ABSTRACT

“There are an estimated 10000 to 125000 new breast cancer cases in India every year” – *Times of India (Oct 21, 2012)*. Breast cancer continues to be one of the most common cancers and a major cause of death among women all over the world. The root cause of breast cancer is still unknown. Hence there are no preventive measures for this dangerous disease. But early detection and treatment can increase the survival rate up to 95%. At present early diagnosis is the only way to reduce the mortality rate of breast cancer. Currently mammography is found to be the golden technique for early diagnostics and reliable results for breast cancer.

At present mammogram readings are performed by radiologists who will examine the mammogram for the abnormal regions which can be interpreted as breast cancer. But due to variety of reasons such as eye fatigue, misdiagnosis of benign as malignant, the efficiency of diagnosis varies from 65% to 85%. Hence second reading is needed here to acquire more reliable result. Computer Aided Diagnosis (CAD) has been introduced to overcome the above problem and to increase the efficiency in diagnosis. CAD can act as a second reader in malignant diagnosis and it could assist the radiologist for improved and efficient mammogram reading.

In the past two decades tremendous researches has been undergone in designing and implementing improved CAD technique to diagnosis breast cancer. Even then the diagnosis rate has not been improved considerably. A standard CAD technique which could read any kind of mammogram invariantly is not available. The goal of this research work is to design and implement the fully automatic, 100% accurate and reliable breast cancer
diagnosis technique with improved efficiency over the other existing techniques.

To achieve the above goals, initially various image processing techniques has been implemented to design and implement a perfect CAD technique. Three different techniques have been implemented as a part of this research to attain the flawless and improved CAD technique. First two methods are acted as a path way to reach the third method which satisfies the goal of the research in all the aspects. In this research work first two methods are developed by just integrating the existing algorithms which has been named as 3SMDT and MDTC based on the steps to design the CAD, which gives the acceptable result. Final approach MDTR has been designed in such a way to overcome the pitfalls of the previous approaches and to achieve the maximum accuracy in diagnosis.

3SMDT has 3 steps such as Preprocessing, Segmentation and Feature extraction. Preprocessing is implemented by down sampling, quantization and Median filtering. Segmentation is implemented by Gaussian and Gabor filtering. Feature extraction is performed by GLCM on pixel basis. There is no separate classification process is implemented in this approach. The classification is performed manually by calculating the large array of feature extracted values which is the major drawback of this method. Hence in MDTC one more step is included for classification. Preprocessing of MDTC is performed with median filtering and some morphological operations. Segmentation is performed by region growing method, Feature extraction is performed with the same GLCM and Classification is implemented by Support Vector Machine. The semi automatic MDTC method did not considerably improve the efficiency of the diagnosis as expected; which led to more research on Segmentation to improve the efficiency.
Improvisation of the segmentation led to the MDTR method, which has 5 steps in diagnosis of Breast cancer such as Region of Interest (ROI) Extraction, Preprocessing, Segmentation, Feature Extraction and Classification. It is named as Malignancy Detection Technique with Region of interest extraction (MDTR). The special algorithm is designed to extract the suspected region alone from the mammogram in ROI extraction process. Due to this extraction process preprocessing load is minimized to noise removal alone which is performed by median filtering. Edge detection is implemented in this approach using Sobel operator. For segmentation new segmentation algorithm has been innovated and implemented in MDTR. The new algorithm is named as ‘Gray Level Gradient Buffering Algorithm’ since the segmentation is performed based on gray Level Gradient value. This new proposed and implemented algorithm is especially designed to segment the mammogram image in effective manner based on the abnormal region. Thus the algorithm helps to get the improved segmentation which in turn results in improved diagnosis. Feature extraction is performed by the combination of GLCM and LBP techniques. Classification is performed by Support Vector Machine.

For comparison purpose all the three approaches were tested with same set of 400 mammograms which were collected from MIAS, DDSM databases and some real time mammograms were collected from Tamilnadu Government Hospital. The performances of the approaches were evaluated using Partest method. Finally derived MDTR method satisfies the goal of this research work in all aspects such as it produced 99.1% accuracy of diagnosis which is far better than the existing algorithms. One more special advantage of this approach is it would allow the radiologists to upload the mammogram directly and can view the result immediately as message box.
Hence it is fully automatic and radiologists can use this as second reader without any manual interruption and coding knowledge. MDTR has been presented in front of the doctors and a detailed demonstration of this method has been given to the radiologists. Received positive feedbacks from the radiologist for MDTR, which has been submitted along with this thesis. Hence assuredly this innovative MDTR could be more useful as second reader in assisting the radiologist for efficient and early diagnosis of breast cancer.