Tumor Delineation in PET to generate Gross Tumor Volume- Comparison and Analysis of Four Image Segmentation Methods

[1]Christy Esther S, [2]Devakumar Devadhas

[1]Assistant Professor, [2]Associate Professor
[1]Department of Biomedical Engineering, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India
[2]Department of Nuclear Medicine, Christian Medical College, Vellore, India

***Abstract*: Numerous image segmentation methods have been developed for the segmentation of tumor in PET. In this article, four image segmentation methods have been compared and evaluated to determine the near accurate method. A series of phantom studies were performed, to determine an most accurate and uniformly applicable method to define the Gross tumor volume (GTV) with PET. The obtained results were affirmative; hence the methods were performed on 10 patient data for accurate delineation of GTV in PET images. Four image segmentation techniques 1) 25% threshold, 2) threshold based Schaefer method, 3) statistical segmentation method, and 4) region growing method, were first tested by a phantom study to determine the GTV in PET images. The methods were then evaluated using patient data in which the segmentation results were compared with the Gross tumor volume (GTV) obtained by manual method. All the four methods showed good results for the phantom study. On the other hand, clinical studies show that, the 25% threshold method gave good segmentation results. The statistical method and region growing method gave consistent good results with a dice similarity coefficient about 96%, while the Schaefer method underestimated the tumor volume.**

***Index Terms*: PET. Image Segmentation. Tumor delineation. Gross tumor volume. Radiotherapy treatment planning.**

I. Introduction

Positron Emission Tomography (PET) images are best suited for quantification and radiotherapy treatment planning due to its sensitivity and specificity [1]. Many studies have proven that the PET images can provide better visibility of the actual tumor volume and the tumor nodes, which is sometimes not accurate in CT [3]. Also the reduction of inter and intraobserver variability is obtained by using PET images for delineation of Gross tumor volume [4-5].
Accurate segmentation of tumor in PET is the critical part in radiotherapy treatment planning. This is mainly due to the low spatial resolution and high noise characteristics of the PET images [6]. Manual Contouring is the usual method used for delineation of tumor in PET images [3]. But the results are found to be less accurate, less precise and are not reproducible. Hence, fast and better segmentation tools are needed for the accurate generation of tumor volume; especially in case of noisy PET images [5, 7]. Phantom studies are generally done before the segmentation algorithms are used on the patient data in order to validate them. Evaluation of the methods has to be done in specific configurations and a comprehensive comparison of different segmentation methods for a wide range of cases should be done. The purpose of this study was to evaluate the performance of four segmentation methods. The segmentation methods were performed using phantom data and 10 patient PET data. An elaborate analysis of the segmentation methods and results were also presented.

II. Type Style and Fonts Materials and Methods

Image Segmentation methods

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts. True Type 1 or Open Type fonts are required. Please embed all fonts, in particular symbol fonts, as well, for math, etc.

Segmentation is the process of separating an image into regions with similar kind of properties like intensity, gray level, texture etc. There are four different categories of PET image segmentation methods like thresholding methods, variational approaches, stochastic modelling-based techniques, and learning methods [6, 11]. In this study, we used 25% threshold method, threshold based Schaefer method, statistical segmentation method and region growing method. We used 3D Slicer and Plastimatch, to perform segmentation by different methods and DICOM RT Export respectively [14-17].

25% Threshold: Threshold segmentation is the simplest method used for PET image segmentation, in which a particular threshold value (T) was chosen to separate the tumor region from the noisy background. The threshold selection depends on the tumor volume. The general condition for simple thresholding is

$$f(x,y) = \begin{cases} 1 \text{ if } g(x,y) > T \\ 0 \text{ if } g(x,y) \leq T \end{cases}$$

Where f(x,y) is the threshold image, T is the threshold value and g(x,y) is the gray level image pixels.

Many different threshold techniques are derived by authors. A 25% threshold method is a percentage threshold method, in which the lower threshold for the segmentation was set at 25% of the maximum intensity and upper threshold, was set at the maximum intensity [10]. The region which lies between lower and upper threshold levels were segmented into tumor volume. This method is more suitable for PET segmentation. Likewise, the value of lower threshold was set at 50% or 40% of the maximum intensity, which are the 50% and 40% threshold methods respectively. The threshold value when set automatically, based on the gray levels and different regions in the image is known as the automatic threshold method.

Threshold based- Schaefer method: This method is based on a contrast oriented algorithm. A particular threshold value cannot be fixed for every PET image. Hence this method was formulated to determine the threshold value, based on the background intensity and the FDG accumulation in the tumor. This was formulated by Schaefer et al, in which the mean standardised uptake value (SUV) of the region of interest which is encircled by a 70% isocontour (mSUV70) was used to represent the Fluro-deoxy glucose (FDG) accumulation of each tumor [2].. The threshold was calculated from the formula

$$TS = a \times mI70 + b \times BKG$$

where TS is the estimated threshold value. mI70 is the mean intensity incurred by adjusting the lower threshold value as 70% of maximum intensity (Imax) . BKG is the mean value obtained by drawing an irregular region of interest manually in the background nearby to the tumor. The value of the constants a and b was taken as 0.3 and 0.7 respectively. The estimated threshold value (TS) was given as the lower threshold and the segmentation was carried out [2].

Statistical segmentation method: This method is used to depict the object features, which are adaptively learned by the label map or seeds given by the user. Following that, some active contours develop at the same time. The contour interactions are governed by principles of mechanics [18]. It is a semi-automatic method because a user defined seed or label map has to be given to differentiate each region separately. Then the leakage issues and the contour for all region of interest are taken care. The contours will not overlap with one another which give the advantage of multi object segmentation.

Region growing method: This method segments the region of interest depending on the statistics of the image [9]. This method uses the important fact that pixels which are close together have similar gray values. A seed point when placed on the pixel will grow and extend to the neighbouring pixels with same gray level, thus the segmentation is takes place. It is also a semi automatic method because a user defined seed point starts the algorithm and no leakage issues found even in case of small tumors [9].

III. Phantom Study

A 6 litre cylindrical phantom was filled with 7Mbq 18FDG activity equivalent to the background activity in clinical situation. Two tube containers of 50 ml volume filled with 37MBq 18FDG activity mixed with water were placed inside the cylindrical phantom. To help identifying in the CT scan, one container additionally contained iodine contrast. A CT scan and a PET scan of the phantom were obtained sequentially. A Siemens PET CT scanner (Biograph 6 TP) was used for image acquisition. The PET data was reconstructed with attenuation and scatter compensation. The reconstructed phantom image was used for the study. Each segmentation method was performed on the phantom image. The Gross tumor volume (GTV) was delineated from phantom images using the different segmentation methods and volumes were obtained. The segmentation result was analyzed whether it yields a volume matching that of the known tube volume. Table 1 show the volume comparison of the different segmentation methods with the manual method. The threshold method and threshold based Schaefer method gave accurate segmentation with volumes of 49.3mL and 49.29mL respectively, which is almost equal to the original tube volume (50mL). While statistical segmentation method, over segments the region of interest. Fig 1 shows the example segmentation for the phantom data.

Fig 1. (a) Phantom image before segmentation; (b) Segmented phantom image.

Table 1
Volume comparison for phantom studies

Volume comparison for phantom studies		
S.No	**Method**	**Volume**
1	Original volume	50
2	25% Threshold	49.3
3	Schaefer	49.29
4	Statistical segmentation	51.5
5	Region growing	49.6

IV. Clinical studies

The tumor region is not constrained to a particular region of the body, a diverse range of data was used for this study like tumor in brain, lungs, rectum, and bladder were chosen for this study. Ten patient PET images were acquired. All four image segmentation methods were performed on the patient data retrospectively. Fig 2 shows the segmentation output for the first patient data.

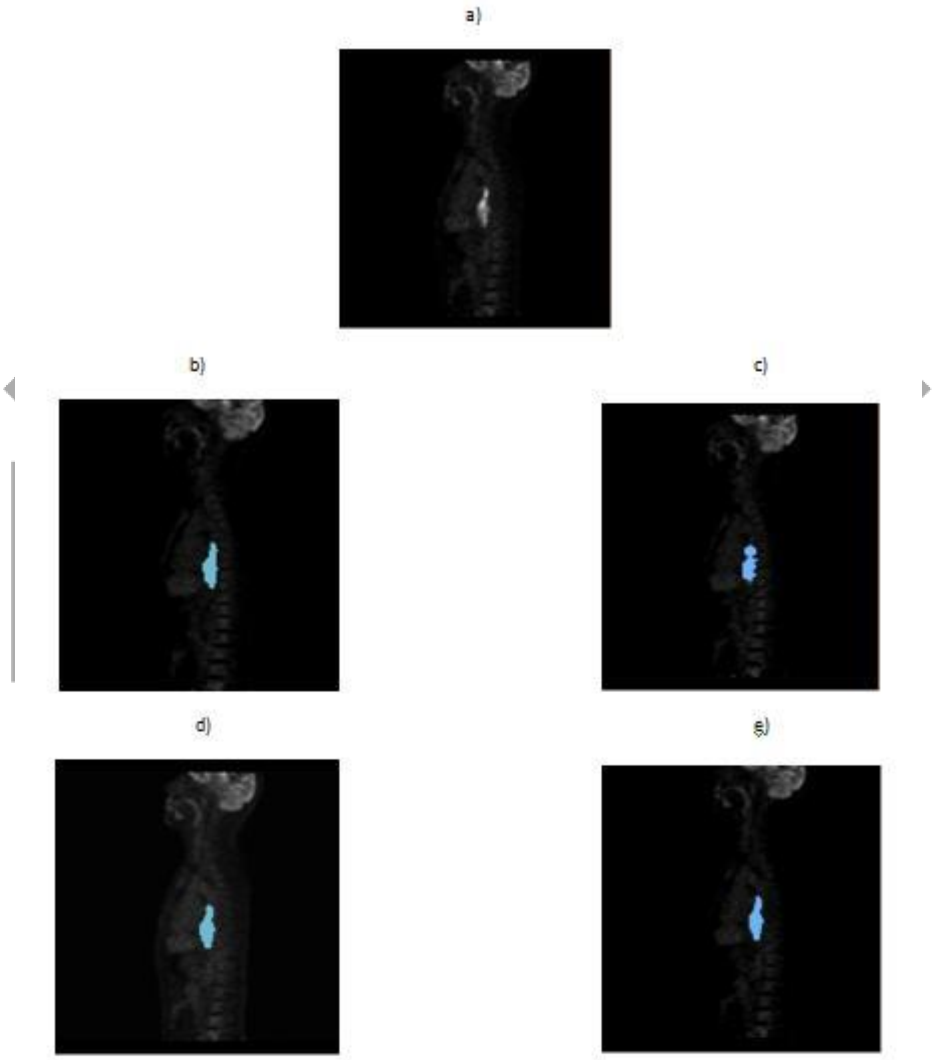

Fig 2. Patient segmentation- a) image before segmentation, b) by threshold method, c) by statistical segmentation method, d) by Schaefer method, e) by region growing method

V. EVALUATION METHODS

To evaluate the accuracy of the segmentation methods, three different parameters were used they are 1) Volume comparison, 2) Dice similarity coefficient and 3) Mean Percentage Error

Firstly in volume comparison, we compared the volume by segmentation methods with the manually obtained tumor volume.

Table 2
Volume comparison for clinical studies

Volume comparison for clinical studies					
Patient Number	M_1	M_2	M_3	M_4	M_5
1	14.8	13.1	15.8	12.8	10.6
2	42.07	28.5	32.7	24	36
3	3.98	1.5	4.5	3.4	1.6
4	60.6	50.7	60.4	58.9	60.1
5	31.56	29.9	36.6	27.8	25.8
6	43.76	40.2	43.8	37.3	28.2
7	61.14	61.8	69.1	55.15	47.5
8	65.1	61.4	50.9	35.9	51.2
9	17.4	14	14.7	10.7	12
10	5.24	4.24	2.95	8.8	6.4

* M_1 – Manual method, M2 – 25% Threshold method, M3 – Statistical segmentation method, M4 – Schaefer method, M5 – Region growing method

Dice similarity coefficient is a statistic used to compare the similarity between two measurements. In this study we use this quantity to measure the similarity between the volumes by segmentation method with the volume by manual method. The similarity index values would range from zero to one. 0 indicates that there is no similarity of manual method with segmentation method and 1 indicates that its exactly similar [12-13]. The formula used to calculate the Dice similarity coefficient is

Dice similarity coefficient = $2(A*B) / (A^2+B^2)$
Where A - Volume by manual method, B - volume by segmentation method

Mean percentage error (MPE), a parameter, used in this study to quantify the mean difference of the volume by manual method and the volume by segmentation method for a particular data.

$$\% \text{ Mean Percentage Error (MPE)} = \frac{1}{n}\sum_{i=1}^{n} \frac{A_i - B_i}{A_i} \times 100$$

where n – total number of patients, i – patient number, A – volume by manual method, B – volume by segmentation method. The method which gives the lowest mean percentage error when compared with others was considered as a better method.

VI. RESULTS AND DISCUSSION

Fig 3 shows the volume comparison of the segmentation with the manual method. The analysis of the comparison shows that the 25% threshold method, region growing and statistical segmentation method gave good results when compared with manual contouring method. The Schaefer method underestimates the region of interest in some patient data. The same problem was found even when this method is used for the phantom studies. The Schaefer method gave accurate output for the phantom studies. But its parameters have to be optimized every time when it is to be used for the patient data.

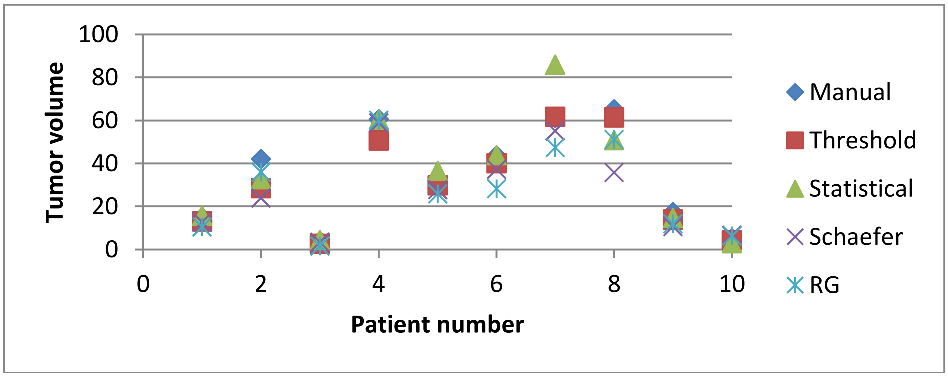

Fig 3. Graph showing the volume comparison of all segmentation methods with the manual method

Next, table 3 shows the dice similarity coefficient values calculated for each segmentation method. Results show that, the statistical segmentation method and region growing method gave an average of 96% for the dice similarity coefficient. The threshold method gave an average of 94.5% and Schaefer method was about 90%.

Table 3
Dice similarity coefficient

Dice Similarity Coefficient				
Patient Number	**DSC_1**	**DSC_2**	**DSC_3**	**DSC_4**
1	0.99	0.99	0.98	0.94
2	0.92	0.96	0.86	0.98
3	0.66	0.99	0.69	0.98
4	0.98	0.99	0.99	0.99
5	0.99	0.98	0.99	0.99
6	0.99	0.99	0.98	0.91
7	0.99	0.94	0.99	0.96
8	0.99	0.93	0.84	0.97
9	0.97	0.99	0.89	0.93
10	0.97	0.85	0.87	0.98

* DSC_1 – Dice similarity coefficient for threshold method, DSC_2 – Dice similarity coefficient for statistical segmentation method, DSC_3 – Dice similarity coefficient for Schaefer method, DSC_4 – Dice similarity coefficient for Region growing method

Table 4 Mean Percentage Error

Mean Percentage Error		
S.No	***Segmentation methods***	***Mean Percentage Error (%)***
1	25% Threshold	5.85
2	Region growing	7.11
3	Statistical segmentation	8
4	Schaefer	13

VII. CONCLUSION

The present study shows that the 25% threshold method gave good results when compared with the manual contouring. The region growing method and statistical segmentation method also showed consistent segmentation, and found better in case of visual interpretation and dice similarity coefficient. The exercise of comparing the segmentation techniques helped us to gain some insight on the highlights & pitfalls of automatic segmentation methods. Irrespective of the segmentation technique used, interventions are required. Optimization of input parameters for the automatic techniques may require the user to do some initial learning. But after that, using the automatic methods is definitely the ideal approach. So even though in this study we considered our base method to be manual contouring, it is ideally not suitable as it represents a very long process and also suffers from low reproducibility. 25% threshold and region growing methods show promising results. Their output can also be used well as the input for further hybrid methods for better volume delineations.

REFERENCES

[1] Arnold Paulino, PET-CT in Radiotherapy Treatment Planning: Elsevier, 2008

[2] Andrea Schaefer, Stephanie Kremp, Dirk Hellwig, Christian Rübe, Carl-Martin Kirsch, Ursula Nestle, "A contrast-oriented algorithm for FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer: derivation from phantom measurements and validation in patient data," Eur J Nucl Med Mol Imaging (2008) 35:1989–1999.

[3] Y.J.Zhang. A Survey on evaluation methods for image segmentation volume, Computer Vision and Pattern Recognition, vol. 29, No.8, pp. 1335–1346, 1996.

[4] A. Gregory Sorensen., Shveta Patel., Carla Harmath., Sarah Bridges., Jennifer Synnott., Amy Sievers., et. Al., Comparison of Diameter and Perimeter Methods for Tumor Volume Calculation, Journal of Clinical Oncology, Vol 19, No 2, 2001: pp 551-557.

[5] Jeremy J. Erasmus., Gregory W. Gladish., Lyle Broemeling., Bradley S. Sabloff., Mylene T. Truong., Roy S. Herbst., and Reginald F. Munden. Interobserver and Intraobserver Variability in Measurement of Non–Small-Cell Carcinoma Lung Lesions: Implications for Assessment of Tumor Response, Journal of Clinical Oncology, Vol 21, No 13, 2003: pp 2574-2582.

[6] Habib Zaidi & Issam El Naqa. PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques, European Journal of Nuclear Medicine Molecular Imaging, 2010, 37:2165–2187.

[7] Patsuree Cheebsumon., Floris H.P. van Velden., Maqsood Yaqub., Virginie Frings., Adrianus J. de Langen., Otto S.Hoekstra., Adriaan A. Lammertsma., and Ronald Boellaard. Effects of Image Characteristics on Performance of Tumor Delineation Methods: A Test–Retest Assessment, The Journal of Nuclear Medicine, 2011, Vol. 52:1550–1558.

[8] Perrine Tylski., Simon Stute., Nicolas Grotus., Kaya Doyeux., S´ebastien Hapdey., Isabelle Gardin., Bruno Vanderlinden., Ir`ene Buvat. Comparative Assessment of Methods for Estimating Tumor Volume and Standardized Uptake Value in 18F-FDG PET, The Journal of Nuclear Medicine, 2010, Vol. 51, No. 2.

[9] Eric Pichon, Allen Tannenbaum, Ron Kikinis. A statistically based flow for image segmentation, Medical Image Analysis 8(3):267–274, September 2004.

[10] Erdi YE, Mawlawi O, Larson SM, Imbriaco M, Yeung H, Finn R. Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. Cancer 1997, 80:2505–9.

[11] Gupta T, Beriwal S. PET/CT-guided radiation therapy planning: From present to the future, Indian J Cancer 2010, 47:126-33.

[12] Hatt M, Cheze Le Rest C, Turzo A, Roux C, Visvikis D. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. IEEE Trans Med Imaging. 2009;28:881–93.

[13] Besl and Jain. Segmentation through variable-order surface fitting. IEEE Trans on Pattern Analysis and Machine Intelligence. 1988, 10:2 167-192.

[14] G Sharp, N Kandasamy, H Singh, M Folkert, "GPU-based streaming architectures for fast cone-beam CT image reconstruction and demons deformable registration, " Physics in Medicine and Biology, 52(19), pp 5771-83,2007.

[15] Pieper S., Lorensen B., Schroeder W., Kikinis R. The NA-MIC Kit: ITK, VTK, Pipelines, Grids and 3D Slicer as an Open Platform for the Medical Image Computing Community. Proceedings of the 3rd IEEE International Symposium on Biomedical Imaging: From Nano to Macro 2006; 1:698-701.

[16] Pieper S., Halle M., Kikinis R. 3D SLICER. Proceedings of the 1st IEEE International Symposium on Biomedical Imaging: From Nano to Macro 2004; 1:632-635.

[17] Gering D.T., Nabavi A., Kikinis R., Grimson W..E.L., Hata N., Everett P., Jolesz F.A., Wells III W.M. An Integrated Visualization System for Surgical Planning and Guidance using Image Fusion and Interventional Imaging. Int Conf Med Image Comput Comput Assist Interv . 1999;2:809-819.

[18] Yi Gao, Allen Tannenbaum, Ron Kikinis. Simultaneous Multi-Object Segmentation using Local Robust Statistics and Contour Interaction, MICCAI 2010, Medical Computer Vision, Volume 6533/2011, 195-203.