ISSN 2278-3091
Volume 9, No.3, May - June 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse55932020.pdf
https://doi.org/10.30534/ijatcse/2020/55932020

Mammogram Breast Cancer Detection using Fast Watershed Segmentation

Osama R.Shahin [1],Rami Ayadi[2], Oussama Ghorbel[3]
[1]Jouf University, KSA, orshahin@ju.edu.sa
[2] Jouf University, KSA, rayadi@ju.edu.sa
[3] Jouf University, KSA, oaghorbel@ju.edu.sa

ABSTRACT

Image Processing represents the main research area within engineering and computing specialization. It is promptly rising technologies today, and its applications found in various aspects of biomedical fields specifically in cancer disease. Breast cancer is taken into account the fatal one in all cancer types in step with recent statistics everywhere the globe. It's the foremost common cancer in women and also the second reason of cancer death between females. In this paper, we recommend implementing a fast segmentation algorithm employing a watershed transformation. This extends the partitioning of the dividing waterline by allowing the mixing of advance information about image objects and therefore the traditional watershed algorithm. Before the watershed transformation can begin, the algorithm needs a way of representing the test image in terms of the amount of change around any particular pixel. We apply the Sobel operator to each pixel in the grayscale representation of the original image. The tumors detected are circular or semicircular in shapes according to the shape, and the brightness of the tumor will be darker as we moved far away from its center. The complement for this prior information can be taken as a local minimum that required beginning the watershed algorithm. So each tumor image can be represented as a lake with minimum value represented by the center in the complement tumor image. After employing the method, the detection of tumor percentage becomes more reliable. Such result indicates that the new technique has improved the performance of our computer aided diagnosis system for mammographic breast cancer detection effectively. The algorithm was implemented in MATLAB and tested under Windows system. The method was tested over several images from MIAS (Mammogram Image Analysis Society, UK), that provides a standardized classification for mammographic studies. In this paper, we recommend implementing a quick segmentation algorithm employing a watershed transformation. This extends the partitioning of the dividing waterline by allowing the mixing of advance information about image objects and therefore the traditional watershed algorithm.

Key words: Mammogram, Geometric Features, Gradient Features, Texture Features, Key point detection, Watershed transform.

1. INTRODUCTION

Breast cancer is an uncontrollable way to grow cells. Breast cancer is the most common cause of death among women, especially in middle age, as the World Health Organization's International Agency for Research on Cancer estimates the number of breast cancer deaths about one million women worldwide per year [1]. Hence, early diagnosis is the main and only factor that helps reduce this alarming rate. Mammography - a special type of imaging with low doses of x-rays - works in early detection of cancer and thus allows for faster recovery [2],[4]. Malignant cells in radiography usually appear as irregular lumps, while those characterized by regular, circular forms are benign [5]. Researchers working in the field of image processing, especially those who discover tumors, extract features from potential areas that may be affected. These features may be based on color characteristics or related to engineering estimates, which, if professionally estimated, serve as the advisory opinion of the treating physician and thus contribute to the early detection of the disease [2, 6]. For this reason, the main research idea is to extract key points for potential cells and then locate these places using watershed algorithm. In this study, digital mammograms provided from the online mammogram database (MIAS database) were used [7]. First, each suspicious region will be located for each image and secondly a number of important features will be calculated. After running the algorithm, the features obtained will be compared to the standard engineering and texture [6],[8],[10]. The remainder of the paper is organized as follows: We discuss the extraction of major and important breast cancer features in Section Two. The proposed algorithm is described in Section III. The experimental results are shown in section four. Shapes and tables are drawn in section five. Finally, we conclude the paper in Section VI.

2. TUMOR FEATURES

In image processing, we distinguish for tumors several classes of entities. In our case we will only look at the methods that help us the most. Therefore, we will use to translate the medical terms the functionalities which will be used more in the phase of calculate. In the next part, we cite the most used categories of entities such as geometric, textures and gradients [6].

2.1 Geometric Features

By definition, they are used to define the geometric dimensions of the area of interest. Then, to be able to use these geometrical dimensions in the recognition phase, we must present the properties that a set of pixels contains in an image [11]. The geometric dimensions of the area of interest represent the basic descriptors of a region which will be used to identify the objects. In fact, in a breast, objects are recognized fundamentally by these geometric dimensions. in a diagnostic, we will use these geometrical dimensions namely the compactness, the outline and the area to try to determine the objects in a breast image and differentiate the areas of interest [12].

2.2 Texture Features

Despite the progress of research in modeling, the texture still remains a blurred aspect for the field of the research of objects in images and this for lack of formal and mathematical foundation given the difficulty of modeling the perception of the human species like the way in which an object like glass can appear and also the way in which we can feel it. This perception varies from one individual to another and depends on his vision or perception of the texture. Regardless of the absence of formal models and a standard definition of the texture to model or synthesize it, there are several techniques for extracting this texture. Indeed, the researchers use descriptors such as coarseness, softness, regularity to extract the texture of an object. The basic idea is that since the image is composed of blocks of texture zone samples, we can therefore deduce that easily its descriptors which reflect the change in gray level of the pixels linked in an image [13].

2.3 Gradient Features

The value of the gradient of an image is determined by the operator Sobel, in fact, this operator calculates the function derived from the local values of an image (this is the definition of the gradient of an image). In the image gradient, the degree of the edge is raised in relation to the original image [14]. Despite its complexity, the use of this operator is more advantageous from the point of view of its sensitivity to the level of the points of isolated variations which are characterized by a higher value of intensity due to the fact that it uses larger regions. We will produce the gradient images using the Sobel operation. The operator's entry is the set of mammography images. Then, starting from the gradient images resulting from this first operation, we will extract the different descriptors of the texture.

In the next step and based on the histogram of the areas of interest of the gradient image, we will deduce the normalized histogram of the gradient image. These two histograms will be used to determine the gradient descriptors.

3. METHODOLOGY

The suspicious area i.e. (ROI) can be determined from a mammogram image according to the process that shown in the flowchart in Figure 1.

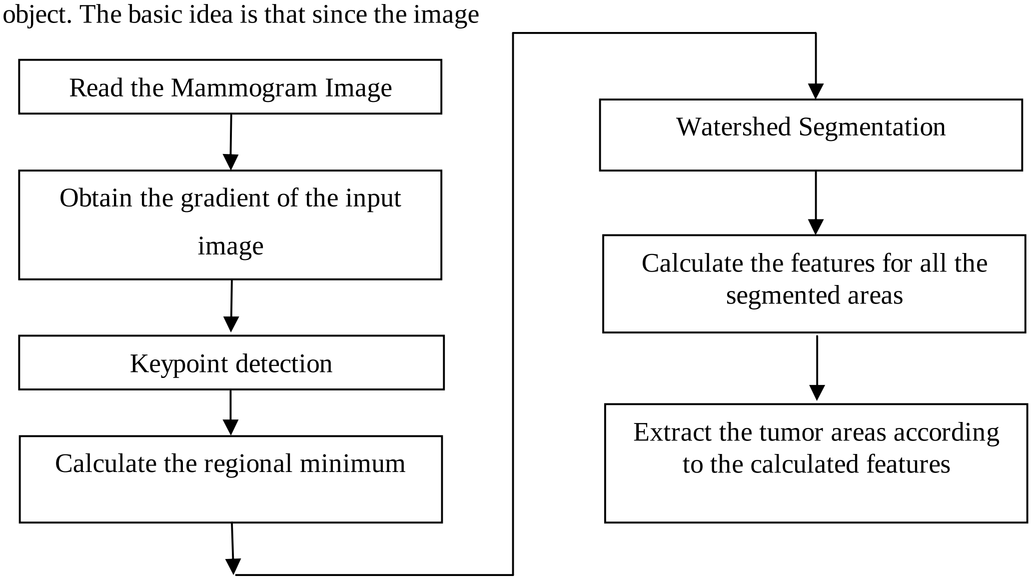


Figure 1: Flow chart for the proposed algorithm

As shown in the last flowchart we can summarize the proposed algorithm as follows:

3.1 Reading the mammogram image

As mentioned before the mammogram images will taken from the MIAS dataset, sample for these images was depicated in Fig.4 (a)

3.2 Obtain the gradient of the input image

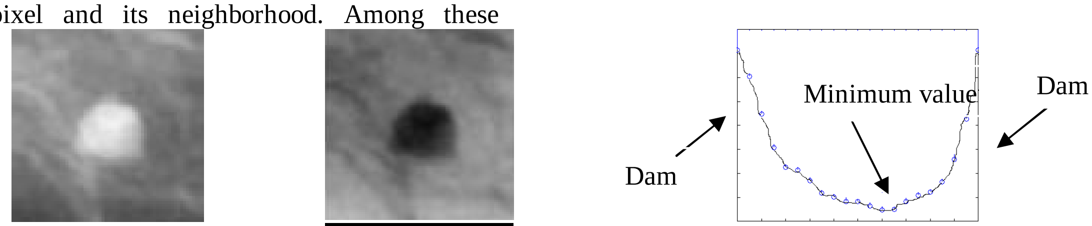
The importance of this step in making a contrast between the expected region of interest (ROI), i.e. tumor region, and the background image. There are different ways to determine the image gradient, Canny, Prewitt, Sobel, Roberts,… etc. This work admit the Sobel detector, because it is simple to implement, and it does not produce a considerable noise.

3.3 Key point detection

The image's properties come from the regions differences from some pixel and its neighborhood. Among these properties, there are texture, color and intensity. These changes occur at specific points. These points called key points. In this work we use Blob and corner detection to detect such points. A blob detector distinguishes points that are darker or brighter than the surrounding. Where, a corner is the intersection of two lines (edges). It can also be defined as a point for which exist changed edge directions in the area of the point.

3.4 Calculate the regional minimum

By complement the image we can calculate the minimum value for any tumor function which represents the center for each complete tumor image. The whole image of the tumor can be described as complete graph of lake catchment basins with a minimum value bounded by two parallel dams. Figure 2.



Input image — Complemented imag***e*** — Graph for the pixel values for the complemented image

Figure 2: Shape of the tumor image

3.5 Watershed Segmentation

Watershed segmentation is a technique of automatically segment apart elements that touch. It begins with a binary image, where black pixels considered as particles. It determines a distance map to discover the fattest portions of the object, here we talk about tumor. It Beginning with the crests as maximal erosion points (MEP's), then it is dilates them as far as possible - one or the other until the edge of the tumor is reached, or the edge the normal tissues (growing) [15],[16],[18]. The idea of watersheds is established on the following assumption that an image in composed from three dimensions: one of them resent the gray levels, the reminder parts characterize the spatial coordinates. Watershed transform can be realized using flooding processes. These flooding processes can be accomplished by using basic morphological operations, such as dilatation and erosion.

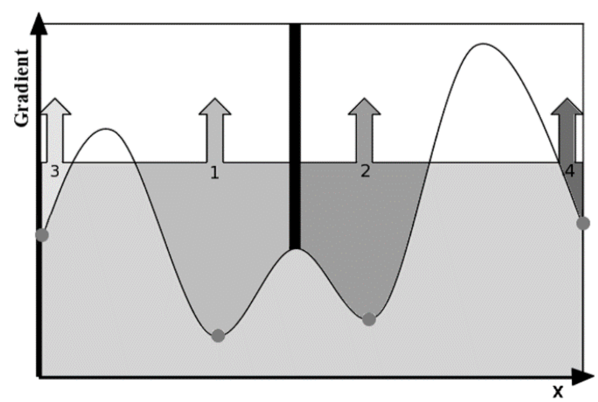


Figure 3: Watershed segmentation[19]

3.6 Calculate the features for segmented areas

In this phase the determination of the geometric, texture, and gradient features will be achieved by using the well-defined equations [6], [12], [13],[1],[16].

3.7 Extract the location of the tumor

According to the region properties for each segmented areas we determine the tumor location according to its features. The determination of the tumor features was based on the features values mentioned in [6], [8],[10].

4. RESULTS AND DISCUSSION

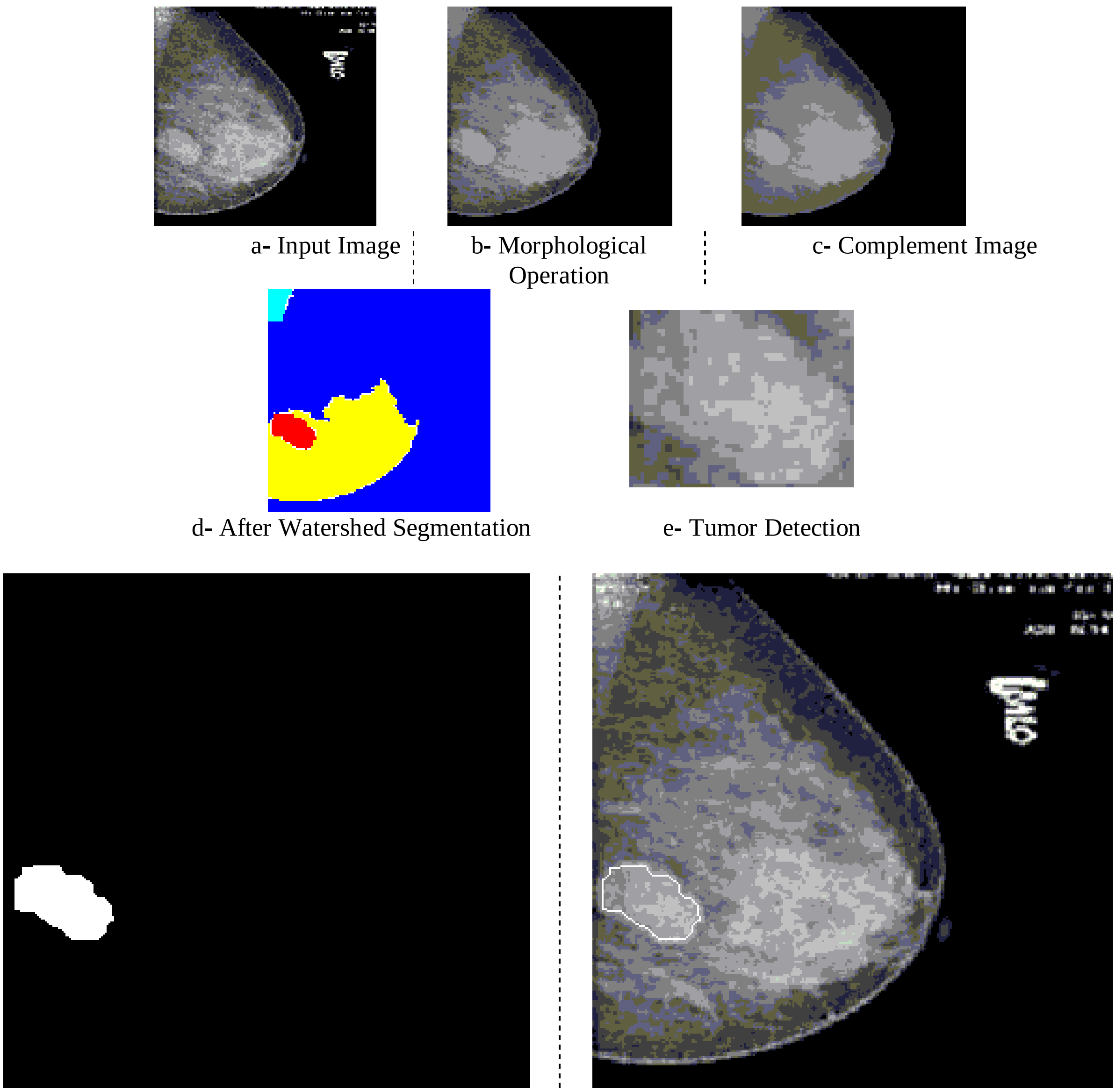
To test the validity of the proposed method, we implemented our algorithm described above and we tested it. The input of the algorithm will be a set composed of 88 real images corresponding to real cases from the standard image database MAIS [7]. There are regions which present differences in the level of the illumination in an image, this inequality is caused by luminous regions in the middle of the mammography which hides certain tissues.

The top hat transformation is applied in order to minimize the inequality in the lighting because we have found that in mammography images the shapes of the objects are lighter than the background of the image. This transformation reflects the level of dissimilarity between the original image and all of the foreground elements that make up the image and that represent the precise structure of an element.

After extracting the entire breast region and detecting the suspect region, a calculation will be made based on the characteristics of these regions. 22 characteristics are used for the calculation operation of which they are going to be recorded in a database which has 88 lines.

We use each line for an image and even 22 columns because a calculation of 22 extraction functionalities is performed in each image.

Thereafter, the summary functionality will be presented in table 1. The latter contains the minimum and maximum value for each functionality. Finally, we rounded up the whole result to three decimal places for easier reading.



a- Input Image

b- Morphological Operation

c- Complement Image

d- After Watershed Segmentation

e- Tumor Detection

f- Segmented Tumor

g- Outlined Original Image

Figure 4: Sample of Tumor Segmentation

Table 1: Feature extraction range values for cancer tumor

Feature	Maximum	Minimum
Area	255	70
Perimeter	3500	265
Compactness	33.53	16.67
Mean	239.77	139.21
Mean Global Area	8.32	1.23
Mean Local Area	12.82	1.35
Uniformity	4.34	0.98
SD	30.32	11.34
Smoothness	0.008	0
Skewness	39.43	-7.41
Entropy	-25.23	-200.01
Correlation	571	52
Inverse	0.56	0.003
Sobel Mean	98.23	46.37
Sobel Mean Global	7.57	1.54
Sobel Uniformity	3.12	0.008
Sobel SD	25.43	4.6
Sobel Smoothness	0.05	0.001
Sobel Skewness	25.005	-4.54
Sobel Entropy	-22.21	-243.23
Sobel Correlation	381	9
Sobel Inverse	7.97	0.003

5. CONCLUSION

Based on the comparison results obtained and shown in Table 1, we note that the totality of the functionalities deduced by the proposed algorithm is similar in terms of functionalities proposed in the following works [6], [8],[10],[16], [17]. Thus, our proposed algorithm shows that it is able to detect the region of tumors. It also makes it possible to determine its eventual locations. Other than the last benefit, our algorithm could detect more than one tumor in the same breast region. This implies the power and robustness of the proposed algorithm, as shown in Figure 4. Based on the experimental results obtained previously, they show good results based on the extraction of characteristics proposed. Consequently, it proves its capacity to implement a CAD system for breast cancer.

REFERENCES

1. Shahin, Osama R., Hamdy M. Kelash, Gmal Mahrous, and Osama S. Farag Allah. **A Novel CAD System for Breast Cancer Detection.** Cancer Biology 2014;4(3)
2. SAMETI, Mohammad. **Detection of soft tissue abnormalities in mammographic images for early diagnosis of breast cancer**. *Diss. University of British Columbia*, 1998.
3. SILVERSTEIN, Melvin J., LAGIOS, Michael D., RECHT, Abram, et al. **Image-detected breast cancer: state of the art diagnosis and treatment**. *Journal of the American College of Surgeons*, 2005, vol. 201, no 4, p. 586-597. https://doi.org/10.1016/j.jamcollsurg.2005.05.032
4. SCHIABEL, Homero, SANTOS, Vivian T., et ANGELO, Michele F. **Segmentation technique for detecting suspect masses in dense breast digitized images as a tool for mammography CAD schemes**. *In : Proceedings of the 2008 ACM symposium on Applied computing*. 2008. p. 1333-1337. https://doi.org/10.1145/1363686.1363996
5. LINDELL, Scott, SHAPIRO, Gary, WEIL, Kenneth, et al. **Development of mammogram computer-aided diagnosis systems using optical processing technology**. *In : Proceedings 29th Applied Imagery Pattern Recognition Workshop. IEEE*, 2000. p. 173-179.
6. AL-SHAMLAN, Hala et EL-ZAART, Ali. **Feature extraction values for breast cancer mammography images**. *In:2010 International Conference on*

Bioinformatics and Biomedical Technology. IEEE, 2010. p. 335-340.
https://doi.org/10.1109/ICBBT.2010.5478947

7. SUCKLING J, P. **The mammographic image analysis society digital mammogram database**. *Digital Mammo*, 1994, p. 375-386.
8. SU, Yanni et WANG, Yuanyuan. **Computer-Aided Classification of Breast Tumors Using the Affinity Propagation Clustering**. *In : 2010 4th International Conference on Bioinformatics and Biomedical Engineering. IEEE*, 2010. p. 1-4..
https://doi.org/10.1109/ICBBE.2010.5518144
9. CHANG, Ruey-Feng, WU, Wen-Jie, MOON, Woo Kyung, et al. **Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors**. *Breast cancer research and treatment*, 2005, vol. 89, no 2, p. 179.
https://doi.org/10.1007/s10549-004-2043-z
10. CHEN, Yunmei, THIRUVENKADAM, Sheshadri, TAGARE, Hemant D., et al. **On the incorporation of shape priors into geometric active contours**. *In : Proceedings IEEE Workshop on Variational and Level Set Methods in Computer Vision. IEEE*, 2001. p. 145-152.
11. VAN DER VELDEN, Arjan P. Schouten, BOETES, Carla, BULT, Peter, et al. **The value of magnetic resonance imaging in diagnosis and size assessment of in situ and small invasive breast carcinoma**. *The American journal of surgery*, 2006, vol. 192, no 2, p. 172-178.
https://doi.org/10.1016/j.amjsurg.2006.02.026
12. MATSUBARA, Tomoko, FUJITA, Hiroshi, KASAI, Satoshi, et al. **Development of new schemes for detection and analysis of mammographic masses**. *In : Proceedings Intelligent Information Systems*. IIS'97. IEEE, 1997. p. 63-66.
13. MENCATTINI, Arianna, SALMERI, Marcello, LOJACONO, Roberto, et al. **Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing**. *IEEE transactions on instrumentation and measurement*, 2008, vol. 57, no 7, p. 1422-1430.
https://doi.org/10.1109/TIM.2007.915470
14. CRUZ, María Victoria Carreras et VILELLA, Patricia Rayon. **Circumscribed mass detection in digital mammograms**. *In : Electronics, Robotics and Automotive Mechanics Conference (CERMA'06). IEEE*, 2006. p. 19-24.
15. Shahin, Osama R., Meshrif Alruily, Mansi Alsmarah, and Musharraf Alruwaill. **Breast cancer detection using modified Hough transform.** Biomedical Research (2018) Volume 29, Issue 16.
https://doi.org/10.4066/biomedicalresearch.29-18-912
16. Shahin, Osama R., and Gamal Attiya. **Classification of Mammograms Tumors Using Fourier Analysis.** IJCSNS 14, no. 2 (2014): 110.
17. NALLAMALA, Sri Hari, MISHRA, Pragnyaban, et KONERU, Suvarna Vani. **Qualitative Metrics on Breast Cancer Diagnosis with Neuro Fuzzy Inference Systems**. International Journal of Advanced Trends in Computer Science and Engineering, 2019, vol. 8, no 2.
https://doi.org/10.30534/ijatcse/2019/26822019
18. Andrews Jose, Dr.D. Sujitha Juliet. **Recent advances and investigation of efficient Computer Aided Diagnosis systems for CT images in Liver cancer detection.** 2019, vol. 8, no 3. p. 343- 348.
https://doi.org/10.30534/ijatcse/2019/02832019