ISSN No. 0976-5697

Volume 1, No. 3, Sept-Oct 2010

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Fuzzy classification model assisted by intensity based approach and segmentation for breast cancer detection and diagnosis

M. A Saleem Durai*
School of Computing Science and Engineering
VIT University
Vellore, India
masaleemdurai@vit.ac.in

N.Ch.Sriman.Narayana.Iyengar
School of Computing Science and Engineering
VIT University
Vellore, India
nchsniyr@vit.ac.in

A.Kannan
Department of Computing Science and Engineering
Anna University
Chennai, India
kannan@annauniv.edu

Abstract: In this paper we propose a method for detection, diagnosis and treatment of breast cancer based on Image processing. Mammograms play a vital role in diagnoses of breast cancer so an intelligent method is needed to differentiate the benign and malignant case of breast cancer from a mammogram. In this work we use an intensity based approach to find whether a mammogram is affected or not using MATLAB. Segmentation is performed to achieve better accuracy, the idea is to partition an image in to smaller images and apply the test condition for each segments. Finally we design a fuzzy classification model to classify the mammogram and to indicate the stage of breast cancer and the appropriate treatment for it

Keywords: Fuzzy classification, Intensity based approach, Segmentation, Mammograms, Breast Cancer.

I. INTRODUCTION

Breast Cancer is one of the most threatening and dreadful diseases in women. In 2009, an estimated 192,370 women were diagnosed with breast cancer. Excluding cancers of the skin, breast cancer is the most common cancer among women in the United States, accounting for more than 1 in 4 cancers diagnosed. An estimated 40,170 women were expected to die from the disease in 2009. Only lung cancer accounts for more cancer deaths in women. There are clear evidences that mammography reduces the risk of dying from breast cancer. After increasing for many years, the percentage of women aged 40 years and above who reported having a mammogram has dropped. According to survey data, 51.2 % of women who are 40 or above had a mammogram within the past year. Modifiable factors associated with an increased risk of breast cancer include postmenopausal obesity, menopausal hormone use, physical inactivity, and alcohol consumption of two or more drinks per day [1] .Breast cancers can be classified by different schemata which include stage (TNM), pathology, grade, receptor status, and the presence or absence of genes as determined by DNA testing. While discussing about stages, stage 0 is a pre-malignant disease, stage 1 to 3 is defined as early cancer and can be cured and stage 4 is advanced/metastatic cancer and incurable.

Mammograms play a vital role in diagnoses of breast cancer there by differentiating the benign and malignant case of breast cancer. Mammography is the best screening method used by the radiologist for the early detection of breast cancer. Benign tumors are not considered cancerous, their cells are close to normal in appearance, they grow slowly, and they do not invade nearby tissues or spread to other parts of the body. Malignant tumors are cancerous they can spread beyond the original tumor to other parts of the body. So it is necessary to develop an intelligent method that helps the radiologist and users to detect malignant cancer at an early stage and diagnose it to prevent from creating harm.

II. RELATED WORK

In this section we review related methods in detection and diagnosis of breast cancer. The research in this area is done for a long time and all the research work depict that an early detection of cancer increases the survival rate and treatment options. Mammography can detect breast tumors at a early stage. It is a special type of x-ray imaging used to create detailed images of the breast and is used to aid in the early detection and diagnosis of breast diseases in women [3]. Maximizer of posterior margin algorithm is used in this segmentation based on texture feature and classifies breast tissue under various categories [12]. A wavelet based image processing technique with MATLAB is used to detect breast cancer [7]. Matrix laboratory(MATLAB) is a high-performance language for technical computing. It integrates computation, visualization and programming. The size of tumor detected by radiologist was calculated using a MATLAB code that reads the number of pixels .The features were tested in combination and is developed by a MATLAB code that divides the input image in to small frames of same size [5]. Content based image retrieval for breast cancer diagnosis is used. It includes machine learning approach for mammogram retrieval where the similarity measure between the lesion mammogram was modeled after expert's observation. This system retrieves similar mammogram from a database and classifies the query mammogram based on retrieved results [16]. A novel soft clustered based direct learning classifier is developed for diagnosing breast cancer that creates soft clusters within a class and learns it using direct calculation of weights [14]. An automatic mass classification of mammograms where designed it includes a preprocessing stage to remove pectoral muscles and to segment region of interest, feature extraction and selection is done through contourlet transform and classification is done through successive enhancement learning [8].

A structured SVM is used to improve the performance of SVM in computer assisted diagnosis for breast cancer. By considering the cluster structure in the training set it determines if each mammographic region is normal or cancerous. It includes preprocessing and feature extraction stage [15]. Mammograms are analyzed using image feature extraction ,for each mammogram of abnormal breast two regions were marked(i) region that corresponds to the development of malignant mass (ii)a region which appeared similar to region one on the same mammogram. Linear classification is used to find patterns in mammogram [11]. Ultrasound images are also used to detect and classify abnormalities of breast cancer. This includes preprocessing stage to enhance the image and to reduce speckle, segmentation ,feature extraction to find a feature set of breast cancer lesions to distinguish between benign and malignant and classification to classify suspicious region[4]. Knowledge based system is used for detection, diagnosis and treatment of cancer. It uses rule based reasoning, case based reasoning, model based reasoning. An intelligent system uses genetic algorithm, artificial neural networks and fuzzy logic. Combination of these techniques with their advantages yielded a better performance [9]. An incremental learning method is used in computer aided diagnosis system of medical images. It performs image analysis, inference and reclassification. Once a medical image is loaded the system automatically extracts image features. It integrates detection and diagnosis system using Computer Aided Diagnosis (CAD) architecture and supports multiple disease and learning of new adaptation knowledge [10]. The features are clustered using fuzzy c-means and the features are selected using the concepts of genetic algorithm. To detect an abnormality the physician performs visual scanning of the images using a CAD architecture and artificial intelligence [13]. An integrated classifier is used in mammogram MR-image for classification of breast cancer and abnormalities using a multistage classifier. This first classifies mammogram in to normal and abnormal cases. For abnormal it determines whether the cancer is benign or malignant and also provides the type of cancer [2].A hybrid system that combines the advantages of fuzzy sets, rough sets and statistical feature extraction technique is used. Fuzzy image processing enhances the image and edges surrounding the region of interest. A grey level co-occurrence matrix is used to extract the features. Rough sets approach and rules are also used, rules are passed to a classifier to test whether they are cancer or non cancer and rough set distance function is used to measure similarity [6].

III. PROPOSED WORK

A. System Architecture

The proposed system architecture consists of the following Phases (i) test mammograms are given as an input to the system and is called as query mammogram. (ii) calculate the Intensity of pixels and resize the mammogram in to 256X256 pixels using segmentation.(iii)apply the diagnosis rules on each segments of mammogram. (iv)display the state as cancerous its stage and necessary treatment or non cancerous. A database is used to store segmented and formulated trained mammograms. This data can be used to get second opinion for the user to confirm the stage and the treatment for breast cancer.

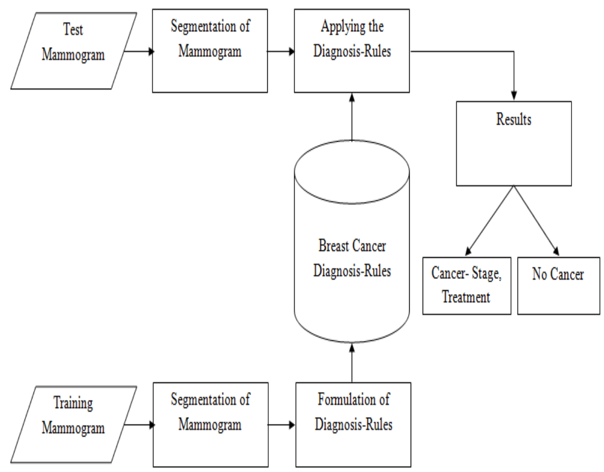

Fig. 1 System architecture

In the proposed system we present an intensity based approach for detecting and diagnosing cancer from a mammogram which has intensity between the range 0 -255. Here 0 indicate black, 1 indicates white and the intermediate values represent grey level. If a mammogram contains a cancerous region then intensity will be high, the normal region in the mammogram will have low intensity. Based on this technique we can easily identify whether the tumor is benign or malignant in the given mammogram. To test whether the given mammogram is affected or not check the intensity of all the pixels and count the number of pixels on a specific region in the mammogram. If the intensity of a pixel is above 140 and if the specific region has more than 100 pixels then we say that the specific region in the given mammogram is affected.

B. Intensity based approach

A snapshot of a mammogram converted in to matrix format is displayed in Fig.2 for 32X32=1024 pixels, it indicates that the affected area in the mammogram has the intensity above 140 and the specific area has more than 100 pixels. This condition is taken as criteria for identifying the suspicious region in the mammogram.

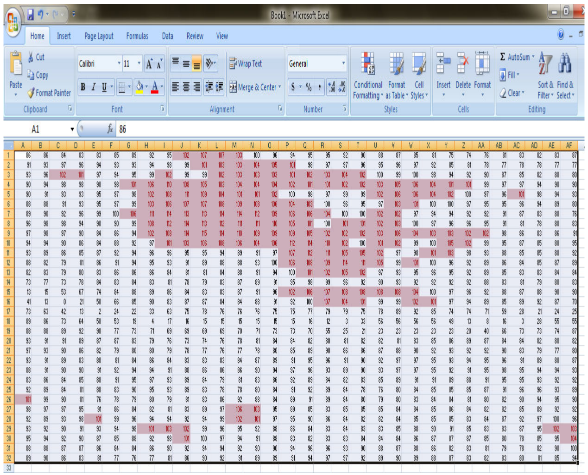

Fig. 2 Pixel intensity measurement

Segmentation is the process of partitioning the given image in to non overlapping region. Any complex problem can be modularized to make it convenient, reliable and easy to interpret, analyze and enhance the image quality. The same technique can be implemented to segment the sample mammogram in to smaller equal size segments. Since images in the MATLAB are in the form of matrix representation we can easily segment the mammograms using operation available in MATLAB. Given a mammogram resize it into a common size of 256X256=65536 pixels, then segment the mammogram in to small segments of 32X32 =1024 pixels. There will be a total of 64 segments and the total number of pixels will be 1024X64=65536 pixels. Segmentation is performed to achieve better accuracy, the idea is to partition an image into smaller image and apply the test condition for each segments. A sample of a mammogram and a part of a segment is displayed in Fig 3.

Fig. 3: Mammogram and a part of mammogram after Segmentation

C. Fuzzy Classification of Tumor in Breast Cancer

While detecting and diagnosing tumor in mammogram we may get unexpected results as the range of intensity may vary and also in our test condition intensity of a pixel in mammogram above 140 is cancerous stage it doesn't mean that intensity 139 of a pixel in a mammogram is non cancerous. To solve this type of problem we make use of fuzzy rule based model to classify and identify the benign and malignant cases in the mammogram.

In our proposed fuzzy model we have membership values between 0 and 1, 0 for non-cancerous, 0.25 for benign, 0.5 and 0.75 for malignant, and 1 for severe case of cancer. In the sample mammogram if the membership value calculated is 0 then the mammogram is not affected and other intermediate values represent the intermediate stages of cancer i.e. mammogram is affected.

1) Calculation of membership function

Based on the intensity level we calculate the membership function for the given range as follows

Table-I Membership function

Input field	Range	Fuzzy Set
Intensity	< 40	Very Low
	40– 80	Low
	81– 120	Medium
	121 – 140	High
	140>	Very high

$$\mu_{\text{very low}}(x) = \begin{cases} 1 & \text{if } x < 36 \\ \frac{40-x}{4} & \text{if } 36 < x < 40 \end{cases}$$

$$\mu_{\text{low}}(x) = \begin{cases} \frac{x-40}{20} & \text{if } 40 \le x < 60 \\ 1 & \text{if } x = 60 \\ \frac{80-x}{20} & \text{if } 60 < x < 80 \end{cases}$$

$$\mu_{\text{medium}}(x) = \begin{cases} \frac{x-80}{18} & \text{if } 80 \le x < 98 \\ 1 & \text{if } x = 98 \\ \frac{120-x}{22} & \text{if } 98 < x < 120 \end{cases}$$

$$\mu_{\text{high}}(x) = \begin{cases} \frac{x-120}{19} & \text{if } 120 \le x < 129 \\ 1 & \text{if } x = 129 \\ \frac{140-x}{11} & \text{if } 129 < x < 140 \end{cases}$$

$$\mu_{\text{very high}}(x) = \begin{cases} \frac{x-140}{14} & \text{if } 140 \le x < 164 \\ 1 & \text{if } x \ge 164 \end{cases}$$

2) Fuzzy Diagnosis-rules

If intensity <40 then n=0 (very low)
Else If intensity ≥40 and intensity ≤80 and pixels≥200 then n=0.25 (low)
Else If intensity >80 and intensity ≤120 and pixels≥100 then n=0.5 (medium)
Else If intensity >120 and intensity <140 and pixels≥50 then n=0.75 (high)
Else If intensity ≥140 and pixels≥30 then n=1 (very high)

3) Classification and necessary treatment

If n=1 then
    Need biopsy
ELSE If n=0.75 || 0.5 then
    Case of Malignant
ELSE If n=0.25 then
    Case of Benign
ELSE
Mammogram is Not – Affected

D. Proposed Algorithm for Breast Cancer Diagnosis

Required input : cancer images
Expected output : classification of image into benign and malignant

Step1. START
Step2. Take a mammogram images and store inside the MATLAB folder
Step3. Read the mammogram from that folder
Step4. Resize the mammogram into 256x256
Step5. Now perform the segmentation on the mammogram
Step6. After segmentation apply the rules on each segments of mammogram
Step7. Identify whether the mammogram is affected or not.
Step8. Apply fuzzy rules for classifying .
Step9. Display the stage and treatment if the mammogram is affected.
Step10.Store the results in database for reference.
Step11:STOP

IV. MAMMOGRAM DATABASE

In our proposed work we make use of the MIAS (Mammographic Image Analysis Society) database. From which we used some images of malignant, benign, and normal. Using these images we have obtained the threshold values for different cases of cancer and non-cancer.

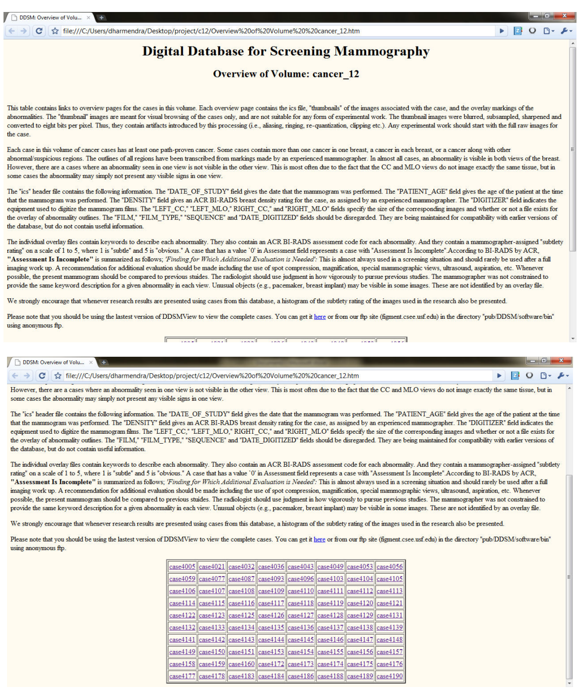

Fig. 4 MIAS database collection

V. RESULTS AND DISCUSSIONS

To evaluate the working of the proposed system the same database is taken in to account. Mammograms are used from MIAS database after using these in our proposed system it yields good results. Here we considered different cases of malignant, benign, and normal. With the help of our proposed system we can identify the cancerous region in the mammogram, and can diagnose the breast cancer in an efficient way at very low cost.

For performance evaluation we considered the following parameters:
Positive: Mammogram is affected
Negative: Mammogram is not affected
True: Correct identification
False: Incorrect identification

Accuracy: Accuracy is used to measure the performance of the technique used in the system. The higher the accuracy the better the performance of the system.
Accuracy = TP+TN/TP+TN+FP+FN

Sensitivity: It is a measure of how well the system correctly identifies a condition. (ie) picking up a disease in a medical test. Sensitivity = TP/TP+FN

Specificity: It is a measure of how well the system correctly identifies the negative cases. (ie) the probability that the test indicates negative if the person does not have a disease.
Specificity = TN/TN+FP

Table II-Parameters Considered

Cancer	Detection	Results
Yes	Yes	TP
No	Yes	FN
No	No	TN
Yes	No	FP

For analysis of the obtained results we are considering a total of 60 mammograms. Out of these 40 are affected and 20 are not affected. We can perform this analysis for case of benign and malignant mammogram also.

Table III-Cases Considered

Mammogram	Affected	Not-affected
Case-1	40	20

Table IV-Results Obtained

Results	Values
TP	38
FP	3
TN	17
FN	2
Accuracy	0.9166
Sensitivity	0.95
Specificity	0.85

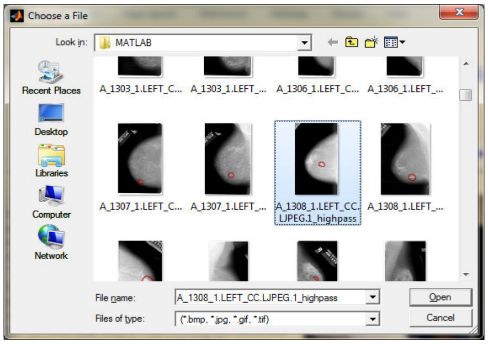

Fig 5 Input mammogram

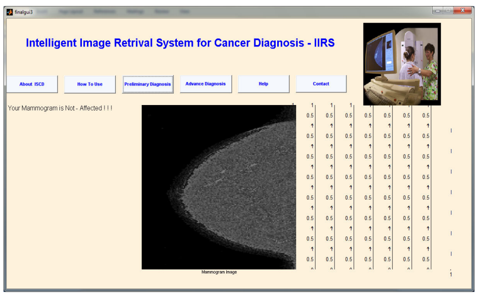

Fig 6 Result of cancer not found in mammogram

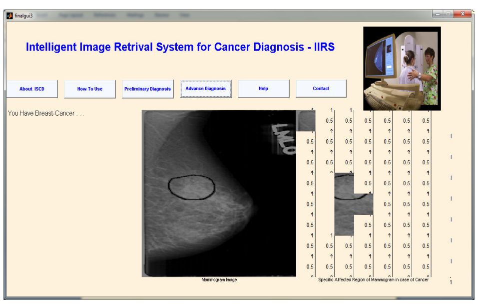

Fig. 7 Results of cancer found in mammogram

VI. CONCLUSION AND RECOMMENDATIONS

In the proposed system we present an intensity based approach for detecting and diagnosing cancer from a mammogram which has intensity between the range 0 -255. Since images in the MATLAB are in the form of matrix representation we easily segmented the mammograms using operation available in MATLAB. Fuzzy classification method is used as an efficient method to detect and diagnose breast cancer. Images used in this approach were obtained from MIAS mammogram database. The designed method uses MATLAB code and when applied to several images and the results are found to be satisfactory and can be used as an assisting model for the radiologist.

In future we plan to further develop and validate the approach using enlarged database, clinical evaluation to improve the performance and accuracy of our system. We can extend our work by including texture analysis, color images to support multiple diseases.

VII .REFERENCES

[1] American cancer society, Breast cancer facts and figures 2009-2010, Atlanta: American cancer society, Inc.

[2] Ardekani Reza Dehestani ,Torabi Meysam and Fatemizadeh Emad , "Breast cancer diagnosis and classification in MR-images using multi-stage classifier". Intl. Conf. on Biomedocal and Pharmaceutical Engineering (ICBPE 2006)

[3] Bhaskar, Jay and Gupta, Poonam Rani, "Design of an automated cancer detection system", Proceedings of ASCNT-2009, CDAC, Noida, India, (2009), pp. 9-15.

[4] Cheng.H.D, Shan,Juan , Ju Wen , Guo Yanhui and Zhang Ling,"Automated breast cancer detection and classification using ultrasonic images: A survey", ELSEVIER Pattern Recognition 43 (2010), pp. 299-317.

[5] El-Sanosi M.D., Habbani A.K., Mustafa N.A., and Hamza A.O, "Computer-Aided Detection of Benign Tumor of the Female Breast". Proceeding of the 2008 IEEE, CIBEC'08.

[6] Hassanien AboulElla , "Fuzzy rough sets hybrid scheme for breast cancer detection",Image and vision computing 25(2007) pp 172-183.

[7] Garge .D.M ,Bapal V.N "A low cost wavelet based mammogram image processing for early detection of breast cancer" Indian Journal of Science and Technology, Vol.2 no 9(2009) pp 63-65.

[8] Moayedi Fatemeh, Azimifar Zohreh, Boostani Reza, Katebi Serajodin, "Contourlet-based mammography mass classification using the SVM family". Computers in Biology and Medicine (Article in Press)

[9] Pandey Babita, and Mishra R.B, "Knowledge and intelligent computing system in medicine". Computers in Biology and Medicine 39 , 215-230, 2009.

[10] Park M., Kang B., Jin S.J., and Luo S,"Computer aided diagnosis system of medical images using incremental learning method". Expert Systems with Applications 36 , 7242-7251, 2009

[11] Sameti Mohammad ,Senior Member, IEEE, Ward Rabab Kreidieh, fellow, IEEE, Morgan-Parkes Jacqueline, and Palcic Branko, "Image Feature Extraction in the Last Screening Mammogram Prior to Detection of Breast Cancer". IEEE Journal of Selected Topics in Signal Processing, Vol. 3, No. 1, February 2009

[12] Sheshadri HS, and Kandaswamy A. "Detection of breast cancer by mammogram image segmentation". Journal Cancer Res Ther- December 2005-Volume 1- Issue 4, www.cancerjournal.net, 2005

[13] Stoitsis John, Valavanis Ioannis, Mougiakakou Stavaoula G., Golemati Spyretta, Nikita Alexandra, and Nikita Konstantina S, "Computer aided diagnosis based on medical image processing and artificial intelligence methods". ScienceDirect Nuclear Instruments & Methods in Physics Research A 569 pp. 591-595, 2006

[14] Verma Brijesh, McLeod Peter, and Klevansky Alan. "Classification of benign and malignant patterns in digital mammograms for the diagnosis of breast mammogram". Expert Systems with Applications 37,pp. 3344-3351, 2010

[15] Wang Defeng, Shi Lin, Heng Pheng Ann "Automatic detection of breast cancers in mammograms using structured support vector machines" Neurocomputing 72(2009) pp 3296-3302.

[16] Yang Yongyi, Wei liyang, and Nishikawa Roberts M. "Microcalcification classification assisted by content-based image retrieval for breast cancer diagnosis". ICIP 2007