

RESEARCH ARTICLE

Un-supervised segmentation and quantisation of malignancy from breast MRI images

Rajeev Ratan^{1*}, P.G. Kohli², Sanjay K. Sharma³ and Amit K. Kohli³

¹ Department of Electronics and Communication Engineering, MVN University, Palwal, Haryana, India.

² Department of Physiology, Punjab Institute of Medical Sciences, Jalandhar, India.

³ Department of Electronics and Communication Engineering, Thapar University, Patiala, Punjab, India.

Revised: 20 April 2015; Accepted: 18 February 2016

Abstract: In this paper, a magnetic resonance imaging (MRI) based image segmentation technique has been proposed, which uses a magnetic resonance parametric information model for breast tumor segmentation. The methodology has been developed on two dimensional MRI datasets. With the help of the proposed technique, breast tumor tissues can be segmented in 6 - 8 minutes with more precision and reproducibility than manual (supervised) segmentation, which takes more than two hours to segment breast tumor tissues. Thus, the proposed semi-automatic (un-supervised) technique can be applied to analyse MRI images, which improves the procedure for diagnosing breast cancer, and it can also be used to generate two-dimensional view of tumor in case of surgical operations.

Keywords: Breast tumor, MATLAB, MRI, region of interest (ROI), watershed segmentation.

INTRODUCTION

The human body comprises different types of cells having specific functions associated with each of them. Almost all the cells are maintained within the human body in an organised way to generate fresh cells as required by the human body for proper functioning. This generation of new cells is carried out in a controlled manner and if disturbed, they divide too frequently and in a disorderly manner resulting in the formation of extra cells, and thus forming a mass of tissue called a tumor. The tumors may be classified as either benign tumors or malignant tumors. To segment these tumors, we use three segmentation techniques: snakes (gradient vector flow), level set segmentation, and watershed segmentation (Ratan *et al.*, 2009a).

Watershed segmentation makes use of the contrast as one parameter for segmenting the entire breast image dataset. Furthermore, the supplementary intricacy of judgment forced to such types of algorithms makes the basis of a propensity towards approaches, which are based on contrast (Ratan *et al.*, 2009b). Amongst the potential techniques for this purpose, watershed segmentation may be applied as an influential tool, which will unreservedly identify the surface of the tumor.

In today's context both supervised (classification) and un-supervised (clustering) methods are used for the detection of tumors that make the system more intelligent. The focus of cancer research these days is at molecular level, which is reflected in MRI scans when treatment therapies and strategies are applied. Therefore at each stage, the treatment is monitored using non-intrusive techniques like MRI scans. Un-supervised methods involve the calculation of edge, gray and contrast parameters. There are many techniques, which use the above mentioned parameters such as watershed, random decision forests and density forest, fuzzy clustering, entropy based segmentation, patch-based segmentation, etc. But none of these methods give results using all three parameters, i.e. edge, gray and contrast parameters. In the proposed method, such algorithms have been incorporated that use all three aforementioned parameters, which are essential for detecting tumors.

This research study presents the design methodology of an un-supervised tool using MRI datasets for the detection of a breast tumor. The proposed technique is a miniscule and practical part of a relatively intricate

* Corresponding author (rajeevratanarora@gmail.com)

system. The presented technique is useful for surgeons to envisage within the human body, so that they can carry out operations without performing open surgery. The endeavour of this proposed methodology is to segment the breast tumor in breast MRI images, so as to make the tumor perceptible, which will ease the treatment. Computed tomography (CT) scan and MRI are the diagnostic instruments, which are usually proposed for this treatment. In this paper, the images used have been taken from MRI (Black *et al.*, 1997; Deng *et al.*, 2001).

Algorithms based on segmentation using the watershed technique have been applied for the detection of tumor in two-dimensional view. MATLAB has been used to carry out the aforesaid.

METHODOLOGY

An un-supervised shape, texture, and content based method has been applied to examine MRI breast images with comparatively lower computational necessities. The proposed methodology is shown in Figure 1.

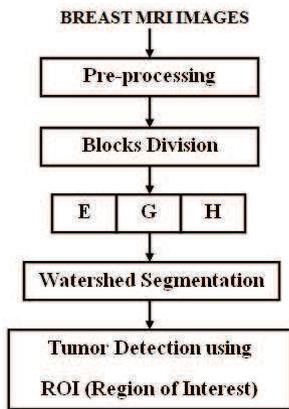


Figure 1: Proposed methodology for breast tumor detection using watershed segmentation

The division of MRI image areas is done using a block-based method (Hinz *et al.*, 2002; Fadi & Abdel, 2007). Each block is examined by scheming its multiple parametric values. In this context, the multi-parameter values mean the subsequent three features: the edges (E), gray values (G), and contrast (H) of the block of MRI image under scrutiny (Jolesz & Blumenfeld, 1994; Dewalle *et al.*, 2008; Jason *et al.*, 2008). For the analysis of such small tissue or structure a block system is devised that divides the region of interest (ROI), which is highly dependent on imaging modalities into block size (B). A block (B) is denoted as a part of the whole image (ROI).

Computed tomography (CT) scan can be acquired only in axial (X-Y) plane, but MRI can be acquired in axial (X-Y), sagittal (X-Z) and coronal (Y-Z) planes. In MRI, each dissimilar type of image is referred to as a sequence. The primary MR sequences include T2, T1, T1 with contrast, diffusion and flair. Each sequence has to be acquired separately, which means that the MRI procedure is consuming more time in comparison to a CT scan.

The first step is pre-processing, which uploads the MRI image sequences to MATLAB workspace to use this input sequences as input data. Thereafter, de-noising algorithms have to be applied on input image sequences, which makes the input data noise free (Bailet *et al.*, 1992; Abbasi & Mokhtarian, 2001).

Current trends in medical image investigations frequently comprise steps, so that an image can be segmented with respect to a small number of factors and into lesser sizes to deal with the various features of examining images into regions, which are anatomically and pathologically significant (Lei & Chan, 2001; Bountris *et al.*, 2005). The investigation in the areas of physiological as well as pathological concerns can be made easier and more definable by categorising regions using multi-parameter values. The multi-parameter factors mean the explicit values for the edges (E), gray (G), and contrast (H) parameters of the pixels present in the images (Krishnan & Atkins, 1998; Lihua *et al.*, 2007; Nacim *et al.*, 2007).

The borders of an object can be determined by the edge information. It is primarily used for investigation to obtain resemblance condition for a pre-determined object. The incidences of mammogram compression reduce the edge. Therefore, the proposed methodology uses the Sobel edge detection technique for the detection of image edges (I_E), because it compensates the effects of mammogram compression in the image sequences. In Sobel edge detection algorithm, the image edges can be perceived with the help of filtering of the input MRI image using two convolution kernels in tandem, one for detecting alteration in vertical contrast (h_x), and the other one for sensing horizontal contrast (h_y) as represented in equation (1). Image output (I_E) can be attained by scheming the gradient magnitude of each and every pixel, which has been shown in equation (2). Thereafter, the edge parameter (E) is calculated, whereby E (r,c) enhances by one, every time when I_E (x, y) = '1' in the slab under supervision, as expressed in equation (3).

$$H_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, H_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \dots(1)$$

$$I_E(x, y) = \sqrt{I_x^2(x, y) + I_y^2(x, y)} \quad \dots(2)$$

$$E_{(r,c)} = \sum_{(x,y) \in B} (I_{E_p} == 1) \quad \dots(3)$$

where, B is the individual block under supervision.

G_D value is calculated using the average pixel value (I_{av_s}) of each image slice (S) for the total image slices (T) of an image dataset as given in equations (4) and (5). The 2-D image obtained into consideration from the MRI machine is of 256×256 (i.e., 65536) pixels in one inch. Therefore in equation (4), the division term 65536 has been used.

$$I_{av_s} = \frac{1}{65536} \sum I_p(x, y) \quad \dots(4)$$

$$G_D = \frac{\sum_{S=0}^T I_{av_s}}{T} \quad \dots(5)$$

The gray parameters (G) for every individual slab of the breast MRI images are accrued, and is described using a binary image (I_T) by means of the G_D value, which is taken as a threshold, as shown in equation (6).

$$G_{(r,c)} = \sum_{(x,y) \in B} I_p(x, y), \forall I_T(x, y) = 1 \quad \dots(6)$$

where, B is the individual block under supervision. The pixel strength for each and every individual slice has been considered for creating the threshold values. Therefore it provides the foundation for examining medical MRI images of the breast tumors patients (Gering *et al.*, 2002; Senaratne *et al.*, 2006).

The contrast (H) may be used frequently for distinguishing the degree of dissimilarity in the pixel strength. In the proposed methodology, the computational algorithms analyse the dissimilarities, particularly in the case of sturdy contrast among items/objects in the MRI image $I(x,y)$.

The ‘min./max. extend programme’ for the 8-neighbourhood connectivity has been adopted in this method, in which min H and max H denote the minimum strength and maximum strength of contrast of the locality pixel $C8(I_H)$, respectively as given in equation (7).

$$I_H(x, y) = \left(\frac{I(x, y) - \min H}{\max H - \min H} \right) \times \max H \mid H \in C_8(I_H) \quad \dots(7)$$

In literature, tumor cells have been frequently linked with the superior value of contrast (H) parameter (Gupta & Undril, 1995). H_d can be obtained by summing the dissimilarity of a slab under supervision, which is as follows:

$$H_{d(r,c)} = \sum_{(x,y) \in B} I_H(x, y) \quad \dots(8)$$

where, B is the individual block under supervision.

The proposed work was implemented and tested on a IBM Think-Centre with i5 processor with 2 GB of RAM. The software tool used for analysing the data was MATLAB. To assess the performance of the projected techniques and algorithms, the research was carried out for two different types of breast MRI datasets.

The two different breast MRI datasets were collected as shown in Figure 2.

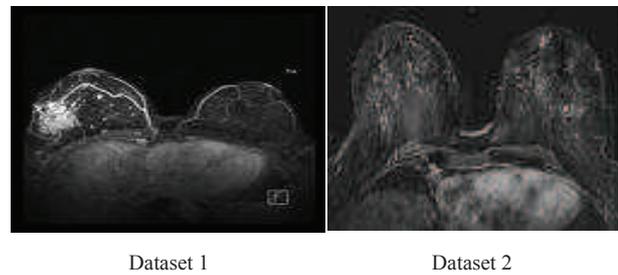


Figure 2: Two different breast MRI datasets

Latest developments in medical image investigation normally comprise processes to divide an image into minute regions in terms of various parameters. These processes address various aspects to analyse the images into regions, which are pathologically and anatomically having an important influence. The study of regions of interest in pathological and physiological sense can be made simpler by categorising the regions using their multi-parameter values. The multi-parameter values are for the edges (E), local contrast (H), and gray values (G) of the pixels.

The objects’ borders can be dogged by its edge information, which is primarily used to obtain the resemblance criterion for any pre-determined object. The reduction of edge count for tumor images is due to the fact that the substance has been pushed aside and compressed by the growth of the tumor within a confined space. This is accomplished by the algorithms’ execution. The corresponding results are depicted in Figure 3.

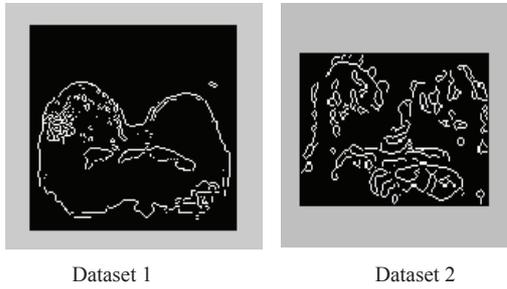


Figure 3: Edge images of two different breast MRI datasets

The threshold value of the binary image can be used to manage the gray parameter (G) for every slab of the breast MRI image sequences. The threshold value was determined by calculating the pixel concentration for each slice, which provides the foundation to analyse medical MRI images of breast tumor patients. The related results are demonstrated in Figure 4.

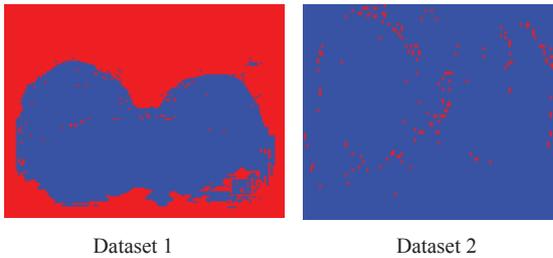


Figure 4: Binary images of two different breast MRI datasets

The amount of difference in pixel concentration is described by using the contrast (H). The differences in the cases of sturdy variation can be explored by a computational programme using watershed segmentation. The malignant tumor cells are illustrated as the intense portion in the MRI images of the breast. These cells comprise a highly proteinaceous fluid (Gupta & Undril, 1995; Haney et al., 2001). Moreover, various contrast segments can be separated out using watershed segmentation by carrying out the algorithm in MATLAB simulation. The results are shown in Figure 5.

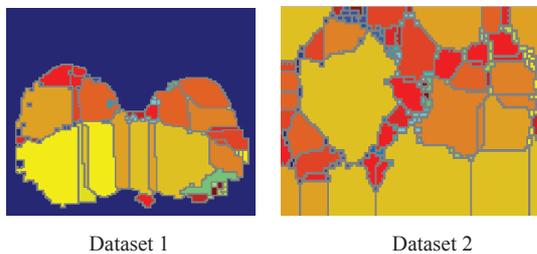


Figure 5: Contrast images of two different breast MRI datasets using watershed segmentation

The above results demonstrate that watershed segmentation splits the high contrast images from the MRI images. The segmentation of the tumor from the MRI images is our primary concern. The tumor is segmented after applying the ROI command in MATLAB. The simulation results are shown in Figure 6.

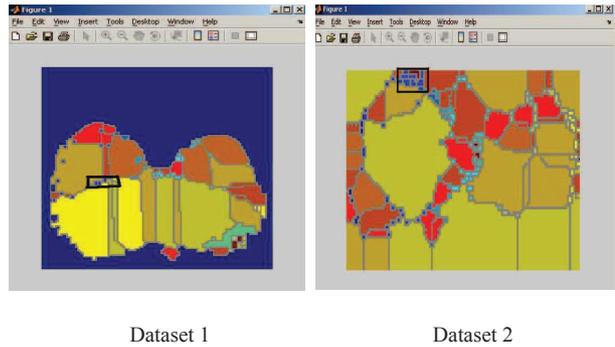


Figure 6: Constructed image after application of ROI command on two different breast MRI datasets

After the application of ROI command, the illustration obtained has been depicted in Figure 7.

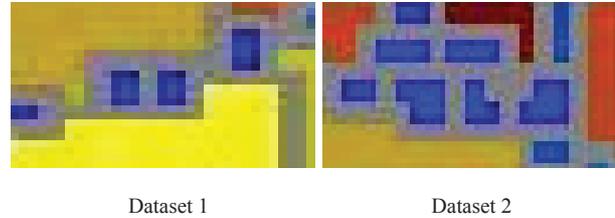


Figure 7: Enhanced image of the area after application of the ROI command of two different breast MRI datasets

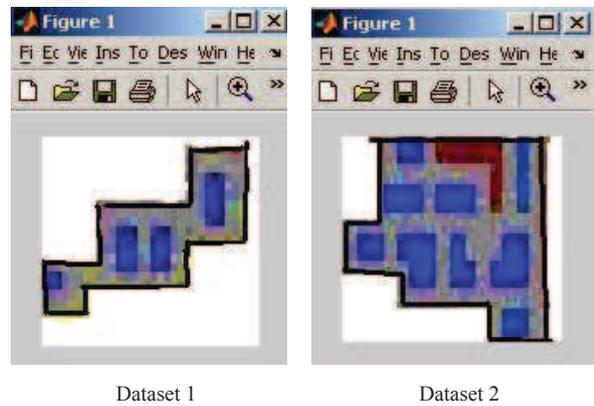


Figure 8: Segmented tumor image-2D of two different breast MRI datasets

Now, the tumor can be segmented by applying the watershed segmentation to Figure 7. The proposed algorithm can be executed to accomplish the aforesaid task, and the results are shown in Figure 8.

RESULTS AND DISCUSSION

Figure 8 depicts the two dimensional result, which has been represented in the shape of pixels in X and Y directions. Since the setting of display is 1028×768 pixels on the screen of the monitor whose dimensions are $280 \text{ mm} \times 210 \text{ mm}$, one pixel's dimension appears to be $0.2734 \text{ mm} \times 0.2734 \text{ mm}$.

Observing the extracted tumor from various positions, the tumor's dimensions in the first dataset appears to be comprised three layers, and the tumor's dimensions in the second dataset appears to be comprised four layers. The dimensions of the various layers for dataset 1 are tabulated (Table 1).

Table 1: Dimensions of the different layers for dataset 1

Serial no.	Name of layer	Dimensions (max) (pixels)	Dimensions (max) (mm,mm)	App. area (mm ²)
1	1 st Layer	27×7	7.38×1.91	14.09
2	2 nd Layer	28×12	7.66×3.28	25.12
3	3 rd Layer	9×7	2.46×1.91	4.69
			Total	43.9

Hence, the whole estimated area of the tumor is 43.9 mm^2 for dataset 1.

The dimensions of the various layers for dataset 2 have been tabulated (Table 2).

Table 2: Dimensions of the different layers for dataset 2

Serial no.	Name of layer	Dimensions (max) (pixels)	Dimensions (max) (mm,mm)	App. area (mm ²)
1	1 st Layer	24×12	6.56×3.28	21.52
2	2 nd Layer	29×7	7.93×1.91	15.15
3	3 rd Layer	21×4	5.74×1.09	6.26
4	4 th Layer	8×6	2.19×1.64	3.59
			Total	46.52

Hence, the whole estimated area of the tumor is 46.52 mm^2 for dataset 2.

CONCLUSION

It is clear from the results that if the parameters are appropriately set, a tumor can be successfully segmented by using watershed segmentation. The initialisation inside the tumor is not required in this method, whereas it is required in other methods. The effectiveness of this method can be demonstrated by visualising and quantitatively evaluating the segmentation results. When the contrast level is significantly different between the non-tumor and tumor regions, the performance of the watershed segmentation algorithm is superior. Non-homogenous tumors can be segmented by this algorithm, which provides the non-homogeneity within the tumor region. It is manifested in this paper that automatic segmentation of the breast tumor can be achieved by the proposed methodology. The segmentation quality is far better than the manual segmentation procedure. It also accelerates the segmentation operation in operative imaging.

Multi-parameter segmentation is found to be the best segmentation method among the various segmentation methods explored. The starting positions from the pointing contrivance can be collected and transferred to the segmentation section by extending the user interface in the main application, which in turn allows activation of the segmentation. The segmented image must be received by the main programme, and finally this programme presents the image as an opaque volume.

REFERENCES

1. Abbasi S. & Mokhtarian F. (2001). Affine-similar shape retrieval: application to multi-view 3-D object recognition. *IEEE Transactions on Image Processing* **10**(1): 131 – 139. DOI: <http://dx.doi.org/10.1109/83.892449>
2. Baillet J.W., Abemayor E., Jabour B.A., Hawkins R.A., Hoh C.K. & Ward P.H. (1992). Positron emission tomography: a new precise modality for detection of primary head and neck tumor and assessment of cervical adenopathy. *Laryngoscope* **102**(3): 281 – 288. DOI: <http://dx.doi.org/10.1288/00005537-199203000-00010>
3. Black P.M., Moriarty T., Alexander E., Stieg P., Woodard E.J., Gleason P.L., Martin C.H., Kikinis R., Schwartz R.B. & Jolesz F.A. (1997). Development and implementation of intraoperative magnetic resonance imaging and its neurosurgical applications. *Neurosurgery* **41**(4): 831 – 845. DOI: <http://dx.doi.org/10.1097/00006123-199710000-00013>
4. Bountris P., Farantatos E. & Apostolou N. (2005). Advanced image analysis tool development for the early stage bronchial cancer detection. *Proceedings of the World Academy of Science, Engineering and Technology (PWASET)* **9**: 151 – 156.

5. Deng Y., Manjunath B.S., Kenney C., Moore M.S. & Shin H. (2001). An efficient colour representation for image retrieval. *IEEE Transactions on Image Processing* **10**(1): 140 – 147.
DOI: <http://dx.doi.org/10.1109/83.892450>
6. Dewalle V.A., Betrouni N. & Makni N. (2008). A new method based on both fuzzy set and possibility theories for tumor volume segmentation on PET images. *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 20 – 24 August, Vancouver Convention and Exhibition Centre, Vancouver, British Columbia, Canada, pp. 3122 – 3125.
7. Fadi A.A. & Abdel Q. (2007). Detection of breast cancer using independent component analysis. *Proceedings of the IEEE International Conference on Electro/Information Technology*, 17 – 20 May, Chicago, USA, pp. 236 – 241.
8. Gering D., Eric W., Grimson L. & Kikinis R. (2002). Recognizing deviations from normalcy for brain tumor segmentation. *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 25 – 28 September, Tokyo, Japan, pp. 388 – 395.
DOI: http://dx.doi.org/10.1007/3-540-45786-0_48
9. Gupta R. & Undril P. (1995). The use of texture analysis to delineate suspicious masses in mammography. *Physics in Medicine and Biology* **40**(5): 835 – 855.
DOI: <http://dx.doi.org/10.1088/0031-9155/40/5/009>
10. Haney S.M., Thompson P.M., Cloughesy T.F., Alger J.R. & Toga A.W. (2001). Tracking tumor growth rates in patients with malignant gliomas: a test of two algorithms. *American Journal of Neuroradiology* **22**(1): 73 – 82.
11. Hinz M., Pohle R., Shin H. & Tonnie K.D. (2002). Region-based interactive 3D image analysis of structures in medical data by hybrid rendering. *Proceedings of the SPIE on Medical Imaging: Visualization, Image-Guided Procedures and Display*, volume 4681, 23 February, San Diego, California, USA, pp. 388 – 395.
DOI: <http://dx.doi.org/10.1117/12.466942>
12. Jason J.C., Eitan S. & Shishir D. (2008). Efficient multilevel breast tumor segmentation with integrated Bayesian model classification. *IEEE Transactions on Medical Imaging* **27**(5): 629 – 640.
DOI: <http://dx.doi.org/10.1109/TMI.2007.912817>
13. Jolesz F.A. & Blumenfeld S.M. (1994). Interventional use of magnetic resonance imaging. *Magnetic Resonance Quarterly* **10**: 85 – 96.
14. Krishnan K. & Atkins M.S. (1998). Segmentation of multiple sclerosis lesions in MRI – an image analysis approach. *Proceedings of the SPIE Conference on Image Processing* **3338**: 1106 – 1116.
DOI: <http://dx.doi.org/10.1117/12.310837>
15. Lei Z. & Chan A.K. (2001). An artificial intelligent algorithm for tumor detection in screening mammogram. *IEEE Transactions on Medical Imaging* **20**(7): 559 – 567.
DOI: <http://dx.doi.org/10.1109/42.932741>
16. Lihua L., Weidong X., Zuobao W., Angela S. & Claudia B. (2007). A new computerised method for missed cancer detection in screening mammography. *Proceedings of the IEEE International Conference on Integration Technology*, 20 – 24 March, Shenzhen, China, pp. 21 – 25.
17. Nacim B., Phillipe P. & Anne S.D. (2007). 3D automatic segmentation and reconstruction of prostate on MRI images. *Proceedings of the IEEE 29th Annual International Conference of Engineering in Medicine and Biology Society*, 22 – 26 August 2007, Lyon, France, pp. 5259 – 5262.
18. Ratan R., Sharma S. & Sharma S.K. (2009a). Brain tumor detection based on multi-parameter MRI image analysis. *ICGST-GVIP Journal* **9**(3): 1 – 7.
19. Ratan R., Sharma S. & Sharma S.K. (2009b). Multi-parameter segmentation and quantization of brain tumor from MRI images. *ISEE-Indian Journal of Science and Technology* **2**(2): 11 – 15.
20. Senaratne G.G., Keam R.B., Sweatman W.L. & Wake G.C. (2006). Solution to the 2 dimensional boundary value problem for microwave breast tumor detection. *IEEE Microwave and Wireless Component Letters* **16**(10): 343 – 348.
DOI: <http://dx.doi.org/10.1109/LMWC.2006.882413>