CHAPTER 1

INTRODUCTION

1.1 OVERVIEW OF BREAST CANCER

Breast cancer is the most prevalent cancer among women and it continues to be a significant public health problem around the world. It is estimated that around 1.67 million women across the world were diagnosed with breast cancer in 2012, accounting for nearly a quarter (25%) of all cancers diagnosed in women. Every year more than 500,000 women die of breast cancer (GLOBOCAN Cancer Fact sheet 2012). In India, breast cancer seems to be a growing problem which estimates to be high, as one in 22 women is predicted to develop the disease (Michigan News, 2013). Even more disturbing is the fact that on an average one out of eight women will develop breast cancer at some point during their lifetime (SEER Cancer Statistics Review 2012).

Since the cause of breast cancer is still unknown, the earlier the cancer is detected, the better the chance for a proper treatment. Breast cancer can be characterized into six different types: Architectural Distortion, Breast Asymmetry, Circumcised mass, Spiculated mass, Ill-defined mass and Microcalcification. Microcalcifications are tiny specks of calcium deposits that are scattered throughout the mammary gland, or occur in clusters. Masses are space-occupying lesions, described by their shapes, margins, and density. Spiculated lesion is characterized by lines radiating from the margins of the mass. Asymmetry of breast in terms of volume
indicates breast cancer. Risk analysis reveals the nature of the cancer, malignant or benign. Benign is a mild and non-progressive tumor that can be cured while Malignant is a dangerous and progressive tumor.

Breast density is an important measure which describes the possibility for the detection of abnormalities in mammograms. Higher breast density usually indicates a higher possibility for the presence of malignant tissue. Based on the tissue density, breasts are classified into Fatty, Fatty glandular and Fatty dense. The following section describes different imaging modalities used for breast cancer detection.

1.2 MEDICAL IMAGING MODALITIES FOR BREAST CANCER DIAGNOSIS

1.2.1 Mammography

X-ray film-screen mammography is a specialized radiographic imaging method used for detection of breast cancer in the early stages. Mammography is a planar radiographic imaging method that provides 2D projected image of breast on a film that is exposed during imaging through an external radiation source of X-rays. The exposed film is then developed to produce a high quality analog image of the organ. Breast tissue is quite vascular and soft with low X-ray attenuation coefficients that impose much more challenging requirements when compared to other anatomical structures. Abnormal tissues have higher X-ray attenuation than the normal ones that lead to difference in contrast in the image. The localization of abnormalities such as microcalcifications is relatively a difficult process, as they appear with low contrast and require high spatial resolution of the order of 50 - 100 microns. Mammographic signs are a useful input to radiologists, that need to have high sensitivity and specificity for early detection of breast cancer, but at the same time mammographic imaging techniques must
minimize scattering radiation due to non-invasive concerns and decrement in the radiation dose. To facilitate optimal diagnostic capability recent advanced X-ray film-screen imaging methods use specialized X-ray tubes, breast compression devices, anti-scatter grids and optimized detector systems along with computer controlled processing. In recent film-screen mammography scanners, a double screen with double film emulsions is used for breast imaging. However, digital X-ray mammogram scanners are also replacing film-screen with scintillation crystal and semiconductor technology based digital detector systems.

1.2.2 Digital Breast Tomosynthesis

The 2D projections may be adequate for many diagnostic applications, however the qualitative and quantitative information about the anatomical structures and associated pathology are not available. Digital Breast Tomosynthesis is a 3D breast imaging technology that acquires images of a stationary compressed breast at multiple angles during a short scan. The principle advantage is that a particular depth-of-field of an abnormal region may be brought into focus while other regions are blurred. Along with advances in image processing algorithms, it provides a very useful and sophisticated imaging tool with possibility of reduced breast compression, improved diagnostic and screening accuracy, 3D lesion localization, and contrast-enhanced 3D imaging. The increased radiation dose, longer acquisition and processing times are the main drawbacks.

1.2.3 Breast Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a complex multidimensional imaging modality that is rapidly becoming a standard of care in breast cancer evaluation. It uses magnetization and radiowaves instead of x-rays to produce very detailed, cross-sectional images. MRI uses the Nuclear Magnetic
Resonance property that is based on the quantum properties of nuclei and protons. MRI provides high resolution images with different contrast features by mapping the spatial location and associated properties of hydrogen protons present in the object being imaged. During breast MRI examination, a contrast material Gadolinium DTPA is injected into a small vein in the arm before or during the examination to improve the ability of MRI to clearly show details of breast tissue. Major challenge of MRI is standardization of inter-scan and intra-scan images. And also movement of patients lead to complications and unreliability, therefore highest level of patient compliance is essential between scans. Breast MRI has the ability to image dense breasts and can give dynamic information about angiogenesis and therefore it is effective for all classes of women. In addition MRI is less accurate in determining the area which is abnormal and which is not. Further MRI is more costly, requires standard examination techniques and interpretation norms, cannot detect microcalcifications, and perhaps has a higher false positive rate (American Cancer Society 2005).

1.2.4 Breast ultrasound

Today safety, portability and low cost aspects makes ultra sound imaging a successful diagnostic imaging modality for imaging anatomical structures, blood flow measurements, and tissue characterization. A conventional ultrasound imaging system comprised of a piezo electric crystal based transducer for transmission and reception of pulses, with a computer processing and display system. The intensity of the received echoes is converted into voltage signal that generates the raw data for the image. The voltage signal is then digitized and processed. Breast ultrasound images are only sectional images, therefore both acquisition and interpretation needs experienced physicians. And also ultrasound imaging is not suitable for screening because too many images would be required for interpretation and
would be time consuming. As ultrasound images appear noisy with speckles and lacks boundary definition of the object structure, the quantification and interpretation of the object is more challenging when compared to other imaging modalities.

1.2.5 Molecular Imaging

It is a functional imaging modality that provides metabolic information about an organ or tissue of the body by involving the tissue itself in the imaging process. These methods are also called emission imaging methods as they utilize radioactivity that disintegrates an unstable nucleus into a stable nucleus by releasing nuclear energy and emitting photons such as gamma photons and/or specific particles such as positrons and alpha particles.

Gamma ray imaging also known as Single Photon Emission Computed tomography is used for 3-D imaging of human organs in which radioisotopes injected into the body metabolize the tissue and make them a source of gamma-ray emissions. The emitted gamma rays captured by the detectors are used as raw data for defining projections and the projected data is reconstructed into images with the help of a computer and high-resolution displays. It is better in assessing metastases, while poor in structural information due to attenuation and scattering problems.

Positron Emission Tomography provides specific information about the biochemical activity and physiological behavior of radioactive material that is metabolized in the tissue. It provides significant information about glucose metabolism and blood flow of the tissue that is critical in determining the heterogeneity and invasiveness of tumors. However, Positron Emission Tomography is much more expensive than Single Photon Emission Computed Tomography and also reconstructed images are noisy and poor in resolution.
1.3 MOTIVATION

Early detection of breast cancer is critical, as therapeutic actions likely to be successful in the early stage of the disease. Mammography is the best radiological technique currently available for early detection of breast cancer. However some of the breast lesions are missed during screening by radiologist as they need to interpret large number of mammograms in screening programs. Computer aided techniques have been developed to help radiologists to improve detection process. The motivation of the thesis is to exploit recent developments of image processing techniques in the different stages of Computer Aided Diagnosis (CAD) systems for accurate detection and characterization of breast cancer.

Texture is characterized by the spatial distribution of gray levels in a neighborhood. In mammograms different tissues have significantly different textures. Therefore local texture characteristics can differentiate benign and malignant tissues. For CAD systems, the diagnosis is done often using decision making algorithms based on extracted features. In breast cancer detection, feature extraction using texture analysis has demonstrated that the efficiency, accuracy and consistency of any decision making algorithm is appreciably improved. But the latest developments in texture analysis methods like curvelet transform, Wave Atom transform has not been used much in mammogram image analysis.

Similarly, though automated decision making using neural networks is shown effective in breast cancer detection, the algorithms developed before two decades are employed in the existing research works. Most of these studies used back propagation neural networks or statistical methods focusing more on feature extraction. Besides optimization of neural network architecture to be used for classification is also essential for efficient classification but until now only some works are published towards optimizing classifiers.
Therefore employing recent developments of image processing algorithms in mammogram analysis, to improve breast cancer detection and reduce diagnosis errors form the motivation of the proposed work. In this thesis advanced tissue texture analysis using multiresolution transforms like Wavelets, Curvelets, Wave Atom and sophisticated classifiers like Extreme Learning Machine (ELM), Circular Complex- Extreme Learning Machine (CC-ELM), Support Vector Machine (SVM) that can be easily optimized to the problem defined are proposed for different stages of mammogram image analysis.

1.4 AIM AND OBJECTIVES

The main goal of the proposed research is to develop an image processing algorithm that can detect the abnormalities at an early stage from screening mammograms. The main modules of the entire work flow are scheduled below.

- Primitive segmentation for Region of Interest (ROI) extraction
- Feature extraction using texture analysis
- Performance study of different classification algorithms
- Detection of mass and microcalcification
- Classification of breast tissue into benign or malignant and different densities

The work proposed here is an automated system for early identification of Breast carcinoma that employs advanced and novel image processing technique, the multiresolution analysis. The digitized mammograms are considered for the analysis and are first preprocessed to eliminate background and noises that may be present in mammograms. Then
features are extracted and various tissue texture analyses have been carried out to prove the effectiveness of the proposed methodology in detecting microcalcifications and masses. The next phase namely detection and classification cover the main processing step of the computer aided detection system which is considered in this research. The intended use of this system is a second opinion strategy for the radiologist.

1.5 METHODOLOGY

Mammography is a standard technique to detect breast cancer. However, x-ray mammography has limited specificity and sensitivity. Texture analysis of mammogram images improves diagnostic and screening accuracy. Multiresolution analysis of image texture is a sophisticated method that aids lesion localization and characterization. In this thesis identification of microcalcified regions from normal regions using texture analysis is first discussed. Then the detected microcalcifications are characterized as malignant or benign using multiresolution analysis. The proposed algorithm has been implemented and tested on images obtained from Mammographic Image Analysis society (Mini-MIAS 1994), and Digital Database for Screening Mammography (DDSM 2001) databases. The following are the specific processes involved in the detection and characterization of microcalcifications and masses present in mammograms.

- Image processing for breast region segmentation
- Extraction of normal and abnormal regions
- Statistical and spectral feature extraction of mammograms
- Implementation of different classification algorithms for normal and abnormal detection
- Selection of best feature and optimal classifier based on experimental investigations
- Identification of abnormal region in mammogram
- Multiresolution analysis of abnormal regions
- Characterization of breast tissue

The performances of designed algorithms for different processes considered in the research are validated using Receiver Operating Characteristics (ROC) analysis (Zweig & Campbell 1993). The various algorithms used in the proposed research work are summarized in Figure 1.1.

**Figure 1.1 The proposed algorithms for mammogram analysis**
1.6 CONTRIBUTION TO THESIS

The proposed research work focused on the development of image processing algorithms for the computer aided detection of breast cancer using mammograms. The major contributions of this thesis are

- A new approach to discriminate the microcalcifications from the normal tissue using ELM and comparison of different feature vectors extracted using Wavelet transform, Gray Level Spatial Dependence Matrix (GLSDM), Gabor transform, Curvelet transform and Maximum Overlapping Wavelet Transform (MOWT).

- Implementation of different classifiers namely Bayes Net classifier, Naïve Bayes classifier, SVM, Sequential Resource Allocation Network (SRAN), ELM and Phase Encoded Extreme Learning Machine (PECELM) in the detection of microcalcifications.

- Automatic characterization of the detected microcalcification regions into benign and malignant using Wave Atom transform and CC-ELM.

- Detection and characterization of both microcalcifications and circumscribed masses from normal breast regions using MOWT and ELM.

- A GUI based comprehensive tool for automated detection, characterization, and risk nature analysis of breast cancer combined with tissue density analysis. Here different transforms like Wavelet, Curvelet and Contourlet transforms
are applied to the image and the resultant image is decomposed into different scales, best coefficients from each scale are retained as feature vector. These feature vectors are classified using SVM.

1.7 THESIS OUTLINE

The thesis is organized as follows:

The chapter 2 presents introduction to breast cancer detection, significance of Mammography in Breast Cancer detection, Mammogram image content details, Mammographic views and abnormalities. It also presents literature review on algorithms proposed by various researchers for mammogram analysis.

In chapter 3 the novel technique developed for automated microcalcification detection using ELM and wavelet transform is presented. The details of database used, image preprocessing and ROI extraction are provided. It also explains the design, implementation and evaluation of proposed CAD system for microcalcification detection. The artificial neural network that is used for automation of the CAD system, texture feature extraction algorithm, comparison of different feature extraction methods and various classification techniques are presented. In addition the chapter describes the performance evaluation techniques used in the proposed method.

The chapter 4 describes the extension of the proposed algorithm for detection and characterization of masses. It presents the design, implementation and evaluation of the CAD system for discriminating mass, microcalcification and normal regions from mammogram images. Employing
MOWT and ELM, the abnormality detection and risk analysis of the detected abnormality are done. It demonstrates the efficacy of the proposed design in improvement of classification accuracy. It also presents the analysis on the experimental results and conveys discussion.

The chapter 5 brief about the GUI developed for breast cancer analysis. In this chapter, a comprehensive system is proposed for complete analysis of breast cancer from primary abnormality detection to final level of risk identification. To accomplish the objective, four cancer analyses namely abnormality, characterization, risk and tissue density analysis are performed using mammogram images.

The chapter 6 proposes the improved microcalcification detection method along with risk analysis. First SRAN based detection of microcalcification accomplished using segmentation of breast region, statistical feature extraction and classification is presented. Then subsequently, Curvelet transform and PECELM employed to improve the performance of detection process are explained. Next the characterization of microcalcification into benign and malignant using Wave Atom transform and CC-ELM are described.

Chapter 7 includes the conclusion and focuses on the root work for the further enhancement of the CAD system. The different breast cancer detection methods considered in the proposed research work are summarized in Figure 1.2. First detection of microcalcifications regions from normal regions is discussed. Then detection mass and microcalcification regions from normal regions are described. Further a GUI developed for cancer analysis, risk analysis and tissue density analysis is explained.
Figure 1.2 The proposed breast cancer detection system