A TEXTURAL FEATURE BASED TUMOR THERAPY RESPONSE PREDICTION MODEL FOR LONGITUDINAL EVALUATION WITH PET IMAGING

J. George[†‡≀], *P. Claes*[†‡≀], *K. Vunckx*[⋆‡], *S. Tejpar*[*], *C.M. Deroose*[⋆], *J. Nuyts*[⋆‡], *D. Loeckx*[†‡◇], *P. Suetens*[†‡≀]
[†]ESAT/PSI/MIC, [⋆]Nuclear Medicine,[*]Gastroenterology, KU Leuven, Belgium
[‡]Medical Imaging Research Center, UZ Leuven, Belgium [◇]icoMetrix NV, Belgium
[≀]IBBT-KU Leuven Future Health Department, Leuven, Belgium

ABSTRACT

Early therapy response prediction, employing biomarkers such as ^{18}F-fluorodeoxyglucose (FDG) followed with positron emission tomography (PET), is an actively researched topic. Traditionally, only the first order intensity based feature estimates are used for the response evaluations. In this work, we focus on the predictive power of lesion texture along with traditional features in follow up studies. Both standard and textural features are extracted after delineating the lesions with state-of-the-art methods. We propose subspace learning to reduce the influence of delineation parameters and to represent each patient as a Grassmann manifold spanned by the extracted feature subspace. We also propose parallel analysis (PA) to find out the optimal subspace dimensionality. Weighted projection distance between longitudinal subspaces is checked for concordance with the progression outcome using time dependent receiver operating characteristics (ROC). The preliminary clinical results suggest that higher order lesion textures have an added value in response evaluations.

***Index Terms*—** PET, tumor delineation, textural features, subspace learning, Grassmann manifold, principal angles, parallel analysis, survival analysis, time dependent ROC, concordance measure

1. INTRODUCTION

Longitudinal evaluation of tumor lesions has been an indispensable routine in clinical trials for early therapy response detection and prediction tasks. A robust predictive model is crucial for the overall success of follow up studies. We investigated PET-based features since PET image metabolic changes faster than anatomical imaging modalities. Over the years, first order intensity estimates like mean activity, total activity, etc have been utilized for correlating with the response outcome. However, texture information could be equally important or even more informative for the therapy outcome prediction. Textural features estimated from images have established their role in several pattern recognition applications [1–5]. Naqa et al. [6] investigated a logistic regressor incorporating textural and shape features for predicting cancer treatment outcome. Recently, Tixier et al. [7] predicted the response to radio chemotherapy from intra tumor heterogeneity characteristics using textural features.

In this work, we extract traditional features along with textural features from tumor delineations [8–10] for establishing a better predictive model. In order to reduce the effects of delineation parameters, a subspace-based prediction is set up. Each subject at a typical time point can be represented by a subspace in a Grassmann manifold [11,12]; the subspace being spanned by an orthonormal basis of extracted feature estimates. PA [13–15] has been used for selecting the significant principal components defining the dimensionality of the subspace. We use the distance between the subspaces defined at the baseline scan and the follow up scan as a measure of response. The time to progression (TTP) for each subject is taken as the ground truth. Time dependent ROCs [16] are computed relating the continuous marker (weighted subspace-subspace distance) with the continuous outcome (TTP). The resulting concordance summary [17] has been used to indicate the significance of the underlying features.

This work is financially supported by KU Leuven's Concerted Research Action GOA/11/006, Research Foundation - Flanders (FWO) and IWT Agency for Innovation by Science and Technology - Applied Biomedical Research (TBM) project 070717.

2. MATERIALS AND METHODS

2.1. PET lesion delineation

Three competent PET volume delineation methods were selected for this study, namely an adaptive threshold based (ATM) [8], a gradient based (GDM) [9] and a stochastic (FLAB) [10] method. These methods were selected as their performance has been reported to be superior to the threshold based region growing method [18], commonly used in clinical trials, over a wide variety of lesions.

The adaptive threshold method (ATM) [8] is optimized for uniform spherical objects: it computes the optimal threshold as a function of the sphere size, the object-to-background activity ratio and the spatial resolution of the scanner. This relation can be computed and/or determined from phantom measurements. The assumption that the tumor is uniform and spherical is its main limitation. Gradient based segmentation (GDM) [9] assumes a PET image acquisition model with noise and blur components. To compensate for that, this framework has 2 preprocessing steps; denoising and deblurring. An adaptive bilateral filter with an additional denoising component serves as the denoising stage. Landweber's deconvolution algorithm forms the deblurring stage. Just as in the ATM method, it is assumed that the PET system resolution can be modeled as a shift invariant Gaussian PSF. The preprocessed image is segmented using the watershed method (a gradient based segmentation algorithm). The granularity of the watershed segments is made coarser using Ward's hierarchical clustering eventually resulting in the binary segmentation map. The FLAB [10] method uses a fuzzy stochastic expectation maximization (SEM) algorithm to model the partial volume affected voxels and then to delineate lesions. From the classification labels, the binary segmentation map is created by allocating the 50% "fuzzy" classes closer to the tumor "hard" class as belonging to the tumor and the rest to the background. FLAB uses an adaptive prior estimation in a local region of interest between SEM iterations.

In addition to the above mentioned automated methods, five threshold based region growing methods were investigated, viz. T30, T40, T50, T60 and T70 using thresholds of 30%, 40%, 50%, 60% and 70% resp.

TrF - Standardized uptake value SUV	
$\mathrm{SUV}_{\mathrm{peak}}$	Mean activity in $1\mathrm{mL}$ sphere
$\mathrm{SUV}_{\mathrm{maxp}}$	Maximum activity in $\mathrm{SUV}_{\mathrm{peak}}$ sphere
$\mathrm{SUV}_{\mathrm{max}}$	Maximum activity in whole lesion
$\mathrm{SUV}_{\mathrm{mean}}$	Mean activity in whole lesion
TLV	Total lesion volume
TLA	Total lesion activity, $\mathrm{TLA}=\mathrm{SUV}_{\mathrm{mean}}\times\mathrm{TLV}$
TLE	Total lesion energy, $\mathrm{TLE}=\mathrm{SUV}^2_{\mathrm{mean}}\times\mathrm{TLV}$

Table 1. Traditional feature (TrF) estimates employed in this longitudinal study

2.2. Feature extraction

From the delineations, several features were estimated. The traditional features (TrF) extracted in the current study are listed in Table 1. The normalized activity in standardized uptake value (SUV) was used for TrF computation. The tumor lesions were indicated by an expert, and the maximum voxel in a $1\mathrm{mL}$ spherical range was automatically identified. $\mathrm{SUV}_{\mathrm{peak}}$ is the mean activity in a $1\mathrm{mL}$ sphere around that maximum activity $\mathrm{SUV}_{\mathrm{maxp}}$.

An overview of all the textural features (TxF) investigated is listed in Table 2. Prior to textural feature extraction, the images were resampled to the same voxel size in all 3 dimensions and delineated. The delineated images were hard thresholded to $0\sim20$ SUV and further quantized $(0,\ldots,2^L-1)$ to compute the higher order statistics. Five quantization levels were used in this study viz. $L=4,5,6,7,8$. Activity level co-occurrence matrix [1] entries ALCC_{ij} express the frequency of co-occurrences of (i,j) activity level pairs in all possible $3\times3\times3$ subimages in the image. Activity level run length matrix [2] elements ALRL_{ij} represent the frequency of i^{th} activity level's j runs in 13 directions. Activity level size zone matrix [4] entries ALSZ_{ij} is updated when i^{th} activity level has a zone with j voxels. For the activity level difference matrix [5] ALDC, a mean image is computed by replacing each voxel by the average in its 3×3×3 neighborhood except for the voxel itself. For each i^{th} activity level in the actual image, corresponding activity in the mean image is averaged to obtain ALDC_i. Activity level histogram descriptive statistics ALHS were also computed and investigated. 7 TrF and 53 TxF computed with 5 quantization modes totaling $7+5\times53=272$ features are investigated in our study.

2.3. Subspace-based response prediction

After feature extraction, we define a subspace in the Grassmann manifold [11, 12] for each subject in the cohort. The subspace is spanned by m orthonormal basis vectors from a set of different delineation methods in $\mathbb{R}^N$. $\mathbb{R}^N$ is the N-dimensional feature vector space. Each delineation will end up with a point in $\mathbb{R}^N$. From such M delineation points in $\mathbb{R}^N$, we incorporated principal component (PC) analysis (PCA) using singular value decomposition (SVD) to construct the subspace. The optimal dimensionality m of the subspace, which represents the number of significant PCs, were derived using parallel analysis (PA) [13–15]. PA employs random data sets to compute the eigenvalues that parallel the actual dataset and those PCs with eigenvalues greater than the mean or 95^{th} percentile of the random sets are selected. Fig. 1 illustrates the subspace-based response prediction methodology. The subspace-subspace distance between the baseline and the follow up scan is used for predicting early tumor response. The blue cloud represents the baseline subspace span(B_i). The green cloud denotes stable disease $(\mathrm{d_S})$ condition with subspace span(S_j), whereas the red cloud depicts progressive disease $(\mathrm{d_P})$ with span(P_k). In this context, we investigate two

TxF - Activity level co-occurrence matrix ALCC [1]	
$\mathrm{ALCC}_{\mathrm{sam}}$	Second angular moment
$\mathrm{ALCC}_{\mathrm{con}}$	Contrast
$\mathrm{ALCC}_{\mathrm{epy}}$	Entropy
$\mathrm{ALCC}_{\mathrm{cor}}$	Correlation
$\mathrm{ALCC}_{\mathrm{idm}}$	Inverse difference moment
$\mathrm{ALCC}_{\mathrm{hom}}$	Homogeneity
$\mathrm{ALCC}_{\mathrm{dis}}$	Dissimilarity
TxF - Activity level run length matrix ALRL [2, 3]	
$\mathrm{ALRL}_{\mathrm{sre}}$	Short run emphasis
$\mathrm{ALRL}_{\mathrm{lre}}$	Long run emphasis
$\mathrm{ALRL}_{\mathrm{aln}}$	Activity level non uniformity
$\mathrm{ALRL}_{\mathrm{rln}}$	Run length non uniformity
$\mathrm{ALRL}_{\mathrm{rp}}$	Run percentage
$\mathrm{ALRL}_{\mathrm{lare}}$	Low activity level run emphasis
$\mathrm{ALRL}_{\mathrm{hare}}$	High activity level run emphasis
$\mathrm{ALRL}_{\mathrm{srlae}}$	Short run low activity level emphasis
$\mathrm{ALRL}_{\mathrm{srhae}}$	Short run high activity level emphasis
$\mathrm{ALRL}_{\mathrm{lrlae}}$	Long run low activity level emphasis
$\mathrm{ALRL}_{\mathrm{lrhae}}$	Long run high activity level emphasis
TxF - Activity level size zone matrix ALSZ [4]	
$\mathrm{ALSZ}_{\mathrm{sze}}$	Short zone emphasis
$\mathrm{ALSZ}_{\mathrm{lze}}$	Large zone emphasis
$\mathrm{ALSZ}_{\mathrm{aln}}$	Activity level non uniformity
$\mathrm{ALSZ}_{\mathrm{zln}}$	Zone length non uniformity
$\mathrm{ALSZ}_{\mathrm{zp}}$	Zone percentage
$\mathrm{ALSZ}_{\mathrm{laze}}$	Low activity level zone emphasis
$\mathrm{ALSZ}_{\mathrm{haze}}$	High activity level zone emphasis
$\mathrm{ALSZ}_{\mathrm{szlae}}$	Short zone low activity level emphasis
$\mathrm{ALSZ}_{\mathrm{szhae}}$	Short zone high activity level emphasis
$\mathrm{ALSZ}_{\mathrm{lzlae}}$	Large zone low activity level emphasis
$\mathrm{ALSZ}_{\mathrm{lzhae}}$	Large zone high activity level emphasis
TxF - Activity level difference matrix ALDC [5]	
$\mathrm{ALDC}_{\mathrm{coa}}$	Coarseness
$\mathrm{ALDC}_{\mathrm{con}}$	Contrast
$\mathrm{ALDC}_{\mathrm{bsy}}$	Busyness
$\mathrm{ALDC}_{\mathrm{cpx}}$	Complexity
$\mathrm{ALDC}_{\mathrm{txs}}$	Texture strength
TxF - Activity level histogram descriptive statistics ALHS	
$\mathrm{ALHS}_{\mathrm{avg}}$	Mean
$\mathrm{ALHS}_{\mathrm{min}}$	Minimum
$\mathrm{ALHS}_{\mathrm{max}}$	Maximum
$\mathrm{ALHS}_{\mathrm{var}}$	Variance
$\mathrm{ALHS}_{\mathrm{std}}$	Standard deviation
$\mathrm{ALHS}_{\mathrm{skw}}$	Skewness
$\mathrm{ALHS}_{\mathrm{kur}}$	Kurtosis
$\mathrm{ALHS}_{\mathrm{gmn}}$	Geometric mean
$\mathrm{ALHS}_{\mathrm{hmn}}$	Harmonic mean
$\mathrm{ALHS}_{\mathrm{tnmn}}$	Winsorized mean
$\mathrm{ALHS}_{\mathrm{med}}$	Median
$\mathrm{ALHS}_{\mathrm{mod}}$	Mode
$\mathrm{ALHS}_{\mathrm{mdr}}$	Midrange
$\mathrm{ALHS}_{\mathrm{trmn}}$	Trimean
$\mathrm{ALHS}_{\mathrm{rge}}$	Range
$\mathrm{ALHS}_{\mathrm{iqr}}$	Interquartile range
$\mathrm{ALHS}_{\mathrm{aadavg}}$	Average absolute deviation from mean
$\mathrm{ALHS}_{\mathrm{aadmed}}$	Average absolute deviation from median
$\mathrm{ALHS}_{\mathrm{aadmod}}$	Average absolute deviation from mode

Table 2. Textural feature (TxF) estimates employed in this longitudinal study; 7 from ALCC, 11 from ALRL, 11 from ALSZ, 5 from ALDC and 19 from ALHS constituted 53 TxF covariates

possible scenarios viz. a single subspace for a subject and a subspace defined for each lesion in a subject. For the first case, the respective features extracted from all lesions are added together ($\mathrm{pat_{cum}}$) just like in response evaluation criteria in solid tumors (RECIST) prior to the subspace creation. In the latter case, the distance between subspaces for each lesion is analyzed to check for concordance with the progression outcome.

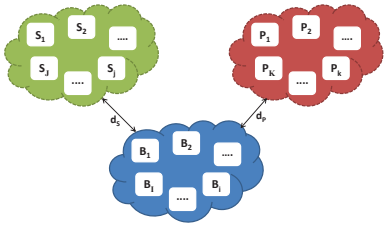


Fig. 1. Blue, green and red subspaces in the Grassmann manifold represent the baseline, stable follow up and progressive follow up cases. The weighted distance between subspaces $\mathrm{d_S}$ (stable) or $\mathrm{d_P}$ (progressive) represents the response measure.

2.4. Grassmann manifold and principal angles

The Grassmann manifold is the set of m-dimensional linear subspaces of the $\mathbb{R}^N$. Let Y_1 and Y_2 be two orthonormal matrices of size N by m. The principal angles $0 \leq \theta_1 \leq \ldots \leq \theta_m \leq \frac{\pi}{2}$ between two subspaces span(Y_1) and span(Y_2) can be computed from the SVD as $\mathrm{Y_1}'\mathrm{Y_2} = \mathrm{U(cos\Theta)V'}$ where $\mathrm{cos\Theta}$ is the diagonal matrix $\mathrm{cos\Theta} = \mathrm{diag(cos\theta_1 \ldots cos\theta_m)}$ The cosines of the principal angles $\mathrm{cos\theta_1, \ldots, cos\theta_m}$ are also known as canonical correlations. In our study, we investigate 3 distance metrics to state the distance between the subspaces, viz. projection metric $\mathrm{d_{proj}(Y_1, Y_2)}$, max correlation $\mathrm{d_{max}(Y_1, Y_2)}$ and min correlation $\mathrm{d_{min}(Y_1, Y_2)}$. $\mathrm{d_{proj}(Y_1, Y_2) = \sqrt{m - \sum_{i=1}^{m} cos^2\theta_i}}$, $\mathrm{d_{max}(Y_1, Y_2) = \sqrt{1 - cos^2\theta_1}}$ and $\mathrm{d_{min}(Y_1, Y_2) = \sqrt{1 - cos^2\theta_m}}$.

2.5. Survival analysis

Proper survival analysis is vital for efficient concordance summary between the observed marker and the disease progression outcome. In our study, the observed marker is the weighted distance metric, $\frac{\|\mathcal{F}\|_{\%}}{1+\mathrm{d}}$, where d is the distance between subspaces. We use % change in $\mathrm{SUV_{max}}$ for $\|\mathcal{F}\|_{\%}$. The efficacy of a binary classifier can be evaluated using a receiver operating characteristics (ROC) curve in terms of the area under the ROC curve (AUC). Nonetheless, we are dealing with a time varying outcome for which a binary classifier is inappropriate. Hence we investigated time dependent ROCs [16] studying the incident sensitivity against the dynamic specificity. The AUCs thus obtained, $\mathrm{AUC}(t)$ depicts the concordance between the marker and the disease status as a time varying function. The concordance summary [17] for a finite follow up period $(0, \tau)$ is computed by the weighted integration of $\mathrm{AUC}(t)$ over time t as given by $C^\tau = \int_0^\tau \mathrm{AUC}(t) w^\tau(t) dt$, where $w^\tau(t) = 2f(t)S(t)/W^\tau$ and $W^\tau = \int_0^\tau 2f(t)S(t)dt = 1 - S^2(\tau)$. Survival function $S(t)$ defines the probability that a subject will survive longer than t and $f(t)$ is the associated density. A random marker will give $C^\tau = 0.5$. In a comparative study, a good marker is one that gives the highest C^τ.

2.6. Clinical data

15 patients suffering from metastatic colorectal cancer and treated with cetuximab were selected for the study. Baseline FDG PET scans were acquired prior to treatment. The follow up scanned at 1 week after 1^{st} dose of the treatment was investigated for early response. The images were maximum likelihood expectation maximization (MLEM) reconstructed with $5\mathrm{mm}$ post smoothing. One to five lesions were identified per patient by an expert nuclear medicine physician. The time to progression (TTP) for the subjects was taken as the ground truth. The Kaplan-Meier estimate depicting the probability that a subject will remain free of progression is given in Fig. 2.

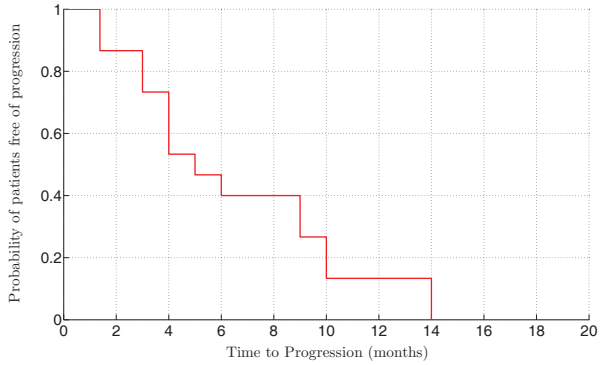


Fig. 2. Kaplan-Meier plot showing the probability of progression free survival as a function of time (in months)

3. EXPERIMENT

^{18}F-FDG PET scans at baseline were expert marked and the lesions were delineated using 8 discussed methods, with the parameters conferred in [18, 19]. Two evaluation directions (EVD) were followed in order to obtain the follow up delineation viz. independent delineation or rigid registration with baseline delineation. We assume that using subspace-based methods, the influence of delineation parameters on the response prediction decision making process will be nullified. To understand the influence of features separately, 3 cases are investigated viz. TrF, TxF, TrF+TxF with $N = 7,265,272$ resp. TxF were extracted after resampling the delineated tumor (from actual $3.38\mathrm{mm} \times 3.38\mathrm{mm} \times 3\mathrm{mm}$) to $3\mathrm{mm} \times 3\mathrm{mm} \times 3\mathrm{mm}$ voxel size. On course of finding orthonormal basis spanning the subspace from the 16 feature vectors (extracted from 16 EVD) in $\mathbb{R}^D$, PCA was used with number of PCs selected after PA. 3 scenarios were investigated viz. $\mathrm{pa_{null}}$, $\mathrm{pa_{mean}}$ and $\mathrm{pa_{perc}}$. In $\mathrm{pa_{null}}$, PCA with the best 5 PCs, without PA, was used to span the subspace. $\mathrm{pa_{mean}}$ and $\mathrm{pa_{perc}}$ opted for PCs with eigenvalue cutoff set based on mean and 95 percentile resp. Distance metrics are calculated for each subject both in the cumulative mode ($\mathrm{pat_{cum}}$) and for each tumor. In the latter case, the observation marker is defined as either average ($\mathrm{tmr_{avg}}$), minimum ($\mathrm{tmr_{min}}$), maximum ($\mathrm{tmr_{max}}$), or summation ($\mathrm{tmr_{sum}}$) of the individual weighted distance metrics. For reference, C^τ of $\mathrm{SUV_{max}}$ under $\mathrm{pat_{cum}}$, $\mathrm{tmr_{avg}}$, $\mathrm{tmr_{min}}$, $\mathrm{tmr_{max}}$ and $\mathrm{tmr_{sum}}$ are 0.5856, 0.6145, 0.6337, 0.5368 and 0.6419 resp.

Table 3 listed C^τ obtained with $\mathrm{pat_{cum}}$ with varying covariates, distance metrics and PA mode. Low C^τ for $\mathrm{pa_{null}}$ indicated the benefit of PA. We selected $\mathrm{pa_{perc}}$ as the optimal setting. In our studies, $\mathrm{pa_{perc}}$ identified 1 to 4 PCs and as a result the $\mathrm{d_{proj}}$ and $\mathrm{d_{min}}$ showed similar trend in concordance with TTP. Considering the low dimensionality of subspace span obtained with PA, it's normal to have similar results with $\mathrm{d_{min}}$ and $\mathrm{d_{proj}}$. Nonetheless, we prefer $\mathrm{d_{proj}}$ as our future setting. The added value of TxF

over TrF was clear from the relatively higher C^{τ} obtained for TxF and TrF+TxF. Table 4 accounted for C^{τ} computed from $\mathrm{tmr}_{\mathrm{avg}}$,

Features		TrF	TxF	TrF+TxF
$\mathrm{d}_{\mathrm{proj}}$	$\mathrm{pa}_{\mathrm{null}}$	0.6166	0.4844	0.4844
	$\mathrm{pa}_{\mathrm{mean}}$	0.5789	**0.6817**	**0.6817**
	$\mathrm{pa}_{\mathrm{perc}}$	0.5789	**0.6817**	**0.6817**
$\mathrm{d}_{\mathrm{max}}$	$\mathrm{pa}_{\mathrm{null}}$	0.5856	0.5856	0.5856
	$\mathrm{pa}_{\mathrm{mean}}$	0.5789	**0.6817**	**0.6817**
	$\mathrm{pa}_{\mathrm{perc}}$	0.5789	**0.6817**	**0.6817**
$\mathrm{d}_{\mathrm{min}}$	$\mathrm{pa}_{\mathrm{null}}$	0.6166	0.4844	0.4844
	$\mathrm{pa}_{\mathrm{mean}}$	0.5789	**0.6817**	**0.6817**
	$\mathrm{pa}_{\mathrm{perc}}$	0.5789	**0.6817**	**0.6817**

Table 3. C^{τ} obtained with $\mathrm{pat}_{\mathrm{cum}}$

$\mathrm{tmr}_{\mathrm{min}}$, $\mathrm{tmr}_{\mathrm{max}}$ and $\mathrm{tmr}_{\mathrm{sum}}$ with distance $\mathrm{d}_{\mathrm{proj}}$ and PA setting $\mathrm{pa}_{\mathrm{perc}}$. The best concordance summary was obtained for $\mathrm{tmr}_{\mathrm{min}}$.

Features			TrF	TxF	TrF+TxF
$\mathrm{d}_{\mathrm{proj}}$	$\mathrm{pa}_{\mathrm{perc}}$	$\mathrm{tmr}_{\mathrm{avg}}$	0.6387	0.6490	0.6490
		$\mathrm{tmr}_{\mathrm{min}}$	0.6617	**0.6676**	**0.6676**
		$\mathrm{tmr}_{\mathrm{max}}$	0.5368	0.5892	0.5892
		$\mathrm{tmr}_{\mathrm{sum}}$	0.6130	0.6011	0.6011

Table 4. C^{τ} obtained with $\mathrm{tmr}_{\mathrm{avg}}$, $\mathrm{tmr}_{\mathrm{min}}$, $\mathrm{tmr}_{\mathrm{max}}$ & $\mathrm{tmr}_{\mathrm{sum}}$

4. CONCLUSION

In this work, we primarily focused on the predictive power of lesion texture along with traditional features in follow up studies. Delineation schemes are known to be heavily influenced by the parameter tuning and our subspace-based prediction is a way to reduce that influence. In addition to that, we solve the problem of defining the optimal dimensionality of the subspace by incorporating PA; which to our best knowledge has not been done before in this context. Our predictive model using weighted subspace-subspace distance metrics resulted in a concordance summary C^{τ} of 0.68 with a combination of textural and traditional features. This result obtained with the current data denotes an added value in using textural information for therapy response evaluation. Moreover, therapy response could be predicted as early as 1 week after 1st dose of treatment from the feature estimates. We also found that parallel analysis is necessary for selecting the optimal PCs to construct the orthonormal basis. In future, we plan to investigate this with more clinical trials to further validate our predictive model.

5. REFERENCES

[1] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man and Cybern.*, vol. 3, no. 6, pp. 610–621, nov. 1973.

[2] M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graph. Image Process.*, vol. 4, no. 2, pp. 172–179, June 1975.

[3] X. Tang, "Texture information in run-length matrices," *IEEE Trans. Image Process.*, vol. 7, no. 11, pp. 1602–1609, nov 1998.

[4] G. Thibault, B. Fetil, C. Navarro, S. Pereira, P. Cau, N. Levy, J. Sequeira, and J. Mari, "Texture indexes and gray level size zone matrix: application to cell nuclei classification," *Pattern Recognition In Process.*, pp. 140–145, 2009.

[5] M. Amadasun and R. King, "Textural features corresponding to textural properties," *IEEE Trans. Syst., Man and Cybern.*, vol. 19, no. 5, pp. 1264–1274, sep/oct 1989.

[6] I. El Naqa, P. W. Grigsby, A. Apte, E. Kidd, E. Donnelly, D. Khullar, S. Chaudhari, D. Yang, M. Schmitt, R. Laforest, W. L. Thorstad, and J. O. Deasy, "Exploring feature-based approaches in PET images for predicting cancer treatment outcomes," *Pattern Recogn.*, vol. 42, pp. 1162–1171, June 2009.

[7] F. Tixier, C. C. Le Rest, M. Hatt, N. Albarghach, O. Pradier, J. P. Metges, L. Corcos, and D. Visvikis, "Intratumor heterogeneity characterized by textural features on baseline ^{18}F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer," *J. Nucl. Med.*, vol. 52, no. 3, pp. 369–378, Mar 2011.

[8] J. A. Van Dalen, A. L. Hoffmann, V. Dicken, W. V. Vogel, B. Wiering, T. J. Ruers, N. Karssemeijer, and W. J. G. Oyen, "A novel iterative method for lesion delineation and volumetric quantification with FDG PET," *Nucl. Med. Communications*, vol. 28, no. 6, pp. 485–493, 2007.

[9] X. Geets, J. A. Lee, A. Bol, M. Lonneux, and V. Grégoire, "A gradient-based method for segmenting FDG-PET images: methodology and validation," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 34, pp. 1427–1438, 2007.

[10] M. Hatt, C. C. Le Rest, A. Turzo, C. Roux, and D. Visvikis, "A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET," *IEEE Trans. Med. Imag.*, vol. 28, no. 6, pp. 881–893, 2009.

[11] J. Hamm, "Subspace-based learning with Grassmann kernels," *PhD Dissertation, The University of Pennsylvania*, 2008.

[12] J. Hamm and D. D. Lee, "Grassmann discriminant analysis: a unifying view on subspace-based learning," *Proc. Int. Conf. Mach. Learning*, pp. 376–383, 2008.

[13] J. L. Horn, "A rationale and test for the number of factors in factor analysis," *Psychometrica*, vol. 30, pp. 179–185, 1965.

[14] S. B. Franklin, D. J. Gibson, P. A. Robertson, J. T. Pohlmann, and J. S. Fralish, "Parallel analysis: a method for determining significant principal components," *Journal of Vegetation Science*, vol. 6, no. 1, pp. 99–106, 1995.

[15] B. P. O'Connor, "SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test," *Behavior Research Methods, Instruments, & Computers*, vol. 32, no. 3, pp. 396–402, 2000.

[16] P. J. Heagerty, T. Lumley, and M. S. Pepe, "Time-dependent ROC curves for censored survival data and a diagnostic marker," *Biometrics*, vol. 56, no. 2, pp. 337–344, 2000.

[17] P. J. Heagerty and Y. Zheng, "Survival model predictive accuracy and ROC curves," *Biometrics*, vol. 61, no. 1, pp. 92–105, 2005.

[18] J. George, K. Vunckx, S. Tejpar, C. M. Deroose, J. Nuyts, D. Loeckx, and P. Suetens, "Quantitative comparison of automated PET volume delineation methodologies using simulated tumor lesions," *Proc. IEEE ISBI*, pp. 653–656, 2011.

[19] J. George, K. Vunckx, S. Tejpar, C. M. Deroose, J. Nuyts, D. Loeckx, and P. Suetens, "Fuzzy statistical unsupervised learning based total lesion metabolic activity estimation in positron emission tomography images," *Proc. MICCAI MLMI*, pp. 233–240, 2011.