Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 9, September 2014, pg.526 – 532

RESEARCH ARTICLE

AUTOMATED MASS DETECTION SYSTEM IN MAMMOGRAMS

M.MONICA
manoharanmonica8@gmail.com

K.SRIDEVI
sece.sri@gmail.com

PG Scholar, Department of Electronics and Communication Engineering,
Sri Eshwar College Of Engineering, Coimbatore
Assistant Professor, Department of Electronics and Communication Engineering,
Sri Eshwar College Of Engineering, Coimbatore

ABSTRACT: Breast cancer remains as a leading cause of cancer deaths among women in many parts of the world. Cancer is an abnormal, continual multiplying of cells. Most types of cancer cells eventually form a lump or mass called a tumor. In this paper, an automatic CADe system for mammographic mass detection that uses complex texture features is employed. The presence of ambiguous margins of lesions and visual fatigue causes radiologists to miss roughly 10-30% of tumor in breast cancer screening. Texture, the pattern of information or arrangement of the structure found in an image, is an important feature in image processing. It also refers a method for the feature extraction of mammograms using Grey Level Co-occurrence Matrix (GLCM) and Optical Density Co-occurrence Matrix (ODCM) for suspicious areas of breast to identify tumor. Thus the radiologists are able to segment masses on mammogram which are surrounded by the perplexed tissues.

Index Terms — Computer-aided detection (CADe) system, mammographic mass detection, optical density

I. INTRODUCTION

BREAST cancer death rates are higher than that of any other cancers for women. Women in U.S. died from breast cancer in 2011 enormously, although death rates have been decreasing. Early breast cancer is seen in a variety of forms: a few particles of micro calcification, a small ill-defined or speculated mass; abnormal asymmetry between right and left breast images. These abnormalities vary in size, shape, structure, brightness and location and may with normal mammographic appearances.

Mammography is the commonly used imaging modality for breast cancer to enhance the breast radiographers to detect and diagnose cancer at an early stage for high survival rate. The digital mammography with

computer aided detection (CAD) system leads mistrust to radiologists in detection of tumor due to large number of false positive (FP) marks is obtained when there is high sensitivity. When using a CADe system with a mammography, a radiologist still reads the mammogram, but a computer program also evaluates the mammogram and highlights suspicious regions. The goal was to reduce the human error and improve detection accuracy.

II. SURVEY WORK

A. The Cognition Network Technology (CNT) has evolved with automated image analysis for complex scenes. CNT generates automatically a semantic object network from unstructured data such as images and database tables. This enabled the detection of breasts, nipples and calcifications. The nipples were found in all images at the correct location, although the exact position is sometimes difficult to determine even for a human expert. The found malign calcifications in 11 patient cases indicate a true positive rate of 100% (no masses detected yet).

B. Image feature extraction was utilized to analyze screening mammograms taken for the detection of a malignant mass of breast cancer. A discriminate analysis showed that six features could be used to best distinguish between the normal and abnormal regions. Screening mammograms prior to the diagnosis of cancer, there exist differences between the region that subsequently was detected to have a malignant mass, and the other regions of the breast. The future versions of this system will automatically examine the whole mammogram, region by region.

C. An evaluation and comparison of the performance of three different feature extraction methods for classification of normal and abnormal patterns in mammogram. The performance of the each feature extraction method is evaluated on Digital Database for Screening Mammography (DDSM) breast cancer databases

D. The most effective method for the early detection of breast cancer is mammography screening. Many computer aided diagnosis (CADx) systems have been developed as a second opinion to assist radiologists to detect or diagnose abnormalities in mammography screening. Mass and micro calcification are the two most common types of abnormalities associated with breast cancer. We find that mass instances of Old age—small ROI size subsets have much lower classification accuracies than other groups. In future work, we need to further investigate the segmentations and features in this group to improve the classification. Currently, only 61 to 77% of ROI images were successfully segmented by each of weak segmentors.

III. RELATED WORK

This methodology is projected with a technique of a system for detecting the masses of mammograms from the scanned (MRI) image of breast consists of three processing levels: (a) Image acquisition for pre-processing stages. (b) Image segmentation for the localization of regions of interest (ROIs). (c) Characteristics of the each square ROI are extracted based on Grey Level and Optical Density images respectively.

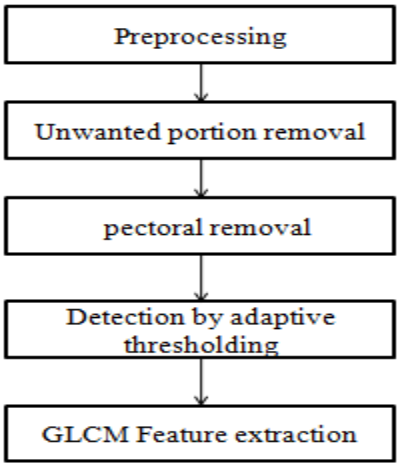


Fig.3.1 Block Diagram

In general, texture refers to surface characteristics and appearance of an object given by the size, shape, density, arrangement, proportion of its elementary parts. Due to the signification of texture information, texture feature extraction is a key function in various image processing applications, remote sensing and content- based image retrieval. The input dataset was separated into a train set and a test set. The train set was used for feature selection and classifier training and the test set for classifier evaluation .A train set contains 100 ROIs that include 50 ROIs of mass and 50 ROIs of normal tissues. These ROIs were selected manually with the various typical of types from the automatic ROI segmentation results of 50 train mammograms with ground truth.

IV. IMAGE PROCESSING

Images play a beneficiary role in fields such as science, medicine, journalism, advertising, design, education and entertainment. Therefore image analysis with the aid of computer becomes more and more substantial in all research fields. Image analysis involves investigation of the image data for a specific application. Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. The goal of this manipulation can be divided into three categories.

- Image processing: image in-> low-level image out
- Image analysis: image in-> measurements out
- Image understandings: image in-> high-level description out

Normally, the raw data of a set of images is analyzed to gain insight into what is happening with the images and how they can be used to extract desired information. The image analysis involves image segmentation, image transformation, pattern classification, and feature extraction.

- ✓ IMAGE SEGMENTATION

  It divides the input image into multiple segments or regions, which show objects or meaningful parts of objects. It segments image into homogeneous regions thus making it easier to analyze them.

- ✓ IMAGE TRANSFORMATION

  It is used to and the spatial frequency information that can be used in the feature extraction step.

- ✓ PATTERN CLASSIFICATION

  It aims to classify data (patterns) based either on a priori knowledge or on statistical information extracted from the image.

- ✓ FEATURE EXTRACTION:

  It is the process of acquiring higher level information of an image, such as color, shape, and texture. Features contain the relevant information of an image and will be used in image processing (e.g. searching, retrieval, storing).

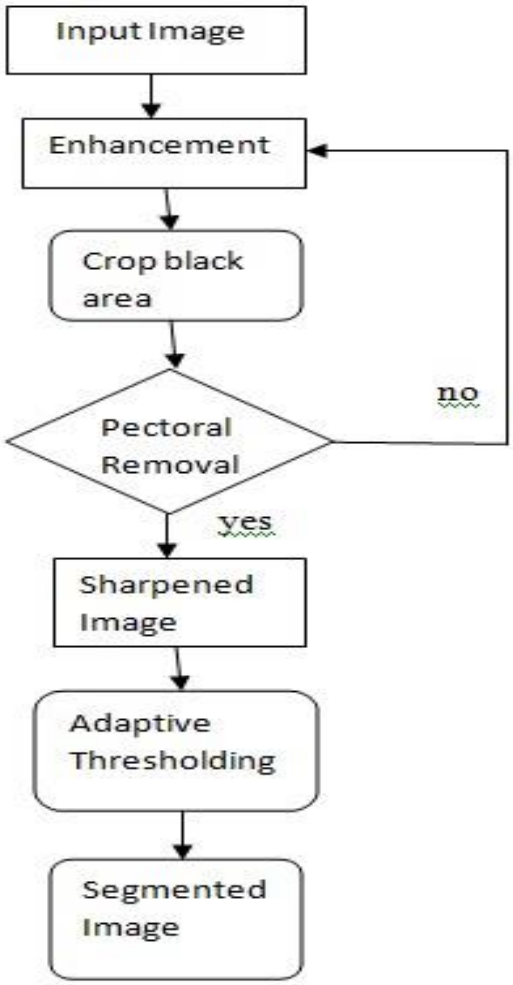


Fig.4.1 Flow Process

V. TYPES OF IMAGE SEGMENTATION

Some types of segmentation:

- Histogram based methods
- Edge detection methods
- Intensity histogram method
- Threshold based segmentation

HISTOGRAM-BASED METHODS:

Histogram based methods are very efficient when compared to other image segmentation methods because they typically require only one pass thought the pixels. In this technique, a histogram is computed from all of the pixels in the image, and the peaks and valleys in the histogram are used to locate the clusters in the image. Color or intensity can be used as the measure.

An image having an object on a contrasting background has a bimodal histogram. The two peaks correspond to the relatively large number of points inside and outside the object. The valley is commonly used to select the threshold gray level. If the image containing the object is noisy and degraded due to illumination artifacts the histogram itself will be noisy and will not be sharp. This can introduce error in selecting the threshold value T. This effect can be

overcome to some extent by smoothing the histogram using either a convolution filter or the curve-fitting procedure. Histogram based thresholding is applied to obtain all possible uniform regions in the image.

Let P_1 and P_2 be the gray value of the peaks of the histogram. The threshold value T is given by

$$T = \frac{P_1 + P_2}{2}$$

Or T may be the gray level at the minimum between the two peaks

$$T = \min_{u \in [P_1, P_2]} H(u)$$

where H(u) is the histogram value at gray level u between P1 and P2

EDGE DETECTION METHODS:

Edge detection is a well-developed field on its own within image processing. Region boundaries and edges are closely related, since there is often a sharp adjustment in the region boundaries. Edge detection techniques have therefore been used to as the base of another segmentation technique. The edges identified by edge detection are often disconnected. To segment an object from an image however, one needs closed region boundaries.

INTENSITY HISTOGRAM FEATURES

Histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit gray scale image, there are 256 intensity values are possible. The intensity histogram features are first order statistics. The histogram is plotted from the image and from the histogram a four features are extracted that can discriminate between the two classes of mammogram.

THRESHOLD BASED SEGMENTATION

Segmentation involves separating an image into regions (or their contours) corresponding to objects. The simplest property that pixels in a region can share is intensity. So, a natural way to segment such regions is through *thresholding*, the separation of light and dark regions. Thresholding creates binary images from grey-level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one.

VI. ADAPTIVE THRESHOLDING

Adaptive thresholding technique is used when images are captured under unknown lightning condition and it is required to segment a lighter foreground object from its background or whenever the background gray level is not constant and object contrast varies within an image. This technique allows the threshold value T to change based on the slowly varying function of position in the image or on local neighboring hood statistics. Threshold T depends on the spatial coordinated (x, y) themselves.

VII. EXTRACTION METHODS

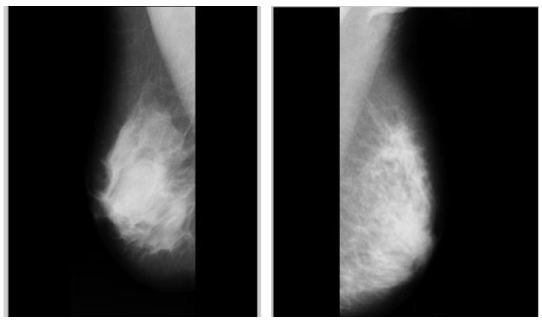
- Gray Level Co-Occurrence Matrix (GLCM)
- Optical Density Co-Occurrence Matrix (ODCM)

GRAY LEVEL AND OPTICAL DENSITY CO-OCCURRENCE MATRIX:

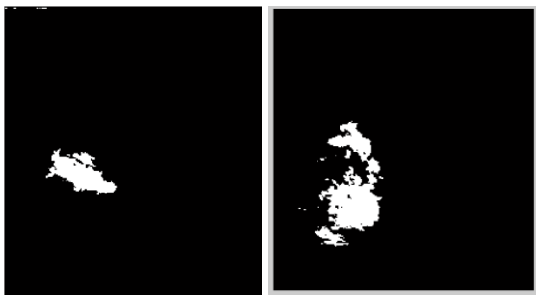
Gray-level co-occurrence matrix (GLCM) is the statistical method of examining the textures that considers the spatial relationship of the pixels. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. The gray co-matrix function in MATLAB creates a gray-level co-occurrence matrix (GLCM) by calculating how often a pixel with the intensity (gray-level) value *i* occurs in a specific spatial relationship to a pixel with the value *j*. By default, the spatial relationship is defined as the pixel of interest and the pixel to its immediate right (horizontally adjacent), but you can specify other spatial relationships between the two pixels. GLCM contains the second-order statistical information of spatial relationship of pixels of an image. From GLCM, many useful textural properties can be calculated to expose details about the image content. However, the calculation of GLCM is very computationally intensive and time consuming.

The ODCM is a co-occurrence matrix of the optical density image. The ODCM is a co-occurrence matrix of the optical density image. An optical density image can be obtained by converting the intensity of the gray-level image into optical density and linearly mapping each optical density value to an image with 8-bit depth information. The minimum optical density value was mapped to 0, and the maximum optical density value was mapped to 255. The intensities were high and close and making it difficult for human eyes to determine the difference in a dense density mammogram

VIII. EXPERIMENTAL RESULTS



(a)



(b)

Fig.9.1 Segmented (a) Input and (b) Output images

IX. CONCLUSION

The system is employed by a method for segmenting masses of breast cancer in mammograms and to detect the abnormalities using the two feature extraction modules combining co-occurrence matrix texture feature (GLCM) and optical density texture features (ODCM).This study preprocess a mammogram to obtain the breast region and suppress the effects of blood vessels, tissues and noises using filters. This method gives pixel correlating with mass and correlation threshold which decides a pixel belongs to suspicious mass or not. Compared to region based

method, the pixel based threshold method identifies the suspicious mass regions. The breast density is based on fatty tissues, fibrous and glandular tissues which expressed better performance by two feature extraction methods in detecting the masses in mammogram. The work projects for the radiologist to diagnose the masses in the breast with dense glandular tissue are difficult to detect lesions than the fatty breasts. The extraction methods can increase the mass detection rate and to reduce the burden of radiologist.

REFERENCES

[1] Günter Schmidt, Alexander Horsch, Rüdiger Schulz, Sukhbansbir Kaur, Matthias Elter, Harald Sittek, Thomas Wittenberg, Maria Athelogo, "Cognition Network Technology for Automated Holistic Analysisin Mammography" Springer-Verlag, Berlin Heidelberg, 2007: 282-287.
[2] Mohammad Sameti, "Image Feature Extraction in the Last Screening Mammograms Prior to Detection of Breast Cancer" VOL. 3, NO. 1, FEBRUARY 2009
[3] R.Nithya, B.Santhi, "comparative study on feature extraction method for breast cancer classification" Vol. 33 No.2, 2011
[4] Nitish Zulpe and Vrushen Pawar," GLCM Textural Features for Breast Tumor Classification", Vol.9, Issue 3, 3 May 2012.
[5] Yu Zhang , Noriko Tomuro, Jacob Furst, Daniela Stan Raicu,"Building an ensemble system for diagnosing masses in mammograms",Vol.7,2012.
[6] T.Romen Singh, Sudipta Roy, Imocha Singh, Tejman Sainam,Manglam Singh," A New Local Adaptive Thresholding Technique in Binarization",Vol.8,Issue 6,No.2,2011.
[7] P. Mohanaiah, P. Sathyanarayana, L.GuruKumar "Image Texture Feature Extraction using GLCM approach" Volume 3, Issue 5, 2013
[8] M. P. Nguyena,, Q. D. Truonga, D. T. Nguyena, T. D. Nguyena, V. D. Nguyena," An Alternative Approach to Reduce Massive False Positives in Mammograms Using Block Variance of Local Coefficients Features and Support Vector Machine" 399 – 405, 2013
[9] Asmat Ara Ainapur, D.A.Kulkarni, G.R.Udupi "Classification of mammograms on zone based statistical features" Volume No.02, Issue No. 07, 2014.

