Analysis of Digital Mammogram Based on Statistical Parameter's Direction Verses Offset Plot

Pravin Palkar[#1], Dr. Pankaj Agrawal[*2]

[#]Research Scholar & RTMNU Nagpur, (M.S.) India

Abstract — *The objective of this paper is to analyze the digital mammogram to classify the masses on digitized screening mammograms as benign or malignant with statistical parameters plot. Many CAD tools were developed over the past two decades to help radiologists in detecting and diagnosing breast cancer. The proposed method removes and deletes unwanted signs present in the background of the mammogram and applies enhancement process to eliminate noise and extract the breast region. Then segmentation phase is performed for automatic mass detection. After the mass is detected, second order texture features from Gray Level Co-occurrence Matrix (GLCM) are extracted. These extracted parameters for different offset are plotted and based on the nature of direction's plots and relation between the curves of different directions; the mass is classified as benign or malignant.*

Keywords — *Breast cancer, Mammogram, Statistical parameters, Segmentation, GLCM, Mass classification.*

I. INTRODUCTION

Cancer is a disease that causes cells in the body to change and grow out of control. It is uncontrolled multiplication of a group of cells in a particular location of the body [1]. Mostly the cancer cells is a tumor and the cancer are named after the part of the body where the tumor originates [2]. Breast cancer begins in breast tissue. In the recent years the incidence of breast cancer has increased significantly and deaths occurred due to it is a recognized world health problem. It is one of the leading causes of fatality, with approximately 1 out of 12 women affected by the disease during their lifetime. In India, breast cancer is the second most common cancer in females after lung cancers and the death rate of one in eight women has been reported due to breast cancer [3].

The most effective way to reduce deaths caused by breast cancer is to detect and treat it early as there are chances of successful treatment. Currently, many imaging techniques for breast cancer detection are available [4] and the mammogram is the most efficient system to detect abnormal masses [5], recommended for breast cancer screening. Detection of suspicious abnormalities is a repetitive and fatiguing task. For every thousand cases analyzed by a radiologist, only 3 to 4 are cancerous and thus an abnormality may be overlooked. Computer-Aided Detection (CADe) 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. Some studies have shown that CADe systems, when used as an aid, have improved radiologists' accuracy for detecting breast cancer [6].

Computer-aided detection (CAD) systems use a digitized mammographic image that can be obtained from either a conventional film mammogram or a digitally acquired mammogram. The computer software then searches for abnormal areas of density, mass, or calcification that may indicate the presence of cancer. The CAD system highlights these areas on the images by pre processing it and alerts the radiologist to the need for further analysis. The work is divided into three stages. The first stage covers the mammogram pre-processing and detection of the breast region. The second stage uses the output of the first stage to detect suspicious densities (abnormalities) in the breast region [7]. This stage then finds the ROI and the statistical features of ROI. The set of statistical features obtained from second stage specifies some quantifiable property and significant characteristics of an image is used as an input for the third stage. The third stage is the classification stage based on the statistical features obtained from the previous stage. Any suspicious density is classified further as one having a benign or malignant tumour. Feature extraction is the first step in breast cancer detection. This stage helps in the diagnosis of breast cancer. Benign is not cancerous. Benign tumors may grow larger but do not spread to other parts of the body. Malignant is cancerous. Malignant tumors can invade and destroy nearby tissue and spread to other parts of the body.

The objective of this paper to investigate the classification of masses on digitized screening mammograms as benign or malignant with second order statistical parameters. Pre-processing techniques are implemented using MATLAB. It is preliminary stage used in mammogram image enhancement. The method removes and deletes unwanted signs present in the background of the mammogram and to extract the breast area. It suppresses the undue influence from the background of the mammogram. Then an enhancement process is applied to improve appearance and increase the contrast of images and to eliminate noise. Once the breast region has been found, a segmentation phase is performed for detection of various types of masses in mammograms [9]. Then second order texture features from Gray Level Co-occurrence Matrix (GLCM) of mass are extracted. These extracted features play very important role in classifying masses of mammograms as benign or malignant.

II. PREPROCESSING AND MASS DETECTION

In our proposed method, we are using MIAS's database images [9]. This database includes both right and left. To make processing easier we put whole the breast image to the left side of the image. Then we deleted the unwanted signs, labels and noises of the image. For deleting these, we act as follow; first we performed median filtering. It is mainly used to reduce or eliminate noise present in the image. Median filter not only reduces the noise but also preserves useful information in the image [10]. A median filter is a nonlinear filter and is efficient in removing salt and pepper noise median tends to keep the sharpness of image edges while removing noise [11]. Then compared median filtered image with threshold intensity value of 32 to get processed digital mammogram image by doing following process [12]. From the binary objects in the mammogram image, the regions with fewer than 50 pixels are removed in each image by using the *beware open* function in MATLAB. This process morphologically opens a binary image and removes all objects in the binary image, except the largest object (breast profile). This operation uses an 8-connected neighbourhood. Next, a morphological operation to reduce distortion and remove isolated pixels (individual 1's surrounded by 0's) is applied to the binary images using the *bwmorph* function in MATLAB with parameter 'clean'.

Above steps checks all pixels in a binary image and sets a pixel to 1 if five or more pixels in its 3-by-3 neighbourhood are 1's, otherwise, it sets the pixel to 0. Upon getting binary image after removing all redundant parts, we used it as a mask and multiplied it with original gray scale image matrix to get gray scale image free from noises and labels. After this, we removed the background part of the image. These redundant parts do not cover breast region and are the right and left part of the image. Now it is also important to detect the pectoral muscle and defines the region of interest (ROI) for further analysis. The pectoral muscle, slightly brighter compared to the rest of the breast tissue as shown in Fig 1, can appear in the mammogram.

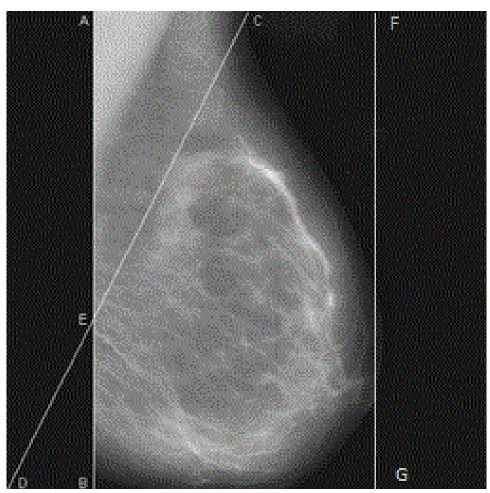

Fig 1. Mammogram with pectoral muscle and background

For removal of redundant image parts and for detecting and removing the pectoral muscle of the breast, we act as follow; At first, to get a straight line cut (AB), we traced image from left side and found first non-zero column between the left background of the mammogram and starting of actual part of breast image. In second step, is to determine middle point (C) at the top margin of the mammogram and plot a straight line (CD) from the middle point (C) to lower left corner point of the mammogram. The line CD cross the line AB at point E resulting an inverted right angle triangle (ACE) that is our region of interest (ROI) to detect the pectoral tissues from mammogram. Except the ROI rest part of mammogram is converted to background colour to make said ROI prominent. Finally to reduce the computational complexity inverted right angle triangle (ACE) is cropped from the original mammogram shown in above figure for further processing. Now the question will arise that in all the cases they said triangle will cover the entire region of pectoral muscle [13] or not.

Next, to get straight line FG, we traced image from right side and found first non-zero column. In this way we successfully removed unwanted pectoral muscle and left hand, right hand black stripe from the image. This method gives absolute success ratio on 80 percent on different pairs of mammogram of different shapes and size [14]. The mammogram carries the information in the form of different intensities and textural variations. Our approach is to isolate the spatially interconnected structures in the image to form regions concentrated around prominent intensities.

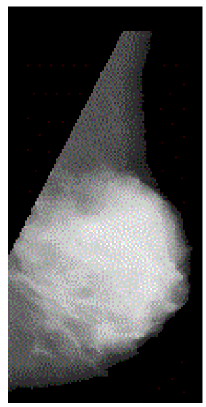

Fig 2. Mammogram without pectoral muscle and background

Hence after getting interested breast contour as shown in above fig 2, we found brightest pixel to get threshold value for getting spicules. For removing some false spicules (fewer pixels) we filtered it out by doing morphological operation. Once getting intended spicules, we identified seed point and grown that dynamically to get mass region as shown in Fig 3.

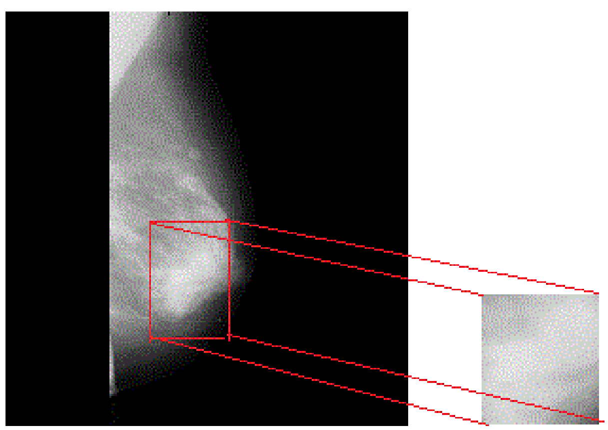

Fig 3. Automatic Detection of Mass Region in Digital Mammograms

In the figure we observe that abnormal areas of density, mass, or calcification is easily detected and with respect to the size of mass, size of ROI is defined dynamically. This defined dynamic ROI of digital mammogram is used to extract the statistical features to detect the breast cancer.

III. FEATURE EXTRACTION

The feature specifies some quantifiable property of an image and is computed such that it quantifies some significant characteristics of the image. Transforming the input data of the image into the set of features is called features extraction. Feature extraction is a method of capturing visual content of an image. The objective of feature extraction process is to represent the image in the form to facilitate decision making process.

The mammogram tissues are characterized by random and no homogeneous structures. For this reason the analysis by the statistical methods will be preferentially used [15].Statistical approach texture is a quantitative measure of the arrangement of intensities in a region. It is easier to compute and is used more often in practice. Statistics that describe a texture can be computed from the gray tones (or colors) themselves. This approach is less intuitive, but is computationally efficient and can work well for classification of textures. The extraction methods of texture feature play very important role in detecting abnormalities of mammograms because of the nature of mammograms. Texture features have been proven to be useful in differentiating masses of breast tissues as benign and malignant.

In our experimental work, the statistical features are extracted using gray level co-occurrence matrices (GLCM).These features provide information about the texture of an image. The GLCM matrices are constructed at a distance of d = 1 to 10 and for direction of θ given as 0°, 45°, 90° and 135°. A single direction and at single distance might not give enough and reliable texture information. Hence four directions and different offset are used to extract the texture information for each masses and non-masses. The texture descriptors derived from GLCM are contrast, energy, homogeneity and correlation of gray level values. The contrast measures the amount of local variations present in an image, while energy is the sum of squared elements in GLCM. Energy may also be referred as uniformity or the angular second moment. The homogeneity measures the closeness of the distribution of elements in GLCM to the GLCM diagonal. The correlation measures the joint probability occurrence of the specified pixel pairs[16].

IV. EXPERIMENT RESULTS

After extraction of features, the graphical representation of second order statistical features with respect to different offset are very interesting. The observation of variation of each features with different direction θ given as 0°, 45°, 90° and 135° at different offsets are very useful to classify the mass.

Based on the MIAS database and resultant images derived from preprocessing and mass detection stage, the detected mass are tested with the range of values of contrast, homogeneity, energy and correlation with respect to different direction at d=1 for benign and malignant and are shown in Table I and Table II respectively.

Direction (Degree)	Statistical Parameters at d=1			
	Energy	Contrast	Homogeneity	Correlation
0	0.4689	0.0323	0.9839	0.938
45	0.4526	0.0525	0.9738	0.8989
90	0.4633	0.0382	0.9809	0.9269
135	0.459	0.0443	0.9779	0.9148

Table I: Statistical Parameters of Benign Mass

Direction (Degree)	Statistical Parameters at d=1			
	Energy	Contrast	Homogeneity	Corelation
0	0.3482	0.065	0.9675	0.9294
45	0.3263	0.0964	0.9518	0.8955
90	0.3516	0.0591	0.9705	0.9362
135	0.3428	0.0716	0.9642	0.9223

Table II: Statistical Parameters of Malignant Mass

The graphical representation of the energy contrast, homogeneity and correlation values for benign and malignant masses are also site in Figure 4 and Figure 5 respectively.

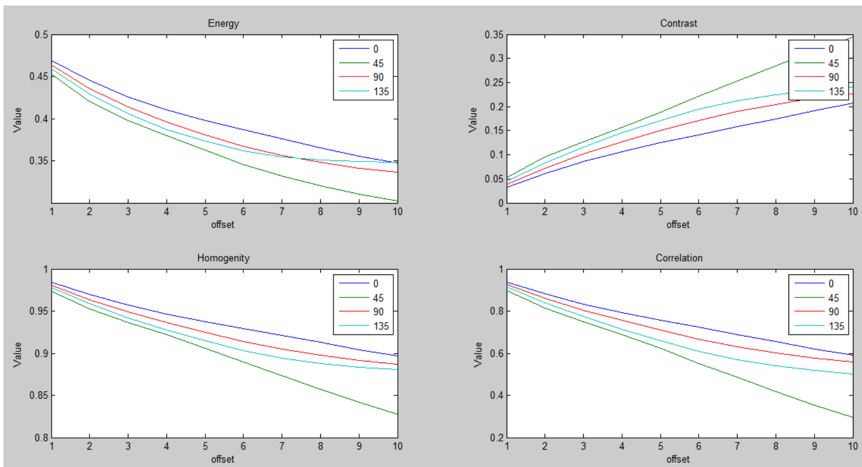

Fig 4. Relation of Second order statistical parameters at different direction for Benign Mass

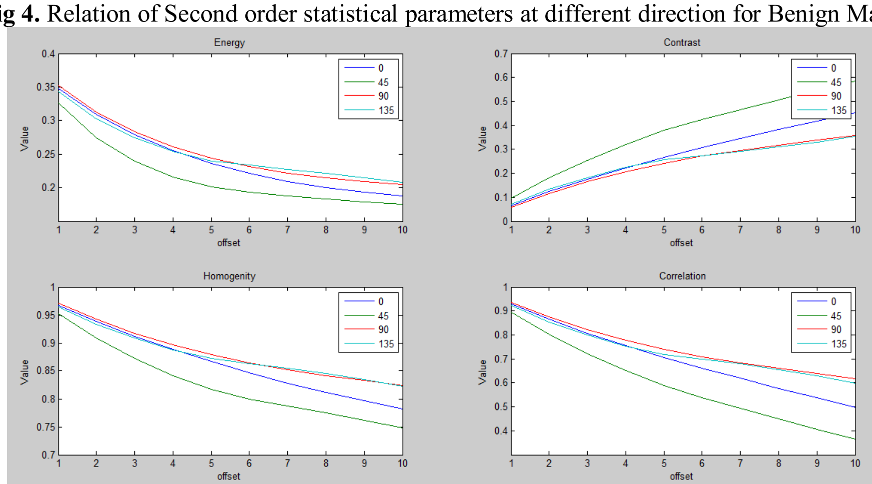

Fig 5. Relation of Second order statistical parameters at different direction for Malignant Mass

It is observed that the variation in values of energy contrast, homogeneity and correlation for image, containing benign and malignant masses are highly different as per expectation and the variation is significantly discriminated. In case of benign, less crossing is there between the plots of directions in all parameters with respect to the malignant case. This has proven the usefulness of the proposed method using these texture descriptors in differentiating the masses as benign and malignant.

V. CONCLUSION

In this paper, we have presented a method to detect the suspicious mass of a digital mammogram and characterize this mass region based on statistical parameter's plot of different direction verses different offset. Investigation of the preliminary results obtained from graphical representation reveals that the variation in mentioned statistical parameters with respect to offset is different for both the cases and clearly distinguishes the mass between benign and malignant. Thereafter, the variation in parameters with different direction at different offset could be used to define criteria of decision that will permit to distinguish a mammogram as benign or malignant.

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 REVIEWS 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, April 2011.

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

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

[6] D.Narain Ponraj, M.Evangelin Jenifer: A Survey on the Preprocessing 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] Sharanya Padmanabhan and Raj i Sundararajan :Enhanced Accuracy of Breast Cancer Detection in Digital Mammograms using Wavelet Analysis , fN 47907, USA 978-1-4673-2322-2112/$31.00, 2012 IEEE

[8] Pradeep N. , Girisha H. , Sreepathi B. And Karibasappa K.,"Feature Extraction Of Mammograms, " International Journal Of Bioinformatics Research ,ISSN: 0975–3087 & E-ISSN: 0975–9115 , Volume 4, Issue 1, 2012, pp.-241-244

[9] Samir Kumar Bandyopadhyay: Detection of Abnormal Masses in Mammogram Images, International Journal of Computer Science and Information Technologies, Vol. 1 (5) , 2010.

[10] J. Sharma, J. K. Rai and R. P. Tewari: Identification of pre-processing technique for enhancement of mammogram images, 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom), Greater Noida, pp. 115-119, 2014.

[11] R. Ramani, Dr. N.Suthanthira Vanitha: The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images, MECS (http://www.mecs-press.org/), 5, 47-54, April 2013.

[12] J. Nagi, S. Abdul Kareem, F. Nagi and S. Khaleel Ahmed: Automated breast profile segmentation for ROI detection using digital mammograms," 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, 2010, pp. 87-92.

[13] Prof. Samir Kumar Bandyopadhyay: Pre-processing of Mammogram Images, International Journal of Engineering Science and Technology, Vol. 2(11), 2010.

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

[15] L. Sheng, "The analysis of digital mammograms : speculated tumor detection and normal mammogram characterisation", PhD. Thesis,Purdue University, 1999

[16] Pravin Palkar,Dr. Pankaj Agrawal: A Benign and Malignant Mass Classification Based on Second-Order Statistical Parameters at Different Offset,Proceedings of International Conference on Recent Advancement on Computer and Communication, Lecture Notes in networks and System, Vol 34, Springer, Singapore, pp 143-150, 2018