Journal of Medical and Biological Engineering (2024) 44:711–721
https://doi.org/10.1007/s40846-024-00906-3

ORIGINAL ARTICLE

Prediction of the Gleason Score of Prostate Cancer Patients Using ^{68}Ga-PSMA-PET/CT Radiomic Models

Zahra Vosoughi[1] · Farshad Emami[2] · Habibeh Vosoughi[3] · Ghasem Hajianfar[4] · Nima Hamzian[1] · Parham Geramifar[3] · Habib Zaidi[4,5,6,7]

Received: 3 July 2024 / Accepted: 19 September 2024 / Published online: 12 October 2024

Abstract

Purpose To predict Gleason Score (GS) using radiomic features from ^{68}Ga-PSMA-PET/CT images in primary prostate cancer.

Methods 138 patients undergoing ^{68}Ga-PSMA-PET/CT imaging were categorized based on GS, with GS above 4 + 3 as malignant and under 3 + 4 as benign tumors. radiomic features were extracted from tumors' volume of interest in both PET and CT images, using Feature Elimination with cross-validation. Fusion features were generated by combining features at the feature level; average of features (PET/CT$_{AveFea}$) or concatenated features (PET/CT$_{ConFea}$). The performance of various models was compared using area under the curve, sensitivity and specificity. Wilcoxon test and F1-score test were used to find the best model. Predictive models were developed for CT-only, PET-only, and PET/CT feature-level fusion models.

Results Random Forest achieved the highest accuracy on CT with 0.74 ± 0.01 AUC$_{Mean}$, 0.75 ± 0.07 sensitivity, and 0.62 ± 0.08 specificity. Logistic regression (LR) exhibited the best predictive performance on PET images with 0.74 ± 0.05 AUC$_{Mean}$, 0.7 ± 0.13 sensitivity, and 0.78 ± 0.14 specificity. The best predictive PET/CT$_{AveFea}$ was achieved by LR, resulting in 0.72 ± 0.07 AUC$_{Mean}$, 0.74 ± 0.12 sensitivity, and 0.63 ± 0.02 specificity. In the case of PET/CT$_{ConFea}$, LR showed the best predictive performance with 0.78 ± 0.08 AUC$_{Mean}$, 0.81 ± 0.09 sensitivity, and 0.66 ± 0.15 specificity.

Conclusion The results demonstrated that radiomic models derived from ^{68}Ga-PSMA-PET/CT images could differentiate between benign and malignant tumors based on GS.

Keywords Prostate Cancer · ^{68}Ga-PSMA PET/CT · Radiomics · Gleason Score

1 Introduction

Prostate cancer (PCa) is the most prevalent form of cancer in men and is rapidly increasing worldwide [1]. Patients with prostate cancer are categorized into three risk groups, mainly low, intermediate-, and high-risk based on their Gleason score. The Gleason score is a grading system that assesses the aggressiveness of prostate cancer cells and is determined through a biopsy of the prostate gland [2]. The Gleason score ranges from 6 to 10, with higher scores indicating more aggressive cancer cells and placement in the high-risk group. To calculate the Gleason score, the two

✉ Nima Hamzian
hamzian.nima@gmail.com

✉ Parham Geramifar
pgeramifar@tums.ac.ir

1 Department of Medical Physics, School of Medicine, Shahid Sadoughi University of Medical Sciences, Shohada Gomnam Blv, Yazd, Iran

2 Nuclear Medicine and Molecular Imaging Department, Imam Reza International University, Razavi Hospital, Mashhad, Iran

3 Research Center for Nuclear Medicine, Tehran University of Medical Science, Tehran, Iran

4 Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva CH-1211, Switzerland

5 Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, Netherlands

6 Department of Nuclear Medicine, University of Southern Denmark, Odense DK-500, Denmark

7 University Research and Innovation Center, Óbuda University, Budapest, Hungary

most common patterns observed in the biopsy sample were added together. However, there is a major flaw in the Gleason scoring system as Gleason scores of 3 + 4 and 4 + 3 are often considered within the same prognostic group as Gleason score 7. Gleason score 7 is divided into two separate groups, Gleason score 3 + 4 (grade 2) and Gleason score 4 + 3 (grade 3) [3]. Gleason score 3 + 4 tumors still have a favorable prognosis. Yet, it is not as good as Gleason score 6 tumors (low-grade). On the other hand, a Gleason score 4 + 3 tumor is more likely to grow and spread compared to a 3 + 4 tumor but not as likely as a Gleason score 8 tumors (intermediate grade). A Gleason score of 8 to 10 indicates high-grade or poorly differentiated cancer.

Different modern medical imaging techniques provide valuable clinical information for various purposes, including clinical diagnosis, prognosis, treatment planning and response evaluation [4]. Gallium-68 prostate specific membrane antigen (^{68}Ga-PSMA) PET/CT as a specific imaging modality, has the potential to enhance lesion detectability and provide accurate staging in patients with PCa [5]. By the end of the twentieth century, the vast amount of data and advanced algorithms have opened a new era in medical imaging. Radiomics is an evolving field of research aiming at extracting and analyzing advanced quantitative features from medical images [6]. Indeed, radiomics is a reliable, non-invasive, and cost-effective approach that may potentially predict diagnosis, monitoring treatment response, prognosis and personalized therapy based on the correlation between the extracted features and specific disease characteristics or outcomes [7, 8]. Recently, much research and development efforts focused on demonstrating the capability of combining radiomics and machine learning methods for the intriguing purposes of cancer imaging like prognosticating histopathological parameters and treatment response. Numerous studies have assessed the potential of MRI and CT radiomic features as a non-invasive approach to prognosticate the Gleason score of prostate cancer [8–10]. Yet, reports on the use of PET imaging are very sparse, with only a handful of studies. Aksu et al. [11], Ghezzo et al. [12] and Khateri et al. [13] predicted Gleason grade in patients with PCa using ^{68}Ga-PSMA-PET radiomic features. Compared to previous studies, most of them segmented the whole prostate gland. In addition, it is noticeable that models combining PET and CT features together were rarely explored before [14]. The aim of this study was to evaluate radiomic features derived from ^{68}Ga-PSMA PET/CT images, as PET only, CT only and then as a combination of PET and CT features to predict benign and malignant tumors with regards to pathological GS in patients with PCa.

2 Materials and Methods

Figure 1 illustrates the schematic workflow utilized in our retrospective study.

2.1 Patients

The study included 138 patients that were divided into two groups corresponding to low- and high-risk PCa, classified as malignant above 4 + 3 and benign tumors under 3 + 4 according to the National Comprehensive Cancer Network (NCCN) classification (NCCN.org) [15]. Clinical information was extracted from the hospital database. Patients with pathologically proven primary PCa, referred between December 2018 and March 2023 underwent ^{68}Ga-PSMA-PET/CT for initial staging of PCa prior to giving any medical treatment, surgery, radiotherapy and chemotherapy. The patients' characteristics are summarized in Table 1.

2.2 Imaging Protocol

Imaging was performed on a Biograph 6 TrueV PET/CT scanner (Siemens Healthineers, Erlangen, Germany) equipped with a 6-slice CT scanner 60 min post-injection of ^{68}Ga-PSMA. PET/CT images from skull to mid-thigh were acquired for all patients. A filtered backprojection (FBP) algorithm was used for CT image reconstruction. The reconstruction matrix size of CT images was 512×512 resulting in a 0.97×0.97 mm^2 pixel size and 5 mm slice thickness. PET images were reconstructed using 3D-OSEM algorithm, 2 iterations and 21 subsets with 3 mm FWHM post-smoothing Gaussian filter. Resolution recovery, normalization, attenuation, scatter and decay corrections were applied. The reconstructed matrix size was 168×168 resulting in a $4.07 \times 4.07 \times 2.027$ mm^3 voxel size.

2.3 Image Segmentation

Volumes of interest (VOIs) of prostate tumors on ^{68}Ga-PSMA PET/CT images were toggled using 3D-Slicer software (version 5.2.1). Tumors in the prostate bed on PET and CT images were delineated and segmented manually using the 3D-Slicer software by an experienced nuclear medicine physician. Standardized uptake value (SUV) conversion factors were computed based on body weight, lean body mass, body surface area, and ideal body weight on PET images. Using resample image module in 3D-Slicer, CT images were resampled to a new resolution and spacing, applied a transformation (using an ITK transform) and wrapped. A sample of segmented prostate tumor using 3D-Slicer is illustrated in Fig. 2.

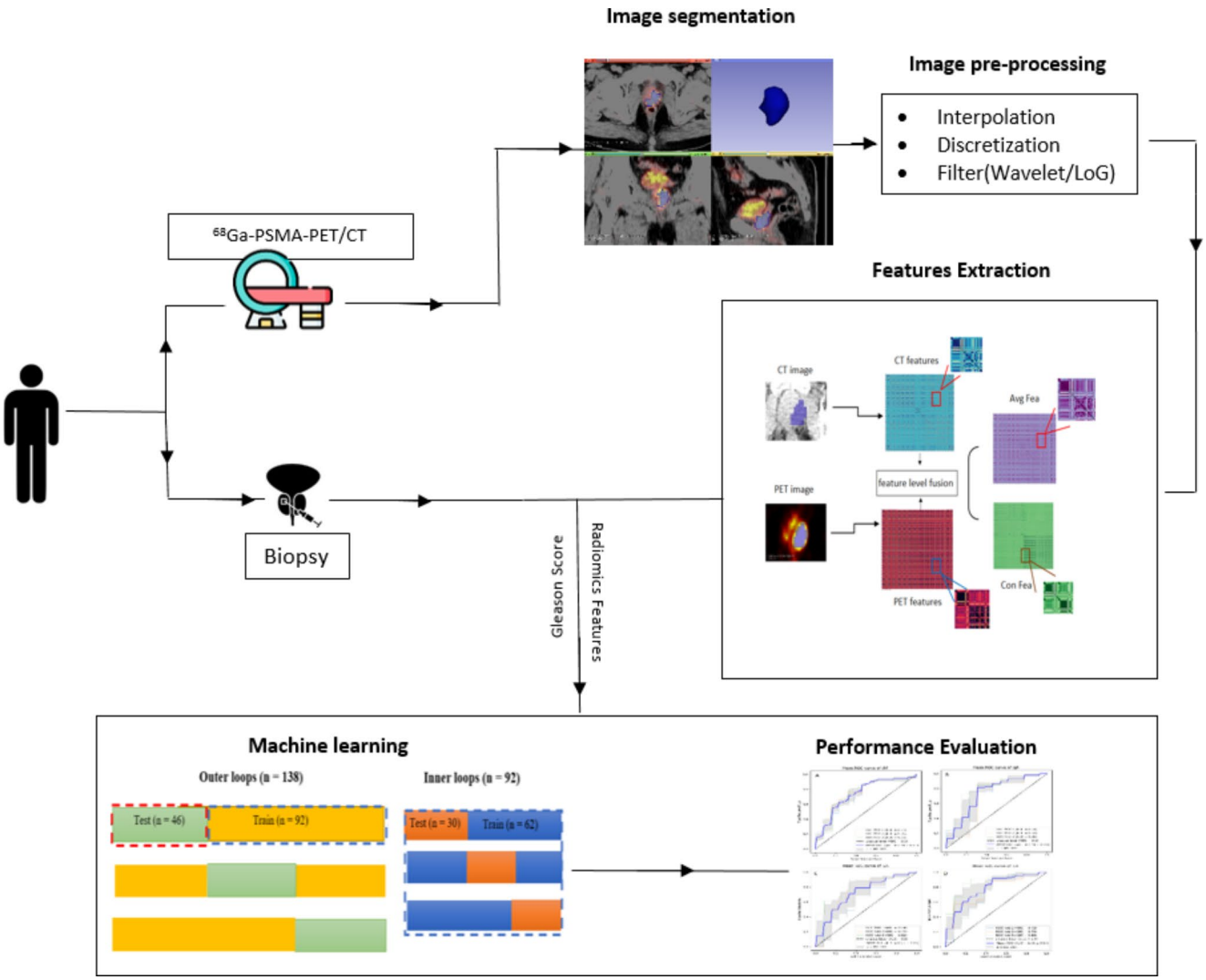

Fig. 1 The framework adopted in this study protocol

Table 1 Patient characteristics. Data reflect median or n (%). PSA: prostate-specific antigen; GS: Gleason score

	All eligible patients ($n=138$)
Age (years)	68 (48–83)
Weight (kg)	75 (104 − 48)
GS	
$\leq 3+4$	54 (39%)
$\geq 4+3$	83 (61%)
PSA (ng^{-1} ml)	84.69 (167.3–2.08)

2.4 Image Preprocessing

Image preprocessing is an important component that aims to standardize and enhance image quality, minimize the variability and improve the accuracy of radiomics analysis. Depending on the envisaged application, it is usually carried out in several steps. In our study, interpolation was utilized for resampling to avoid incomparable information from heterogeneous voxel spacing settings [16]. Therefore, $1\times1\times1$ mm^3 isotropic cubic voxels in CT and $4\times4\times4$ mm^3 in PET images were made using interpolation to isotropic voxel spacing [17, 18]. Discretization was performed with a fixed bin width setting of 64 bins in CT and 0.1 in SUV PET images to decrease noise and improve image contrast [18, 19]. Additionally, Laplacian of Gaussian (LoG) as a filter was applied to all input images. This filter serves the purpose of detecting both edges and noise. Therefore, the images were smoothed using Gaussian kernels with sigma values of 0.5, 1, 1.5, 2 and 2.5. Eventually, three-dimensional wavelet filter was applied to calculate features information in a variety of frequency of an image volume (x, y, z). First, the image is filtered along the x-dimension, resulting in a low-pass image (L) and a high-pass image (H) then filtered along the y-dimension, resulting in four subvolumes (LL, LH, HL, HH). Finally, each of these four subvolumes was filtered along the z-dimension, resulting in eight

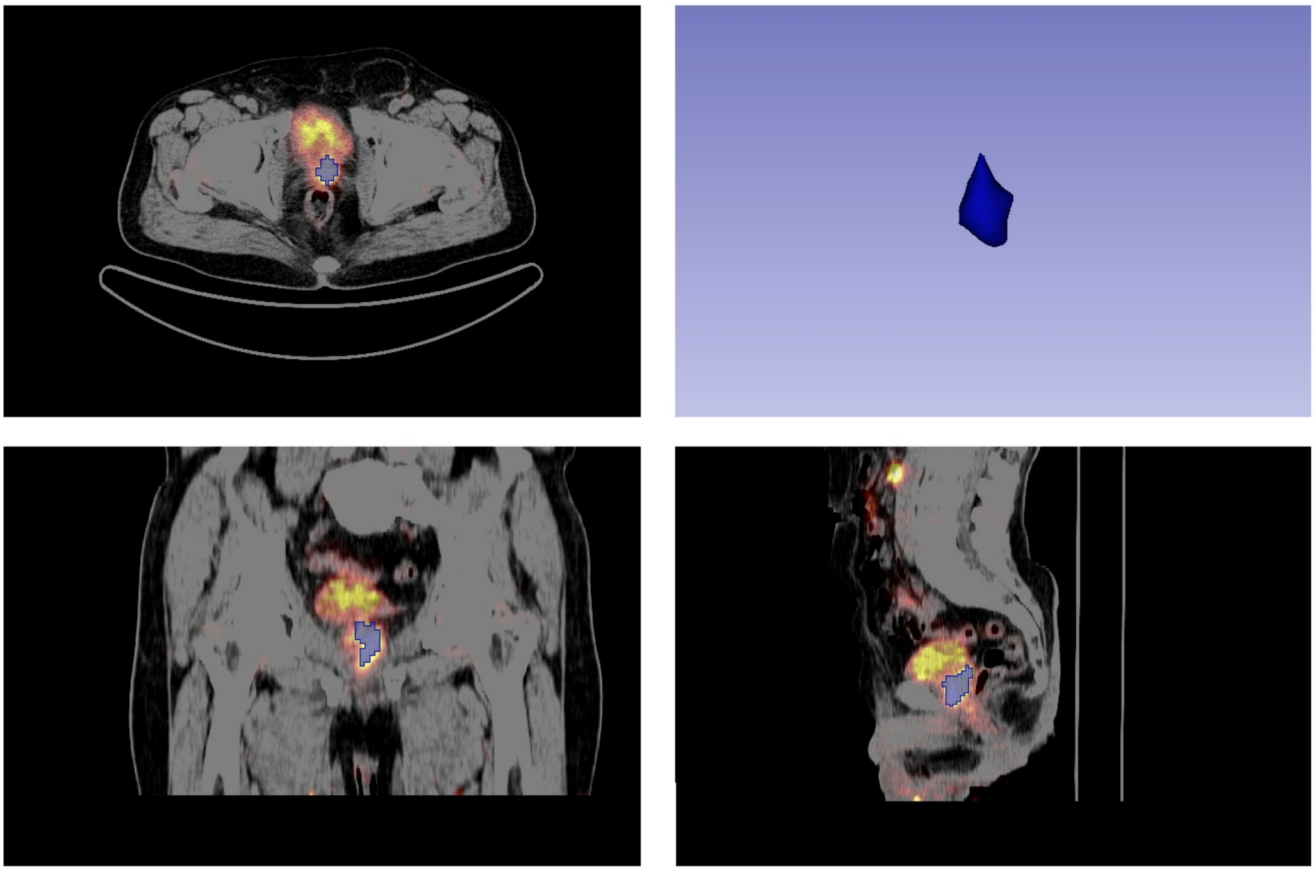

Fig. 2 Representative example of prostate tumor segmentation from a PET/CT image using 3D-slicer software

sub-volumes: LLH, LHL, LHH, HLL, HLH, HHL, HHH and LLL [18].

2.5 Radiomic Features Extraction

Radiomic features were extracted from a particular tumor VOI on PET and CT images following the guidelines of the image biomarker standardization initiative (IBSI) [20]. Features extraction was performed after preprocessing steps. In total, 1288 radiomic features were extracted from CT and PET images using PyRadiomics package, version 3.0.1, in Python [21]. This version contains various features; first-order ($n=18$), shape (3D) ($n=16$), Gray Level Co-occurrence Matrix (GLCM) ($n=24$), Gray Level Run Length Matrix (GLRLM) ($n=16$), Gray Level Size Zone Matrix (GLSZM) ($n=16$), Neighboring Gray Tone Difference Matrix (NGTDM) ($n=5$) and Gray Level Dependence Matrix (GLDM) ($n=14$). For feature-level fusion, two different methods were considered. Firstly, concatenation of features from PET and CT images (PET/CT_{ConFea}) and in the second way, the extracted features from PET and CT were averaged (PET/CT_{AveFea}) [14]. Figure 1 shows the schematics of feature extraction.

2.6 Feature Selection and Model Performance Evaluation

Nested cross-validation was applied with 3-fold cross-validation for the outer and inner loops [22, 23]. The dataset consisting of 138 patients was considered for the outer loop. In each fold, a set of 15 optimal features having the most correlation with GS and the least correlation between the extracted features were found by MRMR. Next, these selected features were fed to Recursive Feature Elimination with cross-validation (RFEcv) which operates by fitting SVM inside. Five machine learning models including Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF) and Support Vector Machine (SVM) were trained to determine the best model and best hyperparameters in the inner loops. In each cross-validation outer loop, two-thirds of the patients ($n=92$) were selected as training data, leaving the remaining ($n=46$) as the test data. The training data was split into training and validation datasets with a 2:1 patient ratio for training ($n=62$) and validation ($n=30$) in each inner fold. Considering the imbalanced distribution in classes, minority class oversampling was performed using the synthetic

minority oversampling technique (SMOTE) to oversample all training features in the less class (GS ≤ 3 + 4) up to a 1:1 proportion with the majority class (GS ≥ 4 + 3) [24, 25]. Subsequently, the best model of each outer loop was determined using the area under the receiver-operating characteristic (ROC) curve (AUC). Models were trained for CT, PET, AveFea and ConFea features. The trained models for GS prediction were tested in Nested cross-validation with the non-augmented validation data in three outer loops [23]. The performance of training and test set of the models is illustrated in Fig. 3.

2.7 Statistical Analysis

All statistical analysis was performed using scipy.stats library in Python [24, 26]. Wilcoxon signed-rank test was used to compare the AUC_{Mean} of the GS predictive classifiers. Statistical analysis was performed using R statistical software (version 4.2) [27]. The AUCs of all obtained models from four different imaging modalities were compared with each other two by two. Three modes including significantly lower, significantly higher and non-significant differences were defined based on the *P*-value. In addition, AUC-ROC was evaluated. The predicted GS was assessed with F_1 score, sensitivity and specificity of the best prediction for each outer fold using the test data [28, 29]. Finally, the model with the highest level of predictive performance was selected according to the statistical results.

3 Results

A total of 138 patients (one image for each) with a median age of 68 (range 48 to 83) were included in the study. Patients' characteristics are summarized in Table 1. Out of 138 patients, 55 (39%) had GS classes ≤ 3 + 4 (6 and 3 + 4) and 83 patients (61%) had GS classes ≥ 4 + 3 (4 + 3 to 10).

3.1 Hybrid Modalities Radiomics Analysis

In all models trained with an average of features derived from PET and CT, LR provided the highest AUC_{Mean} (0.72 ± 0.08), sensitivity of 0.74 ± 0.12, specificity of 0.63 ± 0.02 and F1-score of 0.71 ± 0.09 trained by averaging PET and CT radiomics features. Moreover, concatenated radiomic features of PET and CT were selected. The models were trained and the results showed that LR achieved the best predictive performance in the ConFea model with AUC_{Mean} of 0.78 ± 0.07, sensitivity of 0.81 ± 0.09, specificity of 0.66 ± 0.15, and F1-score of 0.77 ± 0.06.

3.2 Single Modality Radiomics Analysis

CT radiomic features using RF achieved the best performance with AUC_{Mean}, sensitivity, specificity and F1-score of 0.74 ± 0.01, 0.75 ± 0.07, 0.62 ± 0.08 and 0.72 ± 0.03, respectively. Likewise, by training models based on selected radiomic features of PET images, the best performance was obtained for LR classifier with

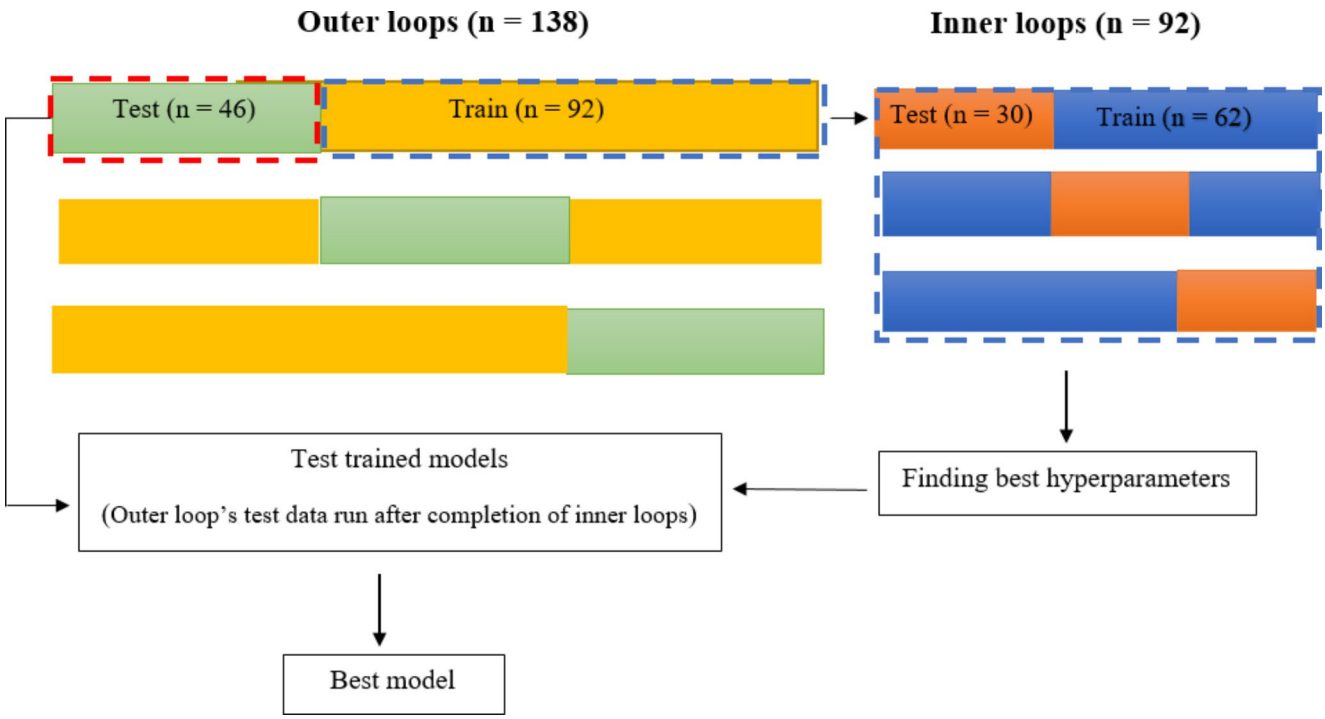

Fig. 3 The performance of nested cross-validation algorithm. The three best models were achieved in each outer loop

Table 2 Selected features obtained from PET/CTConFea images in three outer loops

No.	Loop 1	Loop 2	Loop 3
1	PET original glcm Sum Squares	PET wavelet LLH gldm Low Gray Level Emphasis	PET wavelet HHL glszm Large Area Emphasis
2	PET log sigma 2 mm 3D ngtdm Coarseness	PET log sigma 2.5 mm 3D gldm Small Dependence Emphasis	PET wavelet HHH first-order Skewness
3	PET wavelet LLH gldm Large Dependence Low Gray Level Emphasis	PET wavelet HHL firstorder Mean	PET wavelet HLL first-order Median
4	PET log sigma 2.5 mm 3D glszm Size Zone Non-Uniformity Normalized	PET wavelet LLH glrlm Low Gray Level Run Emphasis	PET log sigma 0.5 mm 3D glszm Small Area Emphasis
5	CT wavelet HHH glrlm Gray Level Variance	PET log sigma 2.5 mm 3D gldm Dependence Non-Uniformity Normalized	PET wavelet HHL gldm Dependence Non-Uniformity Normalized
6	CT wavelet LHL gldm Large Dependence Emphasis	PET wavelet LLL glszm Small Area Low Gray Level Emphasis	PET wavelet HHL glcm Idmn
7	CT wavelet HHL glszm High Gray Level Zone Emphasis	PET wavelet HHL glcm Correlation	CT wavelet LHH first-order Mean Absolute Deviation
8	CT wavelet LHH glrlm Short Run Low Gray Level Emphasis	PET wavelet LHH glszm Size Zone Non-Uniformity Normalized	CT wavelet LHH glcm Joint Energy
9	CT wavelet HLH firstorder Skewness	PET wavelet LLL glrlm Low Gray Level Run Emphasis	CT wavelet HHH glszm Low Gray Level Zone Emphasis
10	CT wavelet HLH gldm Large Dependence High Gray Level Emphasis	CT wavelet HHH glszm Small Area High Gray Level Emphasis	CT wavelet LHH glszm High Gray Level Zone Emphasis
11	CT original glszm Zone%	CT wavelet HHL glrlm Low Gray Level Run Emphasis	CT wavelet HLH glcm Imc2
12		CT wavelet LHH glszm High Gray Level Zone Emphasis	CT wavelet HHH glszm Small Area High Gray Level Emphasis
13		CT wavelet HHL glszm Low Gray Level Zone Emphasis	CT wavelet LHL first-order Variance

Table 3 Performance of the trained classifiers on the test data sets models for each modality, in terms of Mean AUC, Mean sensitivity, Mean specificity and Mean F1 score

Modality	classifier	Mean AUC	Mean sensitivity	Mean specificity	Mean F_1 score
CT	Random Forest	0.74 ± 0.01	0.75 ± 0.07	0.62 ± 0.08	0.72 ± 0.03
PET	Logistic Regression	0.74 ± 0.05	0.75 ± 0.13	0.78 ± 0.14	0.7 ± 0.08
AveFea (PET/CT)	Logistic Regression	0.72 ± 0.08	0.74 ± 012	0.63 ± 0.02	0.71 ± 0.09
ConFea (PET/CT)	Logistic Regression	0.78 ± 0.07	0.81 ± 0.09	0.66 ± 0.15	0.77 ± 0.09

AUC_{Mean}, sensitivity, specificity and F1-score of 0.74 ± 0.05, 0.75 ± 0.13, 0.78 ± 0.14 and 0.70 ± 0.08, respectively.

3.3 Selected Features

The LR classifier trained with concatenated selected radiomic features by RFEcv achieved the best predictive performance. The selected features in the outer loops are provided in Table 2.

3.4 Statistical Analysis

The best predictive model (achieving the highest AUC score in the test) for each imaging modality (PET, CT and PET/CT fusion) was adopted for comparison. The characteristics of the best predictive models are summarized in Table 3. ROC curves of the best radiomic model of all images (LR in PET images, ConFea and AveFea and RF in CT images) are presented in Fig. 4. Supplementary Table 1S summarizes the results achieved by all trained models.

As shown in Fig. 5, five different models, namely DT, KNN, LR, RF and SVM, obtained from PET, CT, AveFea and ConFea, were compared using the Wilcoxon rank sum test. This figure is encoded with three colors. A significantly higher *P*-value occurs when the AUC of a particular model in a row is higher than the AUC of another model in a column. Conversely, a significantly lower *P*-value occurs when the AUC of a particular model in a row is lower than the AUC of another model in a column. Otherwise, the *P*-value is considered non-significant if the AUCs in both the row and column are approximately similar. The results of statistical analysis showed that PET/CT_{ConFea}-LR radiomics model achieved the best predictive performance with statistically

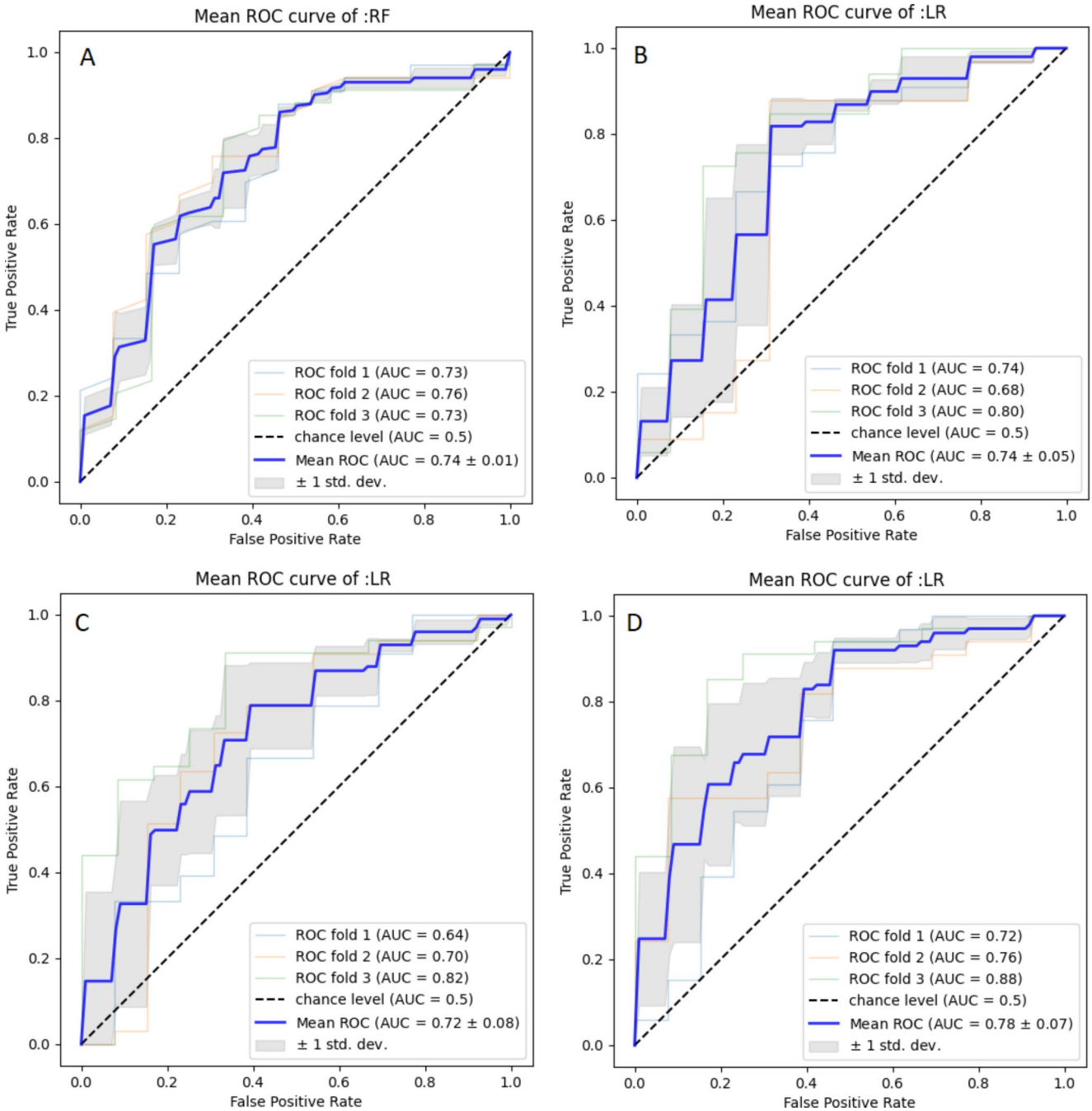

Fig. 4 ROC curve and Mean AUC of the best predictive models LR and RF for all 3 outer loops in **(A)** PET-only (LR), **(B)** CT-only (RF), **(C)** AveFea (LR) and **(D)** ConFea (LR)

significant difference (P-value < 0.05). In addition, the fifteen different models were compared.

4 Discussion

Recent literature has demonstrated the potential of combining radiomics and machine learning techniques for a number of applications in prostate cancer imaging, including predicting histopathological parameters and treatment response [8, 26, 30]. Several studies have evaluated the use of MRI radiomic features as a non-invasive approach to predict the Gleason score [2, 8, 24, 31, 32]. Abdollahi et al. introduced several radiomic models extracted from MRI to predict IMRT response [8]. They also demonstrated that using radiomic features was effective to predict GS ≥ 7 in patients with PCa. Two related studies using PET/CT imaging are worth mentioning. Ghezzo et al. analyzed the role of

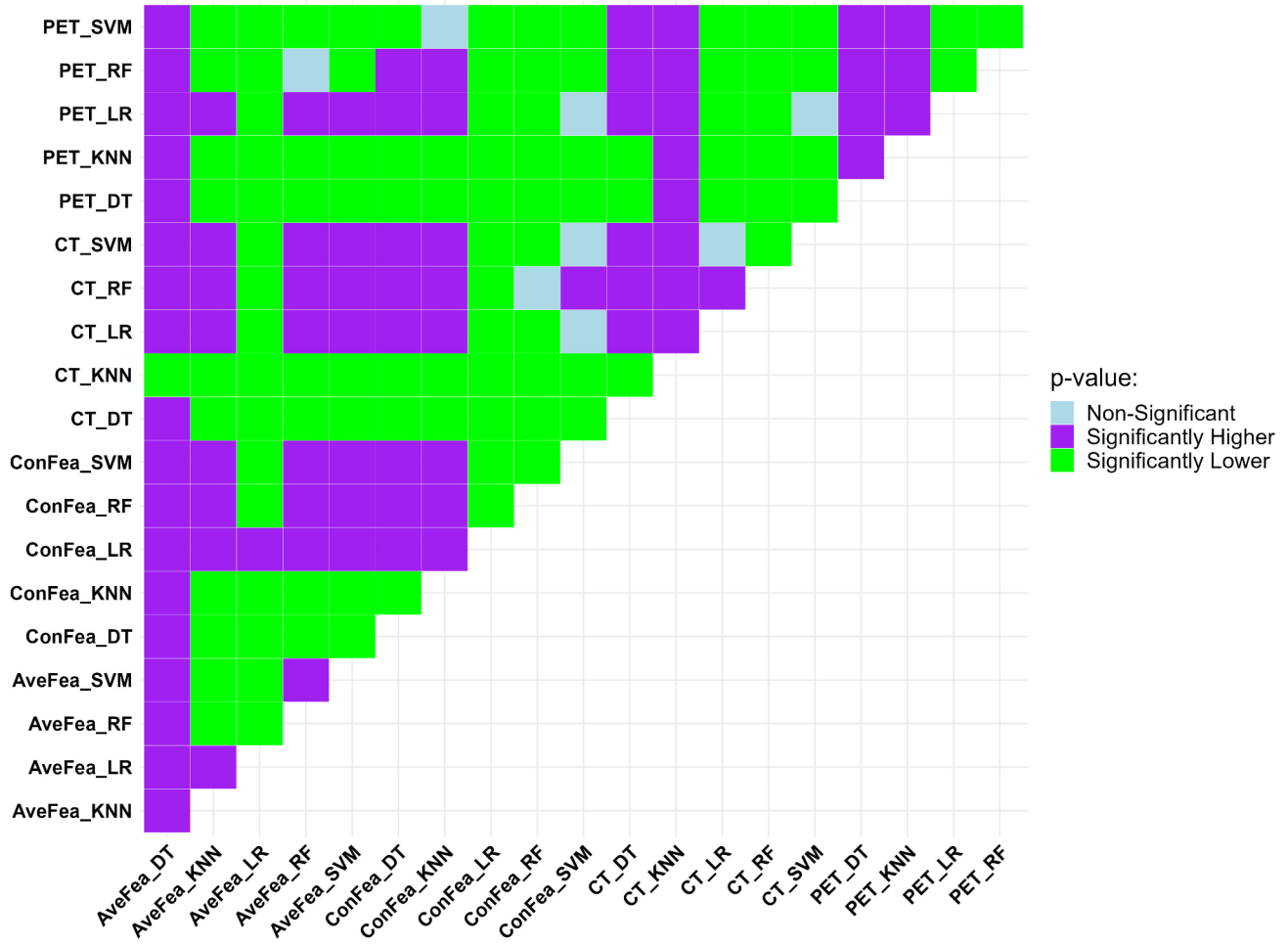

Fig. 5 The models on columns and rows were evaluated against each other. Light blue: if the row model outperformed non-significantly the column model in terms of *P*-value. Purple: in case the row model outperformed significantly higher the column model in terms of *P*-value. Green: when the row model outperformed significantly lower the column model in terms of *P*-value

^{68}Ga-PSMA-PET radiomics for the prediction of the International Society of Urological Pathology (ISUP) grade [12] whereas Aksu et al. evaluated the performance of models based on PET radiomic features to predict Gleason Grade [11].

Tumor segmentation as a noticeable step in radiomics studies might impact the results. Solari et al. [24], Ghezzo et al. [12] and Aksu et al. [11] segmented the whole prostate gland to extract radiomic features, while the malignant part of the prostate or only lesions in the bed of the prostate was segmented in our study. Indeed, selecting normal tissue could impact the correlation between the extracted radiomic features and GS.

In the present study, nested cross-validation algorithm was utilized to train five classifier models by handcrafted radiomic features extracted from ^{68}Ga-PSMA-PET/CT images to predict biopsy Gleason Grade. In this study, we used MRMR to select the most effective features. RFEcv was applied on the selected feature by MRMR method to find the best combination of features. RFEcv operates by fitting an SVM model to all the features, and the SVM taking off the feature that is considered less relevant to leave only the highest ranked and less interdependent features. For comparison with other studies, Solari et al. used RFE method for the selection of the radiomic features [24] whereas Osman et al. applied the function '*find correlation*' in the R package Caret to select the features [33]. Among the features selected in three loops, less than 20% of them have been influenced by the Gaussian filter and the wavelet filter was affected on a significant percentage of features (the first loop 70%, the second loop 84% and the third loop more than 90%). As opposed to other methods, the wavelet filter does not consider that the coefficients are independent, and the computation time is modest. Hence, the wavelet filter could give striking results compared to other filters [34].

The repetitive feature included in all outer loops in the models obtained was GLSZM-HGLZE, which defines the measurement of the distribution of the high gray-level values, with a high value indicating a greater proportion of higher gray-level values in the image. GLRLM-LGLRE were repeated in different frequencies three times in the second loop which measures the distribution of low gray-level values. These findings support the claim that machine learning models trained with a feature of non-uniformity (i.e. GLSZM, GLRLM) might be sufficient to accurately characterize PCa in the staging phase. Solari et al. showed that GLSZM strongly correlates with GLDM, GLCM, NGTDM, SUV_{mean}, 90th Percentile, and GLDM [24]. Aksu et al. reported that both GLRLM-RLNU and Shape_compacity features were higher, indicating greater heterogeneity of high-grade tumors, in which Shape_compacity shows how compact the area of interest and GLRLM-RLNU indicated more homogeneity among run lengths in the image [11].

In contrast to previous studies that focused only on one imaging modality, we used CT-only, PET-only, PET/CT_{AveFea} and PET/CT_{ConFea} models. For instance, Chaddad et al. used mpMRI-based radiomics model to predict GS [2], whereas Cuocolo investigated the correlation between features derived from MRI and GS above (4 + 3) [31]. Osman et al. investigated the role of CT-based radiomic features in risk group classification of PCa patients [33].

Solari et al. evaluated Gleason categories using radiomic features of PET/MR images by training only SVM as a multi-class model. They reported balanced accuracy (bAcc) of single and hybrid imaging modalities and concluded that the combination of PET + Apparent Diffusion Coefficient (ADC) was the best classifier model with bAcc of 0.81 [24]. In our study, the highest performance achieved by the hybrid modality, PET/CT_{ConFea}, was 0.78. The reason for the difference between the two studies can be explained by the higher differentiation capability of the extracted features from PET and MRI that were used for the fusion of two modalities.

Zamboglou et al. proposed a radiomics model derived from ^{68}Ga-PSMA PET images that reached an AUC of 0.84 for the discrimination of GS and pelvic lymph node status [35]. Aksu et al. achieved an AUC of 0.90 for the prediction of the grade group of prostate cancer patients [11]. Our results showed that LR built based on PET/CT_{ConFea} features achieved the best predictive performance with AUC_{Mean} of 0.78 ± 0.07.

In the above referenced study, 41 features were extracted from 41 male patients [11]. It has been suggested that datasets containing 10–15 patients per feature are ideal for evaluating the radiomics predictor models [36]. Increasing the number of features could rise the chance of overfitting, and as such, it might be that their model was overtrained because of the number of extracted features, thus increasing the AUC value. In addition, utilizing cross-validation decreases overfitting. Aksu et al. did not employ cross-validation in their study, potentially leading to overtraining, whereas our study used nested cross-validation to mitigate overfitting [37].

Ghezzo et al. aimed to predict GS > 8 and GS < 8 in 47 PCa patients who underwent ^{68}Ga- PSMA PET from two different scanners before radical prostatectomy, where the SVM model achieved a bAcc of 87% using radiomic features extracted from ^{68}Ga-PSMA PET images [12]. Our results demonstrated that the LR model trained with PET/CT_{ConFea} data was able to predict GS with the highest performance, achieving an AUC_{Mean} of 0.78 ± 0.07 in PCa patients. In addition, the assessment of single modality images showed that logistic regression built based on PET and CT features achieved the best predictive performance with AUC_{Mean} of 0.74. The reason why the AUC was less in the present study compared to the above-mentioned studies could be attributed to the following reasons. Ghezzo's et al. performed a multicentric study that used ComBat harmonization method. Applying ComBat removes all radiomic features with significantly different distributions. Hence, on the whole, the predictive ability of radiomics models was improved through harmonization by ComBat [12]. Therefore, the AUC obtained from Ghezzo's study was higher than the AUC achieved in the present study.

Despite these bright results, the lack of interpretability and standardization in the field of radiomics remains a concern hindering the adoption of radiomics analysis in clinical setting. The approach requires multi-institutional external validation on a larger scale before wider clinical implementation. We considered 46 images to evaluate our model's performance. For this reason, the methodology of the present study was developed to ensure realistic and explainable results, building on previously acquired knowledge. The performance of deep learning is generally better than machine learning, at the expense of the requirement of huge data. Therefore, it is suggested that future studies use multicenter data for deep learning analysis.

5 Conclusion

The findings of our study demonstrated that utilizing radiomic features extracted from ^{68}Ga-PET/CT images could potentially be an effective non-invasive approach to predict pathological indices, such as GS in primary PCa patients.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40846-024-00906-3.

Author Contributions Farshad Emami, Habibeh Vosoughi and Parham

Geramifar contributed to the study conception and design. Material preparation, data collection and analysis were performed by all authors. The first draft of the manuscript was written by Zahra Vosoughi, Ghasem Hajianfar and Nima Hamzian, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding We are grateful to the staff and management of PET/CT centers for their cooperation This study was part of a MSc thesis supported under grant number 14056 by Yazd University of Medical Sciences, Yazd, Iran.

Declarations

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethics Approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Shahid Sadoughi University of Medical Sciences IR.SSU.MEDICINE.REC.1401.113.

References

1. Siegel, R. L., Miller, K. D., Wagle, N. S., Jemal, A., & Cancer statistics (2023). CA Cancer J Clin. 2023;73(1):17–48. doi: 10.3322/caac.21763. PMID: 36633525.
2. Chaddad, A., Niazi, T., Probst, S., Bladou, F., Anidjar, M., & Bahoric, B. (2018). Predicting Gleason score of prostate Cancer patients using Radiomic Analysis. *Frontiers in Oncology*, *8*, 630. https://doi.org/10.3389/fonc.2018.00630PMID: 30619764; PMCID: PMC6305278.
3. Zhu, X., Gou, X., & Zhou, M. (2019). Nomograms Predict Survival Advantages of Gleason Score 3 + 4 over 4 + 3 for prostate Cancer: A SEER-Based study. *Frontiers in Oncology*, *9*, 646. https://doi.org/10.3389/fonc.2019.00646PMID: 31380282; PMCID: PMC6646708.
4. Spohn, S. K. B., Bettermann, A. S., Bamberg, F., Benndorf, M., Mix, M., Nicolay, N. H., Fechter, T., Hölscher, T., Grosu, R., Chiti, A., Grosu, A. L., & Zamboglou, C. (2021). Radiomics in prostate cancer imaging for a personalized treatment approach - current aspects of methodology and a systematic review on validated studies. *Theranostics*, *11*(16), 8027–8042. https://doi.org/10.7150/thno.61207PMID: 34335978; PMCID: PMC8315055.
5. Jochumsen, M. R., Bouchelouche, K., & PSMA PET/CT for Primary Staging of Prostate Cancer - An Updated Overview. (2024). *Seminars in Nuclear Medicine*. ;54(1):39–45. doi: https://doi.org/10.1053/j.semnuclmed.2023.07.001. Epub 2023 Jul 22. PMID: 37487824.
6. Mayerhoefer, M. E., Materka, A., Langs, G., Häggström, I., Szczypiński, P., Gibbs, P., & Cook, G. (2020). Introduction to Radiomics. *Journal of Nuclear Medicine*, *61*(4), 488–495. https://doi.org/10.2967/jnumed.118.222893Epub 2020 Feb 14. PMID: 32060219; PMCID: PMC9374044.
7. Shiri, I., Sorouri, M., Geramifar, P., Nazari, M., Abdollahi, M., Salimi, Y., Khosravi, B., Askari, D., Aghaghazvini, L., Hajianfar, G., Kasaeian, A., Abdollahi, H., Arabi, H., Rahmim, A., Radmard, A. R., & Zaidi, H. (2021). Machine learning-based prognostic modeling using clinical data and quantitative radiomic features from chest CT images in COVID-19 patients. *Computers in Biology and Medicine*, *132*, 104304. Epub 2021 Mar 3. PMID: 33691201; PMCID: PMC7925235.
8. Abdollahi, H., Mofid, B., Shiri, I., Razzaghdoust, A., Saadipoor, A., Mahdavi, A., Galandooz, H. M., & Mahdavi, S. R. (2019). Machine learning-based radiomic models to predict intensity-modulated radiation therapy response, Gleason score and stage in prostate cancer. Radiol Med. ;124(6):555–567. https://doi.org/10.1007/s11547-018-0966-4. Epub 2019 Jan 3. PMID: 30607868.
9. Khateri, M., Babapour Mofrad, F., Jenabi, E., Hajianfar, G., Jafari, E., Dadgar, H., et al. (2022). Non-invasive prostate Cancer Histopathological Subtype Decoding using 68Ga-PSMA PET/CT Radiomics features: A multi-center study. *Journal of Nuclear Medicine*, *63*(supplement 2), 3242.
10. Bagheri, S., Hajianfar, G., Saberi, A., Oveisi, M., Shiri, I., Zaidi, H., Pathological Prostate Gleason Score Prediction Using MRI Radiomics and Machine Learning Algorithms, 2021 IEEE Nuclear Science Symposium and Medical Imaging, & Conference (2021). (NSS/MIC), Piscataway, NJ, USA, pp. 1–3, https://doi.org/10.1109/NSS/MIC44867.2021.9875791
11. Aksu, A., Vural Topuz, Ö., Yılmaz, G., Çapa Kaya, G., & Yılmaz, B. (2022). Dual time point imaging of staging PSMA PET/CT quantification; spread and radiomic analyses. *Annals of Nuclear Medicine*, *36*(3), 310–318. Epub 2022 Jan 6. PMID: 34988888.
12. Ghezzo, S., Mapelli, P., Bezzi, C., Samanes Gajate, A. M., Brembilla, G., Gotuzzo, I., Russo, T., Preza, E., Cucchiara, V., Ahmed, N., Neri, I., Mongardi, S., Freschi, M., Briganti, A., De Cobelli, F., Gianolli, L., Scifo, P., & Picchio, M. (2023). Role of [68Ga]Ga-PSMA-11 PET radiomics to predict post-surgical ISUP grade in primary prostate cancer. *European Journal of Nuclear Medicine and Molecular Imaging*, *50*(8), 2548–2560. https://doi.org/10.1007/s00259-023-06187-3Epub 2023 Mar 18. PMID: 36933074.
13. Khateri, M., Babapour Mofrad, F., Geramifar, P., & Jenabi, E. (2024). Machine learning-based analysis of 68Ga-PSMA-11 PET/CT images for estimation of prostate tumor grade. *Phys Eng Sci Med*, *47*(2), 741–753. https://doi.org/10.1007/s13246-024-01402-3Epub 2024 Mar 25. PMID: 38526647.
14. Amini, M., Nazari, M., Shiri, I., Hajianfar, G., Deevband, M. R., Abdollahi, H., Arabi, H., Rahmim, A., & Zaidi, H. (2021). Multi-level multi-modality (PET and CT) fusion radiomics: prognostic modeling for non-small cell lung carcinoma. Phys Med Biol. ;66(20). https://doi.org/10.1088/1361-6560/ac287d. PMID: 34544053.
15. Kryvenko, O. N., Lyapichev, K., Chinea, F. M., Prakash, N. S., Pollack, A., Gonzalgo, M. L., Punnen, S., & Jorda, M. (2016). Radical prostatectomy findings in White Hispanic/Latino men with NCCN very low-risk prostate Cancer detected by Template Biopsy. *American Journal of Surgical Pathology*, *40*(8), 1125–1132. https://doi.org/10.1097/PAS.0000000000000656PMID: 27158756; PMCID: PMC4949597.
16. Philippe Thévenaz, T., Blu, & Unser, M. (2000). *Image interpolation and resampling. Handbook of medical imaging* (pp. 393–420). Academic Press, Inc.
17. van Timmeren, J. E., Cester, D., Tanadini-Lang, S., Alkadhi, H., & Baessler, B. (2020). Radiomics in medical imaging-how-to guide and critical reflection. *Insights Imaging*, *11*(1), 91. https://doi.org/10.1186/s13244-020-00887-2PMID: 32785796; PMCID: PMC7423816.
18. Chou, Y., Peng, S. H., Lin, H. Y., Lan, T. L., Jiang, J. K., Liang, W. Y., Hu, Y. W., & Wang, L. W. (2023). Radiomic features derived from pretherapeutic MRI predict chemoradiation response in locally advanced rectal cancer. *Journal of the Chinese Medical Association : Jcma*, *86*(4), 399–408. Epub 2023 Jan 24. PMID: 36727777.
19. Schick, U., Lucia, F., Dissaux, G., Visvikis, D., Badic, B., Masson, I., Pradier, O., Bourbonne, V., & Hatt, M. (2019). MRI-derived radiomics: Methodology and clinical applications in the

field of pelvic oncology. *British Journal of Radiology*, *92*(1104), 20190105. https://doi.org/10.1259/bjr.20190105Epub 2019 Oct 10. PMID: 31538516; PMCID: PMC6913356.
20. Hatt, M., Le Rest, C. C., Tixier, F., Badic, B., Schick, U., & Visvikis, D. (2019). Radiomics: Data are also images. *Journal of Nuclear Medicine*, *60*(Supplement 2), 38S–44S.
21. van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J. C., Pieper, S., & Aerts, H. J. W. L. (2017). Computational Radiomics System to Decode the Radiographic phenotype. *Cancer Research*, *77*(21), e104–e107. https://doi.org/10.1158/0008-5472.CAN-17-0339PMID: 29092951; PMCID: PMC5672828.
22. Rundo, L., Beer, L., Escudero Sanchez, L., Crispin-Ortuzar, M., Reinius, M., McCague, C., Sahin, H., Bura, V., Pintican, R., Zerunian, M., Ursprung, S., Allajbeu, I., Addley, H., Martin-Gonzalez, P., Buddenkotte, T., Singh, N., Sahdev, A., Funingana, I. G., Jimenez-Linan, M., Markowetz, F., Brenton, J. D., Sala, E., & Woitek, R. (2022). Clinically interpretable Radiomics-based prediction of histopathologic response to Neoadjuvant Chemotherapy in High-Grade Serous Ovarian Carcinoma. *Frontiers in Oncology*, *12*, 868265. https://doi.org/10.3389/fonc.2022.868265PMID: 35785153; PMCID: PMC9243357.
23. Blüthgen, C., Patella, M., Euler, A., Baessler, B., Martini, K., von Spiczak, J., Schneiter, D., Opitz, I., & Frauenfelder, T. (2021). Computed tomography radiomics for the prediction of thymic epithelial tumor histology, TNM stage and myasthenia gravis. *PLoS One*, *16*(12), e0261401. https://doi.org/10.1371/journal.pone.0261401PMID: 34928978; PMCID: PMC8687592.
24. Solari, E. L., Gafita, A., Schachoff, S., Bogdanović, B., Villagrán Asiares, A., Amiel, T., Hui, W., Rauscher, I., Visvikis, D., Maurer, T., Schwamborn, K., Mustafa, M., Weber, W., Navab, N., Eiber, M., Hatt, M., & Nekolla, S. G. (2022). The added value of PSMA PET/MR radiomics for prostate cancer staging. *European Journal of Nuclear Medicine and Molecular Imaging*, *49*(2), 527–538. Epub 2021 Jul 13. PMID: 34255130; PMCID: PMC8803696.
25. Min, X., Li, M., Dong, D., Feng, Z., Zhang, P., Ke, Z., You, H., Han, F., Ma, H., Tian, J., & Wang, L. (2019). Multi-parametric MRI-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: Cross-validation of a machine learning method. *European Journal of Radiology*, *115*, 16–21. Epub 2019 Mar 15. PMID: 31084754.
26. Zamboglou, C., Bettermann, A. S., Gratzke, C., Mix, M., Ruf, J., Kiefer, S., Jilg, C. A., Benndorf, M., Spohn, S., Fassbender, T. F., Bronsert, P., Chen, M., Guo, H., Wang, F., Qiu, X., & Grosu, A. L. (2021). Uncovering the invisible-prevalence, characteristics, and radiomics feature-based detection of visually undetectable intraprostatic tumor lesions in 68GaPSMA-11 PET images of patients with primary prostate cancer. *European Journal of Nuclear Medicine and Molecular Imaging*, *48*(6), 1987–1997. https://doi.org/10.1007/s00259-020-05111-3Epub 2020 Nov 18. PMID: 33210239; PMCID: PMC8113179.
27. Hajianfar, G., Sabouri, M., Salimi, Y., Amini, M., Bagheri, S., Jenabi, E., Hekmat, S., Maghsudi, M., Mansouri, Z., Khateri, M., Hosein Jamshidi, M., Jafari, E., Bitarafan Rajabi, A., Assadi, M., Oveisi, M., Shiri, I., & Zaidi, H. (2024). Artificial intelligence-based analysis of whole-body bone scintigraphy: The quest for the optimal deep learning algorithm and comparison with human observer performance. *Zeitschrift Fur Medizinische Physik*, *34*(2), 242–257. Epub 2023 Mar 15. PMID: 36932023; PMCID: PMC11156776.
28. Long, L., Sun, J., Jiang, L., Hu, Y., Li, L., Tan, Y., Cao, M., Lan, X., & Zhang, J. (2021 Jul-Aug). MRI-based traditional radiomics and computer-vision nomogram for predicting lymphovascular space invasion in endometrial carcinoma. *Diagnostic and Interventional Imaging*, *102*(7–8), 455–462. Epub 2021 Mar 23. PMID: 33741266.
29. Xie, H., Ma, S., Guo, X., Zhang, X., & Wang, X. (2020). Preoperative differentiation of pancreatic mucinous cystic neoplasm from macrocystic serous cystic adenoma using radiomics: Preliminary findings and comparison with radiological model. *European Journal of Radiology*, *122*, 108747. https://doi.org/10.1016/j.ejrad.2019.108747Epub 2019 Nov 14. PMID: 31760275.
30. Delgadillo, R., Ford, J. C., Abramowitz, M. C., Dal Pra, A., Pollack, A., & Stoyanova, R. (2020). The role of radiomics in prostate cancer radiotherapy. *Strahlentherapie Und Onkologie*, *196*(10), 900–912. https://doi.org/10.1007/s00066-020-01679-9Epub 2020 Aug 21. PMID: 32821953; PMCID: PMC7545508.
31. Cuocolo, R., Stanzione, A., Ponsiglione, A., Romeo, V., Verde, F., Creta, M., La Rocca, R., Longo, N., Pace, L., & Imbriaco, M. (2019). Clinically significant prostate cancer detection on MRI: A radiomic shape features study. *European Journal of Radiology*, *116*, 144–149. Epub 2019 May 7. PMID: 31153556.
32. Gaudiano, C., Mottola, M., Bianchi, L., Corcioni, B., Cattabriga, A., Cocozza, M. A., Palmeri, A., Coppola, F., Giunchi, F., Schiavina, R., Fiorentino, M., Brunocilla, E., Golfieri, R., & Bevilacqua, A. (2022). Beyond Multiparametric MRI and towards Radiomics to detect prostate Cancer: A machine learning model to Predict clinically significant lesions. *Cancers (Basel)*, *14*(24), 6156. https://doi.org/10.3390/cancers14246156PMID: 36551642; PMCID: PMC9776977.
33. Osman, S. O. S., Leijenaar, R. T. H., Cole, A. J., Lyons, C. A., Hounsell, A. R., Prise, K. M., O'Sullivan, J. M., Lambin, P., McGarry, C. K., & Jain, S. (2019). Computed tomography-based Radiomics for risk stratification in prostate Cancer. *International Journal of Radiation Oncology Biology Physics*, *105*(2), 448–456. Epub 2019 Jun 26. PMID: 31254658.
34. Ganapathi, P., Subashini, P., Kumar, M., & Thakur, S. (2011). Comparison of filters used for underwater image Pre-processing. *International Journal of Computer Science and Network Security*, *10*, 58–65.
35. Zamboglou, C., Carles, M., Fechter, T., Kiefer, S., Reichel, K., Fassbender, T. F., Bronsert, P., Koeber, G., Schilling, O., Ruf, J., Werner, M., Jilg, C. A., Baltas, D., Mix, M., & Grosu, A. L. (2019). Radiomic features from PSMA PET for non-invasive intraprostatic tumor discrimination and characterization in patients with intermediate- and high-risk prostate cancer - a comparison study with histology reference. *Theranostics*, *9*(9), 2595–2605. https://doi.org/10.7150/thno.32376PMID: 31131055; PMCID: PMC6525993.
36. Sollini, M., Cozzi, L., Antunovic, L., Chiti, A., & Kirienko, M. (2017). PET Radiomics in NSCLC: State of the art and a proposal for harmonization of methodology. *Scientific Reports*, *7*(1), 358. https://doi.org/10.1038/s41598-017-00426-yPMID: 28336974; PMCID: PMC5428425.
37. Shinozaki, T., & Kawahara, T. (2008). GMM and HMM training by aggregated EM algorithm with increased ensemble sizes for robust parameter estimation. 4405–4408. https://doi.org/10.1109/ICASSP.2008.4518632