Radiomics for identification of active bone marrow from CT: an exploratory study

S. Rosati, G. Balestra, Member, IEEE
Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
gabriella.balestra@polito.it

P. Franco, C. Fiandra, F. Arcadipane, P. Silvetti, U. Ricardi
Department of Oncology, Radiation Oncology
Università di Torino
Torino, Italy

E. Gallio
Medical Physics Unit,
A.O.U. Città della Salute e della Scienza di Torino,
Torino, Italy

***Abstract*—The radiation dose received by the pelvic Bone Marrow (BM) is a predictive factor for Hematologic Toxicity (HT) occurrence in the treatment of anal cancer. For this reason it is important to avoid BM during radiotherapy. In particular, the standard strategy in these cases consists in the identification of hematopoietically active BM (*actBM*), i.e. the part of BM in charge of blood cells generation, on ^{18}FDG-PET, FLT-PET or MRI, but no approached have been developed for identifying actBM from CT images. This exploratory study aims to use radiomics for detecting *actBM* on CT sequences. Our approach is based on the extraction of 36 first-order and texture (second-order) features for each CT slice. These features are used as input of a Decision Tree (DT) classifier able to discriminate between active and inactive BM regions on the images. This method was applied to five patients affected by carcinoma of the anal canal and the obtained *actBM* segmentation was compared with the standard *actBM* identification from ^{18}FDG-PET (reference standard, RS). Our results show that *actBM* identification in lumbosacral and iliac structures using radiomics overlaps the RS for more than 75% in 4 out of 5 patients.**

Keywords—texture features, radiomics, decision tree, computed tomography (CT), hematopoietically active bone marrow.

I. Introduction

Actually, the standard treatment option for anal cancer patients is concurrent chemo-radiation [1]. Despite the use of high-tech delivery approaches such as Intensity-Modulated Radiation Therapy (IMRT), toxicity remains not negligible. Acute Hematologic Toxicity (HT) may affect compliance to therapy, increasing the probability to develop asthenia, bleeding, or infections [2]. Chemotherapy is the most important trigger for HT, but also radiation plays an important role [3]. In adults, pelvic bones and lumbar spine contain almost 60% of total Bone Marrow (BM) [4]. Since the dose received by the BM comprised within the pelvic bones is a predictive factor for HT occurrence, a crucial point in radiotherapy is to implement strategies able to selectively spare BM, by correctly identifying and delineating this structure. Although the use of the whole pelvic bones as surrogates is the most inclusive method with respect to BM, other strategies have been developed. One consists in selectively avoiding the hematopoietically active BM (*actBM*), that is the part of BM in charge of blood cells generation. Several methods exist to identify active BM, based on morphological and functional imaging[5], such as magnetic resonance imaging (MRI), single-positron emission tomography (SPECT), fluorodeoxyglucose (^{18}FDG)-labeledor fluorothymidine (FLT)-labeled positron emission tomography (PET) [6–8]. ^{18}FDG-PET is considered an optional exam by international guidelines in the diagnostic work-out and hence not all patients have it. Computed tomography (CT), however, is a widespread exam and all patients have to undergo it before starting radiotherapy. Thus, an approach based on CT images to define *actBM* would be very useful and would have a potential broader applicability. To the best of our knowledge, no approaches have been proposed to identify *actBM* using CT.

Radiomics identifies the "process designed to extract a large number of quantitative features from digital images" [9]. Its final aim is to extract hidden information from images that can be useful to develop Computer-Aided Diagnosis (CAD) systems and decision support. Since images are processed as matrices of data, radiomics can be potentially applied to all kinds of images, such as CT, MRI, PET. In the study by Rosati et al. [10] it was proved that applying radiomics to carotid ultrasound images allows for identifying high cardiovascular risk subjects and capturing the progressive development of pathological conditions in vessel wall layers. Giannini et al. [11] used radiomics to increase specificity of a CAD system for prostate cancer detection based on MRI. The basic step of radiomics is feature extraction, consisting in the calculation of variables able to quantitatively describe the characteristics of the anatomical structure of interest. Three main groups of features can be used: first-order statistical features, simply describing the statistical distribution of the intensities in a given region, second-order or texture features, describing the spatial distribution of the intensities in a given region, and higher-order features, aiming to find specific patterns.

The aim of this study is to evaluate the feasibility of detecting *actBM* by applying radiomics to CT images.

II. Materials and Methods

A. Population

Five patients with locally advanced squamous cell carcinoma of the anal canal and/or margin were involved in this

study, all belonging to the Oncology Dept. institutional database. The patients were treated with concurrent chemo-radiation employing volumetric modulated arc therapy and concomitant 5-fluorouracil and mytomicin C. All patients underwent ^{18}FDG-PET-CT exam to complete the diagnostic and staging work-up, and a non-contrast enhanced CT of the pelvic region for planning purposes before the beginning of radiotherapy. The study agreed with the ethical principle of the Helsinki Declaration and all participants signed a written informed consent to be included in the study.

B. CT and ^{18}FDG-PET acquisition

For the CT acquisition, patients had a virtual simulation procedure in supine position with both an indexed shaped knee rest and ankle support (CIVCO Medical Solutions, Kalona, IA, USA), without custom immobilization. Planning CT was performed according to standard clinical scanning protocols at the Oncology Dept. with a Philips ''BigBore'' CT scanner (Philips Medical System, Eindhoven, NL). The most common pixel spacing was (0.93 mm, 0.93 mm, 3 mm) for CT. ^{18}FDG - PET scans were performed with a Philips Gemini PET/CT tomography. Data acquisition started 90 min after an intravenous injection of approximately 30 MBq/kg body weight of ^{18}F-glucose. First, a total body CT scan was acquired; then, PET scans of the same area were collected for 2.5 min/bed position. A dedicated fused imaging workstation (Extended Brilliance Workspace 2.0) was adopted for PET clinical interpretation.

C. Reference Standard

- *Bone marrow segmentation on planning CT*

The procedure described in [12] was applied to segment the whole pelvic bone marrow (PBM) on the planning CT. Then, the PBM was manually divided into three subregions by a dedicated radiation oncologist: a) the iliac BM (IBM), comprising the area between the iliac crests and the upper border of femoral head; b) lower pelvis BM (LPBM), including bilateral pube, ischia, acetabula and proximal femura, from the upper border of the femoral heads to the lower aspect of the ischial tuberosities; c) lumbosacral BM (LSBM), comprising the area between the superior border of L5 somatic body to the lower aspect of the coccyx.

- *Active bone marrow segmentation on ^{18}FDG-PET*

We exported all images derived from planning CT on the VELOCITY platform (Varian Medical Systems, Palo Alto, CA), together with treatment volumes, organs at risk (OARs) and dose references. Given that ^{18}FDG-PET images were acquired separately, a deformable co-registration between CT and PET images was performed. Then, we calculated the ^{18}FDG-PET standardized uptake values (SUVs) for the PBM volume of all subjects, after correcting for body weight. To standardize SUVs among all patients, we normalized BM and liver SUVs. For each patient, we labeled as *actBM* the portion of PBM presenting SUV values higher than the patient mean SUV [8]. The difference between PBM and *actBM* was labeled as inactive BM (*inactBM*). Figure 1 highlights PBM (green line) and *actBM* (red line) as identified with the use of ^{18}FDG-PET in a specific patient. The difference between these two areas identified *inactBM.*

D. Active bone marrow identification using radiomics

The process starts by removing the cortical bone from the regions of interest (ROIs); in the second step the image was divided in overlapping 5 by 5 pixels elements and 36 features were computed for each element; the third step consisted in the training set and classifier construction; pixel classification was obtained in the fourth step by means of a voting system; finally a post-processing was applied to the identified regions to enhance the classification accuracy. All the steps are described in details in the following.

- *Removal of cortical bone*

From the whole PBM volume, we removed the cortical bone since it less likely contains BM. For this purpose, we processed each CT slice as a single image. First, we applied the k-means algorithm with $k=2$ to the intensities of CT pixels within the PBM. Then, since the cortical bone appears lighter than the cancellous bone on CT images, we considered as cortical bone the cluster with the highest mean intensity and removed these pixels from the ROIs. This step allowed us to keep only the cancellous bone.

- *Feature Extraction from cancellous ROIs*

The remaining ROIs belonging to PBM were analyzed using radiomics. In particular, we considered an element made of 5-by-5 pixels. Subsequent elements were obtained moving by 1 pixel at time in both directions across the CT slice. For all elements overlapping bone marrow ROIs, a set of 36 features was calculated: 4 first-order statistical features (mean, standard deviation, skewness and kurtosis of the pixels intensities) and 32 second-order (texture) features. We calculated 22 texture features from the grey-level co-occurrence matrices (GLCM) [13], that counts the number of occurrences for which a pixel with a gray level i appears at specific distance from a pixel with grey level j, in a given direction. Starting from the GLCM, we calculated the following variables: autocorrelation [14], contrast [14], correlation1 [15], correlation2 [14], cluster prominence [14], cluster shade [14], dissimilarity [14], energy [14], entropy [14], homogeneity1 [15], homogeneity2 [14], maximum probability [14], sum of squares [13], variance [13], sum average [13], sum variance [13], sum entropy [13], difference variance [13], difference entropy [13], information measure of correlation1 [13], information measure of correlation2 [13], inverse difference normalized [15], inverse difference moment normalized [15]. A set of 5 texture features were obtained from the Gray Level Difference Method (GLDM) [16], taking into account the difference of the intensities of two pixels at specific distance and in a given direction. From this analysis we extracted: contrast, angular second moment, entropy, mean, and inverse difference moment [16]. Finally, 5 texture features were extracted from the Grey-Level Run Length Method (GLRLM) [16], counting the number of contiguous pixels with a specific grey level in a given direction. In this case, the extracted variables were the following: short run emphasis, long run emphasis, gray level distribution, run length distribution, and run percentage [16]. Since in this application we cannot identify

a preferential texture direction, the GLCM, GLDM and GLRLM were evaluated for the four main directions (0°, 45°, 90°, 135°) with a distance of one pixel. Then, the four matrices were averaged to extract the texture features.

- *Training Set and Classifier construction*

Taking into account the purpose of this study, that is to understand the feasibility of identifying *actBM* from CT imaging, we decided to implement a classifier for each patient. First, we constructed a training set for each subject considering five slices equispaced across the subject's CT sequence. For each slice we randomly extracted 1/5 of all valid elements (i.e. overlapping bone marrow ROIs). Each element was characterized by the 36 features and a label indicating the class it belongs to: active or red marrow (RM) if the element overlaps an *actBM* region obtained from ^{18}FDG-PET, inactive or yellow marrow (YM) otherwise. Then, the elements extracted from the five slices were pooled together to create a balanced training set for each patient. Before constructing the classifier, each training set was discretized by means of the *Chi2 algorithm* [17], in order to reduce the effect of noise and improve the classification performance [18]. The *Chi2 algorithm* is a supervised and bottom-up discretization method based on the *$\chi 2$ statistic.* For each variable to be discretized, firstly elements are sorted in ascending order. Then, the algorithm iteratively merges, in a discrete interval, those adjacent elements having a class independent from the variable values. This discretization proceeds until the $\chi 2$ value for all adjacent intervals is greater than the $\chi 2$ value calculated for a give significance level α and for a number of degrees of freedom equal to the total number of classes minus1. In order to avoid to set a significance level *a-priori,* the discretization is performed with decreased α values, until the consistency rate of the discrete dataset is above a desired value. In particular, we run the *Chi2 algorithm* for significance levels from 0.5 to 0.001 and we used the *dependency degree* [19], based on the Rough Set Theory, as consistency measure. The algorithm was stopped when the dependency degree of the discrete dataset for a specific α value was lower than the one of the original continuous dataset.

A Decision Tree [20] (DT) for each patient was obtained based on correspondent discrete training set. For the classifier construction we adopted the CART algorithm [20], while the identification of the best splitting rule for each node was based on the Gini Index [20].

- *Pixel Classification and Post-processing*

All elements extracted from the bone marrow ROIs of the entire CT sequence were firstly discretized using the cut-off points obtained from the *Chi2 algorithm*. Then, they were classified in one of the two classes (RM or YM) using the DT constructed for the specific subject. Finally, since each pixel was included in more than one element, the majority voting was used to classify the single pixel: the pixel class was the one that appears most often among the including elements. This procedure allows for reducing errors due to a single classification and to improve the overall accuracy [21].

Considering only the pixel classified as RM, we obtained a binary mask of the *actBM* for each CT slice. This mask was further processed in order to fill small holes and remove very small regions. For achieving this objective, we performed a morphological closing operation on the binary mask, using a disk-shaped structuring element with radius equal to 3.

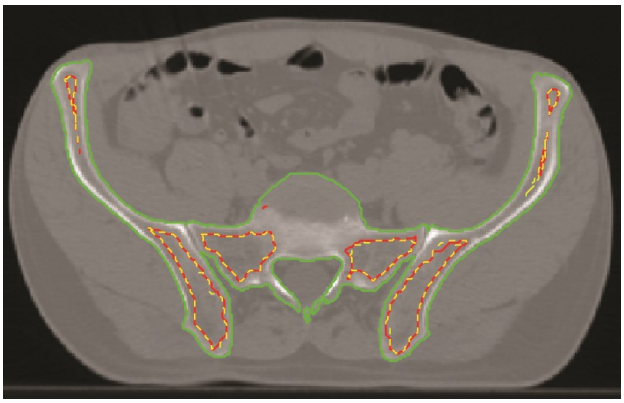

Fig. 1. Example of segmentation obtained using radiomics on CT. Green line: pelvic bon marrow segmentation. Red line: active bone marrow segmentation delineated from ^{18}FDG-PET without cortical bone (reference standard). Yellow dashed line: active bone marrow identified using radiomics on the CT image.

E. Validation

To validate our approach, we compared the *actBM* segmentation obtained from ^{18}FDG-PET (reference standard, *RS*) with those returned by radiomics applied to the CT sequences (*CT*). The comparison was carried out using three indices:

- *Dice index*, measuring the overall overlap between the two segmentations: $Dice = 2 \cdot (RS \cap CT)/(RS + CT)$
- *Precision*, measuring the over-segmentation: $Precision = (RS \cap CT)/CT$
- *Recall*, measuring the under-segmentation: $Recall = (RS \cap CT)/RS$

Dice, *Precision* and *Recall* were calculated for each slice of a CT sequence and for the three subregions separately.

III. Results

Figure 1 shows an example of segmentation obtained using radiomics on CT (yellow dashed line) with the RS delineated from ^{18}FDG-PET (red line).

For each of the 3 subregions we processed a number of slices ranging from 30 to 44 for every patient. The mean and standard deviation of the three indices are reported in Table I for the five patients and the three subregions separately. As it emerges from the table, the highest indices values are obtained for LSBM and IBM subregions, where we have an overlap between RS and CT above 75% in 4 out of 5 patients. The subregion with the lowest over-segmentation is the IBM, reaching precision above 80% for all patients. For the LPBM, not satisfying results are achieved, as the dice index is always below 0.5. However, the recall for this structure is higher than 0.7 in 4 out of 5 patients, meaning that at least the 70% of the RS segmentation is correctly recognized by the DT on CT images. These differences among

TABLE I. RESULTS OF THE COMPARISON BETWEEN ACTIVE BONE MARROW IDENTIFICATION FROM ^{18}FDG-PET AND CT

	IBM			**LPBM**			**LSBM**		
	Dice	***Precision***	***Recall***	***Dice***	***Precision***	***Recall***	***Dice***	***Precision***	***Recall***
Patient #1	0.88 ± 0.12	0.86 ± 0.17	0.93 ± 0.06	0.40 ± 0.19	0.27 ± 0.15	0.86 ± 0.19	0.87 ± 0.16	0.89 ± 0.12	0.87 ± 0.17
Patient #2	0.83 ± 0.13	0.89 ± 0.11	0.80 ± 0.17	0.36 ± 0.18	0.24 ± 0.15	0.79 ± 0.23	0.75 ± 0.31	0.75 ± 0.35	0.86 ± 0.19
Patient #3	0.57 ± 0.26	0.80 ± 0.22	0.48 ± 0.26	0.40 ± 0.20	0.37 ± 0.19	0.55 ± 0.27	0.40 ± 0.32	0.74 ± 0.29	0.37 ± 0.33
Patient #4	0.75 ± 0.18	0.90 ± 0.16	0.66 ± 0.21	0.43 ± 0.18	0.34 ± 0.19	0.72 ± 0.19	0.82 ± 0.22	0.85 ± 0.27	0.82 ± 0.13
Patient #5	0.93 ± 0.05	0.95 ± 0.07	0.91 ± 0.05	0.42 ± 0.10	0.28 ± 0.11	0.89 ± 0.11	0.78 ± 0.15	0.77 ± 0.21	0.82 ± 0.11

the three subregions might be due to the different amount of *actBM* within them, with the LSBM containing almost entirely hematopoietically active bone marrow. Moreover, the *actBM* identification in the LPBM might be influenced by the presence of the femura. Possible improvements in this direction could be obtained by constructing a classifier specific for each subregion, so that the characteristics of the three structures can be captured more accurately.

IV. CONCLUSIONS

This work proposes an exploratory study aiming to understand if radiomics is able to identify hematopoietically active BM from CT imaging. This is the first study in this direction, as the standard strategies for active BM detection are based on PET and MRI. Our results are very promising, above all for lumbosacral and iliac structures where our approach is able to correctly identify more than 75% of *actBM*. However, it is very likely that the obtained results were not bias from the use of a specific classifier for each subject, since the number of elements included in the training set of each DT was extremely lower than the amount of elements to be classified in all CT slices. A larger population of patients will be included in future studies, to better test the generalization capability of our approach. From the feature extraction viewpoint, the use of texture features, focusing on the spatial distribution of the pixel intensities on the image instead of the single intensity value, allows for constructing classifiers that are robust to modifications in the acquisition parameters or protocol. Future works will be aimed to develop a classifier specific for each subregion, that could be applied for all subjects and to explore the ability of other classifiers, also in the field of deep learning.

REFERENCES

[1] P. Franco et al., "Intensity-Modulated Radiation Therapy with Simultaneous Integrated Boost Combined with Concurrent Chemotherapy for the Treatment of Anal Cancer Patients: 4-Year Results of a Consecutive Case Series," Cancer Invest., vol. 33, no. 6, pp. 259–266, Jul. 2015.

[2] D. A. R. Julie et al., "Predictors of acute toxicities during definitive chemoradiation using intensity-modulated radiotherapy for anal squamous cell carcinoma.," Acta Oncol., vol. 55, no. 2, pp. 208–16, 2016.

[3] A. R. Filippi, P. Franco, and U. Ricardi, "Is clinical radiosensitivity a complex genetically controlled event?," Tumori, vol. 92, no. 2, pp. 87–91.

[4] P. Mauch et al., "Hematopoietic stem cell compartment: Acute and late effects of radiation therapy and chemotherapy," Int. J. Radiat. Oncol., vol. 31, no. 5, pp. 1319–1339, Mar. 1995.

[5] J. S. Blebea et al., "Structural and Functional Imaging of Normal Bone Marrow and Evaluation of Its Age-Related Changes," Semin. Nucl. Med., vol. 37, no. 3, pp. 185–194, May 2007.

[6] J. C. Roeske, A. Lujan, R. C. Reba, B. C. Penney, S. Diane Yamada, and A. J. Mundt, "Incorporation of SPECT bone marrow imaging into intensity modulated whole-pelvic radiation therapy treatment planning for gynecologic malignancies," Radiother. Oncol., vol. 77, no. 1, pp. 11–17, Oct. 2005.

[7] S. M. McGuire et al., "Spatial mapping of functional pelvic bone marrow using FLT PET," J. Appl. Clin. Med. Phys., vol. 15, no. 4, pp. 129–136, Jul. 2014.

[8] B. S. Rose et al., "Correlation Between Radiation Dose to 18F-FDG-PET Defined Active Bone Marrow Subregions and Acute Hematologic Toxicity in Cervical Cancer Patients Treated With Chemoradiotherapy," Int. J. Radiat. Oncol., vol. 83, no. 4, pp. 1185–1191, Jul. 2012.

[9] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images Are More than Pictures, They Are Data," Radiology, vol. 278, no. 2, pp. 563–577, Feb. 2016.

[10] S. Rosati, K. M. Meiburger, G. Balestra, U. R. Acharya, and F. Molinari, "Carotid wall measurement and assessment based on pixel-based and local texture descriptors," J. Mech. Med. Biol., 2016.

[11] V. Giannini, S. Rosati, D. Regge, and G. Balestra, "Specificity improvement of a CAD system for multiparametric MR prostate cancer using texture features and artificial neural networks," Health Technol. (Berl)., vol. 7, no. 1, pp. 71–80, Mar. 2017.

[12] L. K. Mell et al., "Dosimetric predictors of acute hematologic toxicity in cervical cancer patients treated with concurrent cisplatin and intensity-modulated pelvic radiotherapy," Int. J. Radiat. Oncol., vol. 66, no. 5, pp. 1356–1365, Dec. 2006.

[13] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," IEEE Trans. Syst. Man. Cybern., vol. 3, no. 6, pp. 610–621, Nov. 1973.

[14] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," IEEE Trans. Geosci. Remote Sens., vol. 37, no. 2, pp. 780–795, Mar. 1999.

[15] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," Can. J. Remote Sens., vol. 28, no. 1, pp. 45–62, Jun. 2014.

[16] R. W. Conners and C. a Harlow, "A theoretical comparison of texture algorithms.," IEEE Trans. Pattern Anal. Mach. Intell., vol. 2, no. 3, pp. 204–222, 1980.

[17] H. L. H. Liu and R. Setiono, "Chi2: feature selection and discretization of numeric attributes," Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence, 1995.

[18] S. Rosati, G. Balestra, V. Giannini, S. Mazzetti, F. Russo, and D. Regge, "ChiMerge discretization method: Impact on a computer aided diagnosis system for prostate cancer in MRI," in 2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings, 2015, pp. 297–302.

[19] R. Jensen and Q. Shen, Computational Intelligence and Feature Selection. Hoboken, NJ: Wiley-IEEE Press, 2008.

[20] J. Han, M. Kamber, and J. (Computer scientist) Pei, Data mining : concepts and techniques. Elsevier/Morgan Kaufmann, 2012.

[21] L. Lam and S. Y. Suen, "Application of majority voting to pattern recognition: an analysis of its behavior and performance," IEEE Trans. Syst. Man, Cybern. - Part A Syst. Humans, vol. 27, no. 5, pp. 553–568, 1997.