadhyay. & J. E. L. Desautels. (2013). Measures of divergence of oriented patterns for the detection of Architectural distortion in prior mammograms, *Int J CARS.*, 8, 527–545.

Rumelhart, David, E., Hinton, Geoffrey, E. & Williams, Ronald, J. (1986). Learning representations by back-propagating errors, *Nature* 323 (6088), 533–536.

Chapter 4: Detection and Classification of Architectural Distortion

Sampat.M. P., G. J. Whitman., M. K. Markey. & A. C. Bovik. (2005). Evidence-based detection of spiculated masses and Architectural distortions,” in Proc. of SPIE Medical Imaging: Image Processing, San Diego, 5747, 26-37.

Saurabh, Karsoliya. (2012). Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture, International Journal of Engineering Trends and Technology, 3(6).

Sickles. E. A. (1986). Mammographic features of 300 consecutive nonpalpable breast cancers. American Journal of Roentgenology, 146(4), 661–663.

Smith, S. M. (1995). SUSAN - a new approach to low level image processing. Technical Report TR95SMS1, Defence Research Agency, Chobham Lane, Chertsey, Surrey, UK.

Smith, S. M. (1996). Flexible filter neighbourhood designation. In Proc. 13th Int. Conf. on Pattern Recognition, 1, 206-212.

Suckling, J. (1994). The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. International Congress Series, 1069, 375-378.

Uzer, Mustafa Serter., Inan, O. & Yılmaz, N. (2013). A hybrid breast cancer detection system via neural network and feature selection based on SBS, SFS and PCA, Neural Computing & Applications. 23(3), 719-728.

Venkatesh, S. & R. A. Owens. (1989), An energy feature detection scheme. Proceeding of the International Conference on Image Processing,

553– 557.

Wang, L. (2008), Feature selection with kernel class separability. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(9), 1534–1546.

Xin-She, Yang. (2013). Bat algorithm and cuckoo search: a tutorial, *Artificial Intelligence, Evolutionary Computing and Metaheuristics*, (Eds. X. S. Yang), *Studies in Computational Intelligence*, 427, 421–434.

Xin-She, Yang. (2014). *Cuckoo Search and Firefly Algorithm: Theory and Applications*, *Studies in Computational Intelligence*, (1st edition), Published by Springer International, 516.

Zhiqiang, Lao. Xin, Zheng. (2011). Multiscale quantification of tissue spiculation and distortion for detection of architectural distortion and spiculated mass in mammography, Published in *SPIE Proceedings, Computer-Aided Diagnosis* (Ronald M. Summers; & Bram van Ginneken: Editors), 7963

Chapter 5

Enhancement of microcalcification

Visibility of the commonly occurring breast cancer symptom, microcalcification, can be improved by using computer aided techniques. In the proposed research work, approaches based on modulus maxima and zero crossings of contourlet transform were utilized to enhance microcalcification features. Parent-child relationship among the contourlet coefficients at various levels was employed to retain the strong edge information that corresponds to the relevant features and to minimize the artifacts. Contrast, target to background contrast ratio and contrast improvement index parameters are considered for performance evaluation of the enhancement algorithm. The MIAS mammographic image database was employed for testing the accuracy of the proposed method. The results illustrate the efficacy of the proposed method.

5.1. Introduction

Mammography, the golden standard for breast cancer detection, is efficient in imaging breast changes up to two years before a physician can identify them [Hajar, 2012]. Mammograms are x-ray images consisting of various tissue structures, ligaments, blood vessels, ducts, abnormal structures etc. [Bao-Ning Zhang, 2015]. Enhancement techniques on mammography improve the visibility and quality of image thereby revealing the presence of abnormalities like microcalcification (MCC). The enhanced appearance of MCC helps radiologists in distinguishing the abnormal mammogram from the normal. When the x-ray beam passes through the compressed breast, the absorption of x-ray photons by calcium is ten times more than that of breast tissue [Cornelius, 2005]. Hence the visibility and capturing capacity of MCC is more in mammography than with other imaging modalities such as MRI and ultrasound [Cornelius, 2005]. MCC is a non-palpable symptom of breast cancer.

Even though the calcium deposits accumulated in breast duct appear on mammogram, detection through naked eye is difficult for the radiologist. MCCs are tiny white spots which appears brighter than the surrounding breast tissues and background. The size of MCC ranges from 0.05 mm up to 1 mm [Tomislav, 2013]. The small size, subtle nature, low contrast, presence of noise and dense tissues reduce the visibility of MCC in mammogram. A cluster of MCC consists of approximately 10 to 100 single MCC findings with different size and irregular shape [Cornelius, 2005].

Le Gal's classification [Le Gal, 1973] of MCC is shown in figure 5.1. According to Le Gal, there are five major types (type 1- 5) of classification with an increasing probability of malignancy. Type 1 is a benign MCC with 0% probability for malignancy. It usually appears as a 'ring' or 'tea cup' shaped MCC. The sediment calcification usually appears in micro cysts or in ecstatic ducts. Type 2 has 19% probability of malignancy with round shape, regularly punctiform but of variable size.

Type 3 is fine, powdery MCCs with a malignancy rate of 39%. Type 4 MCCs are similar to those of type 2 but with variability in size, shape and irregular border. It has got 59% probability of malignancy. Type 5 is of worm like, linear or branching pattern of MCC with 100% probability of abnormality.

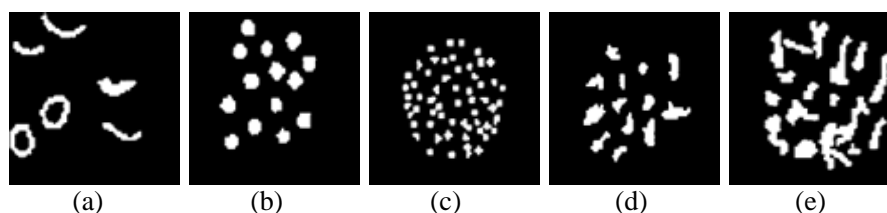


Figure 5.1: Le Gal's classification of MCC

5.2. Literature review

There are several research works related to MCC detection [Cornelius, 2005]. All methods for the enhancement of MCC features try to improve the visibility of MCC using features such as discontinuous change in intensity, appearance at multiple resolution and orientation, global variation in texture, roughly circular shape etc. The literature survey shows that most of the MCC enhancement methods are based on various techniques including statistical measures, fuzzy based methods, multi-scale analysis, morphological operators, fractal analysis, modeling, background subtraction, unsharp masking etc.

Strickland [Strickland, 1996] developed a two stage method based on undecimated Wavelet Transform (WT) using prewhitening matched filters for detecting gaussian objects such as MCCs. Chen [Chen, 1997] developed a method combining multiscale wavelet analysis and gaussian markov random field technique for detecting MCCs. Kim [Kimz, 1997] utilized first derivative and local statistics for enhancing MCCs in mammogram. Ted [Ted, 1998] applied a nonlinear thresholding method on high frequency sub-bands of daubechies 4 and daubechies 20 wavelets to

Chapter 5: Microcalcification Enhancement

enhance the MCCs present in mammogram. The low frequency sub-band is suppressed to utilize the high frequency characteristics of the image. The local maxima locations in an image were used to represent MCCs as they appear as bright spots in mammogram [Bagci, 2002]. A fourth order statistic around maxima locations over sub-bands in adaptive WT isolates MCCs. Peter [Peter, 2003] suggested the use of integrated wavelets for the enhancement of multiscale structures such as MCCs in mammograms. A modified form of spot enhancement function is applied on wavelet coefficients. Wael [Wael, 2008] applied a method based on fractal modeling for enhancing MCCs embedded in inhomogeneous breast tissues of mammogram.

A top-hat morphological operation proposed by Xinsheng [Xinsheng, 2009] with structuring element of different sizes and shapes highlight MCCs. Even though the method improves the visibility of MCCs, it also enhances the noise present in the image. Tomislav [Tomislav, 2010] recommended two different methods for MCC identification. In the first method of multifractal analysis, MCCs are detected in regions with high values of hölder exponent representing high local intensity variation and low values of its distribution indicating the occurrence of rare events. The second method extracts MCCs by iteratively performing a highpass filtering technique using top-hat and bottom-hat morphological operation. Balakumaran [Balakumaran, 2010] proposed a MCC enhancement method using dyadic WT. Weighted detail sub-bands with low frequency sub-band assigned as zero are used for reconstruction. Fuzzy shell clustering algorithm is used for detecting MCCs from the reconstructed image. The method proposed by Mohanalin [Mohanalin, 2010] enhances MCCs with the help of fuzzy algorithm using tsallis entropy. The enhancement of mammary densities and the high time consumption are the disadvantages of the techniques.

In the method proposed by Hashemi [Hashemi, 2012], the soft and hard thresholds based on statistical measures such as skewness and kurtosis

on coefficients of daubechies-2 (db2) WT were used to extract MCC. Hajar Moradmamand [Hajar, 2012] analysed four different methods for enhancing MCCs: WT using five scales of asymmetric daubechies of order 8 decomposition, contrast limited adaptive histogram equalization, morphological operation and unsharp masking. Among the four methods, wavelet based method reduced noise with MCC enhancement. Jaganath [Jaganath, 2012] compared two morphological techniques including top-hat and h-dome morphological processing for the enhancement of MCC. Chao [Chao, 2012] proposed an MCC detection method by retaining local maxima points obtained from the result of mean shift algorithm. A preprocessing method based on gaussian smoothing is used to remove noise from the mammogram. Marrocco proposed Watershed segmentation method Marrocco [Marrocco, 2012] to automatically detect MCCs. However the over segmentation due to watershed processing generates false positive results. Tay et al, [Tay, 2013] proposed a multilevel pyramidal decomposition to enhance the MCC features in mammograms. Instead of the usual low pass filter, a squeeze box filter is used for smoothing. The details are enhanced using a nonlinear function to enhance low contrast features without edge blurring. Zouari [Zouari, 2014] introduced a nonlinear stretching process for the identification of MCCs. A global and local application of nonlinear stretching increased the visual appearance of MCCs.

Sundaram [Sundaram, 2014] proposed a technique to improve the visibility of MCCs by finding the malignancy in low contrast region as weak edges and high contrast region as strong edges. The strong and weak edges are obtained with average filter and unsharp mask respectively. Discrete WT is applied separately on images with weak and strong edges. Coefficients of sub-bands are modified on the basis of energy before reconstruction to yield an enhanced image with visible MCCs. Discrete WT coefficients of all sub-bands are enhanced by energy based mapping on strong and weak edge positions. Arya [Arya, 2015] proposed a

Chapter 5: Microcalcification Enhancement

mammographic image enhancement algorithm using modulus maxima on coefficients of stationary WT. A high boost filtering operation on thresholded detail image with approximation coefficients set as zero is used for reconstruction. Addition of reconstructed image on original image yields an enhanced image.

5.3. Proposed methods

Mammographic image composed by edges, textures and edge associated details oriented in different directions. While processing the image, all these details must be retained. An image transform is characterized by a set of basis functions. Efficiency of an image representation refers to the ability to capture significant information about an object of interest using a sparse description with few non-zero coefficients. Vanishing moments property is the key for the sparse expansion of piecewise smooth signals. For one dimensional piecewise smooth signals, wavelet is a right tool as they provides an optimal representation of signals. As natural images are not simply stacks of 1D piecewise scan lines; discontinuity points (i.e., edges) are usually located along smooth curves (ie. contours) owing to smooth boundaries of physical objects. Wavelets in two dimension are good at isolating the discontinuities at edge points, but will not sense the smoothness along the contours. Two dimensional separable wavelets have directional vanishing moments in the horizontal and vertical directions which make wavelets especially good in capturing horizontal and vertical edges. Thus the horizontal edges, vertical edges and edges oriented at $\pm 45^\circ$ alone are properly detected, as the wavelets employed have aspect ratio equal to one. This disadvantage is overcome with directional transforms such as contourlet transform (CT), which provide directionality and anisotropy property in addition to multiresolution and localization. Research in human visual system shows the importance of directional information in human visual perception. CT

has basis functions oriented at many directions in comparison with the only three directions of wavelets, are able to detect edges oriented along more directions as the basis functions appear at different aspect ratios depending on the scale. A contourlet expansion that satisfies the anisotropy scaling law and sufficient directional vanishing moments achieves the optimal non-linear approximation behavior for two dimensional piecewise smooth functions with discontinuities along smooth curve [Do, 2005].

MCCs can be considered as objects defined by discontinuities or edges [Highnam, 1999]. Calcifications represent high spatial frequencies in the image [Sampat, 2005]. Thus, one approach to the calcification detection task is to localize the high spatial frequencies of the image. Many of the works in literature for the enhancement of MCCs are based on WT. Mallat and Zhong [Mallat, 1992] calculated the local maxima of discrete WT at each scale and formed a multiscale edge representation of an image which characterizes the shape of irregular structures while zero crossings of the WT were considered by Mallat [Mallat,1991]. The general wavelet based procedure for enhancing the MCCs present in a mammogram are decomposition using WT, modification of wavelet coefficients at various scales and reconstruction of image with modified coefficients [Arianna, 2008]. Even though WT is good in capturing the points of discontinuities; there are some drawbacks for the WT. The directional information of WT is limited and confined only to the vertical, horizontal and diagonal directions [Do, 2001].

The proposed algorithms make use of the multiscale and directionality properties of the CT [Do, 2005] to enhance clinically important MCC features in mammograms. The detail coefficients of CT obtained from pyramidal decomposition having large absolute values correspond to sharp intensity changes indicating edges. The proposed methods enhance the sharp variation points indicating the possible occurrence of MCCs at the same time suppressing various artifacts. The algorithm makes use of the multiscale and directionality properties of the

Chapter 5: Microcalcification Enhancement

CT at various levels. The high directional perceptiveness and anisotropy property of the needle-shaped elements of CT are very proficient in representing edges. Singularities can be identified by the presence of zero crossings/ modulus maxima of contourlet coefficients. The zero crossings/ modulus maxima of the CT at different scales with the directional characteristics satisfying the parent-child relationship can be used to locate the sharp variation points indicating MCCs in mammographic image. The algorithm of the proposed methods is given below.

Step1: Extract contourlet coefficients using CT with four levels of directional decomposition on each of the three scale laplacian pyramid (LP).

Step2: Apply zero crossings/ modulus maxima on corresponding directional sub-bands at different levels.

Step3: Coefficients that hold parent-child relationships from finer to coarser levels are retained and others are pruned out.

Step4: Contourlet coefficients of step1 are modified on the basis of retained coefficients in step3.

Step5: Perform contourlet reconstruction with modified detail coefficients to yield enhanced image.

5.3.1. Extraction of contourlet coefficients

A region of MCC is distinguished from the normal area by identifying the tiny bright spots which are represented as sharp intensity discontinuities in the mammogram. An efficient frequency domain representation of mammogram is possible through transforms such as WTs, CTs etc. Compared to the WT, CT exhibits better approximation precision, excellent noise reduction, improved edge representation, efficiency in capturing fine details and high directionality in multiple scale with flexible aspect ratio. A singular curve is represented with less number of contourlet coefficients than wavelet coefficients [Do, 2001].

CT is a two dimensional transform with a capability to capture and link the discontinuity points to form linear structures. A smooth contour is obtained from the point of discontinuities. The contourlets have elongated supports at various scales, directions and aspect ratio. As the image is expanded with basis image as contour segments, the transform is coined as CT. Figure 5.2 describes the difference in the representation of a curve using basis elements of CT and WT.

CT is also known as pyramidal directional filter bank (PDFB) [Do, 2005]. The implementation of CT consists of two stages: multiscale decomposition and multidirectional decomposition. The multiscale decomposition is achieved using LP whereas the multidirectional decomposition is implemented with directional filter bank (DFB).

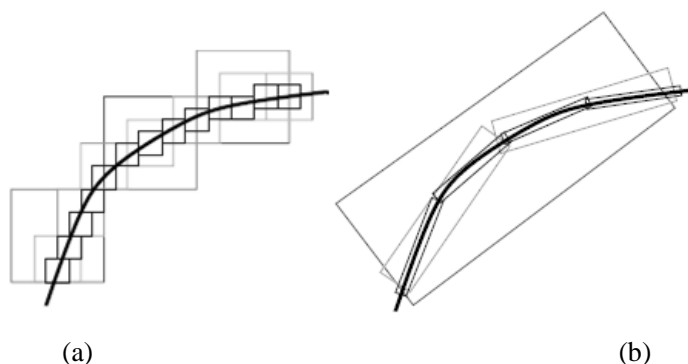


Figure 5.2: Representation of a curve using basis elements of (a) wavelet transform and (b) contourlet transform ^[Do, 2001]

A representation of the analysis filter bank for computing CT is shown in figure 5.3. LP proposed by Burt and Adelson [Burt, 1983] is used for multiscale decomposition where a sampled lowpass result of original image and the difference between the original and the prediction are generated. Figure 5.4(a) and 5.4(b) depicts the analysis and synthesis of LP decomposition. A coarse approximation ‘ c ’ is obtained by decimating a low pass filtered version of original image, ‘ x ’. A prediction image, ‘ p ’ is generated by interpolating a low pass filtered version of coarse

Chapter 5: Microcalcification Enhancement

approximation. A band pass image, ' d ' is obtained by subtracting prediction image from original image.

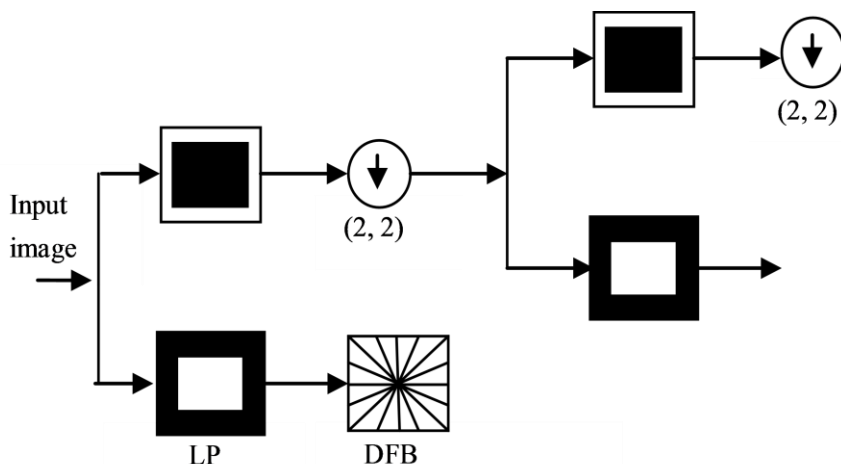


Figure 5.3: Block diagram of contourlet transform

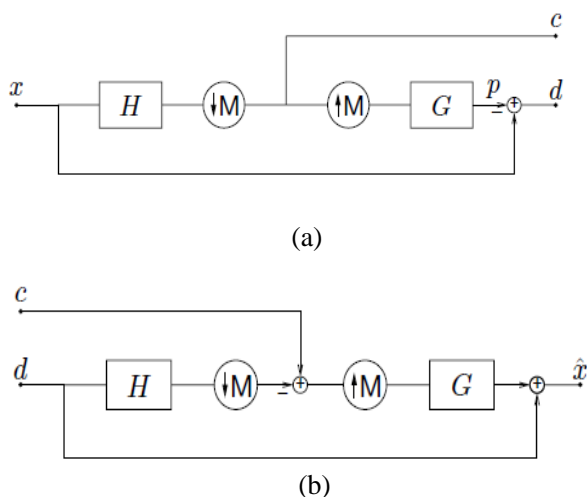


Figure 5.4: LP Decomposition. (a) Analysis. (b) Synthesis ^[Do, 2001]

Further coarser approximations if needed, are obtained by decomposing the approximation image ' c ' using additional levels of pyramidal decomposition as needed. Directional Information is obtained from the band pass image, ' d ', which is given as input to the DFB [Bamberger, 1992]. Efficient directional information at various scales can be obtained by the appropriate combination of the LP and DFB

decomposition. LP separates low frequency information to avoid leaking of low frequencies into directional sub-bands.

DFB is implemented by an m -level binary tree decomposition using quincunx filter banks with fan filters leading to 2^m sub-bands of wedge-shaped frequency partitioning as shown in figure 5.5.

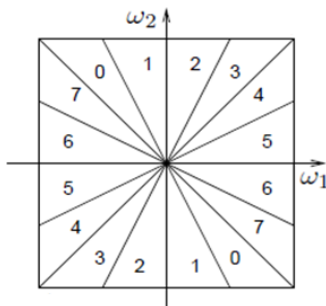


Figure 5.5: Wedge shape frequency partition for a 3-level DFB decomposition that leads to 8 (2^3) sub-bands [Do, 2001].

Quincunx filter bank partitions the frequency spectrum of input signal into horizontal and vertical channel with the help of fan filters [Do, 2001]. DFB is constructed from two building blocks. Figure 5.6 depicts the two stages of DFB using fan filters.

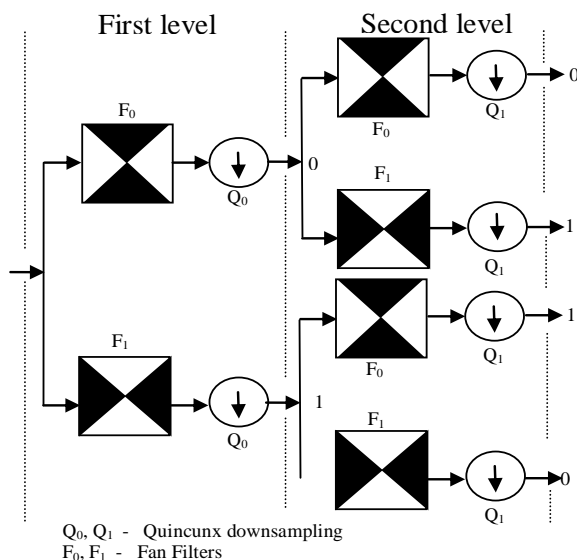


Figure 5.6 DFB decomposition blocks in two level directional filterbank of figure 5.6.

Chapter 5: Microcalcification Enhancement

The first building block, a two-channel quincunx filter bank with fan filters divides the two dimensional spectrum into two directions, horizontal and vertical. The sampling matrices, Q_0 for the first and Q_1 for second level is given as below.

$$Q_0 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad (5.1)$$

$$Q_1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (5.2)$$

The overall sampling after first and second level is $Q_0 Q_1 = 2I_2$ or downsampling by two in each dimension. The generation of four sub-bands after the first two levels is shown in figure 5.7 (a). The quadrant response at the second stage is obtained using multi rate identity [Do, 2001]. The multirate identity for swapping the filtering and downsampling operation is shown in figure 5.8. The fan filters at the second level with the sampling matrix Q_0 at the first level generates an equivalent filter with a quadrant frequency response. The fan filters at the first level with the filters with quadrant response generates the four directional sub-bands with wedge shape as shown in figure 5.7(b).

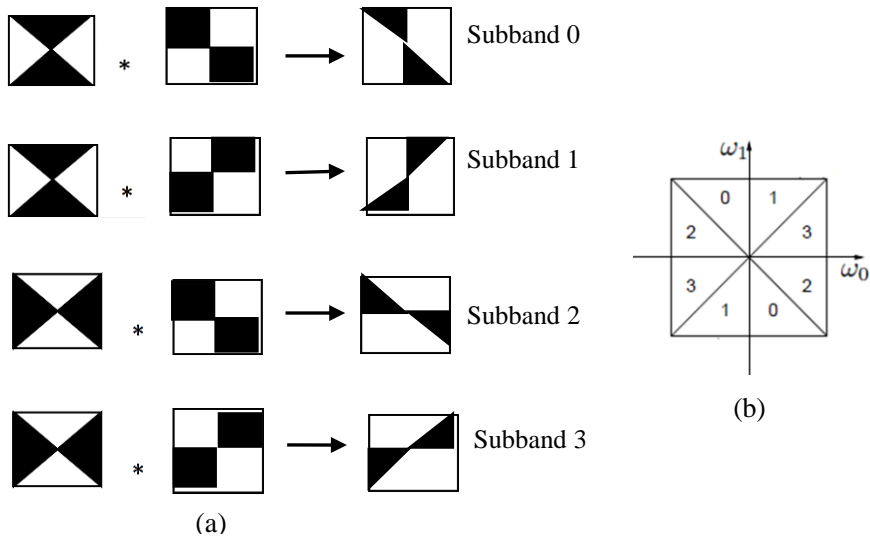


Figure 5.7: Directional subbands. (a) Impulse responses at first two level directional filter bank, (b) directional subbands of two level directional filter bank

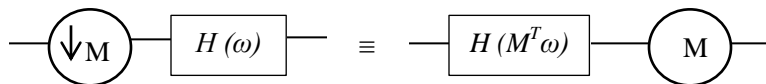


Figure 5.8: Swapping of filtering and downsampling operation using multirate identity^[Do, 2001].

Figure 5.9 represents a resampling operation on quincunx filter bank from third level of DFB in CT to obtain finer frequency division. The four unimodular matrix that are used for rotation or rearranging operation are R_0 ,

$$R_1 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad (5.3)$$

R_1 , R_2 and R_3 where

$$R_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad (5.4)$$

$$R_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad (5.5)$$

$$R_3 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \quad (5.6)$$

Another property of the resampling matrix is $R_0 \cdot R_1 \cdot R_2 \cdot R_3 = I_2$, where I_2 is a two dimensional identity matrix.

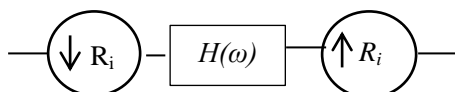


Figure 5.9: Quincunx filter bank with resampling operation from third level of DFB in CT^[Do, 2001].

Figure 5.10 shows the resampling operations for a mammographic region of mdb241 from MIAS database [Suckling, 1994] by resampling matrices R_0 , R_1 , R_2 and R_3 . Figure 5.11 illustrates the various steps involved in a three stage DFB as shown in figure 5.10. There are four different cases of QFB using resampling matrices R_0 , R_1 , R_2 and R_3 to obtain wedge shaped frequencies as shown in figure 5.11. A multidimensional multirate identity

Chapter 5: Microcalcification Enhancement

rule is applied from the third level onwards to obtain parallelogram filter response as shown in figure 5.12.

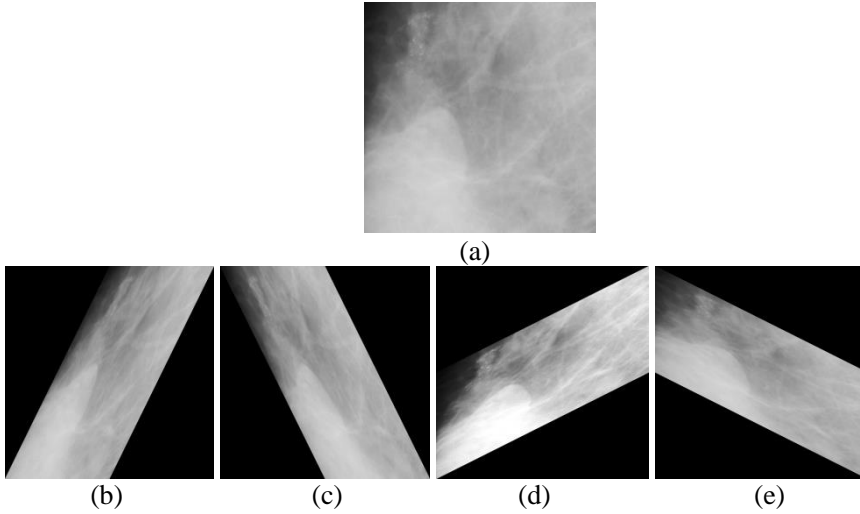


Figure 5.10: Resampling operation on (a) a mammographic region of mdb241. Resultant image by resampling matrix (b) R_0 , (c) R_1 , (d) R_2 and (e) R_4 .

According to smith form [Bamberger, 1992], the quincunx matrix can be represented as

$$Q_0 = R_1 D_0 R_2 = R_2 D_1 R_1 \quad (5.7)$$

$$Q_1 = R_0 D_0 R_3 = R_3 D_1 R_0 \quad (5.8)$$

where D_0 and D_1 are two dimensional diagonal matrix representing dyadic sampling.

$$D_0 = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \quad (5.9)$$

$$D_1 = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \quad (5.10)$$

Hence the overall sampling matrices P_0 and P_1 of figure 5.12 are obtained as given in equation 5.11 and 5.12.

$$P_0 = R_0 Q_0 = D_0 R_2 \quad (5.11)$$

$$P_1 = R_1 Q_1 = D_0 R_3 \quad (5.12)$$

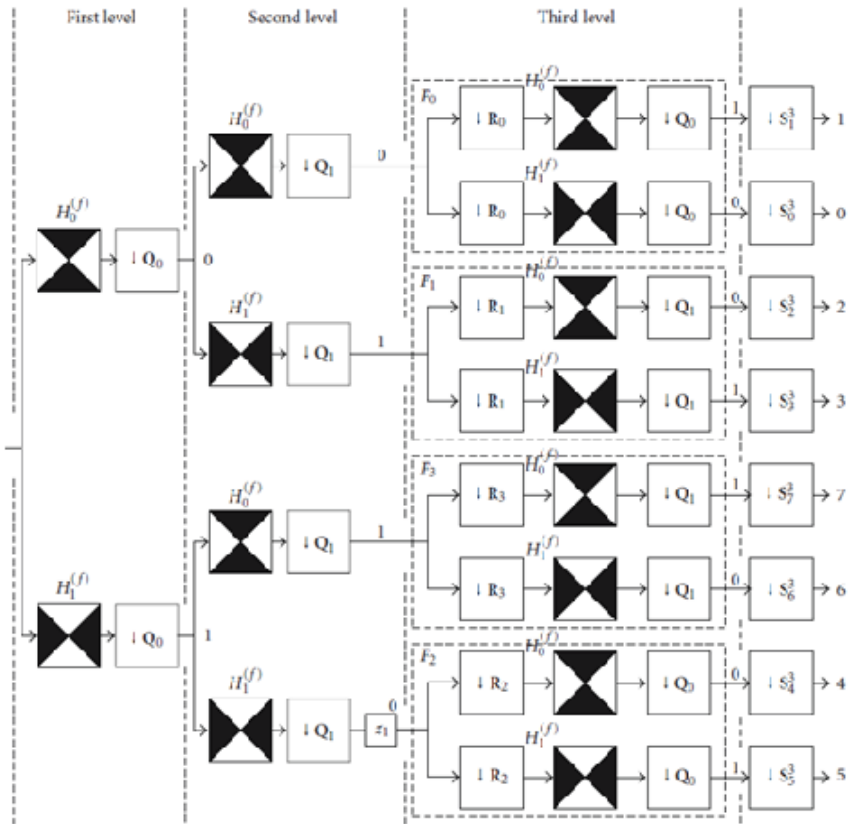


Figure 5.11: Block diagram for a three level DFB of contourlet transform ^[Do, 2001]

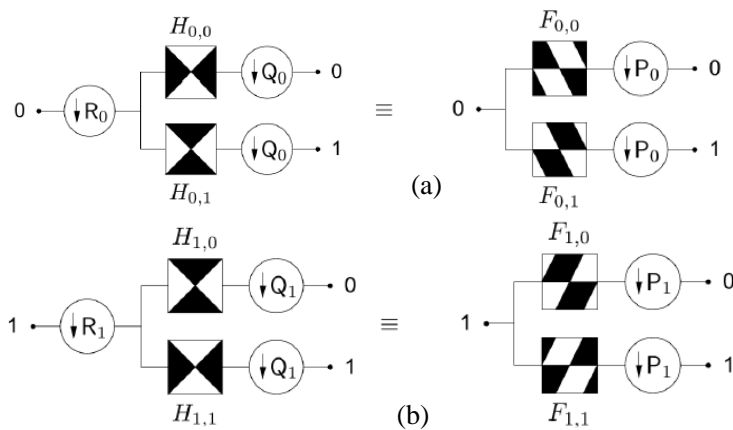


Figure 5.12: Resampled responses of quincunx filter bank on third level of DFB. Parallelogram filters obtained after applying (a) resampling matrix, R_0 and quincunx matrix, Q_0 (b) resampling matrix, R_1 and quincunx matrix, Q_1 ^[Do, 2001].

Chapter 5: Microcalcification Enhancement

Sinc type filters in CT provide the contourlet coefficients an infinite directional vanishing moment [Cunha, 2007]. Compared to WT, CT provides an optimal approximation of piecewise smooth signals due to the presence of directional vanishing moments [Do, 2001].

Efficiency of the proposed methods for enhancement is attributed to the preservation of directional information in CT. Impulse response of a bandpass arm at one level of contourlet decomposition constitutes the combined impulse responses of filters in LP as well as DFB. The direction of the combined impulse response is decided by the orientation of directional filters and the specified number of directional decomposition levels. The low pass filters employed at the LP stage are derived from the PKVA filters (Phoong, 1995) and the fan filters at the DFB stage are derived from the biorthogonal 9, 7 filters (Vetterli, 1992). The 9/7 filters in LP decomposition reduces interscale and interlocation mutual information of contourlet coefficients whereas PKVA filter in DFB is efficient in localizing edge direction by reducing interdirection mutual information of contourlet coefficients [Po, 2006].

Figure 5.13 depicts the combined impulse response of bandpass arm of LP and the 8th directional filter in the DFB, which is smooth enough, to be modelled as the second derivative of a smooth function, oriented in the direction of the impulse response of the corresponding directional filter.

Figure 5.14 shows the resultant contourlet coefficients using a single level LP with four level of directional decomposition. The resultant sub-bands are oriented in sixteen different directions with each providing information in the corresponding direction. Contourlet coefficients of mammographic images can be used to locate the sharp variation points due to MCC features.

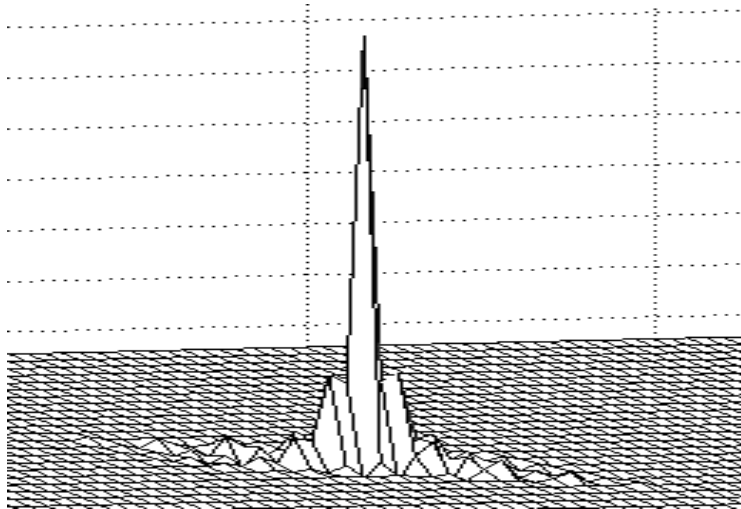


Figure 5.13: Combined impulse response of bandpass arm of LP and the 8th directional filter in the DFB

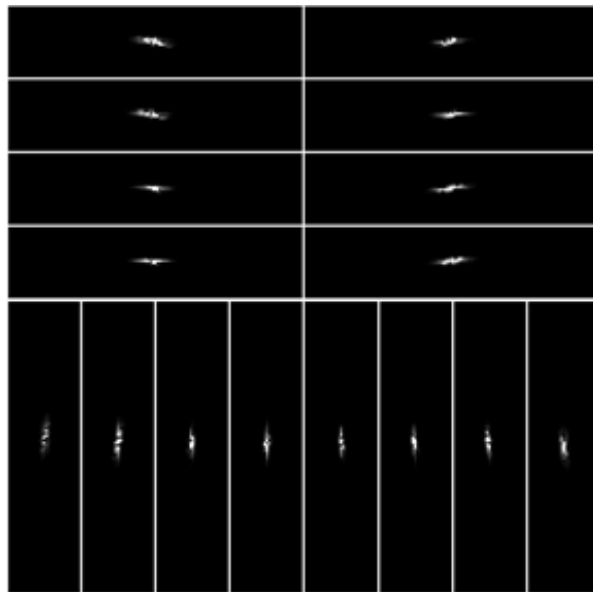


Figure 5.14: Impulse response of the sub-bands after 4 level directional decomposition of CT

5.3.2. Singularity detection of contourlet coefficients using zero crossings/ modulus maxima

MCCs which appear as bright spots in a mammogram are detected by finding the singularities or edges. The geometrical irregularities of an image such as edges are usually located along the smooth contours [Hubel, 1962]. Usually the sharp discontinuities in an image are obtained by finding its first or second order derivatives [James, 2011]. The modulus of the first derivative has a local maximum in accordance with the zero crossings of the second derivative and hence to the singularities [Gonzales, 2007]. Mallat [Mallat, 1991] has shown that the WT has a sequence of local maxima that converges to a point at a finer scale even though the function is regular at that point.

A simple method to find the sharp variations points are by finding the zero crossings/ local extrema of CT coefficients. The identification of zero crossings/ local extremas in the CT enable multiresolution implementations of the Marr-Hildreth [Marr, 1980] and canny edge detectors [Canny, 1986] respectively. For the case of identifying the zero-crossings of the CT, only those pixels whose contourlet coefficient values at a particular level having sign changes in the direction that correspond to the particular sub-band, are retained. The edge information is formed using these retained zero crossing coefficients. In a similar manner edge information from the modulus maxima points in the direction corresponding to the particular sub-band may also be formed. A point x_0 is modulus maximum if $|Cf(x)| \leq |Cf(x_0)|$ when x belongs to neighborhood of x_0 where $Cf(x)$ is the CT of function $f(x)$. In order to detect the contours of small structures as well as the boundaries of larger objects, several researchers in computer vision have introduced the concept of multiscale edge detection [Marr 1980], [Rosenfield 1971], [Witkin, 1980]. The bandpass image obtained after LP decomposition is partitioned into wedge shaped frequency sub-bands in various orientations by the DFB. Each sub band contains high

frequency information such as sharp intensity variation in the corresponding direction of orientation. While enhancing MCC features in the high frequency sub-bands, noise needs to be eliminated as it may interfere with the judgement by the radiologist. For this the relationships among the high frequency sub-band coefficients at various levels were exploited as described in the following section.

5.3.3. Suspicious MCC detection using parent-child relationship in CT

A parent-child relationship exists among contourlet coefficients along various scales similar to the spatial orientation tree concept of wavelet coefficients [Shapiro, 1993]. The parent-child relationship among strongly dependent coefficients at adjacent directions implemented using the DFB [Po 2006] is as follows: If the scale in which the children lie has twice as many directional sub-bands as the scale in which the parent lies, the four children will be in two adjacent directional sub-bands [Sudhakar, 2006]. Figure 5.15 shows the case when the number of directional sub-bands considered is the same at two adjacent levels of decomposition of the LP.

In this example, a four level DFB was constructed at every stage of the LP decomposition, resulting in sixteen directional sub-bands at each resolution. Each of the sub-bands will contain information about edge features that are oriented in a particular direction. If there is a significant edge feature in a coarse sub-band, its children in the finer sub bands will also be significant. Clinically important MCC structures form edge features are present at more than one level of resolutions in the contourlet decomposition whereas edges due to image artifacts are very fine structures that are significant only at the finest resolutions. Accordingly, a zero crossing/ modulus maxima tree is constructed based on the parent-child relationship among pixels as in figure 5.15.

Chapter 5: Microcalcification Enhancement

The zero crossings/ modulus maxima in each directional sub-band are found out only in the direction of orientation of the impulse response of the filter, to be included in the zero crossing/ modulus maxima trees. The tree retains only those zero crossings/ modulus maxima that propagate to the coarser levels of the contourlet decompositions, while discarding those zero crossings/ modulus maxima that exist only at the finer resolutions. The coefficients in the sub-band images are selectively weighted to enhance the calcifications after considering the parent-child relationship. The isolated single pixel intensity changes, i.e., noise, are eliminated by retaining only those zero crossings/modulus maxima that hold a parent– child relationship from the coarsest level to the finest level of detail. These coefficients are added with decomposed coefficients before reconstruction. Thus if either a local modulus maxima or zero crossing is present at a coarse resolution, all child local modulus maxima or zero crossings will be searched for in the immediate finer resolution. If such features are detected, they will be retained, and other modulus maxima or zero crossings that do not propagate to the finest resolutions will be pruned. Thus the significant coefficients among directional sub-bands that potentially correspond to MCC features are retained and very minute structures attributed to noise are removed. The coefficients that correspond to significant edge features including MCC features are enhanced. The inverse CT step is performed now with the modified directional sub-band coefficients to obtain an enhanced image.

The features of an edge are described differently in different directions and at different resolution scales. This is decided by the level of the LP at which the DFB decomposition is performed. A three level LP decomposition with four level DFB decomposition in each scale was performed on the mammographic image as it was observed that further decompositions to coarser levels of the LP failed to capture MCC features.

It was observed that enhancement using the local modulus maxima outperformed that with the local zero crossings. The maxima of the absolute value of the first derivative are sharp variation points, whereas the minima

correspond to slow variations. Zero crossings give position information but do not differentiate small amplitude fluctuations from important discontinuities.

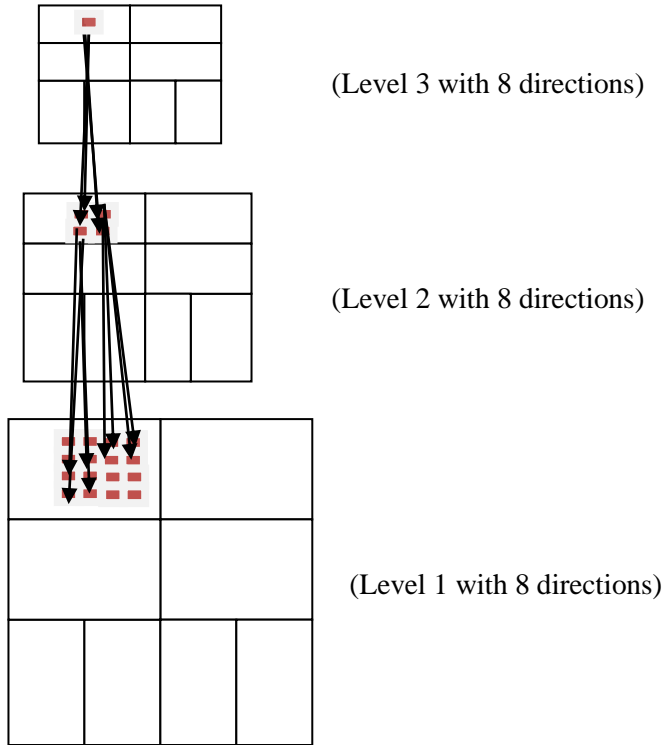


Figure 5.15: Parent-child relationships among the CT coefficients

5.4. Performance measures

The performance measures for analyzing the enhancement of breast region to identify MCC features while removing very tiny structures corresponding to noise are contrast [Morrow, 1992], contrast improvement index (*CII*) [Laine, 1994], target to background contrast (TB_C), [Xinsheng, 2009] and tenengrad criterion [Chen, 2006]. The definitions of *CII* and contrast are defined as,

$$CII = \frac{C_{enhanced}}{C_{original}} \quad (5.13)$$

Chapter 5: Microcalcification Enhancement

where $C_{enhanced}$ and $C_{original}$ are the contrasts for a mammogram in the enhanced and original images, respectively.

The contrast, C of a region is defined by

$$C = \frac{(\max_I - \min_I)}{(\max_I + \min_I)} \quad (5.14)$$

where \max_I is the maximum luminance in the image and \min_I is the minimum luminance. In order to evaluate the visual appearance of the enhanced image, a quantitative measure, target to background Contrast Ratio using Variance (TB_c) is considered, computed as below,

$$\delta_\mu = \left(\frac{\mu_f^e}{\mu_b^e} \right) - \left(\frac{\mu_f^o}{\mu_b^o} \right) \quad (5.16)$$

where δ_μ is the difference between the ratios of the mean gray in the foreground and background, μ_f^e , μ_b^e , μ_f^o , μ_b^o represents mean gray-level value in the foreground of enhanced image, mean gray-level value in the background of enhanced image, mean gray-level value in the foreground of original image, mean gray-level value in the background of original image, and

$$\sigma = \frac{\sigma_f^e}{\sigma_f^o} \quad (5.17)$$

where σ_f^o and σ_f^e are the standard deviations of the target in the original and enhanced image.

The sharpness of the MCC enhanced mammogram image is also computed using the tenengrad criterion, which is based on gradient magnitude maximization. It is considered as one of the most robust and functionally accurate image quality measures [Chen, 2006]. The tenengrad value of an image, I is calculated from the gradient $\Delta I(x, y)$ at each pixel (x, y) , where the partial derivatives are obtained by a high-pass filter, e.g.,

the sobel operator, with the convolution kernels i_x and i_y . The gradient magnitude is given as

$$S(x, y) = \sqrt{(i_x \otimes I(x, y))^2 + (i_y \otimes I(x, y))^2} \quad (5.18)$$

and the tenengrad criterion is formulated as

$$TEN = \sum_x \sum_y S(x, y)^2 \quad (5.19)$$

for $S(x, y) > T$ where T is a threshold. The image quality is usually considered higher if its Tenengrad value is larger.

5.5. Results and discussion

The database used for the proposed work is the MIAS mammographic database (Suckling, 1994). The MIAS database contains 322 images with 25 images having MCC features, with 13 cases diagnosed as malignant and 12 as benign. The proposed algorithm was tested on the 25 images with MCCs and the performance measures were computed.

The performances of the enhancement techniques using zero crossings/ modulus maxima of CT were compared. The low pass filters employed at the LP stage are derived from the PKVA filters [phoong, 1995] and the fan filters at the DFB stage are derived from the BIOR 9, 7 filters [Vetterli, 1992]. Three levels of the LP with four levels of the directional filter bank (i.e., 16 directional sub-bands) at each level of the LP decomposition were employed for the contourlet decomposition of the mammogram images. The PKVA filter is designed by Phoong, Kim, Vaidyanathan, and Ansari with support size of (23, 23) and (45, 45). The directional sub-band coefficients on either side of the selected modulus maxima are boosted by a factor of four. Only those elements in the directional sub-bands that lie on either side of the selected modulus maxima in the appropriate direction (decided by the orientation of the directional filter for that sub-band) are retained. In the figures, 5.16, 5.17 and 5.18, the

Chapter 5: Microcalcification Enhancement

results with enhancement of the MCC features on various images from the MIAS mammographic image database are shown.

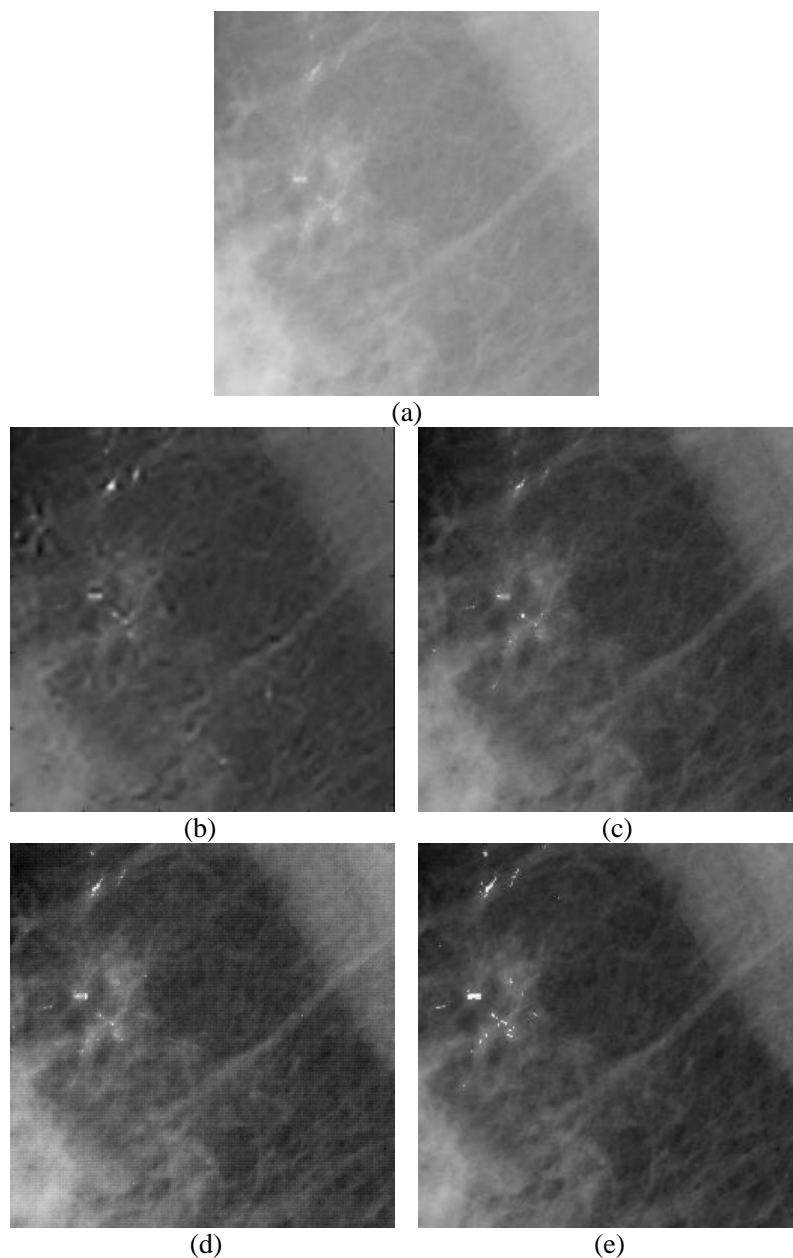


Figure 5.16: Enhanced result of MDB 249. (a) Original. Enhanced results using (b) zero crossings coefficients of WT, (c) zero crossings coefficients of CT, (d) modulus maxima coefficients of WT and (e) modulus maxima coefficients of CT.

Chapter 5: Microcalcification Enhancement

The enhancement using zero crossing/ modulus maxima features of the discrete WT coefficients, with the BIOR 6.8 wavelet and CT are shown for comparison. Figure 5.16 shows enhanced results of mdb 249 using singularities of WT and CT. Figure 5.17 shows enhanced results of mdb 241 based on zero crossings of WT and CT along with modulus maxima of WT and CT. The enhanced results in figure 5.16 and 5.17 shows that the visibility of MCC is more in the method based on modulus maxima of CT compared to that of zero crossing of CT. The visibility of MCC is more for methods based on CT compared to that of corresponding approaches on WT. Figure 5.18 provides a graphical plot of various performance measures such as Tbc , contrast, contrast improvement index and TEN for few enhanced mammographic images based on zero crossings and modulus maxima of wavelet coefficients and contourlet coefficients. Visual quality of the enhanced images for the proposed approaches is obtained using the Tbc parameter. The contrast obtained in the reconstructed images is compared using the contrast and contrast improvement index (CII) parameters. The TEN value, a sharpness measure, which is found useful in medical images is found to be more in enhanced images compared to the original mammographic images. There is a significant improvement in the values of Tbc and contrast with the proposed method using modulus maxima technique, indicating that the enhanced image is visually superior with a larger contrast, to those with the method to which it was compared. The improvement in sharpness of modulus maxima technique is verified through the analysis of tenengrad.

The proposed methods are effective in enhancing the regions of MCCs. These regions were formed by replacing the coarse sub image by an all-zero array in the reconstruction step. The enhanced locations of malignancy are clearly visible on the reconstructed image with suppressed approximation coefficients. A global thresholding operation is performed on the reconstructed image that is obtained after suppressing the coarse sub image in the reconstruction by the contourlet synthesis filter bank to

Chapter 5: Microcalcification Enhancement

indicate the regions that have been enhanced. The high intensity MCCs above the threshold value can be highlighted from the rest of mammographic regions

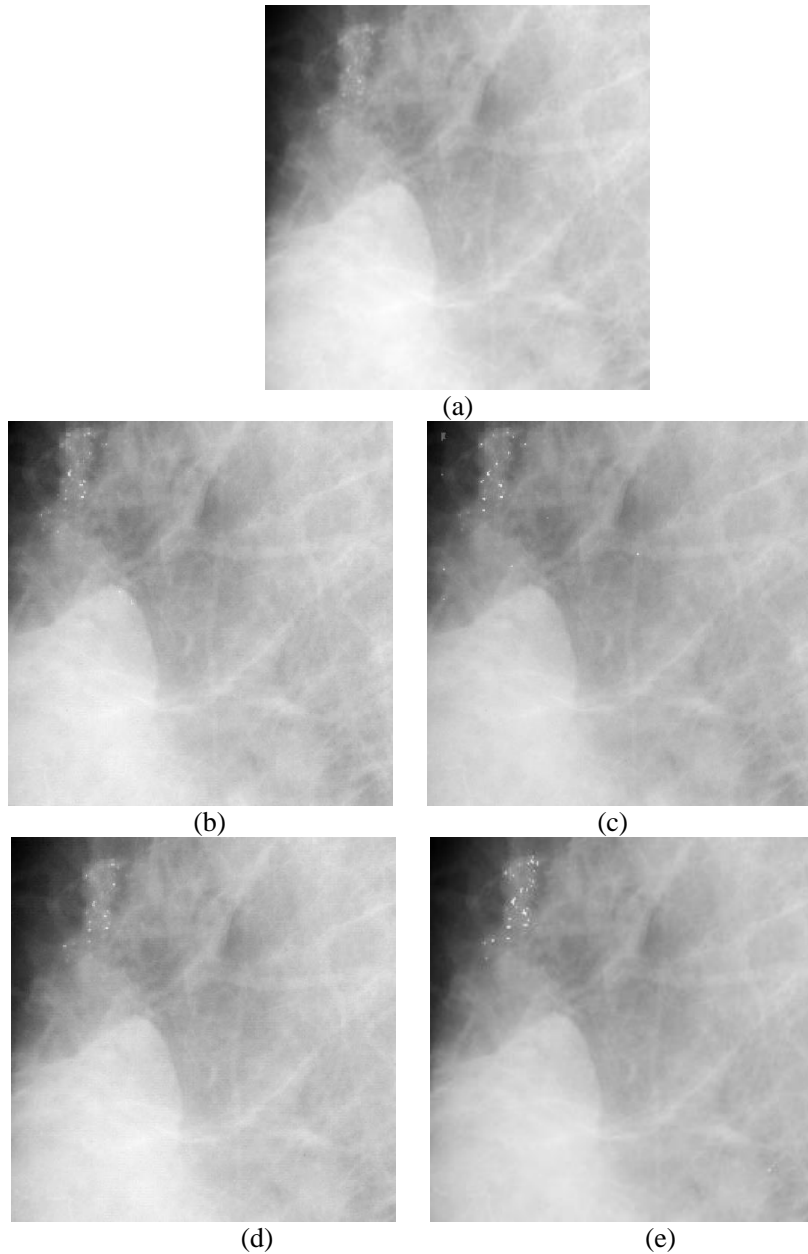
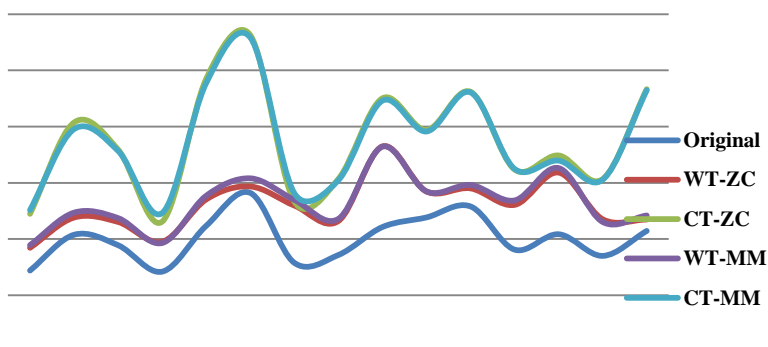


Figure 5.17: Enhanced result of MDB 241. (a) Original. Enhanced results using (b) zero crossings coefficients of WT, (c) zero crossings coefficients of CT and (d) modulus maxima coefficients of WT and (e) modulus maxima coefficients of CT.

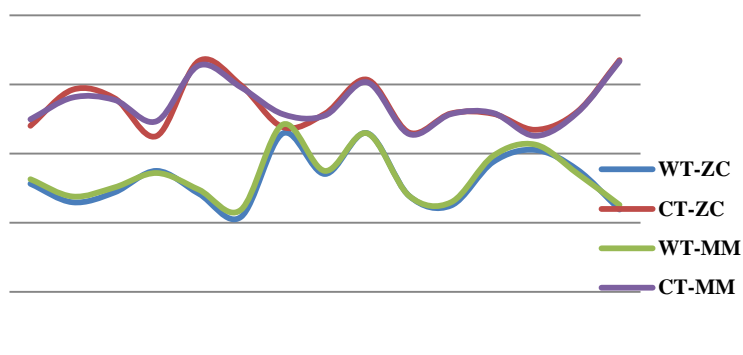
Chapter 5: Microcalcification Enhancement

The images in figure 5.19, 5.20 and 5.21 indicate the regions where the enhancement has taken place. Figure 5.19, 5.20 and 5.21 shows the results of segmented MCCs in mdb 241, mdb 249 and mdb 245 using zero crossings of WT and zero crossings/ local extremas of CT respectively.

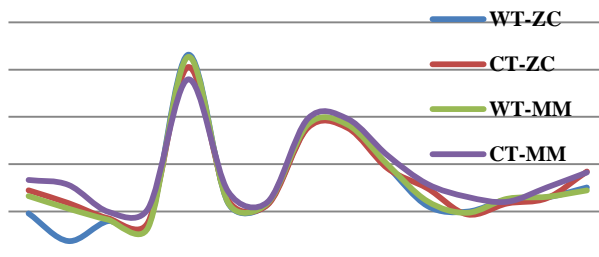
Even though the segmented MCCs appears to lie within the ground truth circle represented by red color in figure 5.19 and 5.20, more MCCs within the malignant area is captured by the local extrema method using CT. Even though the ground truth information of distributed MCCs in mdb 245 is not provided in the database, the local extrema method is found successful in extracting most of the MCCs in the region as verified by the radiologist.



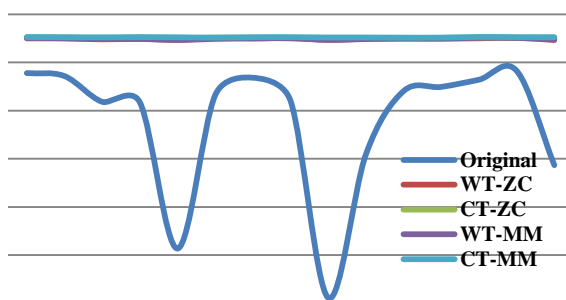
(a) Contrast



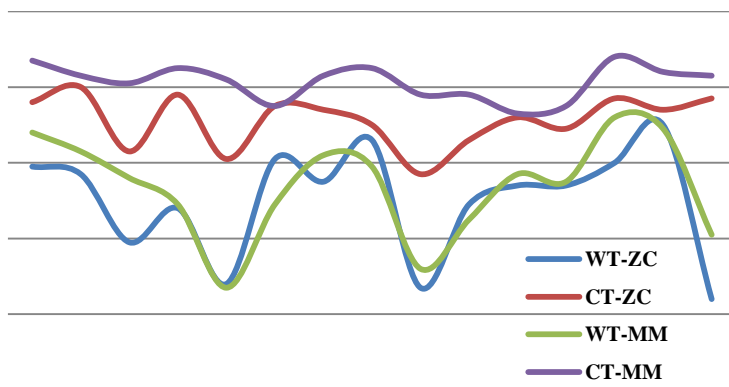
(b) Contrast Improvement index (CII)



(c) Target to Background Contrast Ratio (TBC)



(d) Tenegrad Criterion (TEN)



(e) Tenegrad Criterion (TEN)

Figure 5.18: Comparison of performance measures to evaluate contrast, visual quality and sharpness of enhanced mammographic images using zero crossings and modulus maxima on wavelet and contourlet coefficients. (a) contrast, (b) contrast improvement index, (c) target to background contrast ratio, (d) tenegrad criterion and (e) expanded plot of (d) excluding original.

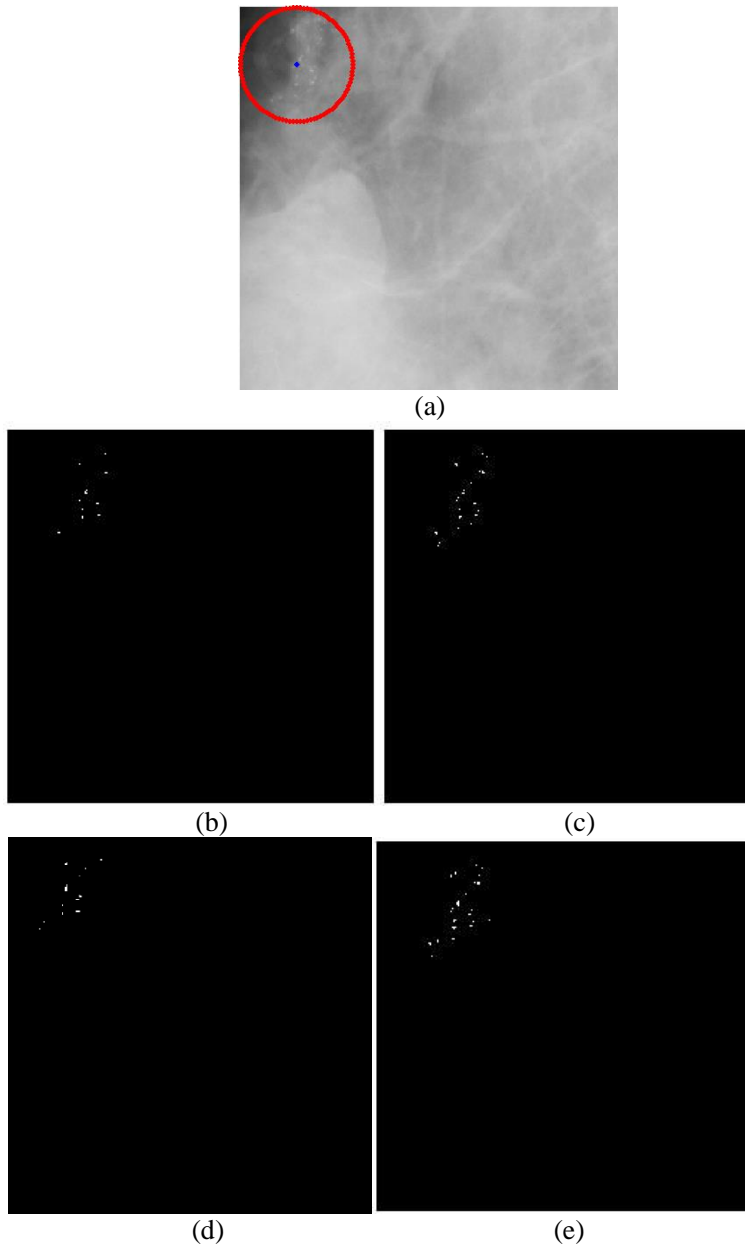


Figure. 5.19: Segmented MCCs. obtained for (a) mdb241 using (b) zero crossings of WT, (c) zero crossings of CT, (d) modulus maxima of WT and (e) modulus maxima of CT. Red circle shows ground truth information.

Chapter 5: Microcalcification Enhancement

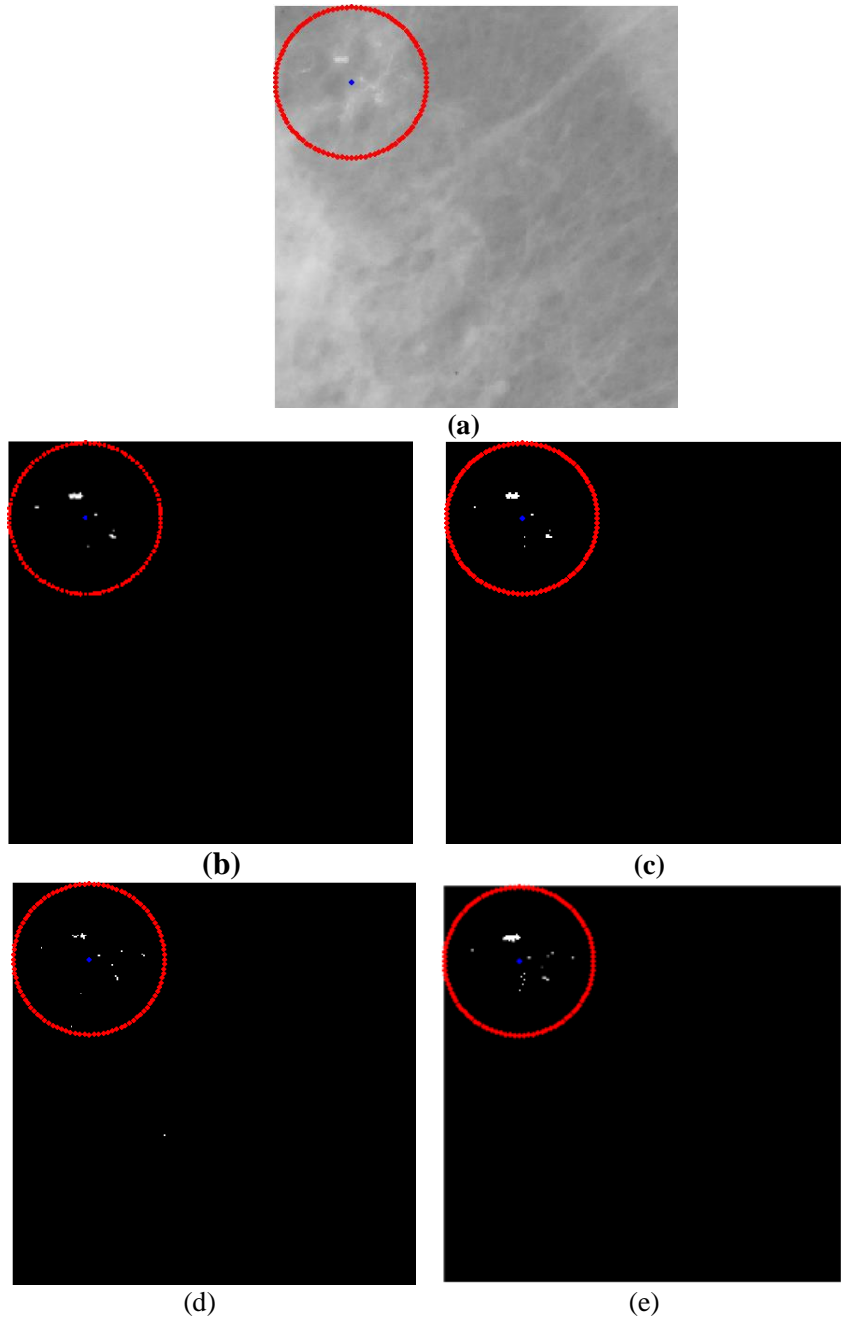


Figure. 5.20: Segmented MCCs. obtained for (a) MDB 249 using (b) Zero crossings of WT, (c) zero crossings of CT, (d) modulus maxima of WT and (e) modulus maxima of CT. Red circle shows ground truth information.

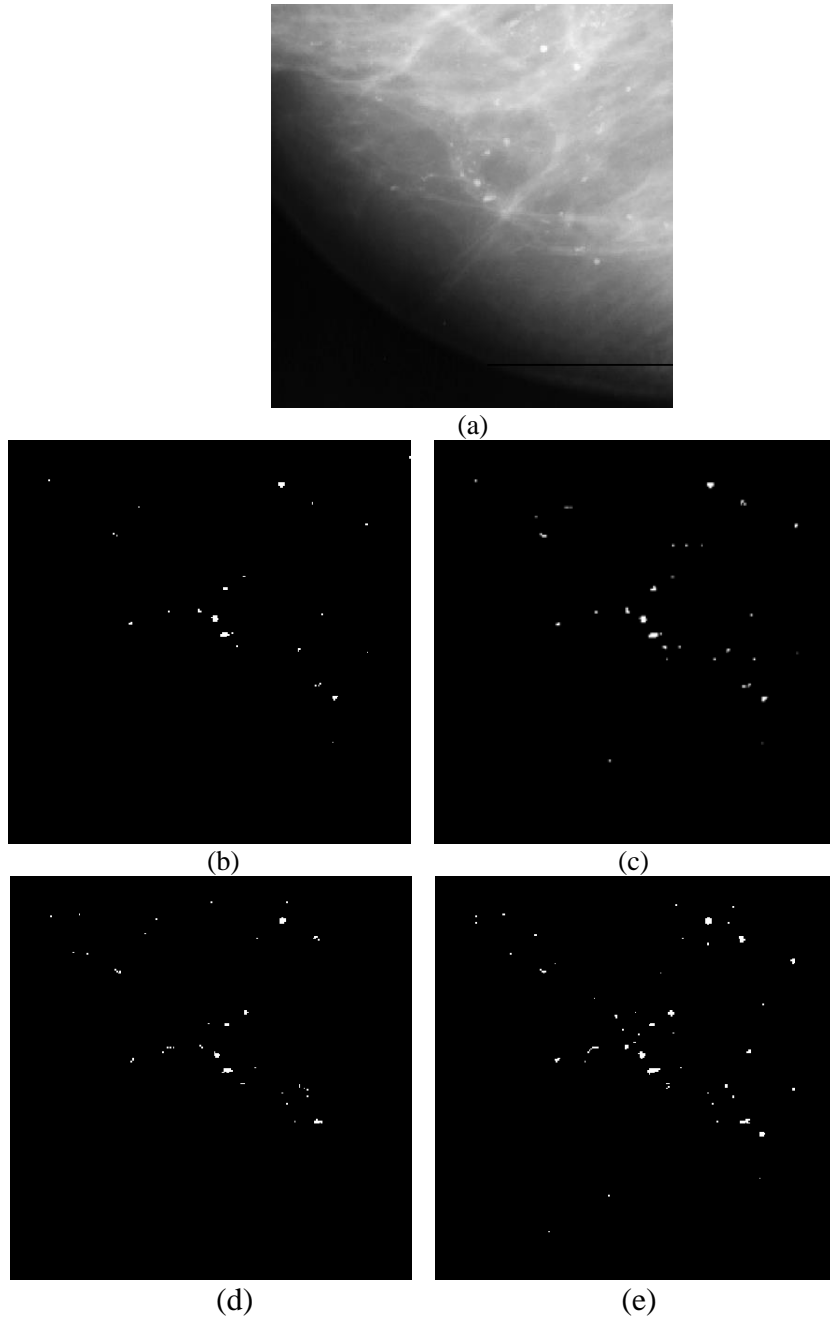


Figure. 5.21: Segmented MCCs. obtained for (a) mdb245 using (b) zero crossings of WT, (c) zero crossings of CT, (d) modulus maxima of WT and (e) modulus maxima of CT.

Chapter 5: Microcalcification Enhancement

Table 5.1 illustrates a comparison of contrast improvement Index (CII) for the proposed methods with that of various methods proposed by Raj [Raj Kumar, 2011] and Arya [Arya, 2015]. It shows the variation of CII across various methods including proposed techniques used for images in MIAS as well as DDSM database. It is found that the proposed technique using modulus-maxima of CT techniques outperforms the other methods.

Table 5.1: Comparison of contrast improvement index (CII) in various methods for MIAS and DDSM database

Top Hat [Raj Kumar, 2011], (MIAS database)	1.027
Wavelet decomposition (Sure Shrink) [Raj Kumar, 2011], (MIAS database)	1.165
Top Hat +Sure Shrink[Raj Kumar, 2011], (MIAS database)	1.242
Top Hat + level dependent Wavelet Shrink[Raj Kumar, 2011], (MIAS database)	1.203
Shrink 1.203 Top Hat +Visual Shrink[Raj Kumar, 2011], (MIAS database)	1.186
Top Hat +level dependent Visual Shrink[Raj Kumar, 2011], (MIAS database)	1.188
Top Hat +modified level independent Visual Shrink[Raj Kumar, 2011], (MIAS database)	1.185
Top Hat +Bit Plane decomposition[Raj Kumar, 2011], (MIAS database)	1.069
Stationary WT (SWT) [Arya, 2015], (MIAS database)	1.64
CT – Zero Crossing [proposed 1], (MIAS database)	1.84
CT – Modulus Maxima [proposed 2], (MIAS database)	1.85
Modified Homomorphic Filtering based on background noise [Ravi, 2013], (DDSM database).	1.1085
CT – Zero Crossing [proposed1], (DDSM database)	1.39447
CT – Modulus Maxima [proposed2], (DDSM database)	1.401979

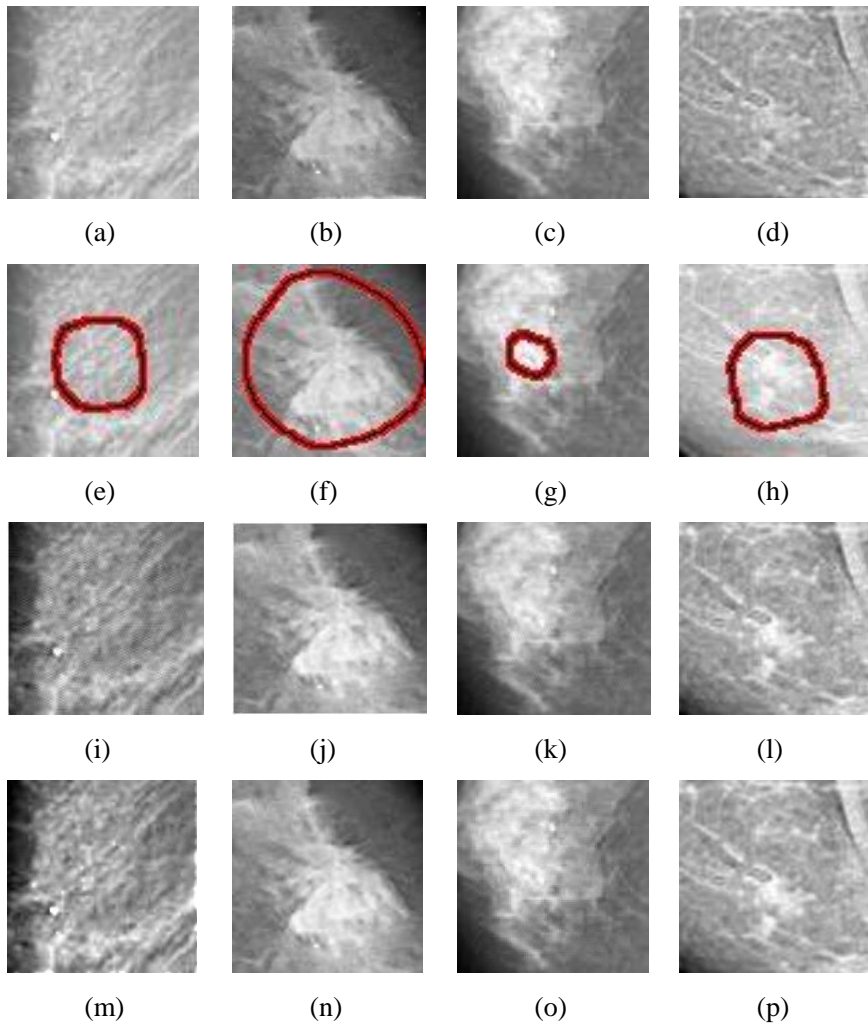


Figure. 5.22: Enhanced MCCs. obtained for mammograms including (a) 4710_rmlo, (b) 1236_rcc, (c) 1152_lcc, (d) 4103_rcc using (i- l) zero crossings of CT and (m-p) modulus maxima of CT. The ground truth information of corresponding mammograms marked by radiologist is shown in (e-h).

The algorithms were also tested on images obtained from the DDSM database. However statistics for enhancement of MCC features in mammograms have not been reported with the DDSM database other than for the parameter CII. For the MIAS database, the modulus maxima method achieved an improvement of 17.9% in visual appearance, 11.3% in contrast

Chapter 5: Microcalcification Enhancement

and CII, 21.7% in noise reduction and 0.26% in sharpness compared to that of zero crossings method. For the DDSM database, the modulus maxima method achieved 0.29% improvement in visual appearance, 0.71 % increase in contrast and 4.57% in CII, 0.29% improvement in noise reduction and improvement of 0.29% in sharpness compared to that of zero crossings method.

5.6. Summary

The proposed methods for enhancing MCCs in mammographic images show promising results. The images from MIAS database are employed to validate the results. The processed mammographic images enhanced using the modulus maxima/ zero crossings of CTs are significantly better in visual appearance, noise reduction, contrast and sharpness compared to the other methods.

5.7. Bibliography

Hajar, Moradmand., Saeed, Setayeshi., Ali, Reza, Karimian., Mehri, Sirous., Mohammad, Esmail, Akbari. (2012). Comparing the performance of image enhancement methods to detect microcalcification clusters in mammography. *Iran J Cancer*, 5(2), 61–68.

Arianna, Mencattini., Marcello, Salmeri., Roberto, Lojacono., Manuela, Frigerio., Federica, Caselli. (2008). Mammographic Images Enhancement and Denoising for Breast Cancer Detection Using Dyadic Wavelet Processing, *IEEE Transactions on Instrumentation and Measurement*, 57(7).

Arya, Devi, P. S., & M. G. Mini.(2015). Mammographic Image Enhancement Based on SWT and High Boost Filtering, *International Journal of Computer Theory and Engineering*, 7(5).

Bagci, A. M. & A. E. Cetin. (May 2002). Detection of microcalcification clusters in mammogram images using local maxima and adaptive wavelet transform analysis. in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '02)*, 3856–3859, Orlando, Fla, USA.

Balakumaran, T., Ila, Vennila. & C.Gowri, Shankar. (2010). Detection of Microcalcification in Mammograms Using Wavelet Transform and Fuzzy shell Clustering, *International Journal of Computer Science and Information Security*, 7(1), 121-125.

Chapter 5: Microcalcification Enhancement

Bamberger, R. H. & M. J. T. Smith. (1992). A filter bank for the directional decomposition of images: Theory and design, *IEEE Trans. Signal Proc.*,40(4), 882–893.

Bao-Ning, Zhang., Xu-Chen, Cao., Jia-Yi, Chen., Jie, Chen., Li, Fu., Xi-Chun, Hu., Ze-Fei, Jiang., Hong-Yuan, Li., Ning, Liao., Dong-Geng, Liu., Ouyang, Tao., Zhi-Min, Shao., Qiang, Sun., Shui, Wang., Yong-Sheng, Wang., Bing-He, Xu. & Jin, Zhang. (2012). Guidelines on the Diagnosis and Treatment of Breast Cancer. *Gland Surgery* 1.1, 39–61.

Burt, P.J., Adelson, E.H. The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, 31(4), 532-540, 1983.

C. H. Chen., & G. G. Lee. (1997). On Digital Mammogram Segmentation and Microcalcification Detection Using Multiresolution Wavelet Analysis, *Graphical models and image processing*, 59(5), 349–364.

Canny, J. (1986).A computational approach to edge detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8, 679-714.

Chao, Li., Kin, Man, Lam., Lei,Zhang., Chun, Hui. & Su, Zhang. (12-15 Aug. 2012). Mammogram microcalcification cluster detection by locating key instances in a Multi-Instance Learning framework.*IEEE International Conference in Signal Processing, Communication and Computing (ICSPCC)*, 175-179.

Chen, Zhi Yu., Bisma, R., Abidi., David, L. Page., & Mongi, A. Abidi. (2006) Gray-Level Grouping (GLG): An Automatic Method for Optimized Image Contrast Enhancement—Part I: The Basic Method.

Chapter 5: Microcalcification Enhancement

IEEE Transactions on image processing, 15 (8).

Cornelius, T. Leondes. (2005). Medical Imaging Systems Technology: Methods in diagnosis optimization, World Scientific.

Cunha, Da., A.L., Do, M.N. (2007). On Two-Channel Filter Banks With Directional Vanishing Moments, IEEE Transactions on Image Processing, 16(5), 1207-1219.

Do, M. N. & M. Vetterli. (2005). The contourlet transform: an efficient directional multiresolution image representation, IEEE Transactions on Image Processing, 14 (12), 2091-2106.

Gonzalez, Rafael, C. & Richard, E. Woods. (2007). Digital Image Processing (3rd Edition).

Hashemi, Eliza. Aghjekandi. (2012). Microcalcification Detection in Mammography using Wavelet Transform and Statistical Parameters, Thesis.

Highnam, R., Brady, J.M. (1999). Mammographic Image Analysis, Computational Imaging and Vision.

Hubel, D. H. & T. N. Wiesel. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, J. Physiology, 160, 106–154.

Jagannath, H.S., J. Virmani., & V. Kumar, Morphological Enhancement of Microcalcifications in Digital Mammograms, J. Inst. Eng. India Ser. B (September–November 2012) 93(3):163–172.

Chapter 5: Microcalcification Enhancement

James, Park., Han-Chieh, Chao., Mohammad, S. Obaidat., & Jongsung, Kim. (2011), Computer Science and Convergence: CSA 2011 & WCC 2011 Proceedings, Volume 114 of Lecture Notes in Electrical Engineering, Springer Science & Business Media.

Kim, J. K., Park, J. M., Song, K. S., & Park, H. W. (1997). Adaptive mammographic image enhancement using first derivative and local statistics, *IEEE Trans Med Imaging*, 16 (5), 495-502.

Lain, F., S. Schuler., J. Fan. & W. Huda. (1994). mammographic Feature Enhancement by Multiscale Analysis, *IEEE Trans. On Med. Imaging*, 13(4), 725-740.

Le Gal, M., J. C. Durand., M. Laurent., & D. Pellier. (1973). Conduite a tenir devant une mammographie revelatrice de microcalcifications groupees sans tumeur palpable, *La Nouvelle Presse Medicale*, 5, 1621–1627.

M. N. Do. (2001). Directional Multiresolution Image Representations, Department of Communication Systems, Swiss Federal Institute of Technology Lausanne, Ph.D. Thesis.

M. Sundaram, K. Ramar, N. Arumugam. & G. Prabin. (2014). Efficient Edge Emphasized Mammogram Image Enhancement for Detection of Microcalcification. World Scientific.

Mallat, S. & S. Zhong. (1992). Characterization of signals from multiscale edges, *IEEE Trans on Pattern Analysis and Machine Intelligence*, 14(7), 710-732.

Chapter 5: Microcalcification Enhancement

Mallat, S. (1991). Zero-Crossings of a Wavelet Transform, *IEEE Transactions on Information Theory*, 37(4).

Marr, D., & Hildreth, E. (29 Feb 1980), "Theory of Edge Detection", *Proceedings of the Royal Society of London, Series B, Biological Sciences* 207 (1167): 187–217.

Marrocco, C., Molinara, M., Tortorella, F., Rinaldi, P., Bonomo, L., Ferrarotti, A., Aragno, C., Schiano, I., Moriello, S. (20-22 June 2012). Detection of cluster of microcalcifications based on watershed segmentation algorithm. *25th International Symposium on Computer-Based Medical Systems (CBMS)*, 1-5.

Mohanalin, Beenamol., Prem Kumar Kalra., & Nirmal, Kumar. (2010). A novel automatic microcalcification detection technique using Tsallis entropy & a type II fuzzy index, *Computers & Mathematics with Applications*, Science Direct, Elsevier, 60(8), 2426–2432.

Peter, Heinlein, Johann, Drexl. & Wilfried, Schneider. (2003). Integrated Wavelets for Enhancement of Microcalcifications in Digital Mammography, *IEEE Transactions on Medical Imaging*, 22(3), 402-413.

Phoong, S.M., C. W. Kim., P. P. Vaidyanathan. & R. Ansari. (1995). A new class of two-channel biorthogonal filters banks and wavelet bases. *IEEE Transactions on Signal Processing*, 43(3), 649–665.

Po, D. D. Y. & M. N. Do. (2006). Directional multiscale modeling of images using the contourlet transform, *IEEE Transactions on Image Processing*, 15(6).

Chapter 5: Microcalcification Enhancement

Raj Kumar. K. K. & Raju. (2011). Enhancement of mammograms using tophat filtering and wavelet decomposition. *Journal of Computer and Mathematical Sciences*, 2(6), 780-898.

Ravi, M., Babu, E., K. Nagaiah, E. (2013). Advanced Enhancement Method for Micro Calcification in Mammography. *International Journal of Science and Research (IJSR)*, 2(3).

Rosenfield, A., & Thurston, M. (1971). Edge and curve detection for visual scene analysis, *IEEE Trans. Comput.*, C-20, 562-569).

Sampat, M. P., M. K. Markey., & A. C. Bovik. (2005). Computer-aided detection and diagnosis in mammography. *Handbook of Image and Video Processing*, A.C. Bovik: Editor, 2nd edition. New York: Academic, 1195–1217.

Shapiro, Jerome, M. (1993). Embedded Image Coding Using Zerotrees of Wavelet Coefficients, *IEEE Trans. on Signal Processing*, 41(12), 3445-3459.

Strickland, R. N., Hee, Il, Hahn. (1996). Wavelet transforms for detecting microcalcifications in mammograms, *IEEE Trans. on Medical Imaging*, 15(2), 218 – 229.

Suckling, J. (1994). The Mammographic Image Analysis Society Digital Mammogram Database *Exerpta Medica. International Congress Series*, 1069, 375-378.

Chapter 5: Microcalcification Enhancement

Tay, P.C., Hongda, Shen., Adams, R.D., Zhang, J. Z. (15-18 Sept. 2013). A microcalcification enhancement method for mammogram images. 20th IEEE International Conference on Image Processing (ICIP), 232-236.

Ted, C. Wang., & Nicolaos, B. Karayiannis. (1998). Detection of Microcalcifications in Digital Mammograms Using Wavelets, IEEE Transactions on Medical Imaging, 17(4).

Tomislav, Bujanovic., Ikhlas, Abdel-Qader. (2013). On Wavelet Transform General Modulus Maxima Metric for Singularity Classification in Mammograms. Open Journal of Medical Imaging, Scientific Research, 3(1), 17-30.

Tomislav, Stojić. & Branimir, Reljin. (2010). Enhancement of Microcalcifications in Digitized Mammograms: Multifractal and Mathematical Morphology Approach, FME Transactions, 38(1), 1-9.

Vetterli. M. & C. Herley. (1992). Wavelets and filter banks: Theory and design. IEEE Transactions on Signal Processing, 40(9), 2207–2232.

Wael, A. Mohamed., Mohamed, A. Alolfe. & Yasser, M. Kadah. (2008). Microcalcification enhancement in digital mammograms using fractal modeling, in Proceedings of the 4th Cairo International Biomedical Engineering Conf., Cairo, Egypt.

Witkin, Andrew, P. (1984), Scale-space filtering: A new approach to multi-scale description, ICASSP, 150-153.

Chapter 5: Microcalcification Enhancement

Xinsheng, Zhang., & Hua, Xie. (2009). Mammograms Enhancement and Denoising Using Generalized Gaussian Mixture Model in Nonsampled Contourlet Transform Domain. *Journal Of Multimedia*, 4(6).

Zouari, M., Masmoudi, A.D., Masmoudi, D.S. (17-19 March 2014). A non-linear stretching image enhancement technique for microcalcification detection. 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 193-197.

Chapter 6

Concluding remarks and future work

The concluding remarks of various computer aided techniques discussed in earlier chapters for PM boundary delineation, MCC enhancement and AD detection and the scope of further work are presented in this chapter.

6.1. Concluding remarks

Breast cancer stands as the second most threatening cancer among women in India. Detection of breast cancer in its early stage responds more effectively to treatments which in turn reduce the mortality rate. Mammography, the gold standard as a screening tool is very helpful for the radiologist in providing an expert opinion. A second opinion using mammograms and computer aided techniques can assist the radiologists in providing accurate opinion thereby avoiding unnecessary biopsies. The GLOBOCON statistics clearly indicate that the lack of timely screening test in developing countries is the reason for high rate of mortality even though the incidence rate is less than that of developed countries [Ferlay, 2014]. As a result of this research, a database of mammograms for the Indian populace could be created with the help of mammogram images collected from the Lakeshore hospital. For evaluating the results, standard databases used by researchers worldwide, the MIAS database and the DDSM database were employed.

The goal of the proposed thesis is to develop computer aided detection techniques to assist radiologists in providing a second opinion for identifying two major anomalies of breast cancer: MCC, the most commonly occurring anomaly and AD, the most commonly missed abnormality. The delineation of the PM region, so as to extract the breast region in the mammogram, is also addressed.

The anatomy of breast, various breast imaging modalities, and the need for computer aided techniques for detection of breast cancer symptoms were discussed. The different types of breast cancer, incidence and mortality statistics, causes and the Bi-rads categories were also described.

Two novel methods for detecting the PM boundary coordinates were investigated in this thesis. The difference in two methods lay in the extraction of strong edges including the PM boundary. The contours of the

homogeneous region were applied to generate the strong edges from an artifact removed mammogram in method 1. Edge detection using the canny edge detector on an approximation image obtained from the LP decomposition of the mammogram was employed in method 2. The PM boundary is selected from the edge structures obtained, after an analysis of the geometrical properties of such structures. The nearness of the computed PM boundary to the ground truth marked by the radiologists was verified using the performance metrics Hausdorff distance, false positive average and false negative average. Around 97% of images had an average row normalized error of 4 mm for method 1 and 3 mm for method 2. The obtained values indicated that the PM border extracted by the method 1 was closer to the ground truth compared to the method 2. Both methods were found to be superior to the works in current literature to which they were compared.

A novel method for detecting architecturally distorted region is also proposed. In order to reduce the area of searching, homogeneous regions in mammograms were extracted using modified SUSAN filter. The strong edge structures of these regions were obtained using an energy based model. A geometrical evaluation of these thinned structures generated 13 features. These features were used by a BPNN to classify the ROIs as normal or architecturally distorted. The specificity, sensitivity and accuracy were significantly improved by initializing the input weights and bias using metaheuristic algorithms: CS and Bat algorithm. The overfitting was addressed by k-fold cross validation for feature selection and reduction of hidden layer neurons. The CS initialized BPNN classification yielded a sensitivity of 89%, 89.8% and 97.6% and specificity of 90.9%, 85% and 96.7% respectively for the MIAS, DDSM and Lakeshore databases. Accuracy of the CS initialized BPNN classification for Lakeshore database was 97.6% with 0.2% FP/ image. For the MIAS database it was 89.1% with 1.2% FP/ image and for the DDSM database it was 89.1% with 2.5% FP/image. The bat initialized BPNN classification yielded a sensitivity of

90%, 89.3% and 97.2% and specificity of 90.9%, 83.6% and 96.7% respectively for the MIAS, DDSM and Lakeshore databases. Accuracy of the bat initialized BPNN classification for the Lakeshore database was 97.1% with 0.2% FP/ image, for the MIAS database 90.1% with 1.2% FP/image and for the DDSM database it was 88.6% with 2.6% FP/ image. The values obtained were superior to those for the methods in current literature to which the algorithms were compared, for the MIAS and DDSM databases.

The directionality properties of contourlet transform were employed to enhance the MCC edge features in the mammogram. The parent child relationship of the zero crossings and modulus maxima of the contourlet coefficients across the different scales were utilized to enhance the abnormality. An analysis of the visual quality, noise reduction and contrast after the enhancement of the MCC features were performed using the Tenengrad criterion, peak signal to noise ratio, average signal to noise ratio, contrast measure and contrast improvement index. When the modulus maxima of the CT were used in MIAS and DDSM database it yielded an improvement in visual appearance, contrast, noise reduction and sharpness. These results were better than those obtained with the zero crossings of the CT.

6.2. Future work

Some proposals for future work are listed below.

- Expansion of the Lakeshore database with more number of mammograms, with accompanying ground truth information
- Evaluation of the MCC enhancement methods with larger databases and to develop a classifier for detection of malignancy from the MCC clusters
- Classification techniques based on BPNN were employed in the research. Other more efficient classifiers could be tried out.

Chapter 6: Concluding Remarks and Future Work

- Other areas that could be addressed include detection of bilateral asymmetry, spiculated lesions and masses.

6.3. Bibliography

J. Ferlay., I. Soerjomataram., R. Dikshit., S. Eser, C. Mathers., M. Rebelo., D. M. Parkin., D. Forman., & F. Bray. (2014). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN-2012. *International Journal of Cancer*, 136(5), E359-E386.

Appendix 1:

Database Specifications

Research in computer-aided detection for mammography has been a challenging and active area of research. The publicly available mammography datasets such as the Mammographic Image Analysis Society (MIAS) (Suckling et al. 1994) and Digital Database for Screening Mammography (DDSM) (Heath et al. 1998) databases have been employed by researchers to develop and evaluate CAD schemes. This research work also includes a collection of mammographic images of Indian populace created from images obtained from the Lakeshore Hospital.

1. MIAS DATABASE: Mammographic Image Analysis Society

The Mammography Image Analysis Society (MIAS), based out of the University of Essex, U.K., is available at: <http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html>. The X-ray films in the database were selected from the United Kingdom National Breast Screening Programme and digitized with a Joyce-Lobel scanning microdensitometer to a resolution of $50 \mu\text{m} \times 50 \mu\text{m}$, a device linear in the optical density range 0-3.2 and representing each pixel with an 8-bit word. The database contains left and right breast images for 161 patients, and is available on a DAT-DDS tape. It consists of 322 images which belong to Normal, benign and malignant classes. There are 208 normal, 63 benign and 51 malignant (abnormal) images. It also includes radiologist's 'truth'-markings on the locations of any abnormalities that may be present. The database has been reduced to a 200 micron pixel edge and padded/clipped so that all the images are 1024 by 1024. Mammographic images are available via the Pilot European Image Processing Archive (PEIPA) of the University of Essex. Four types of abnormalities (architectural distortions, stellate lesions, circumscribed mass and calcifications) are found in the

database. Consider the following description about an image, described using legends in 7 columns as shown in the example:

mdb001	G	CIRC	B	535	425	197
---------------	----------	-------------	----------	------------	------------	------------

The Legends are described below.

1st column:

MIAS database reference number.

2nd column:

Character of background tissue:

F Fatty

G Fatty-glandular

D Dense-glandular

3rd column:

Class of abnormality present:

CALC Calcification

CIRC Well-defined/circumscribed masses

SPIC Spiculated masses

MISC Other, ill-defined masses

ARCH Architectural distortion

ASYM Asymmetry

NORM Normal

4th column:

Severity of abnormality;

B Benign

M Malignant

5th, 6th columns:

x,y image-coordinates of centre of abnormality.

7th column:

Approximate radius (in pixels) of a circle enclosing the abnormality

Other information regarding the database includes:

- The list is arranged in pairs of films, where each pair represents the left (even filename numbers) and right mammograms (odd filename numbers) of a single patient.
- The size of *all* the images is 1024 pixels x 1024 pixels. The images have been centered in the matrix.
- When calcifications are present, centre locations and radii apply to clusters rather than individual calcifications. Coordinate system origin is the bottom-left corner.
- In some cases calcifications are widely distributed throughout the image rather than concentrated at a single site. In these cases centre locations and radii are inappropriate and have been omitted.

2. DSM Database: Digital Database for Screening Mammography (DDSM)

DDSM database is a resource for use by the mammographic image analysis research community. The images are available at: <http://marathon.csee.usf.edu/Mammography/Database.html>. It is a collaborative effort between Massachusetts General Hospital, Sandia National Laboratories and the University of South Florida Computer Science and Engineering Department. The database contains approximately 2,500 studies, each one includes two images of each breast, along with some associated patient information (age at time of study, ACR breast density rating, subtlety rating for abnormalities, and ACR keyword description of abnormalities) and image information (scanner, spatial resolution, etc.). Images containing suspicious areas have associated pixel-level "ground truth" information about the locations and types of suspicious regions.

The Digital Database for Screening Mammography is organized into "cases" and "volumes." A "case" is a collection of images and information corresponding to one mammography exam of one patient. A "volume" is simply a collection of cases collected together for

purposes of ease of distribution. All volumes are available on 8mm tape. Each case in this volume of cancer cases has at least one path-proven cancer. Some cases contain more than one cancer in one breast, a cancer in each breast, or a cancer along with other abnormal/suspicious regions. The outlines of all regions have been transcribed from markings made by an experienced mammographer. Software provided in the site are useful both for accessing the mammogram and truth images and for calculating performance figures for automated image analysis algorithms.

Each case consists of between 6 and 10 files, classified as four categories:

- "ics" file: contains some information about the images, such as the age of the patient, the size of the mammograms, whether or not a file exists for the overlay of abnormality outlines, etc.
- "16-bit PGM" file: overview of the real mammograms.
- "ljpeg" file: contains four image files that are compressed with lossless JPEG encoding.
- "overlay" files: gives the keyword description for a given abnormality in each view, while normal cases will not have any overlay files.

3. Lakeshore Hospital Database

This database was created for the purpose of this research from mammograms obtained from Lakeshore Hospital Kochi. All images are digital mammograms obtained using the Siemens MAMMOMAT 3000 NOVA mammograph. The results of mammograms obtained from Lakeshore Hospital, India are verified by the senior radiologists of Lakeshore Hospital who collaborated with this work.

Appendix 2:

Pseudocode of CS/ Bat algorithm.

1. CS Algorithm:

The pseudo code for CS algorithm can be summarized as follows.

Generate population x_i , where $i=1,2,\dots,n$ for n number of host nests

Evaluate the fitness function, $F_i = f(x_i)$ for all x_i

While (iteration < MaxGeneration) or (stopping criterion)

Generate a cuckoo egg, x_j randomly from host nest using levy flight

Evaluate the fitness function, $F_j = f(x_j)$

Get a random nest i among host nest n

If $F_i < F_j$

Replace x_i with x_j and F_i with F_j

Eliminate a fraction (p_a) of worst nests.

Replace lost nests with new one

Evaluate the fitness of new nests. Find the best nest.

End of while

2. Bat algorithm:

The pseudo code for Bat algorithm following the above assumption is given below.

Initialize population such as position, x_i and velocity, v_i for n number of Bats where $i=1,2,\dots,n$

Initialize parameters including frequency f_i , pulse rate r_i and loudness A_i

Evaluate the fitness function, $F_i = f(x_i)$ for all x_i

While ($t < \text{MaxNoOfIteration}$)

Generate new solutions by adjusting frequency and updating velocities and locations through local search

Evaluate new solutions according to the fitness function

Update echolocation parameters

*Find the current best Bat, x^**

End of while

3. CS/ bat BPNN Algorithm

The pseudocode for a CS/ Bat initialized BPNN classification is given below.

CS/ Bat is initializes and passes the best weights to BPNN

Load the training data

While MSE < stopping criteria

Initialize all cuckoo nests /Bat position & velocity

Pass the cuckoo nests/Bat position & velocity as weights to network

*Feed forward neural network runs using the weights initialized with CS/
Bat*

Calculate the error backward

*CS/Bat keeps on calculating the best possible weight at each epoch until
the network is converged.*

List of Publications

Journal Papers

1. **Rekha Lakshmanan**, Rahila Ramesh, Vinu Thomas, “Enhancement of Microcalcification Features in Mammograms using Zero-Crossings of the Contourlet Transform”, Journal of Emerging Trends in Engineering and Applied Sciences (JETEAS), Vol. 6, No.2, pp:1009-1014, 2011.
2. **Rekha Lakshmanan**, Vinu Thomas, “Mammographic Feature Enhancement using Morphology and Spatial Filter”, proceedings published in International Journal of Computer Application, pp: 29-33, 2012.
3. **Rekha Lakshmanan**, Vinu Thomas, “Microcalcification Detection by Morphology, Singularities of Contourlet Transform and Neural Network”, Bonfring International Journal of Networking Technologies and Applications, Vol. 1, Issue1, pp: 14-19, December 2012.
4. **Rekha Lakshmanan**, Vinu Thomas, “Mammographic Feature Enhancement using Singularities of Contourlet Transform”, Int. J. on Recent Trends in Engineering and Technology, Vol. 8, No. 2, pp: 87-90, by ACEEE, USA, Jan 2013.
5. **Rekha Lakshmanan**, Shiji T P, Vinu Thomas, Suma M. Jacob and Thara Pratab, “Detection of Architectural Distortion in mammogram”, Journal of Image and Graphics, Vol. 2, No.2, December 2014.

6. Accepted, to be published in volume 22, 2016.

Rekha Lakshmanan, Shiji T.P, Suma Mariam Jacob, Thara Pratab, Chinchu Thomas and Vinu Thomas Detection of Architectural Distortion in mammograms using geometrical properties of thinned edge structures, Intelligent Automation & Soft Computing Journal" (Autosoft Journal), published by Taylor and Francis.

International Conference

1. **Rekha Lakshmanan**, Vinu Thomas, "Mammographic Feature Enhancement using Morphology and Spatial Filter", International Conference on Emerging Technological Trends in Advanced Engineering Research (ICETT 2012), February 20-21, 2012. **[Paper Presented]**
 - Proceedings published by International Journal of Computer Applications(IJCA)

2. **Rekha Lakshmanan**, Vinu Thomas, "Investigation of Modulus maxima of Contourlet Transforms for enhancing mammographic features", Second International Conference on Advances in Information Technology and Mobile Communication (AIM 2012), pp: 293-296, April 27-28, Bangalore, India, 2012. **[Paper Presented]**
 - Published in Mobile Communication and Power Engineering, Communications in Computer and Information Science (CCIS) series, Published by Springer, Volume 296, pp: 408-411, 2013.

3. **Rekha Lakshmanan**, Vinu Thomas, “Enhancement of microcalcification features using morphology and contourlet transform”, International Conference on Advances in Computing and Communications, ICACC 2012, published in IEEExplore, pp:14-17, 2012. [**Paper Presented**]
4. **Rekha Lakshmanan**, Vinu Thomas, “Mammographic Feature Enhancement using Singularities of Contourlet Transform”, Fourth International Joint Journal Conference in Computer, Electronics and Electrical, CEE 2012.
5. **Rekha Lakshmanan**, Shiji T P, Vinu Thomas, Suma M.Jacob and Thara Pratab, “Pectoral Muscle Boundary detection - A preprocessing method for early breast cancer detection”, 9th International Forum on Multimedia and Image Processing, August 3-7, World Automation Congress, Waikoloa, Hawaii, USA, 2014, published in IEEExplore, pp:258 – 263, 2014. [**Paper Presented**]
6. **Rekha Lakshmanan**, Shiji T P, Vinu Thomas, Suma M.Jacob and Thara Pratab, “A preprocessing method for reducing search area for Architectural Distortion in mammographic images”, Fourth International Conference on Advances in Computing and Communications, ICACC 2014, published in IEEExplore, pp:101-104, 2014. [**Paper Presented**]
7. **Rekha Lakshmanan**, Shiji T P, Vinu Thomas, Suma M.Jacob and Thara Pratab, “A novel method to detect Architectural Distortion in Mammogram”, International Conference on Information Science, ICIS-2014, July 4-5, 2014. [**Paper Presented**]

8. **Rekha Lakshmanan**, Shiji T P, Vinu Thomas, Suma M.Jacob and Thara Pratab, “Detection of Architectural Distortion in mammogram”, International Conference on Advances in Image Processing , ICAIP-2014, Bangkok, Thailand, October 10-12, 2014. **[Paper Presented]**

- **Received BEST PRESENTATION AWARD**

9. **Rekha Lakshmanan**, Shiji T. P, Vinu Thomas, Suma Mariam Jacob, Thara P, Pectoral Muscle Boundary Detection in Mammograms Using Homogeneous Contours, Fifth International Conference on Advances in Computing and Communications (ICACC), published in IEEEExplore, pp: 354 – 357, 2015. **[Paper Presented]**

National Conference

1. **Rekha Lakshmanan**, Shiji T P and Vinu Thomas, Pectoral Muscle Boundary Detection in Mammograms, National Conference on Emerging Vistas of Engineering and Management (NCEVEM’14), Nov.24-24, Kerala, India, 2014.

- **Received BEST PAPER AWARD**

INDEX

- A
Anatomy, 3, 25, 26, 27
Architectural Distortion, 18, 23,
35, 37, 38, 43, 44, 45, 47, 51,
99, 101, 103, 107, 109, 113,
106, 159, 160, 162, 163, 164,
165, 166, 167, 169
Area Normalized Error, 85
- B
BackPropagation NeuralNetwork,
99, 128, 129, 130, 131, 132,
136, 142, 152, 154, 156, 157,
158, 167
Bat Algorithm, 134, 168
Bilateral Asymmetry,
Bi-Rads categories, 29, 39
Breast Cancer, 2, 11, 30, 33, 35,
40, 48, 49, 50
- C
CAD, 18, 19, 20, 27, 47, 58, 160
Canny, 53, 76, 77, 78, 98, 107,
119, 161, 184, 203
CC view, 16, 17, 46, 142
Classification, 18, 20, 22, 23,
100, 99, 103, 104, 105, 106,
107, 122, 127, 129, 130, 132,
136, 137, 138, 141, 142, 152,
154, 156, 157, 159, 160, 165,
167, 208
Computer Aided Detection, 2, 18,
22, 99
Computer Aided Diagnosis, 18,
165
Computer aided techniques, 1,
29, 43
Contour, 55, 56, 57, 59, 64, 65,
66, 70, 71, 72, 73, 86, 99, 100,
112, 113, 173
Contourlet Transform, 23, 165,
171, 172, 173, 174, 177, 178,
179, 181, 183, 184, 187, 189,
191, 193, 196, 197, 198, 199,
200, 201, 204, 207
Contrast, 6, 8, 12, 26, 36, 38, 43,
44, 104, 105, 165, 167, 169,
170, 187, 188, 191, 198, 199,
201, 204
Contrast enhancement, 57, 59,
64, 65, 99
Contrast Improvement Index,
165, , 187, 191, 198, 199, 200
Crossvalidation, 137
Cuckoo Search Algorithm, 99,
107, 128, 132, 133, 136, 142,
152, 154, 156, 157, 158
- D
Database, 29, 45, 52, 102, 152,
168, 208
Directional Filter Bank,
Ductal Carcinoma, 30, 31, 32
Ductogram, 8
- E
Eccentricity, 66, 72, 73, 74, 75,
76, 80, 81, 82, 99, 122, 138
- F
False Negative, 7, 18, 19, 20, 42,
43, 44, 53, 54, 83, 84, 85, 86,
96, 101, 102
False Negative Average, 83, 85,
96

False Positive, 6, 7, 18, 19, 42, 43, 53, 54, 55, 83, 84, 85, 86, 96, 159, 209
 False positive average, 83, 85, 96, 209
G
 Gaussian pyramid, 76, 77, 82
 Gray level grouping, 57, 64H
 Hausdorff distance, 53, 84, 85, 96, 99, 209
 Incidence rate, 41
L
 Laplacian Pyramid, 209
 Le Gal's classification,
 Lobular Carcinoma, 33
M
 Mass, 35, 37, 38, 39, 44, 45, 101, 103, 160, 169
 Metaheuristic algorithm, 23, 99, 132, 134, 156, 209
 Microcalcification, 6, 12, 21, 22, 23, 35, 36, 38, 43, 44, 45, 102, 101, 103
 MLO view, 15, 16, 17, 45, 53, 54, 55, 58, 62, 63
 Modulus maxima, 210
 Mortality rate, 41, 43
 MRI, 5, 6, 7, 21, 58, 99
O
 Orientation, 56, 58, 59, 62, 66, 72, 73, 74, 75, 76, 80, 81, 82, 105, 106, 118
P
 Parent-child relationship,
 Pectoral Muscle, 15, 16, 17, 22, 23, 43, 44, 53, 54, 55, 56, 57, 58, 59, 60, 62, 63, 64, 65, 66, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 207, 208
 PET, 5, 8, 9, 22
 Pyramidal Directional Filter Bank,
Q
 Quincunx filter bank,
R
 RNE, 84, 85, 87, 88, 96, 209
 ROC, 20, 21, 25, 27, 50, 105, 138, 142, 154, 162
S
 Screening, 14, 17, 27, 29, 161
 Sensitivity, 17, 19, 20, 22, 28, 99, 103, 104, 105, 106, 132, 141, 152, 156, 157, 158, 209
 Sequential backward selection, 136, 137, 168
 Specificity, 20, 99, 132, 141, 142, 152, 156, 157, 158, 209
 Statistics, 42
 SUSAN, 65, 66, 67, 68, 69, 99, 101, 107, 108, 109, 110, 112, 142, 168, 209
 Symptoms, 24, 47
T
 Target to Background contrast ratio,
 Tenengrad, 210
 Thermal imaging, 9, 10
 True negative, 18, 19, 20
 True positive, 18, 19, 102
U
 UltraSound, 5, 7, 21, 49
Z
 Zero crossings, 210