ABSTRACT

One of the outcomes of the recent advancements in science and technologies is the increase of life expectancy of humans. Apart from the mortality due to natural aging, ailment related deaths are also of major concern. In this regard, there is a necessity to detect and diagnose any disease at an early stage which could be appropriately tackled.

The Indian Council for Medical Research (ICMR), in its 2016 report on breast cancer research and analysis, has stated that the breast cancer victims are estimated to be over 1.5 lakhs, which covers 10% of all other cancers. Hence, the breast cancer demands primary attention among its counterparts [ICMR].

Microcalcifications are the tiny specks of calcium deposits found in the breast as single or in clusters. Generally the digital representation of breast cancer has been made up by low dense X-rays and the images are stored in a set of digital forms taken by different viewing/angled positions. These digital forms referred as mammograms, have become a reliable source of breast cancer screening, the analysis of which needs some sequence of processes in order to detect the breast cancer at an early stage.

The computer aided techniques support the radiologists as a second opinion for the early detection of cancerous regions. The computer-assisted diagnostic system for digital mammograms assist in addressing the complexity lies in separating the cancerous regions from non-cancerous ones.
The low energy exposure of mammograms results in poor portraying of the important information on the region of interest. Hence, it is necessary to enhance the image contrast to improve the poor effects prevailing in the image. This contrast enhancement would help in the subsequent processing of digital mammograms.

Mammogram segmentation is the next step following contrast enhancement, which segments the test mammogram into its constituents/objects based on the similarity/dissimilarity of the intensity values. It partitions the mammogram into foreground objects and background which simplifies the computational process of feature extraction and classification.

The phase of feature extraction exposes the diverse characteristics of the segmented regions, which in turn serves as the necessary parameters for the classification analysis.

Classification phase in computer-aided diagnosis classifies the mammogram into either as normal or abnormal, based on its intensity characteristics. The devised classification techniques explore the image features extracted through the developed feature extraction techniques using the features depicted by the intensities of pixels.

The methods developed out of this thesis are corroborated using Matlab R2010a. The performance of the methods are validated on the datasets of Mammographic Image Analysis Society (MIAS) and Digital Database for Screening Mammography (DDSM).

The framework of this proposed research is formulated based on the essential requirements of a holistic computational system for breast cancer detection using digital mammograms. It involves
thirteen algorithms that can be grouped under four distinct phases namely

I. Contrast Enhancement Techniques;
II. Mass Segmentation;
III. Feature Extraction; and
IV. Classification of Masses.

The algorithms devised for those phases are listed hereinunder:

**I. Contrast Enhancement Techniques**
1. Local Thresholding based Contrast Enhancement (LTCE)
2. Adaptive Histogram Equalization based Contrast Enhancement (AHECE)
3. Black Top-Hat Transformation and Gauss Distribution based Local Contrast Enhancement (BHTGD)
4. Modified Full Width Half Maximum based Contrast Enhancement (MFCE)
5. Modified Saturation based Contrast Enhancement (MSCE)

**II. Mass Segmentation**
1. Local Threshold based Segmentation (LTS)
2. Nearest Neighbourhood based Region Grown Segmentation (NRGS)
3. Combinational Local Threshold based Segmentation (CLTS)
4. Local Threshold based Seeded Region Growing Segmentation (LTRGS)

**III. Feature Extraction of Masses**
1. Region Characteristics based Feature Extraction (RCFE)
2. Feature Extraction using Statistical and Textural Characteristic(FESTC)

**IV. Classification of Masses**
1. Statistical Measures based Modified KNN Classifier (SMMK)
2. Normalized Measures based Bayesian Classifier (NMBC)
The outcomes of the classification analysis are evaluated using confusion matrix, as True Positive and True Negative and the results are further authenticated in comparison with the ground truth specifications given by MIAS descriptions like Breast Imaging Reporting and Data System (BIRADS) and DDSM.

The methods developed in this research work have been found to be superior over the available/reported methods. It is certain that the application of these techniques for the prognosis of breast cancer would find its place in the domain of medical imaging.