

## A Review of Segmentation of Mammographic Images Based on Breast Density

Anusree Mohan\*, Devesh D Nawgaje\*\*

\* (Department of Electronics and Telecommunication Engineering, S.S.G. M. College of Engineering, Shegaon, India, 444203.)

\*\* (Department of Electronics and Telecommunication Engineering, .S.S.G. M. College of Engineering, Shegaon, India, 444203)

### Abstract

Breast cancer is one of the leading causes of fatality in women. Mammogram is the effectual modality for early detection of breast cancer. Increased mammographic breast density is a moderate independent risk factor for breast cancer, Radiologists have estimated breast density using four broad categories (BI-RADS) swearing on visual assessment of mammograms. The aim of this paper is to review approaches for segmentation of breast regions in mammograms according to breast density. Studies based on density have been undertaken because of the relationship between breast cancer and density. Breast cancer usually occurs in the fibroglandular area of breast tissue, which appears bright on mammograms and is described as breast density. Most of the studies are focused on the classification methods for glandular tissue detection. Others highlighted on the segmentation methods for fibroglandular tissue, while few researchers performed segmentation of the breast anatomical regions based on density. There have also been works on the segmentation of other specific parts of breast regions such as either detection of nipple position, skin-air interface or pectoral muscles. The problems on the evaluation performance of the segmentation results in relation to ground truth are also discussed in this paper.

**Keywords:** Image segmentation, breast density, mammogram, medical image processing, medical imaging.

### I. INTRODUCTION

Breast cancer is the most prevalent cancer and is the leading terminal illness among women worldwide. Early detection of breast cancer is crucial and for that, mammography plays the most essential role as a diagnostic tool. Breast cancer usually occurs in the fibro glandular area of breast tissue. Fibro glandular tissue attenuates x-rays greater than fatty tissue making it appear bright on mammograms. This appearance is described as 'mammographic density' or also known as breast density [1]. The breast density portion contains ducts, lobular elements and fibrous connective tissue of the breast. Breast density is an important factor in the interpretation of a mammogram. The proportion of fatty and fibro glandular tissue of the breast region is evaluated by the radiologist in the interpretation of mammographic images. The result is subjective and varies from one radiologist to another.

In the study conducted by Martin *et al.* [2], hormone therapies, including estrogen and tamoxifen treatments have been found to be able to change mammographic density and alter the risk of breast cancer. Therefore, a method for measuring breast density can provide as a tool for investigating breast cancer risk. Subsequently, the association of breast density with the risk of breast cancer can be more definitive and will allow better monitoring response

of a patient as preventive or interventional treatment of breast cancers.

Breast cancer is the leading cause of death for women in their 40s in the United States. In developing Asian countries, most breast cancer patients are younger than those in developed Asian and Western countries. Younger patients mean that the mammographic images would be denser. In a dense breast, the sensitivity of mammography for early detection of breast cancer is reduced. This may be due to the tell tale signs being embedded in dense tissue, which have similar x-ray attenuation properties. Therefore, it is most appropriate to focus on density based research of mammograms involving younger aged patients having denser breast and thus are difficult to diagnose.

### II. Segmentation of Breast Regions in Mammogram based on Density

Image segmentation means separating the image into similar constituent parts, including identifying and partitioning regions of interests. Segmentation is an important role and also the first vital step in image processing, which must be successfully taken before subsequent tasks such as feature extraction and classification step. This technique is important in breast applications such as localizing suspicious regions, providing objective quantitative assessment and monitoring of the onset and progression of breast

diseases, as well as analysis of anatomical structures. Many researchers had focused on image processing, including segmentation technique to identify masses and calcifications in order to detect early breast cancer. The medical community has realized breast tissue density as an important risk indicator for the growth of breast cancer. Wolfe has noticed that the risk for breast cancer growth is determined by mammography parenchymal patterns [3], and it has also been confirmed by other researchers, such as Boyd *et al.* [4], van Gils *et al.* [5] and Karssemeijer [6]. Before classification or segmentation is performed, a proper understanding of breast anatomical regions is essential

### 2.1 Mammogram and Breast Regions

A mammogram is an x-ray projection of the 3D structures of the breast. It is obtained by compressing the breast between two plates. Mammograms have an inherent "fuzzy" or diffuse appearance compared with other x-rays or Computed Tomography images. This is due to the superimposition of densities from differing breast tissues, and the differential x-ray attenuation characteristics associated with these various tissues. A mammogram contains two different regions: the exposed breast region and the unexposed air-background (non-breast) region. Breast regions can be partitioned into:

1. Near-skin tissue region, which contains uncompressed fatty tissue, positioned at the periphery of the breast, close to the skin-air interface where the breast is poorly compressed.
2. Fatty region, which is composed of fatty tissue that is positioned next to the uncompressed fatty tissues surrounding the denser region of fibroglandular tissue.
3. Glandular regions, which are composed of non uniform breast density tissue with heterogeneous texture that surrounds the hyperdense region of the fibroglandular tissue.
4. Hyperdense region, which is represented by high density portions of the fibroglandular tissue, or can be a tumor.

Breast density is a measurement of the dense structure of fibroglandular tissue, which appears white on a mammogram. Fibroglandular tissues appear to have disc or cone shapes and extend through the interior of the breast from the region near the chest wall to the nipple. The breast density part contains ducts, lobular elements, and fibrous connective tissue of the breast. Fatty tissues are less dense and appear as darker regions. So, if the tumor is in the fatty region, it is easier to be interpreted compared to if it is in the fibroglandular region. Tumors generally appear similar to hyperdense parts compared to their surroundings tissues. The density of dense structures such as the milk ducts is similar to the tumor making it difficult to interpret. It is tedious

to differentiate between normal, dense tissue and cancerous tissue when the tumor is surrounded by glandular tissues. It is difficult to compare the two regions having similar intensities using the naked eyes, but it is possible to do this using computer-aided detection through segmentation.

Wolfe categorized breast density into four patterns. Quantitative classification of breast density into six categories has been developed by Byng *et al.* [7] and Boyd *et al.* [4]. According to Byng *et al.* [7], in the quantification, it is difficult to evaluate a volume of dense tissue because it is highly dependent on the compressed thickness during the mammographic examination and also on the spectrum of the x-ray beams. Optionally, the proportion of the breast area representing dense tissue is used for the quantification of mammographic density. Byng *et al.* [8] performed segmentation using an interactive thresholding technique of the dense tissue. Quantification is then obtained automatically by counting pixels within the regions recognized as the dense tissue. The segmentation using thresholding technique in the study by Byng *et al.* [8] is limited to the cranio-caudal view of the mammogram image. However, for the media-lateral oblique view, the study suggested the option of suppressing the pectoral muscle. Breast Imaging Reporting and Data System (BIRADs), which was developed by the American College of Radiology (ACR) is the recent standard in radiology for categorizing the breast density. BIRADs classify breast density into four major categories: (1) predominantly fat; (2) fat with some fibroglandular tissue; (3) heterogeneously dense; and (4) extremely dense.

### 2.2 Segmentation of Fibroglandular Tissue

Automated segmentation of glandular tissue or parenchymal pattern can provide as the beginning step in mammographic lesion detection. Segmentation of abnormal structures in the breast, consequently, depends on breast tissue density. Segmentation of the glandular tissue can also supply as a primary step in order to detect the suspicious mass and to reduce false positives. Usually, mass is represented by hyperdense structure. Overlapped fibroglandular tissue also has similar intensity with mass. Hence, by focusing on glandular area and highlighting the hyperdense regions of the glandular area, it can assist and contribute as a second opinion for experts in diagnosis. According to Miller & Astley [9], identification of glandular tissue in a mammogram is necessary for assessing asymmetry between the left and right breasts. According to Matsubara *et al.* [10] the assessment of fibroglandular tissue can be used to estimate the degree of risk that the lesions are obscured by normal breast tissue and also to suggest another examination

such as breast ultrasound. The combination of mammogram and ultrasound is effective in depicting breast cancer. Therefore, there is a need to develop a system, which can segment the glandular tissue area automatically.

Ferrari *et al.* [11] segmented the fibroglandular disc with a statistical method based on a Gaussian mixture modeling. Mixtures of up to four weighted Gaussians represent a particular density class in the breast. Grey-level statistics of the pectoral muscles were used to determine the tissue region that represents the fibro-glandular disc. Ols'en & Mukhdoomi used Minimum Cross-Entropy to obtain an optimum threshold for detecting glandular tissue automatically. The idea of Masek [12] is used for fully automated segmentation algorithm extracting the glandular tissue disc from mammograms

### 2.3 Classification of Breast based on Density

There exists numerous classification research based on breast density. Miller and Astley [8] used granulometry and texture energy to classify breast tissue into fatty and glandular breast types. Taylor *et al.* [10] classified fatty and dense breast types using an automated method of extracting the Region of Interest (ROI) based on texture. Karssemeijer [6] used four categories in the classification of the density. Bovis and Singh [11] analysed two different classification methods, which are four-class categories according to the BIRADS system and two-class categories, differentiating between dense and fatty breast types. Sets of classifier outputs are combined using six different classifier combination rules proposed by Kittler *et al.* [12] and the results were compared. The results showed that the classification based on BIRADS system for the four-class categories (average recognition rate, 71.4%) is a challenging task in comparison to the two-class categories (average recognition rate, 96.7%). Zhou *et al.* classified breast density into one of four BIRADS categories according to the characteristic features of gray level histogram. They found that the correlation between computer-estimated percent dense area and radiologist manual segmentation was 0.94 and 0.91 with root-mean-square (RMS) errors at 6.1% and 7.2%, respectively, for CC and MLO views. Matsubara, *et al.* [10] divided breast mammogram images into three regions using variance histogram analysis and discriminant analysis. Then, they classify it into four categories, which are (1) fatty, (2) mammary gland diffuseness, (3) non-uniform high density, and (4) high density, by using the ratios of each of the three regions.

### 2.4 Segmentation of Breast Anatomical Regions

Only a small group of researchers have done segmentation based on breast tissue anatomy. By doing segmentation based on the breast anatomy,

more detailed divisions can be made. For example, with the detection of breast edge, distortion in breast structure and the nipple position in the breast will be detectable. This will also help in diagnosis. The segmentation method proposed by Karssemeijer [6] allowed subdivision of a mammogram into three distinct areas: breast tissue, pectoral muscle and background. For research on segmentation of breast regions into different densities, the suppression of pectoral muscle is not so significant. Instead, pectoral muscle can be used as a reference in estimating the area of glandular tissue. According to Karssemeijer, the density of the pectoral can be used as a reference for interpretation of densities in the breast tissue area, where regions of similar brightness with the pectoral will most likely correspond to fibro-glandular tissue. Saidin *et al.* used graph cut algorithm on mammograms to segment breast regions into the background, skin-air interface, fatty, glandular and pectoral muscle. Aylward *et al.* segmented the breast into five regions using a combination of geometric (Gradient magnitude ridge traversal) and statistical (Gaussian mixture modeling) method [12]. The five regions that they segmented are the background, uncompressed fat, fat, dense tissue and muscle. Most of the work done on segmentation of breast anatomical regions, automatically will detect the fibroglandular disc..

### 2.5 Segmentation of Other Specific Breast Region in Breast Density Research

Most of the density based breast segmentation system involves pre-processing. Image processing technique is usually employed to detect the boundary of the breast region and to remove markers in background area of mammograms. Breast boundary detection (breast contour, breast edge, skin-air interface detection or also called skin-line estimation) is considered as an initial and essential pre-processing step. The purpose is to enable abnormality detection to be limited to the breast area without influenced from the background. By limiting the area to be processed into a specific region in an image, the accuracy and efficiency of segmentation algorithms could be increased. However, failure to detect breast skin-line accurately, could lead to the situation whereby a lesion which is located near to the breast edge may be missed. Usually, research carried out on the segmentation and classifications of glandular tissue based on density would also give rise to the suppression of pectoral muscle in order to avoid incorrect segmentations. Several studies have been conducted on the suppression of pectoral muscle in the segmentation and classification of glandular tissue. In 1998, Karssemeijer proposed an automatic classification of density patterns in mammograms, including a method for automatic segmentation of the pectoral muscle in oblique mammograms, using the

Hough transform. This is due to the fact that in some mammograms, the pectoral muscle has similar intensities with the glandular tissues. Some of the work applied background and annotation subtraction to correctly focus the algorithm on the glandular tissues.

### III. Performance Evaluation

The most essential requirement from a radiologist point of view for image processing algorithms is the ability to achieve enhanced visualizations of anatomical structure, while preserving the detail of the structure. There are numerous researches, which worked on the classification and segmentation of glandular tissues. Each classification and segmentation result needs evaluation of its performance. There are three types of performance evaluations. The first type involves qualitative assessment, the second is quantitative assessment involving the ground truth evaluation, and the third is a statistical evaluation. Performance evaluation for research on classification of breast density involves comparison of research result with density class that has been given by radiologist, while performance evaluation for segmentation of breast density usually is done in qualitative analysis. This is because of the difficulty in obtaining the ground truths from radiologist. For the quantitative analysis, usually the performance of the segmentation results is compared with the ground truth by the radiologist. Ground truth in these density based research means, a correct marking of the glandular tissue or density area by the radiologist in a digital mammogram. For statistical evaluation, Receiver operating characteristic (ROC) analysis is commonly employed to ensure the validity of computer aided diagnosis systems. The ROC analysis allows for a plot of the sensitivity (True Positive Fraction, TPF) against the specificity (False Positive Fraction, FPF). The area under the ROC curve (Az) represents a quantitative measure of the accuracy of the segmentation or classification technique. When the value is 0, it indicates poor segmentation or classification performance while 1 indicates high segmentation or classification performance. However, it has certain restrictions and also suffers from weaknesses. Since, it is a pixel based assessment, for region based analysis, the Free Response Operating Characteristics(FROC) works better.

### IV. Conclusion

The aim of this paper is to review approaches for segmentation of breast regions in mammograms according to breast density. Studies based on density have been undertaken because of the relationship between breast cancer and density. Classification of glandular tissue is beneficial for estimation of breast density for categorizing it and also to establish an

optimal strategy to follow if there is suspicious region, while segmentation of glandular tissue can visualize the suspicious region. Furthermore, segmentation of breast anatomical region can give more specific delineation of breast tissue to help radiologist in the interpretation.

### V. Future Work

It is important to combine segmentation of the breast into anatomical regions with the segmentation of glandular tissue for general breast cancer screening. Then, focusing on the dense component, specific segmentations of glandular tissue areas should be adapted for breast lesion characterizations. Finally, breast density estimation for breast cancer risk assessment or for monitoring the changes in breast density as prevention or intervention procedure, should also be incorporated. Therefore, future works should combine all the steps in the Computer Aided Diagnosis System

### References

- [1] M. J. Yaffe, "Mammographic density: Measurement of mammographic density", Breast Cancer Research, Vol. 10, No. 3, 2008.
- [2] K. E. Martin, M. A. Helvie, C. Zhou, M. A. Roubidoux, J. E. Bailey, C. Paramagul, C. E. Blane, K. A. Klein, S. S. Sonnad, H-P Chan, "Mammographic Density Measured with Quantitative Computer-aided Method-Comparison with Radiologists' Estimates and BI-RADS Categories", Radiology, vol. 240, 2006, pp. 656-665.
- [3] J. N. Wolfe, "Risk for breast cancer development determined by mammographic parenchymal pattern", Cancer, Vol. 37, 1976, pp. 2486-2492.
- [4] N. F. Boyd, J. W. Byng, R. A. Long, E. K. Fishell, L. E. Little, A. B. Miller, G. A. Lockwood, D. L. Tritchler and M. J. Yaffe, "Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening study", J. Nat. Cancer Inst., Vol. 87, 1995, pp. 670-675.
- [5] C. H. van Gils, J. H. Hendriks, R. Holland, N. Karssemeijer, J. D. Otten, H. Straatman, A. L. Verbeek, "Changes in mammographic breast density and concomitant changes in breast cancer risk", Eur J Cancer Prev., Vol. 8, No. 6, 1999, pp. 509-515.
- [6] N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms", Phys. Med. Biol., Vol. 43, 1998, pp. 365-378.
- [7] J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, M. J. Yaffe, "Automated analysis of

- mammographic densities*”, Phys. Med. Biol., Vol. 41, 1996.
- [8] P. Miller and S. M. Astley, “*Classification of breast tissue by texture analysis*”, Image Vision Comput., Vol. 10, 1992
- [9] T. Matsubara, D. Yamazaki, M. Kato, T. Hara, H. Fujita, T. Iwase, T. Endo, “*An automated classification scheme for mammograms based on amount and distribution of fibroglandular breast tissue density*”, International Congress, series 1230, 2001.
- [10] K. Bovis and S. Singh, “*Classification of mammographic breast density using a combined classifier paradigm*”, in Proc. Med. Image Understanding Anal. Conf., 2002, pp. 177–180.
- [11] J Kittler, M. Hatef, R. P. W Duin, and J. Matas, “*On combining classifiers*,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, 1998, pp. 226–239.
- [12] S. R. Aylward, B. M. Hemminger, E. D. Pisano, “*Mixture Modeling for Digital Mammogram Display and Analysis*”, in *Digital Mammography*, N. Karssemeijer, M. A. O. Thijssen, J. H. C. L. Hendriks, L. J. T. O. Van Erning, editors, Computational Imaging and Vision Series, Vol. 13, Kluwer Academic Publishers, Dordre.