Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 7, July 2021: 8050 - 8058

Research Article

Suspicious Mass Detection Algorithms In Mammograms

Pravin Palkar[1] and Pankaj Agrawal[2]
[1]Electronics and Communication Department, RTM Nagpur University, India
E mail: p_palkar@yahoo.com
[2]Electronics and Communication Department, RTM Nagpur University, India
E mail: pankaz_agr@rediffmail.com

Abstract
In the recent years the death rate due to breast cancer among women has increased significantly and now it is a recognized world health problem. Early detection and treatment can reduce the death rate of breast cancer effectively. Presently numbers of imaging techniques are available for detection of breast cancer. Mammography test is the most efficient and reliable to find the breast cancer early. But finding and detecting breast cancer on mammogram is repetitive, tiring and fatigue obligation to radiologist; hence sometimes it may be overlooked. Therefore, smart Computer-Aided Detection system require to be extended and combined in new way in order to provide automatic detection of suspicious mass that meets the needs of medical application to point out the occurrence of breast cancer. Suspicious mass detection accuracy can be improved, which will assist the radiologist to classify the breast cancer. This paper presents algorithms to detect the suspicious mass in mammogram image, and also extract GLCM features of suspicious mass. These extracted GLCM features are graphically represented and based on the variation of these features; the mammogram is analyzed and classified as malignant and non-malignant.

Keywords:
Mammography, GLCM features, Graphical Representation, Malignant and Non-Malignant.

1. INTRODUCTION

The cancer causes cells of the body to change and grow randomly out of control [1]. Mostly the cancerous cells are tumor and the cancer is named after the part of the body where the tumor originates [2]. If the tumor originates in breast tissues, then it is a breast cancer. In the recent decade, breast cancer death rate has increased significantly among women and now it has become one of the recognized world health problems and also become one of the leading causes of fatality. One out of 12 women is affected approximately by breast cancer in their life time. In India also, breast cancer is reported common now in women after lung cancer with a one in eight fatality rate [3]. Breast cancer cases patients and the number of people killed is rising faster all over the world. To minimize the fatality rate, early detection and treatment is the only effective solution so that it can be treated successfully. Presently, many imaging techniques are available to detect breast cancer [4] and mammography is one of the most effective and efficient way to detect suspicious abnormalities [5] recommended for breast cancer screening. But however, for detecting suspicious mass, mammography is not a completely suitable and perfect way. For radiologist, detection of suspicious abnormalities on mammograms is repetitive, tiring and fatigue obligation; hence sometimes may be

overlooked. But for radiologist, detecting suspicious masses on mammograms is a repetitive, tiring and fatigue obligation; hence sometimes it may be overlooked. In recent decades, number of Computer-Aided Detection system are developed to support and assist radiologist for detection of suspicious abnormalities on mammograms. Such CAD system can act as support system but the final and judgmental call has to take by radiologist only. Scientific studies and research shown that such system have improved the detection accuracy of breast cancer [6]. CAD systems use the digitized mammogram image. The computer software application searches the abnormal mass which may signify the occurrence of breast cancer. Further the classification of suspicious mass is done using classification methods as either malignant or non-malignant. Benign i.e., non-malignant is non-cancerous and malignant means cancerous. The non-malignant tumor may grow but do not spread and destroy other parts of the body but the malignant tumor is danger, can invade and harm nearby tissues. Also, malignant tumor can spread and destroy other parts of the body [7], [8].

2. PROPOSED ALGORITHM FOR BREAST CANCER DETECTION

In our proposed breast cancer detection algorithms, MIAS's database images are used as input digital mammogram image. The MIAS [9] database includes both right and left direction mammograms as shown in Fig. 1.

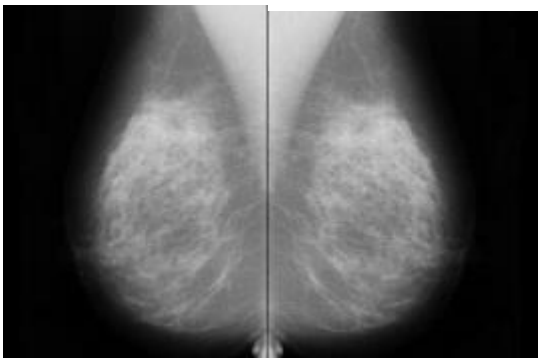

Fig.1. Left direction mammogram and Right direction Mammogram

For simplifying our processing, only right direction mammograms are considered. If encounter any left direction mammogram, swap it to right direction by flipping them in horizontal. The input mammogram is divided in two equal sections vertically and gray level threshold limit of both the section is calculated. The threshold limit of the section, which has breast region in it, obviously is greater than the other section. The algorithm for flipping input mammogram image.

Algorithm: Breast direction detection and flipping Input Mammogram Image

1. Initialize Mammogram A ← Input Mammogram Image
2. Divide in to two halve vertically AL and AR
3. Set grey level Threshold = Th
4. If Th(AR) > Th(AL)
5. B ← A
6. Else
7. B ← flipdim(A)
8. Output Mammogram Image ← B
9. End

After that we have to delete unwanted labels, signs and noises from the mammogram image. Median filtering is performed on the input mammogram image to reduce noise. The small regions are erased by using beareaopen MATLAB function. Then to minimize distortion and to eliminate isolated points, morphological operation is performed. After getting the binary image, it is used as mask and multiplied with original image matrix to get gray scale image without any noise and labels.

Algorithm: Mammogram Preprocessing

1. Initialize Mammogram A ← Input Mammogram Image
2. B ←Resize (A,512,512)
3. C ←ones(512x512)*32
4. B ← Medfilt2(B)
5. D ←gt(B,C); Returns 1 if element of B > 32 else 0
6. D bware ← bwareaopen(D,50)
7. D morph ← bwmorph(D bware,'clean')
8. E ← B.*d morph
9. Output Mammogram Image ← E
10. End

Then following process is implemented to detect and remove the pectoral muscle of the breast and to remove background redundant parts of the mammogram. As depicted in Fig. 2, mammogram is traced initially from left to find first non-zero column. It is marked as Line AB and portion of mammogram before the line AB is cropped. Then at the top margin of mammogram, middle point C is determined and Line CD is traced from middle point C to lower-left corner point D. The vertical cut Line AB is crossed by Line CD at point E and resulted in inverted right angled triangle ACE. This triangle ACE is cropped from the breast portion of mammogram. In this manner detection and removal of pectoral muscle is performed [10]. Then, the mammogram is traced from right side to find the first non-zero column from right side of the mammogram. It is marked as Line FG and the portion of mammogram after the Line FG is cropped. In this manner, unwanted pectoral muscle and unwanted background are cropped and removed effectively [11]. The above process is executed on sample mammograms of MIAS database and as expectancy, the pectoral muscle and redundant background are detected and removed with good success ratio [12], [13]. The algorithm for this method is Pectoral Muscle Detection and Elimination and removal of background redundant parts.

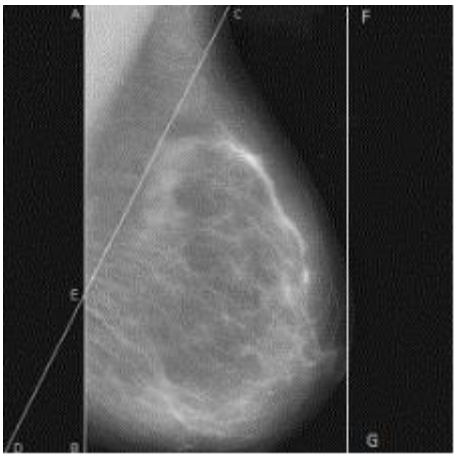

Fig.2. Tracings to detect pectoral muscle and background on mammogram

Algorithm: Pectoral Muscle Detection and Elimination and removal of background redundant parts.

1. Initialize Mammogram E ← Input Preprocessed Mammogram Image
2. j ← first nonzero column of E;
3. Line ab ← E(: , j)
4. Line cd ← E(1 , 256) to E(512,1)
5. Point e ← Intersection of ab and cd
6. i ← first nonzero column from right hand side of matrix E
7. mask image ← matrix containing 0s for triangle aec, columns before j and after i and 1s for all other elements
8. F ← E.*mask image
9. ROI Image ← F(: , j:i)
10. End

After appropriate preprocessing, segmentation is the key step in interpreting ROI i.e., breast contour of the preprocessed mammogram image. The SRG segmentation to detect suspicious mass algorithm determine the seed point as highest intensity pixel in the breast contour region, to get threshold value for finding the spicules. The false spicules are filtered by performing morphological operation. After obtaining the intended spicules, the seed point is detected and dynamically grown based on the region growing segmentation technique [13]. In this way, the suspicious mass region is automatically detected in digital mammograms as shown in Fig. 3.

Algorithm: SRG Segmentation to detect suspicious mass.

1. Initialize Mammogram A ← Input ROI
2. TH ← highest intensity present in A(m , n) - 5
3. B ← ones(m , n)* TH
4. D ← gt(A,B)
5. D bware ← bwareaopen(D,30)
6. (x,y) ← seed point coordinate of D bware grey image
7. J ← regiongrowing(A,x,y,0.03)**
8. [1R1C2R2] ← region grow(J) ##
9. Suspicious mass ← imcrop(A,[C1 R1 C2 R2])
10. End

**Here region is grown iteratively by comparing all unallocated pixels. The measure of similarity is the difference between intensity value of a pixel and mean of the region. The smallest difference measured pixel is allocated to the region. The process continues till the intensity difference between mean and new pixel is smaller than a certain threshold. And process stops, if it becomes larger than the threshold [14].

Calculate non-zero row and column coordinates as (C1, R1) and (C2, R2) for final region grown around seed point of ROI image.

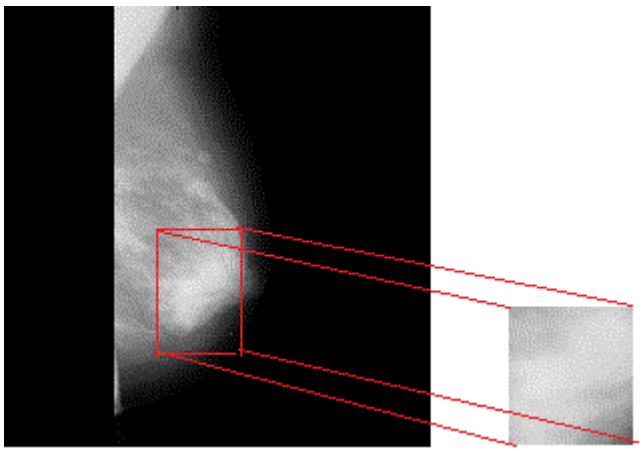

Fig.3. Automatic suspicious mass detection

3. GLCM FEATURES EXTRACTION AND GRAPHICAL REPRESENTATION

To facilitate the decision-making process of the mammogram as benign or malignant, GLCM features are extracted and are represented graphically. The characteristics of mammogram tissues are random and non-homogeneous structure. Hence statistical methods can be preferred to analyze the mammogram [15]. Statistical features are categorized as first-order, second-order and higher-order statistics. The first-order statistics estimate properties like variance, average, standard deviations. It do not consider pixel neighbor relationship. The second-order statistics estimate consider relationships between groups of two neighbor pixels. Higher-order statistics consider relationships between three or more pixels and are only theoretically possible but commonly not implemented due to lot of calculation time and difficulty in interpretation. The statistical methods are quantitative measures of the intensity's arrangements in a region and easy to compute from the intensities themselves. Also, statistical methods are computationally efficient and can be effectively used for texture classification [16]. Hence these statistical features can play very important role in classification of mammogram abnormalities [17]. The statistical features are very useful in differentiating the mammograms abnormalities as malignant and non-malignant and it has been proven [18]. Hence, we can use second-order statistical features to classify the mammogram abnormalities as malignant and non-malignant.

In our research work, second-order statistical features of GLCM (Grey Level Co-occurrence Matrix) are extracted. The GLCM features provide the information about texture of the mammogram abnormalities by considering the relationship between groups of two pixels in the suspicious mass with respect to offset value. The GLCM is a tabular representation of different combination of gray levels of pixels occurs in an image. GLCM considers the relation between two pixels (reference pixel and neighbor pixel) at a time. The distance between these two pixels is called offset (d). The relation between these pixels may be in any one of eight directions (0°, 45°, 90°, 135° and 180°, 225°, 270°, 315°). Both set are opposite to each other and if we select any one set of directions 0°, 45°, 90°, 135° or 180°, 225°, 270°, 315°, all pixels of image can be covered. Also, at any single direction or at single offset, insufficient information about the mammogram abnormalities might be provided, hence all four directions 0°, 45°, 90°, 135° with variation of offset from one to ten are considered for classifying the mammogram abnormalities as malignant or non-malignant. The GLCM matrices of second order statistical texture measures such as contrast, energy, homogeneity and correlation are considered to classify the mammogram abnormalities. The contrast feature measures the amount of local variation in the mammogram abnormalities. It is difference between the highest and lowest values of a contagious set of pixels. The energy feature is referred as uniformity i.e. pixel pair

repetition. It will be high for the mammogram abnormalities, which has very similar pixels. The homogeneity feature takes larger values for smaller gray intensity difference in pair element. It will be high for the mammogram abnormalities, which has uniform gray level intensities. The correlation feature measures the joint probability occurrence of the specified pixel pairs. It measures the linear dependency of gray levels on those of neighboring pixels. The feature correlation between pixels means that there is a predictable and linear relationship between the two neighboring pixels within the window. It will be high for the mammogram abnormalities, which has high predictability of pixel relationships [19]. In our research, these four second order statistical GLCM parameters are extracted from the mammogram abnormalities and are plotted for different direction with respect to different offset as shown in Fig. 4.

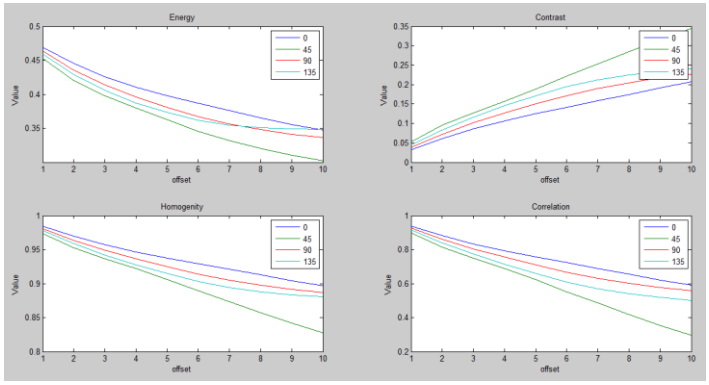

Fig.4. Graphical representation of GLCM statistical features (Values Vs Offset distance) for different direction

Based on the values of features in different directions with respect to offset distance in the graphical representation of statistical features, the mammogram abnormalities are classified as malignant and non-malignant. As per our expectation, the variations in values of statistical features of mammogram abnormalities are dissimilar for malignant and non-malignant mammograms and these differences are easily identifiable in graphical representation. It is observed that the graphical representation of GLCM features plotted with the help of algorithms is effective in discrimination of suspicious mass in mammogram as malignant or non-malignant [13], [20].

4. RESULT AND DISCUSSION

We have adopted the conventional and modified methods for image enhancement and preprocessing of digital mammogram, which suitably removes the noises, labels and other redundant parts and pectoral muscle from the digital mammogram. We have implemented adaptive region growing segmentation algorithm, which is automatically detecting the suspicious mass of variable shape and size from the digitized mammogram, hence is preferred and found more proper over the static/fixed segmentation. The extracted GLCM statistical features of these suspicious mass for malignant and non-malignant cases are different. Hence this adaptive region growing segmentation algorithm improves the mammogram classification performance. Now, by using features extraction algorithm, we can easily transform this suspicious mass data of the mammogram into the set of features, which can facilitate decision making process. This will make mammogram classification simple and easy for the radiologist. Our approach does not require second set of images for deciding whether it is suspicious or not as oppose to methods proposed in [21]. Also, our approach doesn't require database for training, to get the results as proposed in [22], [23]. Hence our proposed method comparatively is fast, easy to use and simple for radiologist to classify and interpret the suspicious mass as malignant or non-malignant at early stage, so that complex biopsy can be avoided. And

necessary diagnosis treatment can be done at early stage. The implemented graphical representation and classification method is a new and unique approach, which easily discriminating the suspicious masses as benign and malignant. Also, our proposed algorithms can detect and analyze multiple suspicious masses for critical case of breast cancer, where multiple cancerous masses are possible.

REFERENCES

[1] Karthikeyan Ganesan, U. Rajendra Acharya, Chua Kuang Chua, Lim Choo Min, K. Thomas Abraham, and Kwan- Hoong Ng., "Computer-Aided Breast Cancer Detection Using Mammograms: A Review", Methodological Review IEEE Re-Views in Biomedical Engineering, Vol. 6, 2013.

[2] Indra K Maitra, Sanjay Nag, Prof. Samir K Banyopadhyay, "Identification of Abnormal Masses in Digital Mammography Images", International Journal of Computer Graphics, Vol. 2, No. 1, May, 2011.

[3] Sara Dehghani, Mashallah Abbasi Dezfooli, "A Method for Improve Preprocessing Images Mammography", International Journal of Information and Education Technology, Vol. 1, No. 1, pp. 90-93, 2011.

[4] U. Bick, F. Diekmann. Berlin, "Digital mammography", Springer, 2009.

[5] H. Gholam, "Registration of vibro-acoustography images and x ray mammography", in Proceeding IEEE, 2005.

[6] D. Narain Ponraj, M. Evangelin Jenifer, "A Survey on the Pre- processing Techniques of Mammogram for the Detection of Breast Cancer", Journal of Emerging Trends in Computing and Information Sciences, Vol. 2, No. 12, December 2011.

[7] National Cancer institute, https://www.cancer.gov/publications/dictionaries/cancer-terms/def/benign, Accessed on 01/01/2021.

[8] IGI Global Publisher of Timely Knowledge, https://www.igi-global.com/dictionary/internet-use-health-decision-making/17739, Accessed on 01/01/2021.

[9] Suckling, J., Parker, J., "The Mammographic Images Analysis Society Digital Mammogram Database", in Proceedings of 2nd International Workshop on Digital Mammography, UK, pp. 375378, 1994.

[10] Samir Kumar Bandyopadhyay, "Pre-processing of Mammogram Images", International Journal of Engineering Science and Technology, Vol. 2, Issue 11, pp 6753-6758, 2010.

[11] Pravin Palkar, Pankaj Agrawal, "Preprocessing of Mammograms for Computer-Aided Diagnosis of Breast Cancer", International Journal of Modern Computer Science (IJMCS), Vol. 4, Issue 5, ISSN: 2320-7868 (Online), October, 2016.

[12] L. Sheng, "The analysis of digital mammograms: speculated tumor detection and normal mammogram characterization", Ph.D. Thesis, Purdue University, School of Electrical and Computer Engineering, 1999.

[13] Pravin Palkar, Pankaj Agrawal, "Analysis of Digital Mammogram Based on Statistical Parameter's Direction Verses Offset Plot", International Journal of Latest Engineering Science (IJLES), Vol. 2, Issue 1, 2019.

[14] Dirk, Jan Kroom, "Segmentation by growing a region from seed point using intensity mean measure algorithm", https://in.mathworks.com/matlabcentral/fileexchange/19084-region-growing, Accessed on 01/03/2018.

[15] Pradeep N., Girisha H., Sreepathi B., and Karibasappa K., "Feature Extraction of Mammograms", International Journal of Bioinformatics Research, Vol. 4, Issue 1, pp.-241-244, 2012.

[16] Pravin Palkar, Pankaj Agrawal, "A Technique to Extract Statistical Parameters of Digital Mammogram to Detect Breast Cancer", International Journal of Advanced Research in Science, Engineering and Technology Vol. 3, Issue 12, 2016.

[17] Pravin Palkar, Pankaj Agrawal, "An Approach to Analyse Digital Mammogram Using Statistical Parameters", International Journal for Research in Applied Science and Engineering Technology, Vol. 5, Issue 1, 2017.

[18] Pravin Palkar, Pankaj Agrawal, "ROI Detection in Digital Mammograms for Detection of Breast Cancer", International Journal of Engineering Science and Computing, February 2017, Vol. 7 Issue No. 2.

[19] https://www.researchgate.net/publication/246325910_GLCM_texture_a_tutorial?pfb Accessed on 01/03/2018.

[20] Palkar P., Agrawal P., " A Benign and Malignant Mass Classification Based on Second -Order Statistical Parameters at Different Offset", in Proceedings of International Conference on Recent Advancement on Computer and Communication. Lecture Notes in Networks and Systems, Springer Vol. 34, pp. 143-150, 2018.

[21] Bandyopadhyay S. K., I. K. Maitra, and T. Kim, "Identification of Abnormal Masses in Digital Mammography Images", in Proceedings of International Conference on Ubiquitous Computing and Multimedia Applications (UCMA), Vol. 2, pp. 35-41, IEEE 2011.

[22] Jyoti Deshmukh and Udhav Bhosle, "GLCM based Improved Mammogram Classification using Associative Classifier", I.J. Image, Graphics and Signal Processing, Vol. 7, pp. 66-74, 2017.

[23] S. Julian Savari Antony, "Graph Based Density Evaluation for Mammographic Image Classification Based on Intensity Values", International Journal of Advancements in Research and Technology, Vol. 6, Issue 3, 2017.