Tumor heterogeneity and gene mutation: A combined study for analysis of radiation therapy efficacy in head-and-neck carcinoma patients

Apurva Singh[1], Sharad Goyal[2], Yuan James Rao[2], and Murray Loew[3]

Departments of [1]Electrical and Computer Engineering, [2]Radiation Oncology, and [3]Biomedical Engineering, George Washington University, Washington, D.C. USA

INTRODUCTION

The purpose of this study was to determine if the information present in the texture of tumor regions in the pre-treatment PET-CT scans of Head-and-Neck Squamous Cell Carcinoma patients can be a useful measure of the efficacy of radiation therapy in their treatment. We have now extended our study to include the gene mutation information of a group of patients to see if it can be used as an additional feature in the determination of treatment efficacy.

DATABASES USED

1. HNSCC
2. Head-Neck
3. TCGA-HNSC
4. TCGA-LUSC and TCGA-LUAD

Clinical data accompanying the database are used to divide the patients into two categories: Recurrent and Non-recurrent. The number of PET-CT scans included for each patient varied from 4 to 20 depending on the tumor size.

Database	Patients		Slices	
	Recurrent	Non-recurrent	Recurrent	Non-recurrent
HNSCC	6	14	71	120
Head-Neck	17	36	153	211
Combined	23	50	224	331
TCGA-HNSC	6	5	30	25
TCGA-LUSC	3	12	15	62
TCGA-LUAD	6	3	34	7

METHODS

Tumor boundary information is present in the form of RT Structure DICOM files. These RT Structure boundaries are overlain on the PET scans of the patients. Screenshots of the scans are then compared with the original database images viewed in RadiAnt DICOM Viewer.
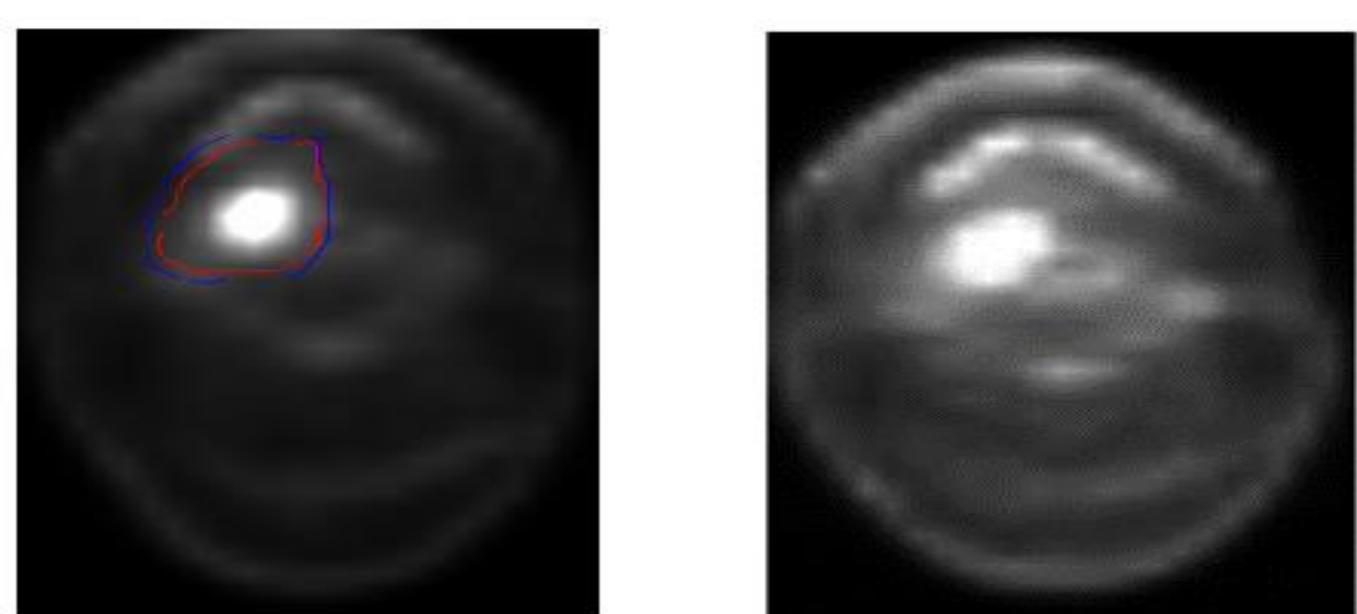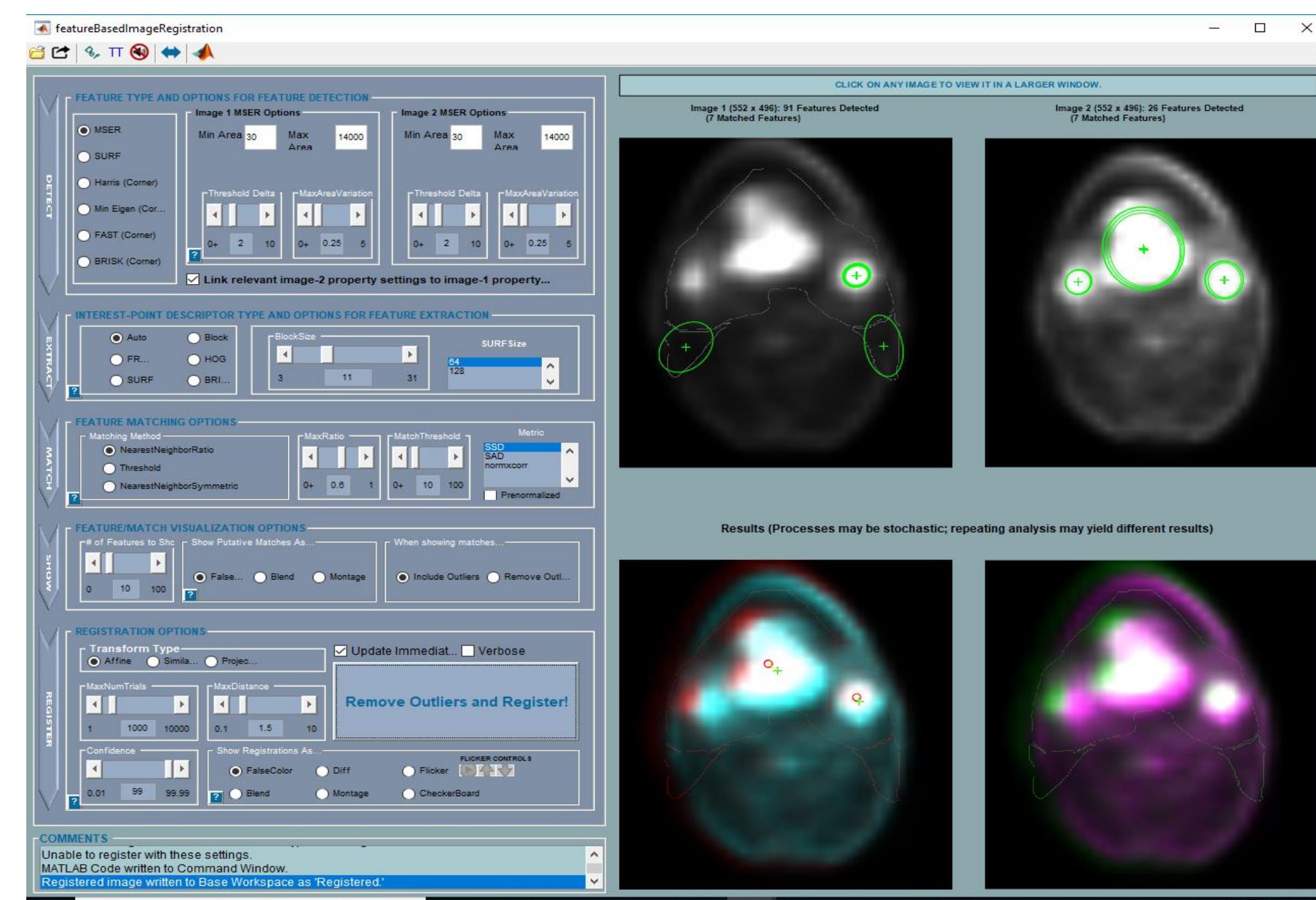

(a)   (b)

Fig1: (a) is the screenshot which shows the tumor region marked in the PET image and (b) shows the corresponding original PET slice as viewed in RadiAnt

Both these images are registered using the feature-based image registration app in MATLAB.

An algorithm is developed in MATLAB to extract the region of the PET scan that is included within the tumor boundaries indicated on the scans.
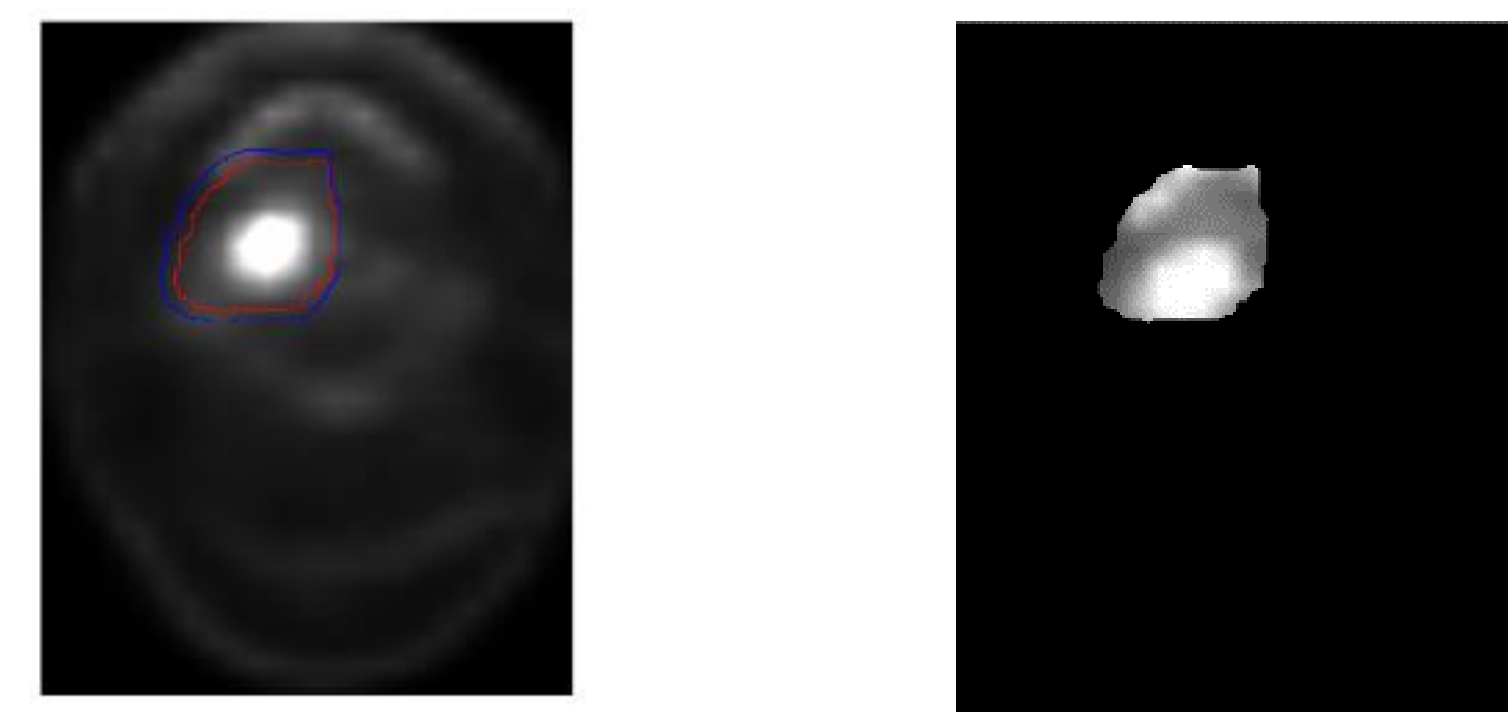

Fig 2 : The images show the tumor region marked in the PET image and the region as extracted from the original image.

Texture analysis of the extracted tumor regions is performed in the following two ways [1]:
a) Treating each scan as a 2D image: Features: Laws, GLCM, Fourier, Hu's
b) Treating each scan as a 3D volume: Features: 3D GLCM, GLRLM, GLSZM, NGTDM

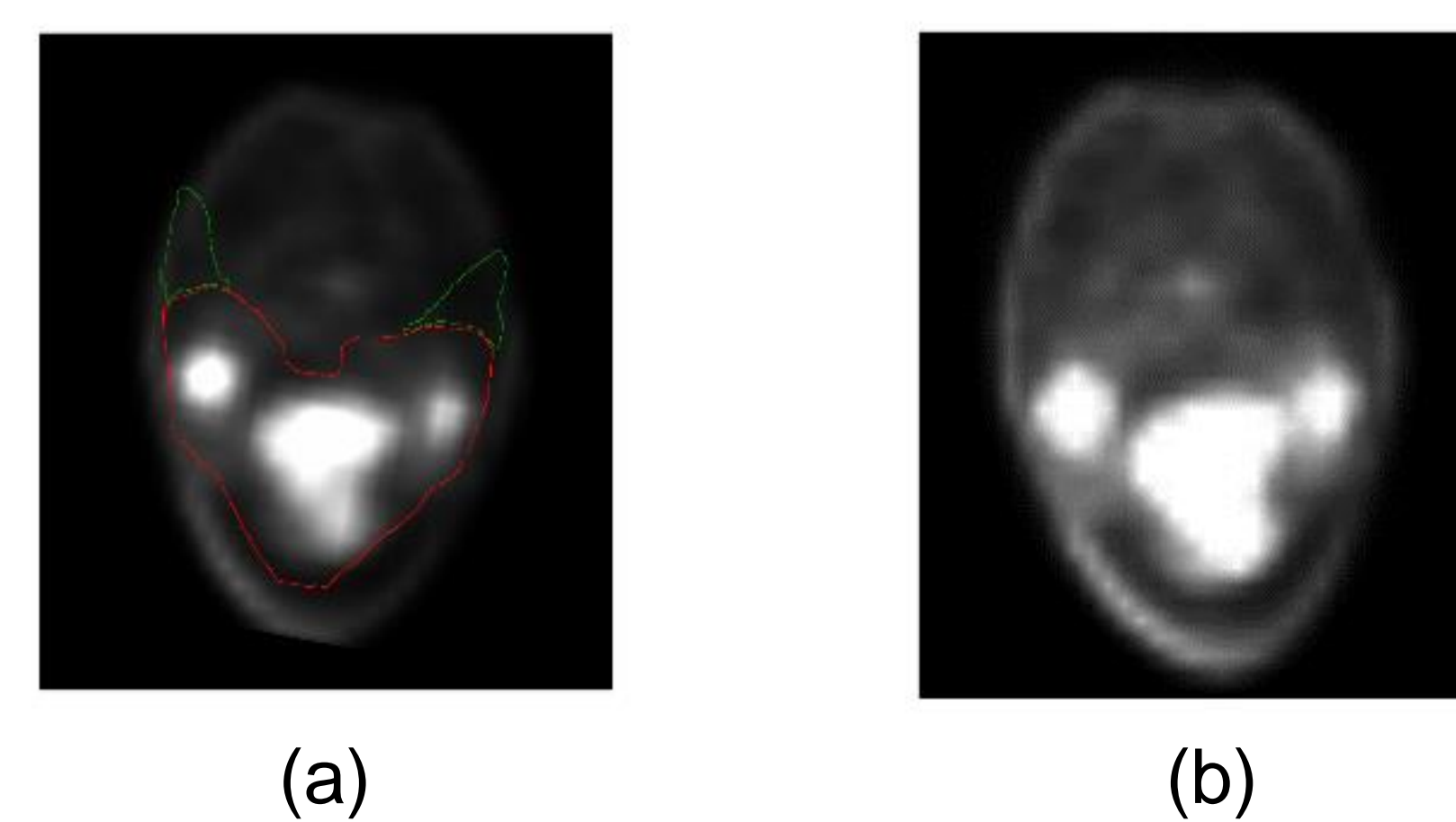For database augmentation, we perform geometric transformations (translation and rotation on the original images).

(a)   (b)

Fig 3: Fig (a) shows the transformed hospital screenshot image and Fig (b) shows the transformed original RadiAnt image

RESULTS

Slice-wise-2D

Database	Recurrent	Non-recurrent	Accuracy	Classifier
HNSCC	63%	84%	76%	Weighted KNN
Head-Neck	61%	84%	72%	Weighted KNN
Combined	71%	71%	71%	Bagged Trees

Slice-wise-3D

Database	Recurrent	Non-recurrent	Accuracy	Classifier
HNSCC	55%	92%	78%	Bagged Trees
Head-Neck	50%	73%	61%	Linear SVM
Combined	53%	78%	68%	Bagged Trees

Patient-wise-2D

Database	Recurrent	Non-recurrent	Classification accuracy
HNSCC	83%	86%	85%
Head-Neck-PT-CT	41%	97%	79%
Combined	61%	86%	78%

Patient-wise-3D

Database	Recurrent	Non-recurrent	Classification accuracy
HNSCC	50%	93%	80 %
Head-Neck-PT-CT	47%	89%	75 %
Combined	61%	92%	82 %

Genomic features-TCGA-HNSC

Patients	PIK3CA	CDKN2A	TP53
TCGA-BA-4077	1	0	0
TCGA-CN-5370	0	1	1

Patient-wise accuracies- 2D:3D

Feature	Recurrent	Non-recurrent	Classifier used
Texture	80:100	60:100	Fine tree: Fine tree
Texture + gene	100:100	100:100	Fine tree: Fine tree

TCGA-LUSC and TCGA-LUAD

The following table shows a comparison of results when binary features describing patterns of the following genes:TP53,FGRL1,PIK3CA,NRF2,DDR2,PTEN,LKB1 are combined with texture features.

Patient-wise accuracies-2D:3D

Feature	Recurrent	Non-recurrent	Overall accuracy	Classifier used
Texture	44:33	66:53	58:45	Medium KNN: RusBoosted Trees
Texture + gene	40:55	60:60	54:58	Medium KNN: RusBoosted Trees

1. Wavelet decomposition features were included in the texture feature set.
2. Non-rigid (shear and rotation) transformation used as a different database augmentation technique.

(a)   (b)

Fig 4 : The image (a) show the original tumor region and (b) shows the tumor region after application of shear and clockwise rotation.

Database augmentation is carried out to have equal number of patients (9) and slices (70) in both recurrent and non-recurrent cases. Classification accuracy results are compared using texture features alone (Case A)and on combining wavelet features with them (Case B).

Slice-wise-2D

Case A		Case B
Overall accuracy	59%	84%
Recurrent	51%	83%
Non-recurrent	56%	86%
AUC	0.60	0.87

CONCLUSION

Our experiments show that identification of gene expression patterns in head and neck carcinoma patients can provide information which, when combined with tumor heterogeneity measures, can improve therapy response prediction scores. Our work with lung cancer patients and our current analysis of patients of TCGA-CESC (Cervical cancer patients) is an effort to develop a more comprehensive approach of therapy personalization.

REFERENCES

[1] Vallières, Martin, et al. "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities." *Physics in Medicine & Biology* 60.14 (2015): 5471.
[2] Hill, Derek LG, et al. "Medical image registration." Physics in medicine and biology 46.3 (2001): R1.