Volume 4, Issue 2, February 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

MammoAPPS: Digital Mammogram Mobile Application for Diagnosing Breast Cancer and Retrieval Reports for Decision Making to Medical Experts

Valliappan Raman*, **Dr.Putra Sumari**
School of Computer Science
University Sains Malaysia

Abstract— ***The objective of the paper is to make a background study and develop a framework on computer aided diagnosis in mammogram dataset to deploy it in smartphone and make it convenient for the users. The possibilities of radiology departments to modify the radiological service process in a regional, international and mobile setting were plan to investigate in the research to build mobile application and deploy it in the android and IOS platform. The main idea of the research is to use smartphone and communicate for MRI scans in a feasibility study of sample patient cases in the domain of breast cancer. All transmitted image series were suitable for giving a preliminary consultation to the clinic, and final report is retrieved for expert decision. The working concept of CAD mammogram prototype was implemented in matlab and validated, research method applied was on best case classifier is to classify the mass tumour in MRI mammogram image, and implemented work of acquisition, pre-processing, segmentation feature extraction and classification will be deployed in mobile platform. Therefore paper highlights the research background, objectives, research methods and framework for the mobile deployment.***

Keywords— ***Mammogram, Classification, Radiology, Mobile Application.***

I. INTRODUCTION

Breast cancer is the most common cancer among women in the western world and in Malaysia, other than skin cancer. It is the second leading cause of cancer death in women, after lung cancer. In Malaysian context for example, figure 1 shows that breast cancer is the leading cancer compare to other cancer type and majorilty happen to women. Survival from breast cancer is directly related to the stage at diagnosis in which the earlier the detection, the higher chances of successful treatment [1, 2]. Screening mammography (x-ray of the breast), is currently the most effective tool for early detection of breast cancer. Radiologists who are responsible on the screening activities visually search mammograms for specific abnormalities. Some of the important signs of abnormalities in breast cancer that radiologists look for are masses and microclacification (tumours) within the mammogram. Figure 2 show example of the mammogram that shows tumors at right cancerous breast. These mammograms (as example in fig. 2) are referred to as mammogram abnormalities in which there are signs of masses (tumour) for possibility to be cancerous. The advances in telecommunication and computer technology have given us widely standardized tools, which can be applied into clinical practice and mobile phones and smart phones have changed the world to access information from anywhere around the globe.

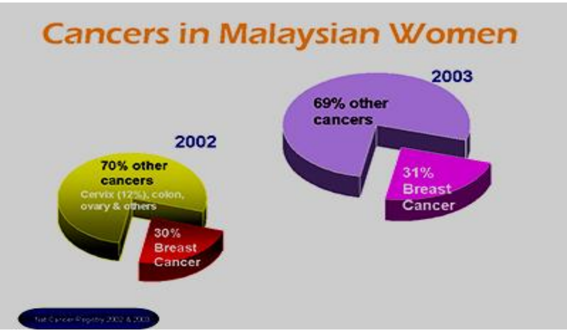

Fig 1(A)

Fig 1(B)

Fig 1(a) illustrates the Percentage of woman affected by breast cancer in Malaysia Fig 1(B) illustrates the Mammogram of right and left breast with cancerous area

Today there are many commercial computer assisted mammogram screening system in the literature to help radiologists. These systems have gone from crude tool in research laboratory toward commercial system such as R2 Technology Inc., Hewlett Packard Co., Sterling Diagnostic Imaging, Siemens, GE and Med Detect/Lockheed Martin.

Detection of suspicious abnormalities (masses & microclacification) within mammogram is a repetitive and fatiguing task. For every thousand cases analysed by a radiologist, only 3 to 4 are cancerous and the rest could be overlooked, very hard to make decision on accuracy of tumour, expert and physicians feedback is important for further process of treatment . Based on the Cancer Survey reports in Malaysia, we identified that nearly 50% of the women responded were not examining their breasts regularly for bumps or other signs of breast cancer. Mainly in rural areas, there has not been any awareness created on breast cancer diagnosis.

To encourage regular checking and create awareness, we have proposed a MammoApps, mobile application software that will run on smartphones (Apple IPhone & Android Based Mobile Phones) where the functionalities will be performed as CAD (Computer Aided Diagnosis) systems. Mobile implemented software component will consists of 4 modules which include image acquisition and display module, image transfer module, tumour identification module and feedback report module. The main aim of MammoApps is to collect DICOM image from scan workstation to smartphone and transfer it to doctors for their feedbacks and decisions, then transfer the feedback report to radiologist, whereas radiologist send the report through mobile phone to patients, therefore patients can retrieve the report and request their query related on breast cancer and risk factors based on the report to radiologist. The proposed apps will walk through the breast cancer diagnosis report and assist the doctor from anywhere around the globe to take accurate decision on their treatment and provide consultation to the patients located in urban and rural areas, it will act as an interface between radiologist and doctors for accurate diagnosis and reduce time complexity in making decisions.

The main aim of the research is to deploy the developed prototype of CAD (Computer Aided Diagnosis System) on breast cancer in mobile platform for easier use by radiologist and research experts. Earlier research on CAD system were investigated and developed in matlab by best case classifier method. Currently this paper highlights the proposed framework of implementing the mobile application; background study and research justification were briefly explained in section 2. The research objectives were highlighted in section 3 for mobile deployment. Previous prototype development which includes existing works, research methods and output were briefly explained in section 4. In section 5, based on the earlier prototype development, brief description on proposed research framework for deployment is explained, section 6 describes the method to develop the prototype in mobile environment, expected output of the research was explained in section 7. Finally conclusion was made in section 8.

II. **BACKGROUND**

In Malaysia breast cancer is the leading cause of cancer-related death among women aged 15-54. Survival from breast cancer is directly related to the stage at diagnosis in which the earlier the detection, the higher chances of successful treatment. Given a mammography, radiologist examines carefully and looking for features that may indicates the potential cancerous problem. The two most common features that are looking for are masses and micro-calcifications. Mass is a cyst, which is non-cancerous collection of fluid, may appear as a mass in the film whereas micro-clarification is a tiny clustered particles and probably the smallest structures that can be identified in the mammogram. The intensities difference between these features of masses are not clear when compare to normal tissues. Both look very much similar and detecting them is the most challenge in mammogram screening. Benign lesions are well circumscribed, compact, and roughly circular or elliptical. But malignant lesions usually have a blurred boundary, an irregular appearance, and sometimes are surrounded by a radiating pattern of linear spicules. Radiologists need to be experience enough on subjective criteria to be successfully detecting these masses on a mammogram and make right decisions with consultation of doctors. Several Computer Aided Diagnosis (CAD) systems for automatic classification of breast lesions as either benign or malignant have been developed. Most of the computer aided diagnosis systems proved to be powerful tools that could assist radiologists in diagnosing a patient. Currently medical community is considered to be both an adopter and innovator in the use of the PDA (Personal Digital Assistant) and the smartphone devices in the workplace. The practice of medicine is based on accumulating clinical patient data, applying data analysis and medical references to diagnose, and defining appropriate therapy. One of the interesting points in the development of mobile application is available to everyone, and fulfils the expectation of radiologist and doctors.

Socially, using recent data it is estimated that due to not having awareness in screening, breast cancer mortality has been increasing in Malaysia. Mobile phone or smartphone have created new revolution to the modern technology where the information's can be accessed anytime around the globe. Current proposed new diagnosis methods based on image processing applications in mobile will lead to higher sensitivity and earlier detection of lesions which is essential for increasing the effectiveness of mammography screening and consequently a further reduction of breast cancer mortality. Scientifically, improving the accurate classification of diseased versus non diseased screenings is one of the on-going challenges in the world of medical informatics. We earlier proposed to develop the system by using region growing algorithm and CBR classifiers in detecting malignancies, will support the interpretation process of mammograms by radiologists. CAD system implementation deployed in smartphones will be a greater assistance of radiologist and experts.

III. **OBJECTIVE**

The Objective of this research was to evaluate the use and implementation of Mobile application for diagnosing breast cancer and creating awareness among the public with help of everyday use mobile phones which are part of the human life. The aim of the research focus on few challenges to overcome: 1) to research whether smart phone system such as IOS and Android OS could combine acceptable image quality with reasonable costs and serve primary health centres and local hospitals for image consultation. 2) To research whether new ubiquitous communication tools, such as mobile data communication, are suitable for radiological purposes.3) to identify whether smartphones could be used for image interpretations and how the reports can be linked to hospital database and can be retrieved through smart phones. 4) To implement the Mobile Apps with tumour Identification Module where it supports the region growing image processing algorithm to identify the tumour and provide feature extraction reports.

IV. PREVIOUS PROTOTYPE DEVELOPMENT AND RESULTS

A. Literature survey

Machine learning techniques are applied to diagnose breast cancer and it becomes a very active research area. Several Computer Aided Diagnosis (CAD) systems for automatic classification of breast lesions as either benign or malignant have been developed. Some of them are based on Bayesian networks learned on mammographic descriptions provided by radiologists or on features extracted by image processing. Other classifying techniques that are used for the diagnosis of breast lesions are Support Vector Machines, Artificial Neural Networks, Linear Classifiers and Association Rule based classifiers. Most of the computer aided diagnosis systems proved to be powerful tools that could assist radiologists in diagnosing a patient. There are few interesting research work are on mass detection and classification are shown below.

Zhang et al[3] noted that the presence of speculated lesions led to changes in the local mammographic texture. They proposed that such a change could be detected in the Hough domain, which is computed using the Hough transform. They partitioned an image into overlapping ROIs and computed the Hough transform for each ROI. The Hough domain of each ROI was threshold to detect local changes in mammographic texture and to determine the presence or absence of a speculated mass.Brzakovic et al[4] use a two stage multi-resolution approach for detection of masses. First they identified suspicious ROIs using Gaussian pyramids and a pyramid linking technique, based on the intensity of edge links. Edges were linked across various levels of resolution. This was followed by a classification stage, where the ROI were classified as malignant, benign or normal based on features like shape descriptors, edge descriptors and area. Petrick et al [5] developed a two-stage algorithm for the enhancement of suspicious objects. In the first stage they proposed an adaptive density weighted contrast enhancement filter (DWCE) to enhance objects and suppress background structures. The central idea of this filtering technique was that it used the density value of each pixel to weight its local contrast. In the first stage the DWCE filter and a simple edge detector (Laplacian of Gaussian) was used to extract ROIs containing potential masses. In the second stage the DWCE was re-applied to the ROI. Finally, to reduce the number of false positives, they used a set of texture features for classifying detected objects as masses or normal. They further improved the detection algorithm by adding an object-based region-growing algorithm to it. Lai made an approach based on a multi-resolution Markov random field model detect mass lesions. Its initial window size for segmentation influences the sensitivity of detection.Li [6] proposed a method on iris filter was developed to detect mass lesions of rounded convex regions with low contrast. The iris filter enhances most round malignant masses. However, some malignant masses are shaped irregularly. Hutt [8], work is based on fuzzy pyramid linking algorithm which was used to detect micro-calcifications in mammograms. The links between various levels were determined by a fuzzy membership function. The pyramid structure is formed by producing images of decreasing resolution with the highest resolution image at the bottom of the pyramid.Tarassenko et.al [9] proposed an image segmentation technique based on region clustering. The mammogram is partitioned into clusters on the basis of data density. In each region the probability density is calculated using Parzen estimator, and the result of the image segmentation procedure is an image containing all possible regions of interest. The regions of interest are then presented to the human expert for further analysis.

Comer et al.[10] utilized an EM technique to segment digitized mammograms into homogeneous texture regions by assigning each pixel was to one of a set of classes such that the number incorrectly classified pixels was minimized. Kupinski and Giger [11] developed a method, which combines region growing with probability analysis to determine final segmentation.Ramon Lopez De Mantaras [12], review a representative selection of CBR research in the past few decades on aspects of retrieval, reuse, revision, and retention. CBR cycles are explained in detailed in his work.Bottigli et al. [13] presented a comparison of some classification system for massive lesion classification. An algorithm based on morphological lesion differences was used to extract the features. The two classes (pathological or healthy ROIs) were differentiated by utilizing the features. A supervised neural network was employed to check the discriminating performances of the algorithm against other classifiers and the ROC curve was used to present the results. In comparison with the other recent studies [15]; the results of the new representation applied are comparable or better, owing to its better ability to distinguish pathological ROIs from the healthy ones.Vibha L [14],proposes a method for detection of tumor using Watershed Algorithm, and further classifies it as benign or malignant using Watershed Decision Classifier (WDC).Experimental results show that this method performs well with the classification accuracy reaching nearly 88.38%.

Serhat Ozekes.et.al[16] proposed to develop a new method for automated mass detection in digital mammographic images using templates .Masses were detected using a two steps process. First, the pixels in the mammogram images were scanned in 8 directions, and regions of interest (ROI) were identified using various thresholds. Then, a mass template was used to categorize the ROI as true masses or non-masses based on their morphologies. Each pixel of a ROI was scanned with a mass template to determine whether there was a shape (part of a ROI) similar to the mass in the template. The similarity was controlled using two thresholds. If a shape was detected, then the coordinates of the shape were recorded as part of a true mass. To test the system's efficiency, we applied this process to 52 mammogram images from the Mammographic Image Analysis Society (MIAS) database. The results of this experiment showed that using the templates with these diameters achieved sensitivities of 93%, 90% and 81% with 1.3, 0.7 and 0.33 false positives per image respectively.

Giulia Rabottino1.et.al[17] worked on an effective algorithm for massive lesions segmentation based on region–growing technique and classification based on fuzzy logic. In this work, a fast and optimized region growing algorithm for the segmentation step aimed at finding the contour of the mass. This procedure is fundamental for the classification of massive lesions and can strongly influence its performance. Consequently, a specific and effective mass contour extraction algorithm is needed in order to really aid radiologists in the detection and diagnosis of cancer. Nevertheless, the algorithm should be fast since the Computer Aided Detection system (CAD) should be able to identify tumoral lesions as a second human reader.

Wei Qian [18], developed a new adaptive module to improve their computer-assisted diagnostic (CAD) method for mass segmentation and classification. The goal was an adaptive module that used a novel four-channel wavelet transform with neural network rather than a two-channel wavelet transform with manual sub image selection. The four-channel wavelet transform is used for image decomposition and reconstruction, and a novel Kalman-filtering neural network is used for adaptive subimage selection. The results of this study confirm the importance of using a new class of adaptive CAD methods that allow a more generalized application for larger image databases or images generated from different sensors or by means of direct x-ray detection, as required for clinical trials.13. Rabi Narayan Panda.et.al [20], proposed technique is based on a three-step procedure: regions of interest (ROI) specification, two dimensional wavelet transformation, and feature extraction based on OTSU thresholding the region of interest for the identification of microcalcifications and mass lesions. ROIs are preprocessed using a wavelet-based transformation method and a thresholding technique is applied to exclude microcalcifications and mass lesions. The above methods show less than five false positives per image with a true positive detection rate of approximately 70%. It is difficult to compare the performance of these methods because their databases are different.

B. RESEARCH METHOD DESCRIPTION

The system methodology has 4 distinct steps: Initialization module, segmentation module, feature extraction module and classification module as shown in figure 2 below:

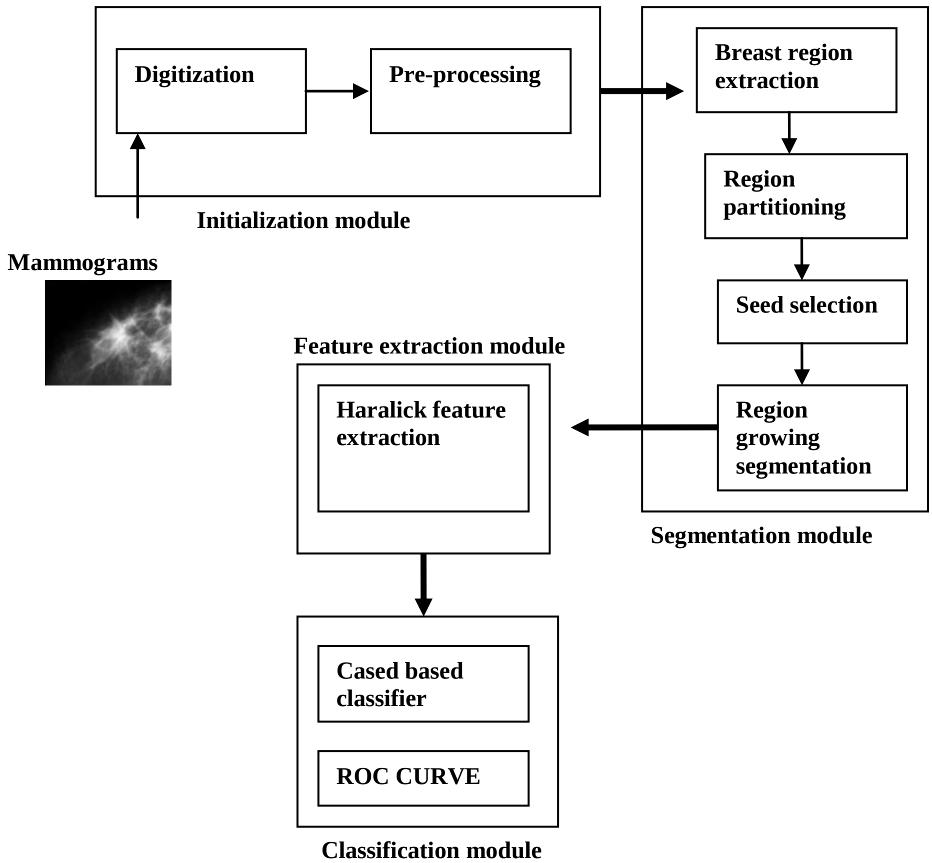

Fig 2 illustrates the prototype development method

1. Initialization module: The initialization module consists of two main process, the digitization and pre-processing. These two processes basically convert the X-ray form into a suitable digital form for the segmentation module.

2. Digitization: X-ray mammograms are digitized with an image resolution of 100 × 100 μm2 and 12 bits per pixel by a laser film digitizer. To detect mass/micro-calcifications on the mammogram, the X-ray film is digitized with a high resolution. Because small masses are usually larger than 3mm in diameter, the digitized

mammograms are decimated with a resolution of 400 × 400 mm 2 by averaging 4 × 4 pixels into one pixel in order to save the computation time.

3. Pre-processing module: Preprocessing is an important issue in low-level image processing. The underlying principle of pre-processing is to enlarge the intensity difference between objects and background and to produce reliable representations of breast tissue structures. An effective method for mammogram enhancement must aim to enhance the texture and features of tumors. The reasons are: (1) low-contrast of mammographic images; (2) hard to read masses in mammogram because it is highly connected to surrounding tissues; the enhancement methods are grouped as global histogram modification approach and local processing approach as shown in table below. Current work is carried out in global histogram modification approach.

4. Table 1 illustrates the pre-processing approach

Preprocessing Approach	Description	Advantage
Global Histogram Modification Approach	Re-assign the intensity values of pixels to make the new distribution of the intensities uniform to the utmost extent	Effective in enhancing the entire image with low contrast
Local Approach	Feature-based or using nonlinear mapping locally	Effective in local texture enhancement

5. Segmentation Module: Segmentation module is extremely important module as it is segmented the digital image into distinted layer, each contains homogenous elements/features. This module consists of four main processes; breast region extraction, region portioning, seed selection and region growing segmentation.

   a) Breast region extraction: After the digitization and pre-processing of the X-ray mammograms, breast region Br is extracted from the digitized image. In general, the non-breast region of the digitized mammograms has a very low intensity and a maximum peak in the histogram. The threshold level Bt for the breast region is determined as follows in equation 1. Where $I_{\max}$ is the pixel intensity at the maximum peak count in the histogram of the decimated image and σ_{bg} is a standard deviation of all pixel values less than $I_{\max}$ under the assumption that the histogram of the background has Gaussian distribution centered at $I_{\max}$.

$$B_t = I_{\max} + 2.5\sigma_{bg} \quad (1)$$

   b) Region Partitioning : The purpose of region partitioning is to apply the different threshold values to each partitioned region. Therefore the region partitioning of the extracted breast region is identified. We adopted Otsu method of region partitioning for extracted breast region. Assume that *x(i, j)* is the pixel of an image *I* at position *(i, j)* and that the gray level of the image ranges from *0 to L - 1*. With a given threshold t, the image can be partitioned into two regions, R_0 and R_1. That is, if *x (i, j) <= t, then (i, j) =* R_0, otherwise, *(i, j) ϵ* R_1. Probability functions of regions R_0 *and* R_1.With respect to a given threshold t is given, respectively, as follows in equation (2) and (3). Where *p(x)* denotes the probability of gray level *x* in the image I. Gray level means and variances of region R_0 and R_1 is calculated as follows in equation (4)(5)(6)(7). The in-class variance $\sigma_w^2(t)$ and the between class variance $\sigma_b^2(t)$ of these regions are described as in equation (8) and (9)

$$W_0(t) = \sum_{x=o}^{t} p(x) \quad (2)$$

$$W_1(t) = \sum_{x=t+1}^{L-1} p(x) \quad (3)$$

$$\mu_0(t) = \sum_{x=0}^{t} xp(x)/W_0(t) \quad (4)$$

$$\mu_1(t) = \sum_{x=t+1}^{L-1} xp(x)/W_1(t) \qquad (5)$$

$$\sigma_0^2(t) = \sum_{x=0}^{t} \left(x - \mu_0(t)\right)^2 p(x)/w_0(t) \qquad (6)$$

$$\sigma_1^2(t) = \sum_{x=t+1}^{L-1} \left(x - \mu_1(t)\right)^2 p(x)/w_1(t) \qquad (7)$$

$$\sigma_w^2(t) = W_0(t)\sigma_0^2(t) + W_1(t)\sigma_1^2(t) \qquad (8)$$

$$\sigma_b^2(t) = W_0(t)W_1(t)(\mu_0(t) - \mu_1(t))^2 \qquad (9)$$

C) Seed Selection: Selection of seed points is important and will influence the quality of the segmentation. Selecting one point allows only for the growing of one region, multiple points allows multiple regions to be grown. With our focus on breast-background segmentation, it is typical to choose to grow either the background or the breast itself. Therefore we want an initial seed point located in a broadly typical part of the background, or in a broadly typical part of the breast. Selection may be automatic, or may involve inviting the user to select a point. Automatic methods may involve both positional and intensity-based factors, such as selecting a pixel with a characteristic intensity in a region of the image which we know is likely to contain either breast or background. When the breast region is partitioned into three regions, three sets of seed pixels are selected from the partitioned regions, respectively. Seed selection considers not only local maximum intensity value but also local contrast between a seed pixel and its neighbors. Let us denote an ascending path *P* of length *l* between two normalized pixels U_s and U_e in a partitioned region, which is a sequence of eight-neighbor connected pixels. Each pixel value U_{pk} is the largest pixel within a 3 × 3 window centered at U_{pk-1} *for 1< k <= l*. Every pixel in the partitioned regions becomes a start pixel *Us* of a path and the end pixel *Ue* largest pixel in its eight neighbors. In order to compute a local contrast at the end pixel *Ue*, we define an intensity difference at Ue that is the difference between the start pixels us and the end pixel Ue as in equation (10). Seed pixels for region growing are selected from the end pixels *Ue* by considering their normalized intensity and seed contrast value. Therefore seed pixels are selected for each partitioned regions and threshold values are also assigned.

$$D(U_sU_e) = U_e - Us \qquad (10)$$

D) Region Growing Segmentation: We chose region growing process for segmentation of a mammographic mass, where a boundary pixel is joined to the current region provided it has the highest gray level among the neighbors of the region. The region growing process starts from seed pixels. The gray level mapping shows local valleys at the boundary of two neighboring regions. The local peak just after the local valley in the gray level mapping gives a sign of the switch between the absorption of pixels in the boundary of the current region and the absorption of pixels in the neighboring region. When the grown region size is equal to or greater than a minimum region size with the stopping condition such as speckle noise, touching previous region, new adjacent region, contrast limitation. Once the stopping condition is achieved, region growing is applied and it segmented.

6. Feature Extraction Module**: After** segmenting the masses in mammogram, The ROI hunter provides the "regions of interest" without giving further information. To this purpose suitable features should be selected so that a decision making system can correctly classify possible pathological regions from healthy ones. Feature extraction plays a fundamental role in many pattern recognition tasks. In this project 18 Haralick texture features (global and local features) are extracted from the segmented masses, currently nine are shown below. The criteria for the feature selection are based on morphological differences between lesions and healthy regions. In particular, the excessive lengthening is often symptom of absence of pathology while the loss of chaotic dynamics in the lesion structures can mark a tumor.

a) Haralick feature texture extraction: After breast region extraction, check whether mass is slightly brighter than its surrounding areas, produces a sharp peak of unusual gray level intensity pixels, peak analysis of histogram is applied for the extraction of significant peak regions as Region of Interest (ROI). Divide the identified ROI into R×R block. Check whether the block is too small or large, the difference of the mass textures from normal textures cannot be well characterized. If it is too large, the result may be too coarse, so calculate the Haralick texture features (shown table below) from Spatial Gray Level Dependence Matrix (SGLD) of each block. From that select the significant features that can easily discriminate mass and non mass region in the image. Next select the block that contains mass based on the features we used the following haralick texture features from the segmentation:

Table 2 illustrates the Local and Global Features

Feature of Selection	Description
Skewness	$\frac{1}{N}\frac{\sum_{i,j=0}^{N-1}\left[g(i,j)-\overline{g(i,j)}\right]^3}{\sqrt{\sum_{i,j=0}^{N-1}\left[g(i,j)-\overline{g(i,j)}\right]^3}}$
Kurtosis	$\frac{1}{N}\frac{\sum_{i,j=0}^{N-1}\left[g(i,j)-\overline{g(i,j)}\right]^4}{\sqrt{\sum_{i,j=0}^{N-1}\left[g(i,j)-\overline{g(i,j)}\right]^4}}$
Circularity	$\frac{A_1}{A}$
Compactness	$\frac{P^2}{A}$
Contrast	$\frac{P_1-P_2}{P_1}$
Standard deviation	σ^2
Intensity	$\overline{g(i,j)}=1/N\sum_{i,j=0}^{N-1}g(i,j)$
Area	*tumor area*
Length	*True Length of Mass*
Breadth	*True Breadth of Mass*
Convex Perimeter	*Perimeter of the convex hull of the mass*
Roughness	*Perimeter/Convex Perimeter*

7. Classification Module**:** Classifiers play an important role in the implementation of computer-aided diagnosis of mammography. The features or a subset of these features are employed by classifiers to classify mass into benign and malignant. Case Based Reasoning (CBR) seems as a potential method in making computerized screening decision. Case-Based Reasoning (CBR) integrates in one system two different characteristics: machine learning capabilities and problem solving capabilities. CBR uses a similar philosophy to that which humans sometimes use: it tries to solve new cases (examples) of a problem by using old previously solved cases. The process of solving new cases contributes with new information and new knowledge to the system. This new information can be used for solving other future cases. The basic method can be easily described in terms of its four phases .The first phase retrieves old solved cases similar to the new one. In the second phase, the system tries to reuse the solutions of the previously retrieved cases for solving the new case. The third phase revises the proposed solution. Finally, the fourth phase retains the useful information obtained when solving the new case. In a Case-Based Classifier System, it is possible to simplify the reuse phase. Classifying the new case with the same class as the most similar retrieved case can do reuse. Performance based Case Base Classifier is a classifier where the re-use phase can be simplified. The kernel in a Case-Based Reasoning system is the retrieval phase (phase 1). Phase 1 retrieves the most similar case or cases to the new case. Obviously, the meaning of most similar will be a key concept in the whole system. Similarity between two cases is computed using different similarity functions. For our purpose in this paper, we use the similarity functions based on the distance concept. The most used similarity function is the Nearest Neighbor algorithm, which computes the similarity between two cases using a global similarity measure. The future practical implementation (used in our system) of this function is based on the Minkowski's metric. Minkowski's metric is defined as:

$$Similarity\left(Case\_x, Case\_y\right)=\sqrt[r]{\sum_{i-1}^{F} W_i \times \left|x_i - y_i\right|^r} \quad \textbf{(11)}$$

Where *Case_x*, *Case_y* are two cases, whose similarity is computed; F is the number of features that describes the case; $x_i y_i$ represent the value of the ith feature of case *Case_x* and *Case_y* respectively; and w_i is the weight of the ith feature. In this study we test the Minkowsky's metric for three different values of r: Hamming distance (r = 1), Euclidean distance (r = 2), and Cubic distance (r = 3). This similarity function needs to compute the feature relevance (w_i) for each problem to be solved. Assuming an accurate weight setting, a case-based reasoning system can increase their prediction accuracy rate. We use also the Clark's and the Cosine distance, both are based on distance concept and also use weighting features. Sometimes human experts can not adjust the feature relevance, automatic method can solve this limitation.

8. Experimental Setup

a) Dataset: A total of 100 mammograms were considered for this study and experimental testing. Mammogram data's are acquired from the MIAS database of mammograms containing cancerous masses and also from private radiologist centers, university medical centre and hospitals. In this program two mammographic views were obtained in the initial screening: medio-lateral oblique (MLO) and cranio-caudal (CC). At subsequent screenings only a MLO was obtained, unless there was an indication that CC views could be beneficial. For every digitized mammogram, a certain number of suspected regions have been indicated. For every region, specific features such as size, shape and speculation have been calculated and a classifier that indicates whether the region is a tumor or not. The number of features differs per case: some cases have prior mammograms while others do not; some cases have lesions visible in two views and others in only one. Therefore multiple classifiers have to be constructed. We will use this kind of datasets for the construction of learning and evaluating case base classifiers.

b) Testing and Evaluation: The mammograms are scanned from X-rays with a maximum resolution of 512x512 pixels. Mammographic image is reduced to an m × n matrix. This matrix contains as many rows, m, as the number of Mass\Mc (Micro-calcification) present in the image, and as many columns (n=4) as the number of features that describe one mass. Next, this m × 4 matrix is transformed into a vector. This transformation computes the average value for each column (feature) across all the rows (mass in the image). Finally, the computed vector is labelled using the class (benign or malign) obtained from the diagnosis done by surgical biopsy. We performed two kinds of experiments in order to compare the performance of the different algorithms. First, we maintained the proportion of original images - now, a set of features for each image- as training and test sets proposed by human experts. Thus, we compared the results obtained by other classifiers with those achieved by human experts, and the statistical modeling terms of classification accuracy. We also evaluate it by comparison of true positive (malignant cases) rate of classified examples *(sensitivity)* and the true negative rate of classified examples *(specificity)* and evaluate it by *ROC & FROC curve*.

c) Results: From the developed prototype in matlab, the results of mammogram segmentation, feature extraction and best case classification are shown in the below figures.

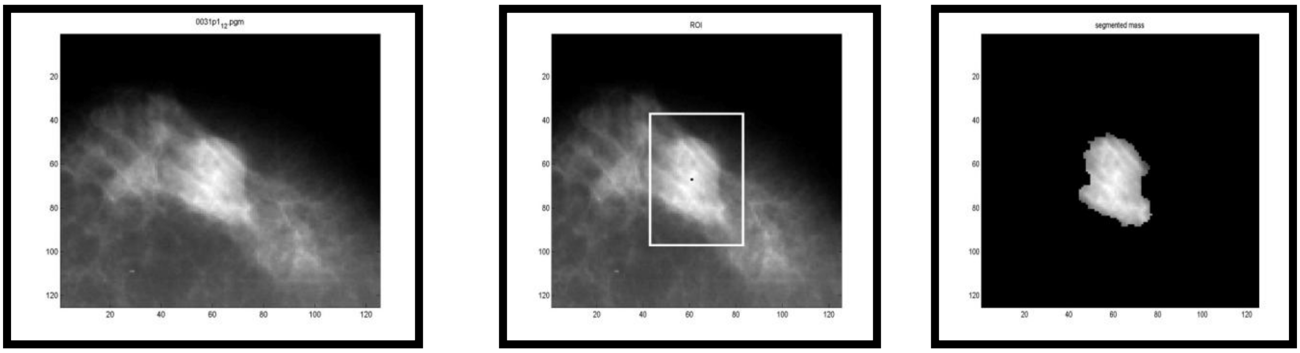

Fig 5 illustrates the sample segmentation of mass benign

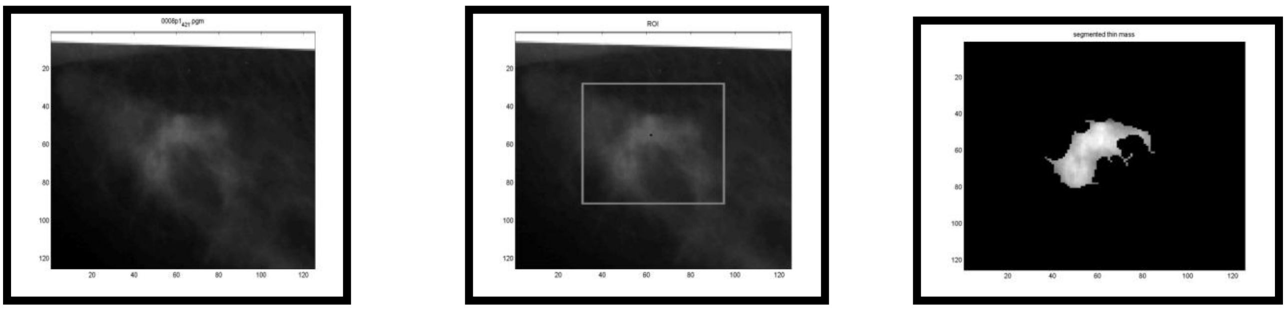

Fig 6 illustrates the sample segmentation of mass malignant

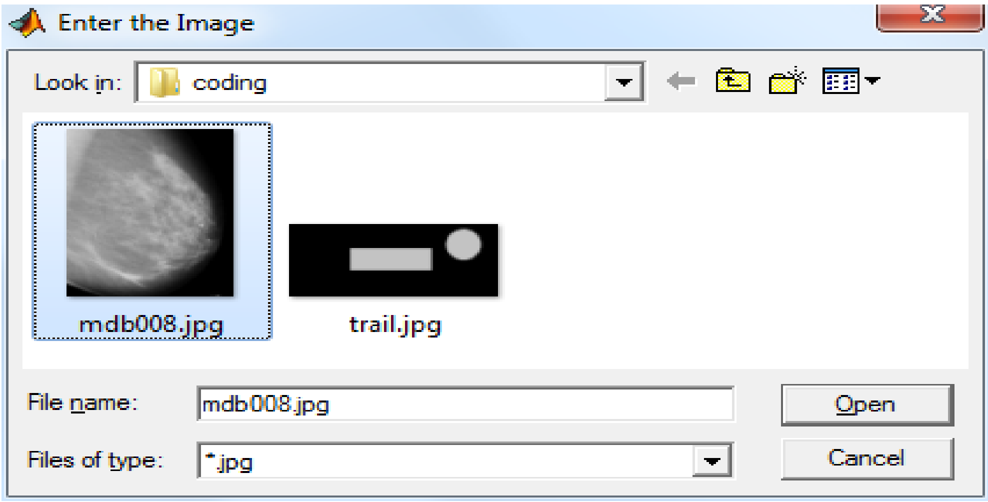

Figure 7 illustrates the input image for segmentation

Figure 8 illustrates the input image for preprocessing

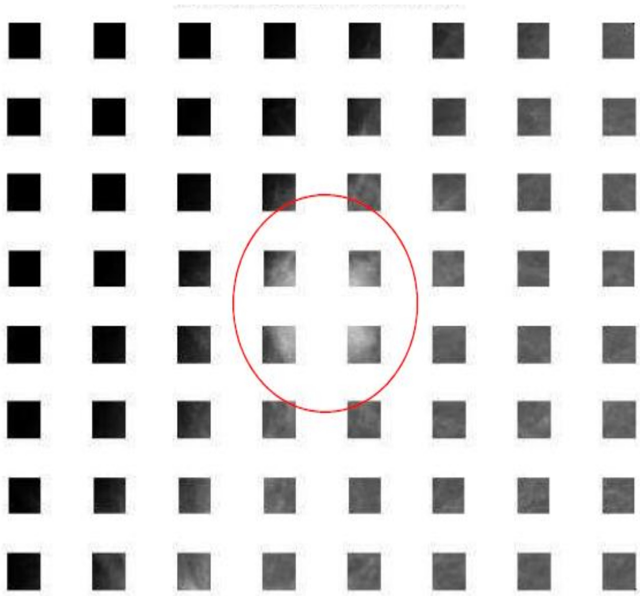

Figure 9 illustrates the enhanced image with non-overlapped regions

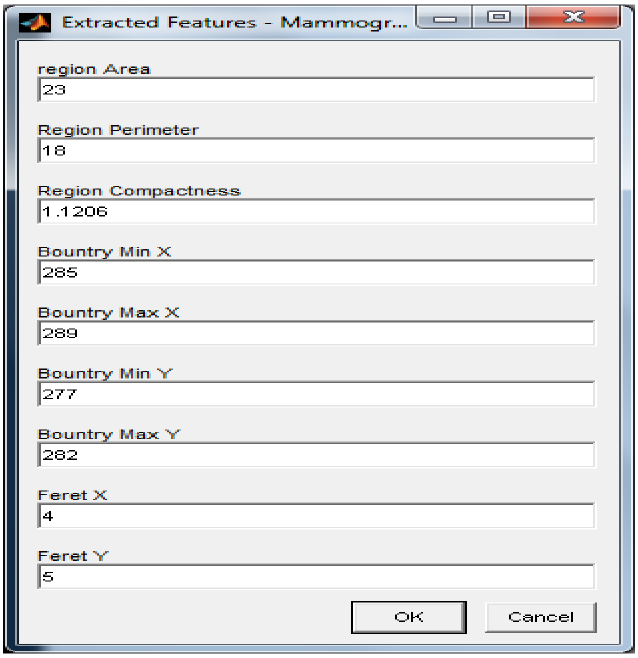

Figure 9 illustrates the feature extraction

Variant	Not Classified	SENSITIVITY	SPECIFICITY	ACCURACY
HAMMING	0.00	83.9	80.2	81.5%
CUBIC	0.00	81.5	82.8	83%
EUCLIDEAN	0.00	82.7	85.2	84%
CLARK	0.00	86.5	90.7	88%
AVERAGE		82%	84%	85%

Figure 10 illustrates the sensitivity and specificity

Correct Classification		Misclassification		Radiologist Misclassification	
Benign	Malignant	Benign	Malignant	Benign	Malignant
84% (20/26)	94% (24/26)	15% (3/30)	10% (2/20)	60-90%(from source)	No Data Source

Figure 11 illustrates the best case classification and misclassification results

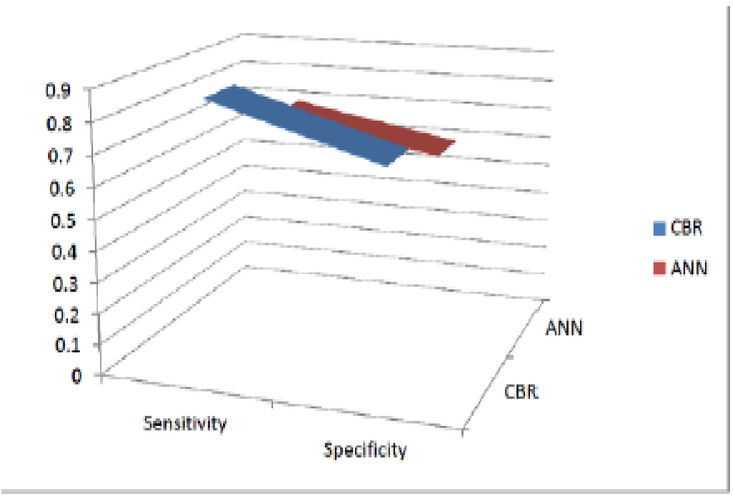

Figure 12 illustrates the ROC curve of classification compare to ANN and Best Case Classifier(CBR)

V. RESEARCH FRAMEWORK FOR MOBILE DEPLOYMENT

The proposed research framework is to improve the access to screening mammography by the use of mobile facilities that can reach rural areas. Digital mammography has been shown to be equivalent in cancer detection to analog mammography for older women. With digital mammography, images can be transmitted, archived, and interpreted off-site, which allows for increased access to high-quality imaging. The implementation of mobile application to to display, store and transfer images, it is challenged by the very demanding spatial resolution requirements of mammography as compared with other forms of medical imaging, because of the need to detect mass ,micro-calcifications and the fine details of lesion margination. The larger size of the image data sets places significant demands on the local and wide area networks to be able to transmit the studies. Our proposed research work has mobile application which will be an interface between physician and radiologist. First, the patients were imaged with a CT scanner (GE HiSpeed Advantage or GE Sytec 3000, GE, Milwaukee, USA). The images were routinely transferred to a visualization workstation for radiologist interpretation. After manipulation, the images were captured from the screen using public domain xv software with a special script. All the captured images from the visualization workstation were stored in a network directory that was physically mounted on To reduce the bandwidth required for transmission, the DICOM images were compressed stored in the dial-up server. The images of patients were transferred to the portable terminal by encrypted GSM network and it is transferred to smartphone which includes combined digital phone and messaging device with short message service (SMS), telefax and even Internet/intranet capability. The transferred image will be reached to doctor's smart phone where the doctors can make analysis and report and send it to radiologist and hospital information system database for further treatment. Radiologist will send the report to patient for further treatment. The system architecture is shown below in fig 4:

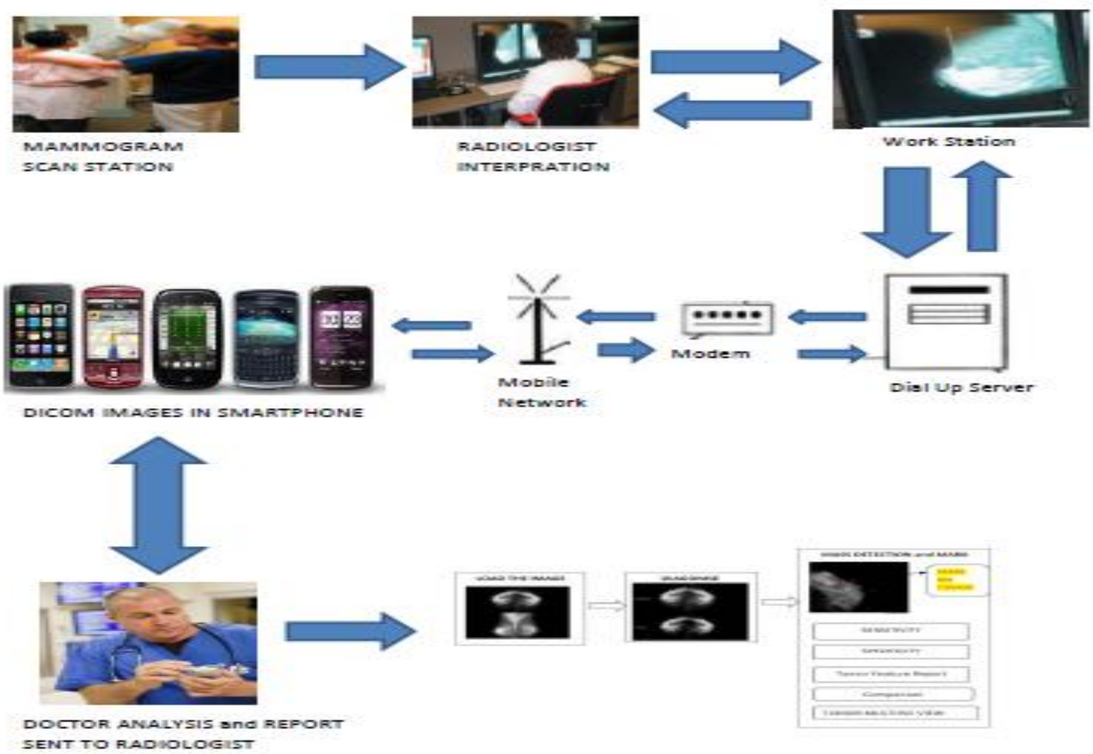

Figure 5 illustrates the research framework of Mobile Apps Deployment in Smartphone's

VI. RESEARCH METHOD FOR MAMMOAPPS

The MaammoApps Mobile application system methodology has 4 distinct steps: Image Acquisition & Display module, Transfer and storage module, Software development module and Report Feedback Module. The MaammoApps Mobile application system methodology has 4 distinct steps: Image Acquisition and Display module, Transfer and storage module, Software development module and Feedback Report Module.

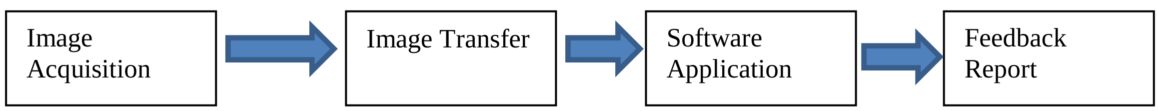

Figure 2 illustrates the research method flow for developing the application

A. Image Acquisition and Display

Once the patient completes the digital mammogram screening and images are collected for processing. The small matrix size images which includes computed tomography (CT), magnetic resonance (MR), ultrasound (US),nuclear medicine (NM), digital fluorography (DF), the data set should provide a minimum of 512 x 512 matrix size at a minimum 8-bit pixel depth i.e.256 grey levels for processing or manipulation with no loss of matrix size or bit depth at display(Ref: American College of Radiology). Modern smartphones have surpassed classical 240x320 resolutions. Smartphones screen height goes up to 854 pixels; width goes up to 480 pixels. They still lack 32 pixels to achieve 512 pixel widths. Image has to be resized to fit screen, or in full size with panning support. Display sizes vary from 3 to 4 inch, where larger display size is better. Color depth can be 16 bit, 18 bit or 24 bit, providing 40 to 256 grey levels. For practicing radiology many other display parameters of smartphone are relevant: color quality, absence of artefacts, peak brightness and black level brightness (Ref: Display mate).

B. Image Transfer Storage

By including IP protocol in mobile networks, mobile phones are capable of accessing Internet by mobile web browsers, email and other network services. Image transfer from PACS archive to mobile phone can be achieved by TCP/IP-based network protocols: which includes DICOM, HTTP, FTP, SMTP or proprietary network protocol. There are already mobile applications that implement DICOM network protocol. DICOM protocol is complex process and developers might consider using HTTP protocol. Benefit is that mobile web browser can be used to read DICOM images from PACS archive with WADO - web-based method of retrieving DICOM images from PACS archive by using HTTP/HTTPS protocol (Ref: DICOM in medicine). By using WADO method, PACS sends, on web browser HTTP request, DICOM image snapshot with specific W/L present in web browser readable formats such as JPEG/PNG/TIFF/GIF/BMP. Transferring or downloading speed varies by type of wireless network, which can be: short range Bluetooth which has speed of 2 Mbps and Wi-Fi has 54 Mbps; long range HSDPA has 14 Mbps, EDGE has 400 kbps and GPRS has 114 kbps. After transferring, storing of data on mobile devices is no longer a problem, since most devices have external storage on removable flash memory card, which can store up to 32 GB. Therefore, one memory card can hold several hundred CT studies on a mobile phone.

C. Software Development Module: (Image Processing and Tumour Identification)

Once images are transferred and stored, we carry the research on exploring two methods of application development: one is mobile operating system (OS) i.e. Apple IOS, Android OS, Windows Mobile, Blackberry RIM and mobile web browser. In Mobile OS, It is not the same if one develops mobile application for Asian, European or American market. There arises a main question is which devices radiologists use, or are about to use for their work, because new devices are coming on daily basis and every new device brings new features better than competitors. Therefore development should be based towards devices with best characteristics for DICOM image viewing, or for platform that provides best devices, or for platform that provides best developing tools. The main objective is to develop software that reads DICOM files, where main programming challenges are different image compression techniques (RLE, JPEG, JPEG2000) in reversible and irreversible form. DICOM Office toolkit and PixelMed Java DICOM Toolkit can be used for processing the DICOM images. Web-based applications are gaining popularity since mobile web browsers are capable of executing complex and dynamic tasks. Web-based application development is good way of bringing radiological images to mobile phones. Mobile web browsers are preferred for executing mobile radiological applications because:1) application is available to all mobile OS 2) support for Cascading Style Sheets (CSS) standard for changing image size, panning and UI development; 3) support for SVG (Scalable Vector Graphics) for vector based drawing on images (annotations, measurements) 4) support for main parts of HTML5 standard for web browser based image storage, dynamic image manipulation, applying various filters 5)mobile web browsers can handle large data sets and loading big web pages. In web based applications, tumour identification image processing algorithm will be implemented as developed for CAD systems.

D. Feedback Report Module

Once the doctor analysed the images and final conclusion report can be keyed in through the smartphone and sent to radiologist for further process of treatment. The feedback reporting system is a database useful for patient tracking. Pathology information on biopsies performed at the radiology practice is entered remotely into the individual databases. The same database is also used to send recall letters to patients when they are due for follow-up studies and create awareness among risk factor of breast cancer.

VII. EXPECTED OUTPUT

The expected output of the proposed system will be a Mobile application called as Mammo Apps which can be used in smartphone's and can assist the radiologist, doctors, experts and patients anywhere around the globe with internet connection. The major outputs achieved after developing the application are 1) Time factor 2) Technical quality 3) Diagnostic ability. Application will be first implemented in android based smart phone with Dicom office tool kit and it will tested with small dataset from MIAS database(Mammogram Image Analysis database), later it will be tested in Apple IOS and other advanced smartphones.

VIII. CONCLUSIONS

The proposed MammoApp application will have improved efficiencies in patient care; mammography interpretation; reporting and follow-up can be achieved. The advanced characteristics of today smartphone devices challenge application developers to develop new medical imaging applications. Mobile based digital mammogram applications can satisfy with image display, images retrieval and tumor identification toolset with basic and advanced requirements of radiologists in everyday medical image analysis. The MammoApps will fill the gap on interactivity between doctor and radiologist to make decisions. First time user may lack in interactivity but win in availability will be advantage of the proposed system.

ACKNOWLEDGMENT

We would like to thank the Ministry of Science and Technology, Malaysia for funding the project under E-Science grant scheme for developing the prototype and mobile application. Would like thank for the facilities and support provided by Institute of Post Graduate studies and School of Computer Science, University Sains Malaysia, Penang for the project.

REFERENCES

[1] Fajardo, L.J, M.B. Williams, The Clinical Potential of Digital Mammography. In Proceedings of the 3rd International Workshop on Digital Mammography, pp. 43-52.Chicago (USA), 1996.

[2] Vyborny, C.J., and M.L. Giger, Computer Vision and Artificial Intelligence in Mammography. American Journal of Roentgenology, Vol 162 pp. 699-708, 1994.

[3] W. Zhang et.al, “An improved shift-invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms," Medical Physics., vol. 23, pp. 595-601, 1996.

[4] D.Brzakovic et.al, "An approach to automated detection of tumors in mammograms," Medical Imaging, IEEE Transactions on, vol. 9, pp. 233-241, 1990.

[5] N.Petrick, H et.al, “Combined adaptive enhancement and region growing segmentation of breast masses on digitized mammograms, Medical Physics”, Vol. 26, pp.1642-54, 1999.

[6] S. M. Lai, X. Li, and W. F. Bischof, "On techniques for detecting circumscribed masses in mammograms," IEEE Transactions on Medical Imaging, vol. 8, pp. 377-386, 1989.

[7] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans System Man Cabernet SMC-9 , , 62–66, 1979.

[8] Hutt I, The computer-aided detection of abnormalities in digital mammograms, University of Manchester, Faculty of Medicine, UK, 1996.

[9] Tarassenko L., Hayton P., Cerneaz N., and Brady M, “Novelty detection for the identification of masses in mammograms”, Fourth International Conference on Artificial Neural Networks, Cambridge, 442-447, 1995.

[10] M.L. Comer, S. Liu, E.J. Delp, “Statistical segmentation of mammograms, Digital Mammography ’: Proceedings of the 3rd International Workshop on Digital Mammography, Chicago, IL, pp. 475-478 (9-12 June 1996).

[11] M.A. Kupinski, M.L. Giger, “Automated Seeded Lesion Segmentation on Digital Mammograms”, IEEE Transactions on Medical Imaging, 17(4), pp. 510-517, 1998.

[12] Ramon Lopez De, David Macsherry. et.al," Retrieval, reuse, revision, and retention in case based reasoning "The Knowledge Engineering Review, Vol.1–2. 2005.

[13] U. Bottigli, D.Cascio, F. Fauci, B. Golosio, R. Magro, G.L. Masala, P. Oliva, G. Raso, and S.Stumbo, "Massive Lesions Classification using Features based on morphological Lesion Differences", Proceedings Of World Academy Of Science, Engineering And Technology Vol 12, March 2006 .

[14] Vibha L, Harshavardhan G M. " Lesion Detection Using Segmentation and Classification in Mammogram”, 25th IASTED Conference on AI and its Applications", Austria, Feb 2007.

[15] Chuin-Mu Wang et al., "Classification for Breast MRI Using Support Vector Machine", IJCSNS International Journal of Computer Science and Network Security, Vol 9 No.5, May 2009.

[16] Serhat Ozekes,Onur Osman,A.Yilmaz Çamurcu. " Mammographic Mass Detection Using a Mass Template", Korean Journal of Radiology, Vol 6[4],December, 2005.

[17] Giulia Rabottino, Arianna Mencattini, Marcello Salmeri,Federica Caselli2, Roberto Lojacono,"Mass Contour Extraction in Mammographic Images for Breast Cancer Identification",16th IMEKO TC4 Symposium,Exploring New Frontiers of Instrumentation and Methods for Electrical and Electronic Measurements,Sept. 22–24, Florence, Italy, 2008.

[18] Wei Qian,Dansheng Song,Clark,MD. "Digital Mammography:Wavelet Transform and Kalman-filtering Neural Network in Mass Segmentation and Detection”, Academic Radiology, Vol 8, No 11, November 2001.

[19] Raman Valliappan and Putra Sumari “Digital Mammogram Segmentation : An Initial Stage “in 4th IASTED International conference on Advanced Computing Science and Technology, Langawi, Malaysia, 2008.

[20] Rabi Narayan Panda,Dr. Bijay Ketan Panigrahi,Dr. Manas Ranjan Patro, "Feature Extraction for Classification of Microcalcifications and Mass Lesions in Mammograms",International Journal of Computer Science and Network Security, VOL.9 No.5, May 2009

[21] Mohammed J. Islam et.al.: Computer-Aided Detection and Classification of Masses in Digitized Mammograms Using Artificial Neural Network, Lecture Notes in Computer Science, Volume 6146/2010, 327-33, 2010.

