

# ACCURATE AUTOMATIC DELINEATION OF HETEROGENEOUS FUNCTIONAL VOLUMES IN POSITRON EMISSION TOMOGRAPHY FOR ONCOLOGY APPLICATIONS

Citation for published version (APA):

Hatt, M., le Rest, C. C., Descourt, P., Dekker, A., De Ruyscher, D., Oellers, M., Lambin, P., Pradier, O., & Visvikis, D. (2010). ACCURATE AUTOMATIC DELINEATION OF HETEROGENEOUS FUNCTIONAL VOLUMES IN POSITRON EMISSION TOMOGRAPHY FOR ONCOLOGY APPLICATIONS. *International Journal of Radiation Oncology Biology Physics*, 77(1), 301-308.  
<https://doi.org/10.1016/j.ijrobp.2009.08.018>

## Document status and date:

Published: 01/05/2010

## DOI:

[10.1016/j.ijrobp.2009.08.018](https://doi.org/10.1016/j.ijrobp.2009.08.018)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

Download date: 30 Sep. 2023

## PHYSICS CONTRIBUTION

# ACCURATE AUTOMATIC DELINEATION OF HETEROGENEOUS FUNCTIONAL VOLUMES IN POSITRON EMISSION TOMOGRAPHY FOR ONCOLOGY APPLICATIONS

MATHIEU HATT, PH.D.,\* CATHERINE CHEZE LE REST, PH.D., M.D.,\*<sup>†</sup> PATRICE DESCOURT, PH.D.,\*  
ANDRÉ DEKKER, PH.D.,<sup>‡</sup> DIRK DE RUYSSCHER, PH.D., M.D.,<sup>‡</sup> MICHEL OELLERS, PH.D.,<sup>‡</sup>  
PHILIPPE LAMBIN, PH.D., M.D.,<sup>‡</sup> OLIVIER PRADIER, PH.D., M.D.,\*<sup>§</sup> AND DIMITRIS VISVIKIS, PH.D.\*

\*Institut National de la Santé et de la Recherche Médicale U650 Brest, France; <sup>†</sup>Centre Hospitalier Universitaire, Morvan, Brest, France;  
<sup>‡</sup>MAASTrict Radiation Oncology Clinic, Maastricht, The Netherlands; <sup>§</sup>Institute of Oncology, Centre Hospitalier Universitaire  
Morvan, Brest, France

**Purpose:** Accurate contouring of positron emission tomography (PET) functional volumes is now considered crucial in image-guided radiotherapy and other oncology applications because the use of functional imaging allows for biological target definition. In addition, the definition of variable uptake regions within the tumor itself may facilitate dose painting for dosimetry optimization.

**Methods and Materials:** Current state-of-the-art algorithms for functional volume segmentation use adaptive thresholding. We developed an approach called fuzzy locally adaptive Bayesian (FLAB), validated on homogeneous objects, and then improved it by allowing the use of up to three tumor classes for the delineation of inhomogeneous tumors (3-FLAB). Simulated and real tumors with histology data containing homogeneous and heterogeneous activity distributions were used to assess the algorithm's accuracy.

**Results:** The new 3-FLAB algorithm is able to extract the overall tumor from the background tissues and delineate variable uptake regions within the tumors, with higher accuracy and robustness compared with adaptive threshold ( $T_{\text{bckg}}$ ) and fuzzy C-means (FCM). 3-FLAB performed with a mean classification error of less than  $9\% \pm 8\%$  on the simulated tumors, whereas binary-only implementation led to errors of  $15\% \pm 11\%$ .  $T_{\text{bckg}}$  and FCM led to mean errors of  $20\% \pm 12\%$  and  $17\% \pm 14\%$ , respectively. 3-FLAB also led to more robust estimation of the maximum diameters of tumors with histology measurements, with  $<6\%$  standard deviation, whereas binary FLAB,  $T_{\text{bckg}}$  and FCM lead to 10%, 12%, and 13%, respectively.

**Conclusion:** These encouraging results warrant further investigation in future studies that will investigate the impact of 3-FLAB in radiotherapy treatment planning, diagnosis, and therapy response evaluation. © 2010 Elsevier Inc.

Heterogeneous functional volumes delineation, Automatic segmentation, Image-guided radiotherapy, Dose painting.

## INTRODUCTION

Although most clinical applications of positron emission tomography (PET) rely on manual and visual analysis, accurate functional volume delineation in PET is crucial for numerous oncology applications. These include the use of tumor volume and associated determination of semiquantitative indices of activity concentration for diagnosis and therapy response evaluation (1) or the definition of target volumes in intensity-modulated radiation therapy (IMRT) (2). Subjective (1) and tedious manual delineation cannot perform accurate and reproducible segmentation, particularly when considering complex shapes and nonhomogeneous

uptake. This results from the low quality of PET images due to statistical noise and partial volume effects (PVE) (3), arising from the scanner's limited spatial resolution.

Most of the previously proposed methods for PET volume definition are semiautomatic and threshold-based, using either fixed (30%–75% of the maximum activity) (2, 4, 5) or adaptive approaches incorporating the background activity (6–10). Unfortunately, these approaches often require additional *a priori* information and are user- and system-dependent. They require manual background regions of interest (ROIs), and their performance depends on parameters requiring optimization using phantom acquisitions for

Reprint requests to: Mathieu HATT, Ph.D., INSERM U650, LATIM, CHU MORVAN, 5 avenue Foch, 29609 Brest. Tel: (+33) 298018111; Fax: (+33) 298018124; E-mail: [hatt@univ-brest.fr](mailto:hatt@univ-brest.fr)

Conflict of interest: none.

**Acknowledgments**—This work was supported by the Brittany Region grant program (Grant No. 1202-2004), the French National Research Agency (Grant Nos. ANR-06-CIS6-004-03 and ANR-08-ETEC-005-01), and Cancéropôle Grand Ouest (R05014NG).

Received Mar 31, 2009, and in revised form July 9, 2009.  
Accepted for publication Aug 13, 2009.

each scanner and reconstruction. Finally, all of these approaches are strictly binary and were not validated considering heterogeneous volumes.

Numerous works have addressed PET lesion segmentation using more advanced image segmentation methodologies (11–19). However, the majority of these approaches often depend on pre- or postprocessing steps such as deconvolution or denoising, are often binary only, and are validated on phantom acquisitions or clinical data without rigorous ground truth.

We previously developed an algorithm for PET volume definition by combining a fuzzy measure with a locally adaptive Bayesian-based classification (FLAB) that has been shown to perform better with respect to fixed thresholding, fuzzy C-means (FCM), or fuzzy hidden Markov chains (FHMC) for PET volume definition, as far as homogeneous spheres or slightly heterogeneous and nonspherical tumors are concerned (20). Preliminary results show that FLAB is also robust with respect to variability of the acquisition and reconstruction parameters (24).

Clinical tumors may be characterized by heterogeneous uptake, thus demanding a nonbinary approach for an accurate segmentation that may have a significant impact in defining biological target volumes for dose painting (21). The goals of this work were to (1) improve the FLAB model by incorporating the use of three hard classes and three fuzzy transitions and (2) evaluate its accuracy on real (with known diameter measured in histology) and simulated (with known ground truth) data sets containing inhomogeneous tumors.

## METHODS AND MATERIALS

### Three-class fuzzy Bayesian segmentation (3-FLAB)

The 3-FLAB algorithm is an extension of our previous work considering only a binary segmentation (20). FLAB automatically estimates parameters of interest from the image, maximizing the probability of each voxel to belong to one of the considered classes. This probability is estimated for each voxel as a function of its value and the values of its neighbors relative to the voxels' statistical distributions in the image, which corresponds to an estimation of the noise within each class. Hence, each voxel of the volume is considered a random variable within a Bayesian framework:

$$P(X|Y) = \frac{P(X, Y)}{P(Y)} = \frac{P(Y|X)P(X)}{P(Y)}, \quad (1)$$

where  $P(X|Y)$  is the probability of belonging to Class X knowing Observation Y. This probability is obtained by the product of  $P(Y|X)$  and  $P(X)$ , corresponding to the noise model and the spatial model, respectively.  $P(Y|X)$  is estimated considering the statistical distribution of the voxels within each class, whereas  $P(X)$  is estimated using a sliding cube of  $3 \times 3 \times 3$  voxels; hence, each voxel's classification is influenced by its neighbors. The parameters to estimate are the mean and variance of each class and the spatial probabilities of each voxel with respect to its neighbors. This is performed iteratively using a stochastic version (SEM) (25) of the Expectation Maximization (EM) (26) initialized with K-means (27) or fuzzy C-means (28). In addition, a fuzzy measure between the classes was added to account for the blur between regions, assuming each voxel may contain a mixture of classes (22, 23).

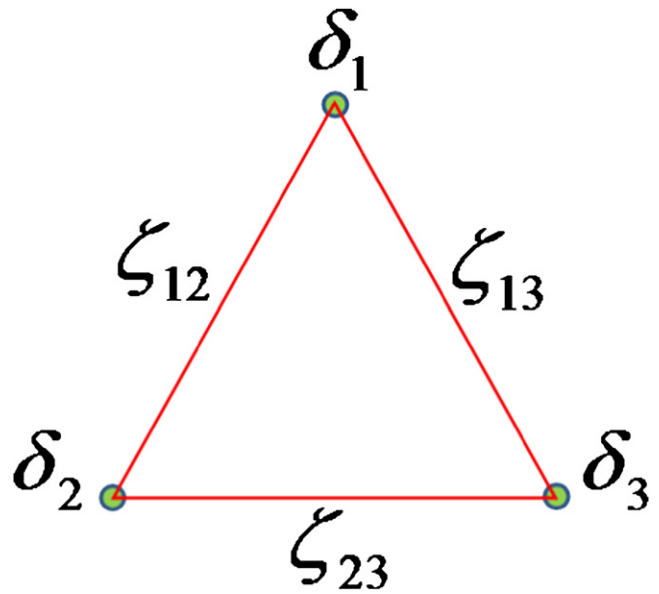


Fig. 1. The fuzzy scheme in the three-class fuzzy locally adaptive Bayesian (3-FLAB) implementation.

The difference between 3-FLAB and the previously developed binary-only FLAB (20) is the use of three classes and three fuzzy transitions within the model (see Fig. 1), to deal with both homogeneous and heterogeneous activity distributions. Figure 2 demonstrates the inability of FLAB to handle highly nonuniform activity distributions, where the lower uptake part of the lesion is erroneously considered as part of the background (see Fig. 2b), emphasizing the need to better model heterogeneous activity distributions. 3-FLAB should retain the accuracy and robustness of the original model, while also being able to handle the challenging heterogeneous activity distributions that are frequently seen in clinical lesions. The 3-FLAB segmentation workflow is summarized as follows, and the implementation and mathematical details can be found in the Appendix.

1. Initialization of both the spatial and noise models parameters: means and variances of each class are obtained using the K-means or fuzzy C-means. The prior probabilities are fixed at one third for each class.
2. Iterative estimation is performed using the SEM by stochastic sampling for each voxel according to its posterior probability.
3. Segmentation is done by selecting for each voxel the class or fuzzy level that maximizes its posterior probability and fusion of fuzzy levels with each hard class to generate a two- or three-class segmentation map.

### Alternative segmentation methodologies used for comparison

We compared the results of the 3-FLAB algorithm with the binary FLAB approach and the fuzzy C-means (with two or three clusters) clustering introduced by Dunn (28) and used to segment PET brain tumors in (13), as well as an adaptive thresholding (6) ( $T_{\text{bckg}}$ ):

$$I_{\text{threshold}} = \alpha \times I_{\text{mean}} + I_{\text{background}}. \quad (2)$$

$I_{\text{mean}}$  was obtained by computing the mean of all voxels contained inside an initial threshold at 70% of the maximum and  $I_{\text{background}}$  by computing the mean of the voxels inside a ROI manually drawn on the background.  $I_{\text{mean}}$  and  $I_{\text{background}}$  were subsequently used to

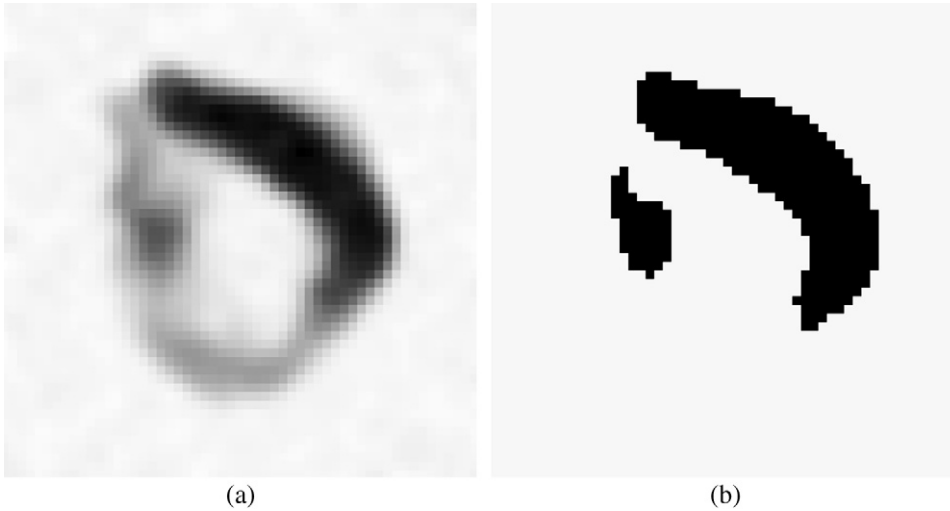


Fig. 2. Binary fuzzy locally adaptive Bayesian (FLAB) model applied to a heterogeneous simulated tumour (a). The segmentation result (b) clearly misses parts of the tumour.

derive a first approximation of the source-to-background contrast. The parameter  $\alpha$  was optimized using phantom acquisitions on each scanner used to obtain the data. The adaptive thresholding algorithm was implemented using a region-growing approach with the maximum intensity voxel as a seed and iteratively adding three-dimensional (3D) neighboring voxels if their value was above the threshold calculated using Eq. 2.

### Validation studies

*Data sets:* Data Set 1 was used to evaluate the performance of the algorithm under realistic imaging conditions. It consists of 20 3D simulated tumors with variable levels of irregular shape and homogeneous or nonhomogeneous uptake distributions derived from tumors in patients undergoing 18F-fluorodeoxyglucose PET/CT investigations for radiotherapy treatment planning purposes. These images were acquired in 2D and 3D mode using the GE Discovery LS and Philips Gemini PET/CT scanners, respectively. Three of these tumors illustrating the range of sizes, shapes, and heterogeneities considered are shown in Fig. 4a–4c. The goal was to produce realistic images of PET tumors while retaining a voxel-based ground truth to compute accurate voxel-based classification errors. Half of the tumors were simulated considering a homogeneous uptake distribution, whereas the other half was simulated using significant heterogeneity within the tumor. The procedure followed to generate these images is illustrated in Fig. 3 and detailed in the following paragraphs.

Each clinical tumor is first manually delineated on the PET image by a nuclear medicine expert, thus creating a voxelized volume that represents the ground truth of the simulation. The activity levels attributed to each of the tumor parts were derived from the average activity measured in the same areas of the tumor in the corresponding patient images. This ground truth tumor structure is subsequently transformed into a nonuniform rational B-splines (NURBS) volume using Rhinoceros (CADLINK software, Morangis, France), for insertion into the NURBS-based CArdiac-Torso (NCAT) phantom (29) attenuation maps at the approximate position where it was located in the patient (30). No respiratory or cardiac motions were considered. Simulations using a model of the Philips PET/CT scanner previously validated with Geant4 Application for Tomography Emission (GATE) (31) were carried out. Forty-five million coincidences were simulated corresponding to the statistics

of a clinical acquisition over a single-axial 18-cm field of view (31). Images were subsequently reconstructed using OPL-EM (seven iterations, one subset) (31) with two voxel sizes ( $4 \times 4 \times 4$  for the Philips Gemini and  $2 \times 2 \times 5$  mm<sup>3</sup> for the GE Discovery LS) to match those used in the corresponding clinical images.

Data Set 2 contains 18 images of lung tumors from patients with histologically proven non-small cell lung cancer (clinical Stage Ib–IIIb), acquired on the Siemens Biograph PET/CT scanner and reconstructed using OSEM (four iterations, eight subsets), with scatter and CT-based attenuation correction, and  $5.31 \times 5.31 \times 3.38$  mm<sup>3</sup> voxels. These tumors were surgically extracted for a histology study in which their maximum diameter was measured by macroscopic examination (32). These diameters range from 15 to 90 mm ( $44 \pm 21$ ). One of these tumors is shown in Fig. 4d.

*Analysis:* Because our goal is not the detection of a lesion in the whole image but the accurate estimation of its volume and shape, we assume it has been detected and isolated by the clinician within a 3D “box” encompassing the tumor.

Because a ground truth was available, classification errors (CE) were computed. In the case of a two-class ground truth, the CE is:

$$CE = \frac{\text{card}\{t|c_t \neq x_t\}}{\text{card}\{t|x_t = 1\}} \times 100, \quad (3)$$

where  $c_t$  is the classification of voxel  $t$ , and  $x_t$  is the true class. Card is the number of elements. This error measurement takes into consideration the spatial distribution of the tumor by considering both background voxels classified as object and object voxels classified as background. Consequently, this measure is more appropriate than simple volume estimation, which could lead to overall small volume errors associated with largely inaccurate segmentations. In addition, the errors are computed relatively to the size of the object, to avoid biases relative to the size of the processing box. In the case of a three-class ground truth, CE may be computed for each of the three classes using Eq. 4 or with respect to a binarized ground truth (second and third class merged) using Eq. 3.

$$CE_c = \frac{\text{card}\{t|x_t = c, c_t \neq c\} + \text{card}\{t|x_t \neq c, c_t = c\}}{\text{card}\{t|x_t = c\}} \times 100, \quad (4)$$

where  $CE_c$  stands for the classification error associated with a given class  $c$ .

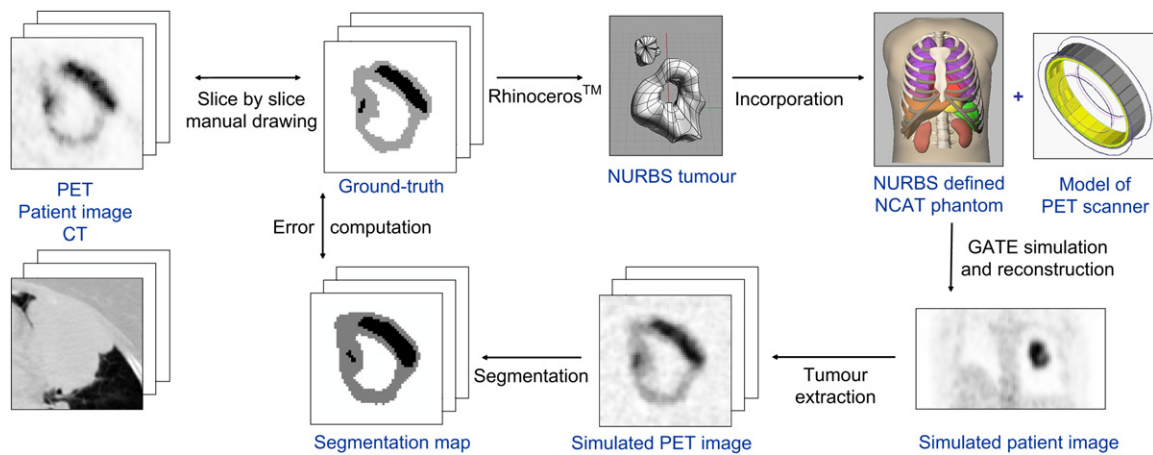


Fig. 3. The simulation of realistic positron emission tomography images.

Two analyses were conducted using Data Set 1. The first considered the entire data set (both homogeneous and heterogeneous tumors) and CE computed using Eq. 3 to compare overall performances of FLAB (binary only), 3-FLAB, FCM, and  $T_{\text{bckg}}$ . The second considered only the 10 heterogeneous tumors to compute  $CE_2$  and  $CE_3$  using Eq. 4 for 3-FLAB and FCM with three clusters.

The segmentation accuracy on the tumors with histology (Data Set 2) was assessed by segmenting the clinical image and subsequently measuring the maximum diameter on the segmented volumes to compare it with the histology measurement.

## RESULTS

Figure 5 contains one axial slice of the segmentations obtained on three simulated tumors of Data Set 1 and one tumor of Data Set 2. Figure 6a contains the mean classification errors and standard deviation obtained by all the methods on the 20 tumors of Data Set 1. FLAB (binary only) performed well on homogeneous tumors but failed as expected on strongly heterogeneous lesions, leading to overall errors of  $15\% \pm 11\%$ . 3-FLAB, in contrast, produced segmentation maps closer to the ground truth, both visually and quantitatively, with errors between 5% and 15% ( $9\% \pm 8\%$ ). FCM (with two or three clusters) was competitive with respect to 3-FLAB for some tumors but showed a higher variability (10%–40%) and mean error ( $20\% \pm 12\%$ ). This translated qualitatively in FCM being unable to differentiate two regions within the tumor as well as being unable to detect discontinuities in the contours (e.g., Fig. 5d, first row). In addition, for the regions where a transition was present between the high uptake region and the background (e.g., Fig. 4d), the 3-FLAB approach was the only one giving accurate representation of this transition (Fig. 5c vs. Fig. 5d, last row).  $T_{\text{bckg}}$  was not able to produce satisfactory segmentation in several cases. Tumors with high overall contrast were approximately extracted from the background (e.g., Fig. 5e, rows 2–4). However, as a binary method, it is unable to delineate uptake distributions within the tumor. In several cases, the heterogeneity was significant, and  $T_{\text{bckg}}$  lead to significant underevaluation of the tumor volume (CE up to 60% with a mean of  $17\% \pm 14\%$ ) because

it tends to extract the high-activity region or parts of the reduced uptake region only (e.g., Fig. 5e, first row).

Figure 6b compares 3-FCM (using three clusters) and 3-FLAB concerning the three-class segmentation of the 10 heterogeneous simulated tumors of Data Set 1. 3-FCM is less accurate and robust compared with 3-FLAB, especially in the delineation of higher activity regions (third class), with about twice the mean error and standard deviation ( $24\% \pm 20\%$ ) of 3-FLAB ( $11\% \pm 8\%$ ).

Figure 7 contains the mean error and standard deviation with respect to the maximum diameter, computed on the tumor histology database (Data Set 2). Whereas all methods gave relatively low mean errors ( $\leq 3\%$ ), the standard deviation associated with FCM and  $T_{\text{bckg}}$  (13% and 12%, respectively) is about twice that of 3-FLAB ( $<6\%$ ), and binary FLAB showed a standard deviation of almost 10%. The low mean error for all these algorithms is explained by the fact that there were about the same amount of under- and over-estimation of the diameters in this data set, resulting in an overall low mean error. Here the standard deviation is a better indicator of the accuracy obtained on the data set and demonstrates higher accuracy and robustness for 3-FLAB.

## DISCUSSION

Functional volume delineation represents an area of interest for multiple clinical applications (routine and research) of PET. Such areas include response to therapy studies and the use of biological tumor volumes in radiotherapy treatment planning. Although several fully automatic algorithms have recently been proposed (11–20), segmentation methodologies currently used in clinical practice are based on the use of fixed and adaptive thresholding (4–10). These algorithms have been shown to determine functional volumes accurately under specific imaging conditions of spherical and homogeneous activity distribution object in phantom studies and have been evaluated on clinical images for which the ground truth is unknown. In clinical practice, lesions are often heterogeneous in shape and uptake. To address these issues, we have extended a previously developed algorithm to evaluate

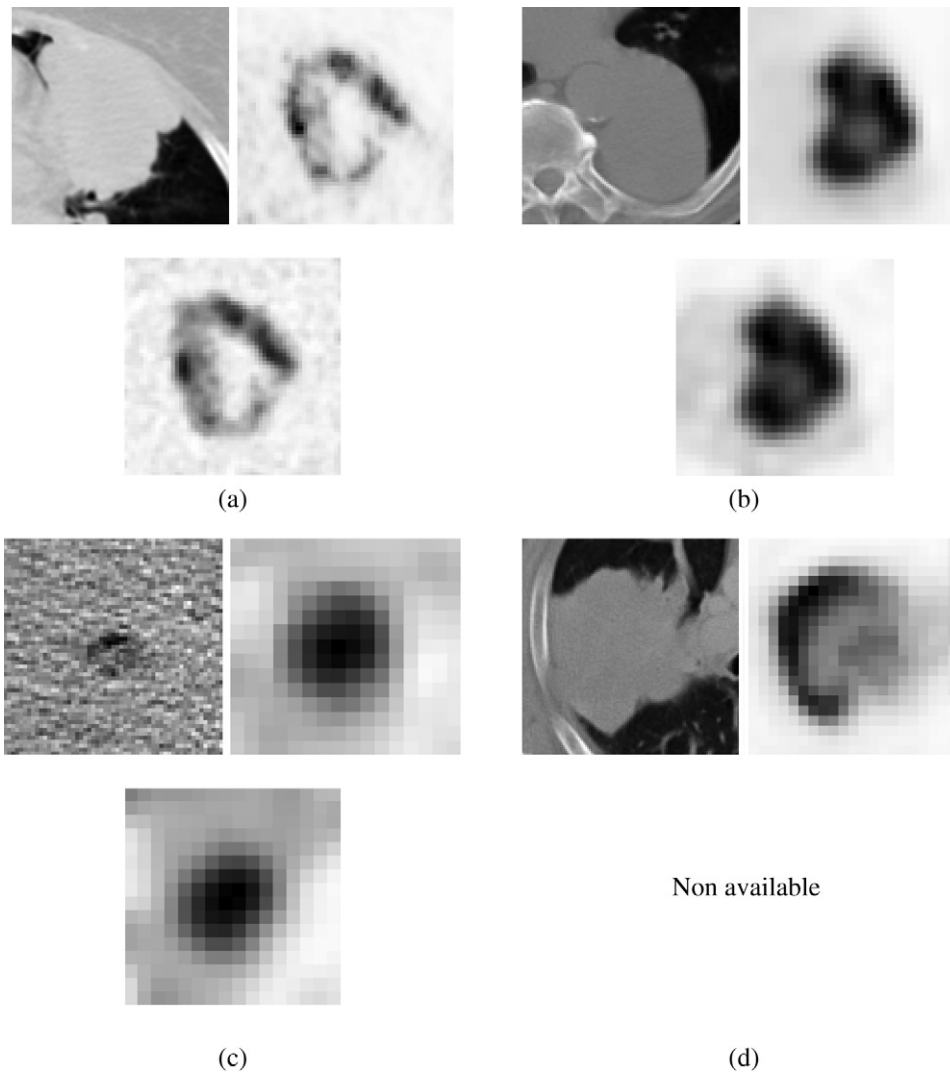


Fig. 4. Data sets illustration. (a–d) Examples of clinical tumors (up) with CT (left) and PET (right), and the corresponding simulated PET (down): (a–c) Data Set 1; (d) Data Set 2.

lesions with nonuniform uptake and nonspherical forms. In addition, we have proposed an evaluation framework including both realistic simulated patient lesions and histological assessment of tumor diameters, allowing for evaluation of segmentation algorithms under standard imaging conditions and the added advantage of knowing the ground truth.

The inability of the adaptive thresholding considered in this study to segment complex tumors accurately is demonstrated by its poor performance. This is explained by the fact that in cases of heterogeneous uptake, the 70% threshold used for the initial estimation of the tumor-to-background contrast may retain only the high uptake region, thus leading to incorrect contrast estimation. However, if the lesion is small or has a small contrast, the 70% threshold may lead to an initial overestimation of the volume of the tumor, and hence an underestimation of its uptake and an incorrect estimation of the contrast, for which the subsequent adaptive thresholding may not be able to compensate. In addition, the background ROI is user-dependent with a potentially high impact on the result, especially with heteroge-

neous background. In such cases, we systematically selected the ROI that resulted in the lowest error. Finally, the region growing implementation avoids incorporating false positives of the background if they are not connected to the main tumor, especially when the contrast is low or the background is noisy and heterogeneous. However, it also makes the algorithm dependent on the seed location and can lead to missing parts of the tumor when several high-uptake regions are connected by low-uptake regions. FCM can produce binary or three-class segmentations, but its robustness and accuracy are much lower compared with FLAB because it incorporates neither spatial correlation nor noise modeling. One advantage of the  $T_{\text{bckg}}$  over FCM is its region growing implementation that makes it less susceptible than FCM to the inclusion of high-intensity voxels of the background. Therefore, FCM usually performs poorer than  $T_{\text{bckg}}$  for low-contrast lesions and noisy images but better for heterogeneous activity distributions within the tumor. In contrast, 3-FLAB performed accurately even under challenging contrast, noise, and heterogeneity conditions, with

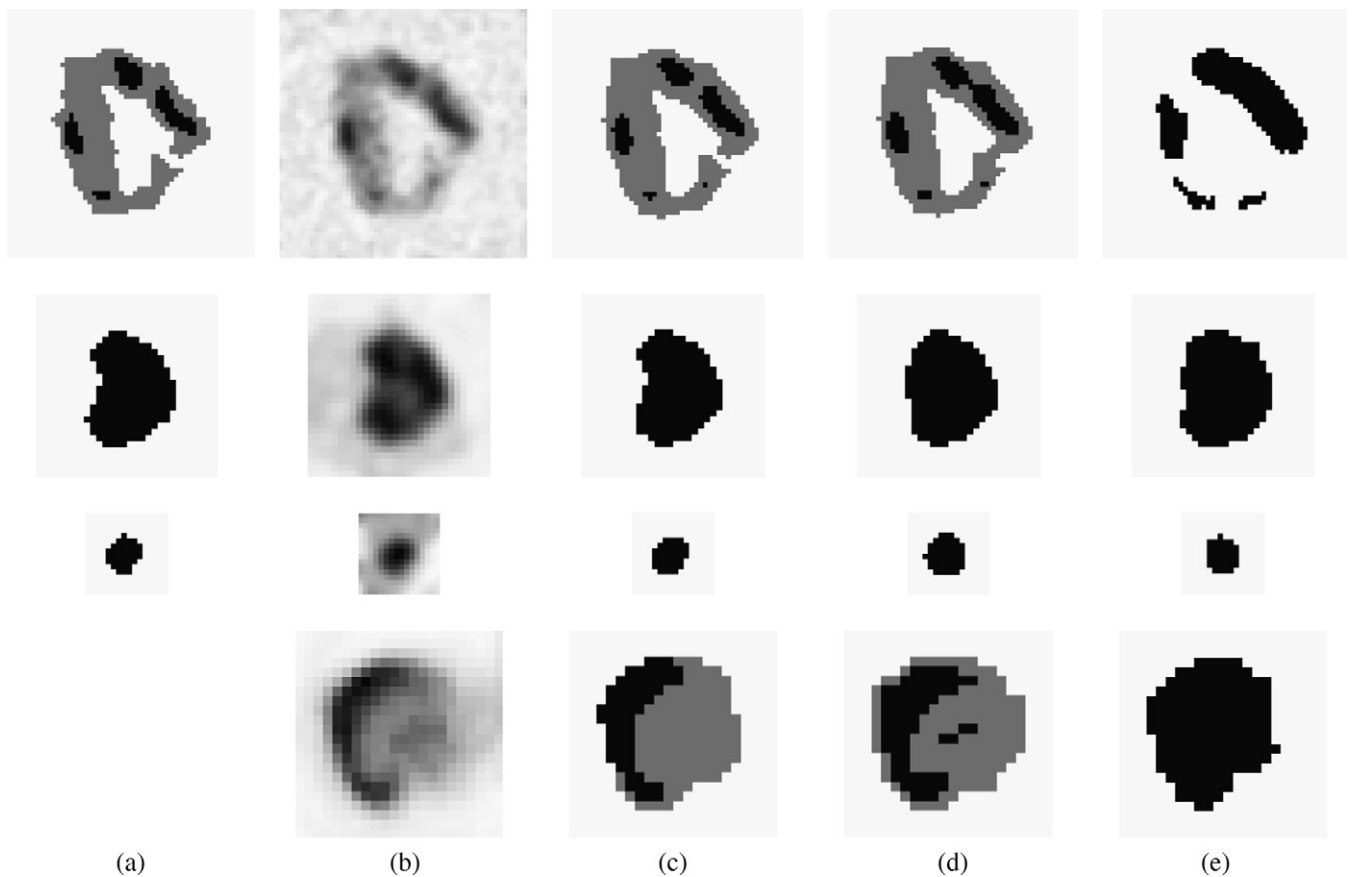


Fig. 5. Segmentations of the tumors in Fig. 4a–4d: (a) ground truth; (b) positron emission tomographic image; segmentations for (c) three-class fuzzy locally adaptive Bayesian, (d) fuzzy C-means, and (e) adaptive threshold models.

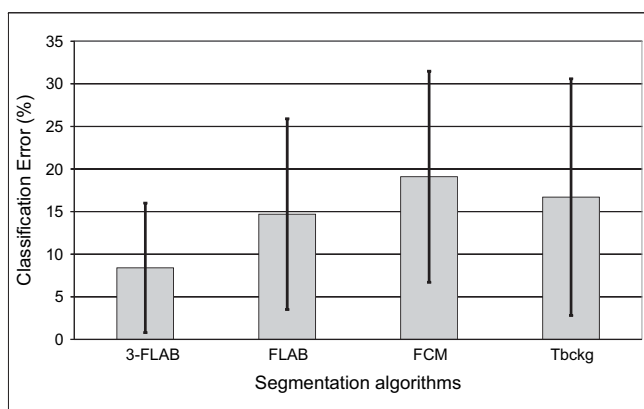
overall superior performance compared with the other algorithms considered here.

The need for more than three classes may arise for heterogeneous tumors on a heterogeneous background. However, all the clinical tumors considered in this study were correctly delineated using two or three classes because the contrasts between the heterogeneities within the tumor are usually much higher than those occurring in the background. Hence, only one hard class may be sufficient to deal with the background, whereas two are required to correctly handle the significantly different uptakes occurring inside the tumor. Eventually the 3-FLAB algorithm could be extended to more than three classes assuming that only pairs of hard classes generate fuzzy transitions. One also has to keep in mind that using more classes will lead to smaller regions, but those regions within the tumor will subsequently be used for quantification or radiotherapy dose boosting and/or painting and should therefore be kept reasonably large. The potential impact of using three classes proposed by 3-FLAB should therefore be investigated before more complex segmentations using additional classes can be considered.

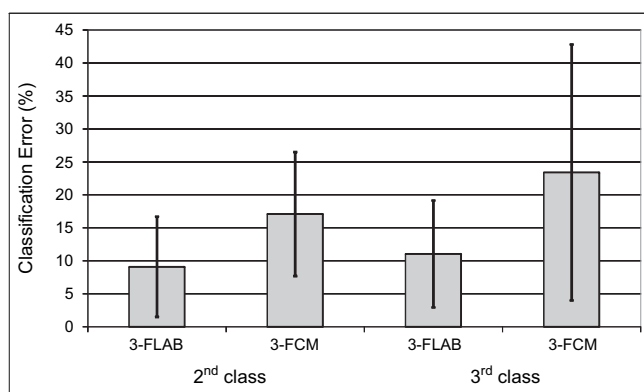
We have already demonstrated that FLAB performs well for small lesions down to 13 mm in diameter (20), and this study was not designed to investigate specifically the ability of 3-FLAB to deal with small tumors because these rarely

exhibit heterogeneous uptake that can be detected on the PET image considering the existing resolution limits. 3-FLAB retains all the characteristics of FLAB but also has the ability to consider a third class and therefore handle non-uniform lesion activity distributions. Thus, 3-FLAB does not as such improve the delineation of small (<2 cm) lesions. However, the higher/lower uptake regions within the larger tumors are often of small size, comparable to that of small lesions, with PVE affecting them with respect to their “background,” which is, in fact, the other part of the tumor with a different uptake. As Fig. 6b demonstrates, 3-FLAB is capable of accurately segmenting these regions.

An application that could greatly benefit from the use of FLAB is radiotherapy treatment planning (33). It is now acknowledged that planning based on PET/CT volumes improves tumor delineation by reducing inter- and intraobserver variability (32, 34). It can also lead to the inclusion of regions not visible on CT or the exclusion of regions without significant uptake (35). Using the 3-FLAB algorithm could help lower inter- and intraobserver variability, as well as shorten the time-consuming delineation process associated with currently implemented algorithms given the need for multiple phantom studies in the use of adaptive thresholding. 3-FLAB takes a few seconds per iteration even for the largest tumors considered in this study (on a single 2-GHz core processor in C++ implementation). Further,



(a)



(b)

Fig. 6. Mean classification errors and standard deviation for (a) all methodologies considering all 20 tumors of Data Set 1, (b) three-class fuzzy locally adaptive Bayesian, and 3-class fuzzy C-means considering the second and third classes of the 10 heterogeneous tumors of Data Set 1.

“dose painting” can be facilitated by the nonbinary nature of the proposed segmentation, allowing for automatic definition of ROIs inside the tumor—for example, in dose-escalation studies (36)—in addition to the external contour information for optimized dosimetry, potentially reducing the dose delivered to healthy surrounding tissues and organs. The impact of such improved accuracy on overall patient outcome remains to be demonstrated in clinical studies, which are planned for the future. Finally, FLAB robustness with respect to the noise characteristics associated with the use of different scanners, acquisition protocols, and reconstruction algorithms has been demonstrated in a preliminary study (24) and should allow its use with any type of PET images without the need for time-consuming preprocessing optimization.

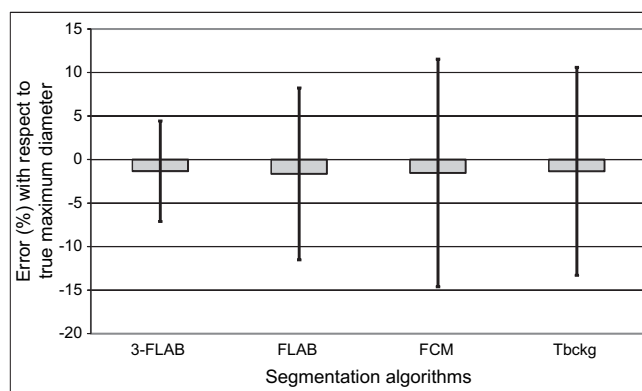


Fig. 7. Mean errors and standard deviation for each methodology, with respect to known maximum diameter of Data Set 2 tumors.

The proposed algorithm may also have an impact on diagnosis and therapy response assessment when combined with PVE correction (PVC) for accurate quantification. With various PVC approaches, anatomic information from MRI or CT is used to improve the quantitative and qualitative accuracy of functional images (37, 38). Unfortunately, when no anatomic image is available or no correlation exists between the anatomic and functional structures, such approaches are not easy to use (3). This is especially true in cases of large heterogeneous tumors for which there is little to no correlation between the anatomic and functional information. A potential solution will be the use of the FLAB result instead of the anatomic image in combination with one of the previously proposed PVC algorithms. This should lead to improved contrast at the object’s borders as well as improved quantification in the regions within the tumor. Such combination recently demonstrated encouraging results (39) and warrants further investigation regarding the potential impact in clinical therapy response studies.

## CONCLUSION

A modified version of the FLAB algorithm has been developed to include the estimation of three hard classes and three fuzzy transitions. This automatic approach combines statistical and fuzzy modeling to address specific issues associated with 3D-PET images, such as noise and PVE. Its accuracy has been assessed on both simulated and clinical images of complex shapes containing inhomogeneous activities and small regions. The results demonstrate the ability of 3-FLAB to delineate such lesions, for which the threshold-based methodologies suggested until now have failed.

## REFERENCES

1. Krak NC, Boellaard R, Hoekstra OS, *et al.* Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *Eur J Nucl Med Mol Imaging* 2005;32:294–301.
2. Jarritt PH, Carson KJ, Hounsel AR, Visvikis D. The role of PET/CT scanning in radiotherapy planning. *Brit J Rad* 2006; 79(Suppl):S27–35.
3. Soret M, Bacharach SL, Buvat I. Partial-volume effect in PET tumour imaging. *J Nucl Med* 2007;48:932–945.
4. Erdi YE, Mawlawi O, Larson SW, *et al.* Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. *Cancer* 1997;80(Suppl 12):2505–2509.
5. Greco C, Rosenzweig K, Cascini GL, *et al.* Current status of PET/CT for tumour volume definition in radiotherapy treatment



- planning for non-small cell lung cancer (NSCLC). *Lung Cancer* 2007;57:125–134.
6. Nestle U, Kremp S, Schaefer-Schuler A, et al. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *Jour Nucl Med* 2005;46:1342–1348.
  7. Black QC, Grills IS, Kestin LL, et al. Defining a radiotherapy target with positron emission tomography. *Int J Radiat Oncol Biol Phys* 2004;60:1272–1282.
  8. Davis JB, Reiner B, Huser M, et al. Assessment of 18(F) PET signals for automatic target volume definition in radiotherapy treatment planning. *Radiother Oncol* 2006;80:43–50.
  9. Daisne J-F, Sibomana M, Bol A, et al. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms. *Radiother Oncol* 2003;69:247–250.
  10. Van Dalen JA, Hoffman AL, Dicken V, et al. A novel iterative method for lesion delineation and volumetric quantification with FDG PET. *Nucl Med Commun* 2007;28:485–493.
  11. White CJ, Brady JM. A semi-automatic approach to the delineation of tumour boundaries from PET data using level sets. Society of Nuclear Medicine 52nd Annual Meeting, Toronto: Canada; June 2005. abstract 314.
  12. Tylski P, Bonniaud G, Decenciere E, et al. 18F-FDG PET images segmentation using morphological watershed: A phantom study. *IEEE Neurosci Symp Conference Record* 2006;4:2063–2067.
  13. Zhu W, Jiang T. Automation segmentation of PET image for brain tumours. *IEEE Neurosci Symp Conference Record* 2003;4:2627–2629.
  14. Montgomery DWG, Amira A, Zaidi H. Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model. *Med Phys* 2007;34:722–736.
  15. Demirkaya O. Lesion segmentation in wholebody images of PET. *IEEE Neurosci Symp Conference Record* 2003;4:2873–2876.
  16. Geets X, Lee JA, Bol A, et al. A gradient-based method for segmenting FDG-PET images: Methodology and validation. *Eur J Nucl Med Mol Imaging* 2007;34:1427–1438.
  17. Li H, Thorstad WL, Biehl KJ, et al. A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours. *Med Phys* 2008;35:3711–3721.
  18. Yu H, Caldwell C, Mah K, et al. Co-registered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging* 2009;28:374–383.
  19. Hatt M, Lamare F, Boussion N, et al. Fuzzy hidden Markov chains segmentation for volume determination and quantitation in PET. *Phys Med Biol* 2007;52:3467–3491.
  20. Hatt M, Turzo A, Roux C, et al. A fuzzy Bayesian locally adaptive segmentation approach for volume determination in PET. *IEEE Trans Med Imaging* 2009;28:881–893.
  21. Ling CC, Humm J, Larson S, et al. Towards multi-dimensional radiotherapy (MD-CRT): Biological imaging and biological conformality. *Int J Radiat Oncol Biol Phys* 2000;47:551–560.
  22. Caillol H, Pieczynski W, Hillon A. Estimation of fuzzy Gaussian mixture and unsupervised statistical image segmentation. *IEEE Trans Image Processing* 1997;6:425–440.
  23. Salzenstein F, Pieczynski W. Parameter estimation in hidden fuzzy Markov random fields and image segmentation. *Graphic Models Image Processing* 1997;59:205–220.
  24. Hatt M, Turzo A, Bailly P, et al. Automatic delineation of functional volumes in PET: A robustness study. Presented at the Society of Nuclear Medicine 2009. Annual Meeting Toronto: Canada; June 1317.
  25. Celeux G, Diebolt J. L'algorithme SEM: un algorithme d'apprentissage probabiliste pour la reconnaissance de mélanges de densités. *Revue Statistique Appliquée* 1986;34:35–52.
  26. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc B* 1977;39:1–38.
  27. McQueen J. Some methods for classification and analysis of multivariate observations. *Proc 5th Berkeley Symp Math Stat Prob* 1967;1:281–297.
  28. Dunn JC. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *J Cybernet* 1974;31:32–57.
  29. Segars WP. Development and application of the new dynamic NURBS-based cardiac-torso (NCAT) phantom [Ph.D. thesis]. Chapel Hill, NC: University of North Carolina; 2001.
  30. Le Maitre A, Segars WP, Marache S, et al. Incorporating patient specific variability in the simulation of realistic whole body 18F-FDG distributions for oncology applications. *Proc IEEE*, in press.
  31. Lamare F, Turzo A, Bizais Y, et al. Validation of a Monte Carlo simulation of the Philips Allegro/Gemini PET systems using GATE. *Phys Med Biol*, 2006;51:943–962.
  32. Van Baardwijk A, Bosmans G, Boersma L, et al. PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumour and involved nodal volumes. *Int J Radiat Oncol Biol Phys* 2007;68:771–778.
  33. Pan T, Mawlawi O. PET/CT in radiation oncology. *Med Phys* 2008;35:4955–4966.
  34. Fox JL, Rengan R, O'Meara E, et al. Does registration of PET and planning CT images decrease interobserver and intraobserver variation in delineating tumor volumes for non-small-cell lung cancer? *Int J Radiat Oncol Biol Phys* 2005;62:70–75.
  35. Ashamalla H, Raa S, Parikh K, et al. The contribution of integrated PET/CT to the evolving definition of treatment volumes in radiation treatment planning in lung cancer. *Int J Radiat Oncol Biol Phys* 2005;63:1016–1023.
  36. Sovik A, Malinen E, Olsen DR. Strategies for biologic image-guided dose escalation: A review. *Int J Radiat Oncol Biol Phys* 2009;73:650–658.
  37. Rousset OG, Ma Y, Evans AC. Correction for partial volume effects in PET: Principle and validation. *J Nucl Med* 1998;39:904–911.
  38. Boussion N, Hatt M, Lamare F, et al. A multiresolution image based approach for correction of partial volume effects in emission tomography. *Phys Med Biol* 2006;51:1857–1876.
  39. Boussion N, Hatt M, Visvikis D. Partial volume correction in PET based on functional volumes. *J Nucl Med* 2008;49(Suppl 1):388.