

Automatic Cancer Detection in Mammographic Images

Jasmeen Kaur¹, Mandeep Kaur²

Research Scholar, ECED, Punjabi University, Patiala, Punjab, India¹

Assistant Professor, ECED, Punjabi University, Patiala, Punjab, India²

Abstract: Breast cancer is the major health problem in women that has increased over years. Early detection is necessary as to treat breast cancer with success and reduce mortality. Towards this, mammography is the main diagnostic test for screening of breast tumors wherein high quality images with low dose x-rays are used to image the breast(s). Due to the immense danger that this disease presents it becomes necessary to develop a Computer-aided diagnosis (CAD) systems to improve radiologist's diagnosis. This paper has implemented such a system. The developed method proceeds in four steps: pre-processing mammograms for enhancement, segmentation to detect breast abnormalities, then texture feature(s) are extracted and finally image is classified.

Keywords: Breast Cancer, Mammography, image pre-processing, segmentation, feature extraction, image classification.

1. INTRODUCTION

Cancer is a disease that causes cells in the body to change and grow out of control. Most types of cancerous cells eventually form lump or mass called tumor, and are named after the part of the body where the tumor originates like: Non-melanoma skin cancer, Prostate cancer, Lung cancer, Breast cancer, Colorectal cancer, melanoma, Kidney cancer, Bladder cancer, Leukemia and many more. The vast majority of breast cancers begin in the parts of the breast tissue that are made up of glands for milk production called lobules and ducts that connect the lobules to the nipple. The rest of the breast is made up of fatty, lymphatic and connective and tissues. Breast cancer is typically detected either during screening examination before symptoms have developed or after a woman notices a lump [1]. Breast cancer is the second most common cancer in the world and, by far, the most frequent cancer among women with an estimated 1.55 million new cancer cases diagnosed worldwide and 76,000 in India (2016). Breast cancer remains a significant scientific, clinical and a societal challenge. The decline in breast cancer mortality rate is attributed to both improvements in breast cancer treatment and its early detection. Various techniques have been developed for breast cancer detection: Mammography, MRI, Ultrasound, Biopsy. Mammography is most preferred as it uses a low dosage of x-rays and takes a 20 second scan. It is more sensitive in accessing fatty and dense breasts.

This paper presents the attempt towards interpreting mammographic images and checking for the presence of tumorous area(s) and its severity index employing computer-aided detection (CAD). The CAD system puts a mammogram into digital form and then computer software searches for abnormal areas, mass or calcification. The system then highlights suspicious areas, alerting our

radiologist to the need for further analysis. The CAD system can improve detection of cancer in the breast by acting as a second set of eyes to find abnormal areas on a mammogram [2]. Breast Cancer Stage is usually expressed as a number on a scale of 0 through IV with stage 0 describing non-invasive cancer that remain within their original location and stage IV describing invasive cancers that have spread outside the breast to other parts of the body. Image processing has great potential to improve image quality and secondarily diagnostic accuracy and even reduces the radiation dosage necessary to achieve an image of acceptable. Digital mammograms have a wide dynamic range and have the ability to process the image data to provide an opportunity to display data more effectively [3].

2. LITERATURE SURVEY

Sangyun Park et al, in 2007[10] presented segmentation of solid nodules using texture information. To create features reflecting the texture information 1-level 2-D Discrete Wavelet Transform algorithm (DWT) was employed. DWT decomposed the original image into lower and higher sub-bands which included information of both frequency and space domain that aided in extracting features. Optimal threshold from the acquired information is optimized to make a decision rule. Finally, solid nodule region is segmented from the classified texture region.

Hala Al-Shamlan et al, in 2010[8] specified 23 significant breast cancer features extraction. The features that translated medical works in computer words were studied namely: Geometric, Texture and Gradient features. Geometric features described geometric features of ROI like area, perimeter and compactness. Texture features

described smoothness, coarseness and regularity; they represented gray level variations between adjacent pixels in the image. The features included: mean, mean global area, mean local area, uniformity, standard deviation, smoothness, skewness, entropy, correlation and inverse. Gradient image is the derivative of local image values. Sobel operator was used to compute the gradient image like: sobel mean, sobel mean global area, sobel mean local area, sobel uniformity, sobel standard deviation, sobel smoothness, sobel skewness, sobel entropy, sobel correlation and sobel inverse. The features were selected under expert guidance and produced excellent results.

Mussarat Yasmin et al, in 2013 [7] proposed algorithms and overviewed the techniques for detection of breast tumor and for interpreting its stage. Digital mammography technique is widely used for early stage breast cancer diagnosis but due to its negative effects on human body other safe techniques like infrared imaging, MRI, Biopsy are also proposed. The most accurate imaging techniques proved include Mammography and Biopsy.

Approach for Breast Cancer Detection

The proposed model includes the layout of the functioning of the CAD system. It includes capturing a mammographic image. Pre-processing is done to enhance the image. Later it segments the suspicious mass region(s) that might contain malignant mass. Feature extraction grabs the features of the image which include the texture features based on the Gray level scheme. Finally the classification of the mass region is done so as to it is normal or abnormal mass. After distinguishing results are brought upon and discussion is done keeping in account the case study. The proposed model is shown in Fig. 1.1.

3. MAMMOGRAPHY IMAGES

In digital mammography, image acquisition is an important initial step. The choice of sensors and acquisition conditions greatly affect the image quality. The more the precision is achieved, the more the treatment will be reliable and efficient. Mammography is an x-ray of the breasts to detect any abnormalities.

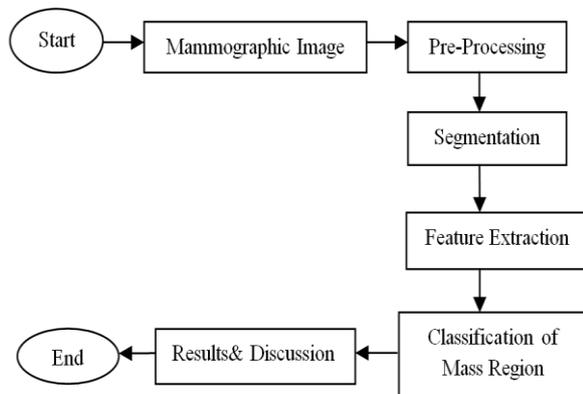


Fig 1.1 Model of Mammography for Cancer Detection

It provides images of interior breast tissue. To get better results several shots of the breast are taken from different angles. The output image is then saved as 2D matrix, where each element represents the X-ray transmission corresponding to a particular path through the breast [3].

4. PREPROCESSING OF MAMMOGRAM

Mammographic images with and masses micro calcifications are usually small and have low contrast thus making the abnormalities hard to be detected. Pre-Processing block involves enhancing image, removal of noises, blood vessels and glandular tissues which become a cause of many False Positives during detection stage.

Fig. 1.2(a) shows a mammogram containing a mass in mediolateral-oblique (MLO) view, and the pre-processing method is described below. The contrast adjustment was first applied to adjust the contrast of the mammogram by linearly scaling the pixel values between upper and lower limits. The pixel values that lie in this range are saturated to the upper or lower limit value, respectively as shown in Fig 1.2(b).

Next, Filtering operation is applied as it transforms pixel intensity values to reveal certain image characteristics. Filters employed improved the: uneven contrast, smoothed the image, removed noise, gray level enhancement, highlights sharp gradients and improves the edges of the mammographic image as shown in Fig. 1.2(c). Because the pectoral muscles represent a brighter region, it can affect the detection process. Thus, the whole foreground is transformed by gamma correction with a decoding gamma to preserve the brighter luminance and suppress the darker luminance. In other words, the gamma expansion enhances the pectoral muscle [4], the enhanced image is as shown in Fig 1.2(d).

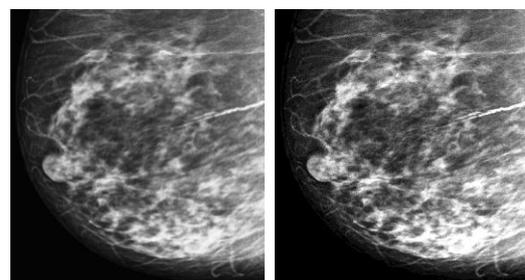
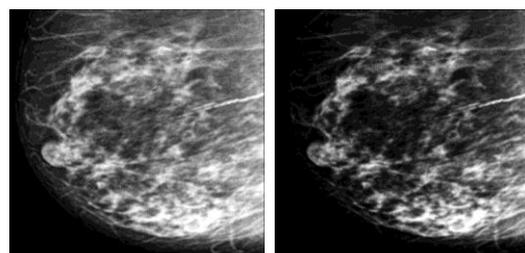


Fig.1.2 (a) Original image (b) Contrast Adjusted



(c) Filtered Image (d) Enhanced Image

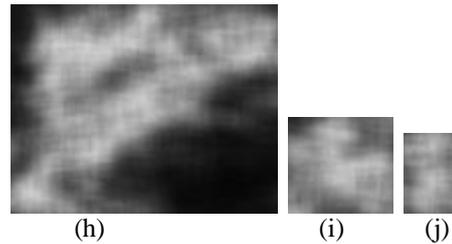
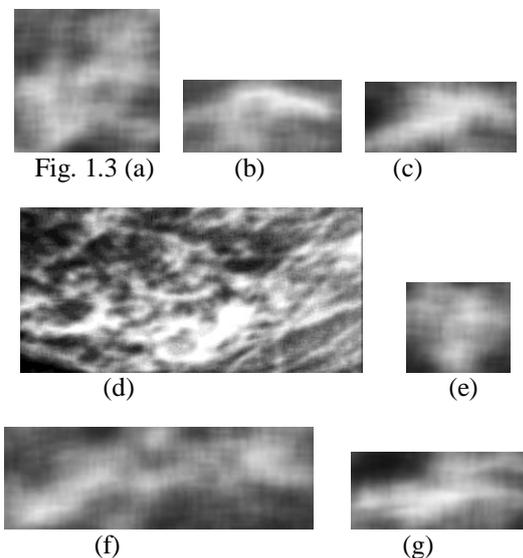
5. SEGMENTATION

This stage uses the output of the first stage to segment the ROIs. Segmentation is the division of the enhanced image Fig 1.2(d) into various non-overlapping regions. It corresponds to the extraction of objects from the background. The segmentation is done to extract locations of suspicious areas to aid and classify the abnormalities as malignant or benign. Segmentation algorithms are based on intensity value, which are discontinuous, based on abrupt changes in the image, as edges and similarity [5].

Thus, depending on the nature of images and the region of interest, the segmentation methods can attempt to detect the ROIs. After the pre-processing of the image we get image in gray format and now we have to segment this enhanced image. According to the images and directions from the radiologist, tumor regions were selected and the regions had varying intensity values. Thus, various morphological operations are applied to extract the required regions as discussed below:

1. Connected Components: Removal of the connected components that have fewer than 50 pixels and produce another BW image. This operation is known as an area opening.
2. Measurement of Properties: After the removal of connected components we are interested in knowing various properties of the regions using 'regionprops' like: Area, Euler Number, Orientation, Bounding Box, Extent, Perimeter, Centroid, Convex Area, Filled Area, Pixel List, Eccentricity etc. Then used the suitable parameters like Area, Centroid and Bounding Box.
3. Segmentation: Rectangular boxes of defined values as specified from above properties are segmented and saved. Thus, segmentation of image(s) is done now further process of tumor severity needs to be unlocked.

The segmented regions from the enhanced image Fig. 1.2(d) are shown in Fig. 1.3(a,b,c...j).



6. FEATURE EXTRACTION AND CLASSIFICATION

In this stage, output from III stage i.e segmented ROIs is used as the input image. Features are extracted from the segmented image by wavelet decomposition method. The decomposition of image using wavelets results in a set of coefficients that acts as features for further analysis. The type of wavelet used here is 2D Haar Discrete Wavelet Transform. The outputs from this wavelet decomposition are the decomposition vector C and bookkeeping matrix S. The decomposition vector C, consists of three detail coefficients: H- horizontal detail coefficient, vertical detail coefficient-V and diagonal detail coefficient-D [6].

Each pixel of the original image is searched for four features:

- Texture Extraction characteristics,
- Threshold Value Determination through the above information,
- Texture Classification and
- Smoothing of classified region and segmentation in the end [7].

The wavelet transformed mammogram is shown in Fig. 1.4.

The four texture features selected and found were: Skewness, Kurtosis, Mean and Standard Deviation [8].

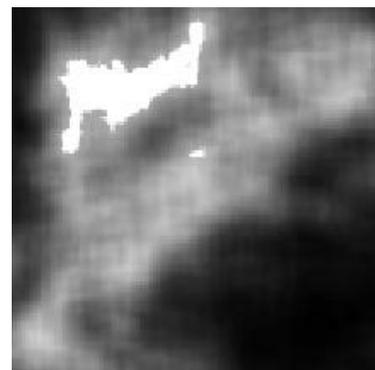


Fig 1.4 Wavelet Transformed Mammogram

The texture features obtained from above step are used as input to the fuzzy logic and reached an output determining the severity index profile (SIP) for the mammogram. The Fuzzy Logic Toolbox™ software is used, as it maps an input space to an output space. The central concept that plays major role is the fuzzy if-then rules. All the rules are evaluated in parallel and order of rules is inconsiderant and the output of each rule is a fuzzy

set. The output fuzzy sets for each rule are then aggregated into a single output fuzzy set. Finally, the resulting set is defuzzified to a single number [9]. Build Mamdani Systems (GUI) has been used.

7. CONCLUSION & FUTURE SCOPE

This paper presented a novel approach to enhance the mammograms and then extracted the ROIs more efficiently. 2D Discrete Wavelet transform is presented to extract texture features for classification purpose. The severity index profile (SIP) was displayed employing Fuzzy Logic. The technique discussed in the paper is implemented on a number of mammogram data and better results were seen. Thus, this technique can very well work as second opinion for the radiologist to automatically detect presence of tumors.

Future works may improve the accuracy of the system by employing more features in the presented work to achieve better accuracy.

REFERENCES

- [1] Breast Cancer Facts & Figures 2015-2016, American Cancer Society.
- [2] Narayan, K. Lalit (2008). Computer Aided Design and Manufacturing. New Delhi: Prentice Hall of India.
- [3] ACR–AAPM–SIIM PRACTICE PARAMETER FOR DETERMINANTS OF IMAGE QUALITY IN DIGITAL MAMMOGRAPHY.
- [4] Shen-Chuan Tai, Zih-Siou Chen and Wei-Ting Tsai, “An automatic mass detection system in mammograms based on complex texture features”, IEEE Journal of biomedical and health informatics, Vol. 18, No.2, March 2014.
- [5] Carlos B. Fiallos, Maria G. Pérez, Aura Conci, Víctor H. Andaluz, “Automatic Detection of Injuries in Mammograms Using Image Analysis Techniques”, IEEE 2015.
- [6] Sharanya Padmanabhan and Raj i Sundararajan, “Enhanced Accuracy of Breast Cancer Detection in Digital Mammograms using Wavelet Analysis”, IEEE 2012.
- [7] Mussarat Yasmin, Muhammad Sharif and Sajjad Mohsin, “Survey Paper on Diagnosis of Breast Cancer Using Image Processing Techniques”, Research Journal of Recent Sciences, Vol. 2(10), pp 88-98, October (2013).
- [8] Hala Al-Shamlan and Ali El-Zaart, “Feature Extraction Values for Breast Cancer Mammography Images”, 2010 IEEE.
- [9] Lee, C.-C., "Fuzzy logic in control systems: fuzzy logic controller-parts 1 and 2," IEEE Transactions on Systems, Man, and Cybernetics", Vol. 20, No. 2, pp 404-435, 1990.
- [10] Sangyun Park, Hyoun-Joong Kong, Woo Kyoung Moon, and Hee Chan Kim, “Segmentation of Solid Nodules in Ultrasonographic Breast Image Based on Wavelet Transform”, Proceedings of the 29th Annual International Conference of the IEEE EMBS Cite Internationale, Lyon, France August 23-26, 2007.
- [11] J.Mohanalin, Prof.P.K.Kalra, and Dr.Nirmal Kumar, “Fuzzy Based Micro Calcification Segmentation”, 5th International Conference on Electrical and Computer Engineering ICECE 2008, 20-22 December 2008, Dhaka, Bangladesh.
- [12] Vikrant Bhateja, Shabana Urooj, Ashutosh Pandey, Aimé Lay-Ekuakille and Mukul Misra, “Improvement of Masses Detection in Digital Mammograms employing Non-Linear Filtering”, 2013 IEEE.
- [13] Shibin Wu, Qingsong Zhu, Yuhan Yang, Yaoqin Xie, “Feature and contrast enhancement of mammographic image based on multiscale analysis and morphology”, IEEE International Conference on Information and Automation Yinchuan, China, August 2013.
- [14] Rafael C. Gonzalez & Richard E. Woods, “Digital Image Processing”, 3rd Edition, 2008.