IPASJ International Journal of Computer Science (IIJCS)
Web Site: http://www.ipasj.org/IIJCS/IIJCS.htm
A Publisher for Research Motivation ....
Email: editoriijcs@ipasj.org
Volume 2, Issue 8, August 2014
ISSN 2321-5992

A survey of Image Processing techniques for Detection of Mass

Lothe Savita A.[1], Dr. Deshmukh Prapti D.[2]

[1]Vasantrao Naik Mahavidyalaya, Aurangabad, MS(India)

[2]Principal, MGM, Dr. G. Y. Pathrikar College of Computer Science & IT, Aurangabad, MS (India).

ABSTRACT

Mammography is the best available examination for the detection of early detection of breast cancer. The most common breast abnormalities that may indicate breast cancer are masses and calcifications. Computer-Aided Detection (CAD) systems have been developed to aid radiologists in detecting mammographic lesions that may indicate the presence of breast cancer. These systems act only as a second reader and the final decision is made by the radiologist. Recent studies have also shown that CAD detection systems, when used as an aid, have improved radiologists' accuracy of detection of breast cancer. This paper gives a survey of image processing techniques for detecting mass from mammogram images.

Keywords : ROI, RBFNN, SGLD, TS, k-nearest neighbors, PCNN, computer aided detection.

1. INTRODUCTION

X-ray mammography is the most common techniqueused by radiologists in the screening and diagnosis ofbreast cancer in women. Although it is seen as the bestexamination technique for the early detection of breastcancer reducing mortality rates by up to 25%, theirinterpretation requires skill and experience by a trainedradiologist [1]. Mammographic interpretation can be considered a two-step process. A radiologist firstly screens the mammograms for abnormalities. If a suspicious abnormality is detected, further diagnostic workup isperformed to estimate the likelihood that the abnormality is malignant. The main mammographicsigns of breast cancer are clustered by microcalcifications and masses [2].A mass can be roughly considered as a circle with aluminance that grows from its border to its center. Recognizea mass in a mammogram is difficult because of their lowcontrast with respect to the surrounding healthy tissue [3].

Following figures shows the masses from mammogram images:

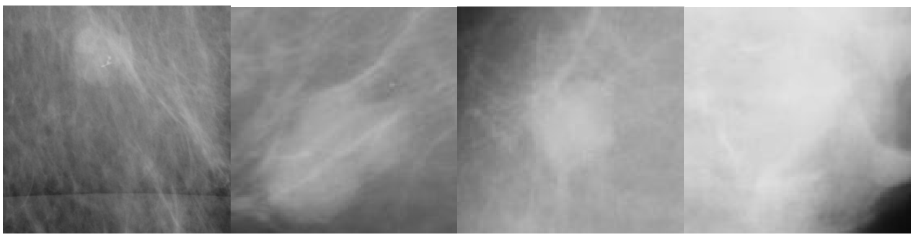

Fig 1: Mammographic images showing Masses

2. LITERATURE SURVEY

Chen et al [4], utilizes the spatial frequency or resolution information provided by the MWA (Multiresolution Wavelet Analysis) in conjunction with the contextual information provided by GMRF (Gaussian Markov Random field) for the analysis of mammographic subtle mass regions and microcalcifications. The mammograms are first studied in acoarse to fine resolution framework. In general larger lesions are characterized by coarser resolutions whereas higher resolutions show finer and more detailed anatomical structures. Nonstationary GMRF or 2D noncausal image models are then used for texture analysis and adaptive feature extraction on the wavelet transformed images. This hierarchy of multiresolution features contains pertinent information concerning the fuzzy edges of the mass regions or microcalcifications that is not otherwise available in the features extracted from a single resolution alone. The discriminating localized features obtained as a result of MWA and GMRF are then used in both the detection of microcalcifications and segmentation of mass regions in conjunction with FCM clustering.Based on a soft decision concept, the clustering resultfurnishes a good initial estimate for the maximum aposteriori (MAP) estimator approximated via constrained optimization and expectation maximization. Theformulation of the Bayesian paradigm using Gibbsdistribution provides a natural image model formammograms with further enhancements of segmentationresults.

Christoyianni et al [5], presented a novel method for fast detection of regions of suspicion (ROS) that contain circumscribed lesions in mammograms. The position and the size of ROS are first recognized with the aid of a Radial-Basis-Function neural network (RBFNN) by performing windowing analysis. RBFNN method is able to make a decision whether a mammogram is normal or not and then detects the masses' position by performing sub-image windowing analysis. In the latter case, with the implementation of a set of criteria, square regions containing the masses are marked as region of suspicion. A fast feature extraction reduces significantly the overall processing time. Then a set of criteria is employed to these regions to make the final decision concerning the abnormal ones. Accelerated estimation of the high-order statistical features decreases the computational complexity 55 times in multiplication operations.

Velthuizen et al [6], presented a techniques to find location of a mass, which is automatically segmented using fuzzy clustering. Features are extracted from the segmentation results using morphological, first-order statistical, and texture measures. Selection of relevant features is done using sequential selection. Fitness functions are based on the scatter matrices, k-nearest neighbors classifier, or neural network classifier using two-fold cross validation. The diagnosis is then provided by a trained three layer neural network. Feature selections provides a dramatic reduction in the number of required measurements to less than 25 as well as improve the accuracy of the results, from about 70% correct to 82% correct. The area under the ROC curve also increased dramatically. Computer vision on mammographic masses results in a very complex data space, that requires careful analysis for the design of a classifier.

Bovis et al [1], uses the texture features for the detection of masses, involved five phases as:(i) Location of a Common Reference Point: which states that alignment of left and right breast image pairs priorto bilateral subtraction a common reference, the spatialposition of the nipple, is located.(ii) Alignment and Bilateral Subtraction of Leftand Right Breast Images: Using aligned left and right breast image pairs,two images are generated by bilaterally subtracting oneimage from the other.(iii) Reduction of False Positives: Each difference image generated from bilateralsubtraction. (iv) Feature Extraction: is done using the quad-tree region model generated for all remaining suspicious regionswithin a difference image, five co-occurrence matricesare constructed in four different spatial orientations,horizontal, left diagonal, vertical and right diagonal, (0^0, 45^0, 90^0, 135^0respectively). A fifth matrixis alsoconstructed as the mean of the first four. Each co-occurrence matrix reflects the joint probability of a pixelpair at a given orientation and distance. (v) Classification: In order to determine the discriminating effectivenessof texture features extracted from co-occurrence matricesconstructed at different distances, a classification isinitially performed using each normalised feature vectorand linear discriminant analysis. The feature vector givingthe best true positive fraction was selected for subsequentclassification using an Artificial Neural Network (ANN)developed using the Stuttgart Neural Network Software (SNNS) package. For classification using an ANN, theselected feature vector is normalised and principalcomponent analysis performed to reduce thedimensionality of the data.

Tweed et al [7], presented a simple method that computes texture and histogram thresholds and selects regions of interest in order to detect tumours in mammographies. The method is based on a combined analysis of texture and histogram. The ROI thus constructed have the three properties: a reduced pixel number in the initial image is retained, no region containing a cancer is excluded, the feature variability is lower in the selected region than in the whole image. Indeed, the quality computed is generally close to 45 % and the ROI thus constructed is twice too big. In some particular cases, the quality computed is between 5 and 10% and the ROI is then not well delimited around the tumour. Moreover several ROI can be selected in a mammography, one of them containing the tumour. Their work consist on segmenting the ROI into objects and extracting shape factors from each segmented ROI. We are also studying on how to determine automatically the two histogram thresholds from the histogram of the whole image. A final analysis of the texture, histogram and shape factors will enable the determination of the factors that are relevant in order to classify the ROI into one of these three groups: normal, benign or malignant.

Hassanien et al [8], the main obstacle to analysis mammogram imageslays in low contrast between normal and malignant glandulartissues and the noise in such images that makes it verydifficult to segment them. Therefore, in digital mammogramthere is a need for enhancing imaging before areasonable segmentation can be achieved. Contrast enhancementis useful when an area of the image that is ofparticular importance has only subtle changes in pixel intensity.In this case, it may be difficult for the human eye tomake out the structures clearly, especially if the image is beingdisplayed on a low quality screen. By exaggeratingthe changes in pixel intensity the image may becomeeasier to interpret. By applying the contrast enhancementfilter will improve the readability of areas with subtlechanges in contrast but will also destroy areas of the image where the intensity of the pixels is outside the rangeof intensities being enhanced .in this paper,we adopt Fuzzy Histogram Hyperbolizationalgorithmto highlight most areas that contains tumors to help the segmentationprocess. develops an automated algorithmfor segmenting spiculated masses of the mammogramimages based on Pulse Coupled Neural Networks (PCNN)in conjunction with fuzzy set theory. The pulse-coupled neural network is a neural networkthat has the ability to extract edges, image segments and textureinformation from images. The PCNN is very generic.Only a few changes are necessary to effective operating ondifferent types of data. Each neuronin the processing layer is directly tied to an image pixelor a set of neighboring image pixels, these are the feedinginputs, and they are also linked to nearby neurons,the linking inputs.

Chuin-Mu et al [9],uses three groups of characteristics related to mass texture are adopted, namely, SGLD (Spatial Gray Level Dependence), TS (Texture Spectrum) and TFCM (Texture Feature Coding Method) to describe the characteristics of masses and normal textures on digitized mammograms. Next, under the testing by classifiers, three Feature Selection Methods**--**SBS (Sequential Backward Selection), SFS (Sequential Forward Selection) and SFSM (Sequential Floating Search Method) are used to find out suboptimal subset from 19 features in order to improve the performance of mass detection. Finally, two classifiers PNN (Probabilistic Neural Network) and SVM (Support Vector Machine) are applied and their performances are compared.

Jahanbin et al [10], determining the size of the ROI depends on the response of a set of unique spiculation Filters (SF). The design of these filters is based on manually annotated physical characteristics of spicules. The accuracy of algorithm is measured in terms of the percentage of spicule pixels located inside the identified ROI. Spicules on each image were identified by an experienced radiologist to serve as a reference to determine the percentage of spicules located in the ROI. On average, 94 percent of spicule pixels were located inside the ROI identified by algorithm. The spiculation filters have three parameters. 1) radius 2) frequency and 3) σ. The radius of the filter is the size parameter measured in pixels and its value corresponds to the length of the spicules. The frequency corresponds to the number of spicules located per circumference of the central mass region. σis the standard deviation in pixels. The primary goal of the paper is to identify a region in the mammogram image that completely contains the candidate mass and its spicules. It is important that the dimensions of the ROI be not too small, or else some information regarding the candidate spiculated lesion will be lost.

M. Hanmandlu et al [11], uses the three steps for segmenting mass – background subtraction, fuzzy texture representation and entropic thresholding. Background subtraction technique estimates the low-frequency background by using the grey level of band of pixels around the perimeter of the ROI. This method is based on visual comparison of the background-corrected images with the original image so that the background is levelled and no artifact is present. Running average of the pixel values along the perimeter of the ROI is calculated using a box filter of a 32 x 16 kernel, of which the long dimension is parallel to the edge of the ROI. For the perimeter pixels that are within 16 pixels of one of the four corners of the ROI, the long dimension of the box filter kernel is reduced on the side that is limited by the ROI edge. Then to convert the spatial domain image into the fuzzy domain, they consider the spatial arrangement of gray levels of pixels over a window. The fuzzy property can be expressed in terms of a membership function, provides the defuzzified response which is subjected to entropic thresholding used for mass segmentation to extract the suspicious tumor from the background.

J. Blanc-Talon et al [12],proposed a Computer Aided Detection (CAD)system for the extraction of the contour of tumoral masses from this ROI. ThisCAD consists in the following steps: artifacts removal, Contrast enhancement, Segmentation by region growing algorithm, Peninsulas removal. In performing segmentation they consider

1) Seed choice: the seed is chosen approximately at the center of the mass and its value is setto the average of the 15x15 neighboring pixels. In this way, the local intensity ofthe seed pixel does not influence the growing process.

2) Similarity criteria: For the inclusion of new pixels in the growing region, weverify that the pixel luminance is between the two thresholds:

$$Th1 = \bar{I} - (0.3 - K) \cdot \bar{I} \qquad (a)$$

$$Th2 = \bar{I} + (0.3 + K) \cdot \bar{I} \qquad (b)$$

where$\bar{I}$ is the average value of luminance in segmented region and it changesat each growing step, and K is a parameter depending on both luminance anddistance from the seed.

3) Optimization: Testing different images with the complete algorithm and the optimized one we note that the processing time of the complete algorithm is 8, 6 and 5 times the processing time of the algorithm with 1, 2 and 3 iterations respectively.

Mencattini et al [3], implements the procedure to indicates the mass regions in the mammographic image for finding tumoral signs as

- Decimate the original image: To reduce the computationaltime of the algorithm, we consider 1 pixel every 10. Inthis way, it can identify smallestmasses that at least have a diameter of 3 mm, that, witha spatial resolution of 42 μm, corresponds approximatelyto70 pixels.
- Segment the background. To reduce the computationaltime, they have not consider the film background, isolating itsimply with a threshold. Considering the characteristicsof the mammographic images, with experimentallyset a threshold so that also the lighter background in thedatabase can be segmented. They prove that thisthreshold value segments the background without alteringthe breast tissue.

- Choose a grid. Applying the algorithm on every pixel ofthe original image requires a lot of time. Applying the algorithm only on some pixel taken by gridding theoriginal image with a stepsthat have to be chosencarefully. And represents a measure of the convergence of gradient vectors in the circle on the pixel of interest (i, j).
- Repeat for different radii and compute, for every point(i, j) of the grid and finally sort the results.

Raman & Patrick et al [13], performs digitizations and decimation of mammogram image, breast region is extracted which is divided into three partitioned regions the fat, glandular and dense region (using otsu's method). Then, automatic seed selection and segmentation by region growing are performed for each partitioned region. Nine features(Global{skewness, kurtosis, circularity, compactness, perimeter} and local features{mean, contrast, standard deviation, area}) are extracted from the segmented masses. The criteria for the feature selection are based on morphological differences between lesions and healthy regions. Perform the classification by casebase reasoning (CBR) 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.

Samma et al [14], presents a method consists of four stages, i.e., localization using hybrid of morphological operations, segmentation using Chunming's algorithm, feature extraction using wavelet transform, and feature selection and classification using a hybrid GA-SVM model.

Wang et al [15], applied three approaches for the task, including feature selection using a neural classifier (back propagation), a clustering criterion and a combined scheme. To evaluate the performance of these feature selection approaches, a same neural classifier is then applied using the selected features and the classification results are then compared. Feature selection is to choose the most representative components among the feature set, and the principle is based on the fact that these selected features should be most discriminative in classifying the training samples. This discriminative ability is determined using two basic classifiers, including neural network and statistical clustering.

Yu-Shun et al [16], applied the Fuzzy theory and linear discriminant Analysis (LDA). First, the images were automatically segmented, and then targeting this feature, the brightness values of suspicious regions were retained while the brightness of other areas was reduced to find the suspected location of the cancer. After that, with the images obtained, Law's Mask and grayscale value momentum-intensive technologies were adopted to find the texture of each area. Finally, the Fuzzy LDA was adopted to identify the texture of the cancers in order to capture images of the cancer areas.

Giordano et al [17], proposed CAD system consist of 3 steps:1) A pre-processing module that aims at eliminating both eventual noise introduced during the digitization and other uninteresting objects; 2) A mass detection module that relies on a contrast stretching method that highlights all the pixels that likely belong to masses with respect to the ones belonging to the other tissues and on a wavelet-based method that extracts the candidate masses taking as input the output image of the contrast stretching part. The selection of the masses (among the set of candidates) to be passed to the classification module is performed by exploiting a-priori information on masses. 3) A mass classification module that works on the detected masses with the end of distinguishing the malignant masses from the benign ones using support vector machine whose features are the spatial moments extracted from the identified masses.

D.-S. Huang et al[18],presented a method to classify masses using level set based segmentation which can be implemented in two ways, explicitly or implicitly. In theexplicit method, the contour is represented with discrete points, and in the implicitmethod, the contour is represented with a level set function in higher dimension. Thelevel set based implementation has the advantage to better accommodate thetopological changes during evolution, normalized accumulative angle feature was extracted from the boundary of a mass. Linear Discriminant Analysis and Support Vector Machine were investigated for the classification. Theexperiments were tested using a database of 292 clinical mammograms. Resultsobtained demonstrate that SVM based classification from Fourier descriptor ofnormalized accumulative angle yielded an encouraging accuracy of Az=0.8803.

Faye et al[19], are used the classifiers K nearest neighbors (kNN) and Discriminant Analysis (DA) with Wavelet transform. A method of classification of images using multiscale transforms and random feature selection method is investigated. The method is based on preselecting features based on their capabilities of differentiating classes using a T test. Random subsets achieving a predefined accuracy rate are then used to generate a final set of features. The method was used in this work with wavelet transform with LDA and kNN classifiers. Although the final accuracy rate obtained in the experiments are relatively low, the improvement when combining classifiers is highly encouraging. This suggests the investigation of various combinations of classifiers as future work. Replacing Wavelet by other multiscale transforms could as well improve the overall accuracy rate.

3. CONCLUSION

Mammography is a low-dose x-ray procedure that allows visualization of the internal structure of the breast, for the detection of early signs of breast caner such as masses and calcifications. Masses are characterized through density, shape, and type of margin. The typical sign of an invasive breast is an irregular or speculated density. This paper presents a survey of image processing techniques to detect masses from mammogram images.

REFERENCES

[1] KeirBovis and Sameer Singh, “Detection of Masses in Mammograms Using Texture Features”, Proceedings of the International Conference on Pattern Recognition (ICPR'00), 2000, IEEE.

[2] Cristina Juarez L., VolodymyrPonomaryov, J. Luis Sanchez R, “Detection of Microcalcifications in Digital Mammograms Images UsingWavelet Transform”, Proceedings of the Electronics, Robotics and Automotive Mechanics Conference (CERMA'06), IEEE, 2006.

[3] Arianna Mencattini, Giulia Rabottino, Marcello Salmeri, and Roberto Lojacono, “An Iris Detector for Tumoral Masses Identification in Mammograms”, International Workshop on Medical Measurements and Applications Cetraro, Italy, IEEE, May 29-30, 2009, pp 215-218.

[4] C. H. Chen’ and G. G. Lee, “A Multiresolution wavelet analysis of Digital Mammograms”, Proceedings of the 1996 International Conference on Pattern Recognition (ICPR '96), pp 710-714.

[5] I. Christoyianni, E. Dermatas and G. Kokkinakis, “FAST DETECTION OF MASSES IN DIGITIZEDMAMMOGRAMS”, 0-7803-5041-3/99, 1999, IEEE, pp 2355-2358.

[6] Robert P. Velthuizen, “Computer Diagnosis of Mammographic Masses”, Proceedings of the 29th Applied Imagery Pattern Recognition Workshop (AIPR.00), IEEE, 2000, pp 166-172.

[7] Tiffany Tweed and Serge Miguet, “ Automatic Detection of Regions of Interest in Mammographies Based on a Combined Analysis of Texture and Histogram”, Proceedings of the 16 th International Conference on Pattern Recognition (ICPR’02), IEEE, 2002.

[8] Aboul Ella Hassanien and Jafar M. Ali, “Digital Mammogram Segmentation Algorithm Using Pulse Coupled Neural Networks”, Proceedings of the Third International Conference on Image and Graphics (ICIG’04), IEEE, 2004.

[9] Chuin-Mu Wang and Sheng-Chih Yang, Pau-Choo Chung,” Comparative Evaluation of classifiers and Feature Selection Methods for Mass Screening in Digitized Mammograms”, IEEE, 2006.

[10] RanaJahanbin, MehulSampat, Gautam S. Muralidhar, Gary J. Whitman, Alan C. Bovik, Mia K. Markey, “Automated Region of Interest Detection of Spiculated Masses on Digital Mammograms”, SSIAI,pp 129-132, IEEE, 2008.

[11] M. Hanmandlu, D. Vineel, Vamsi K. Madasu, S. Vasikarla, “Computer aided detection of suspicious masses and micro-calcifications”, ITNG, pp 789–794, IEEE, 2008.

[12] Arianna Mencattini, Giulia Rabottino, Marcello Salmeri, Roberto Lojacono,andEmanueleColini, “Breast Mass Segmentation in MammographicImages by an Effective Region GrowingAlgorithm”, ACIVS, LNCS 5259, pp. 948–957, Springer-Verlag Berlin Heidelberg, 2008.

[13] Valliappan Raman & Patrick Then, Putra Sumari, “Digital Mammogram Tumor Preprocessing Segmentation Feature Extraction and Classification”, World Congress on Computer Science and Information Engineering, CSIE, IEEE, 2009.

[14] Hussein Samma, CheePeng Lim, Ali Samma, “A Computer-Aided Detection System for Automatic Mammography Mass Identification”, ICONIP, Part II, LNCS 6444, pp. 226–233, Springer-Verlag Berlin Heidelberg, 2010.

[15] Dong Wang, JinchangRen, Jianmin Jiang, Stan S. Ipson, “Applying Feature Selection for Effective Classification of MicrocalcificationClusters in Mammograms”, 10th IEEE International Conference on Computer and Information Technology (CIT 2010), pp 1384-1387, IEEE.

[16] Yu-Shun Cho Chiun-Li Chin Kun-Ching Wang, “Based on Fuzzy Linear Discriminant Analysis for Breast Cancer Mammography Analysis”, Conference on Technologies and Applications of Artificial Intelligence, pp 57- 61, IEEE, 2011.

[17] Daniela Giordano, IsaakKavasidis, and ConcettoSpampinato, “Adaptive Local Contrast Enhancement Combined with 2D Discrete Wavelet Transform for Mammographic Mass Detection andClassification” DICTAP, Part I, CCIS 166, pp. 209–218, Springer-Verlag Berlin Heidelberg, 2011.

[18] Xiaoming Liu, XinXu, Jun Liu, and J. Tang, “Mass Classification with Level Set Segmentation and Shape Analysis for Breast Cancer Diagnosis Using Mammography”, ICIC 2011, LNAI 6839, pp. 630–637, Springer-Verlag Berlin Heidelberg, 2012.

[19] Ibrahima Faye, “A random feature selection method for classification of Mammogram images”, Third International Conference on Intelligent Systems Modelling and Simulation, ISMS, pp 330-333, IEEE, 2012.

AUTHOR

Lothe Savita A., has received M.Sc. degree in Computer Science (2003) from Dr. BabasahebAmbedkarMarathwada University, Aurangabad and M.Phil. (2008) from Y.C.M.O.U. Maharashtra. From 2003 she is working as a Lecturer in VasantraoNaikMahavidyalya, Aurangabad. Presently she is Research Student in Dr. B.A.M. University. Her fields of interests include Image Processing, Bio-medical image processing, Microprocessor, Embedded Systems.

Dr. Prapti D. Deshmukh-Patil, M.Sc. Physics (Elect), M.Sc. (Comp Sci.), Ph.D. presently working as a Principal in MGM Dr. G. Y. Pathrikar College of Computer Science and IT, Aurangabad, MS-India. She is associated with MGM since 16 years imparting teaching instructions to various undergraduate and post graduate courses in computer science and Information Technology. She has published several research papers in national and international conference and journals. Seven students are doing their Ph. D in Computer Science in various areas of multi-biometric and biomedical image processing under her guidance. She is a Fellow of IETE, Life member of CSI, Life member of ISCA. Her research interests are Biometrics, Advanced Pattern Recognition, Neural Network, and Advanced Image Processing. She is recipient of best poster presentation award at ISCA 2013 and Amazing idea award in IEEE Science and information conference, London, UK. She is reviewer and editor of several journals.