Computer Aided Diagnosis In Digital Mammography:
Classification Of Mass And Normal Tissue

Monika Shinde
University of South Florida

Follow this and additional works at: https://digitalcommons.usf.edu/etd

Part of the American Studies Commons

Scholar Commons Citation
https://digitalcommons.usf.edu/etd/1477

This Thesis is brought to you for free and open access by the USF Graduate Theses and Dissertations at Digital Commons @ University of South Florida. It has been accepted for inclusion in USF Tampa Graduate Theses and Dissertations by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.
Computer Aided Diagnosis In Digital Mammography:
Classification Of Mass And Normal Tissue

by

Monika Shinde

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Computer Science
Department of Computer Science and Engineering
College of Engineering
University of South Florida

Co-Major Professor: Sudeep Sarkar, Ph.D.
Co-Major Professor: John Heine, Ph.D.
N.Ranganathan, Ph.D.

Date of Approval:
July 10, 2003

Keywords: mass segmentation, laws' texture features, expectation maximization

© Copyright 2003 , Monika Shinde
# TABLE OF CONTENTS

| LIST OF TABLES | iv |
| LIST OF FIGURES | v |
| ABSTRACT | vii |

## CHAPTER 1 INTRODUCTION

1.1 Motivation 1
1.2 Thesis Goal: Automated Segmentation 3
1.3 Thesis Outline 4

## CHAPTER 2 BACKGROUND

2.1 Mammography: General Image Information 5
2.2 Mammography: Facts and Figures 6
2.3 Screening and Diagnostic Mammography 7
2.4 Mammographic Abnormalities 8
  2.4.1 Calcification 9
  2.4.2 Mass 9
2.5 Density 10
2.6 Present Clinical Protocol 10
  2.6.1 BI-RADS Descriptors and Assessment Categories 10
    2.6.1.1 BI-RADS Mass Descriptors 11
    2.6.1.2 Breast Composition 13
    2.6.1.3 Assessment Categories 13
  2.6.2 Mammogram Interpretation 14
2.7 Advances in Mammography 15
2.7.1 Digital Mammography
2.7.2 Computer Aided Detection
2.7.3 Computer Aided Diagnosis  

CHAPTER 3 COMPUTER ANALYSIS OF MAMMOGRAMS: LITERATURE REVIEW

3.1 Automated Detection 18
3.2 Automated Classification 19
   3.2.1 Texture Analysis 20
   3.2.1 Clustering Analysis 21

CHAPTER 4 PROPOSED AUTOMATED MASS CLASSIFICATION SYSTEM

4.1 Automated Mass Classification System 22
   4.1.1 Image Data Acquisition 23
   4.1.2 Detection 24
   4.1.3 Segmentation 24
   4.1.4 BI-RADS Predication 25
   4.1.5 Feature Extraction 25
   4.1.6 Classification 25

4.2 Automated Mass Classification: Example Case 25

4.3 Segmentation Method 28

CHAPTER 5 ALGORITHMS

5.1 Laws’ Texture Features 30
   5.1.1 Laws’ Texture Feature Extraction Algorithm 32
   5.1.2 Feature Vector 33

5.2 Expectation Maximization 34
   5.2.1 Mixture Model Estimation 34
   5.2.2 Expectation Maximization Algorithm 36

5.3 Results of EM 37
CHAPTER 6 EXPERIMENTAL SETUPS AND RESULTS

6.1 EM with Intensity 41
6.2 EM with Laws’ Texture Features 43
6.3 EM with Laws’ Texture Features and Intensity 43
   6.3.1 EM with Selected Laws’ Texture Features and Intensity 43
   6.3.2 EM with Wave and Ripple Feature with Intensity 44

CHAPTER 7 CONCLUSION 46

REFERENCES 47

APPENDICES 53
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 4.1</th>
<th>Training Data Set Distribution</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4.2</td>
<td>Assessment Transformation Table</td>
<td>27</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Example Encoding for Single Mass</td>
<td>27</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Example Training Data File</td>
<td>27</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>36 2-D Filter Masks</td>
<td>31</td>
</tr>
<tr>
<td>Table 6.1</td>
<td>Classification Results for BIRADS (Specified by Radiologist)</td>
<td>39</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Classification Results for Automated Feature Extraction Using Manual Segmentation</td>
<td>40</td>
</tr>
<tr>
<td>Table 6.3</td>
<td>Classification Results for Segmentation Using EM with Intensity Feature</td>
<td>42</td>
</tr>
<tr>
<td>Table 6.4</td>
<td>Classification Results for Segmentation Using EM with Ripple Feature</td>
<td>44</td>
</tr>
<tr>
<td>Table 6.5</td>
<td>Classification Results for Segmentation Using EM with Wave Feature</td>
<td>45</td>
</tr>
<tr>
<td>Table A.</td>
<td>Biopsy Results</td>
<td>56</td>
</tr>
<tr>
<td>Figure Number</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Figure 1.1</td>
<td>Block Diagram for Automated Mass Classification System</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Mammographic Breast Anatomy</td>
<td>6</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Views Taken in Screening Mammography</td>
<td>7</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Latero medial (LM) Mammographic View (Left)</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Mediolateral (ML) Mammographic View (Right)</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>BI-RADS Mass Descriptors for Shape</td>
<td>11</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>BI-RADS Mass Descriptors for Margin</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>Suspicious Areas Marked on the Digitized Mammogram by a CAD System</td>
<td>16</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Flow Chart for Overall Automated System</td>
<td>23</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Region of Interest of a Mammogram Showing Mass (Left)</td>
<td>26</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Manual Outline Marked by a Radiologist (Right)</td>
<td>26</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Flow Chart for Segmentation Approach</td>
<td>29</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Filter Mask obtained by Convolving the L7 and W7 Vectors</td>
<td>31</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>Resultant Set of 21 Images after Laws’ Texture Feature Analysis on ROI</td>
<td>33</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>ROI Showing Mammographic Mass (Left)</td>
<td>35</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Histogram Plot of Intensity for Respective ROI (Right)</td>
<td>35</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>ROI Showing Mammographic Mass (Left)</td>
<td>38</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>Segmentation Result Using EM on Ripple and Intensity Feature (Right)</td>
<td>38</td>
</tr>
<tr>
<td>Figure 6.1</td>
<td>Sample Dataset with Manual Outline</td>
<td>41</td>
</tr>
<tr>
<td>Figure 6.2</td>
<td>Sample Image Set with EM Segmentation Results</td>
<td>42</td>
</tr>
<tr>
<td>Figure 6.3</td>
<td>Sample Image Set with Segmentation Results for Ripple Feature</td>
<td>44</td>
</tr>
<tr>
<td>Figure 6.4</td>
<td>Sample Image Set with Segmentation Results for Wave Feature</td>
<td>45</td>
</tr>
</tbody>
</table>
COMPUTER AIDED DIAGNOSIS IN DIGITAL MAMMOGRAPHY:

CLASSIFICATION OF MASS AND NORMAL TISSUE

Monika Shinde

ABSTRACT

The work presented here is an important component of an ongoing project of developing an automated mass classification system for breast cancer screening and diagnosis for Digital Mammogram applications. Specifically, in this work the task of automatically separating mass tissue from normal breast tissue given a region of interest in a digitized mammogram is investigated. This is the crucial stage in developing a robust automated classification system because the classification depends on the accurate assessment of the tumor-normal tissue border as well as information gathered from the tumor area. In this work the Expectation Maximization (EM) method is developed and applied to high resolution digitized screen-film mammograms with the aim of segmenting normal tissue from mass tissue. Both the raw data and summary data generated by Laws’ texture analysis are investigated. Since the ultimate goal is robust classification, the merits of the tissue segmentation are assessed by its impact on the overall classification performance.

Based on the 300 image dataset consisting of 97 malignant and 203 benign cases, a 63% sensitivity and 89% specificity was achieved. Although, the segmentation requires further investigation, the development and related computer coding of the EM algorithm was successful. The method was developed to take into account the input feature correlation. This development allows other researchers at this facility to investigate various input features without having the intricate understanding of the EM approach.
CHAPTER 1

INTRODUCTION

Breast cancer (BC) is the second leading cause of cancer deaths among women in United States and it is the leading cause of cancer deaths among women in the 40 – 55 age group [1-5]. According to American College of Radiology (ACR) statistics, one out of nine women will develop breast cancer during her lifetime. Between 1973 and 1999, breast cancer incidence rates increased by approximately 40% [6]. In the year 2003, 40,200 deaths (39,800 women, 400 men) are anticipated from breast cancer in 2003 [3]. However, during 1989-1995 the BC mortality rates declined by 1.4% per year and by 3.2% afterwards [6]. These declines have been attributed, in large part, to early detection [3]. Also, survival through BC is found to be stage-dependent and the best survival is observed when diagnosed at early disease-stage. Mammography is an effective tool for early detection because in many cases it can detect abnormalities such as masses, calcifications, and other suspicious anomalies up to two years before they are palpable.

1.1 Motivation

Although radiographic breast imaging and screening has allowed for more accurate diagnosis of breast disease at earlier stages of development, 10-30% of malignant cases (biopsy proven cancerous) are not detected for various reasons such as technical problems in the imaging procedure, abnormalities that are not observable, and abnormalities that are misinterpreted [7]. This group of “non-detected” cancers is generically referred to as missed cancers (MC). Evidence indicates that somewhere between 7-20% of mammograms with abnormalities currently detected also show signs in the previous mammogram when viewed in retrospect, which may be considered as false negative (FN) errors [7,8]. Since it is better to error on the side of safety, about
65-80% of breast biopsies result in benign diagnosis, which may be considered as false positive (FP) biopsies [7]. In addition to the physical trauma, there is undo emotional stress associated with the FP reading. Likewise, the cost of the FN misinterpretation is enormous.

The diagnosis errors discussed above form the foundation for the work presented here. That is, we believe that computer aided decision methods can improve both the FP and the FN diagnosis rates. Although there are a few commercially available automated detection systems (discussed below) that are used as imaging checking systems in conjunction with the radiologist interpretation, the idea of automated classification has not been used clinically as of yet to any extent [9].

There are important distinctions between detection and classification of suspected abnormalities when considering computer applications. The detection process always precedes classification and may be implemented by some automated method or by a radiologist through conventional methods, as in the normal mammography protocol. Once there is a detected abnormality, by whatever means, it must be classified, which may be achieved by human assessment, pathology analysis, with automated methods, or some combination of the three.

The work presented here may be considered as the groundwork for an overall automated classification system for use in digital mammography (DM). This system, which is under development at this facility, may be considered as a complement to the radiologist’s assessment. That is, the radiologist does the detection task and then cues the system to region of the suspected abnormality for either selected subjects or all applicable subjects. The system then provides a probabilistic figure of merit relating to the degree of malignancy. The intended uses of this system includes both stand-alone classification or as a second opinion strategy. Since the classification system is designed via modular programming techniques, it may also be joined with a given automated detection method, where automated techniques find the abnormal areas that warrant further classification.
1.2 Thesis Goal: Automated Segmentation

Once the system is cued to the abnormality location, the classification consists of three main processing steps:

- Separate the abnormality from normal tissue
- Feature analysis.
- Classify the degree of malignancy.

The Figure 1.1 shows the elements of the overall classification scheme. The mammograms are digitized in the acquisition phase while the abnormality is located in the second phase called detection. The next three phases: segmentation, feature extraction and classification cover the main processing steps of the automated mass classification system that is under development here at this facility.

![Figure 1.1 Block Diagram for Automated Mass Classification System](image-url)
The segmentation step is the crucial stage addressed here; if it fails, the entire classification analysis fails. The goal of this work is to develop a robust method of segmenting breast masses from the normal background breast tissue. The success of automated classification requires knowledge of the mass, ambient normal-tissue, background border region, and the tumor area. The specific aims of this project are (1) develop the computer code for implementing the Expectation Maximization (EM) procedure for mass segmentation applications in (DM), and (2) apply the method to mammograms as an initial feasibility study. The EM method is applied to both the raw data as well as summary data derived by applying Laws’ texture analysis method. For evaluation purposes, it is necessary to discuss elements of the overall classification scheme; the segmentation performance is assessed by its impact on the overall classification process. In order to develop an understanding of the problem and for algorithm training purposes, input from experienced radiologist is necessary. The mass-border region is not exactly known from assessing the mammogram. Mammographers have electronically hand-labeled many cases for training purposes, but there is significant inter-radiologist variability in the border descriptions. Although, we will use these subjective measures as a guide for developing our ideas, the segmentation’s impact on the overall classification performance will be used as the analytical assessment.

1.3 Thesis Outline

The manuscript is organized as follows. In Chapter 2 breast cancer facts-statistics, the role of mammography, and advancements in computer aided detection and diagnosis (CAD) are discussed. A literature review of related computer methods applied in mammography is provided in Chapter 3. The proposed segmentation approach and the necessary preliminary mathematical developments are developed in Chapter 4. In Chapter 5, special attention is given to the well known Laws’ texture feature analysis and the mathematical nuances of the EM approach are presented. The experimental procedures and the results are presented in Chapter 6. The conclusions and possible extensions of this work are discussed in Chapter 7.
CHAPTER 2

BACKGROUND

The early detection of breast cancers by screening mammography greatly improves a woman's chance of survival [6]. In many cases mammography can detect abnormalities up to two years before they become palpable. The current guidelines from the U.S. Department of Health and Human Services (HHS), the American Cancer Society (ACS), the American Medical Association (AMA) and the American College of Radiology (ACR) recommend screening mammography every one to two years for women beginning at age 40. However, this commencement age is often debated. Young women are apt to have greater proportions of dense breast tissue, which gives rise to low contrast mammograms that may be difficult to interpret. This is another area where CAD may play an important role in increasing screening efficacy via digital manipulation.

This chapter provides the general mammographic image formation information and relevant statistical facts related to mammography. Further, it explains the types of mammograms, mammographic abnormalities and also discusses the current advancements in mammography.

2.1 Mammography: General Image Information

Mammography is a transmission planar x-ray image formed by a diverging x-ray beam. Thus, the breast volume attenuation is represented by light and dark shadows captured in a film-screen combination process; the resulting image is planar projection of the three dimensional breast. The image is very similar to observing a light beam after passing through the canopy of an oak tree. There are many sources of uncertainty in the captured image due to (for example) scattering, beam hardening, diverging x-rays, and signal derived form x-rays leaving the x-ray
tube through areas other than the focal spot. There are additional uncertainties due to the nature of photon counting statistics and detection process. Thus the resulting image is less than perfect. A more complete exposition of the image process may be found elsewhere [10].

In current mammography imaging practice, there are basically two types of normal tissue distinguishable in the images. One is dense tissue, which is a two component mixture of stromal and epithelial tissue, appearing bright in the image and other is fatty tissue, which appears dark. The fundamental difficulty in either human or computerized breast image analysis is that dense normal tissue and abnormal tissue often have similar x-ray attenuations with respect to the x-ray spectrum in conventional imaging practice, which results in similar image intensities; also the textures are similar. A sample mammogram displaying the breast anatomy is shown in Figure.2.1

Figure 2.1 Mammographic Breast Anatomy [16]

2.2 Mammography: Facts and Figures

Some important aspects of mammography are provided here.
• The FDA reports that mammography can find 85-90% of breast cancers in women over 50 and show some lumps (masses) up to 2 years before it can be felt [3].

• Breast cancers found by screening mammography of women in their forties were smaller and at an earlier stage with less spread to lymph nodes or other organs than cancers found in women not having mammography [3].

• The results reported by American Cancer Society of the recent compilation of eight randomized clinical trials found 18% fewer deaths from breast cancer among women in their forties who had mammography [3].

2.3 Screening and Diagnostic Mammography

In practice mammograms are taken in two different environments: regular screening mammography and diagnostic mammography. Screening mammography aims to find cancers early under regular periodic surveillance. Diagnostic mammography is an extended intervention that may apply to screen-detected abnormalities, abnormalities that are palpable and not observable under normal imaging protocol, or for further analysis including serial surveillance.

*Screening Mammography* is a low-dose x-ray examination of the breasts in a woman who is asymptomatic. The *Screening Mammograms are two x-ray views for each breast, typically cranial-caudal view, (CC) and mediolateral-oblique (MLO) as shown below.

![Figure 2.2 Views Taken in Screening Mammography](image)
Diagnostic Mammography is an x-ray examination of the breast in a woman who is symptomatic. This includes a breast lump found via self-examination or during regular screening and nipple discharge. Diagnostic Mammography is more involved and time-consuming than screening mammography. The goal of diagnostic mammography is to pinpoint the size and location of breast abnormality and to image the surrounding tissue and lymph nodes or to rule-out the suspicious findings. Typical views for diagnostic mammograms include lateromedial (LM) and mediolateral view (ML) along with the CC and MLO views as defined on previous page. For specific problems additional special views such as exaggerated cranial-caudal, spot compression, and magnified may be taken. (Spot compression, magnification views often to evaluate microcalcifications and Ductogram / Galactogram for imaging the Breast Ducts are some of the special mammographic views.)

Figure 2.3 Lateromedial (LM) Mammographic View (Left) [50]
Figure 2.4 Mediolateral (ML) Mammographic View (Right) [50]

2.4 Mammographic Abnormalities

Mammography is used to detect a number of features that may indicate a potential clinical problem, which include asymmetries between the breasts, architectural distortion, confluent densities associated with benign fibrosis, calcifications and masses. By far, the two most common features that are associated with cancer are clusters of micro calcifications and masses, which are discussed below.
2.4.1 Calcification

Calcifications are small mineral (calcium) deposits within the breast that appear as localized high-intensity regions (spots) in the mammogram. There are two types of calcifications: micro-calcifications and macro-calcifications. Macro-calcifications are coarse, scattered calcium deposits. These deposits are usually associated with benign conditions and rarely require a biopsy. Micro-calcifications may be isolated, appear in clusters, or found embedded in a mass. Individual micro-calcifications typically range in size from 0.1-1.0 mm with an average diameter of about 0.5 mm. A cluster is typically defined to be at least three micro-calcifications within a 1cm$^2$ region; the clusters are important cues for the mammographer in determining if the reading is suspicious. About 30-50% of non-palpable cancers are initially detected due to the presence of micro-calcifications clusters [10]. Similarly, in a large majority of the ductal carcinoma in situ (DCIS) cancers, calcification clusters are present [4].

2.4.2 Mass

Breast cancer often presents as a mass with or without presence calcifications [3]. A cyst, which is non-cancerous collection of fluid, may appear as a mass in the film. However, ultrasound or fine needle aspirations can distinguish the difference. The similarity in intensities with the normal tissue and in morphology with other normal textures in the breast makes it more difficult to detect masses compared with calcifications [10].

The location, size, shape, density, and margins of the mass are useful for the radiologist in evaluating the likelihood of cancer [5]. Most benign masses are well circumscribed, compact, and roughly circular or elliptical [5]. Malignant lesions usually have a blurred boundary, an irregular appearance, and sometimes are surrounded by a radiating pattern of linear spicules [5]. However, some benign lesions may have a spiculated appearance or blurred periphery.
2.5 Density

The glandular tissue of the breast, or breast density, appears as bright (clearer) areas on film or higher intensity areas in the digitized images. Thus, the increasing areas of dense breast tissue on mammogram can make it more difficult to interpret, although the tissue is generally normal. In general, younger women have a greater proportion of dense breast tissues compared with older women. After menopause, the glandular tissue of the breasts is replaced with fat, typically making abnormalities easier to detect with mammography.

2.6 Present Clinical Protocol

The section covers the current clinical protocol followed by the radiologists for mammographic examination and interpretation. The standardized interpretations follow from the BI-RADS lexicon. The ACR developed BI-RADS lexicon is an acronym for the Breast Imaging Reporting and Data System [11]. This standard provides a mechanism for describing the characteristics of a given abnormality including the final pre-pathology finding.

For mass classification purposes the borders, shape and relative intensities are important descriptive features. In the following subsection, the relevant BI-RADS descriptors and assessment categories are provided. This discussion is constrained to the mass assessment only. A more complete exposition of the rating system with examples can be found elsewhere [11].

2.6.1 BI-RADS Descriptors and Assessment

BI-RADS descriptors are important factors for predicting malignancies that are assessed and provided by the radiologist. The mass narratives include the overall shape description, the border region margin regularity, and the relative intensity of the mass region compared with the ambient normal tissue intensity. The BI-RADS also provides a four-category rating for assessing the overall breast tissue characteristic in terms of the fibro-glandular composition. The composition categories relate to the degree of interpretation difficulty. Similarly, the BI-RADS
gives a 5-point overall assessment that is related to the degree of probable malignancy or necessary follow-up work.

2.6.1.1 BI-RADS Mass Descriptors

1. Shape

The mass shape is described with a four-point assessment: round, oval, lobular and irregular as shown in Figure 2.4, which gives the overall impression

![Figure 2.4 BI-RADS Mass Descriptors for Shape](image)

2. Margin

The mass margins modify the boundaries. For example the overall shape may be round, but close inspection may reveal scalloping along the border, which may indicate a degree of irregularity or a lobular characteristic. The margins are rated with a 5-point system: circumscribed (well-defined or sharply-defined) margins, microlobulated margins, obscured margins, indistinct margins and spiculated margins as shown in Figure 2.5
3. Density

The intensity or the x-ray attenuation of the mass tissue region is described as density. The density here is the relative density, i.e. higher, lower or similar relative to the surrounding tissue. The density is rated on 4-point system:

- High density
- Equal density
- Low density (lower attenuation, but not fat containing)
- Fat containing - radiolucent. The “d” includes all lesions containing fat such as an oil cyst, lipoma, or galactocele.
2.6.1.2 Breast Composition

This is an overall assessment of the global tissue composition, which indicates the relative possibility that the normal tissue could hide a lesion. Generally, this includes fatty, mixed or dense. The four class breast composition ratings are:

- Almost entirely fat.
- Scattered fibro glandular densities.
- Heterogeneously dense, which may lower the sensitivity.
- Extremely dense, which could obscure a lesion.

2.6.1.3 Assessment Categories

Assessment categories are defined for standardized interpretations of mammographic findings. Each category provides the overall assessment related to the findings and the necessary follow up. The 5-point assessment categories are described as follows,

- **Category 0 Incomplete Assessment**: Needs additional imaging evaluation.
- **Category 1 Negative**: The breasts are symmetrical and no abnormalities are present.
- **Category 2 Benign Finding**: This is also a negative mammogram. But the interpreter may describe the finding such as calcified fibro adenomas, fat-containing lesions such as oil cysts etc. that showed no mammographic evidence of malignancy.
- **Category 3 Probably Benign Finding - short interval follow-up suggested**: A finding has a high probability of being benign.
- **Category 4 Suspicious Abnormalities - biopsy should be considered**: These are lesions that do not have the characteristic morphologies of breast cancer but have a definite probability of being malignant.
• **Category 5 Highly Suggestive of Malignancy - appropriate action should be taken:** These lesions have a high probability of being cancerous.

### 2.6.2 Mammogram Interpretation

The radiologist interprets the mammographic examination in the form of a mammogram report. The mammogram report describes the findings (i.e. breast abnormalities), provides the radiologist’s impression based on BI-RADS, and recommends appropriate course action. The following subsections discusses elements of the mammogram reporting discussed above:

- **Findings:** These are the description of breast abnormalities (i.e. of mass, calcification etc) found from the mammogram in terms of their size, location, and characteristics. Primary signs of breast cancer may include spiculated masses or clustered pleomorphic microcalcifications. Secondary signs of breast cancer may include asymmetrical tissue density, skin thickening or retraction, or focal distortion of tissue.

- **Impression:** This contains the radiologist’s overall assessments (findings/breast abnormalities) using the BI-RADS as explained in section 3.1.

- **Recommendation:** Depending on the assessments, this section contains specific instructions on what actions should be taken next. For example, the radiologist could recommend: (a) additional imaging such as spot views, breast ultrasound, MRI etc for category 0, (b) no action necessary if the assessment is category 1 and 2, (c) a six month follow-up mammogram to establish the finding’s stability for category 3, or (d) a biopsy in case of category 4 and 5. A biopsy is a surgical procedure where a sample of tissue is removed by a surgeon and analyzed by a pathologist to determine whether it is cancerous or benign.
2.7 Advances in Mammography

2.7.1 Digital Mammography

X-ray film screen mammography generally has greater sensitivity and specificity for detection of BC than other non-invasive diagnostic technique currently in use [10]. However, there are limitations to its ability to display subtle details and simultaneously produce the image while maintaining safe radiation doses. Moreover, in recent years the trend is moving slowly towards DM applications with the aim of improving some of the performance and quality issues related to film based images. There are two forms of digital mammograms one that is derived from digitizing film-screen images and other that is acquired with digital detection without film. In the latter case, commonly referred to as full field digital mammography (FFDM,) the x-ray film is replaced by solid-state detection that converts signals to digital form.

2.7.2 Computer Aided Detection

The term CAD is commonly used to refer both computer aided detection and computer aided diagnosis. Computer aided detection refers to locating or finding the abnormality and the computer aided diagnosis refers to evaluation or assessment of mammographic abnormality. A detailed discussion on computer aided detection is provided in the following paragraph while the computer aided diagnosis is discussed in section 2.7.3 in reference to mass diagnosis.

Computer-aided detection technology may be used as a second opinion for reviewing a subjects film after the radiologist has already made an initial interpretation. The CAD unit highlights any detected breast abnormalities on the digital mammograms. Figure 2.6 shows a mammogram with abnormalities highlighted by CAD.
Figure 2.6 Suspicious Areas Marked on the Digitized Mammogram by a CAD System

The digitized mammograms are displayed on high-resolution monitors. Based on the results of the CAD marker information, the radiologist may choose to re-examine the original mammogram and possibly modify the initial findings. Thus, the CAD technology works as an "image checking system" for radiologists, alerting them to areas that may require more attention [7, 10].

2.7.3 Computer Aided Diagnosis

Although the BI-RADS forms a standardized interpretation or reporting scheme (as discussed in section 2.5), they are somewhat subjective in nature because they are determined without quantitative methods. That is, the assessment follows from the mammographers’ experience and opinion. Another goal for CAD is to overcome this variability in image assessment [10, 12]. In particular, an ongoing project at this imaging facility includes developing automated methods for calculating the BI-RADS mass descriptors.

Briefly, CAD may serve as a diagnostic tool in a few varied capacities. It may help in reducing the variations in the BI-RADS assessments, which could render non-experts as experts. CAD may be useful for improving the actual diagnosis performance, improving the overall detection performance, or some combination of both.
CHAPTER 3

COMPUTER ANALYSIS OF MAMMOGRAMS: LITERATURE REVIEW

Although mammographic screening is a cost effective BC detection method, there are many interpretation problems [10, 13, 14]. We must keep in mind that the mammographic image is a planar representation of a projected volume and is a poor abstraction of the complicated attenuation properties of the breast. Often, there is a blurred distinction between a suspected abnormality and normal breast tissue in the vicinity, which may give rise to interpretation errors. With the development of low cost computing and economical storage capacity, many researchers have been investigating methods of incorporating computer analysis in the detection and diagnosis of BC. To date there are few commercially available systems that are used for detection purposes in conjunction with the mammographer’s assessment. Currently the R2 Imagechecker, CADx Second Look Inc, and MammoReader are the three CAD systems approved by the FDA, which may assist radiologists in the mammographic image interpretation [9]. Basically, these systems are used in an image checking capacity.

Essentially, the computerized analysis methods in mammography can be divided into two areas that may be conjoined to form a total analysis system: automated abnormality detection and abnormality classification [10, 15]. The automated abnormality detection methods locate the abnormality and leave the assessment task to mammographer, whereas automated classification or diagnosis methods may help the radiologist in the final assessment (benign or malignant prediction).
Apart from these, mammogram registration is another important research area. Mammogram registration is an automated analysis that involves comparison of either bilateral mammograms obtained at the same screening session, or mammograms of the same breast obtained at different screening sessions [16, 17].

3.1 Automated Detection

The majority of automated abnormality detection systems usually involve three steps: noise removal, filtering or some method of rendering the data more useful and the decision analysis with a binary outcome of either yes (abnormality present) or no (not present). The first two steps may be considered as pre-processing. Reviews of important work in preprocessing, image processing, and statistical methods relevant to calcification and mass detection are provided below.

The purpose of preprocessing is to “enhance” the image, which may be achieved by either increasing the contrast or by removing the background tissue or suppressing noise. Contrast enhancement techniques include local or global area thresholding [12, 18], density-weighted contrast enhancement and segmentation [19]. For noise removal, often non-linear methods are applied. These researchers have used median filtering, edge preserving smoothing, half neighborhood and directional smoothing methods for noise removal [20-22]. Locating the breast region relative to the off-breast image area (background) is a pre-processing task common to most image analysis approaches.

Once the breast region is located and pre-processed, various image processing and statistical methods are applied to detect the abnormality. Depending on the type of abnormality, the detection methods can be divided into two groups: mass detection and micro-calcification detection. The micro-calcification detection task is normally not arduous compared with mass detection. There are three reasons for this:
• Dissimilar to calcifications, which are characterized by small high-contrast spots, masses may assume varied shapes and sizes, e.g., spiculated, round, or irregular [1, 5,13].

• Apart from the variability in shape and size, masses are also variable in density and poor in image contrast [1, 5, 23].

• Further, masses are highly connected to the surrounding parenchymal tissue density, especially in case of spiculated lesions and are often surrounded by non-uniform tissue background with similar characteristics [23].

Hence, a wide variety research as been devoted to automated calcification detection. These approaches include the use of wavelet transforms, watershed transforms, and clustering analysis [4, 18, 23, 25-31].

Due to the characteristic similarity of the surrounding tissue with actual mass, the mass detection methods concentrate on extracting features that may differentiate the mass. These features may include asymmetry measures between the breasts, local textural changes, or radiating density patterns. Hence, the approaches used for mass detection include left-right breast comparisons, directional wavelet analysis, rubber band straightening transforms, a variety of texture analysis methods and clustering analysis [8, 9, 26-27 32-34, 36-39].

3.2 Automated Classification

The automated classification or diagnosis methods are developed to assist radiologist in making final assessment [12]. They may be used to estimate the likelihood of malignancy for the given abnormality. The abnormality under consideration may be marked by a radiologist or obtained through automated detection procedure. The classification is carried out using various classifiers such as artificial neural networks or linear discriminant analysis [1, 23,40]. These classifiers are often based on characteristic features of a given abnormality derived from
computerized feature extraction schemes. The features required to distinguish the benign from malignant mass are (may be) abnormality dependent. There has been some effort expended to extract a quantitative set of features that can help in this automated diagnosis quest. For example, texture features have been found useful in classifying masses. Whereas for calcification classification, numbers of calcifications within a given area, individual length, and cluster width of are considered important [4,5]. Detailed information on the distinguishing features for mass and calcification can be found elsewhere [4, 5, 41]. The literature review indicates that Law’s texture analysis may be important for the analysis of masses [2].

3.2.1 Texture Analysis

Texture is a feature that may be useful for partitioning images into regions of interest and to classify those regions; since, it is generally believed that one of the main visual cues are differences in textural properties between the regions. Texture provides information about the spatial distribution of intensity levels in a neighborhood. So, it cannot be defined for a point. It can be viewed as a repeating pattern of local variations in image intensity. Texture analysis methods may be used for segmentation as well as classification [37].

Texture classification is concerned with identifying a given textured region from a given set of texture classes. Texture segmentation is based on determining the boundaries between various texture regions in an image. Texture segmentation can be divided into two major categories: region based, which attempts to group or cluster pixels with similar texture properties and boundary based, which is an attempt to find “texture-edges” between pixels from different texture distributions. Most of the methods for mass separation follow region based texture segmentation [1, 2, 40]. Laws’ texture feature based segmentation method may be considered as the region-based segmentation method.
3.2.2 Clustering Analysis

Clustering analysis is based on partitioning a collection of data points into a number of subgroups, where the objects inside a cluster (a subgroup) show a certain degree of closeness or similarity. In abnormality detection, the image pixels are partitioned and the pixels corresponding to abnormality are clustered on similarity basis. Different clustering methods are applied depending on the criteria used to find the similarity measure. The K-means\cite{25,26}, FCM \cite{28} and Expectation Maximization \cite{42-46} are often used clustering methods for abnormality detection.
CHAPTER 4

PROPOSED AUTOMATED MASS CLASSIFICATION SYSTEM

This chapter presents the overall project approach for the automated system. A special emphasis is given on the segmentation method since it is the first and essential component of the proposed automated mass classification system. Also, this section includes the database description and remarks concerning the importance of BI-RADS assessment in mass classification and malignancy estimation.

4.1 Automated Mass Classification System

Once a region of suspicion is marked by any means, the goal of the automated mass classification system (AMCS) is to perform the following three operations:

- Separate the mass tissue from normal tissue
- Extract the features from segmentation results.
- Classify the mass as benign or malignant, based on the extracted features.

To achieve the goal of an AMCS, the following project approach is used. Figure 5.1 shows the block diagram for this project approach.
4.1.1 Image Data Acquisition

The images were acquired using an Image Clear3000 (DBA, Melbourne, FL) film digitizer. The dynamic rage for the DBA scanner is 16 bits per pixel with a 30 micron pixel resolution. However, for storage purposes the images are half-band filtered, down sampled and stored with 60-micron resolution. For this thesis work, 300 mammographic masses are considered. This includes 203 benign and 97 malignant cases. The benign masses, where appropriate, and malignant masses are pathology proven; 100 normal masses were not biopsied but were followed for two years before pronounced normal. Table 4.1 shows the training dataset distribution.

For ground-truth comparisons the BI-RADS are derived from the mammography reports, and the masses have hand drawn boundaries in electronic copy referenced to the raw image in size and spatial location, which were provided by experienced radiologists. The truth files
provide the region of suspicion and the manual outline represents the radiologists’ opinion (hand drawn image) based on years of experience.

Table 4.1 Training Dataset Distribution

<table>
<thead>
<tr>
<th>Shape</th>
<th>Margin</th>
<th>Density</th>
<th>Breast Composition</th>
<th>Pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round (97)</td>
<td>Circumscribed (113)</td>
<td>High (60)</td>
<td>Fat (32)</td>
<td>Benign (97)</td>
</tr>
<tr>
<td>Oval (47)</td>
<td>Micro-lobulated (64)</td>
<td>Equal (218)</td>
<td>Scattered Dense (120)</td>
<td>Malignant (203)</td>
</tr>
<tr>
<td>Lobulated (69)</td>
<td>Indistinct (48)</td>
<td>Low (19)</td>
<td>Heterogeneous (109)</td>
<td></td>
</tr>
<tr>
<td>Irregular (70)</td>
<td>Obscured (43)</td>
<td>Radio-lucent (3)</td>
<td>Dense (39)</td>
<td></td>
</tr>
<tr>
<td>Arch. Dist. (17)</td>
<td>Spiculated (32)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.1.2 Detection

Although not a part of this thesis work, the mass must be detected prior to the segmentation process. Specifically for the system under development at this facility, the radiologist cues the system to the suspected abnormality and draws a box around the region of the abnormality. This defines the region of interest. Likewise, for an AMCS, the radiologist may be replaced by an automated algorithm to generate the bounding box.

4.1.3 Segmentation

Once given the region containing the mass, the task of separating the mass tissue from the normal surrounding tissue is accomplished by some segmentation method. The segmentation results are used as a guide to extract the features that may distinguish benign and malignant masses. Further, the segmentation results may be used to predict the BIRADS descriptors, which are good predictors of malignancy. The segmentation can be manual (radiologist drawn outline)
or automated (some segmentation method). The performance of any segmentation method can be compared by replacing it with another, for example, a manual outline provided by a radiologist.

4.1.4 BI-RADS Prediction

Radial distance patterns, Laws’ texture features and density detection algorithms may be useful in predict the BIRADS features (shape, margin, density, breast composition etc) from the segmentation results.

4.1.5 Feature Extraction

Various other features are extracted from the segmentation results. These include measures like the mean, contrast, lucency, homogeneity, texture measures, and wavelet features. In an automated method these features may also be useful estimating the BI-RADS ratings specified by the radiologist for classifying the mass as benign or malignant.

4.1.6 Classification

Classifier used in this work is a quick propagation neural network with the leave-one-out validation method. The training data is formed using the features extracted and the BI-RADS assessment along with the pathological assignment of malignancy. A sample of training data file is shown and explained in section 4.2.

4.2 Automated Mass Classification: Example Case

A classification mock trial is provided in this section using the AMCS. The radiologist provides the detection, segmentation and BI-RADS descriptors. This implies the truth files are used for the detection, the manual outline provides the segmentation, the shape, margins etc for BI-RADS are given by a radiologist, and the malignancy is known through the pathological report. Figure 4.2 shows an example mass and a manually specified segmentation is shown in Figure 4.3. This is followed by the radiologist’s BI-RADS interpretation.
Figure 4.2 Region of Interest of a Mammogram Showing Mass (Left)
Figure 4.3 Manual Outline Marked by a Radiologist (Right)

Shape: Round
Margin: Circumscribed
Density: High
Breast Composition: Heterogeneous

This mass is pathologically proven benign mass.

The automated system input requires coded data. The assessment information is transformed into a computer readable format using Table 4.2. Class 0 is benign while class 1 defines the malignant mass. Table 4.3 shows the assessment transformation for the above-described mass. Table 4.4 shows an example of training file generated for five mammographic masses using the assessment transformation table.
### Table 4.2 Assessment Transformation Table

<table>
<thead>
<tr>
<th>Shape</th>
<th>Margin</th>
<th>Density</th>
<th>Breast Composition</th>
<th>Pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round(0)</td>
<td>Circumscribed(0)</td>
<td>High(0)</td>
<td>Fat(1)</td>
<td>Benign(0)</td>
</tr>
<tr>
<td>Oval(1)</td>
<td>Micro-lobulated(1)</td>
<td>Equal(1)</td>
<td>Scattered Dense(2)</td>
<td>Malignant(1)</td>
</tr>
<tr>
<td>Lobulated(2)</td>
<td>Indistinct(2)</td>
<td>Low(2)</td>
<td>Heterogeneous(3)</td>
<td></td>
</tr>
<tr>
<td>Irregular(3)</td>
<td>Obscured(3)</td>
<td>Radio-lucent(3)</td>
<td>Dense(4)</td>
<td></td>
</tr>
<tr>
<td>Arch. Dist.(4)</td>
<td>Spiculated(4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.3 Example Encoding for Single Mass

<table>
<thead>
<tr>
<th>Image</th>
<th>Shape</th>
<th>Margin</th>
<th>Density</th>
<th>Breast Compos.</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>812</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 4.4 Example Training Data File

<table>
<thead>
<tr>
<th>Image</th>
<th>Shape</th>
<th>Margin</th>
<th>Density</th>
<th>Breast Compos.</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>628</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>634</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6351</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>6352</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>6353</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>
The training file for whole dataset is fed into the quick propagation neural network for classification. The leave one out method is used to obtain the benign-malignant prediction for each mass and the overall prediction rates.

In the totally automated classification scheme, the segmentation by the radiologist (manual outline) is replaced by proposed segmentation method, and BI-RADS features are replaced by other various features described in section 4.1.5. That is, instead of using the features provided by the radiologist; the features are extracted from the automated segmentation.

In both the cases the results for sensitivity and specificity are obtained as explained in Appendix A. An increased sensitivity and specificity may be used as a quality measure for the applied segmentation method relative to the radiologist’s segmentation. Thus, the classification method acts as a relative validation method for the segmentation method.

4.3 Segmentation Method

The approach used for segmentation is important because the malignancy estimation is dependent on the segmentation guided feature extraction. These features include the shape, margin, density features which are relative to the surrounding tissue. Hence any over or under segmentation may affect the malignancy prediction. Thus, it is essential that a segmentation method should retain the shape, and margin features intact. Keeping all these things in mind, here is a detailed approach for the proposed segmentation method.

As shown in Figure 4.3 segmentation approach consist of following steps,

- Apply the Laws’ Texture Features to the region of interest
- Apply the Expectation Maximization Algorithm.
- Segment using morphological operators.
The details for Law’s Texture Feature Analysis and the Expectation Maximization Algorithm are provided in Chapter 5.
CHAPTER 5

ALGORITHMS

The EM approach may be applied to raw data as well as summary data derived from the raw data or to a combination of both data types. Summary data implies data obtained from filtering the raw data in this case. In this chapter the foundation for Laws’ texture analysis and the theoretical EM framework are provided. In particular, the mechanisms that define a feature vector and how it is related to the EM approach are discussed.

5.1 Laws’ Texture Features

The application of Laws’ texture features is amounts to filtering the data with various filter kernels related to three image (or signal) features developed by the Kenneth Ivan Laws at the University of Southern California [47]. The three fundamental kernels are defined as: (1) $L_3 = [1, 2, 1]$, the level detector, (2) $E_3 = [-1, 0, 1]$, the edge detector, and (3) $S_3 = [-1, 2, -1]$, the spot detector. These three filter kernels are then convolved with each other to provide a set of six one dimensional filter kernels referenced as level, edge, spot, wave, ripple, and oscillation.

$$L_7 = [1, 6, 15, 20, 15, 6, 1]$$

$$E_7 = [-1,-4,-5, 0, 5, 4, 1]$$

$$S_7 = [-1,-2, 1, 4, 1,-2,-1]$$

$$W_7 = [-1, 0, 3, 0,-3, 0, 1]$$

$$R_7 = [1,-2,-1, 4,-1,-2, 1]$$

$$O_7 = [-1, 6,-15, 20,-15, 6,-1]$$
For example the L7 is obtained by repeated convolutions: L7 = L3*L3*L3. For more details see Laws’ work [47]. The approach is easily extended for two-dimensional applications by forming the direct product of the 1-D kernels resulting in 36 two dimensional filters, which are as shown in Table 5.1. Figure 5.1 shows the details of L7W7 2-D filter mask. Each of these 2-D kernels is then used to perform the texture analysis on an image by standard convolution.

Table 5.1 36 2-D Filter Masks

<table>
<thead>
<tr>
<th></th>
<th>L7L7</th>
<th>E7L7</th>
<th>S7L7</th>
<th>W7L7</th>
<th>R7L7</th>
<th>O7L7</th>
</tr>
</thead>
<tbody>
<tr>
<td>L7E7</td>
<td></td>
<td>E7E7</td>
<td>S7E7</td>
<td>W7E7</td>
<td>R7E7</td>
<td>O7E7</td>
</tr>
<tr>
<td>L7S7</td>
<td></td>
<td>E7S7</td>
<td>S7S7</td>
<td>W7S7</td>
<td>R7S7</td>
<td>O7S7</td>
</tr>
<tr>
<td>L7W7</td>
<td></td>
<td>E7W7</td>
<td>S7W7</td>
<td>W7W7</td>
<td>R7W7</td>
<td>O7W7</td>
</tr>
<tr>
<td>L7R7</td>
<td></td>
<td>E7R7</td>
<td>S7R7</td>
<td>W7R7</td>
<td>R7R7</td>
<td>O7R7</td>
</tr>
<tr>
<td>L7O7</td>
<td></td>
<td>E7O7</td>
<td>S7O7</td>
<td>W7O7</td>
<td>R7O7</td>
<td>O7O7</td>
</tr>
</tbody>
</table>

So the filter mask for L7W7 will be,

$$\begin{bmatrix}
-1 \\
0 \\
3 \\
0 \\
-3 \\
0 \\
1
\end{bmatrix} \times \begin{bmatrix}
1 & 6 & 15 & 20 & 15 & 6 & 1
\end{bmatrix} = \begin{bmatrix}
-1 & -6 & -15 & -20 & -15 & -6 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 18 & 45 & 80 & 45 & 18 & 3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-3 & -18 & -45 & -80 & -45 & -18 & -3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 6 & 15 & 20 & 15 & 6 & 1
\end{bmatrix}$$

Figure 5.1 Filter Mask obtained by Convolving the L7 and W7 Vectors
5.1.1 Laws’ Texture Feature Extraction Algorithm

The algorithm used here is a combined result of various approaches often applied for mammogram images [2, 47,48].

- **Apply convolution kernels:** On a given ROI the texture analysis is performed by convolving the image with each of the 36 2-D filter kernel.

- **Perform Windowing Operation:** In order to calculate the Texture Energy Measure at each pixel, we average out the absolute values in a 15 X 15 square window (box-car averaging of the absolute value data)

- **Combine Similar Features:** Thus a set of 36 2-D filter kernels produces 36 resultant images for single ROI. In order reduce the number of resultant images per ROI; a similar features combination technique is used. Combining the similar features removes the bias from ‘directionality’, e.g. L7E7 is sensitive to vertical edge and E7L7 to horizontal on combination results in a single component for the edge. Hence, all the features are added together with their transpose convolution kernels (This indicates switch the order of the direct product and combine: L7E7 combined with E7L7 in accord with the previous step. Thus, a set of 21 images is obtained without loss of texture information.)
5.1.2 Feature Vector

As shown in Figure 5.2 the result of Laws Texture Analysis after combination of similar features consists of 21 texture feature images. Thus, a single pixel in ROI will have 21 relevant
features in addition to the intensity feature. And this information about each pixel is represented using a vector of size 1X22. This vector is called as feature vector. And the set of feature vectors formed by each ROI is called feature space. Thus a given ROI of size D1 X D2 is now a cluster of points represented by a 3–D vector of size D1 X D2 X 22. A mixture model representation may be used to model this feature space. then is clustered into two groups one belonging to mass and other to a normal tissue using EM algorithm (as explained in following section).

5.2 Expectation Maximization

The feature space obtained above is derived using the statistical properties. And there might be a possibility of interdependency in the features derived. So, a clustering method based on a statistical model may help than any other distance-based method e.g. k-means. Clustering methods based on statistical model are called as a Probabilistic Clustering Method [25]. EM Algorithm is one such technique for probabilistic clustering.

The Expectation Maximization Algorithm was first introduced by Dempster, Laird, and Rubin [42]. It is an iterative method, which tries to estimate the probabilities for a data point to be in a cluster and then updates the parameters (mean and covariance matrices) to maximize the mixture likelihood. The algorithm is randomly initialized and continues with its iterations as long as the parameter estimate differs by a certain amount from one to the next iteration (stopping criteria). The result is the means and covariance matrices for the clusters.

5.2.1 Mixture Model Estimation

Figure 5.3 and Figure 5.4 shows an example mass and its pixel value distribution (histogram plots). The mixture model provides a better approximation for these non-Gaussian distributions as compared to a single Gaussian model [43].
For this thesis work, we assume a mixture model formed by the feature space to be the combination of K Gaussians. In other words, the model can be broken into k classes, \{1, 2...k\}, with some prior probabilities \(w_1, w_2, ..., w_k\) of a random point belonging to the associated class. And since each class represents a Gaussian distribution, the probability of each point in image data is given as

\[
f(x | \phi) = \sum_{h=1}^{k} w_h f_h(x | \phi_h)
\]

where x is a feature vector, \(w_h\) represents mixing weights or the priors \(w_h\) for \(k^{th}\) classes \((\sum_{h=1}^{k} w_h = 1)\), \(\phi\) represents collection of parameters \((\phi_1, ..., \phi_k)\) means and covariance matrix in this case and \(f_h\) is a multivariate Gaussian Density Function given as,

\[
f_h(x | \phi_h) = f_h(x | \mu_h, \Sigma_h)
\]

\[
f_h(x | \mu_h, C_h) = \frac{1}{\sqrt{(2\pi)^d | C_h |}} \exp\left(-\frac{1}{2}(x - \mu_h)^T C_h^{-1}(x - \mu_h)\right)
\]
where $\mu_h$ stands for mean and $C_h$ for covariance matrix of size $d \times d$. (‘d’ is the dimension of the Laws’ texture feature space.)

### 5.2.2 Expectation Maximization Algorithm

- **Initialization:** The algorithm can start with any initial values for $K$ mean vectors, and $K$ covariance matrices to represent each of the $K$ groups. However, to achieve better results compared to random initialization we set the initial co-variances to the identity matrix and the mean vector to a mean value from the data. This does not mean that exact initialization is necessary for the success of segmentation.

- **Update Equations:** The updating combines the two steps of EM: (1) the expectation and (2) the maximization [44].

For mixing weights,

$$w_{n}^{new} = \frac{1}{N} \sum_{v=1}^{N} p(h \mid x_v, \phi^{old}),$$

for the means,

$$\mu_{h}^{new} = \frac{\sum_{v=1}^{N} x_v p(h \mid x_v, \phi^{old})}{\sum_{v=1}^{N} p(h \mid x_v, \phi^{old})},$$

and for the covariance matrices,

$$\sum_{h}^{new} = \frac{\sum_{v=1}^{N} p(h \mid x_v, \phi^{old})(x_v - \mu_{h}^{new})(x_v - \mu_{h}^{new})^T}{\sum_{v=1}^{N} p(h \mid x_v, \phi^{old})},$$

where $N$ is the total number of feature vectors, and the $p(h \mid x_v, \phi)$ is the probability that the pixel $x_v$ is from class $h$, given the data $\phi$.
\[ p(h \mid x_v, \phi) = \frac{\sum_{k=1}^{K} w_k f_k(x_v \mid \phi_k)}{\sum_{k=1}^{K} w_k f_k(x_v \mid \phi_k)} \]

- **Stopping Criteria:** The algorithm starts with random initialization and continues with its iterations as long as the parameter estimate differs by a certain amount from one to the next iteration. This certain amount in our case is determined using the log likelihood as shown below,

\[
\log L (\phi \mid x) = \log \prod_{k=1}^{N} f(x \mid \phi_k)
\]

The above update equations are repeated until the log likelihood increases by less than 1% from one to the next iteration [44].

### 5.3 Result of EM

The EM Clusters the pixels in two groups with the result as the means and covariance matrices for the clusters. Using these parameters, the posterior probability for each pixel is calculated. The pixel gets assigned to the cluster giving greater posterior probability than the other. This also means that if the posterior probability of a pixel for class 1 is greater than 0.5 then the pixel gets assigned to class 1 and so on. Thus, the resultant image is a binary image with the assumed mass separated from the assumed normal tissue as shown in Figure 5.5 and 5.6.
Figure 5.5 ROI Showing Mammographic Mass (Left)
Figure 5.6 Segmentation Result Using EM on Ripple and Intensity Feature (Right)
CHAPTER 6

EXPERIMENTAL SETUPS AND RESULTS

We have evaluated two aspects of our algorithm, namely the accuracy of segmentation of mass and the usefulness of the Laws’ texture feature. Both of these are evaluated in context of final classification of either benign and malignancy. First, we study how good the automated feature extraction is by comparing the performance of BI-RADS features (without any segmentation) with the automatically extracted features using manual segmentation, including Laws’ texture features. Secondly, we study the performance of automated segmentation method by replacing it with the manually specified segmentation. The performance is measured based on sensitivity and specificity. The sensitivity refers to the probability of detecting cancer when a cancer exists divided by all cancers present in the population at the same time. Specificity refers to the number of normal cases in the population divided by all normal cases.

Table 6.1 provides the classification results obtained using BIRADS, which shows the importance of the shape and boundary features.

<table>
<thead>
<tr>
<th></th>
<th>Benign</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign Prediction</td>
<td>182</td>
<td>26</td>
</tr>
<tr>
<td>Malignant Prediction</td>
<td>21</td>
<td>71</td>
</tr>
</tbody>
</table>

Sensitivity= 73.2 %    Specificity= 89.7 %
The Table 6.2 shows the classification results for automated feature extraction using manual segmentation. These results are obtained using the manually specified segmentation by a radiologist and an automated feature extraction method. 198 different features are extracted by the automated feature extraction method. These features include statistical measures: mean, contrast, lucency, homogeneity, texture measures, wavelet features. As can be seen from the results in Table 6.1 and 6.2, The results obtained using 198 features with the manual segmentation are comparable to the results obtained based on BI-RADS features as observed in Table 6.1 and 6.2. Hence, our goal will be achieved if we can arrive at a segmentation method that produces results comparable to the manual outline.

Before applying the method to whole dataset, a sample of 12 masses out of 300 are selected in order to visually test the performance of the automated segmentation method. Masses were selected so that they form a representative group of the whole collection. Figure 6.1 shows few images from the sample image set along with their manual outline.

<table>
<thead>
<tr>
<th></th>
<th>Benign</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign Prediction</td>
<td>146</td>
<td>19</td>
</tr>
<tr>
<td>Malignant Prediction</td>
<td>9</td>
<td>64</td>
</tr>
</tbody>
</table>

Sensitivity = 77.10%  Specificity = 94.14%
6.1 EM with Intensity

The obvious starting point is to apply the EM method to the pixel (intensity) values without considering additional features. Figure 6.2 shows the results for the sample images. As observed, the EM segmentation separates the image in two parts: (1) one containing mass and (2) the other is normal tissue. But, the method is unsuccessful in separating the only mass tissue from normal tissue. The sensitivity and specificity obtained over the full dataset of 300 masses with automated segmentation is shown in Table 6.3.
In comparison with the manual segmentation results the classification results obtained using intensity based EM were not acceptable especially the sensitivity. The method was unsuccessful in evaluating a malignant mass.
6.2 EM with Laws’ Texture Features

As the intensity based EM was not able to provide acceptable results in case of malignant masses, a Laws’ texture feature analysis was applied as a pre processing technique for mass enhancement. And segmentation is carried out using Expectation Maximization Algorithm. 21 Laws’ Textures Features were considered as explained in chapter 5. This experiment was performed on the sample data set. The results were not satisfactory.

6.3 EM with Laws’ Texture Features and Intensity

Thus, another experiment was carried out with intensity added as 22nd feature to the Laws’ texture features with EM method. The results for this experiment showed no improvement.

6.3.1 EM with Selected Laws’ Texture Features and Intensity

As discussed above, the initial survey indicated that the wave and ripple features (RR and WW) may be useful, although there may be others. Computing time, memory required for processing and the storage space for 300 images (and all the feature images) were a major concern. Evaluation of covariance matrices indicated that these features were independent of each other. Likewise, the RW, RS, SS, and SW were independent. Based on initial findings, the ripple and wave features were picked for further experimentation.

6.3.2 EM with Wave and Ripple feature with Intensity

Two more experiments were also carried out one with WW filter and intensity feature with EM and other with the RR filter and intensity with EM. The classification results for both are as shown in Tables 6.4 and 6.5.
Table 6.4 Classification Results for Segmentation Using EM with Ripple Feature

<table>
<thead>
<tr>
<th></th>
<th>Benign</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign Prediction</td>
<td>138</td>
<td>30</td>
</tr>
<tr>
<td>Malignant Prediction</td>
<td>17</td>
<td>53</td>
</tr>
</tbody>
</table>

Sensitivity=63.85% Specificity=89.03%
Figure 6.4 Sample Image Set with Segmentation Results for Wave Feature

Table 6.5 Classification Results for Segmentation Using EM with Wave Feature

<table>
<thead>
<tr>
<th></th>
<th>Benign</th>
<th>Malignant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign Prediction</td>
<td>178</td>
<td>45</td>
</tr>
<tr>
<td>Malignant Prediction</td>
<td>18</td>
<td>47</td>
</tr>
</tbody>
</table>

Sensitivity=51.07%

Specificity=92.80%
CHAPTER 7

CONCLUSION

In this work, the EM algorithm was developed and implemented in the most general terms. The method is completely coded in the IDL programming language. The algorithm is fully automated and modular. The method takes the raw data, summary data, or any combination as the input and considers the correlation properties of the input data in the decision process.

EM method was applied in conjunction with Laws’ Texture Features with the aims of (1) developing a robust segmentation method if possible and (2) discovering useful features for mass segmentation. Evaluation of the performance of the mass segmentation was based on its classification impact. We analyzed the effect of various Laws’ Texture Features on the mass segmentation and found two features in combination with the intensity feature that produced encouraging results of 90% specificity and 64% sensitivity. The misclassified malignant masses were the more difficult cases in the dataset.

The worked performed here showed that the laws features were generally not useful for mass segmentation purposes in these specific circumstances. However, two features look promising. Given 22 features, there are $2^{22}$ different feature sets that could be considered. Hence, we cannot imply that the Laws’ features are not useful without further experimentation. This work was successful in the development of the EM algorithm, which may be actuated by any user in the future without the intricate understandings of the methods. It is now an easy step to apply the EM on other data at this facility. Future work includes using wavelet-generated features.
REFERENCES


50. Screening and Diagnostic Mammography, URL= “http://www.imaginis.com/”.
APPENDICES
APPENDIX A: GLOSSARY OF STATISTICAL TERMS

Following is a glossary of statistical terms that are used for the basic and advanced audit of a mammography practice, both of which follow the glossary:

- **True Positive (TP):** Cancer diagnosed within one year after a biopsy recommendation based on mammographic examination with abnormal findings (BI-RADS® category 4 and 5).

- **True Negative (TN):** No known diagnosis of cancer within one year of a mammographic examination with normal or probably benign findings (BI-RADS® category 1, 2, and 3).

- **False Negative (FN):** Diagnosis of cancer within one year of a mammographic examination with normal or probably benign findings (BI-RADS® category 1, 2, and 3).

- **False Positive (FP):** No known cancer diagnosis within one year of a positive screening mammographic examination (BI-RADS® category 0, 4, and 5).

- **Positive Predictive Value (PPV) (biopsy recommended):** The percentage of all screening or diagnostic cases recommended for biopsy or surgical consultation (BI-RADS® category 4 and 5) that resulted in the diagnosis of cancer.

\[
PPV_2 = \frac{TP}{TP + FP}
\]

- **Sensitivity:** The probability of detecting a cancer when a cancer exists, or the number of cancers diagnosed after being identified at breast imaging examination in a population within one year of their imaging examination, divided by all cancers present in that population in the same time period.

\[
Sensitivity = \frac{TP}{TP + FN}
\]

[FN is actually a malignant case as per the radiologist]

- **Specificity:** The number of mammographically normal cases in a population divided by all normal cases in the population; or the number of true negative mammograms in a
APPENDIX A (Continued)

population divided by all actual negative cases (those who do not show pathologically proven breast cancer within one year of their screening mammogram) in the population.

Specificity = TN / (FP + TN) Table A. shows the Biopsy results in tabulated form.

Table A. Biopsy Results

<table>
<thead>
<tr>
<th>SCREENING TEST FOR CANCER</th>
<th>Positive (Biopsy demonstrated malignancy)</th>
<th>Negative (Biopsy is benign)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammogram positive (BIRADS® categories 0, 4, 5)</td>
<td>TP</td>
<td>FP</td>
</tr>
<tr>
<td>Mammogram negative (BIRADS® categories 1, 2, 3)</td>
<td>FN</td>
<td>TN</td>
</tr>
</tbody>
</table>

To understanding of above terms, consider following example

Say the radiologist examines 100 abnormality cases and outcome after biopsy is as given below,

True Positive (TP) = 20       False Positive (FP) = 10
False Negative (FN) = 10     True Negative (TN) = 60

Now, the specificity and sensitivity are calculated as follows,

Sensitivity = TP / (TP + FN) = 20 / (20+10) = 0.66
Specificity = TN / (FP + TN) = 60 / (10+60) = 0.85