

AUTOMATIC HOT SPOT DETECTION AND SEGMENTATION IN WHOLE BODY FDG-PET IMAGES

Haiying Guan¹ *, Toshiro Kubota², Xiaolei Huang², Xiang Sean Zhou², Matthew Turk¹

¹ Computer Science Department, Univ. California, Santa Barbara
{haiying, mturk}@cs.ucsb.edu

² Computer Aided Diagnosis and Therapy Group, Siemens Medical Solutions, USA
{xiaolei.huang, toshiro.kubota, xiang.zhou}@siemens.com

ABSTRACT

We present a system for automatic hot spots detection and segmentation in whole body FDG-PET images. The main contribution of our system is threefold. First, it has a novel body-section labeling module based on spatial Hidden-Markov Models (HMM); this allows different processing policies to be applied in different body sections. Second, the Competition Diffusion (CD) segmentation algorithm, which takes into account body-section information, converts the binary thresholding results to probabilistic interpretation and detects hot-spot region candidates. Third, a recursive intensity mode-seeking algorithm finds hot spot centers efficiently, and given these centers, a clinically meaningful protocol is proposed to accurately quantify hot spot volumes. Experimental results show that our system works robustly despite the large variations in clinical PET images.

1. INTRODUCTION

Positron emission tomography (PET) using fluorine-18 deoxyglucose (FDG) is a nuclear medicine medical imaging technique that produces a three dimensional image of functional processes in the body. Tumors in FDG-PET appear as “hot spots” due to increased FDG uptake, and by late 1990s, a large body of literature have clearly shown that FDG-PET imaging is essential and effective for detection, diagnosis and prognosis of tumors. It is also shown applicable to neurological disorders and coronary artery disease [1].

It is expected that the number of FDG-PET images performed in most facilities will exceed all other procedures performed with radioactive compound in the coming years [1]. As the amount of PET data generated every day increases, the burden on the radiologists in reading and diagnosing the cases increases. This may result in a negative effect, as the rate of overlooking anomalies can increase due to fatigue, and the risk of over-diagnosis may also rise. Computer-aided detection (CAD) tools have constantly demonstrated effectiveness in supplementing the task of radiologists; they improve the sensitivity of anomaly detection and reduce the interpretation variation among radiologists. Currently, to our knowledge,

there are no commercially available CAD tools for FDG-PET images. In this paper, we present a system design for a PET-CAD tool that detects suspicious regions where the amount of FDG uptake values are abnormally high.

Previous works on PET image studies have been presented in both clinical study and medical image processing. From a clinical perspective, the main challenge in hot spot detection and segmentation is that normal tissues also uptake different levels of FDG [2]. There is a great variance in the normal physiology of FDG uptake. For example, FDG is cleared primarily through the renal system. Thus, the normal bladder might be the highest FDG uptake organ. Brain have high FDG uptake. Heart and bowel have variant levels of FDG uptake. Liver, bone marrow and spleen sometimes also have mildly high FDG uptake which is higher than some tumors. Considering all these normal tissue uptakes, the intensity range of the normal tissue overlaps with the tumors significantly. In addition, the shape, size, and location of tumors are unpredictable.

From a medical image processing perspective, there are several challenges in hot spot detection and segmentation. First, Whole-body PET images usually have low resolution, typically 5-10mm, which is much lower than CT whose resolution can be in the sub-millimeter level. Second, the intensity range of PET images is large, and intensity values of a PET image heavily depends on the clinical factors, such as applied radiopharmaceutical, injection dose and time, and patients factors, such as body weight and the tissue uptake.

In this paper, we present a PET-CAD system for automatic and robust detection and segmentation of hot spots in whole body FDG-PET images. The system is designed to exploit knowledge from the clinical literature and provides solution to three problems: body context interpretation, detection of volumes with high FDG uptake, and accurate segmentation of hot spots. Rough labeling of the body allows the system to adaptively adjust its behavior to each section. A solution to the second problem provides the candidate volume (volume of interest or VOI) that serve as input to the segmentation algorithm. The results of the segmentation algorithm are the CAD hot-spot findings and can be presented to the radiologist in both 2D and 3D views.

*The first author performed the work while at SIEMENS

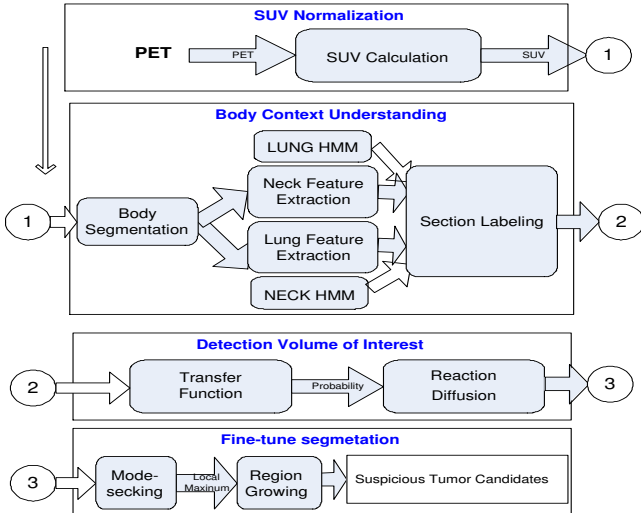


Fig. 1. The system diagram

The paper is organized as follows. Section 2 gives an overview of our CAD system. Section 3 describes a pre-processing step where the PET values are converted to more clinically meaningful and stable values called *standard uptake values*. Section 4 proposes a body context interpretation method using Hidden Markov Models. Section 5 describes the competition-diffusion algorithm for VOI detection. Section 6 proposes a mode-seeking region growing algorithm for hot spot segmentation. Section 7 illustrates experimental results. Finally, conclusions are drawn in Section 8.

2. SYSTEM OVERVIEW

Figure 1 shows an overview of our PET-CAD system by organizing the system components with the processing flow. Because the raw intensity of PET depends on many factors and is very unstable, we first convert PET values to Standardized Uptake Values (SUV). Since FDG intakes of different organs are different, for accurate detection and segmentation, it is essential to understand the body context. Thus the next step in the processing flow is to derive a rough body context interpretation using a spatial *Hidden Markov Model* (HMM). Next, we extract suspicious regions with abnormal FDG intakes. We employ a computationally efficient *competition-diffusion* (CD) algorithm for the task. The CD is sensitive to initialization, and we use the body section information obtained above to introduce different initialization strategies to lung and other organs. Finally, according to the body context knowledge and detection results, we fine-tune the segmentation by a model-seeking region growing algorithm.

3. STANDARDIZED UPTAKE VALUE (SUV)

The big challenge for PET image processing is that the intensity of PET images heavily depends on the clinical and patient factors. In addition, the intensity value is widely distributed from 0 to 32767. The histogram of PET intensity typically shows no peak. Thus, the traditional Gaussian Mixture Model (GMM) is not suitable for modeling the histogram

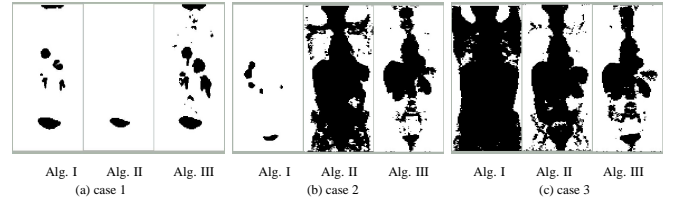


Fig. 2. Comparisons of three segmentation algorithms distribution of intensity values of PET images, even after the normalization.

In clinical literature, SUV is widely used by radiologists for normalization across time points, and different patients¹. It normalizes the measured FDG concentration by decay corrected injection dose per gram body mass. We convert a raw PET value to the SUV using the following equation.

$$v_{suv} = A_s v_{pet} W D^{-1} e^{-\log(2)T/\Delta T} \quad (1)$$

where v_{suv} is the SUV value, v_{pet} is the raw PET value, A_s is the scaling coefficient of the reconstruction protocol, W is the patient weight in gram, D is the amount of radioactive dose in Bq, T is the interval between injection time and the scan start time, and ΔT is the half-life of the radio isotope. The detail information about SUV can be found in [3].

Figure 3 illustrates the difficulties associated with working directly on PET values. We segment a PET volume using three different techniques: fixed thresholding on PET values, adaptive thresholding on PET values, and fixed thresholding on SUV values². In Figure , (a) (b) (c) are three test volumes, and (b) and (c) are volumes from the same patient taken at different time point. For each cases, from left to right, the fixed thresholding on PET values (Alg. I), the adaptive thresholding on PET values (Alg. II), and fixed thresholding on SUV values (Alg. III) are shown. PET values show high variations between two volumes and fixed thresholding cannot cope with the variations; case (b) Alg. I was under segmented and case (c) Alg. I was over segmented. The adaptive thresholding failed to handle inter-patient variations of PET values; Case (b) Alg. II and (c) Alg. II are over-segmented, while case (a) Alg. II is severely under-segmented. SUV values exhibit stability against both inter and intra-patient variations. As shown in Figure Alg. III, a simple thresholding drawn at 2.5 could effectively and efficiently extract regions of high FDG uptake.

4. BODY CONTEXT INTERPRETATION

Malignant cells in different organs exhibit different levels of FDG uptake. For example, cancer cells in liver tend to take more FDG than cancer cells in lung, while a normal liver may take as much FDG as cancer cells in lung. Thus, to accurately

¹For example, the cutoff SUV threshold for differentiating malignant and benign lung tumors is 2.5 [1]. The maximum SUV values are used for grading tumors.

²Fixed threshold levels are set to 3000 for PET values (Alg. I) and 2.5 for SUV values (Alg. III). The adaptive thresholding (Alg. II) uses the MATLAB implementation of thresholding level computation (*graythres*) applied only to the body part. This threshold level is multiplied by two to get the threshold level used to derive the results shown here.

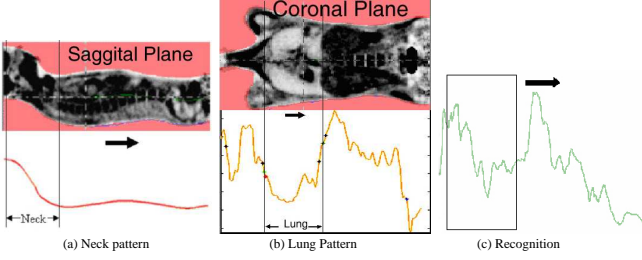


Fig. 3. HMM training and recognition

detect abnormal FDG uptake in a PET volume, it is important that we have a rough anatomical interpretation of the volume.

It is extremely difficult to extract anatomical information from a PET volume, as the PET data does not present detailed structures of the body. One way to accomplish the task is to use a CT volume of the same patient and register it to the PET volume. Many recent PET scanners are combined with a CT scanner, providing a pair of accurately registered PET-CT volumes. It is thus feasible to provide anatomical interpretation of the PET volume by first extracting bones, lungs and other organs from the CT volume and mapping them onto the PET counterpart. The approach, although highly effective, does require a registered CT data, demand more storage, and make the system less portable. Due to these disadvantages, in the present work, we explore a way to achieve the goal with a PET volume alone.

Our body section labeling algorithm is based on two observations. First, the neck curve when the 3D medial axis of the body is projected to the sagittal plane exhibits a distinctive pattern as shown Fig. 3 (a). Second, the FDG uptake of the normal tissue of the lung is generally less than other tissues such as liver and kidney. Fig. 3(b) shows one example. We compress the 3D body into two feature curves that preserve the above observations. The feature curve of the first observation is a trace of the centroid of the body in each axial slice. The feature curve of the second observation is, after removing SUVs above 2.5, a trace of the mean intensity in each axial slices (black arrows in Fig. 3).

We then employ an Hidden Markov Model(HMM) to divide the body into three sections: neck, lung and abdomen. HMM is a powerful parametric model and is feasible to characterize the stochastic processes: the measurable observation process and the immeasurable hidden states. It has been successfully applied to speech recognition and gesture recognition in the temporal space [4]. In our work, we consider a spatial discrete HMM along the head to toes direction, t .

First, two HMM models, neck model and lung model, are obtained from the observation patterns of necks and lungs. The training samples of the two feature curves are extracted by manual markings of the lungs. After vector quantization, for the given training samples, the models are obtained by maximizing the likelihood $\Pr(Observ.|Model)$ with Baum-Welch algorithm.

In recognition stage, with the models, the locations of a neck and lungs are obtained by searching for the the best

match point in the whole body feature curves to the models. To do so, we compute $\Pr(Model|Observ.)$ within a sliding window with the forward-backward algorithm and locate the maximum point. In our system, we post-process the probability curve with the spatial integration, and output the locations of the maximum likelihood point as our final neck or lung point. Figure 3(c) illustrates the recognition process for the lung feature.

5. DETECTION OF THE VOLUME OF INTEREST

The Competition-diffusion (CD) algorithm [5] is a segmentation algorithm which assigns to each voxel a class label from a set of n classes. It contains two processes: competition and diffusion. The competition process selects the most fitted label and prevents over smoothing. The diffusion process brings a good spatial coherence to the detection map. Each voxel is given a state vector $\vec{x} \in S^n$, where $S^n = \{\vec{x} | x_i \geq 0, \sum_i x_i = 1\}$. Each component in the state vector gives a normalized fitness of the particular class. The closer the state value is to 1, the more fitting the location is to the class.

For our system, $n = 2$, and x_1 and x_2 are state values for the voxel being normal and suspicious tissues, respectively. The initialization step needs to translate the observed voxel value (SUV value in our case) to x_1 and x_2 . For lungs, SUVs higher than 2.0 are considered abnormal, while for liver and kidney, SUVs higher than 3.0 are considered abnormal. From the body context interpretation step, we know roughly where the lungs are, and use this knowledge to employ different initialization strategies for lungs and other organs. In particular, we use $x_2(v_{suv}) = \min(1, v_{suv}/3.8)$ for the lung section and $x_2(v_{suv}) = \min(1, v_{suv}/5)$ for other sections to derive the initial value of x_2 . Throughout, x_1 is initialized to $1 - x_2$.

After the initialization, x_i at each voxel obeys the following dynamics.

$$\dot{x}_i = x_i(f_i - \bar{f}) + \mu \nabla^2 x_i \quad (2)$$

where \dot{x}_i is the time derivative of x_i , f_i is the fitness measure associated with x_i , $\bar{f} = \sum_i x_i f_i$ is the average fitness, μ is a constant, and ∇^2 is a linear diffusion operator. When f_i s are linear, symmetric and positive definite, the system converges to a state where each voxel has one dominant state. We use an identity mapping $f_i = x_i, \forall i$. We assign the label corresponding to the dominant state to generate a segmentation map.

Figure 4 shows a result of the VOI detection in the PET image of a patient with Bone Metastasis. The left image is a result of applying thresholding at 2.5 to the SUV volume. The middle image is a result of applying CD with uniform initialization. We used the non-lung initialization strategy as described above. The right image is a result of applying CD with section dependent initialization. CD removed some small clusters whose SUV values are not significantly high, and the body section dependent initialization revealed additional hot spots in the ribs.



Fig. 4. VOI detection results

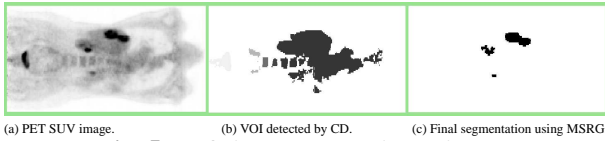


Fig. 5. Refining segmentation using MSRG
6. MODE-SEEKING REGION GROWING

The CD algorithm with a low cut-off threshold produces VOI that have high probabilities of being hot spots. Because this is a conservative step to detect VOI, the volumes tend to be larger than the true extent of the hot spots (see Fig. 5(b)), and several nearby hot spots may be merged in one volume. In order to achieve accurate quantitative volume measurement of each hot spot, we further apply a *Mode-Seeking Region Growing* segmentation algorithm (MSRG). The first step is to find the intensity modes (or local maximums), which correspond to the primary locations of hot spots, where the FDG uptake is the maximum. Radiologists often use the maximum SUV value of a hot spot to grade tumors, and help to determine the extent of a hot spot [6]. Hence a segmentation algorithm based on the maximum SUV value produces clinically meaningful and accurate segmentation results that are consistent with a doctor's interpretation. In our mode-seeking step, we start from a seed point as a probe. At the probe point, we find the maximum SUV location within the neighborhood of the point, we then move the probe to the point with the maximum SUV value in the neighborhood. This process continues until the probe reaches an intensity mode in the spatial domain where the probe itself would have the maximum SUV value and stop moving. If we view the SUV image as a 4D surface $f(x, y, z)$ where (x, y, z) covers the 3D image domain, then the path of the probe is a path moving uphill in the function field $f(x, y, z)$ to the nearest mode. This mode-seeking procedure is a more straightforward version of mean-shift [7]. By starting the mode-seeking procedure from many seed points uniformly distributed within a VOI, we can detect all the SUV-maxima points that correspond to the primary locations of the true hot spots. The second step of the algorithm is to apply 3D region growing with seed points at the detected SUV-maxima points. The region growing process stops when the SUV values of the boundary points drop to 40% of the maximum SUV value at the initial seed point; this is a clinically meaningful threshold that have been utilized in the literature [6]. We apply the mode-seeking region growing algorithm on all the VOI detected by CD, and an example segmentation result is shown in Fig. 5(c).

7. EXPERIMENTAL RESULTS

Our data are acquired by SIEMENS Biograph LSO PET/CT scanners. We tested 30 cases for body volume segmentation

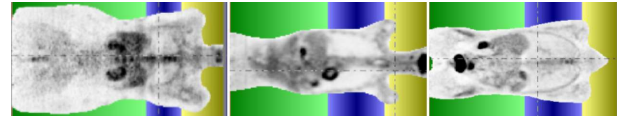


Fig. 6. Body segmentation and section labeling results

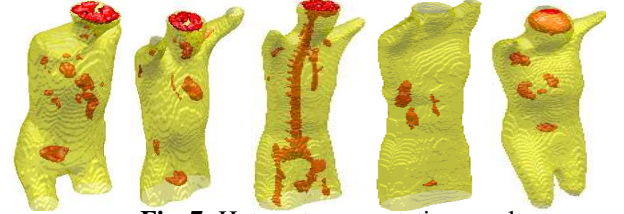


Fig. 7. Hot spot segmentation results

and body section labeling (11 cases are used for training). Figure 6 shows some example results and the algorithm is robust to noise and can adapt to patients with different diseases. We applied Competition Diffusion and mode-seeking region growing algorithms to segment hot spots in all cases. Some example 3D-view segmentation results are shown in Figure 7. The segmentation process is completely automatic and the results are reproducible.

8. CONCLUSION

We have presented a PET-CAD system, which mainly includes the following contributions: the adoption of SUV for PET data normalization, HMM for body section recognition, Context-based VOI detection with the Competition Diffusion algorithm, and clinically meaningful hot spot segmentation using mode-seeking region growing. In the future, each step of our system could be improved by utilizing CT information using PET-CT data and PET-CT registration, in order to improve system performance and provide more information to the radiologists. The recognition of abnormal vs. normal hot spots also needs more elaborate solutions.

9. REFERENCES

- [1] A. Alavi, Ed., *Radiologic Clinics of North America, PET Imaging II*, vol. 43, ELSEVIER, SAUNDERS, 2005.
- [2] P. S. Conti and D. K. Cham, *PET-CT: A Case-Based Approach*, Springer, Singapore, 2004.
- [3] S. C. Huang, "Anatomy of SUV," *Nuclear Medicine and Biology*, vol. 27, pp. 643–646, 2000.
- [4] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, , no. 2, pp. 257–286, Feb 1989.
- [5] Toshiro Kubota and Fausto Espinal, "Reaction-diffusion systems for hypothesis propagation," in *ICPR'00*, vol. 3, p. 3547.
- [6] Q. C. Black, I. S. Grills, L. L. Kestin, C. Y. O. Wong, J. W. Wong, A. A. Martinez, and Yan. D., "Defining a radiotherapy target with positron emission tomography," *Int. J. Radiation Oncology Biol. Phys.*, , no. 4, pp. 1272–1282, 2004.
- [7] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis Machine Intell.*, vol. 24, no. 5, pp. 603–619, 2002.