CHAPTER 2

BACKGROUND AND STATE OF THE ART 
OF BREAST CANCER DETECTION

2.1 INTRODUCTION

Cancerous lesions are difficult to visually detect when they are embedded in breast parenchymal tissues. Therefore CAD methods are being used clinically in breast cancer detection. As the accuracy of CAD is determined by the image interpretation process, it is considered to be the subject of detailed study. The existing CAD systems approved by the U.S. Food and Drug Administration include: Image Checker, Mammokeader and Second Look for x-ray mammography. It is evident that the current CAD systems increase the accuracy of mammography, for some types of breast lesions. The sensitivity and specificity improvements are needed to enable more efficient interpretation that avoid missed cancers and obviate benign biopsies. Thus, there is a substantial room for improvement in CAD systems and more research is needed to improve algorithms at every stage of the detection process.

2.2 SIGNIFICANCE OF MAMMOGRAPHY IN BREAST CANCER DETECTION

Mammography is the main tool used for the detection and diagnosis of breast malignancies. It is a radiographic examination technique that uses X-rays to image breast tissue to detect breast pathology. Though breast imaging can be performed using number of imaging methods available, the
most effective and economical breast imaging modality so far has been X-ray mammography because of its simplicity, portability and low cost. Mammography is a well-developed technology that provides high-quality images at low radiation doses for the majority of patients. Denser tissue absorbs more X-rays and results in a bright region on the image. Mammography attempts to identify structural or morphological differences that can indicate presence of cancer, such as mass, microcalcification and architectural distortion. Conventionally, the image has been recorded on film and of late digital mammography techniques that record the image in digital format have been adopted.

Currently X-ray mammography is the only medical imaging modality used in screening. With about 70% sensitivity and 30% positive predictive value, mammography screening has been shown in clinical trials to reduce breast cancer mortality by 25% to 30% for women in the age group of 50 to 70.

2.2.1 Mammographic Breast Anatomy

The breast is a modified skin organ with its main mass stretched between two layers. The frame work of the organ consists of fibrous and fatty tissue. The relation of the fibrous and fatty tissue within the breast determines the pattern of the mammographic image. The fatty tissue is radiolucent and becomes black on the mammogram, while the fibrous or dense tissue that is radio opaque remains white. The various sections of the mammographic image (A_0412_1.LEFT_MLO) are depicted in the Figure 2.1. The mammographic image contains labels, black background and the breast region. It may also contain scribbling and other noises in addition. The labels are high intensity noises that impede the performance of the detection of breast cancer while black background increases the size of the image. These
irrelevant data present has to be removed and breast region should be extracted to reduce the computation time and for an efficient analysis.

![Mammographic image and its contents](image)

**Figure 2.1** Mammographic image and its contents

### 2.2.2 Mammographic Views

Mammography is a science of imaging and the art of positioning abnormality. A high quality mammogram should exhibit correct positioning, optimal compression, good contrast, adequate exposure, low noise, high sharpness, absence of artifacts, and overall quality. Prime among factors that influence the quality of final image is correct positioning which enables the best possible compression (Heywang-Köbrunner et al. (2001)). In most of the countries mammographic examination consisting of two views Medio Lateral Oblique (MLO) and Carnio Caudal (CC) of right and left breasts are followed as a standard and are depicted in Figure 2.2. In MLO view, X-ray radiations
are provided from the sides of breasts with an oblique angle and image the upper outer quadrant of the breast. While in CC, X-ray radiations are provided from the top of the breasts and image the lower quadrant. Of the two views, MLO is considered more important than CC as in the upper outer quadrant majority of cancers occur. Mini-MIAS and DDSM databases are considered in this research, in which Mini-MIAS consists of only MLO views and DDSM presents both the views.

**Figure 2.2 Mammographic views**

### 2.2.3 Mammographic Abnormalities

Calcifications corresponding to small calcium deposits are one of the findings seen on mammograms. Depending on its size, the calcifications could be categorized into two types: macrocalcifications and microcalcifications. Macrocalcifications are large calcium deposits and are not usually linked to breast cancer. In contrast, microcalcifications are tiny bits of calcium that may show up in clusters or patterns. They consist of brighter spots compared with neighboring normal regions. Tight clusters found in area of rapidly dividing cells can indicate early breast cancer. In breast cancer screening, the characteristics of microcalcifications are the key to estimate the degree of malignancy. However, the image background, the tissue area, and informative marks lead microcalcifications difficult to be recognized and positioned. Even the most experienced radiographer misses presence of cancer that are visible in retrospect due to breast density, human
factors such as fatigue, distraction, oversight and technical limitations. Estimates indicate that 10% to 30% of breast cancers are missed by radiologists during routine screening (Bird et al 1992). Tackling obstacles of perception and interpretation of microcalcifications can be easily achieved by CAD systems. A variety of CAD systems have been proposed to provide a valuable second look to radiologist.

Mass is a space occupying lesion seen in two different projections. The masses are categorized as spiculated, circumscribed and ill defined based on the margin shape. Circumscribed masses have well-defined margins. The margins are sharply demarcated with an abrupt transition between the lesion and the surrounding tissue. Without additional modifiers there is nothing to suggest infiltration. In indistinct or ill defined masses the margin are poorly defined and raises concern that there may be infiltration by the lesion. In spiculated masses the lesion is characterized by lines radiating from the margins of a mass (D'Orsi et al 2003).

One of the most missed signs of breast cancer is Architectural Distortion. According to BI-RADS, Architectural Distortion is defined as the normal architecture distorted with no definite mass visible. This includes spiculations radiating from a point and focal retraction or distortion at the edge of the parenchyma. Architectural Distortion was the most commonly missed abnormality in false-negative cases. Architectural distortion can also be an associated finding.

Bilateral Asymmetry is another indication to detect breast cancer by radiologist. This is a density that cannot be accurately described using the other shapes. It is visible as asymmetry of tissue density with similar shape on two views, but completely lacking borders and the conspicuity of a true mass. It could represent an island of normal breast, but it lacks the specific benign characteristics that may warrant further evaluation. In bilateral asymmetry
cases, the left and right breast regions appear different from each other in the corresponding mammogram images. Additional imaging may reveal a true mass or significant architectural distortion. The regions showing different types of abnormalities that may appear in mammogram are provided in Figure 2.3. Similarly the regions of benign and malignant tissue are depicted in Figure 2.4.

![Mammogram Images](image)

**Figure 2.3** Types of abnormalities in mammograms (a) Normal (b) Circumscribed mass (c) Illdefined Mass (d) Microcalcification (e) Spiculated (f) Architectural Distortion and (g) Breast Assymetry

![Mammogram Images](image)

**Figure 2.4** Benign and malignant microcalcifications (a) Benign & (b) Malignant
2.2.4 BI-RADS Assessment Categories

Mammographic interpretation is ultimately a medical art, used to detect abnormalities in breast tissue that are potential indicators of breast cancer. The current clinical protocol followed by the radiologist for mammographic examination and interpretation are from the American College of Radiology BI-RADS lexicon, (Breast Imaging Reporting and Data Systems) (D'Orsi et al 2003). This standard describes the characteristics of a given abnormality.

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 mammogram images corresponding to 5 point assessment categories of BI-RADS are shown in Figure 2.8 and are described as follows.

- **Category 0: Incomplete Assessment**

  Needs additional imaging evaluation. Figure 2.5 (a) is a mammogram image assessed as BI-RADS 0 at screening.

- **Category 1: Negative**

  In Figure 2.5 (b) the breasts are symmetrical and no abnormalities are present and therefore assessed as BI-RADS 1.

- **Category 2: Benign finding**

  Figure 2.5 (c) is also a negative mammogram. But the interpreter may choose to describe the finding 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**

  Figure 2.5 (d) and (e) both are Category 3 type mammogram images with different final assessments. In Figure 2.5 (d) a mammogram containing non-palpable sharply defined lesion with a cluster of calcifications is shown with follow up images of 6, 12 and 24 months. As there is no change in the follow up images, the final assessment changed to Category 2. Figure 2.5 (e) depicts the two mammogram images, upper containing two calcifications and lower is 12 month follow up. Here the final diagnosis is invasive carcinoma.

• **Category 4: Suspicious Abnormalities**

  Biopsy should be considered. The lesion shown in Figure 2.5 (f) does 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. The lesion shown in Figure 2.5 (g) has a high probability of being cancerous.

![Figure 2.5 (Continued)](image-url)
Figure 2.5 BI-RADS Assessment categories (a) Category 0, (b) Category 1, (c) Category 2, (d) & (e) Category 3, (f) Category 4 and (g) Category 5

2.3 OVERVIEW OF EXISTING METHODS

Though there are various types of abnormalities viewed in mammograms, the most important types of mammographic abnormalities are mass and microcalcifications. In the following section the existing methods to detect microcalcifications and mass are presented in detail.

2.3.1 Microcalcification Detection

Microcalcifications are tiny granule like deposits of calcium. They are relatively small and dense in nature in comparison with the surrounding normal tissue (Monsees 1995). Clusters of microcalcifications are the best
indicators for malignancy in mammograms (Picca & Ellen Shaw de Paredes 2003). In view of the above it is evident that methods for automatic detection of microcalcifications in mammograms are desired in order to assist radiologist in the interpretation of mammograms and the diagnosis of breast cancer. For the past two decades numerous research studies are focused on automated microcalcification interpretation (Karssemeijer et al (1991), Mascio et al (1993), Barman et al (1994) and Netch et al (1999)). Conventionally, simple morphological techniques have been used to detect microcalcifications. Nevertheless, due to the small size and variable appearance of microcalcifications these methods are unable to accurately detect the breast cancer. Therefore advanced image analysis has been and still being developed to increase the rate of interpretation.

Yu et al (2000) proposed an automatic system for detection of clustered microcalcifications using neural network classifiers through two steps. In the first step, potential microcalcification pixels are segmented using mixed features and in the second step, checked the potential objects using 31 statistical features. Mousa et al (2005) proposed a system using wavelet analysis and Adaptive Neuro-Fuzzy Inference System for building a classifier to distinguish normal from abnormal and to determine whether the type of abnormality is mass or microcalcification. The features were extracted by summing a predefined number of energy values together. The given results showed a successful classification rate. Bouyahia et al (2009) proposed wavelet based techniques for automatic detection of microcalcifications using undecimated wavelet transform, multi-scale product and wavelet packets transform.

Recently, the Dual-Tree Complex Wavelet Transform has shown a good performance in applications that involve image processing due to more phase information in data, shift invariance, and directionality than other
wavelet transforms. Alarcon-Aquino et al (2009) proposed an approach to
detect microcalcifications in digital mammograms using the Dual-Tree
Complex Wavelet Transform. Mini-MIAS database images were used to test
the approach and found that it performs better than stationary and discrete
wavelet transform. Rizzi et al (2012) have presented a study on various
methods available for betterment of CAD systems in microcalcification
detection as a motive to improve the diagnostic tools and techniques.

Xiaoyong Zhang et al (2013) proposed a hybrid image filtering
method for Computer Aided Detection of microcalcification clusters in
mammograms. The microcalcifications are first enhanced and then the
candidates of microcalcifications are refined by a multilevel wavelet
reconstruction approach. The proposed method detected 92.9% of true
microcalcification clusters with an average of 0.08 false microcalcification
clusters per image based on the distribution feature of microcalcifications.
Cheng et al (2003) and Sakka et al (2006) provide a comprehensive review of
microcalcification diagnosis algorithms in their work.

2.3.2 Mass Detection

In mammograms, masses can be characterized as bright spherical
objects as the X-ray attenuation coefficient of tumorous tissue is larger than
that of normal breast tissue. The actual contrast of mass lesions is highly
variable due to the variability of background structures and mass sizes.
Masses can be classified into different types based on its shape and margin.
Masses are malignant in nature when it posses irregular and spiculated
margins, architectural distortion and asymmetry in shapes. The detection of
masses is a challenging task in screening mammography when its boundaries
are irregular and fuzzy. In addition noises present in mammograms further
decreases detection rate. A lot of research work based on various theories has
been carried out to tackle the problem of computerized mass detection

Christoyianni et al (2000) presented a method for fast detection of circumscribed masses in mammograms. Their method performs a binary classification of mammographic mass regions into cancerous and normal tissue using a radial basis function neural network, and provides the location of the mass by employing sub-image windowing analysis. The efficiencies are 90.9%, 62.5%, and 33.3% in fatty, glandular, and dense tissue, respectively. The first observable problem is that it was highly sensitive to the background tissue. A second problem is that it was designed to detect only circumscribed masses. Varela et al (2007) proposed a system to detect malignant masses on mammograms using an iris filter at different scales on suspicious regions and a back propagation neural network classifier. The system yielded a sensitivity of 88% for lesion-based evaluation and 94% at 1.02 false positives per image for case-based evaluation.

Rojas & Nandi (2008) presented a method for automatic detection of mammographic masses. First the contrast of the image was enhanced by local statistical measures. Secondly segmentation was done by multilevel thresholding and then the features were extracted. Finally the masses were detected by region-ranking method with 80% detection sensitivity and 2.3% false positives per image.

2.4. REVIEW ON ALGORITHMS FOR MAMMOGRAM ANALYSIS

2.4.1 Introduction

Digital image processing algorithms in CAD tools are intended to help radiologists in improving diagnosis with the aid of computer systems. These CAD tools increase reader’s sensitivity by 10% by improving image
quality, identifying malignant signs, enhancing mammography features, etc. (Nishikawa 1998). Though CAD systems for breast cancer detection have been studied for more than three decades, automated interpretation of abnormalities such as microcalcifications and masses remains a challenge as they have almost the same brightness level as the background tissue. In view of the above problems many works are addressed to improve radiologists’ performance in detecting clustered microcalcifications and masses. Recent research on this topic has been focused on improving following one or more methods stated: image preprocessing and enhancement, segmentation and ROI extraction, feature extraction and selection, detection and classification of abnormality. A review of aforementioned methods is essential and are discussed below.

2.4.2 Mammogram Preprocessing and Enhancement

Usually, the preprocessing includes digitization of the mammograms with different sampling and quantization rates. Then, the ROIs selected from the digitized mammogram are de-noised and enhanced. Fractal modeling of mammograms was proposed by Li et al (1997) for enhancement of microcalcifications. Gurcan et al (1997) explored the idea of applying a sub-band decomposition to enhance microcalcifications. Nam and Choi et al (1998) applied histogram equalization and contrast enhancement to the problem of detecting microcalcifications from mammogram images. Heinlein et al (2003) have improved sub-band decomposition of the image by deriving a wavelet transform specifically adapted to the size and form of the microcalcifications. Wirth et al (2004) enhanced microcalcifications based on morphology. Jiang et al (2005) proposed a structured tensor operator and a fuzzy enhancement operator for the suppression of non-microcalcification regions. Papadopoulos et al (2008) used five image enhancement algorithms such as Contrast Limited Adaptive Histogram Equalization, Local Range
Modification and Redundant Discrete Wavelet linear stretching and shrinkage algorithms to test the effect of image enhancement in detection of microcalcifications.


2.4.3 Segmentation and ROI Extraction

The segmentation and ROI extraction is used to find suspicious areas containing masses from the background that will be used for extracting features from mammograms. Sheshadri et al (2005) proposed a lesion segmentation technique applying watershed algorithm. Rabottino et al (2011) proposed an algorithm for mass lesion segmentation based on region growing technique that is independent of mass size and also provides a performance evaluation procedure.

As the segmentation of microcalcifications from mammogram image strongly influences classification results, it has been studied by various researchers to aid radiologist in detection. Halkiotis et al (2007) segmented microcalcifications from their background through evaluation of their topographic representation using the mathematical morphology. Malek et al (2010) proposed an image segmentation technique by combining seed based region growing and boundary segmentation in sequential order. First an initial seed point is identified in region growing process and then boundary
segmentation technique is implemented in order to improve the segmentation results.

2.4.4 Feature Extraction and Selection

The extraction of discriminating feature that can efficiently detect mass is an important issue in the mammogram analysis. Liu et al (2001) formulated a multi-resolution procedure for the detection of spiculated lesions in digital mammograms using a linear-phase, non-separable, 2-D wavelet transform. Pixel based features were extracted at each resolution level and analyzed to determine the spiculated lesions. Moayedi et al (2010) employed contourlet transform as a feature extractor to obtain the contourlet coefficients for mass detection. Eltoukhy et al (2010) presents an approach for breast cancer diagnosis in digital mammogram using curvelet transform. Feature extraction process is accomplished by decomposing the mammogram images in curvelet basis and a special set of the biggest coefficients is extracted as feature vector.

Several efforts have been directed toward finding features that can adequately detect microcalcifications and characterize them into malignant or benign. Extracted features can also facilitate distinguishing them from other structures in mammograms. Various approaches have been applied in order to achieve a higher accuracy in such analysis. Sheshadri & Kandaswamy (2006) proposed a new technique based on the use of digital filters together with filter response energy measure as texture feature extractors for detection of microcalcification. Alarcon-Aquino et al., (2009) proposed an approach to detect microcalcifications in digital mammograms using the Dual-Tree Complex Wavelet Transform. Mini-MIAS database images were used to test the approach and found that the DT-CWT performs better than stationary and discrete wavelet transform. Amir Tahmasbi et al (2013) used Zernike moments as feature vector to classify masses into benign and malignant.
In the literature, researches related to feature selection mostly use traditional methods like Fisher’s discriminate function (Gulsrud & Husoy 2001, Mu et al (2008)). The selected features are used as inputs to neural networks in order to improve the classification process. Jiang et al (2007), proposed the use of genetic algorithms to select the most discriminate features for classification. Sun et al (2005) compared different feature selection methods for breast cancer detection in mammograms.

2.4.5 Detection and Classification

A lot of works were confined towards detection of microcalcification and masses. After the detection of abnormalities, they need to be classified into benign or malignant by the CAD tools. Extensive research is being conducted towards classification of microcalcifications and mass based on the features selected. Mousa et al (2005) proposed a system based on wavelet analysis and used the Adaptive Neuro - Fuzzy Inference System for building a classifier to distinguish normal from abnormal and to determine whether the type of abnormality is mass or microcalcification. Christoyianni et al (2000) presented a method for fast detection of circumscribed masses in mammograms using a radial basis function neural network, and then provides the location of the masses by employing sub-image windowing analysis. Rojas & Nandi (2008) presented a method for automatic detection of mammographic masses using region-ranking method with 80% detection sensitivity and 2.3% false positives per image. Eltoukhy et al. (2010) presented a study where Euclidean distance classifier is used for classification of cancerous images from normal images.

As the detection of microcalcifications is a very difficult process a lot of research work is under progress in the development of advanced image processing algorithms. Wu et al (1992), were the first author to use neural network to detect potential microcalcifications and reduce false positives.