Breast ultrasound automated ROI segmentation with region growing

CONFERENCE PAPER · AUGUST 2015

Siau-Chuin Liew
Universiti Malaysia Pahang

Breast ultrasound automated ROI segmentation with region growing

Lay-Khoon Lee
Kolej Kediaman 1, Lebuhraya Tun Razak , 26300 Gambang, Kuantan, Pahang Darul Makmur, University Malaysia Pahang.
jesscefyn@hotmail.com

Siau-Chuin Liew
Faculty of Computer System and Software Engineering, Lebuhraya Tun Razak , 26300 Gambang, Kuantan , Pahang Darul Makmur, University Malaysia Pahang.
liewsc@ump.edu.my.

***Abstract*—Image segmentation is an important technology used in different areas ranging from image processing to image analysis. One of the simplest methods for image segmentation that is widely implemented in medical images is the region growing method. Current researches mostly focus on using the region growing method to automatically detect the presence of tumor in MRI (Magnetic Resonance) images instead of ultrasound images. In this paper, we present an algorithm to automatically detect tumors in ultrasound images. Inspired by SergeBeucher and Balasubramanian's road segmentation algorithm, this paper will implement the road segmentation algorithm into medical image segmentation. Results show that, the road segmentation algorithm actually works on the segmentation of medical image. The dice coefficient was used to evaluate the accuracy of the algorithm, eventually getting a value of 0.988 ± 0.00147 as the mean and standard deviation. This value is significant, because the higher the DC value, the more accurate is the segmentation. Besides that, the DC value can use for future reference and comparison.**

Keywords—breast ultrasound; combination; dice coefficient; medical image; segmentation; tumor detection

I. Introduction

Breast cancer is the most common cancer type amongst women in Malaysian. Woman in Malaysia has 1/20 chance of getting breast cancer in her lifetime. According to the National Cancer Registry [1], the leading type of cancers among the general population in Malaysia was breast cancer with 17.8% in 2008. It is essential that early detection of possible breast cancer is done to save lives. Tumors can be classified into two categories: benign and malignant. Benign is non-cancerous whereby malignant is cancerous. Benign tumors are not injurious to health; their cells have close resemblance to normal in appearance, they grow relatively slowly, and do not invade nearby tissues or spread to other parts of the body. On the other hand, malignant tumors are injurious to health. Without a proper diagnosis and treatment, malignant cells eventually can cause human death. The segmentation in ultrasound is important because segmentation is the first stage in any attempt to analyze or interpret an image. An auto-detection of any tumor using segmentation technique may assist physicians in image diagnosis. However segmentation in ultrasound remains to be a challenge due to the poor quality of ultrasound images which are also affected by speckle noise and anatomical structures in the human body. These problems can directly affect the result of the segmentation. Moreover, there are a lot of image segmentation techniques, such as histogram thresholding, edge detection neural network, etc. Each technique has its own disadvantages and advantages. Some of the image segmentation methods may not be suitable implement in the case of ultrasound images. Besides that, there is no universal algorithm for image segmentation. Thus, it is still a challenge for researchers.

There appears to be a gap between the knowledge of region growing road segmentation and ultrasound tumors segmentation. The purpose of this paper is to propose a combined algorithm used in road segmentation algorithm and seed region growing into medical image segmentation, which may lead to the auto-detection of tumor.

II. Proposed Method

The basic idea of the design is based on the works of SergeBeucher [2] and Balasubramanian [3]. The section below shows the flow of experimental process.

A. Pseudo code of modified region growing:

Begin
1. Input image
2. Read image
3. Get divided image
4. Generate ROI
5. While (segmentation is failed)
   5.1 preprocessing (speckle detection and noise reduction)
   5.2 Background separation
   5.3 Region growing implementation

End while
6. Tumor segmented

Firstly ultrasound image is used as an input to perform the analysis. The system will check whether the input is a valid image then it is followed by the next procedure, which is the preprocessing step. Preprocessing starts by dividing the medical image equally into 9 segments, after that, the image will undergo noise filtering, speckle reduction and image preprocessing involves techniques such as noise reduction, contrast enhancement and image sharpening. Both input and output are images. In image segmentation, regions of interest are extracted from the image. Usually, in feature extraction and pattern classification, the inputs are images and the outputs are data (like features of segmented objects) obtained from the images. The pseudocode is then continued by the seed selection process. Seed selection plays a vital role in the process. The nearer the seeds that are selected, the segmentation process will be more accurate and faster. The algorithm will check whether if the tumor is successfully segmented, if the algorithm fails to segment the tumor in ultrasound, the system will go back to previous stage until tumor segmented. The figure 1 below shows the overall flow of region growing segmentation.

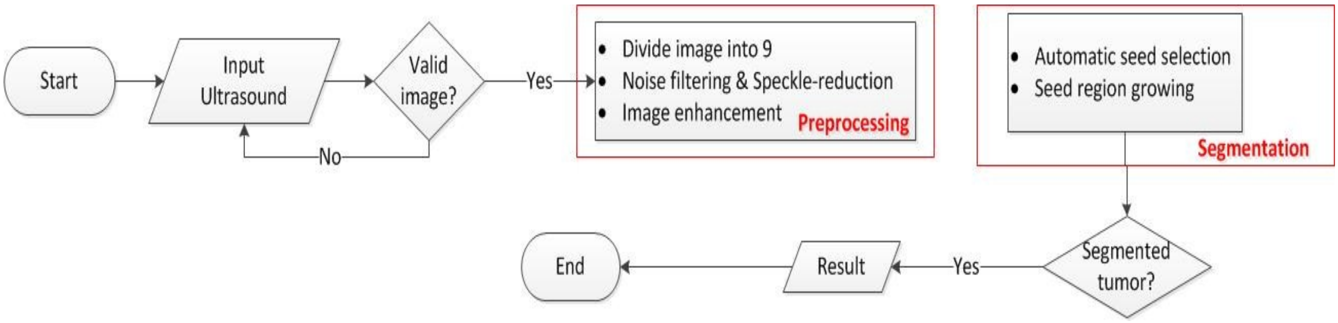

Fig.1. Flow of segmentation

B. *PREPROCESSING STEP*

Median filtering is the chosen preprocessing method to be implemented in the experiment. This is due to Median filtering can remove noise in the medical image, even though it is slow to compute, but it preserves edges of ROI. Median Filtering is the image processing process that will be conducted before ROI is generated. The objective of image preprocessing is to enhance the image and to reduce speckle without destroying the important features of breast ultrasound images for diagnosis. The following is the equation used in the function adpmedian.p , provided by the digital image[4].

Z_{min} = minimum gray level value S_{xy} (1)
Z_{max}= minimum gray level value in S_{xy} (1a)
Z_{med}= median of gray level values in S_{xy} (2)
Zxy= gray level of coordinates (x,y) (3)
S_{max}= maximum allows size of S_{xy} (4)

Level A:

$$A1= Z_{med} - Z_{min} \quad (5)$$
$$A2 = Z_{med} - Z_{maz} \quad (6)$$

If A1 > 0 AND A2 < 0, Go to level ß, Else increase the window size, if window size ≤ S_{xy} repeat level A. Else output zxy

Level A:

$$B1=Z_{xy\,-}\,Z_{min} \quad (7)$$
$$B2 = Z_{xy} - Z_{max} \quad (8)$$

If Ba> 0 AND B2 < 0 output $Z_{xZ.}$ *Else output* Z_{med}

The equation above shows how the preprocessing phase actually works. The output from the preprocessing phase will be taken into the following process which is the seed region growing phase. An ultrasound image has inhomogeneity, noise, and other factors which affect the continuity and accuracy of the images segmentation results. Therefore, the anisotropic diffusion filter which was introduced by Ima et.al will be used to reduce the image noise [5]. This filtering will reduce the image noise while preserving the region edges, and also enhancing the edges by smoothing along them.

B. Generate ROI

As mentioned before, the input of the medical image will be divided into 9 equal parts and converted to gray scale. A seed point is chosen as the starting point for region growing and its selection is very important for the segmentation result. If a seed point is selected outside the region of interests (ROIs), the final segmentation result would definitely be incorrect. Due to the low quality of ultrasound images, most of the region growing methods require the seed point be selected manually in advance. In order to make the region growing segmentation fully automatic, it is necessary to develop an automatic and accurate seed point selection method for US images. In this paper, we implement automatic seed point selection method for breast US images by following this rule [6].

$$X_{seed} = (X_{min} + X_{max}) / 2 \quad (9)$$
$$Y_{seed} - (\equiv y \,|\,) (X_{seed} - Y) \text{ e lesion region} \quad (10)$$

C. Overview of Segmentation

Segmentation is a process that divides the image into non-overlapping regions, and it separates the objects (lesions) from the background The boundaries of the lesions are delineated for feature extraction. Extraction and selection is to find a feature set of breast cancer lesions that can accurately distinguish lesion/non-lesion or benign/malignant. The feature space could be very large and complex, so extracting and selecting the most effective features is very important.

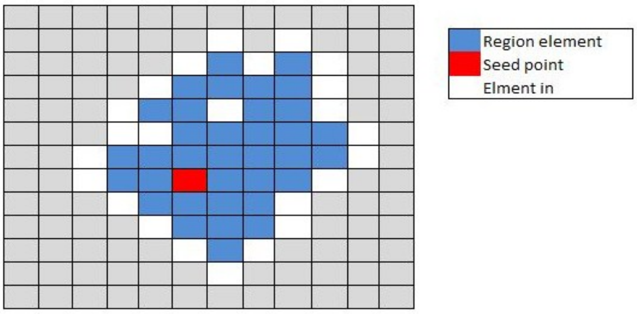

Fig.2. Region growing seed selection

Pseudocode of seed selection:
1. Define seed point
2.Add n-neighbors to list L
3.Get and remove top of
4. Test n-neighbors p
if p not treated
if P (p, R) = True, then p→ Land add p to region
else p marked boundary
5. Go to 2 until L is empty two regions R and ¬ R.

The figure 2 shows how region growing is implemented in the experiment. In Pixel Density Measurement, the image consists of pixels. Each pixel's density is measured among neighboring pixels and itself. A pixel and its neighboring pixels having intensity value are equal, the pixel is highly dense and its value would be zeros. The calculation will stop after the entire pixel surrounded is zero in value. Simple segmentation rule has the form:

$$P(R) : I\ (r,c) < T \qquad (11)$$

A simple approach to image segmentation is to start from some pixels (seeds) representing distinct image regions and to grow them, until they cover the entire image. For region growing we need a rule describing the growth mechanism and a rule for checking the homogeneity of the regions after each growth step.

$$P(Ri\ (k)\ U\ (X)) = TRUE \text{ is valid} \qquad (12)$$

The growth mechanism –at each stage K and for each region $Ri\ (k), i = 1, \ldots, N$, we check if there are unclassified pixels in the 8 neighbourhood of each pixel of the region border. Before assigning such a pixel x to a region $Ri(k)$, we check if the region homogeneity:

$$X = I = 1, \ldots N\ U\ R\ (i) \qquad (13)$$

Refer to the equation (15), it shows that the segmentation must be complete; that is, every pixel must be in a region.

$$R\ (i)\ U\ R\ (j) = 0 \text{ for } 1 \neq 1 \qquad (14)$$

Required that points in a region must connected in some predefines sense like the equation (16) above.

$$P\ (R(i)) = TRUE \text{ for } 1,2\ldots\ N \qquad (15)$$

The equation (17) indicates that the regions must be disjoint.

$$P\ (R\ (i)\ U\ R\ (j)) = FALSE \text{ for } i \neq j \qquad (16)$$

Where the equation of 18 indicates that $P(R(i) \cup R(j))$ get a false statement, due to Ri and Rj will be different in the sense of predicate P. Image enhancement or improving the visual quality of a digital image can be subjective. Saying that one method provides a better quality image could vary from person to person. For this reason, it is necessary to establish quantitative/empirical measures to compare the effects of image enhancement algorithms on image quality. The result of image enhancement will have a direct effect on the segmentation result. Therefore the preprocessing phase is very crucial for medical image segmentation. The results of segmentation are presented in the next section.

III. Experiment Result

A total of 30 ultrasound images were evaluated. There are two types of ultrasound breast tumor images attached in this paper, which is malignant tumor breast and healthy normal breast.

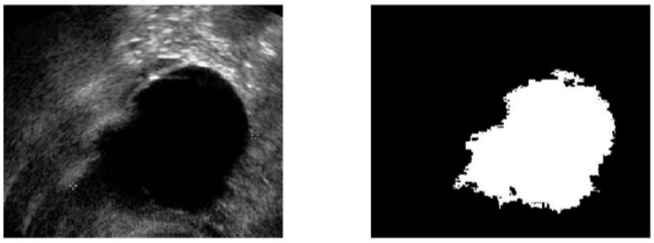

Fig.3. Malignant breast ultrasound Case I

The medical image in figure 3 is the first medical image examined in the experiment, it takes 8.722 second to complete the process, and the dice coefficient is 0.9882, thus it is a successful segmentation.

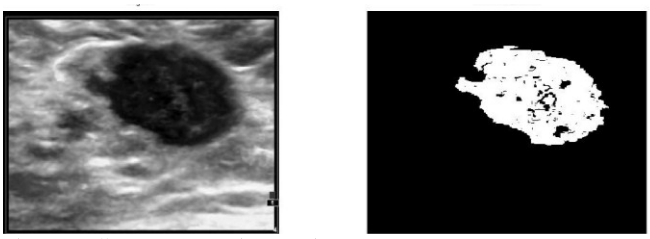

Fig.4. Malignant Breast Ultrasound Case II

This figure 4 above is the second medical image examined in the experiment, it takes 8.344 second to complete the process, and the dice coefficient is 0.9811.

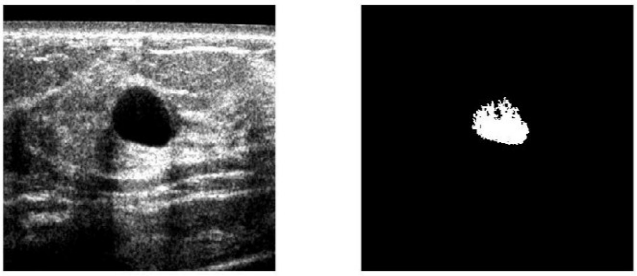

Fig.5. Malignant Breast Ultrasound III

Figure 5 above, is another malignant breast ultrasound examined in the experiment, it takes 8.722 second to complete the process, and the dice coefficient is 0.9882.The next two figure is example of benign tumor of breast ultrasound.

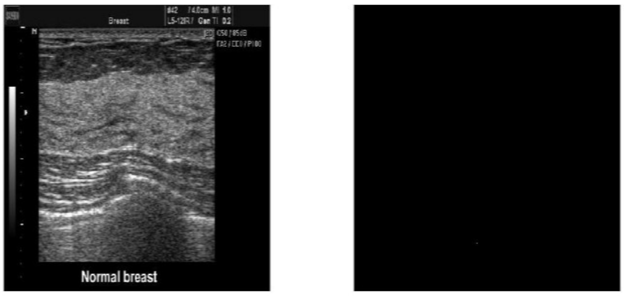

Fig.6. Normal Breast Ultrasound II

The medical image in figure 6 above is first benign breast ultrasound medical image examined in the experiment, it takes 1.606 second to complete the process, and the dice coefficient is 0.9883.

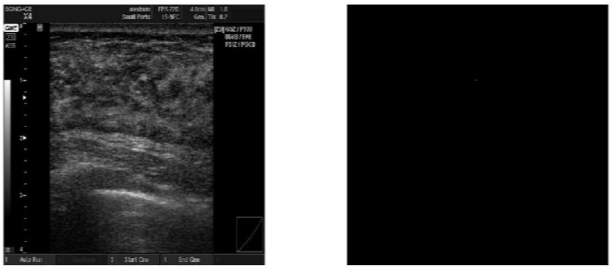

Fig.7. Normal Breast Ultrasound III

The medical image in figure 7 is the medical image examined in the experiment, it takes 1.486 second to complete the process, and the dice coefficient is 0.9886.

In the experiment, different types of breast ultrasound were examined. There are several types used in the experiment, which are breast cancer, breast abscess, breast fibroadenoma (benign), breast ducts and breast carcinoma (malignant) and normal breast ultrasound. Besides normal breast, the rest of the breast ultrasounds were malignant breast type, which is the type that requires treatment.

IV. Discussion

The primary aim of this study is to determine road algorithm can work in medical image. The table below indicated the overall result of the 30 breast ultrasound images.

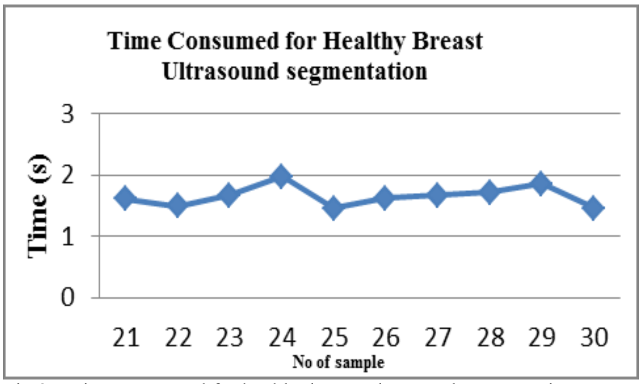

Fig.8. Time consumed for healthy breast ultrasound segmentation

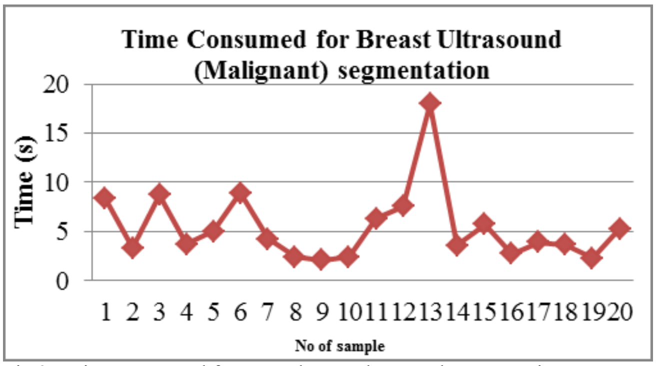

Fig.9. Time consumed for tumor breast ultrasound segmentation

There are a total of 20 malignant breast ultrasounds and 10 normal breast ultrasounds being tested in this experiment. Based on the above experimental result, malignant breast tumors need more time to be segmented which is about 5.4 second; whereby benign tumors need 1.651 second to analysis the medical image. The figure 8 and 9 show the time consumed for both type tumor segmentation. The numbers of samples are corresponding to table 1 which will be presented later. For malignant tumor segmentation, the time used to perform the segmentation is not the longest time used at 18 second. While for benign tumor segmentation, there appears to be more consistency as the time used to segment the image is around 1.5 to 2 second.

The table 1 below shows the result of segmentation time with malignant and benign ultrasound. The result indicates segmentation of benign tumor consumed less time than malignant breast ultrasound. It is because the benign breast ultrasound has nearer neighbor pixel, thus segmentation easier to compute in benign ultrasound images.

TABLE 1. AVERAGE TIME USED AND DC VALUE OF THE RESULT

No	Type of imaging	Average of Time used (s)	Average of DC Value
1.	Malignant Breast Ultrasound	3.34s	0.9453
2.	Benign Breast Ultrasound	2.478s	0.9532
	Maximum	**7.237**	**9.52**
	Minimum	**1.49**	**1.04**

A. Evaluation of Segmentation Performance

Dice coefficient will be used in the experiment as a performance metrics and recently was implemented in MRI segmentation by Belma Dogdas at el, in the year of 2005 [7]. Dice coefficient (DC) is a type of metrics used to measure the accuracy of the region growing algorithm segmentation, it is also a type of statistic used for comparing the similarity of two samples. According to Pratibha's publication in 2013, [8] dice coefficient was independently developed by the Thorvald Sørensen [7] and Lee Raymond [9] who published his work in the year of 1948 and 1945 respectively. The work presents a methodology to model DC as a function of object shapes, sizes, contrasts, noise levels and filters.

Equation of 19 and 20 is the equation that implemented in Matlab. The value of Dice coefficient ranges from 0 to 1. These metrics are commonly used to measure and their values range between 0 (no overlap) and 1 (perfect agreement), which means if the value of Dice coefficient is closer to 1, the more similar the image before and after segmentation.

$$D = 2*\frac{|A \cap B|}{|A| + |B|} \quad (17)$$

Or;

$$D (A , B) = 2((A \text{ intersect } B)/(|A| \ + B|)) \text{ ; [9]} \quad (18)$$

Similarity metric is the basic measurement used by a number of data mining algorithms. It is used to measure similarity between data objects. These objects may have one or more than one attributes related to them. The DSC value is a simple and useful summary measure of spatial overlap, which can be applied to studies of reproducibility and accuracy in image segmentation.

B. Comparison with similar project

This section associates the performance of latest project and author work. Total 4 similar researches have been chosen to conduct a comparison between author's works. According to Eva Kollorz research in 2011, typical clinical evaluation includes the manual and approximate measurement in two section planes in order to obtain an estimate of the nodule's size. Thus they proposed power watershed. In the experiment, they try to input different seeds to evaluate the potential of the applied algorithm. Lastly they achieve 0.81 of Dice coefficient [10].

The second research chosen to compare is Benjamin's work. In 2012, he presented an algorithm based in 3D statistical shape model to segment the fetal cerebellum on 3D ultrasound volume [11]. In the experiments, he tested 20 ultrasound images in between 18 and 24 gestational weeks. Eventually the experiment was obtained mean of dice coefficient in 0.8.

The third chosen research to be compared is done by Hsien-Chi et al research in 2014 [12]. They presented and assessed a method for the three-dimensional (3-D) segmentation of breast masses on dedicated breast computed tomography (bCT) and automated 3-D breast ultrasound images. The proposed segmentation method is enhanced from earlier segmentation method for masses on contrast- enhanced bCT, including two steps, which is initial contour estimation and active contour-based segmentation to further evolve and refine the initial contour by adding a local energy term to the level-set equation. Only ultrasound result will be taken for benchmarking purpose. According to Hsien, segmentation performance was measured in terms of Dice coefficients (DICE) for 98 lesions on 3-D breast ultrasound (US) images. For 3-D breast US, the DICE value was 0.71. Hence, the proposed method acquired promising results for the 3-D imaging modalities, providing a good basic foundation for further quantitative image analysis and possible future expansion to other 3-D imaging modalities. Segmentation of breast masses on dedicated three-dimensional breast ultrasound image.

The last research to be compared is Cerrolaza's work in 2014. In the paper the author proposed a new segmentation method for 3D ultrasound images of the pediatric kidney [13]. Based on the popular active shape models, the algorithm is tailored to deal with the particular challenges raised by US images. Firstly, a weighted statistical shape model allows compensating the image variation with the propagation direction of the US wave front. Second, an orientation correction approach is used to create a Gabor-based appearance model for each landmark at different scales. This multi-scale characteristic is incorporated into the segmentation algorithm, creating a hierarchical approach where different appearance models are considered as the segmentation process evolves. The performance of the algorithm was evaluated on a dataset of 14 cases, both healthy and pathological, in time attained 0.85 as the average of Dice coefficient. The table below shows the comparison of the methods and result.

TABLE 2: Comparison of segmentation result with related publications

Author	Eva Kollorx et al	Benjamin.G.B et al	Hsien-chi et al	Cerrolaza et al	LK lee et al (Author)
Year of publication	2011	2012	2014	2014	-
Method	Watershed	Nelder-Mead simplex algorithm	Active contour-based segmentation	Gabor-based appearance models	Region growing & road segmentation
Modality	Ultrasound	Ultrasound	Breast Ultrasound and Breast CT scan	3D ultrasound	Breast Ultrasound
Dice coefficient	0.81	0.8	0.71	0.85	0.988 ± 0.00147

V. CONCLUSION

In this paper, the authors combined road segmentation algorithm with region growing. The reason for this is that we found out that road segmentation algorithm actually can be applied onto medical image .The table 2 above indicates result with author and other researcher, it shows that our method is efficient and can be run in real time. Author used dice coefficient to measure the accuracy of segmentation. Result shows that the purposed method has higher DC value compare to the previous research.

Ultrasound images have been widely used for the diagnosis and treatment of breast cancer. A lot of work has been done to investigate the automatic segmentation of breast tumors. This design paper combined two different ideas algorithm in road segmentation and algorithm in medical image segmentation to automatically extract breast tumor in an ultrasound image. Results show that this newly proposed design is more convenient compared to manual segmentation. The method combines speckle reduction filter and a prior knowledge.

This purpose of this paper is to show clearly how an algorithm is being implemented in region growing segmentation in ultrasound breast tumor. The expected result of the research is to produce an enhanced algorithm that allows auto detection of tumor in breast ultrasound images.

Acknowledgment

We would like to express our gratitude to the Research and Innovation Department of Universiti Malaysia Pahang for the financial support provided for the research work and also Dr. Ranga for providing the materials used in the experiment.

REFERENCE

[1] S. Beucher, & M. Bilodeau, “Road segmentation and obstacle detection by a fast watershed transformation,” In Intelligent Vehicles' 94 Symposium, Proceedings of the IEEE, pp. 296-301. , October, 1994.

[2] Balasubramanian, C., Ramasamy, K., Srinivasagan, K. G., & Ramakrishnan, K, “Automatic Segmentation of Brain Tumor from MR Image Using Region”, Growing Technique. Journal of Pattern Recognition & Image Processing, 4th ed., vol. 3, pp. 273-282, 2013.

[3] L. Zadeh” Fuzzy sets” Inform. Control, 8 (1965), pp. 338–353

[4] J.C. Bezdek, “Pattern Recognition with Fuzzy Objective Function Algorithms” Kluwer Academic Publishers, Norwell, MA, USA, 1981.

[5] Scott, I.M., Cootes, T.F., Taylor ,C.J., “Improving appearance model matching using local image structure,” In: Proc, IPMI.LNCS, vol.2732, Springer, 2003.

[6] Sørensen, T. "A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons". Kongelige Danske Videnskabernes Selskab vol. 5, issue. 4, pp.1–34, 1948.

[7] Pratibha “Performance Evaluation of Skull Stripping Methods and Tools” International Journal of Computer Trends and Technology (IJCTT) – vol. 4 Issue 9, Sep 2013.

[8] Dice, Lee R. "Measures of the Amount of Ecologic Association between Species". Ecology 26th ed. Vol. 3, pp. 297–302. Doi: 10.2307/1932409. JSTOR 1932409, 1945.

[9] E. Kollorz, E. Angelopoulou, M. Beck, D. Schmidt, D., & T. Kuwert, “Using Power Watersheds to Segment Benign Thyroid Nodules in Ultrasound Image Data. In Bildverarbeitung für die Medizin 2011, Springer Berlin Heidelberg, pp. 124-128, 2011.

[10] Gutiérrez-Becker, B., Cosío, F. A., Huerta, M. E. G., Benavides-Serralde, J. A., Camargo-Marín, L., & Bañuelos, V. M. “Automatic segmentation of the fetal cerebellum on ultrasound volumes, using a 3D statistical shape model.” Medical & biological engineering & computing, vol.51, issues.9, pp. 1021-1030, 2013.

[11] Kuo, H. C., Giger, M. L., Reiser, I., Drukker, K., Boone, J. M., Lindfors, K. K., & Sennett, C. A. “Segmentation of breast masses on dedicated breast computed tomography and three-dimensional breast ultrasound images.” Journal of Medical Imaging, vol.1, issue.1, pp.14501-14502, 2014.

[12] Cerrolaza, J. J., Safdar, N., Peters, C. A., Myers, E., Jago, J., & Linguraru, M. G., “Segmentation of kidney in 3D-ultrasound images using Gabor-based appearance models.” In Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium, pp. 633-636. April 2014