PAPER • OPEN ACCESS

Grey Level Co-occurrence Matrix (GLCM) as a Radiomics Feature for Artificial Intelligence (AI) Assisted Positron Emission Tomography (PET) Images Analysis

View the article online for updates and enhancements.

Grey Level Co-occurrence Matrix (GLCM) as a Radiomics Feature for Artificial Intelligence (AI) Assisted Positron Emission Tomography (PET) Images Analysis

Mahbubunnabi Tamal

Imam Abdulrahman Bin Faisal University, PO Box 1982, Dammam 31441, Kingdom of Saudi Arabia

Email: mtamal@yahoo.com; mtamal@iau.edu.sa

Abstract. Positron Emission Tomography (PET) allows tumour microenvironment to be studied in vivo with high sensitivity and specificity. Inter- and intra-tumour morphological and phenotypic heterogeneity or pattern provided by PET images are of critical importance. The traditional practice of visual interpretation of these images are not sufficient enough to extract all the information embedded in the images. On the other hand, simultaneous development of automated and reproducible analysis methodologies makes it possible to extract large amount of quantitative features from these images which is termed as radiomics. Analysis of these radiomics feature using artificial intelligence (AI) can significantly improve individualized treatment selection and monitoring. Grey level co-occurrence matrix (GLCM), a member of texture based radiomics feature family is widely used as a biomarker of heterogeneity and can provide information of the tumour microenvironment. The GLCM can subsequently be used for artificial intelligence (AI) assisted tumour diagnosis, monitoring of progression and treatment planning as well as for monitoring response to therapeutic intervention. This aim of the study was to investigate the accuracy and robustness of PET based GLCM in varying image acquisition and analysis conditions using phantom data. It has been observed that GLCM based textural features (e.g., correlation, entropy, homogeneity, energy contrast and dissimilarity) are not only dependent on the volume but also on the quantization level. They are also dependent on signal-to-noise ratio (SNR) and image contrast. The dependencies of these features to the varying imaging conditions are also not linear and cannot always be directly related. To use these GLCM derived textural features as biomarkers for AI assisted analysis, all the information regarding the textural features should always be included along with the changes in volumes and contrast of the PET images in the training dataset.

1. Introduction

PET radiotracer uptake in tumour is often heterogeneous due to different biological characteristics of tumour cells (e.g., cell proliferation, cell death, differential metabolic activity, vascular structure etc.) and large amount of quantitative features from these images can be extracted which is termed as radiomics. Artificial intelligence (AI) assisted accurate quantification of tumour radiomic features [1] has the potential to be used as a tumour staging and prognostic biomarker [2-3]. Among a number of radiomic features describing tumour heterogeneity [4-6], textural features (homogeneity, correlation, energy, contrast, dissimilarity and entropy) – a second order heterogeneity metric extracted from quantifier based grey level co-occurrence matrices (GLCMs) [7] accounting for both spatial and

intensity information have shown to be capable of staging tumour [8] as well as to predict response [9, 10] for FDG PET images at varying levels.

GLCMs are generated using quantized or resampled intensities within a volume of interests (VOIs) [10] where intensities are resampled in an integer number of bins with the number of bins being power of 2. Textural features extracted from these GLCMs have been reported to be strongly dependent on the metabolically active volume (MATV) using simulated data [11] and confirmed on clinical data [12-17]. Intensity quantization substantially affects the texture indices and thus should be chosen carefully [12, 18]. Reducing quantization always decreases homogeneity [19] and prognostic impact of the textural features is influenced by quantization level [20]. Several groups have suggested using either quantization level 32 [12] or 64 [10,15]. Quantization level 150 or higher also has been proposed in other studies [11,13]. No statistically significant differences have been reported in an another study [10]. Three textural features - homogeneity, dissimilarity and entropy are found to be robust to delineation method and partial volume effects (PVE) [15]. A separate study suggested that smoothing and segmentation have only a small effect compared to quantization [18].

Non uniform selections of parameters and methods across studies make the choice of best textural feature based on MATV, quantization and segmentation challenging and its relationship with the tumour biological characteristics indistinguishable [12,14]. Relationship between volume and quantization has not been explicitly investigated in these studies. No systematic report is available in the literature regarding the effects of image contrast and noise on segmentation and textural features. This aim of the study was to investigate the accuracy and robustness of PET based GLCM in varying image acquisition and analysis conditions using phantom data.

2. Materials and Methods

The torso NEMA phantom containing six spheres (Figure 1) with 10, 13, 17, 22, 28 and 37 mm diameters correspond to 0.52, 1.15, 2.57, 5.58, 11.49 and 26.52 ml volume respectively was filled with ^{18}F solutions. Two different contrasts (2:1 and 4:1) between the spheres and the background were created by reducing the radioactivity in the background.
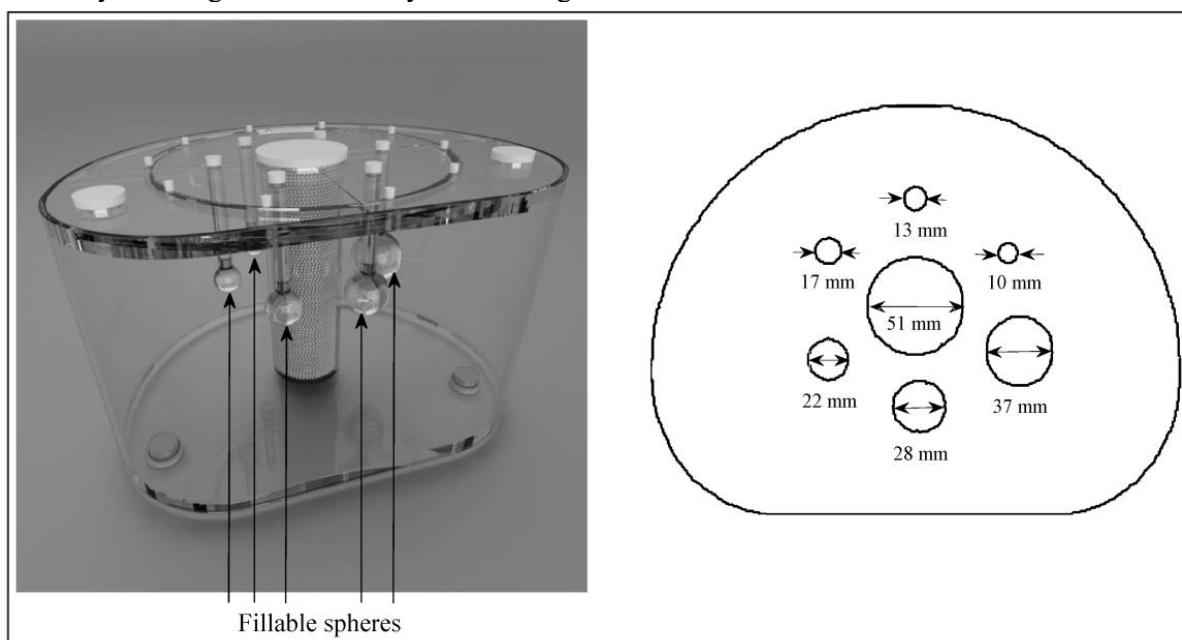

Figure 1: Torso NEMA phantom with a cold insert and six fillable spheres (left) with 10, 13, 17, 22, 28 and 37 mm diameters correspond to 0.52, 1.15, 2.57, 5.58, 11.49 and 26.52 ml volume respectively (right).

The phantom data were acquired in 3D mode on the TrueV PET-CT scanner (Siemens, USA) for 120 minutes which provides 109 image planes or slices covering a 21.6 cm axial FOV (field of view). Images were reconstructed into a 256×256×109 matrix with voxel dimensions of 2.67×2.67×2.00 mm using OSEM reconstruction algorithm with 4 iterations and 21 subsets for five different scan durations (900, 1200, 2000, 4000 and 7200 seconds corresponding to 15, 20, 33.3, 66.6 and 120 minutes respectively) to represent different levels of signal-to-noise ratio (SNR). The starting time of each static frame were shifted to reconstruct five different overlapping realizations for the first four durations. All the reconstructed images were then smoothed with a 4-mm FWHM (full width at half maximum) Gaussian filter after applying decay correction.

All the spheres were delineated using three different segmentation methods. First volume of interest (VOI_{true}) was estimated using the calculated boundaries based on the known diameter and position of each sphere. The second delineation method was a fixed threshold set to 40% (I_{40T}) of the maximum intensity (I_{max}) within the sphere giving a VOI noted as VOI_{40T} [21]. The final volume of interest (VOI_A) was estimated using an adaptive threshold based method as described by (Schaefer et al), where the threshold intensity (I_A) is given by

$$I_A = (\alpha \times I_{70}) + (\beta \times I_{bg}) \tag{1}$$

I_{70} is the mean intensity in a contour containing all voxels with a value greater than 70% of the Imax in the sphere and I_{bg} is the mean background intensity within a sphere of size 26.52 ml located away from all the spheres to avoid partial volume effect (PVE). Both the threshold based methods were applied separately on each roughly delineated VOI containing a sphere to generate the corresponding VOIs. The α and β parameters for the adaptive threshold were calculated using the mean value of optimal cutoff intensities ($I_{optimal}$). $I_{optimal}$ of each hot sphere is calculated using optimal threshold ($T_{optimal}$) and I_{max}. $T_{optimal}$ is estimated as the percentage threshold value of I_{max} which provides the best matched thresholded volume with the VOI_{true} for the uniform sphere phantom.

Quantization of intensities of each VOI was carried out by normalizing the intensities (between 0 and 1) and multiplying the normalized intensities by different quantized values, {Q= 8, 16, 32, 64, 128 and 256}. Grey level co-occurrence matrix (GLCM) was derived for each normalized and quantized VOI data. Several textural features (homogeneity, correlation, energy, contrast, dissimilarity and entropy), a second order heterogeneity measures, were then estimated from these GLCM data.

3. Results

It has been observed that all the textural features are dependent on the quantization value at varying degree. Figure 2 shows the relationships between the mean textural features of five realizations and quantization values for VOI_{true}. Homogeneity exponentially decreases with the increase of quantization levels. Separations among homogeneity for different spheres remain unchanged for different quantization levels. Correlation remains constant with quantization for all spheres from quantization level 32 onwards. However, there are clear separations among correlations for different spheres. Contrast and dissimilarity increase approximately linearly with the increase of quantization levels. Separation among the spheres increases with the increase of quantization levels. Energy decreases and entropy increases with quantization levels. For volumes less than 5.58 cm^3 both entropy and energy remain unchanged after quantization 32. However, it keeps on changing with quantization for bigger spheres and requires higher quantization level to remain unaffected. Higher quantization level would make volumes appear as heterogeneous. On the contrary, low quantization level would make them appear as homogeneous. A compromise is required while choosing appropriate quantization level. Considering all six textural features, quantization level 64 or 32 appears to provide the best compromise. Quantization level 64 has been chosen in this study to generate all the textural features unless mentioned otherwise.
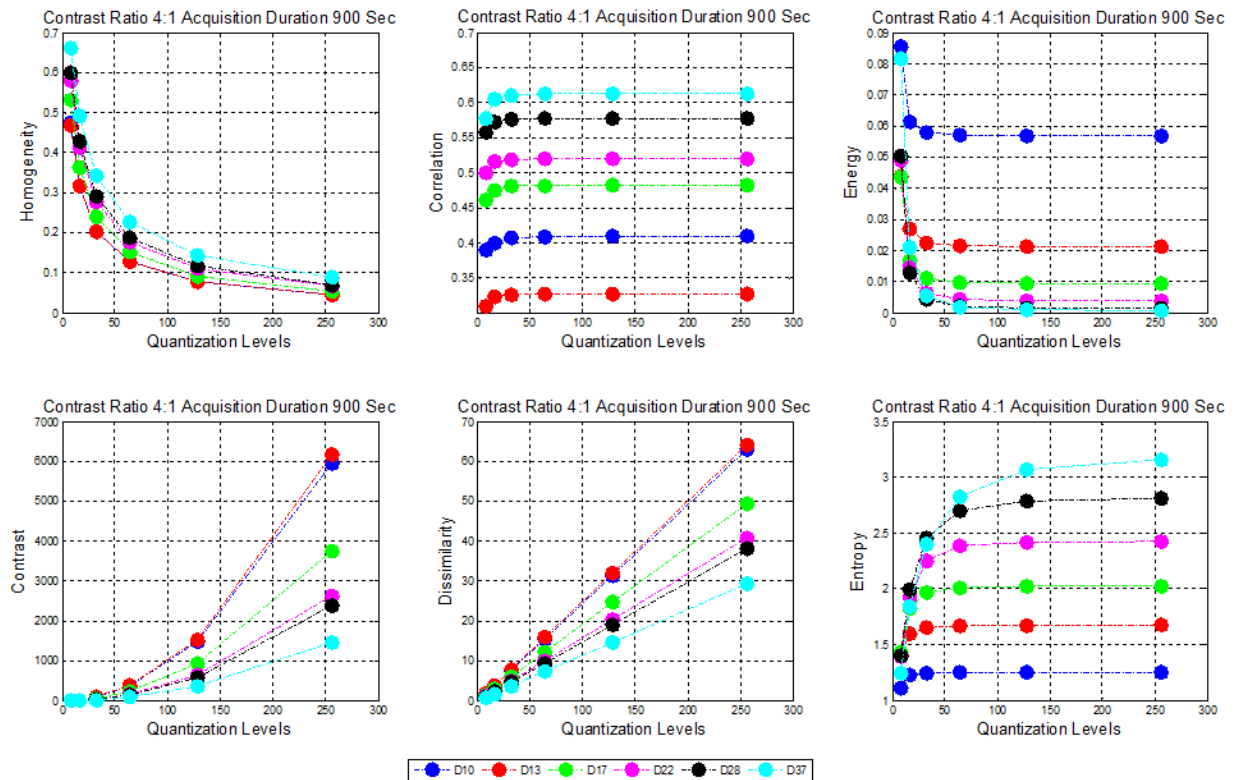
Figure 2: Mean textural features of five realizations against quantization for VOI$_{true}$ for all six spheres. The features shown here are for 4:1 contrast and 900 seconds acquisition duration.

Dependency of textural feature on sphere volume for contrast 4:1 is shown in Figure 3. Features for the spheres located at the background also show dependency on the volumes. Homogeneity and entropy increase with volumes, whereas contrast, dissimilarity and energy decrease. There are subtle differences between the spheres and backgrounds for homogeneity, contrast and dissimilarity showing their dependency on the volume edge. Entropy and energy are robust to edge as shown by very good agreement between the sphere and background. The separation between sphere and background for correlation indicates that it is more dependent on the intensity variations. All the features reaches plateau with the increase of volume at varying rates.
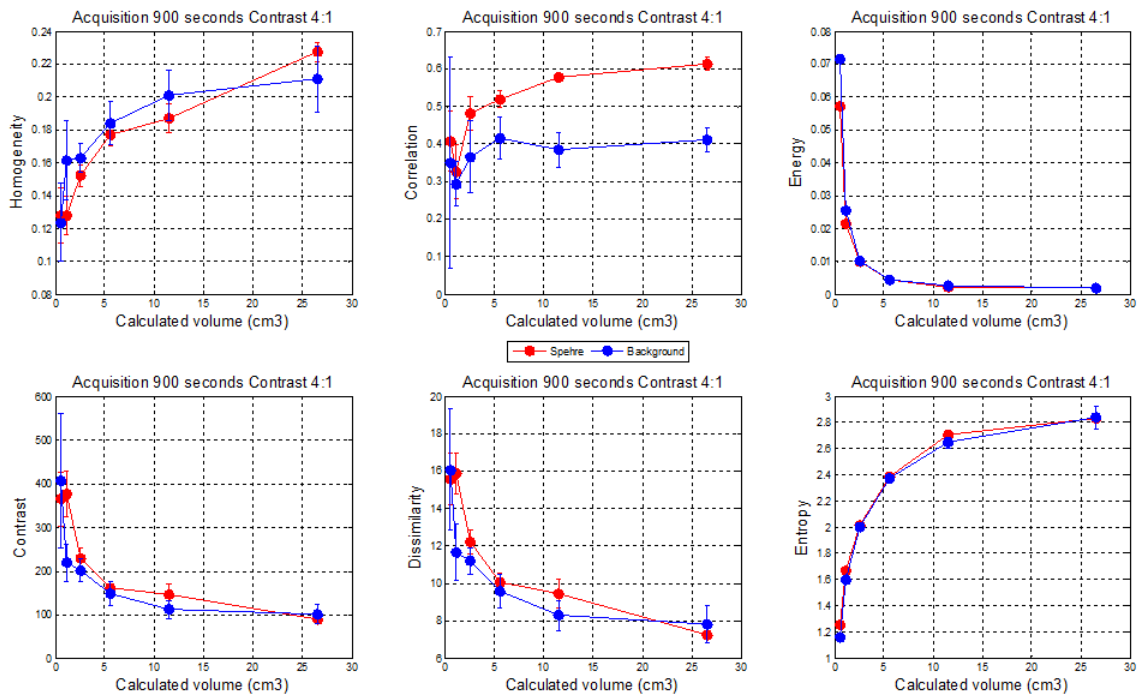
Figure 3: Textural features against volumes for both sphere and background. To analyze the texture in the background same volume as the sphere were used.

Figure 4 compares the relationships between textural features and acquisition durations for VOI_{true} and VOI_{40T} for contrast 2:1. Textural features derived using VOI_{40T} are significantly different than those of VOI_{40T} for smaller volumes. As the volume increases the differences between them reduces. The textural features also vary with the noise as the VOI_{40T} vary with the noise. With an adaptive segmentation method, all textural features become independent of noise for volume greater than 2.57 cm^3 and match closely with the features generated using VOI_{true} (Figure 5).
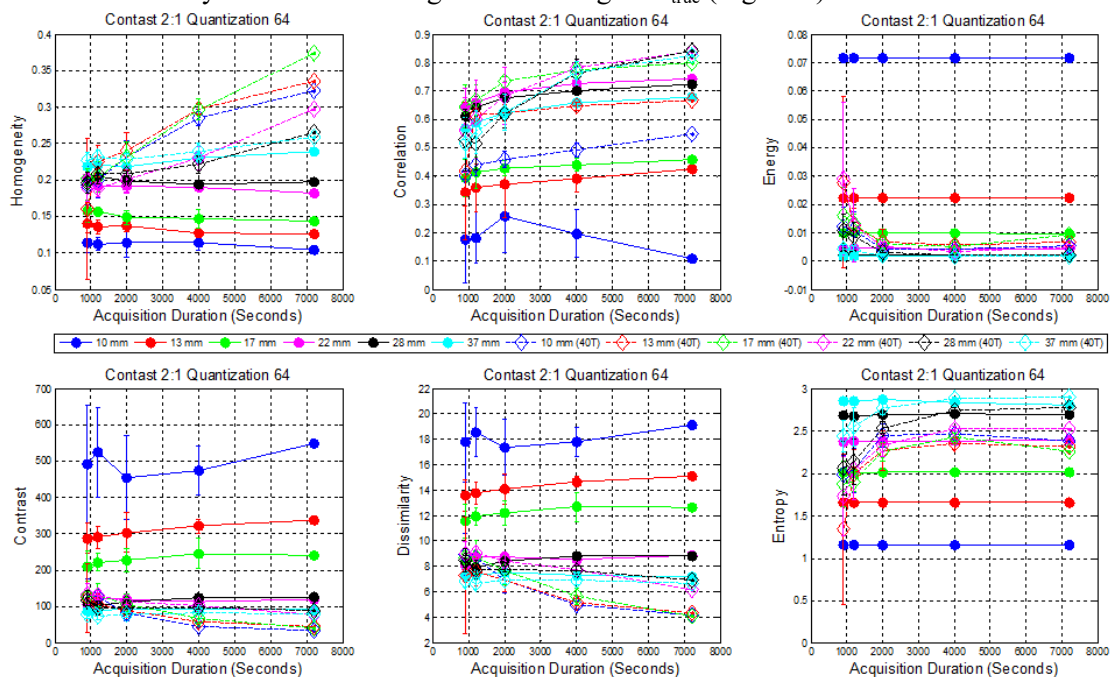

Figure 4: Comparison of textural features extracted using VOI_{40T} and VOI_{true} for different acquisition durations with contrast 2:1 and quantization level 64.

Figure 5: Comparison of textural features extracted using VOI_A and VOI_{true} for different acquisition durations with contrast 2:1 and quantization level 64.
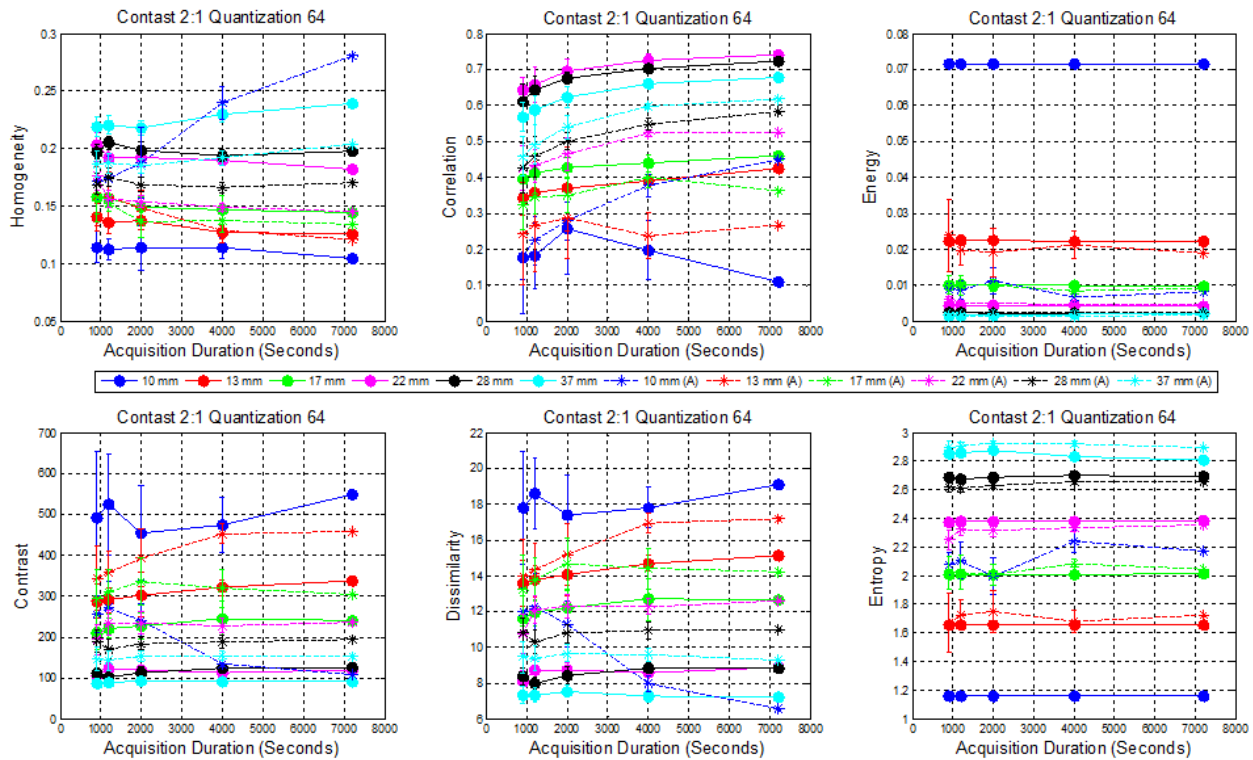
4. Discussion

To use textural features as a tumour staging and prognostic biomarker using AI, better understanding of relationships of textural features with MATV, quantization and segmentation are very important. Investigation of spheres filled with same homogeneous activity reveals that bigger the volume wider the range of intensities making quantization sensitive to the volume of lesions. In such cases, higher quantization makes bigger homogeneous spheres appear as heterogeneous compared to the smaller ones. Lower quantization level removes the dependency on volume by forcing the intensities to be homogeneous and eliminating the heterogeneity information. Considering the characteristics of all the textural features for the homogeneous spheres over a range of volumes, it appears that quantization level 32 or 64 should be preferred and the findings are similar to the finding of previous studies [10,12].

All six textural features are dependent on volumes at varying degrees with entropy and energy being the most sensitive ones. Spheres of similar volumes placed in the background reveals that PVE effect on textural features is far smaller than the effect of volume. Dependency of entropy on MATV significantly reduces for volumes greater than 45 cm^3 for quantization 256 [13]. However, a different study suggested to use volume greater than 10 cm^3 [17] for quantization 64. Dependency of quantization on volume investigated in this study explains the reason for finding two different cut-off volumes. Investigation on heterogeneous spheres suggested that if response occurs as a result of combined changes in volume and heterogeneity, entropy and energy are only able to display changes in volumes (not heterogeneity), making them unsuitable for prognostic biomarkers of heterogeneity. High sensitivity of correlation to intensity also makes it less suitable to report changes in heterogeneity.

Two threshold based delineation methods (40% fixed and adaptive) were employed to investigate the effects of segmentation on textural features. The volumes generated using these two methods are substantially different. Since VOIs delineated using 40% threshold are different from each other, textural features generated using these VOIs are also different with the actual lesion volumes being the

same. However, since VOIs generated using adaptive threshold matches with the VOI_{true}, textural features are closer to the true textural features compare to VOI_{40T}. These results suggested that texture indices are highly sensitive to the segmentation method. The results are consistent with previously published ones [12,15,22].

Volume delineated by a robust segmentation method is capable of generating textural features such as homogeneity, contrast and dissimilarity that are capable of capturing tracer uptake heterogeneity if the volume changes between scans are minimal. Since homogeneity directly related to volume, it can only be used as a feature of image heterogeneity if the changes of volume and homogeneity are in opposite directions, i.e., if the combined multiplicative changes of volumes and homogeneity are either zero or negative. On the other hand, as contrast and dissimilarity are inversely related to volume they can be used as an image heterogeneity feature if the combined multiplicative changes of volumes and homogeneity are either zero or positive. Since contrast is approximately two times more sensitive to volumes compared to dissimilarity, homogeneity and dissimilarity are the two textural features that should be used to measure heterogeneity. These two features also provide complementary heterogeneity information which can be used for cross validation.

5. Conclusion

Homogeneous regions appear heterogeneous on PET images as quantified by textural features. Textural features generated using GLCM depends on quantization and volume. Since these features differentially vary with volume, regions should be segmented using methods are that are robust to variations in contrast and noise using quantization level 64. Small scale heterogeneity phantom studies suggest that homogeneity and dissimilarity are the most suitable textural features to be used as heterogeneity measures where there are combined changes in both heterogeneity and volume due to treatment. Further investigations are required with different heterogeneous phantoms to fully understand the volume effects on these textural indices. Nonetheless, to use these textural features as prognostic biomarkers for an AI assisted system, changes in textural features between baseline and treatment scans should be utilized along with the changes in volumes to train the system.

References

[1]   Choi JY. Radiomics and Deep Learning in Clinical Imaging: What Should We Do? Nuclear Medicine and Molecular Imaging. Mar 2018;52:89–90.

[2]   Eary JF, O'Sullivan F, O'Sullivan J, Conrad EU. Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. Dec 2008;49(12):1973-1979.

[3]   Asselin MC, O'Connor JP, Boellaard R, Thacker NA, Jackson A. Quantifying heterogeneity in human tumours using MRI and PET. European journal of cancer. Mar 2012;48(4):447-455.

[4]   El Naqa I, Grigsby P, Apte A, et al. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. Pattern recognition. Jun 1 2009;42(6):1162-1171.

[5]   Miwa K, Inubushi M, Wagatsuma K, et al. FDG uptake heterogeneity evaluated by fractal analysis improves the differential diagnosis of pulmonary nodules. European journal of radiology. Apr 2014;83(4):715-719.

[6]   van Velden FH, Cheebsumon P, Yaqub M, et al. Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. European journal of nuclear medicine and molecular imaging. Sep 2011;38(9):1636-1647.

[7]   Haralick RM, Shanmugam K, Dinstein IH. Textural Features for Image Classification. Systems, Man and Cybernetics, IEEE Transactions on. 1973;SMC-3(6):610-621.

[8]   Mu W, Chen Z, Liang Y, et al. Staging of cervical cancer based on tumor heterogeneity characterized by texture features on (18)F-FDG PET images. Physics in medicine and biology. Jul 7 2015;60(13):5123-5139.

[9]   Cook GJ, Yip C, Siddique M, et al. Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? Journal of nuclear medicine : official publication, Society of Nuclear Medicine. Jan 2013;54(1):19-26.

[10] Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. Mar 2011;52(3):369-378.

[11] Brooks FJ, Grigsby PW. Low-order non-spatial effects dominate second-order spatial effects in the texture quantifier analysis of 18F-FDG-PET images. PloS one. 2015;10(2):e0116574.

[12] Orlhac F, Soussan M, Maisonobe JA, Garcia CA, Vanderlinden B, Buvat I. Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. Mar 2014;55(3):414-422.

[13] Brooks FJ, Grigsby PW. The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. Jan 2014;55(1):37-42.

[14] Brooks FJ. On some misconceptions about tumor heterogeneity quantification. European journal of nuclear medicine and molecular imaging. Aug 2013;40(8):1292-1294.

[15] Hatt M, Tixier F, Cheze Le Rest C, Pradier O, Visvikis D. Robustness of intratumour (1)(8)F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. European journal of nuclear medicine and molecular imaging. Oct 2013;40(11):1662-1671.

[16] Willaime JM, Turkheimer FE, Kenny LM, Aboagye EO. Quantification of intra-tumour cell proliferation heterogeneity using imaging descriptors of 18F fluorothymidine-positron emission tomography. Physics in medicine and biology. Jan 21 2013;58(2):187-203.

[17] Hatt M, Majdoub M, Vallieres M, et al. 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. Jan 2015;56(1):38-44.

[18] Doumou G, Siddique M, Tsoumpas C, Goh V, Cook GJ. The precision of textural analysis in F-FDG-PET scans of oesophageal cancer. European radiology. May 21 2015.

[19] Brooks FJ, Grigsby PW. FDG uptake heterogeneity in FIGO IIb cervical carcinoma does not predict pelvic lymph node involvement. Radiation oncology. 2013;8:294.

[20] Cheng NM, Fang YH, Chang JT, et al. Textural features of pretreatment 18F-FDG PET/CT images: prognostic significance in patients with advanced T-stage oropharyngeal squamous cell carcinoma. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. Oct 2013;54(10):1703-1709.

[21] Erdi YE, Mawlawi O, Larson SM, et al. Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. Cancer. Dec 15 1997;80(12 Suppl):2505-2509.

[22] Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. May 2012;53(5):693-700.