Segmentation of Masses in Mammographic Images

Himani Deone[1], P. B. Bhalerao[2]

[1]Department of Computer Science and Engineering Marathwada ShikshanPrasarak Mandal's Deogiri Institute of Engineering & Management Studies, Aurangabad Maharashtra state, India 2019-2020

[2]Assistant Professor Department of Computer Science and Engineering Marathwada ShikshanPrasarak Mandal's Deogiri Institute of Engineering & management Studies, Aurangabad Maharashtra state, India 2019-2020

Abstract: ***Image Processing is a process of performing few operations on an image either to obtain an enhanced image or to extract useful information from it. In image processing input provided is an image and output can be image or features associated with particular image. Image processing is one of the most rapidly growing technology in computer science field which also forms the core research area. Image processing techniques are widely used in detection of breast cancers. Segmentation is a process of dividing an image into distinct regions with each pixel having similar properties. Region should strongly relate to the feature of interest to make it more meaningful and useful image analysis. Segmentation is the first step in transforming a greyscale image into one of more other images i.e. from low level image processing to high level image processing. The success of image analysis depends on segmentation but partitioning of an image accurately is a challenging task. This paper presents an approach of automatic breast mass segmentation which has three stages: contour initialization, construction of fuzzy contour and estimation of fuzzy membership maps.***

Keywords: Computer aided detection (CAD), Mammographic Image Analysis Society (MIAS), Region of interest (ROI), Otsu thresholding, Fuzzy contour.

1. Introduction

Image Processing is performed to improve the visual appearance of images to human viewer and to prepare images for measurement of features and structures present. Breast cancer is the most frequent causing cancer in woman, which is also considered as the second most fatal type of cancer after lung cancer. In 2012 there were 1.7 million diagnosed cases which is around 25% of all types of cancer caused in women. Hence detecting breast cancer at early stage is necessary to increase recovery rate and also to decrease mortality rates.Mammography is considered as the most effective and trusted tool to detect breast cancer among women at a very early stage. Mammography is a technique which uses low energy X-rays(around 30 kVp) to study human breast for detecting masses or tumor in it. The image produced after mammography techniques are called Mammograms. Mammograms are basically X-ray picture of human breast. Microcalcification and masses are two most common types of abnormalities in mammograms [1].Based on level of suspicion of abnormality doctors recommend the further tests. Mammographic Mass Interpretation is basically based on various factors such as its shape, density, size etc. Hence accurate Mass Segmentation is a criticalstep in CAD system because it would affect the performance of further analysis steps. Breast Masses are of very low contrast with different sizes and shapes. They also have ambiguous margins which makes to differentiate from surrounding parenchymal. These all factors make Mass Segmentation a Challenging Task. Depending upon its shape masses are of two types Benign and Malignant. Benign tumorareround and oval in shape and are usually non-cancerous. Malignant tumors are partially rounded with spiked or irregular outline. Malignant tumors are cancerous in nature and are whiter than any tissue surrounding it. Mammogram image enhancement is a technique in which mammographic images are manipulated to increase their contrast and decrease the unwanted noise present to help radiologist in detecting the abnormalities. Mammogram image segmentation is the process of dividing homogenous regions into region of interest. Digital mammography is also referred to Full Field digital Mammography (FFDM). The digital mammography has more advantages when compared to film mammography because it has better image quality and also higher spatial resolution.

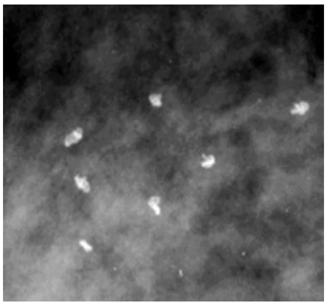


Figure 1: Calcium Deposit

The figure 1 shows the tiny calcium deposits in the nearest which are also referred as micro calcification. Micro calcification are tiny bits of calcium deposits which look like circle or fine lines in the breast tissues. In image processing first step is to detect region of interest. After once the region of interest is detected area is segmented followed by classification process. There are number of mass segmentation techniques used which are region growing, Watershed, Active Contours, Clustering, Cellular Automata.

2. Literature Survey

Breast cancer is the most common cancer caused among women. According to World Health Organization in the year 2012, there were around 1.7 million newly diagnosed cases in breast cancer sector.In year between 2008 and 2012 breast cancer increased by 20% amd mortality also increased bgy 14%. Such tragedy caused researchers to design new and advanced tools to detect and diagnois breast

cancer at an very early stage.Computer Aided Diagnosis system are developed to reduce the workload of the researchersand as well to hep them detect cancer at early stage. These system include four different phases : preprocessing, segmentation, feature extraction, selection. There is an extensive literature on Mass segmentation methods. Thresholding techniques are widely adapted by researchers to detect masses as well as for segmentation.
Thresholding based techniques of mass segmentation are classified into global thresholding and local thresholding. Local thresholding determines a local threshold value for each and every pixel based on intensity values of its neighbouring pixels. Whereas global thresholding uses global information such as histograms. In 2015, T.L.V.N Swetha, CH. Hima Bindu used Otsu thresholding method with 10 different threshold levels. They used Otsu thresholding with Hybrid image segmentation to yeild better result. By using this combination of methods artifacts were eliminated and the breast tumors were detected[3]. Also in 2016, Jonathan Hernandez-Capsitran, Jorge F.Martinez-Carballido reviewed microcalcification segmentation on mammograms in obvious, subtle and cluster categories. The paper uses four various thresholding algorithms and the comparison between them[4].

Region growing based technique starts from initial seed point and it then groups the pixels which have similar properties to divide mammographic image into homogenous regions. Segmentation of images using region growing technique is becoming moe and more popular because of high level knowledge involvment in seed selection process. In 2010, B. Senthilkumar, G. Umammaheshwari proposed a novel region growing segmentation algorithm for detection of breast cancer. They used selective median filter for preprocessing and CLAHE i.e Contrast Limited Adaptive Histogram Equalization technique for image enhancement[5]. In 2012, S. Meenalosini, Dr.J. Janet proposed segmentation technique using region growing method and Gabor features. Alarm regions are generated with region growing method to segment suspicious regions. Later these segmented regions are examined by Gabor filters from different frequency levels and angles[6].

Clustering based techniques creates a set of clusters of mammograms pixels which have similar properties. The only disadvantage of clustering based technique is that they need to set number of clusters manually. In 2015, Jaya Sharma, J.K. Rai proposed a combined water shed segmentation approach using K-means clustering for mammograms. Initial region of interest is gained which is further enhanced by Adaptive histogram equalization. Segmentation is done using watershed technique folloed by K-means clustering which is applied to get foregrounf markers. These markers are fed as input to watershed segmentation step[7]. In 2012, R. Subash Chandra Boss, K. Thangvel proposed mammogram image segmentation using Fuzzy C-means clustering algorithm. Image is pre processed with the help of median filter. Then using gray level co-occurrence matrix 14 Haralick features are extracted. K-means and FCM algorithms are used to segment region of interest for further classification[8].

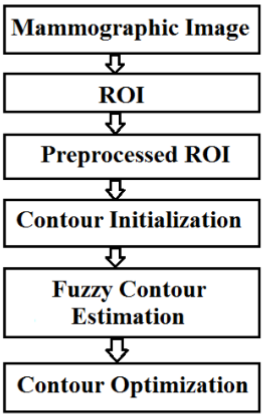
Active contour based techniques include snakes and level sets. The only difference is that these two have different mathematical implementation. The boundary in snakes evolves explicitly whereas in level sets it evolves implicitly. In 2017, Shafilluah Soomro, Kwang Choi proposed robust active contours for image segmentation of mammograms. In the formulated energy function, a characteristic function limits the cotour evolution inwards. Using phase shifted Heaviside function a new SPF function is defined that helps to get optimumsolution in very few number of iteration. Active contour are widely used because they have an advantageous properties like topology adaptability, robustness to initialization etc[9]. In 2013, Ilke Tunali, Erdal Kilic proposed mass segmentation of mammograms using active contours which is used to find benignand malignant masses which are further segmented. ChanVese active contour algorithm with new stopping criteria was implemented after adding low valued pixels. This prevented lacking of segmentation. This proposed method had Area Overlap ratio of 75.1%[10].

3. Database

The mammographic images required to do the study are been taken from MIAS i.e. Mammographic Image Analysis Society database. It is an organization of UK researchers. The MIAS database is reduced to 200 micron pixel edge and is also clipped/padded so that every image is 1024*1024 pixels. It in total contains 322 mammographic images which are true to the given information. The database contains seven columns, each column has specific meaning. First column indicates MIAS database reference number. Second column has a character of background tissue i.e fatty (F), fatty-glandular (G), dense-glandular (D). Third column has different classes of abnormality i.e Calcification (CALC), well-defined (CIRC), spiculated masses (SPIC), ill-defined masses (MISC), architectural distortion (ARCH), asymmetry (ASYM), normal (NORM). Fourth column describes the severity of abnormality i.e benign (B), malignant (M). Fifth and sixth column has x,y co-ordinates of center of abnormality. Finally seventh column has approximate radius of a circle enclosing the abnormality (in pixels).

4. Methodology

The proposed mass segmentation is an hybrid method, which is an combination of edge based active contours, region based active contours and also fuzzy reasoning. Mammographic region of interest consist of three classes : Class Mass, Class Background and Class Contour. Class Mass contains pixels belonging to the mass (tumor)[2]. Class Background contains pixels that don’t belong to the mass, they may belong to breast tissues or other lesions. Class Contour contains pixels that separate mass from background region. The proposed method work flow is shown in following figure.



4.1 Mammographic Image

Mammographic image is taken from MIAS database.

4.2 ROI (Region Of Interest)

ROI from mammographic image is extracted based on information provided by database by selecting bounding rectangles of masses. By doing this we avoid false results that occurs because of automatic ROI extraction method.

4.3 Preprocessed ROI

By enlarging the intensity differences between mass and background ROI is preprocessed to improve the contrast level.ROI is preprocessed to remove all the noises and distortions. Image contrast is enhanced by using a gamma correction factor. Gamma correction factor is a parameter controlling shape of transformation curve. It is a non linear transformation process.

4.4 Contour Initialization

To segment an image active contours require an initial contour to start with. To get initial contour ROI is been preprocessed as explained earlier aand then we use Otsu Thresholding method to binarize the mass region.Otsu thresholding algorithm considers that an image has two different classes. It automatically finds out an optimal threshold value to divide gray scale image into these two classes. And finally binarize image of Class mass is obtained after applying Otsu thresholding technique. Mammographic masses are monobloc in nature and also are solid lesions i.e they don't have holes inside them or noise outside them[6]. Only region that represents mass is been applied with morphological opening. Thus initial contour is been extracted from well processed binary image.

4.5 Fuzzy Contour Estimation

Accurate mass segmentation is a challenging task in mammographic images. There are two main reasons for this difficulty. Uncertainity and Imprecision. Uncertainity refers to noisy enviorment whereas imprecision refers to blurred boundaries and complexity in their shape. The conept of Fuzzy Logic is been introduced to deal with these kind of problems. There are six main steps in estimating a fuzzy contour which are explained below

- Evolve initial contour C1 using edge based active contour in order to obtain C2,
- Locate gravity center of mass i.e. GCM,
- Trace radiating rays from gravity center of mass to contours in direction Θ where $\Theta = 0°$ to 360° by keeping angle of variation = 2°,
- Find intersection points between rays and contours, if ray crosses 2 points on contour we consider the farthest point as center,
- Construction of 2D gaussian memebership function $\mu_k(x,y)$,
- Finally fuzzy contour is obtained.

4.6 Contour Optimization

Chan-vese model is used to get final mass contour. But the major disadvantage of Chan-vese is boundary leaking which occurs because the mass contrast is very poor or the margins are ill defined. To avoid this problem we limit the search within the previously calculated fuzzy contour area. Chan-vese usesexternal energy and to act upon this energy estimation of Fuzzy membership maps of class mass and background is required. To construct the fuzzy membership map of class mass the membership value of pixels located insides contour C2 are assigned to value 1 , membership value of pixels located outside fuzzy contour are assigned to 0 and the membership value of contour is kept less than 0.5. Similarly, to construct fuzzy membership map of class background the membership value of pixels located inside contour C2 are assigned to 0, membership value of pixels located outside fuzzy contour are assigned to 1 and the membershiip value of contour is limited to 0.5[2].

References

[1] T. Nakagawa, T. Hara, H. Fujita, T. Iwasc, T. Endo, K. Horita, Automated contour extraction of mammographic mass shadow using an improved actiive contour model, in:International Congress Series, Vol.1268, Elsevier, 2004.

[2] J. Anitha, J. D. Peter, Mass segmentation in mammograms using a kernel based fuzzy level set method, Internationa;l Journal of Biomedical Engineering and Technology 19,2015.

[3] T. F. Chan, L. A.Vese, Active contours without edges, IEEE transaction on Image Processing 10,2001.

[4] T.L.V.N Swetha, CH. Hima Bindu, Detection of breast cancer with Hybrid image segmentation and Otsu's thresholding, International Conference on Computing and Network Communication 16,2015.

[5] Jonathan Hernandez-Capistran, Jorge F. Martinez-Carballido, Thresholding methods review for microcalcification segmentation on mammography images in obvious, subtle and cluster categories, International Conference on Electrical Engineering, Computing Science And Automatic Control,2016.

[6] B. Senthilkumar, G. Umamaheshwari, J. Karthik, A novel region growing segmentation algorithm for the detection for breast cancer, International Conference on Computational Intelligence and Computing Research, 2010.

[7] S. Meenalosini, Dr. J.Janet, Segmentation of Cancer cells in mammograms using region growing method and Gabor features, International Journal of Engineering Research and Applications, 2012.
[8] Ilke Tunali, Erdal Kilic, Mass Segmentation on mammogramsusing active contours, Signal Processingand Communications Application Conference, 2013.
[9] Shafilluah Soomro, Kwang Nam Choi, Robust Active Contours for mammograms image Segmentation, IEEE International Conference on Image Processing, 2017.
[10] Jaya Sharma, J. K. Rai, A combined watershed segmentation approach using K-means clustering for mammograms, International Conference on Signal Processing and Integrated Networks, 2015.

Author Profile

Himani K. Deone,received the B.E. degree in Computer Science and Engineering from Jawaharlal Nehru Engineering College, Aurangabad, Dr. Babasaheb Ambedkar Marathwada University, Maharashtra, India and pursuing M.Tech in CSE from Deogiri Institute of Engineering & Management Studies, Aurangabad, Maharashtra, India.

Mr. Pramod B. Bhalerao, received the ME degree in Computer Science and Engineering, Aurangabad. Presently he is working as Assistant Professor in Computer Science and Engineering at DIEMS, Aurangabad, Dr.BAMU University, Maharashtra, India