Active bone marrow segmentation based on computed tomography imaging in anal cancer patients: A machine-learning-based proof of concept

Active bone marrow segmentation based on computed tomography imaging in anal cancer patients: a machine-learning-based proof of concept

Fiandra C[a]., Rosati S[b]., Arcadipane F[c]., Dinapoli N[d]., Fato M[e]., Franco P[a,1]., Gallio E[f]., Scaffidi Gennarino D[b]., Silvetti P[a]., Zara S[g]., Ricardi U[a]., Balestra G[b]

a. Department of Oncology, University of Turin, Turin, Italy.
b. Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy
c. Department of Oncology, Radiation Oncology, A.O.U Citta' della Salute e della Scienza, Turin, Italy
d. UOC Radioterapia Oncologica, Dipartimento Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
e. Department of Informatics, Bioengineering, Robotics and System Engineering (DIBRIS), University of Genova, Genova, Italy
f. Medical Physics Unit, A.O.U. Città della Salute e della Scienza, Turin, Italy
g. Tecnologie Avanzate, Torino, Italy.

1. Current address: Department of Translational Medicine, University of Eastern Piedmont and Department of Radiation Oncology, University Hospital "Maggiore della Carità,", Novara, Italy.

Corresponding author: Christian Fiandra, christian.fiandra@unito.it , via Genova 3, Turin, Italy

Abstract

Purpose. Different methods are available to identify haematopoietically active bone marrow (ActBM). However, their use can be challenging for radiotherapy routine treatments, since they require specific equipment and dedicated time. A machine learning (ML) approach, based on radiomic features as inputs to three different classifiers, was applied to computed tomography (CT) images to identify haematopoietically active bone marrow in anal cancer patients.

Methods. A total of 40 patients was assigned to the construction set (training set + test set). Fluorine-18-Fluorodeoxyglucose Positron Emission Tomography (^{18}FDG-PET) images were used to detect the active part of the pelvic bone marrow (ActPBM) and stored as ground-truth for three subregions: iliac, lower pelvis and lumbosacral bone marrow (ActIBM, ActLPBM, ActLSBM). Three parameters were used for the correspondence analyses between ^{18}FDG-PET and ML classifiers: DICE index, Precision and Recall.

Results. For the 40-patient cohort, median values [min; max] of the Dice index were 0.69 [0.20; 0.84], 0.76 [0.25; 0.89], and 0.36 [0.15; 0.67] for ActIBM, ActLSBM, and ActLPBM, respectively. The Precision/Recall (P/R) ratio median value for the ActLPBM structure was 0.59 [0.20; 1.84] (over segmentation), while for the other two subregions the P/R ratio median has values of 1.249 [0.43; 4.15] for ActIBM and 1.093 [0.24; 1.91] for ActLSBM (under segmentation).

Conclusion. A satisfactory degree of overlap compared to ^{18}FDG-PET was found for 2 out of the 3 subregions within pelvic bones. Further optimization and generalization of the process is required before clinical implementation.

Introduction

The standard of care for patients with squamous cell carcinoma of the anal canal is concurrent chemo-radiation (RT-CHT), which is a combination of radiotherapy, usually delivered with intensity modulation, and concurrent 5-fluorouracil and Mitomycin C [1,2]. The acute toxicity associated with combined modality treatment is non-negligible and may cause unplanned treatment breaks and reduced treatment intensity, potentially jeopardizing oncological outcomes [3,4]. Hematologic toxicity (HT) increases the risk of asthenia, bleeding, or infections and may negatively affect patient's compliance to treatment [5,6]. Chemotherapy is one of the most important HT triggers as it induces myelosuppression, but radiotherapy, delivered to circulating blood cells and precursors within bone marrow (BM), also plays an important role. The observation that the radiation dose received by BM comprised within the pelvic bones is a predictive factor for HT occurrence prompted researchers to implement strategies to selectively spare pelvic BM through the accurate identification and delineation of these areas [7]. This may decrease the probability of developing HT, since in the average adult population, pelvic bones and lumbar vertebrae make up almost 60% of the total BM.

Although the use of the whole pelvic bones outlined as the outer contour of bony structures on computed tomography (CT) is the most inclusive method to define BM, other methods have been developed to identify haematopoietically active BM (ActBM). A widely accepted strategy for the contouring of active BM volumes has yet to be defined, but several methods have been proposed: single-photon emission computed tomography (SPECT), ^{18}F-fluorodeoxyglucose-labeled positron-emission tomography (^{18}FDG-PET) and 3'-deoxy-3'-^{18}F-fluorothymidine-labeled positron-emission

tomography (^{18}FLT-PET), having potential advantages and drawbacks [8-10]. A commonly used method for identifying ActBM is ^{18}FDG-PET, frequently prescribed to anal and cervical cancer patients during diagnosis and staging work up. Nevertheless, it is considered an optional exam by international guidelines in the diagnostic work-out of anal cancer patients and it is therefore not available for all patients in daily clinical practice [1].

Magnetic resonance (MR) has also been used for the semi-automatic or fully automatic segmentation of the active BM based on water-fat MR imaging in gynaecologic cancer patients or on patients who received spine MRI [11,12]. Although interesting, these methods could be difficult to be introduced in the routine of radiotherapy treatment as the process of registering images acquired in different times and positions could be both challenging and time consuming.

Conversely, all anal cancer patients planned to receive radiotherapy undergo a CT scan; hence the aim of this work is developing a fast and fully automated system based solely on CT images without the aid of extra imaging.

Because of the exploratory intent of the study and the limited dimension of the data set [13], we chose a supervised machine learning (ML) approach. Machine learning has been widely used to automate segmentation at multiple stages of cancer-related medical workflows. With many available contoured images, ML approaches can aid in segmentation by learning appropriate priors for structures and organs or image context and tissue appearance for voxel classification [14,15] and some experiences were also reported on the use of ML algorithms for auto-segmentation starting from CT images [16,17] and applicable on a broader scale than the segmentation strategy based on ^{18}FDG-PET.

Materials and methods

2.1 Patients' selection

A total of 40 patients affected with locally advanced squamous cell carcinoma of the anal canal was enrolled in the present study. All patients underwent radiotherapy treatment at our department.

2.2 Active bone marrow identification

During the staging work-up, ^{18}FDG-PET/CT on a Philips Gemini tomography (Philips Medical System, Eindhoven, NL) was performed. Data acquisition started 90 min after intravenous injection of

approximately 30 MBq/kg body weight of ^{18}F-glucose. After a full-body CT scan, PET scans were acquired for 2.5 min/bed position. ^{18}FDG-PET images pixel spacing was 4 x 4 x 4 mm^3. A dedicated fusion workstation (Extended Brilliance Workspace 2.0, Philips Medical System, Eindhoven, NL) was used for PET clinical interpretation. For radiotherapy planning use, all patients undergone a non-contrast-enhanced CT of the pelvic region, acquired in the supine position with both an indexed shaped knee rest and ankle support (CIVCO Medical Solutions, Kalona, IA, USA) on a Philips ''BigBore'' CT scanner (Philips Medical System, Eindhoven, NL). CT pixel spacing was 0.93 x 0.93 x 3 mm^3. Using VELOCITY platform (Varian Medical Systems, Palo Alto, CA), planning CT was co-registered with ^{18}FDG-PET images on a rigid registration and the co-registered PET images were then resampled to match the voxel dimensions of the planning CT to up-sample the lower digital resolution image to the higher resolution one [18]. Pelvic bone marrow (PBM) was delineated on planning CT as a whole and then divided in 3 subregions: iliac bone marrow (IBM), lower pelvis bone marrow (LPBM) and lumbosacral bone marrow (LSBM) [19]. The ^{18}FDG-PET standardized uptake values (SUVs) within the PBM volume of all patients were corrected for body weight and considered. ActBM was obtained by segmenting areas within PBM with SUV values higher than the mean SUV within the pelvic bones and then dividing them into 3 subregions: ActIBM, ActLPBM, ActLSBM [20]. The remaining PBM was defined as inactive BM (InActBM).

2.3 Extraction of radiomic features and training set construction

To divide the spongy bone, in which bone marrow is present, from compact bone, a k-means clustering algorithm with k=2 was used in the planning CT images; then, all the spongy bone was divided into 5-by-5 regions of interest (ROIs) centred in each pixel (one ROI for each pixel of the spongy bone) and labelled according to the portion of PBM they belonged to (ActBM or InActBM). Thirty-six radiomic features were extracted for each ROI with home-made code in Matlab software (version R2021a), importing all RT objects from the standard RT-Dicom format. Four first order statistical features (mean, standard deviation, skewness, and kurtosis) and 32 statistical features of the second order (22 regarding the Gray Level Co-occurrence Matrix (GLCM), 5 the Gray Level Dependence Matrix (GLDM) and the last 5 deriving from the Gray Level Run Length Matrix (GLRLM)) were calculated. All the features were implemented according to Image Biomarker Standardisation Initiative (IBSI) [21] guidelines.

The 36 radiomic features were extracted for the first ROI and assigned to its central pixel; then the ROI is moved of one pixel and the next ROI is characterized with the same set of features.

A dataset was built in the form of a matrix containing in the rows the ROIs, and in the columns the 36 features, the three coordinates of the central pixel, the label (ActBM or InActBM), and the patient identification code.

Three training sets were then constructed, one for each subregion, by randomly extracting, for each patient, 5 slices belonging to that anatomical region. From each slice, 25 pixels of both ActBM and InActBM were randomly selected to obtain a balanced training set containing the same number of ActBM and InActBM ROIs; a total of 5000 ROIs from ActBM and 5000 from InActBM were collected for each training set. With such a process, there was a large amount of available data for each patient for each subregion also for validation (about 95% of ROIs are not used for training), so it was decided to evaluate the performance working on the entire dataset, including both training and validation sets, considering performance evaluation sufficiently unbiased.

2.4 Genetic algorithm and classifiers optimization

Three different ML classifiers were compared in this study: the Decision Tree (DT), the K-Nearest Neighbours (KNN) and the feedforward artificial Neural Networks (NN).

For each classifier, three different Genetic Algorithms (GAs), one for each anatomical subregion, were used to simultaneously select the input features and define the classifier parameters [22,23]. Consequently, each solution of the GA was defined as a binary vector made of two subparts: the first subpart allows the selection of the input features (36 bits, one for each feature; a "1" in a given position identified a selected feature), while the second subpart selects the classifier parameters (the number of bits was set according to the number of parameters to be explored for each classifier).

The optimization process was based on a function, called *fitness*, that quantified the goodness of each solution and defined as:

$$fitness = 1 - acc + 0.3 * |sens - spec|$$

where *acc, sens* and *spec* are the values of accuracy, sensitivity, and specificity, respectively, obtained using a classifier with parameters defined by the second subpart of the solution and fed with the features selected by the first subpart. Half of the training set (i.e. ROIs from the first 20 patients) is used for the learning phase, while the other half of the training set was used to calculate the performance. Lower fitness values corresponded to better solutions. The fitness function was

designed so that the selected solutions had high accuracy, but also similar values of sensitivity and specificity. This prevents to obtain classifiers tending to under-segment or over-segment.

Each GA started from a population of 500 randomly generated solutions, iteratively evolved using a sequence of selection, crossover, and mutation. During selection, 350 solutions were extracted from the actual population to be the "parents" of new solutions. The probability of each solution to be chosen was inversely proportional to its fitness value. The selected solutions were randomly paired, and each couple was combined to produce two new solutions (crossover with probability (p_c) equal to 1). Finally, a mutation (a 0 bit was transformed into 1 and vice versa) occurred in the new solutions with a probability (p_m) equal to 0.5. Parents and children were pooled together and a new population of 500 solutions was randomly selected for the next iteration. This procedure was repeated for 100 iterations, with an additional stop condition to account for performance saturation, stopping the GA in case of no improvement of the best solution (i.e., the solution with best fitness) after 30 consecutive iterations.

Each classifier was optimized separately for each subregion and 9 best solutions were obtained (3 classifiers x 3 subregions). Figure 1 illustrates the flowchart how the method is designed.

Finally, the information encoded in each best solution, in terms of feature subset and classifier parameters, were used to construct a classifier able to assign each ROI to InActBM or actBM. Since the final aim of the system was to classify voxels and each voxel fell in more than one ROI, the majority voting was used to label each CT voxel as InActBM or actBM.

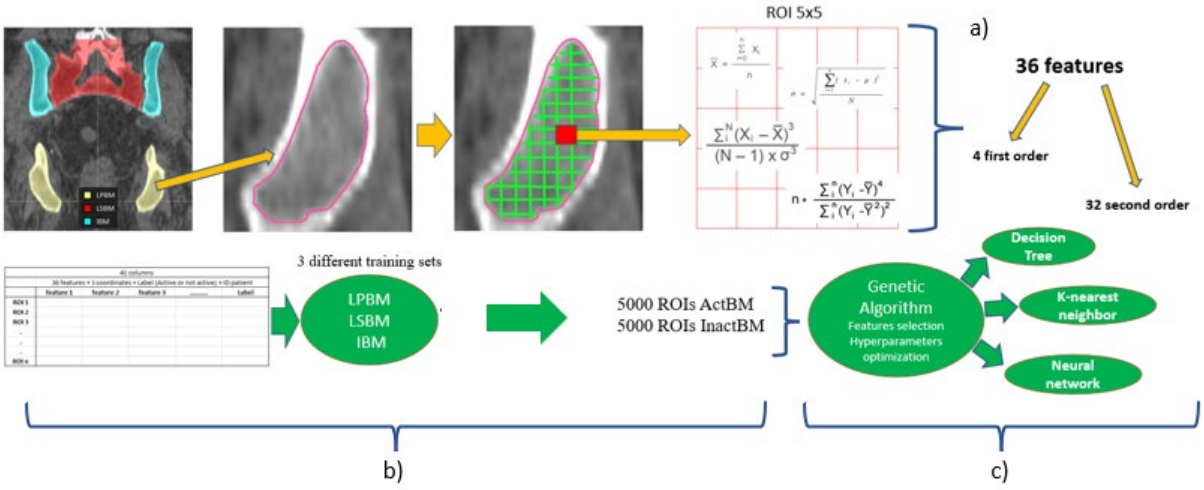

Figure 1. Visual representation of the steps in which the method has been divided: a) extraction of thirty-six radiomic features for each ROI; b) construction of the dataset where the rows are the ROIs, and the columns are the 36 features, the three coordinates of the central pixel, the label (ActBM or InActBM), and last column the patient identification code. One training set for each anatomical subregion was created by randomly extracting 5 slices for each patient and 25 pixels

from each slice, for both ActBM and InActBM. This allows to obtain a balanced training set containing the same number of ActBM and InActBM ROIs, for a total of 5000 ROIs from ActBM and from InActBM; c) for each classifier, three different Genetic Algorithms (GAs), one for each anatomical subregion, were used to simultaneously select the input features and to define the classifier parameters.

2.5 Overlapping parameters

Three parameters were used to evaluate the overlap between the ActBM regions identified by the ^{18}FDG-PET (PETActBM) and those detected on CT sequences (CTActBM): Dice index (DSC), recall (R) and precision (P). They were defined as follow:

$$DICE\ index = 2\ x\ \frac{\text{PETActBM} \cap \text{CTActBM}}{\text{PETActBM} \cup \text{CTActBM}} \quad (1)$$

$$Recall = \frac{\text{PETActBM} \cap \text{CTActBM}}{\text{PETActBM}} \quad (2)$$

$$Precision = \frac{\text{PETActBM} \cap \text{CTActBM}}{\text{CTActBM}} \quad (3)$$

These parameters range from 0, in case of no overlap between the two masks, to 1 in case of perfect match between them. Different studies have recommended a DSC of 0.7-0.8 to be considered a good overlap [24]. The ratio between P and R (P/R) may be used to highlight over-segmentation (P/R<1) or under-segmentation (P/R>1) tendency. Kruskal-Wallis non-parametric test was used for detecting differences among the three classifiers in terms of DSC and P/R ratio.

Results

Our cohort (40 patients) included 14 male (35%; average age 64±17 years) and 26 female (65%; average age 63±10 years) patients.

Figure 2 shows an example of automatic segmentation obtained with DT for ActIBM, ActLSBM and ActLPBM, respectively, for a patient with a good agreement with ^{18}FDG-PET contours and for a different one with lower concordance.

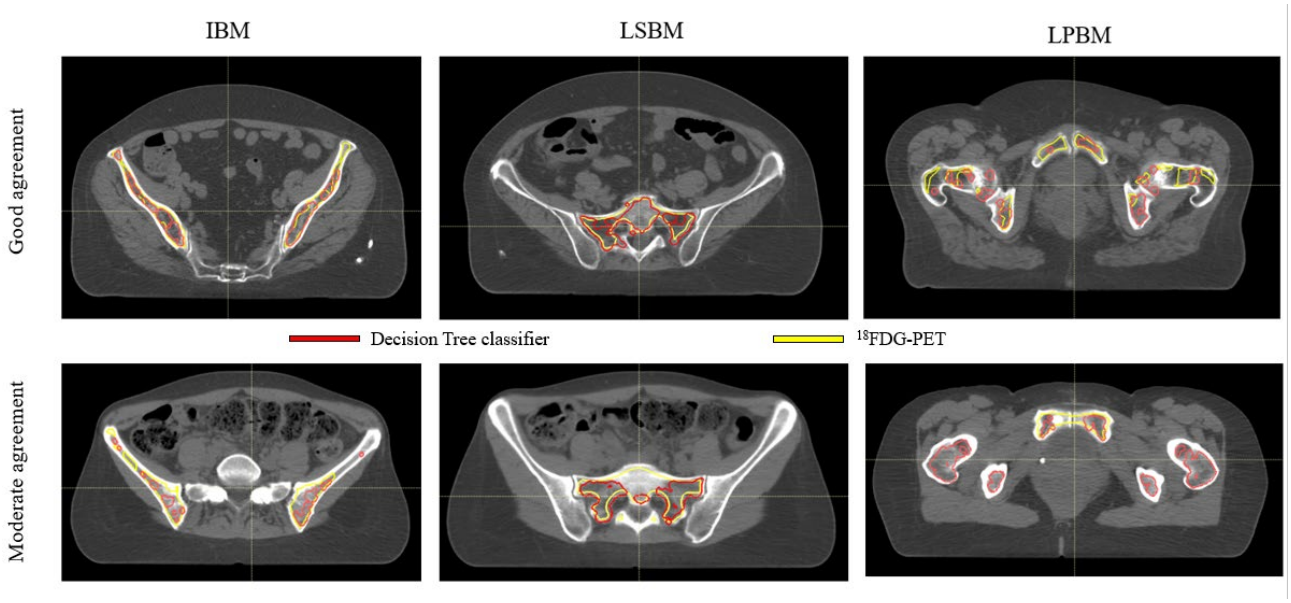

Figure 2. An example of a comparison between ^{18}FDG-PET active bone marrow ROI (yellow) and the segmentation of DT classifier (red) for iliac bone marrow (IBM), lumbosacral bone marrow (LSBM) and lower pelvis bone marrow (LPBM).

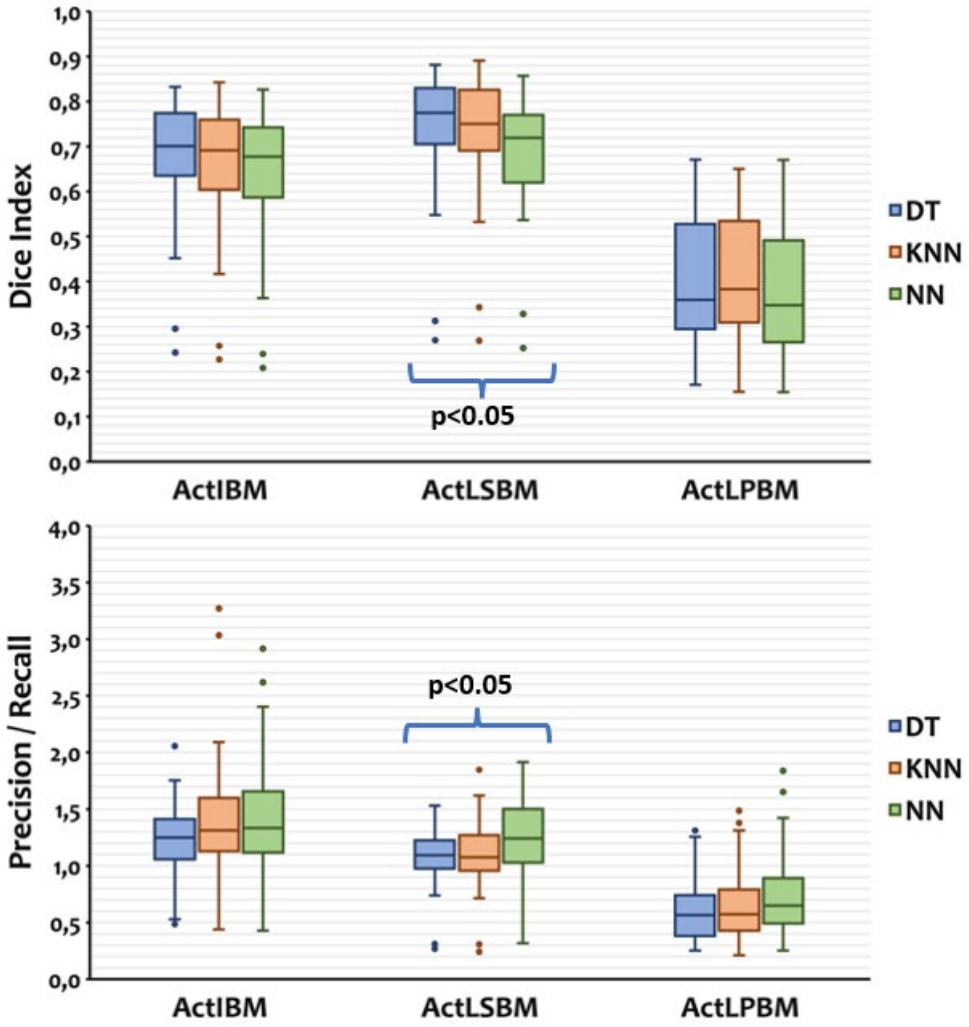

Figure 3. Box plot of the Dice Index and of the ratio Precision/Recall for the three considered classifiers for the 40 patients of the construction set for the tree subregions; a significant difference was found for the classifier NN compared with both DT and KNN for the ActLSBM region for DSC as well as P/R ratio.

Figure 3 shows the boxplots of DSC and P/R for the three classifiers for the three pelvis subregions. The median values of the DSC for ActLPBM for the three classifiers were 0.359 [0.17;067] (DT), 0.383 [0.15;0.65] (KNN) and 0.348 [0.15;0.67] (NN); they are lower than those obtained for the other two subregions (IBM: 0.7 [0.24;0.83], 0.69 [0.23; 0.84] and 0.68 [0.21;0.83]; LSBM: 0.77 [0.27;0.88], 0.75 [0.27;0.89] and 0.719 [0.25;0.86]). If we take as reference values those of the DT classifier, we can observe how the P/R median value for the LPBM structure was 0.56 [0.25;1.31] (over segmentation), while for the others two subregions the P/R median has values of 1.25 [0.49; 2.05] for IBM and 1.093 for LSBM [0.27;1.53] (under segmentation).
No significant difference was found among the three classifiers DT, KNN, and NN in terms of DSC values as well as in terms of P/R ratio values for the two regions ActIBM and ActLPBM; for the region ActLSBM, however, a significant difference was found for the NN compared with both DT and KNN (Kruskal-Wallis test p-value<0.05).

Discussion

The dose received by BM within the pelvic bones is a predictive factor for acute HT occurrence during RT-CHT for squamous cell carcinoma of the anus [7]. Therefore, an accurate identification and delineation of these regions is required to implement targeted approaches during both treatment planning and delivery, to effectively spare these structures and potentially reduce the risk of developing HT [7,8].
Our method attempts to identify and delineate ActBM relying exclusively on CT images implementing a segmentation tool which uses radiomic features as inputs of ML classifiers [25, 26]. The proposed experimental approach was then compared with an established segmentation method based on ^{18}FDG-PET images considered as ground truth. The use of CT imaging in studies employing AI-based approaches is increasing, as reported in Avanzo et al [27], where 12% of survey studies employed CT images (16% focused on segmentation).

Among the three considered classifiers, the only significant difference was reported for ActLSBM region with NN underperforming in terms of DSC respect to DT and KNN. Both ActLSBM and ActIBM

have an acceptable correspondence between the areas identified by PET and those recognized by the ML approaches. This is confirmed by the mean DSC values that are around 0.7 for these 2 subregions.

The component of under-segmentation (P/R>1) observed for the ML when compared to ^{18}FDG-PET, can be partially explained by the PET partial-volume effect, that cause a blurring effect, and the difference in voxel size between CT and PET scans. The average voxel size for our CT dataset was approximately 1 mm on axial images and 3 mm in the inf-sup direction, while the average voxel dimension for PET images was 4 mm. This difference may be responsible for the reduction in correspondence of the BM volumes, and this is particularly evident for the edges of the volumes, in which PET-based delineation would consistently overestimate the edges of BM, since it relied on a larger-sized voxel, leading to the under-segmentation emerging for the ML approaches. This voxel size difference also plays a role for the geometric uncertainties during the rigid co-registration process.

Conversely, if the percentage of ActBM was below 50% (and this is precisely the case with the LPBM structure), we can observe a tendency of ML approaches to over-segment (P/R<1) compared to the ground truth. As previously mentioned, the classification of the CT sequences was achieved using classifiers capable of identifying ActBM and InactBM voxels for each specific subregion; each classifier was constructed using a training set of 5000 voxels extracted from all the voxels belonging to that specific subregion, using the 40 patients enrolled in the study. If there were small active areas in the considered subregion it could happen that the corresponding actBM voxels were not used for constructing the training set, since a threshold of 25 pixel of both ActBM and InactBM was set to include the corresponding slice. This could reduce the ability of the classifier to correctly identify a specific area in the image and could be particularly cogent for patients that present low percentage of actBM. At the same time also the spatial resolution of PET may represent a limit when the area of actBM is smaller.

We have already discussed on how the ActBM can be identified by additional imaging and elaborated of the fact that this approach is hardly applicable on a large scale; however, Li et al. [28] have developed a method based on atlases precisely to overcome this problem and more recently, Yusufaly et al. [29] subsequently gave greater robustness to this method by applying it in a multicentric context. However, this approach assumes that ActBM distributions follow relatively predictable, canonical patterns within the whole pelvic bone marrow space. For these reasons, we

think it can be reasonably useful to develop an alternative method, although still investigational, to improve accuracy as observed in other districts [30,31].

Further studies on various pixel extraction strategies are recommended before a possible clinical implementation of such method, including the effects of different segmentation strategies on the final treatment planning optimization process. Also, to understand the sub-optimal correspondence of the LPBM district studies are currently underway with both the use of deep learning algorithms for CT segmentation [32] and the addition of a different imaging technique, such as MRI, as reference. Finally, for a further generalization of the process, it will be necessary to test the robustness of this approach with a sufficiently large test set completely independent from training.

Conclusion

We investigated the reliability of a ML-based approach to identify active BM within pelvic bones in anal cancer patients submitted to RT-CHT, by comparing the performance with an established method based on ^{18}FDG-PET. A better correspondence was found for 2 out of the 3 subregions evaluated. Both could potentially be identified with ML-based approaches with acceptable performance, while the discrimination ability of these models was suboptimal for ActLPBM.

Conflict of Interest Statement

All authors declare:

- All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.
- This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.
- Stefania Zara is employees of the company “Tecnologie Avanzate TA Srl” that distributes the software RayStation in Italy; this company supports the group in terms of importing contours generated by ML. Data were analyzed objectively and independently from the vendor.
- Remaining authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript

References

[1] Rao, S et al. Anal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow up. Ann. Oncol. 2021. doi: 10.1016/j.annonc.2021.06.015

[2] Arcadipane F et al. Image-guided IMRT with simultaneous integrated boost as per RTOG 0529 for the treatment of anal cancer. Asia Pac J Clin Oncol. 2018;14(3):217-23. doi: 10.1111/ajco.12768.

[3] Franco P et al. Volumetric modulated arc therapy (VMAT) in the combined modality treatment of anal cancer patients. Br J Radiol. 2016; 89(1060): 2015832. doi: 10.1259/bjr.20150832

[4] Ben-Josef E et al. Impact of overall treatment time on survival and local control in patients with anal cancer: a pooled data analysis of Radiation Therapy Oncology Groups Trials 87-04 and 98.11. J Clin Oncol. 2010; 28(34): 5061-6. doi: 10.1200/JCO.2010.29.1351.

[5] Julie DA et al. Predictors of acute toxicities during definitive chemoradiation using intensity-modulated radiotherapy for anal squamous cell carcinoma. Acta Oncol. 2016;55(2):208–16. doi: 10.3109/0284186X.2015.1043396.

[6] Franco P et al. Hematologic toxicity in anal cancer patients during combined chemo-radiation: a radiation oncologist perspective. Expert Rev Anticancer Ther. 2017;17(4):335-45.

[7] Franco P et al. Dosimetric predictors of acute hematologic toxicity during concurrent intensity-modulated radiotherapy and chemotherapy for anal cancer. Clin Transl Oncol. 2017; 19(1):67-75.

[8] Roeske JC et al. Incorporation of SPECT bone marrow imaging into intensity modulated whole-pelvic radiation therapy treatment planning for gynecologic malignancies. Radiother Oncol. 2005 Oct;77(1):11-7.

[9] Franco P et al. Incorporating 18FDG-PET-defined pelvic active bone marrow in the automatic treatment planning process of anal cancer patients undergoing chemo-radiation. BMC Cancer 2017; 17:710

[10] Wyss JC et al. [(18)F]Fluoro-2-deoxy-2-d-glucose versus 3'-deoxy-3'-[(18)F]fluorothymidine for defining hematopoietically active pelvic bone marrow in gynecologic patients. Radiother Oncol. 2016;118(1):72

[11] Andreychenko A et al. The feasibility of semi-automatically generated red bone marrow segmentation based on MR-only for patients with gynecologic cancer. Radiother Oncol. 2017;123(1):164-

[12] Hwang EJ, Kim S, Jung JY. Fully automated segmentation of lumbar bone marrow in sagittal, high-resolution T1-weighted magnetic resonance images using 2D U-NET. Comput Biol Med. 2021 Dec 1;140:105105.

[13] Pin Wang, En Fan, Peng Wang, Comparative analysis of image classification algorithms based on traditional machine learning and deep learning, Pattern Recognition Letters, Volume 141, 2021, Pages 61-67

[14] Harrison K et al. Machine Learning for Auto-Segmentation in Radiotherapy Planning. Clin Oncol (R Coll Radiol). 2022 Feb;34(2):74-88.

[15] Seo H et al. Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. Med Phys. 2020 Jun;47(5):e148-e167.

[16] Michal KOS. Semi-automatic CT image segmentation using random forests learned from partial annotations. Paper presented at: In Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologie 2018.

[17] Rosati S et al. Radiomics for identification of active bone marrow from ct: An exploratory study 2018 IEEE Life Sciences Conference, LSC 2018.

[18] Zhao C et al. Effects of Spatial Resolution on Image Registration. Proc SPIE Int Soc Opt Eng. 2016 Feb 27;9784.

[19] Rose BS et al. Correlation between radiation dose to ^{18}F-FDG-PET defined active bone marrow subregions and acute hematologic toxicity in cervical cancer patients treated with chemoradiotherapy. Int J Radiat Oncol Biol Phys. 2012 Jul 15;83(4):1185-91.

[20] Franco P et al. Dose to specific subregions of pelvic bone marrow defined with FDG-PET as a predictor of hematologic nadirs during concomitant chemoradiation in anal cancer patients. Med Oncol. 2016 Jul;33(7)

[21] Zwanenburg A et al. "Image biomarker standardisation initiative," arXiv preprint. 2019. arXiv:1612.07003v11

[22] S. Rosati, G. Balestra, M. Knaflitz. Comparison of Different Sets of Features for Human Activity Recognition by Wearable Sensors. Sensors, vol. 18, no. 12, p. 4189, Nov. 2018

[23] S. Rosati et al. Radiomics to Predict Response to Neoadjuvant Chemotherapy in Rectal Cancer: Influence of Simultaneous Feature Selection and Classifier Optimization," in 2018 IEEE Life Sciences Conference (LSC), 2018, pp. 65–68

[24] Mattiucci GC et al. Automatic delineation for replanning in nasopharynx radiotherapy; what is the agreement among experts to be considered as benchmark? Acta Oncologia 2013; 52: 1417–22

[25] Maffei N et al. Radiomics classifier to quantify automatic segmentation quality of cardiac sub-structures for radiotherapy treatment planning. Phys Med. 2021 Mar; 83:278-286;

[26] Ubaldi L et al; Strategies to develop radiomics and machine learning models for lung cancer stage and histology prediction using small data samples. Phys Med. 2021 Oct; 90:13-22

[27] Avanzo et al. Artificial intelligence applications in medical imaging: A review of the medical physics research in Italy. Phys Med. 2021 Mar; 83:221-241.

[28] Li N et al. Feasibility of atlas-based active bone marrow sparing intensity modulated radiation therapy for cervical cancer. Radiother Oncol. 2017 May;123(2):325-330.

[29] Yusufaly T et al. A Multi-atlas Approach for Active Bone Marrow Sparing Radiation Therapy: Implementation in the NRG-GY006 Trial. Int J Radiat Oncol Biol Phys. 2020;108(5):1240-1247.

[30] Vrtovec T et al. Autosegmentation of organs at risk for head and neck radiotherapy planning: from atlas-based to deep learning methods. Med Phys 2020;47:e929ee950.

[31] Fu Y et al. A review of deep learning based methods for medical image multi-organ segmentation. Phys Med 2021;85:107e122
.
[32] Chen J et al. Generative models improve radiomics performance in different tasks and different datasets: An experimental study. Phys Med. 2022 Jun;98:11-17.