

A New Local Adaptive Mass Detection Algorithm in Mammograms

Ehsan Koozegar¹, Mohsen Soryani¹ and Ines Domingues²

¹*Iran University of Science and Technology(IUST), Narmak, Tehran, Iran*

²*INESC TEC (formerly INESC Porto) and Faculty of Engineering, University of Porto, Porto, Portugal*

Keywords: Mamograms, Masses, Detection.

Abstract: Mammography is the most effective procedure for an early detection of breast abnormalities. Masses are a type of abnormality which are very difficult to be visually detected on mammograms. In this paper an efficient method for detection of masses in mammograms is introduced and tested. The algorithm is inspired by binary search and was evaluated both on mini-MIAS and INBreast databases. Mini-MIAS results show that our algorithm outperforms other competing methods. For INBreast database there are no other published mass detection results for comparison, but we believe that our algorithm has good performance.

1 INTRODUCTION

Breast cancer is one of the most lethal diseases in various parts of the world especially in western countries. Several reports about the outbreak and severity of breast cancer are published by different organizations (Oliviera et al., 2011). According to some reports, breast cancer is the second most common disease after lung cancer (10.9% of cancer incidence in both men and women) and the fifth most common cause of cancer death (Oliveira et al., 2009). The National Breast Cancer Foundation has estimated that 200,000 people suffer from the disease and 20,000 die every year. Furthermore, according to American National Cancer Institute, every three minutes one woman is diagnosed with a cancerous case and every 13 minutes one woman is killed by the disease (Oliveira et al., 2009).

Among various modalities, mammography is the most popular method to detect different abnormalities in breasts. During the last two decades, many scientists have been attempting to help radiologists in the detection and diagnosis of these anomalies. It is however important to note that Computer Aided Diagnosis (CAD) systems are designed to assist radiologists only as a second interpretation and never as a substitute.

Masses and microcalcifications (MCCs) are two most frequent findings in mammograms. Detecting masses is more difficult than detecting MCCs because mass features can be ambiguous or similar to breasts' parenchyma. Masses are usually located

in the dense regions of the breast. Furthermore, they have smoother boundaries than MCCs and more various shapes as well. These factors make mass detection a challenging problem both for humans (radiologists) and machines (CAD systems). It has been reported that most abnormalities missed by radiologists are related to cancerous masses (Malagelada, 2007). Most of the available commercial CAD systems for detecting MCCs have reached 100% of detection rate, but the detection rate of masses is still below 90%.

Mass detection has a vital role in full CAD systems and many studies have been made during the last two decades. A good review was published in (Oliver et al., 2010) that covers mass detection algorithms until 2008. Some of the more recent works are worth mentioning. Martins team (Martins et al., 2009) presented a methodology for detecting masses in digitized mammograms using the growing neural gas algorithm for image segmentation and Ripley's K function to describe the texture of segmented structures. Growing neural gas is an incremental and non-supervised clustering algorithm while Ripley's K function is a second-order tool to analyze completely mapped spatial-point process data. Using the digital database for screening mammography (DDSM), the methodology reached an accuracy rate of 89.3%, with 0.93 false positive (FP) and 0.02 false negative (FN) per image.

In 2010, Gao (Gao et al., 2010) proposed two concentric layer criteria to detect different types of suspicious regions. After mammograms are

separated into multi-intensity layers by using different intensity thresholds, the real mass regions should contain some concentric layers on different intensity layers. If one of the following criteria is satisfied, the region is deemed as a suspicious region: Multilayer Criterion (focal regions with concentric layers ≥ 1 are considered as mass regions and the confidence increases with the increase in the number of layers) and Single-Layer Criterion (focal regions without concentric layers in their adjacent lower intensity layer are considered as mass regions only if their morphological features satisfy stricter threshold conditions and the additional contrast condition at the same time). The combination was evaluated on DDSM, resulting in a sensitivity of 99% in malignant, 88% in benign, and 95.3% in all types of cases.

Mencattini and Salmeri (Mencattini and Salmeri, 2011) developed a suspicious mass detection scheme where, after smoothing images with a Gaussian filter, it evaluates them with a Gradient and Hessian matrix. FP rejection is achieved by comparing three features against a threshold: condition number, mean eigenvalues intensity map, and Area. For a sensitivity of 0.9, results on DDSM images exhibit a False- Positive per Image equal to 0.6 for cancer cases and 0.2 for normal cases.

Sampaio's team (Sampaio et al., 2011) proposed to use a cellular neural network to segment regions that might contain masses. Two templates were used: Textural template (is able to separate the mass candidates but inserts pixels that do not belong to the candidates) and Blur template (does not insert extra pixels, but might remove several pixels from the candidates). Images resulting from the use of these templates are aggregated using the binary OR operator. Sensitivity of 80% with rates of 0.84 FPs per image and 0.2 FNs per image, and an area under the ROC curve of 0.87 were obtained in the DDSM database.

The remainder of this paper is organized as follows: section 2 covers all the technical details like used databases, mass detection algorithm description and evaluation methodology; in section 3 results of our method are shown and compared with those of other competing methods. The paper closes in section 4 with some final remarks and future work directions.

2 MATERIALS AND METHODS

In this section we detail on the set of mammographic images used, the mass detection algorithm and the

methodology used to evaluate the results.

2.1 Databases

Two mammogram databases were used, mini-MIAS and INBreast. Mini-MIAS (Suckling et al., 1994) consists of 330 images in which every image is of size

1024*1024. These 330 images include 209 normal images, 56 images with at least one mass, and the remaining have other types of anomalies. One of the images (mdb059) was discarded in our experiments because there is no information about the center of the mass present in the mammogram. INBreast database (Moreira et al., 2012) has a total of 115 cases (410 images) of which 90 cases are from women with both breasts (4 images per case) and 25 cases are from mastectomy patients (2 images per case). Several types of lesions (masses, calcifications, asymmetries, and distortions) are included. In this work 107 images were used (all images with at least one mass) with a total of 116 masses.

Note that, while mini-MIAS is a well-known database, with the advantage of being already used in several published works, it is a small database of digitized mammograms having only the center and radius information about the findings' location. INBreast, however, is a recent database having the disadvantage of not being used by many works yet making it more difficult to compare among different algorithms. It has, as advantages, the fact that all the images are Full-field digital mammograms and accurate information on the form of detailed contours on the shape and location of every finding is available.

2.2 Detection

Our mass detection algorithm is a local adaptive thresholding method which has been inspired from binary search to determine an appropriate threshold related to each local region (called cell).

The flowchart of the mammogram mass detection algorithm (applied to each cell of the grid respectively) is shown in Figure 1.

Each image is first divided into equal non-overlapping cells (a grid). In each cell of the grid, the pixel with maximum gray level is found. The location of the maximum pixel is shown as Index and its value is named m .

First and *Last* are the bounds of the range which is being explored and TH is the proper threshold. *First* and *Last* are initialized to 0 and m respectively.

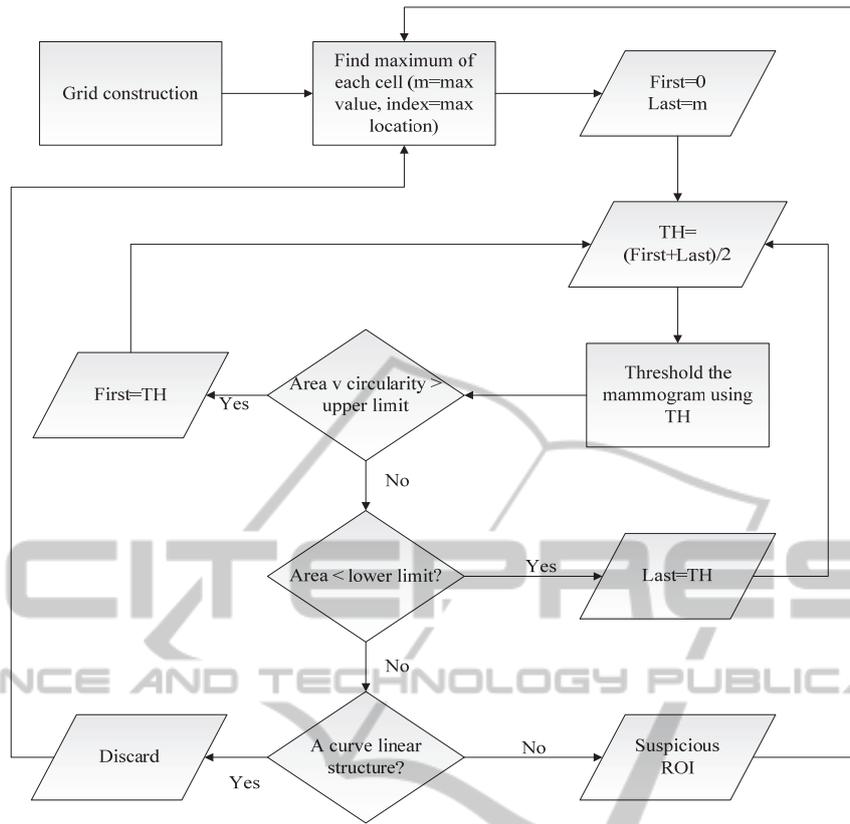


Figure 1: Flowchart of the mammogram segmentation algorithm.

In the first iteration, TH is assigned with the middle value of the range $[0;m]$ and then the threshold is applied to the whole mammogram. After that, the circularity measure is extracted from the region that contains $index$ ($index$ is the location of the maximum value):

$$\text{Circularity} = P^2 / 4\pi A \quad (1)$$

In Equation 1, P is the circumference of the region and A is the area. In this equation, maximum circularity is 1 and the less circular the region, the bigger the circularity value will be. If the area or the circularity of that region exceeds the corresponding upper limit ($Areamax$, $Circmax$) then we should search within the upper half of the previous range (i.e. $[TH;Last]$). Else, if the area of the region is lower than a threshold ($Areamin$), then the TH is too high and we should search the proper TH in the range $[First;TH]$. These operations are iterated until a region with the area between $Areamin$ and $Areamax$ and also less circularity than $Circmax$ is found (if it exists). In fact, masses generally have a radius between a lower limit and an upper limit and are also not very irregular. Although spiculated masses are irregular in shape, their circularity can be lower

than a predefined limit and the irregularity occurs in the margins of those masses and has a fairly small effect on the whole circularity of those regions.

Moreover, we defined another measure as:

$$AC_{ratio} = \text{Area} / \text{Circularity} \quad (2)$$

to filter curve-linear structures such as blood vessels and milk ducts in mammograms. If a region has a value lower than a predefined threshold, it is considered as a curve-linear object and discarded.

2.3 Evaluation

We combine three typical rules to create a very strong tagging rule. These rules are as follows:

- c_1 : if $|Y_b - Y_{cad}| < \max(R_b, LY/2)$ and $|X_b - X_{cad}| < \max(R_b, LX/2)$
- c_2 : if $(X_{cad} - X_b)^2 + (Y_{cad} - Y_b)^2 \leq R_b^2$
- c_3 : if exists 50% overlap between biopsy-proven mass and suspicious region

Where (LX, LY) are the length and width of the detected ROI bounding box, (X_{cad}, Y_{cad}) is the region's center of gravity and $(X_b, Y_b); R_b$ are center and radius of the biopsy-proven mass. An extracted ROI is labelled as True Positive (TP) if all the above

rules are true. Otherwise, that ROI is labelled as FP.

3 EXPERIMENTS AND RESULTS

All images were scaled to 512*512 pixels and the parameters needed for the mass detector were empirically set as: $A_{max} = 8000$ pixels, $A_{min} = 155$ pixels and $C_{max} = 7$. Some detection examples are given in Figure 2 for mini-MIAS and in Figure 3 for INBreast database.

In Table 1 some detection methods are shown for comparison. From Table 1 it can be seen that only the sensitivity of Density-Weighted Contrast Enhancement (DWCE) filter is comparable with the sensitivity of our mass detection algorithm. But, with almost the same sensitivity, our method results in fewer FPs per image (4.77) than those of DWCE filter (12).

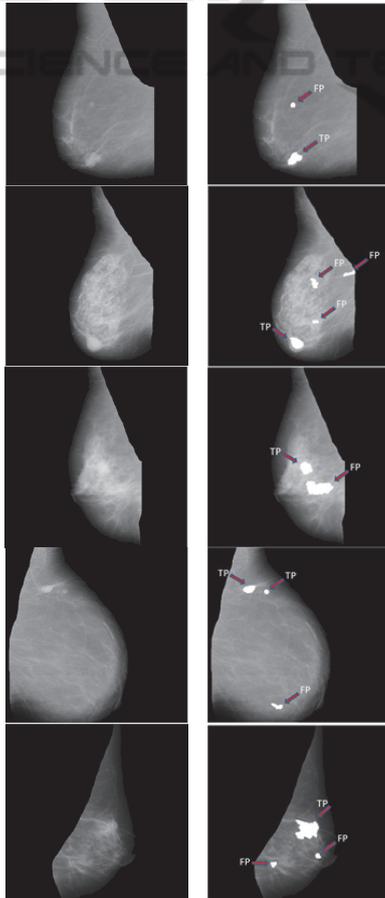


Figure 2: mini-MIAS mass detection examples. Left column: original mammograms; right column: regions obtained after segmenting the original mammograms.

Region growing (Eltonsy et al., 2007), Adaptive

thresholding (Kom et al., 2007) and Difference of Gaussians (DoG) filter (Oliver et al., 2010) have lower sensitivity and also more FPs in comparison with our method. Although Template matching (Nguyen et al., 2010) results in less FP, its sensitivity is very low, making it unreliable.

By using the proposed method, the detection of masses in 261 mammograms lasted 388 minutes, i.e., one minute and 48 seconds for each image.

As far as we know, this work is the first published work presenting mass detection results on INBreast database. We have achieved a sensitivity of 87% on INBreast (the algorithm missed 15 masses) and FP rate per image is 3.67.

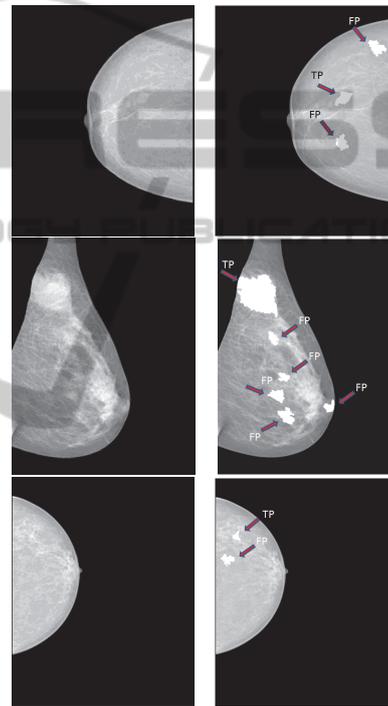


Figure 3: INBreast mass detection examples. Left column: original mammograms; right column: regions obtained after segmenting the original mammograms.

Table 1: Results of different detection algorithms.

Detection method	Sensitivity	FP rate
Template matching (Nguyen et al., 2010)	0.38	2.9
Region growing (Eltonsy et al., 2007)	< 0.6	> 8.5
Adaptive thresholding (Kom et al., 2007)	< 0.65	> 9
DWCE filter (Petrick et al., 1996)	< 0.9	> 12
DoG filter (Oliver et al., 2010)	< 0.72	> 10.5
Our proposed method	0.91	4.77

4 CONCLUSIONS

In this paper, we tested a mass detection algorithm

on two public mammogram databases. The algorithm, inspired by binary search, reached a sensitivity of 91% with false positive rate per image of 4.77 in mini-MIAS database and a sensitivity of 87% with a false positive rate per image of 3.67 in INBreast.

REFERENCES

- Eltonsy, N. H., Tourassi, G. D., and Elmaghraby, A. S. (2007). A concentric morphology model for detection of masses in mammography. *IEEE Transactions on Medical Imaging*, 26(6):880–889.
- Gao, X., Wang, Y., Li, X., and Tao, D. (2010). On combining morphological component analysis and concentric morphology model for mammographic mass detection. *IEEE transactions on information technology in biomedicine*, 14(2):266–273.
- Kom, G., Tiedeu, A., and Kom, M. (2007). Automated detection of masses in mammograms by local adaptive thresholding. *Computers in Biology and Medicine*, 37(1):37–48.
- Malagelada, A. O. (2007). Automatic mass segmentation in mammographic images. Ph.D. dissertation, Department of Electronics-Computer Science and Automatic Control, University of Girona.
- Martins, L. D. O., Paiva, A. C. D., and Gattass, M. (2009). Detection of breast masses in mammogram images using growing neural gas algorithm and ripleys k function. *Journal of Signal Processing Systems*, 55:7790.
- Mencattini, A. and Salmeri, M. (2011). Breast masses detection using phase portrait analysis and fuzzy inference systems. *International Journal of Computer Assisted Radiology and Surgery*, page 111.
- Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., and Cardoso, J. S. (2012). INbreast: toward a full-field digital mammographic database. *Academic radiology*, 19(2):236–248.
- Nguyen, V. D., Nguyen, D. T., Nguyen, T. D., Thi, N., and Tran, D. H. (2010). A program for locating possible breast masses on mammograms. *In Proceeding of the 3rd International Conference on the Development of BME in Vietnam*, pages 11–14.
- Oliveira, M. L. D., Braz, J. G., Cardoso, P., and Gattass, M. (2009). Detection of masses in digital mammograms using k-means and support vector machine. *Electronic Letters on Computer Vision and Image Analysis*, 8(2):39-50.
- Oliver, A., Freixenet, J., Marti, J., Prez, E., Pont, J., Denton, E. R. E., and Zwiggelaar, R. (2010). A review of automatic mass detection and segmentation in mammographic images. *Medical Image Analysis*, 14(2):87–110.
- Oliviera, J. d., Albuquerque, A. d., and Deserno, T. M. (2011). Content-based image retrieval applied to BIRADS tissue classification in screening mammography. *World Journal of Radiology*, 3(1):24-31.
- Petrick, N., Chan, H., Sahiner, B., and Wei, D. (1996). An adaptive Density-Weighted contrast enhancement filter for mammographic breast mass detection. *IEEE Transactions on Medical Imaging*, 15(1):59–67.
- Sampaio, W. B., Diniz, E. M., Silva, A. C., de Paiva, A. C., and Gattass, M. (2011). Detection of masses in mammogram images using CNN, geostatistic functions and SVM. *Computers in Biology and Medicine*, 41(8):653-664.
- Suckling, J., Parker, J., Dance, D. R., Astley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S. L., Taylor, P., Betal, D., and Savage, J. (1994). The mammographic image analysis society digital mammogram database. *In Exerpta Medica. International Congress Series 1069*, volume 1069, pages 375–378.