Learning Robust Graph-Convolutional Representations for Point Cloud Denoising

Comparison of different classifiers to recognize active bone marrow from CT images

Samanta Rosati
Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
samanta.rosati@polito.it

Pierfrancesco Franco
Department of Oncology, Radiation Oncology
Università di Torino
Torino, Italy
pierfrancesco.franco@unito.it

Christian Fiandra
Department of Oncology, Radiation Oncology
Università di Torino
Torino, Italy
christian.fiandra@unito.it

Francesca Arcadipane
Department of Oncology, Radiation Oncology
Università di Torino
Torino, Italy
francesca.arcadipane@gmail.com

Patrick Silvetti
Department of Oncology, Radiation Oncology
Università di Torino
Torino, Italy
patrick.silvetti@edu.unito.it

Elena Gallio
Medical Physics Unit
A.O.U. Città della Salute e della Scienza di Torino
Torino, Italy
egallio@cittadellasalute.to.it

Jovana Panic
Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
jovana.panic.jop@gmail.com

Umberto Ricardi
Department of Oncology, Radiation Oncology
Università di Torino
Torino, Italy
umberto.ricardi@unito.it

Gabriella Balestra
Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
gabriella.balestra@polito.it

***Abstract*—One of the main problems during in the treatment of anal cancer with chemotherapy and radiation is the occurrence of Hematologic Toxicity (HT). In particular, during radiotherapy it is crucial to spare Bone Marrow (BM), since the radiation dose received by BM in pelvic bones predicts the onset of HT. In this direction, the most popular strategies are based on the identification of the hematopoietically active BM (*actBM*), that is the part of BM in charge of blood cells generation, using MRI, SPECT or PET, but no approached have been proposed based on CT. In this study we compare four different classifiers in recognizing *actBM* from CT images using 36 radiomic features. We used Genetic Algorithms (GAs) to simultaneously optimize the feature subsets and the classifier parameters, separately for three pelvic subregions: iliac bone marrow (IBM), lower pelvis bone marrow (LPBM), and lumbosacral bone marrow (LSBM). The obtained classifiers were applied to CT sequences of a cohort of 25 patients affected by carcinoma of the anal canal. Classifiers results were compared with the *actBM* identified from ^{18}FDG-PET (reference standard, RS). It emerged that the performances of the 4 classifiers are similar and they are satisfactory for IBM and LSBM subregions (Dice > 0.7) whereas they are poor for LPBM (Dice < 0.5).**

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

I. Introduction

The usual treatment for anal cancer patients is concurrent chemo-radiation [1]. Even with the use of high-tech delivery approaches such as Intensity-Modulated Radiation Therapy (IMRT), toxicity remains a problem. Compliance to therapy may be affected by acute Hematologic Toxicity (HT) that can produce asthenia, bleeding, or infections [2]. Both chemotherapy and radiation play an important role for HT [3]. In adults, almost 60% of total Bone Marrow (BM) is contained in pelvic bones and lumbar spine [4] and the dose received by BM in these structures predicts the occurrence of HT. For this reason, it is important to develop methods for correctly identify BM in pelvic bones.

One of the most popular strategies is based on the recognition of the hematopoietically active BM (*actBM*), that is the part of BM in charge of blood cells generation, using 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].

International guidelines in the diagnostic work-out of anal cancer consider ^{18}FDG-PET an optional exam. On the contrary, Computed tomography (CT) is mandatory for all patients before starting radiotherapy and the identification of *actBM* from this kind of images would have a potential broader applicability. However, to the best of our knowledge, no strategies in this direction have been developed.

Radiomics is defined as "process designed to extract a large number of quantitative features from digital images" [9]. The extracted information can be used for developing Computer-Aided Diagnosis (CAD) and decision support systems. Radiomics essentially treats each image as a matrix of data and, thus, it can be used to process every kind of images, such as CT, MRI, etc.… For example, in the study by Rosati et al. [10] radiomics was applied to carotid ultrasound images to segment carotid vessel wall layers and recognize subjects with higher cardiovascular risk. In the study by Giannini et al. [11], radiomics was applied to MRI for improving the performances of a CAD system for prostate cancer detection.

Feature extraction is the fundamental step of radiomics, allowing to obtain a set of variables that describes the structure of interest. Variables can be essentially grouped into three categories: first-order statistical features, measuring the statistical distribution of the intensities in a given area, second-order or texture features, capturing the spatial distribution of

the intensities, and higher-order features, able to identify specific patterns.

In a previous work [12] we conducted an exploratory study in order to understand if radiomics could allow the identification of *actBM* from CT. Our results showed that *actBM* identification in lumbosacral and iliac structures using radiomics is feasible. The aim of this study is to compare four different classifiers in recognizing the *actBM* from CT images using radiomics. In this case, the feature subsets and the classifier parameters were simultaneously optimized using a Genetic Algorithm (GA) and larger population of 25 subjects was used for results evaluation.

II. Materials and Methods

A. Population and Image Acquisition Protocol

In this study we involved a total of 25 patients affected with locally advanced squamous cell carcinoma of the anal canal. All the patients were treated with RT-CHT delivered with volumetric modulated arc therapy [13], [14]. A written informed consent was signed by all participants included in the study and the ethical principle of the Helsinki Declaration was adopted in this study.

During the staging work-up, ^{18}FDG-PET-CT on a Philips Gemini PET/CT tomography 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. A dedicated fusion workstation (Extended Brilliance Workspace 2.0) was used for PET clinical interpretation.

A non-contrast-enhanced CT of the pelvic region was acquired in the supine position with both an indexed shaped knee rest and ankle support (CIVCO Medical Solutions, Kalona, IA, USA), which was used for radiotherapy planning on a Philips ‘‘BigBore’’ CT scanner (Philips Medical System, Eindhoven, NL). Voxel spacing was (0.93 mm, 0.93 mm, 3 mm) for CT.

A detailed description of acquisition protocol can be found in [12].

B. Reference Standard Construction

Pelvic bone marrow (PBM) was delineated on planning CT and manually divided in 3 subregions: iliac bone marrow (IBM) including the area between the iliac crests and the upper border of femoral head; lower pelvis bone marrow (LPBM) made up of bilateral pube, ischia, acetabula and proximal femura, from the upper border of the femoral heads to the lower aspect of the ischial tuberosities; lumbosacral bone marrow (LSBM), comprising the area between the superior border of L5 somatic body a nd the lower aspect of the coccyx.

We used the VELOCITY platform (Varian Medical Systems, Palo Alto, CA), to co-register planning CT with ^{18}FDG-PET images.

In order to identify *ActBM,* we calculated the ^{18}FDG-PET standardized uptake values (SUVs) within the PBM volume of all patients, corrected for body weight and considered [8]. Then, we labeled as *ActBM* all areas within PBM with SUV values higher than the mean SUV within the pelvic bones. Finally, we divided them into the 3 subregions: *ActIBM*, *ActLPBM*, *ActLSBM*. The remaining PBM was labeled as inactive BM (*inactBM*) and separated into the three subregions, analogously. These areas we used as reference standard (RS) for classification and validation.

C. Active bone marrow identification from CT

In order to recognize *ActBM* from CT, we first removed the cortical bone from the whole PBM volume. Then we applied radiomics to extract a set of features to be used for classification. Four different classifiers belonging to machine learning were implemented, each optimized in terms of parameters and input feature subset using a GA. Finally, the classifiers were applied for recognizing the areas containing *actBM* on CT sequences and compared among them in terms of agreement with the RS.

- *Removal of cortical bone*

Since cortical bone less likely contains BM, we removed it from the whole PBM volume using a clustering approach. In particular, we applied the k-means algorithm (*k=2)* to the PBM voxels for each CT slice. This procedure divided the voxels into two groups based on their intensity and we assign to cortical bone the group with the highest mean intensity, since it appears lighter than cancellous bone on CT sequences. All these voxels were not further processed in the following steps.

- *Extraction of Radiomic Features*

Radiomics was applied to describe the cancellous bone areas on CT. In particular, regions of interest (ROIs) made of 5-by-5 voxels were moved by 1 voxel at time in both directions across each CT slice. ROIs completely overlapping BM areas were characterized by 36 features: 4 statistical features (mean, standard deviation, skewness and kurtosis of the voxels intensities) and 32 texture (second-order) features.

In particular, we calculated 22 texture features from the grey-level co-occurrence matrices (GLCM), 5 texture features from the Gray Level Difference Method (GLDM) [15], and 5 texture features from the Grey-Level Run Length Method (GLRLM) [15]. Since no preferential texture direction can be identified in the areas of interest, we evaluated the GLCM, GLDM and GLRLM in the four main directions (0°, 45°, 90°, 135°) and we averaged before extracting the texture features. The complete list of 36 features can be found in [12].

- *Training Set Construction*

Since our aim is to construct a classifier for each pelvic subregion, three different training sets were created containing IBM, LPBM and LSBM voxels respectively. First, we randomly extracted 5 patients from our population. Then, for each subregion, 5 slices were randomly extracted from the CT sequence and, for each slice, 1/5 of the ROIs totally included in actBM areas were randomly inserted in the training set. The same number of ROIs completely overlapping *inactBM* areas were added, in order to obtain balance training sets. The final training sets were made up of 2400 ROIs for IBM, 2200 ROIs for LPBM, and 2500 ROIs for LSBM. Every ROI was characterized by the set of 36 radiomic features and labelled with the portion of PBM to which it belonged: active or red marrow (*RM*) if it overlaps an *actBM* area, inactive or yellow marrow (*YM*) otherwise.

- *FS and Classifier Optimization using GA*

In this study, we construct and compare the performances of the following 4 classifiers belonging to machine learning: K-Nearest Neighbors (KNN), Feedforward Neural Network

(FNN), Support Vector Machine (SVM) and Decision Tree (DT).

For each classifier, three Genetic Algorithms (GAs), one for each subregion, were used to simultaneously select the input features and define the classifier parameters [16], [17]. For this purpose, the GA solutions were codified as binary vectors composed of 2 parts: the first part, made of 36 bits (1 bit for each feature), was used for selecting the most relevant features to be used for classification, and the second part was used to set the classifier parameters. Since each classifier requires a different set of parameters, the codification of the second part was adapted to each situation as follows.

1) KNN: only the number of the *K* neighbors must be set for the KNN classifier. Usually, an approximate initial value for *K* can be obtained with the formula $K_{in} = \sqrt{N}$, where *N* was the number of elements in the training set [18]. Since in this study we decided to explore 64 values around this starting value, the second part of the GA solutions was made of 6 bits (2^6=64) and each value was associated to a specific K value for the classifier.

2) FNN: for this application a basic structure was defined for the network: an input layer and a hidden layer with a number of neurons equal to the number of features selected in the first part of the GA solution; an output layer containing one neuron and returning the voxel class between *actBM* and *inactBM*. The second part of GA solutions, made of 3 bits, was used to add from 1 to 8 further hidden layers to this basic structure. Each new layer included half of the neurons if the previous layer. We set the sigmoid transfer function for all hidden layers and the linear transfer function for the output layer.

3) SVM: this classifier requires to set both the penalty term *C* that allows to balance margins and misclassification error, and the kernel function. These two paramenters were codified in the second part of GA solutions: two bits were used for selecting the kernel function among linear, Gaussian, polynomial of order 2 and polynomial of order 3, and four bits were adopted for defining the *C* value according to the following equation:

$$C = \begin{cases} 0.5 & if\ C_{dec} = 0 \\ 1 & if\ C_{dec} = 1 \\ (C_{dec} - 1) * 10 & otherwise \end{cases} \qquad (1)$$

where C_{dec} is the decimal value of the 4 bits codifying the *C* term. Using eq. (1) we were able to explore values between 0.5 and 140.

4) DT: the CART algorithm [19] and the Gini Index [19] were used for the tree construction and the identification of the best splitting rule for each node, respectively. Once these methods have been selected, no other parameters must be set for DT construction and running. For this reason, in our GA the optimization of the DT did not require bits associated to the second part of the solution.

For each GA solution, the corresponding classifier was constructed: it was fed with the feature subset defined by the first part of the solution whereas the second part defines the its parameters. We evaluated the goodness of GA solutions based on accuracy (*acc*), sensitivity (*sens*) and specificity (*spec*) returned by the classifier trained and tested using the training set extracted from a specific subregion. These three parameters were used to evaluate the goodness of GA solution, according to the following fitness function:

$$fitness = 1 - acc + 0.3 * |sens - spec| \qquad (2)$$

Lower fitness values correspond to better solutions.

In our GAs, we started from an initial population of 500 solutions randomly generated. From this initial population, 350 solutions were selected by means of the roulette wheel algorithm [20] and evolved for 100 iterations. We set the mutation probability to 0.5 and the crossover probability to 1.

Each GA was run five times starting from the same initial population of solutions, in order to manage the random nature of the algorithm. Since each run returned the solution with the lowest fitness value (called *best solution*), we obtained five best solutions for each subregion and each classifier. Among them, we identified those with the lowest fitness value, the highest sensitivity and the highest specificity. We constructed the corresponding three classifiers and we used them for voxel classification.

- *Voxel Classification and Post-processing*

For each subregion, all 5-by-5 ROIs were assigned to RM or YM class using the classifiers obtained by GA. Since each voxel belonged to more than one ROI, we applied the majority voting to obtain the final class, reducing errors due to a single classification and improving the overall accuracy [21].

We constructed a binary mask of the *actBM* for each CT slice, taking into account only the RM voxels. These masks were post-processed by means of a morphological closing operation with a disk-shaped structuring element with radius equal to 3. This procedure allowed to fill small holes and remove very small regions.

The described procedure was repeated for the three best GA solutions of each classifier. Finally, the three binary masks were combined by means of the majority voting procedure.

D. Validation

For validation, the masks obtained from classification (*CL*) were compared with the *RS* using the *Dice index*, evaluating the overlap between the two segmentations:

$$Dice = 2 \cdot \frac{RS \cap CL}{RS + CL} \qquad (3)$$

The parameter was measured slice by slice for the three subregions separately.

E. Training set analysis

The difference among the patients were analyzed by means of the characteristics of the training set.

Hierarchical clustering was applied separately to divide the ROIs of each patients in homogeneous groups. Five dendrograms for each of the three subregions were

TABLE I. NUMBER OF VARIABLES INCLUDED IN AT LEAST ONE OF THE 3 BEST SOLUTIONS USED FOR CONSTRUCTING THE 4 CLASSIFIERS

	DT	kNN	FNN	SVM
IBM	27	22	32	31
LPBM	29	24	34	29
LSBM	32	26	32	33

TABLE II. DICE VALUES (AVERAGE AND STANDARD DEVIATION) OBTAINED WITH THE 4 CLASSIFIERS FOR THE THREE SUBREGIONS. VALUES IN (A) REFERS TO THE 5 PATIENTS INCLUDED IN THE TRAINING SET WHILE THOSE IN (B) TO THE 20 PATIENTS OF THE VALIDATION SET.

(a)	TRAINING SET PATIENTS			
	DT	**kNN**	**FNN**	**SVM**
IBM	0,80 ± 0,10	0,80 ± 0,04	0,77 ± 0,10	0,80 ± 0,08
LPBM	0,49 ± 0,08	0,49 ± 0,09	0,49 ± 0,09	0,49 ± 0,07
LSBM	0,88 ± 0,04	0,82 ± 0,09	0,78 ± 0,09	0,84 ± 0,09

(b)	VALIDATION SET PATIENTS			
	DT	**kNN**	**FNN**	**SVM**
IBM	0,75 ± 0,09	0,71 ± 0,14	0,71 ± 0,13	0,74 ± 0,10
LPBM	0,40 ± 0,13	0,38 ± 0,12	0,39 ± 0,12	0,41 ± 0,14
LSBM	0,81 ± 0,08	0,78 ± 0,10	0,72 ± 0,11	0,78 ± 0,09

constructed. The aim of the hierarchical clustering was to look for differences in the training set among the patients.

The best cut every dendrograms in two clusters. To characterize the cluster the two centroids were calculated. For each subregion, the centroids of all the clusters were compared by means of the Euclidean distance.

III. RESULTS AND DISCUSSION

Table I reports the number of variables that are included in at least one of the three best solutions used for constructing the four classifiers. It should be noted that the number of selected variables is high. This fact highlights the difficulties in discriminating the red from the yellow marrow.

Table II presents the dice values (average and standard deviation) obtained with the different classifiers for the three subregions. Table II(a) refers to the 5 patients included in the training set while Table II(b) to the 20 patients of the validation set.

Comparing the two tables we can make several observations. First, there are 2 subregions (IBM and LSBM) with acceptable results while the results for LPBM are not acceptable. The performances decrease on the validation set patients, as it is expected. Looking at the training set patients, the performances of the 4 classifiers may be considered similar. Also, for the validation set patients the differences are very small but in the 2 subregions with acceptable results. Comparing the results of the training set patients with those of the validation set, it seems that the performances obtained with the DT classifiers have a smaller decreasing than those achieved the other classifiers.

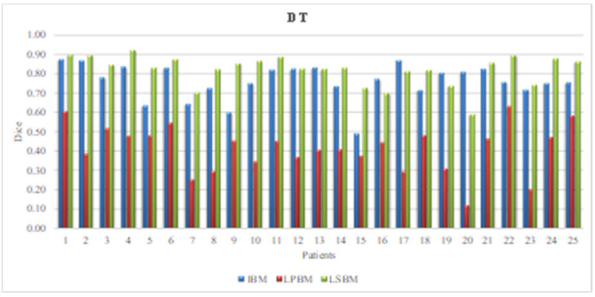

Fig. 1. Dice value obtained patient by patient with the DT classifier for the three subregions. Patients from #1 to #5 were used for constructing the training set.

TABLE III. MEAN DIFFERENCES BETWEEN THE CENTROIDS OF EACH CLUSTER OF A PATIENT AND THE CLUSTERS OF THE OTHER PATIENTS

	ActBM				
	Pat. 1	**Pat. 2**	**Pat. 3**	**Pat. 4**	**Pat. 5**
IBM	0,49 ± 0,18	0,47 ± 0,16	0,46 ± 0,17	0,43 ± 0,15	0,40 ± 0,13
LPBM	0,52 ± 0,16	0,40 ± 0,15	0,38 ± 0,14	0,46 ± 0,17	0,47 ± 0,16
LSBM	0,61 ± 0,29	0,57 ± 0,28	0,52 ± 0,26	0,57 ± 0,27	0,64 ± 0,27

	InactBM				
	Pat. 1	**Pat. 2**	**Pat. 3**	**Pat. 4**	**Pat. 5**
IBM	0,58 ± 0,18	0,48 ± 0,21	0,49 ± 0,20	0,47 ± 0,22	0,63 ± 0,22
LPBM	0,52 ± 0,14	0,45 ± 0,19	0,47 ± 0,19	0,43 ± 0,16	0,47 ± 0,22
LSBM	0,52 ± 0,22	0,46 ± 0,19	0,42 ± 0,22	0,43 ± 0,22	0,52 ± 0,24

Standard deviations are not so small, letting us think that there is variability among the patients. This hypothesis is confirmed by the results showed in Fig. 1 in which the performances of the DT classifier for the three subregions are displayed.

The results on the single patients show that there are a few patients that have results extremely below average. Among the critical patients there is patient #5 for IBM. The fact is particularly strange because this patient is part of the training set.

Looking for an explanation of this anomaly, we decided to compare the ROIs of the training set by means of hierarchical clustering. Table III presents the mean differences between the centroids of each cluster of a patient and the clusters of the other patients. The values are quite similar demonstrating that the variability among the performances are probably due to a training set that does not represent all the possible types of ROIs.

With respect to our previous work [12], in this study we included 20 more subjects in our population, that were used for validation. Moreover, since it was an exploratory study in this direction, in [12] we proposed only one kind of classifier (DT) that was constructed subject by subject and considering all subregions together. Here we compared four kinds of classifier and we proposed one classifier for each subregion, optimized using GA in terms of features subset and parameters.

IV. CONCLUSIONS

In this work we used radiomics to recognize active bone marrow in pelvic bone from CT images. We compared 4 different classifiers optimized using GA and trained with the same training set.

The results show that the performances of the 4 classifiers are similar. On average they are satisfactory for IBM and LSBM subregions, but also in these cases there are patients with poor results.

Looking at the training set we were not able to explain the differences. More work is needed to investigate the differences among the patients and to obtain satisfactory results also for LPBM subregion.

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

[13] F. Arcadipane et al., "Image-guided IMRT with simultaneous integrated boost as per RTOG 0529 for the treatment of anal cancer," Asia. Pac. J. Clin. Oncol., vol. 14, no. 3, pp. 217–223, Jun. 2018.

[14] P. Franco et al., "Volumetric modulated arc therapy (VMAT) in the combined modality treatment of anal cancer patients," Br. J. Radiol., vol. 89, no. 1060, p. 20150832, Apr. 2016.

[15] 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.

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

[17] S. Rosati, C. M. Gianfreda, G. Balestra, V. Giannini, S. Mazzetti, and D. Regge, "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.

[18] A. C. Rencher, Methods of multivariate analysis. J. Wiley, 2002.

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

[20] A. P. Engelbrecht, Computational Intelligence: An Introduction. Wiley Publishing, 2007.

[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.