An Automatic Method for Metabolic Evaluation of Gamma Knife Treatments

Alessandro Stefano[1,2(✉)], Salvatore Vitabile[3], Giorgio Russo[1], Massimo Ippolito[4], Franco Marletta[4], Corrado D'Arrigo[4], Davide D'Urso[1], Maria Gabriella Sabini[4], Orazio Gambino[2], Roberto Pirrone[2], Edoardo Ardizzone[2], and Maria Carla Gilardi[1]

[1] Institute of Molecular Bioimaging and Physiology,
National Research Council (IBFM-CNR) - LATO, Cefalù, PA, Italy
`alessandro.stefano@ibfm.cnr.it`
[2] Department of Chemical, Management, Information Technology and
Mechanical Engineering, University of Palermo, Palermo, Italy
[3] Department of Biopathology and Medical Biotechnologies (DIBIMED),
University of Palermo, Palermo, Italy
[4] Cannizzaro Hospital, Catania, Italy

Abstract. Lesion volume delineation of Positron Emission Tomography images is challenging because of the low spatial resolution and high noise level. Aim of this work is the development of an operator independent segmentation method of metabolic images. For this purpose, an algorithm for the biological tumor volume delineation based on random walks on graphs has been used. Twenty-four cerebral tumors are segmented to evaluate the functional follow-up after Gamma Knife radiotherapy treatment. Experimental results show that the segmentation algorithm is accurate and has real-time performance. In addition, it can reflect metabolic changes useful to evaluate radiotherapy response in treated patients.

Keywords: Segmentation · Random walk · PET imaging · Gamma Knife treatment · Biological target volume

1 Introduction

Gamma Knife (LGK; Elekta) radiosurgery is a mini-invasive technique defined as the delivery of a single, high dose of radiation to obtain a complete destruction of brain lesions. It provides a safe and effective way of treating inaccessible cerebral tumors. An examination to differentiate malign and benign tissue in brain tumors with great preciseness is the Positron Emission Tomography (PET) with the amino acid tracer 11c-methionin (MET). PET is a non-invasive functional imaging technique that shows complementary information with respect to Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). In addition, metabolic changes are often faster and more indicative of the effects of the therapy with respect to anatomical imaging [1]. Numerous studies have shown that the specificity of the MET PET for marking tumor delineation and for the differentiation relapse versus radiation necrosis is higher

compared with MRI. In the paper reported in [2], metabolic imaging was used for biological target delineation in 36 patients that showed a significantly longer median survival compared with the group of patients, in which target volume was merely defined by MRI. Nevertheless, due to the nature of PET images (low spatial resolution, high noise and weak boundary), the Metabolic Tumor Volume (MTV) varies substantially depending on the algorithm used to delineate functional lesions: the choice of a standard method for PET volume contouring is a very challenging yet unresolved step.

Visual delineation is widely used because it is easily applicable, but it is potentially inaccurate being susceptible to the window level settings and subject to both intra and inter-operator variability. In this paper an algorithm based on random walks (RW) on graphs [3] has been adapted for PET imaging. To create an automatic and operator independent method starting from previous work [4], we propose an automated seed localization method to identify RW seeds. Our framework is used in 12 patients with cervical metastases to evaluate therapeutic response in sequential scans for a total of 24 PET scans. Patients undergo MET PET examinations before and 2 months after the Gamma Knife treatment.

The paper is organized as follows: in the next section the current state of the art in PET image segmentation techniques is reviewed. In the "Materials and Methods" section, the RW algorithm adapted for PET imaging and for a clinical environment is described. In the "Results" section, our delineation method is evaluated to assess the accuracy and the Gamma Knife treatment response. We conclude with a discussion of results in the last section.

2   Related Works

Image thresholding and region growing methods are the most widely used due to their simplicity to implement but they are too sensitive to PET image noise and heterogeneity [5, 6]. Variational approaches based on gradient differences between target and background regions are mathematically efficient but sensitive to noise and subject to numerical fluctuation [7, 8]. Learning methods as artificial neural network (ANN), support vector machine (SVM), k-means algorithm, fuzzy C-means algorithm are efficient but require high computational steps and are sensitive to variability of PET radiotracer depending on study protocol, as for example scanner characteristics, radiotracer injected dose and interval between radiotracer injection and exam start, e.g. [9, 10]. In addition, supervised algorithms have limited application in PET imaging, unlike in the MRI or CT fields, due to high heterogeneity that makes the recognition of stable features in the training set very difficult. Graph based methods are used to find the globally optimal segmentation of images. The RW algorithm was developed by Grady [3] in the computer vision domain and then was extended for image segmentation [11, 12]. RW treats the segmentation as the solution to a linear system with an exact solution and it is very accurate in noisy and low contrast images, such as PET images. At last, in several works, e.g. [13, 14], the different tumor contours on PET and on CT are simultaneously segmented. Although our patients undergo a PET/CT

scans, CT has strong limitations for target delineation of cervical metastases and this makes impossible to apply a PET/CT delineation in our study.

3   Materials and Methods

3.1   Phantom Studies

National Electrical Manufacturers Association International Electro-technical Commission (NEMA IEC) phantom with six spheres of different diameters (d1= 10 mm, d2 = 13 mm, d3 = 17 mm, d4 = 22 mm, d5 = 26 mm, d6 = 37 mm) is used to estimate the accuracy of the PET segmentation algorithm. Spheres, and background are filled with a ratio between measured sphere radioactivity concentration and measured background radioactivity concentration (S/B) that ranged from 3 to 7.

PET acquisitions are performed on time of flight PET/CT Discovery 690 by General Electric Medical Systems following the same protocol used in clinical routine. PET images have an in-plane resolution of 256 × 256 voxels with 1.1719x 1.1719x3.27 mm3 voxel size. The proposed segmentation method is evaluated by matching the sphere delineation with the ground truth in the CT images.

3.2   Clinical Study

Twelve patients with cervical metastases are enrolled to evaluate the metabolic response to Gamma Knife treatment, for a total of 24 PET scans. The patient fasts for 4 hours before PET exam and is intravenous injected with MET before and 2 months after the Gamma Knife treatment (see Fig.1). The PET oncological protocol begins 10 minutes after the injection.

Fig. 1. Patients with cervical metastases undergo a MET PET examinations before and 2 months after the Gamma Knife treatment

The handmade segmentation is the gold standard for the clinical study. It is used to compare the results obtained with the automatic proposed method. PET/CT studies are reported by an expert nuclear medicine physician for diagnostic and staging purposes. A time-expansive slice-by-slice manual segmentation is performed on a software platform (Xeleris workstation) from General Electric Medical Systems healthcare. Due to inevitable variability of PET image visualization, the window level setting is changed by the nuclear medicine physician to obtain a clear visual appearance of MET PET positive structures. Then, the active tumor volume is defined including the tumor volume with an intense tracer uptake respect to background MET activity level.

3.3   Automatic Segmentation

3.3.1 RW Method in PET Imaging Field

RW is an efficient and accurate method in low contrast images characterized by noise and weak edges such as PET studies [15]. The graph-based segmentation method represents an image as a graph in which the voxels are its nodes and the edges are defined by a cost function which maps a change in image intensity to edge weights [3]. The weights w_{ij} between nodes are obtained using the following Gaussian function:

$$w_{ij} = \exp(-\beta(g_i - g_j)^2) \qquad (1)$$

where both g_i and g_j are the image intensity values at voxels i and j; β is a free parameter (in our experiments, β is set to 1).

RW parameters have been calibrated to be suited for PET imaging modality. In (1) we replace g_i and g_j with the Standardized Uptake Value (SUV) in the voxels i and j. SUV is the ratio of lesion radioactivity concentration, and administered dose at the time of injection divided by body weight [16]. The PET image is converted into a graph where some nodes are labeled by the user and some nodes are not known.

The segmentation problem is to assign a label to unknown nodes. This is done by trying to find the minimum energy among all possible scenarios in the graph to provide an optimal segmentation. The RW method partitions the nodes into foreground and background subsets from the probability that a "random walker" starting at a source node, first reaches a node with a pre-assigned label, visiting every voxel.

The RW problem has the same solution as the combinatorial Dirichlet problem [3]:

$$D[x] = (x^T L \, x) / 2 \qquad (2)$$

where L indicates the graph's Laplacian matrix and x the vector of the probabilities that each voxel is inner to target.

A probability array is then produced, and a threshold of 50% is chosen to discriminate the foreground from the background creating a voxel binary mask, so that:

- target voxel value = 1 if its probability $\geq$ 50%
- background voxel value = 0 if its probability < 50%.

3.3.2 Automatic Detection of RW Seeds

To create an user independent method starting from previous work [4], we propose an automated seeding process of the RW algorithm. The algorithm automatically identifies the PET slice with the highest SUV (SUV_{max}) and the seeds for each slice. The n voxels with a SUV>95% of SUV_{max} are marked as target seeds. Then, the method explores the neighborhood of the voxel with SUV_{max} through searching in 8 directions to identify the background voxels with a SUV<30% of SUV_{max} (see Fig.2). Once the target and background seeds are localized, RW performs a 3D delineation (see Fig.3).
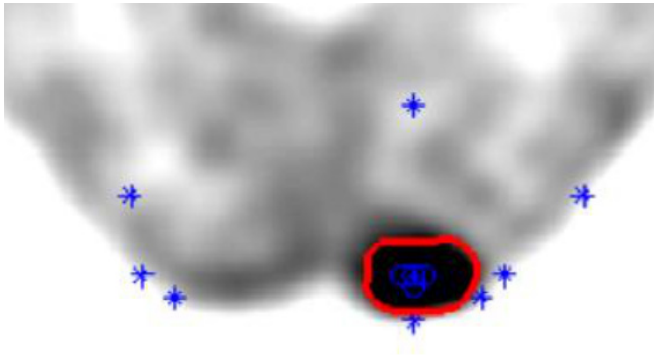

Fig. 2. Background (*) and target seeds (°) automatically identified by our method. After this step, the RW algorithm performs the lesion segmentation (red curve) to discriminate foreground from background voxels
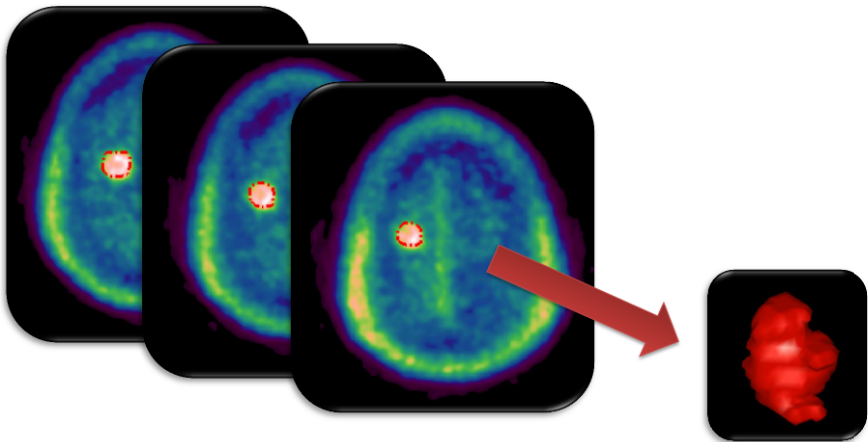

Fig. 3. A metabolic lesion in three different PET slices and the corresponding 3D rendered object derived from RW segmentation are show

3.3.3 Adaptive RW Method

RW method depends on the choice of the β-weighting factor in (1). To overcome this issue (in our experiments, β is set to 1), we change, slice after slice, the probability threshold to discriminate the target from background voxels. Using an adaptive probability, we obtain a method independent of the choice of the β-weighting factor. In fact, β influences how quickly the probability decreases with increasing intensity differences.

The probability map produced by the RW algorithm is processed to obtain an adaptive threshold value (P) using the following method:

- Calculate the mean of the probability values inside a large pre-segmented lesion obtained using a fixed probability of 20%
- Identification of voxels with a probability less than the mean, and voxels with a probability greater than the mean
- Calculate the probability means (P1 and P2) of the two voxel sets
- Calculate P as the mean between P1 and P2

In this way any voxel with less than a P% chance of being in the target is rejected. This method follows the whole lesion volume taking into account the intensity gradient of the PET lesion in different slices. The flow diagram of the proposed method is shown in figure 4.
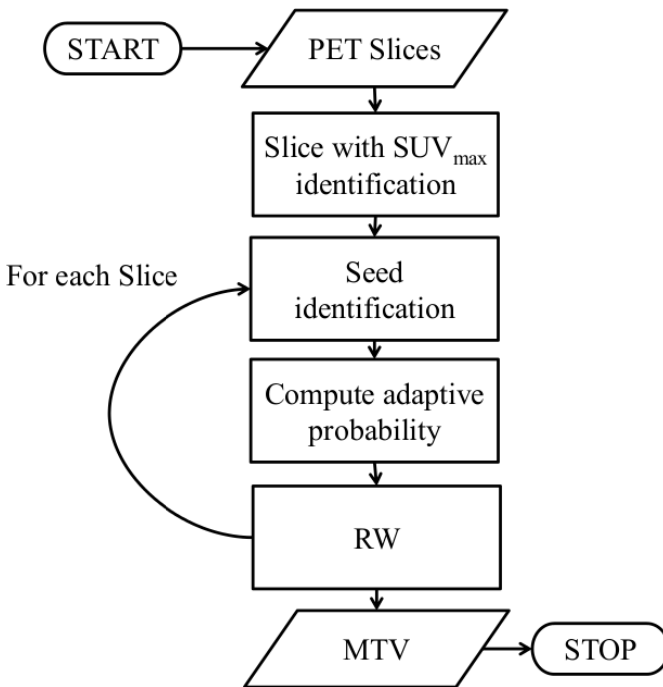

Fig. 4. Flow diagram of the proposed delineation method. The method follows the whole lesion volume taking into account the gradient of intensity and contrast changes of the PET lesion in different slices

3.4   Evaluation

The segmentation performance of the proposed method is evaluated making a comparison with manual MTV segmentation by the dice similarity coefficient (DSC), median Hausdorff distance (HD) and true positive and false positive volume fractions (TPVF and FPVF). DSC is a measurement of spatial overlap between segmented and manual MTV: a DSC of one indicates a perfect match between the two volumetric segmentations, while a DSC of zero indicates no overlap. HD is a shape dissimilarity metric measuring the most mismatched boundary pixels between automatic and manual MTV: a small median of HD values means an accurate segmentation, while a large median of HD values means no accuracy. TPVF indicates the fraction of the total amount of tissue inside the target lesion (sensitivity), and FPVF denotes the amount of tissue falsely identified (specificity=100 - FPVF) [17]. A perfect segmentation algorithm would be 100% sensitive (segmenting all voxels from the target voxels) and 100% specific (not segmenting any from the background voxels).

The average time for delineating brain metastases is recorded to assess algorithm performance.

In addition, lesion segmentation is used to evaluate therapeutic response using SUV_{max}, MTV, and Total Lesion Glycolysis (TLG) variations in sequential scans. MTV provides metabolic volumetric information of the tumors; TLG, defined as MTV x (average SUV within the MTV), combines the volumetric and SUV information, to try to obtain a better evaluation of the treatment response.

Variations (Δ) in SUV_{max}, MTV, and TLG in sequential scans are normalized to baseline:

$$\Delta(\%) = 100 \text{ x (post-treatment – baseline)/baseline} \qquad (3)$$

4   Experimental Results

4.1   Trials and Results on Phantoms

Phantom images are used to validate the proposed method. The size of the spheres is known, and the accuracy of the delineation method can be evaluated. The thresholds to identify target and background voxels in section 3.3.2 are set to 95% and 30% because there give better results among the tested values. The DSC range is found to be from 81.23% up to 99.99% (95.27±2.98%). The HD range is found to be from 1.69mm up to 3.66mm (2.10±0.79mm). The average TPVF is 98.80±2.22% (range: 95.10% - 100%). The average specificity (100 – FPVF) is ~ 100% since the sphere voxel number with respect to the background voxel number is very small.

4.2   Trials and Results on the Clinical study

24 PET studies (2 scans for each patient) are used to assess the accuracy of our method by comparing the automatic to manual segmentation. Fig. 5 shows the lesion segmentation obtained using the two different methods in three patient studies. The RW delineation is not subject to both intra and inter-operator variability.
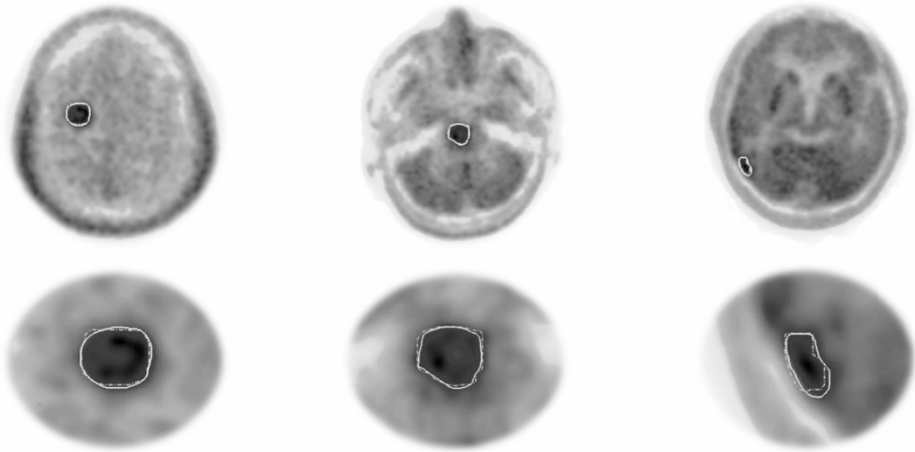
Fig. 5. Three different segmentation examples of uptake regions are shown in each column with a zoom of the lesion. RW (broken line) and manual (continuous line) lesion segmentations are overlaid

The DSC range of PET delineation using the RW algorithm is found to be from 76.35% up to 92.70% with a mean of 87.61±4.47%. The HD range is found to be from 1.00 mm up to 2.85 mm with a mean of 1.86±0.55 mm.

The average TPVF is found to be 89.40±8.09% (range: 71.20% - 100%). The average specificity (100 – FPVF) is ~ 100% since the target voxel number with respect to the background voxel number is very small (a single PET slice consists of 65536 voxels while the largest lesion in a single slice is < 50 voxels).

High DSC and TPVF, and low HD and FPVF values confirm the accuracy of method. An analysis of the time performance shows that our algorithm is fast: the segmentation time for single slice is around 0.3 seconds.

About Gamma Knife treatment response evaluation, results are shown in Table 1. Most patients show the same treatment response trend using SUV_{max}, MTV, and MTV variations in sequential scans. Differently, the third patient shows a minimal change in SUV_{max} (-1.00%) and great variation in the metabolic tumor mass and in TLG (-61.12% and -59.43%, respectively). Patients number 4, and 9 show a positive ΔSUV_{max} (+9.47% and +14.73%, respectively), and a negative ΔTLG, and ΔMTV (-16.89% and -8.50%, +49.98% and +39.74%, respectively).

The sixth patient study shows a positive ΔMTV (+24.85%), and a negative ΔSUV_{max} (-44.36%), and ΔTLG (-26.15%). At last, patients number 8 and 12 show a visual disappearance of metabolic lesions indicating a complete metabolic response (CMR) to treatment.

Table 1. Metabolic response obtained on the segmented lesions for each patient

Patients	ΔSUV_{max}	ΔMTV	ΔTLG
#1	-30.10%	-62.20%	-66.67%
#2	-26.21%	-31.53%	-43.83%
#3	-1.00%	-61.12%	-59.43%
#4	9.47%	-16.89%	-8.50%
#5	-41.83%	-81.16%	-83.86%
#6	-44.36%	24.85%	-26.15%
#7	-62.72%	-20.22%	-64.93%
#8	CMR	CMR	CMR
#9	14.73%	-49.98%	-39.74%
#10	-39.64%	-13.27%	-10.21%
#11	-37.00%	-30.95%	-56.50%
#12	CMR	CMR	CMR

5   Discussions and Future Works

Gamma Knife radiosurgery is a stereotactic treatment defined as the delivery of a single, and high dose of radiation for a precise destruction of target tissues. MRI, the commonly used imaging modality in neuro-radiosurgery, has limitations for the post-operative evaluation. We integrate PET imaging to evaluate the treatment response in 12 patients with cervical metastases. PET has, however, not yet been fully incorporated into routine Gamma Knife procedure: metabolic segmentation is a critical task due to the lack of consistency in tumor contour, low image resolution, and a relatively high level of noise and heterogeneity of uptake within a tumor. Nevertheless, an accurate automatic 3D delineation is desirable because manual segmentation is a prohibitively laborious task. In addition, automated segmentation is also important because of the need for repeatable delineation in patient studies for a proper quantification of therapy response, given the considerable variations within and across nuclear medicine physicians.

In this paper a modified RW method for PET images segmentation has been presented. We propose an extension of our previous method [4] to automatically change the probability value needful to lesion from background area discrimination and to make it fully automatic.

First, our method uses an adaptive probability value instead of a fixed one to make the segmentation performance independent of the choice of the β factor in the Gaussian weighting function (1). Second, the algorithm automatically identifies the RW seeds: our method is an operator independent method, satisfying this critical requirement in a clinical environment. Third, the method accuracy is optimal with high DSC and TPVF values and low HD and FPVF values. At last, the proposed method is very powerful in terms of time performance: the algorithm is fast (one slice in ~ 0.3 seconds) if compared against the time needed for manual segmentation.

In addition, we use our method to calculate MTV and TLG in order to reflect metabolic changes in sequential PET scans after Gamma Knife treatment throughout the entire tumor mass. These parameters should be more accurate methods of detecting global changes than a single-pixel value measurement such as SUV_{max}. The results in Table 1 show that the patient number 3, 4, and 9 could be a demonstration of this theory. On the other hand, in the literature cut-offs for SUV are reported [1, 18], instead there are no data for MTV or TLG evaluations. Analyses of receiver operating characteristic curves to estimate these cut-offs will be evaluated. In addition, MTV and TLG values depend on the delineation processes. Our segmentation approach is fully automatic, and an experienced nuclear medicine physician is considered in the manual lesion definition, to assess our method accuracy. Nevertheless, partial volume effect is one of the most important factors impacting the quality and the quantitative accuracy in PET imaging [19]. The images are blurred due to the limited spatial resolution of PET scanner and small lesions appear larger [20]. Several corrective techniques have been developed and a method of partial volume correction could be included in the algorithm, as that described in [21].

At last, the sixth patient shows a positive ΔMTV, and a negative ΔTLG that is indicative of partial response to treatment. TLG combines the volumetric (MTV) and the semi-quantitative parameter (SUV) information: it probably provides a better evaluation of the prognosis compared to MTV.

In conclusion, PET delineation in neurosurgical radiosurgery appears helpful for assessing the effects of therapy in brain metastases. The developed method could be used as a Medical Decision Support System to help clinicians in treatment response evaluation of oncological patients. Nevertheless, as this is only a preliminary study, further investigations are required with a larger number of patients in order to assess the prognostic usefulness and long-term clinical impact to correlate MTV segmentation with clinical outcomes, progression-free survival and overall survival.

References

1. Wahl, R.L., et al.: From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors. Journal of Nuclear Medicine **50**, 122S–150S (2009)
2. Grosu, A.L., et al.: Reirradiation of recurrent high-grade gliomas using amino acid PET (SPECT)/CT/MRI image fusion to determine gross tumor volume for stereotactic fractionated radiotherapy. International Journal of Radiation Oncology Biology Physics **63**(2), 511–519 (2005)
3. Grady, L.: Random walks for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **28**(11), 1768–1783 (2006)
4. Stefano, A., Vitabile, S., Russo, G., Ippolito, M., Sardina, D., Sabini, M.G., Gallivanone, F., Castiglioni, I., Gilardi, M.C.: A graph-based method for PET image segmentation in radiotherapy planning: a pilot study. In: Petrosino, A. (ed.) ICIAP 2013, Part II. LNCS, vol. 8157, pp. 711–720. Springer, Heidelberg (2013)

5. Jentzen, W., et al.: Segmentation of PET volumes by iterative image thresholding. Journal of Nuclear Medicine **48**(1), 108–114 (2007)
6. Li, H., et al.: A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours. Medical Physics **35**(8), 3711–3721 (2008)
7. Geets, X., et al.: A gradient-based method for segmenting FDG-PET images: methodology and validation. European Journal of Nuclear Medicine and Molecular Imaging **34**(9), 1427–1438 (2007)
8. Wanet, M., et al.: Gradient-based delineation of the primary GTV on FDG-PET in non-small cell lung cancer: A comparison with threshold-based approaches, CT and surgical specimens. Radiotherapy and Oncology **98**(1), 117–125 (2011)
9. Hatt, M., et al.: A Fuzzy Locally Adaptive Bayesian Segmentation Approach for Volume Determination in PET. IEEE Transactions on Medical Imaging **28**(6), 881–893 (2009)
10. Zaidi, H., et al.: Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging. Physics in Medicine and Biology **47**(7), 1143–1160 (2002)
11. Bagci, U., et al.: A Graph-Theoretic Approach for Segmentation of PET Images. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) **2011**, 8479–8482 (2011)
12. Onoma, D.P., et al.: 3D Random walk based segmentation for lung tumor delineation in PET imaging. In: 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1260–1263 (2012)
13. Song, Q., et al.: Optimal Co-Segmentation of Tumor in PET-CT Images With Context Information. IEEE Transactions on Medical Imaging **32**(9), 1685–1697 (2013)
14. Xia, Y., et al.: Dual-modality brain PET-CT image segmentation based on adaptive use of functional and anatomical information. Computerized Medical Imaging and Graphics **36**(1), 47–53 (2012)
15. Zaidi, H., El Naqa, I.: PET-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. European Journal of Nuclear Medicine and Molecular Imaging **37**(11), 2165–2187 (2010)
16. Paquet, N., et al.: Within-patient variability of F-18-FDG: Standardized uptake values in normal tissues. Journal of Nuclear Medicine **45**(5), 784–788 (2004)
17. Udupa, J.K., et al.: A framework for evaluating image segmentation algorithms. Computerized Medical Imaging and Graphics **30**(2), 75–87 (2006)
18. Young, H., et al.: Measurement of clinical and subclinical tumour response using F-18 -fluorodeoxyglucose and positron emission tomography: Review and 1999 EORTC recommendations. European Journal of Cancer **35**(13), 1773–1782 (1999)
19. Soret, M., Bacharach, S.L., Buvat, I.: Partial-volume effect in PET tumor imaging. Journal of Nuclear Medicine **48**(6), 932–945 (2007)
20. Stefano, A., et al.: Metabolic impact of partial volume correction of 18F FDG PET-CT oncological stucies on the assessment of tumor response to treatment. Quarterly Journal of Nuclear Medicine and Molecular Imaging **58**(4), 413–423 (2014)
21. Gallivanone, F., et al.: PVE Correction in PET-CT Whole-Body Oncological Studies From PVE-Affected Images. IEEE Transactions on Nuclear Science **58**(3), 736–747 (2011)