

## ORIGINAL ARTICLE

**Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters**PAULINA E. GALAVIS<sup>1</sup>, CHRISTIAN HOLLENSEN<sup>2,3</sup>, NGONEH JALLOW<sup>1</sup>,  
BHUDATT PALIWAL<sup>1,4</sup> & ROBERT JERAJ<sup>1,4</sup>

<sup>1</sup>Department of Medical Physics, University of Wisconsin, Madison, WI, USA, <sup>2</sup>Department of Informatics and Mathematical Models, Technical University of Denmark, Copenhagen, Denmark, <sup>3</sup>Department of Radiation Oncology, Copenhagen University Hospital-Rigshospitalet, Denmark and <sup>4</sup>Department of Human Oncology, University of Wisconsin, Madison, WI, USA

**Abstract**

**Background.** Characterization of textural features (spatial distributions of image intensity levels) has been considered as a tool for automatic tumor segmentation. The purpose of this work is to study the variability of the textural features in PET images due to different acquisition modes and reconstruction parameters. **Material and methods.** Twenty patients with solid tumors underwent PET/CT scans on a GE Discovery VCT scanner, 45–60 minutes post-injection of 10 mCi of [<sup>18</sup>F]FDG. Scans were acquired in both 2D and 3D modes. For each acquisition the raw PET data was reconstructed using five different reconstruction parameters. Lesions were segmented on a default image using the threshold of 40% of maximum SUV. Fifty different texture features were calculated inside the tumors. The range of variations of the features were calculated with respect to the average value. **Results.** Fifty textural features were classified based on the range of variation in three categories: small, intermediate and large variability. Features with small variability (range  $\leq 5\%$ ) were entropy-first order, energy, maximal correlation coefficient (second order feature) and low-gray level run emphasis (high-order feature). The features with intermediate variability ( $10\% \leq \text{range} \leq 25\%$ ) were entropy-GLCM, sum entropy, high gray level run emphasis, gray level non-uniformity, small number emphasis, and entropy-NGL. Forty remaining features presented large variations (range  $> 30\%$ ). **Conclusion.** Textural features such as entropy-first order, energy, maximal correlation coefficient, and low-gray level run emphasis exhibited small variations due to different acquisition modes and reconstruction parameters. Features with low level of variations are better candidates for reproducible tumor segmentation. Even though features such as contrast-NGTD, coarseness, homogeneity, and busyness have been previously used, our data indicated that these features presented large variations, therefore they could not be considered as a good candidates for tumor segmentation.

Positron emission tomography (PET) with [<sup>18</sup>F]-2-Fluoro-2-deoxy-D-glucose (FDG) is widely used for clinical diagnostic, staging, prognosis and treatment response of cancer [1]. In addition to its use as a staging tool, FDG-PET has also been used to assist with target definition [2]. The incorporation of PET in radiotherapy for tumor delineation provides the physiologic information about the tumor, allowing for a biological radiation therapy treatment [3,4]. Various methods have been used to determined the tumor boundaries in FDG PET images. The first and still widely used, is the manual tumor segmentation by an experienced nuclear medicine physician or a radiation oncologist but, this presents problems of inter-observer variations [5]. Another approach is

based on auto segmentation employing image intensity threshold levels, which are based on phantom studies where the tumors are assumed to be spherical in shape [6–8]. However, tumors are rarely spherical and activity distributions are more complex in patients than in phantoms.

Alternatively, texture is an important property commonly used for image classification in the field of pattern recognition. There are three different approaches used in image processing to find the textural feature of a region of interest in an image: first-order features, second-order features, and higher-order features. The first order features use statistical moments of the intensity histogram of the image [9] and do not contain information about the relative

position of pixels with respect to each other. Second order features employ the angular nearest-neighbor gray tone spatial-dependence matrices, also known as gray level co-occurrence matrices (GLCM) [10]. Higher order features can be obtained using the run-length [11], neighboring dependence [12], and neighborhood difference [13]. Quantitative analysis of medical images using texture features has been used in diagnostic for differentiation between normal and abnormal regions or as a guide for tumor segmentation, using computed tomography (CT) [14], digitized mammography [15,16] and magnetic resonance [17]. More recently, quantitative textural analysis of FDG PET/CT images has been investigated for characterization of head and neck tissues [18] as well as tumor segmentation [19]. Other study employed textural analysis to predict treatment outcomes on patients with cervix and head and neck cancers [20]. The purpose of this work is to study the variability of the textural features in PET images due to different acquisition modes and reconstruction parameters.

## Materials and methods

### *Patients, image acquisition and image reconstruction*

The data for 20 patients with different types of solid tumors were acquired on a PET/CT GE Discovery VCT scanner (Waukesha, WI). Patients were injected with 10 mCi of [ $^{18}\text{F}$ ]FDG and scanned 45–60 minutes post injection. Malignancies included adrenal gland carcinoma, lung, epiglottis, and esophagus cancer. This retrospective study was approved by the University of Wisconsin Institutional Review Board (IRB), under the protocol number M-2010-1010. The PET acquisition for ten patients were performed in 2-dimensional (2D) mode followed by 3-dimensional (3D) mode, while the other ten were performed in the reverse order. The 2D/3D studies were acquired using seven to eight bed positions to cover the area from skull to mid-thigh. Attenuation correction using the CT data was also applied.

The raw PET data were reconstructed in 2D using the ordered subset expectation maximization (OSEM) algorithm with 14 sub-sets  $\times$  2 and four iterations (28 and 56 iterative updates). The 3D reconstruction was done using an Iterative-Vue Point algorithm with two and four iterations. The grid size ( $128 \times 128$  vs.  $256 \times 256$ ) and post-reconstruction filter width (3 mm, 5 mm and 6 mm) in the reconstruction algorithms were also varied. In total, ten different reconstructed images within the clinical setting were used for each patient, listed in Table 1, resulting in two hundred analyzed images. The image reconstruction labeled as 3D-256-ITER2-3mm will be referred as the default image. After reconstruction, all images were normalized to the body mass Standardized Uptake Value (SUV) [21].

### *Segmentation and feature extraction*

Patient diagnostic reports were used to locate the tumors on each patient. Using Amira (Visage Imaging), for image visualization, the tumor regions were identified. Subsequently, the  $\text{SUV}_{\text{max}}$  (maximum SUV) was determined in the tumor region and used to define the tumor volumes using the threshold level of 40% of the  $\text{SUV}_{\text{max}}$ , which was chosen as the reference contours. Eight first order features [9], 23 features using the co-occurrence matrix [10], 11 features employing the gray level run length matrix [11], five features using the neighboring gray level [12] and three features using the neighborhood gray tone difference matrix [13] for a total of 50 features inside each tumor were calculated. For each voxel inside the base contours, a *patch* was extracted, defined as a portion of the image with  $7 \times 7$  (coronal, sagittal) voxels in size, centered on that voxel. Texture features were computed on these patches on all axial slices containing the  $\text{SUV}_{40\%}$  and then the mean value was calculated. All feature extraction and thresholding-based segmentation was performed using an inhouse code developed in MATLAB. Figure 1 shows a diagram with the steps employed to extract the texture features.

Table I. Image parameters.

Image #	Acq. Mode	Grid-Size	Recon. Alg	Iter. number	Post-filter width (mm)	Legend
1	2D	$128 \times 128$	OSEM	2	3	2D-128-OSEM2-3mm
2	2D	$128 \times 128$	OSEM	2	5	2D-128-OSEM2-5mm
3	2D	$128 \times 128$	OSEM	4	5	2D-128-OSEM4-5mm
4	2D	$256 \times 256$	OSEM	2	3	2D-256-OSEM2-3mm
5	2D	$256 \times 256$	OSEM	2	5	2D-256-OSEM2-5mm
6	3D	$128 \times 128$	ITER	2	3	3D-128-ITER2-3mm
7	3D	$128 \times 128$	ITER	2	6	3D-128-ITER2-6mm
8	3D	$128 \times 128$	ITER	4	6	3D-128-ITER4-6mm
9	3D	$256 \times 256$	ITER	2	3	3D-256-ITER2-3mm
10	3D	$256 \times 256$	ITER	2	6	3D-256-ITER2-6mm

Acq. Mode = acquisition mode; Recon. Alg = reconstruction algorithm; Iter = iteration.

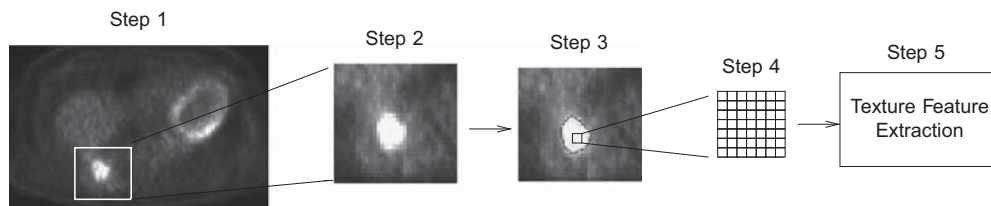


Figure 1. Diagram for Feature Extraction. From left to right: **Step 1:** Identification of the axial slices showing the tumor. **Step 2:** Cropping to tumor region and location of the  $SUV_{max}$ . **Step 3:** Tumor delineation using 40% of the  $SUV_{max}$ . **Step 4:** Definition of the  $patch\ 7 \times 7$  voxels in size. **Step 5:** Feature extraction.

For each texture feature, the percent difference with respect to the mean value was calculated according to Equation 1.

$$\%Diff = \frac{100 \times (X - X_{mean})}{X_{mean}} \quad (1)$$

where  $X$  corresponds to texture feature value for each reconstructed image and  $X_{mean}$  is the average value. The absolute value was omitted in the Equation 1 to account for the deviation above and below the mean.

## Results

For the purpose of textural feature variability evaluation, these were classified as *small* ( $\leq 5\%$ ), *intermediate* ( $10\% \leq \text{range} \leq 25\%$ ) and *large* ( $> 30\%$ ) range of variation with respect to the mean using Equation 1. Following this criteria, the features that presented small variations are the entropy (first order), energy, maximal correlation coefficient (second order feature), and low gray-level run emphasis (high order feature-run length). The features that presented intermediate variations include entropy-GLCM, sum entropy, both second order features, high gray level run emphasis and gray level non-uniformity (high order features-run emphasis), small number emphasis and entropy-GL (high order features-neighboring dependence). Forty features showed a large range of variations, some of these include contrast, coarseness and busyness which have been commonly used in previous studies. Figure 2 shows the maximum and minimum deviations among all patients from the mean for first, second and high order features.

The entropy has been previously calculated employing different matrices for feature extraction. We studied the entropy from the co-occurrence and neighboring matrices (entropy-GLCM and entropy-NGL, respectively). Our data demonstrated that the entropy outcomes are affected by the matrix employed to extract this feature. The entropy displayed higher variation when extracted from neighboring matrices. The maximum entropy variation was 10% and 5% when extracted from neighboring and co-occurrence matrices, respectively. In both matrices the entropy was influenced

more strongly by grid size than filter width. The contrast was extracted from different matrices (contrast-GLCM and contrast-NGTD) but the trends in variation were very similar. The maximum variations of the contrast using both matrices were above 100%.

## Discussion

Texture features in PET images are gaining importance as a tool for tumor discrimination and segmentation [18,19], as well as a potential metric for treatment assessment [20]. In this work, a systematic study of textural feature variability in PET images due to different acquisition modes and reconstruction parameters was presented. The range of variations of 50 textural features in FDG PET images were calculated and classified into small, intermediate, and large variations due to different image reconstruction.

The features exhibiting small variations such as entropy-first order, energy, maximal correlation coefficient, and low gray level run emphasis, are better candidates for reproducible auto segmentation and tumor assesment. For instance, entropy and energy have already been employed to automatically distinguish between tumor and normal regions in computer aid systems for auto-segmentation by [19]. It is important to mention that even though the maximal correlation coefficient and the low gray-level run emphasis features showed small variations, they exhibited image grid size dependency.

Intermediate variations were observed when the entropy was extracted from the co-occurrence matrix (entropy-GLCM) or from the neighboring gray level matrix (entropy-NGL), meaning that this feature is affected by the SUV scaling. The contrast was also extracted from the co-occurrence matrix (contrast-GLCM) and neighborhood gray tone difference matrix (contrast-NGTD) presenting the same trend and comparable variability.

We also studied features presenting large variations, such as contrast, homogeneity, coarseness, and busyness, that have been previously considered to discriminate between tumor and normal tissue in patients with HN and lung cancer [18] and proposed as potential metrics for treatment assessment of HN

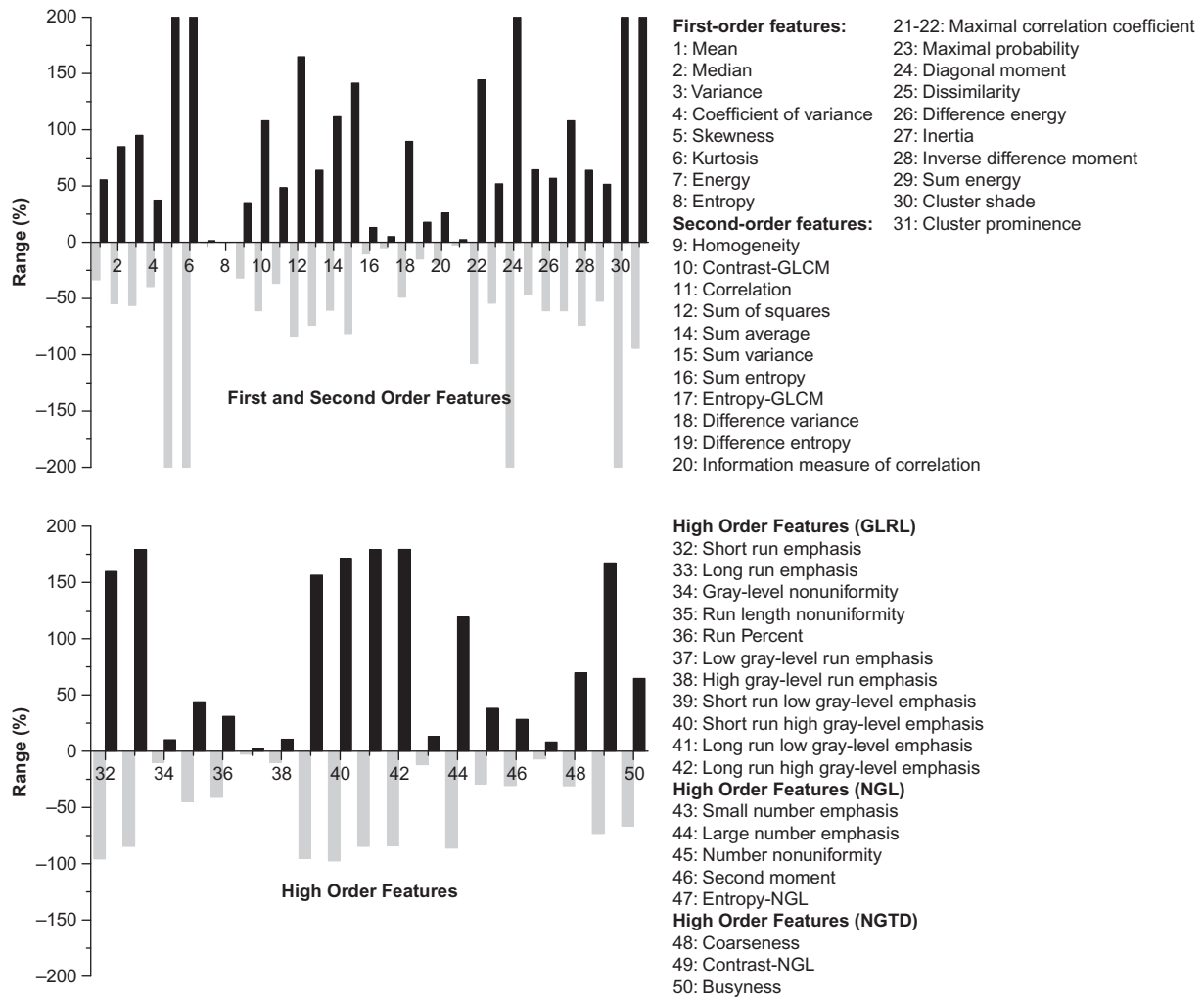


Figure 2. Range of variation of the textural features. **Top:** shows 8 first and 23 second order features. **Bottom:** shows high order features: 11-GLRL, 5-NGL, 3-NGTD. The gray and black shade represent the minimum and maximum deviations from the mean.

and cervix cancer [20]. In the study by Yu et al. [19], where texture features were employed on CT and PET images to develop an auto segmentation system (COMPASS), PET extracted features included, contrast, coarseness, busyness. Our data indicated that these features presented large variations in tumor regions when PET images are acquired and reconstructed with different parameters. Therefore, these features are more prone to errors if employed to quantify the change in tumor texture in response to a therapy.

**Acknowledgements**

This work was supported by NIH grant 1R01CA136927 and the graduate school programme IT Everywhere under contract 645-09-0097 from the Danish Agency for Science Technology and Innovation. We would like to acknowledge Christine Jaskowiak for her assistance in performing research on the clinical PET/CT scanner.

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

**References**

- [1] Heron DE, Andrade RS, Beriwal S, Smith RP. PET/CT in radiation oncology – the impact on diagnosis, treatment planning, and assessment of treatment response. *Am J Clin Oncol–Cancer Clin Trial* 2008;31:352–62.
- [2] Paulino AC, Koshy M, Howell R, Schuster D, Davis LW. Comparison of CT- and FDG PET defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2005;61:1385–92.
- [3] Geets X, Tomsej M, Lee JA, Duprez T, Coche E, Cosnard G, et al. Adaptive biological image-guided imrt with anatomic and functional imaging in pharyngo-laryngeal tumors: Impact on target volume delineation and dose distribution using helical tomotherapy. *Radiother Oncol* 2007;85:105–15.
- [4] Devic S, Tomic N, Faria S, Dean G, Lisbona R, Parker W, et al. Impact of FDG PET on biological target volume (btv) definition for treatment planning for non-small cell lung cancer patients. *Nucl Instrum Methods Phys Res Sect A* 2007;571:89–92.

- [5] Riegel AC, Berson AM, Destian S, Ng T, Tena LB, Mitnick RJ, et al. Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion. *Int J Radiat Oncol Biol Phys* 2006;65:726–32.
- [6] Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, et al. Comparison of different methods for delineation of FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med* 2005;46:1342–8.
- [7] Vauclin S, Doyeux K, Hapdey S, Edet-Sanson A, Vera P, Gardin I. Development of a generic thresholding algorithm for the delineation of FDG PET positive tissue: Application to the comparison of three thresholding models. *Phys Med Biol* 2009;54:6901–16.
- [8] Tylski P, Stute S, Grotus N, Doyeux K, Hapdey S, Gardin I, et al. Comparative assessment of methods for estimating tumor volume and standardized uptake value in f-18-fdg pet. *J Nucl Med* 2010;51:268–76.
- [9] Gonzalez RC, Woods RE. *Digital Image Processing*. 3rd ed. Pearson Prentice Hall Upper Saddle River, New Jersey 07458: 2008.
- [10] Haralick RM, Shanmugan K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern* 1973;3:12.
- [11] Tang XO. Texture information in run-length matrices. *IEEE Trans Image Process* 1998;7:1602–9.
- [12] Sun CJ, Wee WG. Neighboring gray level dependence matrix for texture classification. *Comput Vision Graph Image Process* 1983;23:341–52.
- [13] Amadasun M, King R. Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern* 1989; 19:1264–74.
- [14] Xu Y, Sonka M, McLennan G, Guo JF, Hoffman EA. Mdc-based 3-d texture classification of emphysema and early smoking related lung pathologies. *IEEE Trans Med Imaging* 2006;25:464–75.
- [15] Bellotti R, De Carlo F, Tangaro S, Gargano G, Maggipinto G, Castellano M, et al. A completely automated cad system for mass detection in a large mammographic database. *Med Phys* 2006;33:3066–75.
- [16] Szekely N, Toth N, Pataki B. A hybrid system for detecting masses in mammographic images. *IEEE Trans Instrum Meas* 2006;55:944–52.
- [17] Assefa D, Keller H, Menard C, Laperriere N, Ferrari RJ, Yeung I. Robust texture features for response monitoring of glioblastoma multiforme on T1-weighted and T2-flair MR images: A preliminary investigation in terms of identification and segmentation. *Med Phys* 2010;37: 1722–36.
- [18] Yu H, Caldwell C, Mah K, Mozeg D. Coregistered FDG PET/ CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging* 2009;28(3):374–83.
- [19] Yu H, Caldwell C, Mah K, Poon I, Balogh J, MacKenzie R, et al. Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. *Int J Radiat Oncol Biol Phys* 2009;75:618–25.
- [20] El Naqa I, Grigsby PW, Apte A, Kidd E, Donnelly E, Khullar D, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recognit* 2009;42:1162–71.
- [21] Thie JA. Understanding the standardized uptake value, its methods, and implications for usage. *J Nucl Med* 2004; 45(9):1431–4.