DETECTION OF SUSPICIOUS LESIONS IN MAMMOGRAM USING ZEBRA MEDICAL VISION ALGORITHM

Mrs.LAKSHMI.S[1], ELAKKIYA.G[2], SINDHUJA.R[3], SRINIDHI.U[4]

[1]Assistant Professor, Dept. of Electronics & Instrumentation Engineering, Panimalar Engineering College, Chennai, Tamil Nadu

[2,3,4]Students , Dept. of Electronics & Instrumentation Engineering , Panimalar Engineering College, Chennai, Tamil Nadu

---***---

Abstract - *The main objective of this project is to detect breast cancer which is the second most deadly type of cancer among women. The existing system is based on the Fuzzy C-means algorithm. In Fuzzy clustering, data elements can belong to more than one cluster, and associated with each element is a set of membership levels. One of the widely used clustering algorithms is Fuzzy C-Means algorithm. The clustering based techniques are the techniques, which segment the image into clusters having pixels with similar characteristics. The existing system i.e., FCM is less accurate. FCM is not that much fast, robust and easier to understand.FCM still needs more computational time. Zebra Medical Vision Algorithm detects the presence of lesions in the Mammogram. Zebra's mammography algorithm is developed with the aid of thousands of patient studies, aims to optimize BC screening by reducing both false positive results and false negative results. In this project, MATLAB is used as the key tool for identification and classification of the tumor. In the future, this detection technique can be further enhanced by using Artificial Intelligence. By programming the robots with these algorithms the presence of lesions can be detected and removed with the efficiency equal to that of expert radiologists. Our proposed method has achieved greater accuracy than the previous methods.*

***Key Words*: Mammogram, Fuzzy, Clustering, Shock filter, Sobel , Masking , Marker Controlled Watershed Segmentation**

1. INTRODUCTION

The Zebra algorithm provides superior results compared to current tools, reducing misdiagnosis and false alarms. Women over 45 are advised to have a screening mammogram every two years. Approximately 10% of tests are sent for further evaluation due to suspicious findings, and approximately 5 women out of every 1,000 will develop breast cancer. Unfortunately, one of those 5 will be missed, and discovered too late. Furthermore, most women that are sent for biopsy follow ups turn out to be healthy – subjecting them to unnecessary tests and mental anguish. Zebra's new algorithm helps provide better outcomes in two keys ways by reducing both false negatives and false positives. Less false negatives results in accurately detecting women with cancer and fewer false positives means women will not have to undergo unnecessary tests and stressful procedures.

1.1 Breast Cancer

Breast cancer starts in the cells of the breast. A cancerous (Malignant) tumor is a group of cancer cells that can grow into and destroy nearby tissue. It can also spread (metastasize) to other parts of the body. Cells in the breast sometimes change and no longer grow or behave normally. These changes may lead to non-cancerous (Benign) breast conditions such as atypical hyperplasia and cysts. If the cancer is located only in the breast, the 5-year relative survival rate of people with breast cancer is 99%. Sixty-one percent (61%) of cases are diagnosed at this stage. If the cancer has spread to the regional lymph nodes, the 5-year survival rate is 85%. If the cancer has spread to a distant part of the body, the 5-year survival rate is 26%.

1.2 Mammogram

A Mammogram is an x-ray picture of the breast. Mammography (also called Mastography)is the process of using low-energy X-rays (usually around 30 kVp) to examine the human breast for diagnosis and screening. The goal of mammography is the early detection of breast cancer.

2. EXISTING SYSTEM

The existing system is based on the Fuzzy C-means algorithm. In Fuzzy clustering, data elements can belong to more than one cluster, and associated with each element is a set of membership levels. One of the widely used clustering algorithms is fuzzy C-Means algorithm. The clustering based techniques are the techniques, which segment the image into clusters having pixels with similar characteristics. Data clustering is the method that divides the data elements into clusters such that elements in same cluster are more similar to each other than others. In the threshold based segmentation the image is considered as having only two values either black or white. But the bit map image contains 0 to 255 gray scale values. So sometimes it ignores the tumor cells also. In case of the region growing based segmentation it needs more user interaction for the selection of the seed. Seed is nothing but the center of the tumor cells; it may cause intensity in homogeneity problem. And also it will not provide the acceptable result for all the images.

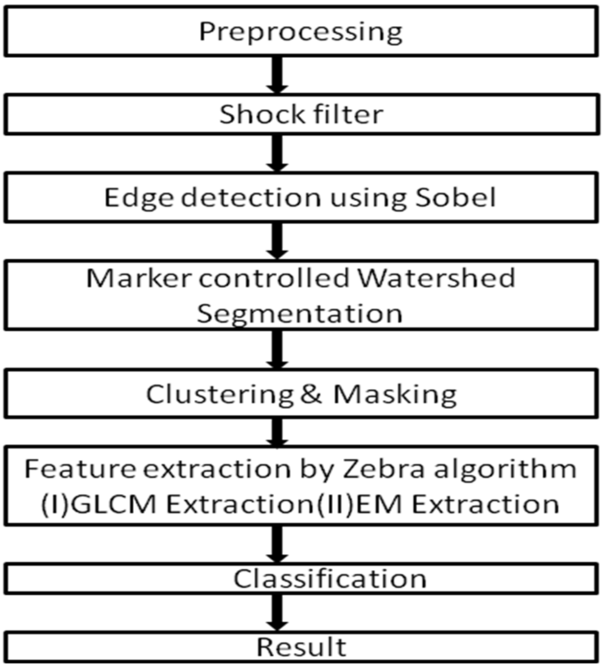
2.1 Limitations Of Existing System

The existing system i.e., FCM is less accurate. It is sensitive to the initial guess (speed, local minima). Also, it is sensitive to noise and one expects low(or even no)membership degree for outliers(noisy points).FCM is not that much fast, robust and easier to understand.FCM still needs more computational time.

3. PROPOSED SYSTEM

The proposed system is based on "Zebra Medical Vision Algorithm". Based on existing system we cannot extract the tumor from the image. This is the main drawback of the existing system. Due to that we go for the proposed method for tumor segmentation. So the proposed system has mainly four modules: Pre processing is done by filtering. Segmentation is carried out by advanced zebra algorithms. This algorithm comes under medical vision learning. This helps to identify the amount of lesions scattered over the body. Feature extraction is by threshold and finally, Approximate reasoning method to recognize the tumor shape and position in MRI image using classification method. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. MATLAB R2014a is used to implement the coding and the output images are obtained from it. MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation.

3.1 Overview Of The Proposed system



3.2 Explanation

1) Pre-processing: According to the need of the next level the pre processing step convert the image. It performs filtering of noise and other artifacts in the image and sharpening the edges in the image. RGB to grey conversion and Reshaping also takes place here. It includes shock filter for noise removal.

2) Shock Filter: Shock filters are based in the idea to apply locally either dilation or erosion process, depending on whether the pixel be-longs to the influence zone of a maximum or a minimum. Shock Filter is applied for preprocessing, in order to improve the contrast, remove the noisy fluctuations and to enhance the edges containing useful information.

3) Sobel Edge Detection: The edge detection techniques are well developed techniques of image processing on their own. The edge based segmentation methods are based on the rapid change of intensity value in an image because a single intensity value does not provide good information about edges. Edge detection techniques locate the edges where either the first derivative of intensity is greater than a particular threshold or the second derivative has zero crossings. Result of these methods is basically a binary image.

4) Marker Controlled Watershed Segmentation: Segmentation is carried out by advanced medical vision learning such as zebra algorithms. The extracted cluster is given to the threshold process. It applies binary mask over the entire image. In the approximate reasoning step the tumor area is calculated using the binarization method.

5) Clustering and Masking: Clustering methods attempt to group together patterns that are similar. This goal is similar to what we are attempting to do when we segment a image. Masking occurs when surrounding breast tissue obscures a cancer. While masking is not a substantial problem in patients with non-dense breasts, mammographic sensitivity is diminished by up to 10-20% in dense breasts. This is a major contributor to the drive for an additional screening modality to be used in conjunction with mammography.

6) GLCM and EM Extraction: Gray Level Co-occurrence Matrix is formulated to obtain statistical texture features. A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels in the image. A common task in image processing is the estimation of parameters of the probability distribution function.

Expectation maximization (EM) is ideally suited to produce maximum likelihood. This algorithm consists of two major steps, first is Expectation step and second is Maximization step. The Expectation step consists of calculating the expected value of the complete data likelihood function. In the Maximization step, first plug in the expected values of the latent variables in the log-likelihood of the given data. Then, maximize this log-likelihood to evaluate the parameters. The main goal of EM algorithm is to facilitate maximum likelihood parameter estimation by introducing the so-called hidden variables which are not observed.

7) Classification: In each segmented part is given to the SVM classifier recognized. We used SVM classifier in order to gain a clear understanding of the relationship between the inputs and the outputs of the models and to facilitate a comparison of the classification performance.

8) Result: From the output images, the area of the lesion is calculated. If the area of the lesion is below or equal to 1000 cm^2, then the result will be displayed as BENIGN cancer. Similarly, if the area of the lesion is above 1000 cm^2, then the result will be displayed as MALIGNANT cancer.

3.3 Output Images

Fig -1: Input grey image

Fig -2: Filtered Image

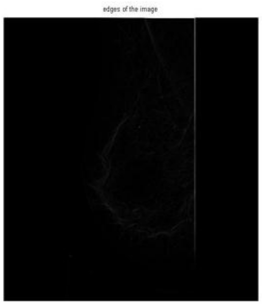


Fig -3: Edge Detection

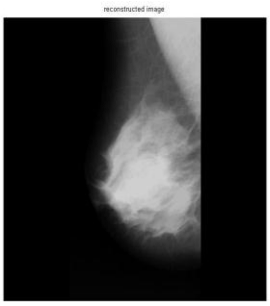


Fig -4: Reconstructed Image

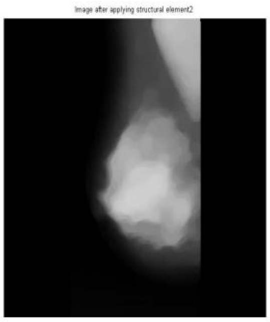


Fig -5: Structural Elemental Image

Fig -6: Complemented Image

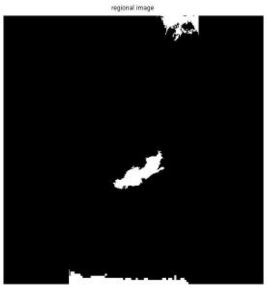


Fig -7: Regional Image

Fig -8: Bordered Image

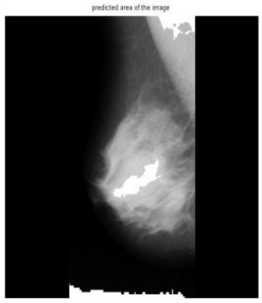


Fig -9: Predicted Area

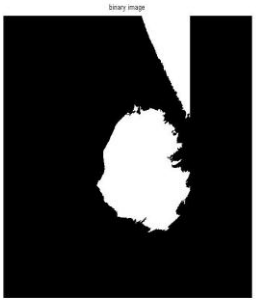


Fig -10: Binary Image

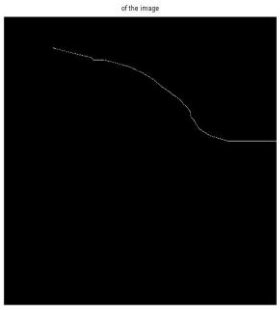


Fig -11: Distance

Fig -12: Magnitude

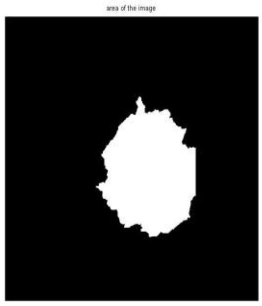


Fig -13: Area

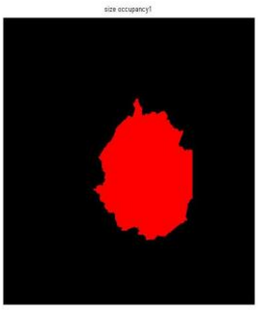


Fig -14: Size Occupany1

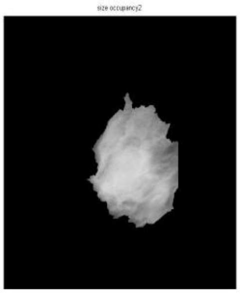


Fig -15: Size Occupancy2

Fig -16: Masked Image

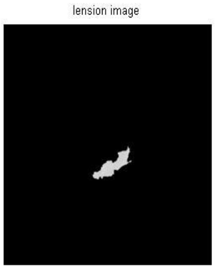


Fig -17: Lesion image

Fig -18: Calculated Area

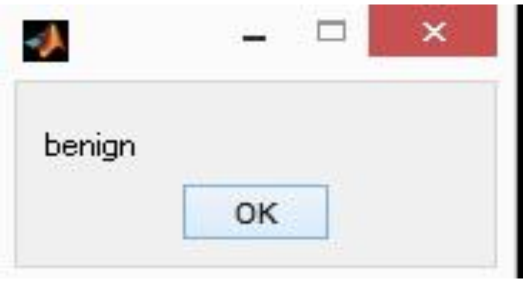


Fig -19: Final Result (Cancer Type)

4. PERFORMANCE ANALYSIS

The performance of the proposed system ie., Zebra Medical Vision algorithm is compared with the proposed system ie., Fuzzy C-Means algorithm. The major parameters taken for comparison are sensitivity, specificity and area of curvature. The sensitivity and specificity are expressed in terms of percentage.

Table -1: Performance Analysis

PARAMETERS	EXISTING METHOD	PROPOSED METHOD
Sensitivity	89.07%	91%
Specificity	76.33%	80%
Area Of Curvature	0.52	0.76

From the table, it is concluded that sensitivity, specificity, area of curvature are higher in the proposed system , when compared to the existing system.

5. FUTURE SCOPE

The most advanced methods that are used for detection of lesions in mammogram are ABC (Advanced Breast Cancer) and Zebra algorithm. Since
Zebra algorithm is much more accurate in prediction of lesions than the other methods, it is used as the proposed method.

In the future, this detection technique can be further enhanced by using Artificial Intelligence. By programming the robots with these algorithms the presence of lesions can be detected and removed with the efficiency equal to that of expert radiologists.

6. CONCLUSIONS

In Zebra medical vision algorithm, the suspicious lesions are detected with high accuracy. This highly accurate result is helpful to identify the type of breast cancer. This algorithm mainly reduces the chances of false prediction.

REFERENCES

[1] Frucci, Maria; Sanniti di Baja, Gabriella (2008). “From Segmentation to Binarization of Gray-level images” . Journal of Pattern Recognition Research.

[2] Jean Cousty, Gilles Bertrand, Laurent Najman, and Michel Couprie. Watershed cuts: thinnings, shortest-path forests and topological watersheds. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[3] Independent UK Panel on Breast Cancer, Screening (17 November 2012). "The benefits and harms of breast cancer screening: an independent review“.

[4] Nass, Sharyl J.; Henderson, I. Craig; Cancer, Institute of Medicine (U.S.). Committee on Technologies for the Early Detection of Breast (2001). Mammography and beyond:developing technologies for the early detection of breast cancer. National Academies Press. pp. 106– .ISBN 978-0-309-07283-0. Retrieved 17 July 2011.

[5] "The Impact of Breast Density on Breast Cancer Risk and Breast Screening". Current Breast Cancer Reports. 4: 161–168. June 2012.