Breast Cancer - Classification And Analysis Using Different Scanned Images

Sasikala S[1], Vasanthakumar V[2]

[1] Professor, Department of ECE
Kumaraguru College of Technology
Coimbatore- 641049. Tamilnadu

[2] PG Scholar, Department of ECE
Kumaraguru College of Technology
Coimbatore- 641049. Tamilnadu
achievervasanth@gmail.com

Abstract

To find the most effective procedure for the early detection of breast cancer, classification and analysis of different scanned images by AJCC/TNM system. In this paper, we develop a K-mean clustering algorithm to detect cancer mass segmentation on a multi resolution representation of the original scanned images of MRI, Ultrasound and Mammogram. The algorithm has been verified with different scanned images by creating database. From the experimental results we analysis that the sensitivity of MRI, Ultrasound, Mammogram scanned images.

Keywords: K-Mean clustering, breast cancer, computer-aided detection, mammography, MRI, Ultrasound, classification, sensitivity

1.INTRODUCTION

Breast cancer is the second leading cause of death for women worldwide, and more than 8% of all women will suffer this disease during their lifetime. According to cancer statistics 2012 [1] This report estimate earlier detection of cancer, the better the treatment that can be provided. Early detection requires an accurate and reliable diagnosis which should also be able to distinguish between benign and malignant tumors.

In order to reduce morbidity and mortality, early detection of breast cancer is essential. The previous study shows that accurate early detection can effectively reduce the mortality rate caused by breast cancer, and mammography is currently the best technique for reliable detection of early no palpable curable breast cancer [3]. However, the appearances of breast cancers are very subtle and unstable in their early stages. Therefore, doctors and radiologists can miss the abnormality easily if they only diagnose by experience. The computer-aided detection technology can help doctors and radiologists in getting a more reliable and effective diagnosis.

This research study does a comparative analysis on the images obtained from different scanning methods such as Ultrasound, Mammogram, MRI and CT Scan and measure their performance through their accuracy and suggest the best scanning method for breast cancer.

2. RELATED WORK

Amandeep Singh and Amanpreet kaur (2012) presents image processing technique to detect tissue information, biomedical images have the ability to assist physicians in detecting disease caused by cells abnormal growth. Developing algorithms and software to analyse these images may also assist physicians in their daily work. The key and hardest task is auto-extracting of tiny modules or tumour from the biomedical image, which if detected at initial stage gives the information of early cancer. This study combines image threshold, edge detection and segmentation helps in detection of cancer .

Automated technique for mammogram segmentation is explored already. This algorithm proposes morphological pre-processing and selected region growing algorithm in order (1) to remove digitization noises (2)suppress radiopaque artifacts (3)separate background region from the breast profile region and (4) remove the pectoral muscle, for accentuating the breast profile region. The evaluation of performance characteristics for two separate sources using around truth images. The first contribution of this algorithm is a breast contour representation of breast profile region and the other one is the performance evaluation. Finally, for images of differing from multiple databases, results are obtained with high accuracy.

A new enhancement process namely binary homogeneity, An automated scheme for the detection of abnormalities by segmentation of breast ROI. Enhancement algorithm is followed by an approach for edge detection. Initially, breast boundary is obtained using breast boundered detection algorithm and then pectoral muscle detection algorithm (PMDA) is done in order to suppress the muscle, thus obtaining the breast ROI. To differentiate the various regions within the breast anatomical segmentation of breast ROI algorithm is used and finally seeded region growing algorithm is used and finally seeded region growing algorithm is used to isolate normal and abnormal region in the breast tissue. Results obtained show a reliable detection rate of abnormal regions and it has also authenticated the accuracy of detection methods.

Bio medical images are processed and detection of tissue information is being done. The task involves auto-extraction of tumor from the image being processed. This involves image enhancement, edge detection and segmentation in the detection of cancer. Global thresholding is incorporated together with the crop segmentation so that the breast cancer detection is done accurately. Enhancement makes image more immune to noise and does not cause any false impression in the image.

Normally this segmentation leads to the weakness, that it is not suitable for multichannel images. Also making noise sensitive of in-homogeneity leading to the corruption of the image histogram. Edge detectors have the difficulty in adapting to different situation and they behave very poorly. Quality of edge detection is highly dependent on lighting conditions, presence of objects similar intensities and noise.

Another algorithm deals with the diagnosis of liver cirrhosis which is ROI based. The procedure involves the generation of ROC curve of K-means clustering. The goodness of K-mean clustering analysis is compared with commonly used classifier including the LDA(linear discriminant analysis)and ANN (artificial neural networks)ROC depicts the tradeoff between the sensitivity and specificity for LDA, multivariate data observations are transformed in to univariates.The output of both LDA and ANN ,K-mean clustering does not transform the multivariate feature input in to a univariate. A non parametric K-means clustering performance well.

TYPES OF SCREENING

1) Mammogram

Mammography is a special type of x-ray imaging technique which is used to create detailed images of the breast. There are two types of mammography exams, screening and diagnostic.

Screening mammography is an x-ray examination of the breasts in a woman who is asymptomatic (has no complaints or symptoms of breast cancer). The goal of screening mammography is to detect cancer when it is still too small to be felt by a woman or her physician. Early detection of small breast cancers by screening mammography greatly improves a woman's chances for successful treatment. Diagnostic mammography is an x-ray examination of the breast in a woman who either has a breast complaint (for example, a breast lump or nipple discharge is found during self-exam) or has had an abnormality found during screening mammography. Diagnostic mammography is more time consuming than screening mammography and is used to determine exact size and location of breast abnormalities and their surrounding tissue and lymph nodes. Typically, several additional views of the breast are imaged and interpreted during diagnostic mammography. Thus, diagnostic mammography is more expensive than screening mammography.

2) Ultrasound

It is an imaging technique in which deep structures of the body are visualized by recording the reflections (echoes) of ultrasonic waves directed into the tissues. There are three steps to create an ultrasound image. First, producing a sound wave, then, receiving echoes and finally, interpreting those echoes. The ultrasonic waves are produced by electrically stimulating a piezoelectric crystal called a transducer. As the beam strikes an interface or boundary between tissues of varying acoustic impedance (e.g., muscle and blood) some of the sound waves are reflected back to the transducer as echoes. The echoes are then converted into electrical pulses that travel to the ultrasonic scanner, where they are processed and transformed into a digital image displayed on the monitor, presenting a picture of the tissues under examination. However, the interference of those echoes, having different phases, add together to give a resultant wave whose amplitude and phase varies randomly resulting in the speckle, which is common in ultrasound images.

3) CT Scan

A CT scan (also called a CAT scan, or computerized tomography scan) is an x-ray technique that gives doctors information about the body's internal organs in 2-dimensional slices, or cross-sections. During a CT scan, you lie on a moving table and pass through a doughnut-shaped machine that takes x-rays of the body from many different angles. A computer puts these x-rays together to created detailed pictures of the inside of the body. Before the test, you need to have a contrast solution (dye) injected into your arm through an intravenous line. Because the dye can affect the kidneys, your doctor may perform kidney function tests before giving you the contrast solution.

Right now, CT scans are not used routinely to evaluate the breast. If you have a large breast cancer, your doctor may order a CT scan to assess whether or not the cancer has moved into the chest wall. This helps determine whether or not the cancer can be removed with mastectomy.

4) Magnetic resonance imaging

Magnetic Resonance Imaging (MRI), nuclear magnetic resonance imaging (NMRI), or magnetic resonance tomography (MRT) is a medical imaging technique used in radiology to visualize internal structures of the body in detail. MRI makes use of the property of nuclear (NMR) to image nuclei of atoms inside the body.

An MRI scanner is a device in which the patient lies within a large, powerful magnet where the magnetic field is used to align the magnetization of some atomic nuclei in the body, and radio magnetic fields are applied to systematically alter the alignment of this magnetization. This causes the nuclei to produce a rotating magnetic field detectable by the scanner and this information is recorded to construct an image of the scanned area of the body. Magnetic field gradients cause nuclei at different locations to process at different speeds, which allows spatial information to be recovered using Fourier analysis of the measured signal. By using gradients in different directions, 2D images or 3D volumes can be obtained in any arbitrary orientation.

MRI provides good contrast between the different soft tissues of the body, which makes it especially useful in imaging the brain, muscles, the heart, and cancers compared with other medical imaging techniques such as computed tomography (CT) or X-rays. Unlike CT scans or traditional X-rays, MRI does not use ionizing radiation.

III. ALGORITHM

MEDIAN FILTER

Median Filter is used to reduce the noises in the scanned images. Median filtering is similar to an averaging filter, in that each output pixel is set to an average of the pixel values in the neighborhood of the corresponding input pixel. However, with median filtering, the value of an output pixel is determined by the median of the neighborhood pixels, rather than the mean. The median is much less sensitive than the mean to extreme values (called outliers). Median filtering is therefore better able to remove these outliers without reducing the sharpness of the image.

K-MEAN CLUSTERING ALGORITHM

he k-means algorithm assigns feature vectors to clusters by the minimum distance assignment principle, which assigns a new feature vector $\mathbf{x}^{(q)}$ to the cluster $\mathbf{c}^{(k)}$ such that the distance from $\mathbf{x}^{(q)}$ to the center of $\mathbf{c}^{(k)}$ is the minimum over all K clusters. The basic k-means algorithm is as follows:

1. Put the first K feature vectors as initial centers
2. Assign each sample vector to the cluster with minimum distance assignment principle.
3. Compute new average as new center for each cluster
4. If any center has changed, then go to step 2, else terminate.
5. The advantages of the method are its simplicity, efficiency, and self-organization. It is used as initial process in many other algorithms.

Given a set of observations ($\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$), where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k sets ($k \le n$) $\mathbf{S} = \{S_1, S_2, \ldots, S_k\}$ so as to minimize the within-cluster sum of squares (WCSS).

K-Means algorithm is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other. The algorithm assumes that the data features form a vector space and tries to find natural clustering in them. The points are clustered around centroids $\mu_i \forall_i = 1 \ldots k$ which are obtained by minimizing the objective

$$V = \sum_{i=2}^{k} \sum_{xj \in si} (x_i - \mu_i)^2$$

where there are k clusters Si, i = 1,2,….,k and μ_i is the centroid or mean point of all the points $xj \in si$

As a part of this project, an iterative version of the algorithm was implemented. The algorithm takes a 2 dimensional image as input. Various steps in the algorithm are as follows:

1. Compute the intensity distribution(also called the histogram) of the intensities.
2. Initialize the centroids with k random intensities.
3. Repeat the following steps until the cluster labels of the image does not change
anymore.
4. Cluster the points based on distance of their intensities from the centroid intensities.

$$c^{(i)} := argmin \left\| x^{(i)} - \mu_j \right\|^2$$

5. Compute the new centroid for each of the clusters.

$$\mu_i := \frac{\sum_{i=1}^{m} 1\{c_i = j\} x^{(i)}}{\sum_{i=1}^{m} 1\{c_i = j\}}$$

where k is a parameter of the algorithm (the number of clusters to be found), i iterates over the all the intensities, j iterates over all the centroids and μ_i are the centroid intensities. *k*-means clustering in particular when using heuristics such as Lloyd's algorithm is rather easy to implement and apply even on large data sets. As such, it has been successfully used in various topics, ranging from market segmentation, computer vision, geo statistics, and astronomy to agriculture. It often is used as a pre-processing step for other algorithms, for example to find a starting configuration.

k-means clustering, and its associated expectation-maximization algorithm, is a special case of a Gaussian mixture model, specifically, the limit of taking all co variances as diagonal, equal, and small. It is often easy to generalize a k-means problem into a Gaussian mixture model. Mean shift clustering - Basic mean shift clustering algorithms maintain a set of data points the same size as the input data set. Initially, this set is copied from the input set. Then this set is iteratively replaced by the mean of those points in the set that are within a given distance of that point. By contrast, k-means restricts this updated set to k points usually much less than the number of points in the input data set, and replaces each point in this set by the mean of all points in the input set that are closer to that point than any other (e.g. within the Voronoi partition of each updating point). A mean shift algorithm that is similar then to k-means, called likelihood mean shift, replaces the set of points undergoing replacement by the mean of all points in the input set that are within a given distance of the changing set. One of the advantages of mean shift over k-means is that there is no need to choose the number of clusters, because mean shift is likely to find only a few clusters if indeed only a small number exist. However, mean shift can be much slower than k-means, and still requires selection of a bandwidth parameter. Mean shift has soft variants much as k-means does.

Principal component analysis (PCA) It was asserted in that the relaxed solution of *k*-means clustering, specified by the cluster indicators, is given by the PCA (principal component analysis) principal components, and the PCA subspace spanned by the principal directions is identical to the cluster centroid subspace. However, that PCA is a useful relaxation of k-means clustering was not a new result, and it is straightforward to uncover counterexamples to the statement that the cluster centroid subspace is spanned by the principal directions.

Bilateral filtering k-means implicitly assumes that the ordering of the input data set does not matter. The bilateral filter is similar to K-means and mean shift in that it maintains a set of data points that are iteratively replaced by means. However, the bilateral filter restricts the calculation of the (kernel weighted) mean to include only points that are close in the ordering of the input data. This makes it applicable to problems such as image denoising, where the spatial arrangement of pixels in an image is of critical importance

Imaging for breast cancer staging

In patients with a known or highly suspected cancer, staging is performed to determine the extent of spread and overall burden of the disease. It includes evaluation of the affected breast, of regional lymph nodes, and of distant or systemic sites. Results of staging are stratified according to the TNM system. Assessment of the T parameter in patients with breast cancer is commonly based on X-ray Mammography, CT and MRI,US as described above for diagnostic characterization of suspicious tumor Lesions.

Lymph node status (parameter N) is a major prognostic factor in early-stage disease and this information is of paramount importance for tailoring patient-specific treatment. In particular, the presence or absence of metastatic infiltration of the axillary lymph nodes must be considered for further treatment after surgery (adjuvant therapy). In the preoperative phase, the axillary status can be assessed by clinical examination (low sensitivity/specificity), US (sensitive and inexpensive, possibly integrated with fine-needle aspiration cytology to increase specificity), and by PET/CT with FDG (over 95% specificity, but low sensitivity for nodes <1 cm in size). Therefore, most patients are scheduled for surgery without the axillary lymph node status having been ascertained accurately.

The traditional approach to axillary nodal staging has for several decades been represented by systematic axillary lymph node dissection. However, this procedure is frequently burdened with important immediate and long-term morbidities (such as wound infection and prolonged healing, sensory/motor nerve damage and, above all, lymphoedema of the upper limb). These drawbacks are especially important when considering that complete axillary dissection is actually necessary in only about one out of three patients with early breast cancer. In the past few years, the procedure of sentinel.

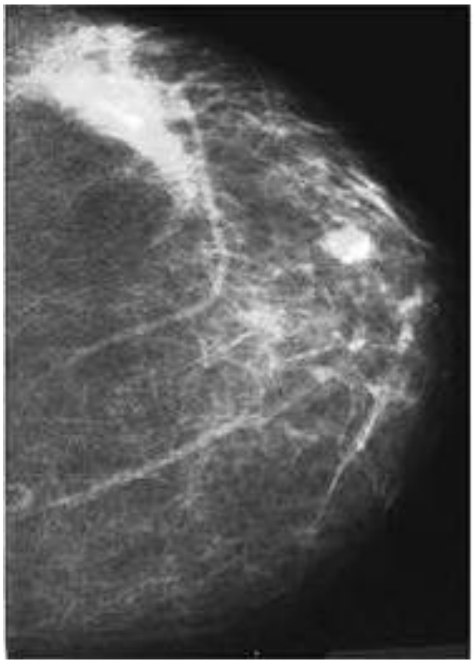
lymph node biopsy for predicting tumor status of the axilla has become the standard of care for breast cancer patients with a clinically negative axilla. This procedure is best performed as radio-guided biopsy of the sentinel lymph node, after lymphoscintigraphic mapping with the use of a radio colloid agent that is injected interstitially at the tumor.

Staging for distant metastases (parameter M) as part of the initial evaluation is recommended for locally advanced breast cancer, especially for patients with advanced axillary nodal disease, because in these conditions the risk of systemic metastases is high. In these cases imaging is designed to survey the chest, abdomen, pelvis and bones; it generally includes standard chest X-ray or computed tomography (CT), abdominal ultrasound or CT.

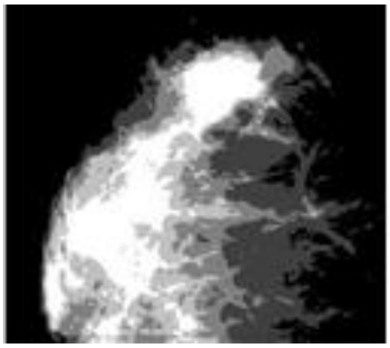
In early-stage breast cancer patients (Stage I or low-end of Stage II), systemic staging is not recommended, unless symptoms are present, since the chance of distant metastases is low and therefore the chance of false positive findings is considerably higher than the chance of true positive findings.

SIMULATATION AND RESULT

The Images of breasts, affected by cancer, taken by various screening methods like mammogram, ultra sound, MRI and CT Scan are filtered with the median filter algorithm for noise reduction. Then the output images of the median filter are simulated using K-mean Clustering algorithm and results are shown.



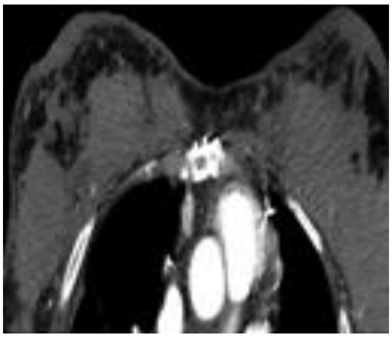
(a)



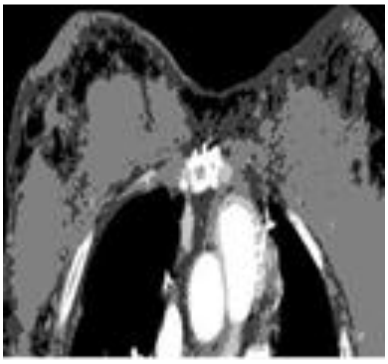
(b)

(c)

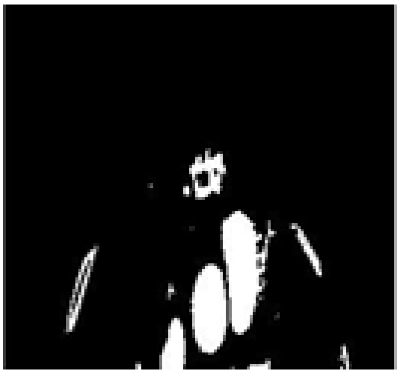
Fig. 1. Simulation results of breast images, affected by cancer: (a) Original Mammogram image (b) Median filtered image (c) Segmented image using k-mean clustering algorithm.



(a)

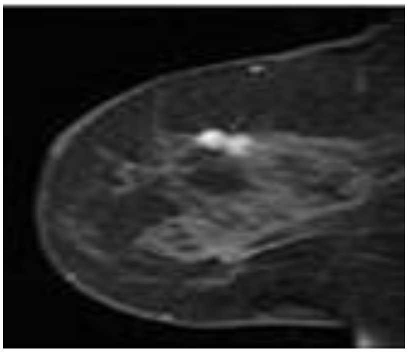


(b)

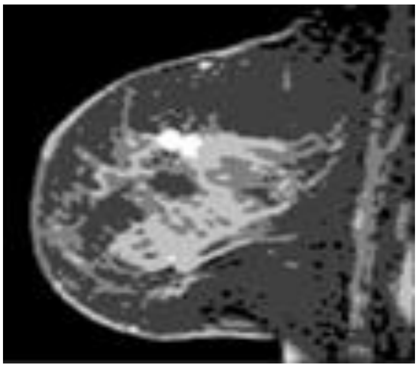


(c)

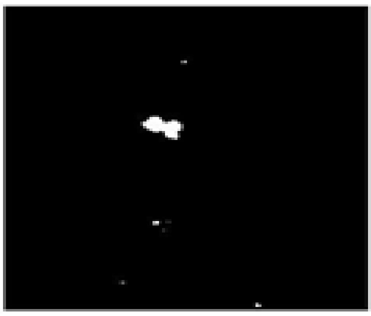
Fig. 2. Simulation results of breast images, affected by cancer: (a) Original CT image, (b) clustered image using k-mean clustering algorithm (c) Segmented image using k-mean clustering algorithm.



(a)

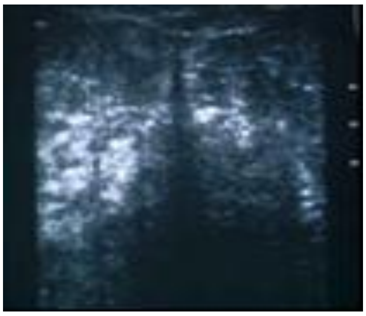


(b)

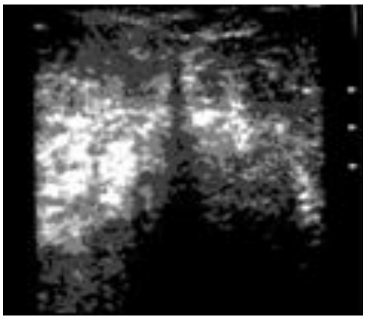


(c)

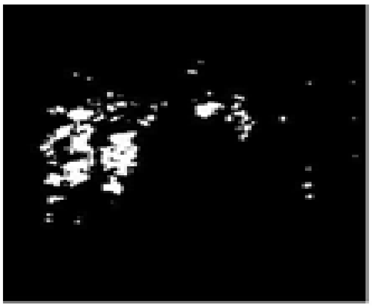
Fig. 3. Simulation results of breast images, affected by cancer: (a) Original MRI image (b) clustered image using k-mean clustering algorithm (c) Segmented image using k-mean clustering algorithm.



(a)



(b)



(c)

Fig. 1. Simulation results of breast images, affected by cancer: (a) Original Ultrasound image (b) clustered image using k-mean clustering algorithm (c) Segmented image using k-mean clustering algorithm.

DISCUSSION:

There are lots of noises in medical images obtained from CT Scan, Ultrasound, Mammogram and MRI which reduces the efficiency and accuracy of identify a breast cancer. To reduce the noises in the medical image we do use median filters. Using K-Mean Clustering algorithm we segmented the image which shows the cancer affected part in all type of scanned images preferably CT Scan, Ultrasound, Mammogram and MRI. Then the area, circumference, diameter, radius are calculated. The above details are calculated in order to find the staging system in breast cancer. And we do follow **AJCC/TNM** System for identifying the stages in breast cancer. Further to go on with research we also classify the types of breast cancer using the segmented images.

The following parameters of scanned images are calculated to identify the accuracy and sensitivity of the various scanned image. This result can suggest the best scanning technology for breast cancer. The various parameters are Mean, SD, Variance , Entropy, RMS, Smoothness, Kurosis, Skewness, Energy, Variance, IDM, Contrast, etc.

V. CONCLUSION AND FUTURE WORK

In this work, an attempt is made to implement K-means clustering algorithm for breast image segmentation for early detection of breast cancer. This attempt is a new idea in image processing technique. This technique is applied on breast cancer image segmentation. Our proposed segmentation algorithm is able to segment lesions(micro-calcifications and masses) efficiently and more accurately. The experimental results show that the developed approach is able to segment breast in CT Scan, Ultrasound, mammography and MRI images. we will extract some features from the regions of segmented breast in order to distinguish the breast cancer from the tissue and identify the stages of the cancer cells. The study is on-verge of completing by calculating the parameters of the all the four segmented scanned image to find out the accuracy and sensitivity. After completing this part, it can suggested which scanning technology is better.

REFERENCES:

1.Amandeep Singh 2amanpreet Kaur Breast Tumour Detection Using Segmentationtechnique From Ct Scan Iracst – International Journal Of Computer Networks And Wireless Communications (Ijcnwc), Issn: 2250-3501vol.2, No.2, April 2012

2. Sami S. Brandt*, Gopal Karemore, Nico Karssemeijer, And Mads Nielsen, An Anatomically Oriented Breast Coordinatesystem For Mammogram Analysis Ieee Transactions On Medical Imaging, Vol. 30, No. 10, October 2011

3.Nalini Singh Ambarish G Mohapatra Gurukalyan Kanungo Breast Cancer Mass Detection In Mammograms Using K-Means And Fuzzy C-Means Clustering International Journal Of Computer Applications (0975 – 8887) Volume 22– No.2, May 2011

4. Jawad Nagi*,1, Sameem Abdul Kareem1, Farrukh Nagi2, Syed Khaleel Ahmed2 "Automated Breast Profile Segmentation For Roidetection Using Digital Mammograms" 2010 Ieee Embs Conference On Biomedical Engineering & Sciences (Iecbes 2010), Kuala Lumpur, Malaysia, 30th November - 2nd December 2010.

5.Bhagwati Charan Patel And Dr. G.R.Sinha An Adaptive K-Means Clustering Algorithm For Breastimage Segmentation International Journal Of Computer Applications (0975 – 8887)Volume 10– N.4, November 2010

6.Jianhua Yao, Jeremy Chen, And Catherine Chow Breast Tumor Analysis In Dynamic Contrastenhanced Mri Using Texture Featuresand Wavelet Transform Ieee Journal Of Selected Topics In Signal Processing, Vol. 3, No. 1, February 2009

7.Mohammad Sameti, Andbranko Palcicimage Feature Extraction In The Last Screeningmammograms Prior To Detection Of Breast Cancer Ieee Journal Of Selected Topics In Signal Processing, Vol. 3, No. 1, February 2009

8.Shakti K. Davis*, Member, Ieee, Barry D. Van Veen, Fellow, Ieee, Susan C. Hagness, Senior Member, Ieee,And Frederick Kelcz Breast Tumor Characterization Based Onultrawideband Microwave Backscatter Ieee Transactions On Biomedical Engineering, Vol. 55, No. 1, January 2008

9.Jacob Levman*, Tony Leung, Petrina Causer, Don Plewes, And Anne L. Martel Classification Of Dynamic Contrast-Enhancedmagnetic Resonance Breast Lesions Bysupport Vector Machines Ieee Transactions On Medical Imaging, Vol. 27, No. 5, May 2008

10.Gobert N. Lee And Hiroshi Fujita K-Means Clustering For Classifying Unlabelled Mri Data 0-7695-3067-2/07 $25.00 © 2007 Ieee Doi 10.1109/Dicta.2007.11