Review on Mammogram Mass Detection by Machine Learning Techniques

Valliappan Raman, Putra Sumari, H.H.Then, and Saleh Ali K. Al-Omari

Abstract— Breast cancer continues to be a significant public health problem in the world and number one cause for death rate in Malaysia. Early detection is the key for improving breast cancer prognosis. Mammography is the most effective tool now available for an early diagnosis of breast cancer. However, the detection of cancer signs in mammograms is a difficult task due to irregular pathological structures and noise which are present in the image. It has been shown that in current breast cancer screenings 8%–20% of the tumors are missed by the radiologists. For this reason, a lot of research is currently being done to develop systems for computer aided detection to improve the accuracy. In this paper, review of mammogram mass detection and segmentation is focused. The main aim of the paper is to summarize and compares the method of mass detection in mammogram images. In specific, preprocessing, segmentation, feature extraction and classifications are discussed, Receiver operating curve and free-response receiver operating curve of each method is highlighted to show the sensitivity and specificity of the tumor detection.

Index Terms— Mammogram, preprocessing, segmentation, feature extraction and classification

I. INTRODUCTION

Early detection of breast cancer is considered as a major public health issue. Breast cancer incidence is the highest among female cancers and the second cause of mortality in the western countries [1]. To address this problem, it is necessary to create the adequate conditions allowing for the installation of mass detection campaigns, i.e. involving the maximum number of women at risk[2][3].

Mammography is the technique of choice to detect breast cancer and it is based on the difference in absorption of X-rays between the various tissue components of the breast such as fat, tumor tissue, and calcifications. Mammography has high sensitivity and specificity, even small tumors and micro calcifications can be detected on mammograms. The projection of the breast can be made from different angles.

The two most common projections are medio-lateral oblique (side view taken at an angle) and cranio-caudal (top to bottom view), as shown in Fig. 1.

Valliappan Raman is with Swinburne University of Technology Sarawak Campus, Kuching, Malaysia, Phone: 006082260872 email: vraman@swinburne.edu.my
Putra Sumari is with University Sains Malaysia, Penang, Malaysia. Email: putras@cs.us.my
H.H.Then is with Swinburne University of Technology Sarawak, Kuching, Malaysia. Email: pthen@swinburne.edu.my
Saleh Ali Omari is with University Sains Malaysia, Penang, and Malaysia. Email: salehalomari2005@yahoo.com
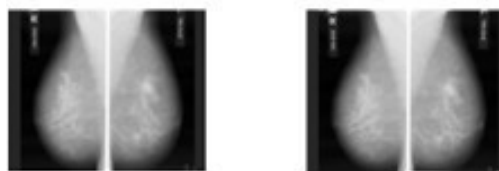

Fig. 1. Illustrates the medio Lateral and cranio-caudal view [4]

The two most important signs of breast cancer that can be seen on a mammogram are focal masses and micro calcifications. In this paper we are mainly interested in focal masses. When a mass is present in a breast, a radiologist will estimate its malignancy by looking at the appearance of the lesion and the surrounding tissue. The most important sign of malignancy is the presence of spiculation i.e. spiky lines radiating in all directions from a central region extending into surrounding tissue. Benign masses have sharp, circumscribed borders where malignant masses have slightly jagged or spiculated borders. The objective of the research is to reduce the error of false negatives and false positives. The presented work will detect mass lesions by analyzing a single view of the breast. The first step detects suspicious locations inside the breast area. In the second step the image at these locations is segmented into regions and several features are calculated for each region. These features are being used to determine whether a lesion is benign or malignant. They are also used to eliminate false positive detections.

In this paper, complete review of mammogram mass detection method is discussed. Challenges in reading mammogram images are highlighted in Section 2. Section 3 describes the various computer vision techniques of detecting the tumor in mammogram images. Section 4 explains the comparison analysis with ROC curve on accuracy of detecting tumor. Finally conclusion is made at Section 5.

II. CHALLENGES

When mammograms are examined retrospectively for signs of malignancy, the abnormality is considered occult or either classified as minimal sign or a screening error. An abnormality is called a minimal sign if something abnormal is found in the region of interest that is not suspicious enough to recall. If signs of malignancy are present that are actionable, it is called a screening error. There are two possibilities of error occurs while handled by radiologist. The first possibility is that the sign was overlooked, and has not been examined at all by the radiologist. The second possibility is that the sign was examined but was considered benign, normal, or not found suspicious enough for further examination. Apart from that there are other challenges

observed in detection mass are as follows:

- Intensity levels vary greatly across different regions in a mammogram, and features for segmentation are hard to formulate.
- Subtle grey level variations across different parts of the image make the segmentation of tumor areas by grey level alone difficult.
- Tumors (masses) are not always obvious.
- Poor illumination and high noise levels in the image that can vary up to 10-15% of the maximum pixel entity

III. METHODOLOGY- EXISTING METHODS TO DETECT MASS IN MAMMOGRAM IMAGES

Most of the mass detection algorithm will be performed in four phases. First, image preprocessing of the digitized mammogram can suppress noise and improve the contrast of the image. Second, image segmentation is defined by most of the articles about mass detection as locating the suspicious regions. It is different from the common definition of segmentation in image processing. In the third phase, features are extracted and selected for classifying lesion types or removing false positives. Finally, the detection/ classification of masses will be conducted. In this review preprocessing of mammogram image, detection of ROI and segmentation, feature selection and extraction, classification and evaluation are studied [5] [6].
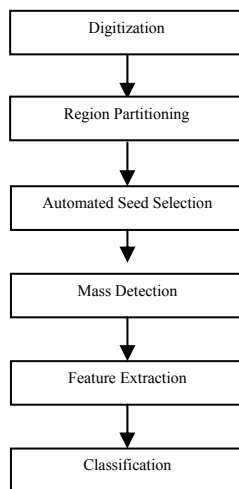

Fig. 2. Illustrates the phases of mammogram mass detection

A. Preprocessing- Digitization and Image Enhancement

In all research studies, first stage is, the 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 & microcalcifications 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.

Preprocessing is an important issue in low-level image processing. The underlying principle of preprocessing is to enlarge the intensity difference between objects and background and to produce reliable representations of breast tissue structures. There are several approaches applied in

enhancing the mammogram images and it is listed below from various research works

Global Histogram Modification Approach: A commonly used global histogram modification approach is the histogram equalization (HE).The main idea is to re-assign the intensity values of pixels to make the new distribution of intensities uniform to the utmost extent. Advantage: Effective in enhancing the entire image with low contrast. Drawback: Cannot enhance the textual information & working only for the images having one object

Local Processing Approach: Local-processing approaches are also studied for image contrast enhancement. There are many methods for contrast enhancement by changing pixel intensities. One way is based on nonlinear mapping methods (local histogram technique, bi-linear, sigmoid, non-continuous, etc.). The implementation can be feature-based, and the local features may be gained by edge detection, or by using local statistic information such as local mean, standard deviation. Advantage: Effective in local texture enhancement. Drawback: Cannot enhance the entire image well.

Histogram-based Intensity Windowing: HIW is a variant of intensity windowing. Intensity windowing allows a selected subrange of the image intensity values to receive the full contrast of the display device. All parts of the image with values outside the selected intensity window range are set to black (values below the minimum value of the intensity window range) or white (values above the maximum value of the intensity window range). HIW customizes standard intensity windowing by individually selecting the intensity window range for each image by statistically analyzing the histogram of each image, locating the "humps" or modes of the histogram, and determining which modes represent the different breast tissue types (fatty, dense, muscle) or other parts of the image (background, labels).

Mixture Model Intensity windowing: MMIW provides region-specific intensity window settings for mammograms. It operates by automatically identifying the five major regions in a mammogram: background, uncompressed fat, compressed fat, dense tissue, and muscle. It identifies these regions using a combination of geometric (i.e., gradient magnitude ridge traversal) and statistical (i.e., Gaussian mixture modeling) techniques.

Contrast-limited Adaptive Histogram Equalization: Adaptive histogram equalization maximizes the contrast throughout an image by adaptively enhancing the contrast of each pixel relative to its local neighborhood. This process produces improved contrast for all levels of contrast (small and large) in the original image. For adaptive histogram equalization to enhance local contrast, histograms are calculated for small regional areas of pixels, producing local histograms.

Unsharp Masking: With unsharp masking, a low-pass filtered version of the original image is created and the image values that result are subsequently multiplied by a weighting factor and subtracted from the original image

B. Detection and Segmentation

The second stage of mass detection CAD schemes is to separate the suspicious regions that may contain masses

from the background parenchyma, i.e., to partition the mammogram into several non-overlapping regions, then extract regions of interests (ROIs), and locate the suspicious mass candidates from ROIs. The suspicious area is an area that is brighter than its surroundings, has almost uniform density, has a regular shape with varying size, and has fuzzy boundaries. As per the survey four approaches are used for detecting and segmenting the tumor.

The global thresholding technique is based on the global information, such as the histogram. Since the masses are brighter than the surrounding tissues, it makes thresholding a useful method for segmentation. The regions with abnormalities impose the extra peaks on histogram while a healthy region has only a single peak. After a global thresholding value is attained, the objects can be separated from the background. Methods depending only on the global thresholding are not good to identify ROIs. Because mammograms are the 2D projections of the 3D breasts, the regions of overlapping tissues including three kinds of tissues: a fat region, a fatty and glandular region and a dense region, may be brighter than the masses. The output of the global thresholding is mainly used as an input to the next step in most of systems.

Local thresholding (LT) can refine the results of global thresholding or identify suspicious areas. LT is better for mass detection than global thresholding, because a local threshold value is determined locally for each pixel based on the intensity values of the surrounding pixels. Two variables of the local thresholding should be considered: the window size and thresholding value.

Region growing is one of the popular techniques for segmenting masses in digitized mammograms [7][8]. The basic idea of the algorithm is to find a set of seed pixels in the image first, and then to grow iteratively and aggregate with the pixels that have similar properties. If the region is not growing any more, then the grown region and surrounding region are obtained. If the grown region of a seed has an average intensity greater than that of the surrounding, the region is classified as the parenchyma, or fat, tissue. The accuracy reaches 90% for classifying the tissue patterns. Region clustering and region growing are very similar. Region clustering searches the region directly without any prior information.

Edge detection is a traditional method for image segmentation. There are a lot of operators, Roberts's gradient, Sobel gradient, Prewitt gradient, Laplacian operator, etc. The combined edge detection method was developed to increase accuracy.

DWCE is used in two stages, first applies it globally to isolate the suspected area, then uses it locally to refine the segmentation. It is in conjunction with LoG filter. Logic filter is a nonlinear filter, and logic operators AND, OR and XOR are used. The concrete logic expressions depend on the prior information, and the filter structure influences the results. Iris filter is an adaptive filter. It is applied locally Gaussian filter ROIs are highlighted by a DOG filter. It can reduce number of FPs. Deformable models it is a contour or interface which after initialization moves according to its local properties and the priori information of the object. It is good in finding the contour of the suspected area. The performance may depend on the initialization.

Template matching is one of the most common approaches for medical image segmentation. This method uses the prior information of mammograms, and segments possible masses from the background using the prototypes. The prototypes of possible masses are created based on the characteristics or physical features of the targeted masses or based on the two-dimensional search function. When the priori information about the size of the masses is not available, a range of sizes for the templates is used.

Fuzzy logical has been introduced for segmenting suspicious masses. The algorithm first assigns a fuzzy membership value to each pixel, and then an error value is calculated in iteration and the fuzzy membership is updated. Effects of the neighboring pixels are also considered in the update rules. The algorithms stops when a zero error is reached, indicating that each pixel was assigned to either the bright or dark region, i.e., the mass region or background region. It has been proved that fuzzy set theory coupled with texture-based algorithms was very useful for the classification of masses. There are basically two kinds of fuzzy methods: fuzzy thresholding and fuzzy region clustering or growing.

C. Feature Selection and Extraction

The third stage of mass detection is the feature extraction and selection. The features can be calculated from the ROI characteristics such as the size, shape, density, and smoothness of borders, etc. The feature space is very large and complex due to the wide diversity of the normal tissues and the variety of the abnormalities. Only some of them are significant. Using excessive features may degrade the performance of the algorithm and increase the complexity of the classifier. Feature selection is the process of selecting an optimum subset of features from the enormous potential features available in a given problem domain after the image segmentation. The feature space can be divided into three sub-spaces: intensity features, geometric features, and texture features. The feature extraction and selection processes for mass detection can base on the principle component analysis, linear discriminate analysis and GA algorithm.

D. Classification of Masses

Once the features related to masses are extracted and selected, the features are input into a classifier to classify the detected suspicious areas into normal tissues, benign masses, or malignant masses. Classifiers such as the linear discriminants (LDA), Bayesian, CBR, artificial neural network (ANN) have performed

E. Existing Research Study

There are several existing approaches were made to detect the abnormal tissues in breast images and to detect the cancer Zhang et al [9] 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 [10] 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 [11] 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 [12] 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 [12] 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 [13], work is based on fuzzy pyramid linking algorithm which was used to detect microcalcifications 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 [14] 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.[15] 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 [16] developed a method, which combines region growing with probability analysis to determine final segmentation. Ramon Lopez De Mantaras [17], 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. [18] 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 [18]; 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 [19], 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%. By applying SVM classifier for classification accuracy reached 89.9% [20].

Serhat Ozekes.et.al [21] 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 [22] 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 (CADe) should be able to identify tumoral lesions as a second human reader.

Wei Qian [23] 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 subimage 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.

Rabi Narayan Panda.et.al [24], 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.

Kegelmeyer et al [25] developed a method to detect spiculated masses using a set of 5 features for each pixel. They used the standard deviation of a local edge orientation histogram (ALOE) and the output of four spatial filters which are a subset of Law's texture features. The idea of using the ALOE feature is that a normal mammogram exhibits a tissue structure that radiates in a particular orientation (from the nipple to the chest). A spiculated mass would change this trend and thus normal tissue would have edge orientations in a particular direction whereas, in suspicious regions containing spiculated lesions, edges would exist in many different orientations. To detect this difference Kegelmeyer et al. [24] computed edge orientations in a window around each pixel and then generated a histogram of edge orientations.

Karssemeijer et al. [26] detected stellate distortions by a statistical analysis of a map of pixel orientations. The orientation at each pixel was computed from the response of three filter kernels, which are second-order, directional derivatives of a Gaussian kernel in the directions $(0, \pi 3, 2\pi 3)$.

Matsubara et al [27] developed an adaptive thresholding technique for the detection of masses. They employed histogram analysis techniques to divide mammograms into 3 categories ranging from fatty to dense tissue. Potential masses were detected using multiple threshold values based on the category of the mammogram. A number of features such as circularity, area, and standard deviation were used to reduce the number of false positives.

Polakowski et al [28] used a single Difference of Gaussian (DoG) filter to detect masses. The DoG filter was designed to match masses which were approximately 1 cm in diameter. ROIs were selected from the filtered image. They used nine features based on size, contrast, circularity and Laws texture features to reduce the number of false positives and to then classify ROIs as malignant or normal. The above methods show less than five false positives per image with a true positive detection rate of approximately 90%. It is difficult to compare the performance of these methods because their databases are different.

F. Commercial CAD Systems

There is Three FDA approved commercially available computer-aided detection (CAD) systems have been developed to aid radiologists to detecting mammographic abnormalities. Currently, there are not FDA approved systems for computer-aided diagnosis (CADx).

R2 Technology, Inc. R2 Technology's Image Checker® was the first commercial mammographic CAD system approved by the FDA [29]. This device is designed to search for signs that may be associated with breast cancer. Masses are marked with an asterisk while microcalcification clusters are marked with a triangle. The detection accuracy of calcifications was reported as 98.5% sensitivity at 0.74 false

positives per case (set of four images). The detection accuracy of masses was reported as 85.7% at 1.32 false positive marks per case.

Intelligent Systems Software, Inc. The FDA approved ISSI's CAD system MammoReader™ in 2002. MammoReader™ was designed to detect primary signs of breast cancer in mammogram images including microcalcification clusters, well and ill-defined masses, spiculated lesions, architectural distortions, and asymmetric densities [30].

CADx Medical Systems. CADx Medical systems were the third company to receive approval for a mammographic CAD system called SecondLook™ [31]. SecondLook™ was designed to mark areas of a mammogram that are indicative of cancer. It marks masses with circles and microcalcification clusters with rectangles. Thus, while above systems now provide radiologists with a powerful aid for calcification detection, improvements are needed in mass detection

IV. EVALUATION

Performance evaluation in algorithm design is a commonly neglected concept. Few mammogram segmentation algorithms have been tested extensively. Abdel-Mottaleb et al [32] test 500 mammograms, with their algorithm finding an "acceptable" boundary in 98% of the images. Méndez et al. [33] test their algorithm on 156 mammograms of which the breast contour is deemed to be "accurate" or "nearly accurate" in 89% of the images. Bick et al. [34] test their algorithm on 740 mammograms, and 97% of the segmentation results were visually rated as "acceptable". Chandrasekhar and Attikiouze [35] use all the images from the MIAS database, with the algorithm providing about 94% acceptable results. Others, like Ojala et al. [36] make no mention of the extent of the testing, and the illustrated mammograms contain visible contours. Yet what constitutes an "acceptable" result differs significantly, and is often based on visual subjective opinion with very little quantitative endorsement. The accuracy of this technique was evaluated through quantitative measures derived through the comparison of each segmented mammogram "mask" with its corresponding "gold standard". A quantitative measure was then derived to describe the accuracy of the segmentation. The region extracted by the segmentation algorithm (mask) which matches the GT is denoted as true positive (TP) emphasizing that the algorithm has indeed found a portion of the breast. Pixels shown in the GT but not shown in the mask are defined as false negative (FN) classifications. These are considered missing pixels in the breast region. Conversely, the pixels not in the GT, but in the mask are deriving two metrics: completeness (CM) and correctness (CR). The completeness is the percentage of the GT region which is explained by the segmented region and Correctness is the percentage of correctly extracted breast regions

$$SENSITIVITY\ (Completeness) = \frac{TP}{TP + FN}$$

$$SPECIFICITY\ (Correctness) = \frac{TP}{TP + FP}$$

Another way to evaluate the constructed systems, the classification performance of each system has to be measured. Often the performance of a system cannot be described by a single value [37]. Performance measure is done by ROC curve. In a Receiver Operator Characteristics (ROC) curve the sensitivity, which in this study is the share of malign tumors that is correctly classified, is plotted against 1-specificity, the share of benign tumors that is falsely classified, for different cut values. Often the ROC analysis is used to find an optimal cut value sometimes referred to as criterion, for use in decision-making. By changing the cut value of the system it is possible to achieve the optimal balance between sensitivity and specificity that is needed for a certain purpose. If the cost of not detecting a particular disease is very high to society, for example a highly contagious disease, one could change the cut value to achieve a very high sensitivity, but consequently lower specificity.

V. CONCLUSION

This paper discusses the mammogram image processing approaches for detection of masses in early diagnosis of breast cancer. In this review paper, every technique that has been employed in each stage of machine learning approach is explained and including the techniques that has been used to measure the performance of proposed systems. ROC and FROC analysis are standard methodologies for measurement of performance of detection and diagnosis algorithms in CAD systems. Currently, many researchers evaluate their system's performance using these evaluation methodologies. The vast amount or research related to analysis of mammography, as well as widespread interest from the medical community stimulates the development of commercial CAD systems. Although there are many outstanding performances have been achieved by mammogram CAD systems, the challenges and future directions of research are still remaining in the Field of CAD for mammogram mass & mcc detections.

REFERENCES

[1] R. Valliappan, P. Sumari, and M. Rajeswari, "A Theoretical Methodology and Prototype Implementation for Detection Segmentation Classification of Digital Mammogram Tumor by Machine Learning and Problem Solving Approach," *IJCSI Journal*, Vol. 7(5), pp.1694-0814, September 2010.

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

[2] C. J. Vyborny and M. L. Giger, "Computer Vision and Artificial Intelligence in Mammography" *American Journal of Roentgenology*, Vol.l62, pp. 699-708, 1994.

[3] Mammogram Image Analysis Database, UK.

[4] R. Valliappan, P. Sumari, and P. Then, "Matlab Implementation and Results of Region Growing Segmentation Using Haralic Texture Features on Mammogram Mass Segmentation" in *IEEE Conference on ICCSEA*, Dubai, UAE, , May 2011, pp. 10-14.

[5] R. Valliappan, P. Sumari, and P. Then, "Mammogram Problem Solving Approach: Building CBR classifier For Classification of Masses in Mammogram" *in IEEE Conference on CNC*, Bangalore, India , Feb 2011, pp.284-289.

[6] R. Valliappan and P. Sumari "Digital Mammogram Segmentation: An Initial Stage*," 4th IASTED International conference on Advanced Computing Science and Technology*, Malaysia, 2008, pp. 200-204.

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

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

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

[10] N .Petrick et.al, "Combined adaptive enhancement and region growing segmentation of breast masses on digitized mammograms", *Journal of Medical Physics*, Vol. 26, pp.1642-54, 1999.

[11] 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.

[12] I. Hutt, "The computer-aided detection of abnormalities in digital mammograms", Ph.D. dissertation, Faculty of Medicine, University of Manchester, UK, 1996.

[13] L. Tarassenko, P. Hayton, N. Cerneaz, and M. Brady, "Novelty detection for the identification of masses in mammograms*," Fourth International Conference on Artificial Neural Networks*, Cambridge, UK, 1996, pp.442-447.

[14] 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, June 1996, pp. 475-478.

[15] M. A. Kupinski and M. L. Giger, "Automated Seeded Lesion Segmentation on Digital Mammograms," *IEEE Transactions on Medical Imaging*, Vol.17 (4), pg. 510-517, 1998.

[16] R. L. d. Mantarasi, D. Mcsherry et.al, Retrieval, reuse, revision, and retention in case based reasoning. *The Knowledge Engineering Review Journal*, Vol.1–2, 2005.

[17] 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, UK*. March 2006, pp.12-14.

[18] L. Vibha, and G. M. Harshavardhan, "Lesion Detection Using Mammogram Segmentation," *Proceedings of 25th IASTED Conference on AI and its Applications*, Austria, Feb 2007, pp 112-115.

[19] C. M. Wang et al, "Classification for Breast MRI Using Support Vector Machine," *International Journal of Computer Science and Network Security,* Vol.9, pp. 115, May 2009.

[20] S. Ozekes, O. Osman, and A. Yilmaz Çamurcu, "Mammographic Mass Detection Using a Mass Template", *Korean Journal of Radiology*, Volume 6[4], December 2005.

[21] G. Rabottino, A. Mencattini, M. Salmeri, F. Caselli, and R. Lojacono, "Mass Contour Extraction in Mammographic Images for Breast Cancer Identification," *16th IMEKO TC4 Symposium on Exploring New Frontiers of Instrumentation and Methods for Electrical and Electronic Measurements* , Florence, Italy, Sept. 22–24, 2008, pp. 11-15.

[22] W. Qian, D. Song, M D. Clark, "Digital Mammography: Wavelet Transform and Kalman-filtering Neural Network in Mass Segmentation and Detection," *Journal of Academic Radiology*, Vol 8, No 11, November 2001.

[23] R. N. Panda, B. K. Panigrahi, and M. R. 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.

[24] W. P. Kegelmeyer, Jr, J. M. Pruneda, P. D. Bourland, A. Hillis, M. W. Riggs, and M. L. Nipper,"Computer-aided mammographic screening for spiculated lesions," *Journal of Radiology*, Vol. 191, pp. 331-7,1994.

[25] N. Karssemeijer and G. M. Te Brake, "Detection of stellate distortions in mammograms" *IEEE Transactions on Medical Imaging*, vol. 15, 1996.

[26] T. Matsubara, H. Fujita, T. Endo, K. Horita, M. Ikeda, C. Kido, and T. Ishigaki, "Development of mass detection algorithm based on adaptive thresholding technique in digital Mammograms," *Digital Mammography Symposium*, Japan, June 1996.

[27] W. E. Polakowski, D. A. Cournoyer, S. K. Rogers, M. P. DeSimio, D. W. Ruck, J. W. Hoffmeister, and R. A. Raines, "Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency," *IEEE Transactions on Medical Imaging*, Vol. 16, pp. 811-819, 1997.

[28] U. S. Food and Drug Administration, "Summary of Safety and Effectiveness Data: R2 Technologies," ISSI Article, PP9700758, 1998.

[29] U. S. Food and Drug Administration, "Summary and Safety of Effectiveness Data", ISSI Article," P010038, 2002.

[30] U. S. Food and Drug Administration, "Summary of Safety and Effectiveness Data: CADx Medical Systems," ISSI Article, P010034, 2002.

[31] M. Abdel-Mottaleb, C.S. Carman, C.R. Hill, and S.Vafai, "Locating the Boundary between the Breast Skin Edge and the Background in Digitized Mammograms*," in proceedings of the 3rd International Workshop on Digital Mammography*, 1996, pp.467-470.

[32] A.J. Mendez, P.J. Tahoces, M.J. Lado, M. Souto, J.L.Correa, and J.J. Vidal, J.J. "Automatic Detection of Breast Border and Nipple in Digital Mammograms", *Journal of Computer Methods and Programs in Biomedicine*, Vol 49, pp.253-262. 1996.

[33] U. Bick, M.L. Giger, R.A. Schmidt, R.M. Nishikawa, D.E, Wolverton, and K. Doi. "Automated Segmentation of Digitized Mammograms", *Academic Radiology Journal*, Vol 2(2), pp.1-9, 1995.

[34] R. Chandrasekhar and Y. Attikiouzel. "Gross Segmentation of Mammograms using a Polynomial Model. In proceedings of the *International Conference of the IEEE Engineering in Medicine and Biology Society,* UK, 1996, pp.1056-1058.

[35] T. Ojala, J. Näppi, and O. Nevalainen, "Accurate Segmentation of the Breast Region from Digitized Mammograms," *Computerized Medical Imaging and Graphics Journal*, Vol 25(1), pp.47-59, 2001.

[36] J. Nagi, A. Kareem, S, F. Nagi, and S. Khaleel Ahmed, "Automated breast profile segmentation for ROI detection using digital mammograms" *IEEE Symposium on Biomedical Engineering and Sciences (IECBES*), Kualalumpur, Malaysia, Nov 2010.

Valliappan Raman is currently doing his PhD in mammogram tumor segmentation and classification at University Sains Malaysia. He has completed his Master's Degree in 2005 at University Sains Malaysia and completed his Bachelor of Engineering in Computer Science in 2002 at Madurai Kamaraj University. He has been working as lecturer in Swinburne University of Technology Sarawak campus.

He have undergone various research projects under major government grants and published papers, articles and journals. His research interest is in medical imaging, watermarking and health informatics

Putra Sumari is currently an Associate professor in School of Computer Science, University Sains Malaysia. He has undergone various research projects under government and university grants. He has supervised many postgraduate and undergraduate students. His research areas are in applied informatics, multimedia and image processing. He has published many papers in highly repotted journal and conferences

H.H Then is currently Senior lecturer in Swinburne University of Technology Sarawak Campus. He is the head of Advanced of Informatics Group Lab. He have undergone and headed various research projects funded by Malaysian national grants and published papers in highly repotted journals and conferences. His research area of interest is on watermarking, computer vision and pattern recognition, data mining & bio-informatics.

Salleh Omari is a post graduate student in University Sains Malaysia. His research interest is in multimedia contents, multimedia storage, MPEG compression and image processing.