





Article

A Combined Radiomics and Machine Learning Approach to Distinguish Clinically Significant Prostate Lesions on a Publicly Available MRI Dataset

Leandro Donisi ^{1,2,†}, Giuseppe Cesarelli ^{2,3,†}, Anna Castaldo ¹, Davide Raffaele De Lucia ¹,
Francesca Nessuno ¹, Gaia Spadarella ^{1,†} and Carlo Ricciardi ^{2,4,*,†}

¹ Department of Advanced Biomedical Sciences, University of Naples “Federico II”, 80131 Naples, Italy; leandro.donisi@unina.it (L.D.); annacastaldo1202@gmail.com (A.C.); dav.delucia@gmail.com (D.R.D.L.); francesca.nessuno@gmail.com (F.N.); gaia.spadarella@gmail.com (G.S.)

² Bioengineering Unit, Institute of Care and Scientific Research Maugeri, 27100 Pavia, Italy; giuseppe.cesarelli@unina.it

³ Department of Chemical, Materials and Production Engineering, University of Naples “Federico II”, 80125 Naples, Italy

⁴ Department of Electrical Engineering and Information Technologies, University of Naples “Federico II”, 80125 Naples, Italy

* Correspondence: carloricciardi.93@gmail.com

† The authors equally contributed to the work.



Citation: Donisi, L.; Cesarelli, G.; Castaldo, A.; De Lucia, D.R.; Nessuno, F.; Spadarella, G.; Ricciardi, C. A Combined Radiomics and Machine Learning Approach to Distinguish Clinically Significant Prostate Lesions on a Publicly Available MRI Dataset. *J. Imaging* **2021**, *7*, 215. <https://doi.org/10.3390/jimaging7100215>

Academic Editor: Leonardo Rundo

Received: 8 September 2021

Accepted: 13 October 2021

Published: 18 October 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Although prostate cancer is one of the most common causes of mortality and morbidity in advancing-age males, early diagnosis improves prognosis and modifies the therapy of choice. The aim of this study was the evaluation of a combined radiomics and machine learning approach on a publicly available dataset in order to distinguish a clinically significant from a clinically non-significant prostate lesion. A total of 299 prostate lesions were included in the analysis. A univariate statistical analysis was performed to prove the goodness of the 60 extracted radiomic features in distinguishing prostate lesions. Then, a 10-fold cross-validation was used to train and test some models and the evaluation metrics were calculated; finally, a hold-out was performed and a wrapper feature selection was applied. The employed algorithms were Naïve bayes, K nearest neighbour and some tree-based ones. The tree-based algorithms achieved the highest evaluation metrics, with accuracies over 80%, and area-under-the-curve receiver-operating characteristics below 0.80. Combined machine learning algorithms and radiomics based on clinical, routine, multiparametric, magnetic-resonance imaging were demonstrated to be a useful tool in prostate cancer stratification.

Keywords: radiomics; machine learning; MRI; prostate cancer

1. Introduction

According to America Cancer Society, the estimated numbers of new cases and deaths from prostate cancer in the USA in 2021 are more than 240,000 and over 30,000, respectively [1]. As the prognosis of prostate cancer is strictly related to its biologically aggressive behavior, early detection and accurate risk stratification play a key role in ensuring the best outcome for patients [2]. In summary, clinically significant prostate cancer needs to be discriminated from low-grade disease to propose an adequate treatment to the patient [3]. To this end, magnetic resonance imaging (MRI) emerged as the most accurate imaging modality for the detection of clinically significant prostate cancer and actually plays a major role in the diagnostic pathway of the disease, since MRI is able to guide targeted biopsies [4,5]. Nevertheless, this technique has some limitations, such as the contrast-agent administration, a moderate specificity and the need for a high level of expertise to be correctly interpreted [6,7].

In recent years, radiomics and machine learning (ML) have shown their potential to extract quantitative features and elaborate them with complex algorithms to improve both the diagnosis and prognosis of patients.

Several authors demonstrated the advantage of the application of radiomics and ML, not only in prostate cancer but also in other fields of oncology [8–10]. In addition, recently, systematic reviews described the promising role of these techniques in prostate cancer [11–15]. The results of these studies suggested that, while MRI radiomics and ML approaches can reach high diagnostic accuracy in detecting severe prostate cancer and thus should be further investigated, the high heterogeneity of these studies has prevented their application in real life, indicating the need for standardized pipelines and the concomitant use of reliable benchmarks.

As a result, the aim of the present study is to evaluate the ability of the combined radiomics and ML approach using several ML algorithms (tree-based, instance-based and based on the a priori probability theory) on a publicly available dataset of MRI images, elaborated by Cuocolo et al. [16], in differentiating a clinically significant from a clinically non-significant prostate lesion.

Figure 1 summarizes the research workflow, which starts with MRI acquisition and ends with ML analysis.



Figure 1. Workflow of the research.

2. Materials and Methods

2.1. Dataset

A total of 299 verified prostate lesions were included in this study. Specifically, the lesion annotation masks were obtained from an online open repository (https://github.com/rcuocolo/PROSTATEx_masks, accessed on 1 July 2020) and coupled with the source MRI images, which can be found in the PROSTATEx training dataset (<https://wiki.cancerimagingarchive.net/display/Public/SPIE-AAPM-NCI+PROSTATEx+Challenges>, accessed on 1 July 2020) [16,17]. The ground-truth of the public dataset is obtained with a manual annotation. The 3×3 lesion and gland zone coordinate masks, freely available on a public repository (https://github.com/rcuocolo/PROSTATEx_masks, accessed on 1 July 2020), were retrieved by slice-by-slice segmentation on T2-weighted (T2w) and apparent diffusion coefficient (ADC) images, by the residents, with a subsequent check and eventual refinement by a radiologist. Of these 299 prostate lesions, 76 harbored clinically significant prostate cancer (cut-off = Gleason grade group ≥ 2) [18]. T2w and ADC maps images were used for the extraction of radiomic features. Images were obtained by two Siemens 3T MRI scanners, the MAGNETOM Trio and Skyra, without an endorectal coil. The acquisition of T2-w images was performed using a turbo-spin echo sequence with a resolution of around 0.5 mm in plane and a slice thickness of 3.6 mm. The ADC map was acquired by the scanner software from the diffusion-weighted imaging (DWI) (a single-shot echo planar imaging sequence with a resolution of 2 mm in-plane and 3.6 mm slice thickness, and with diffusion-encoding gradients in three directions) with three b-values (50, 400, and 800). Several algorithms were used to standardize signal intensity. Specifically, the T2-estimate map was obtained by using the MRI signal equation with an automated process [19] and the ADC map was automatically acquired from the diffusion-weighted images using the MRI scanner software. Figure 2 shows a clinically significant and a clinically non-significant lesion.

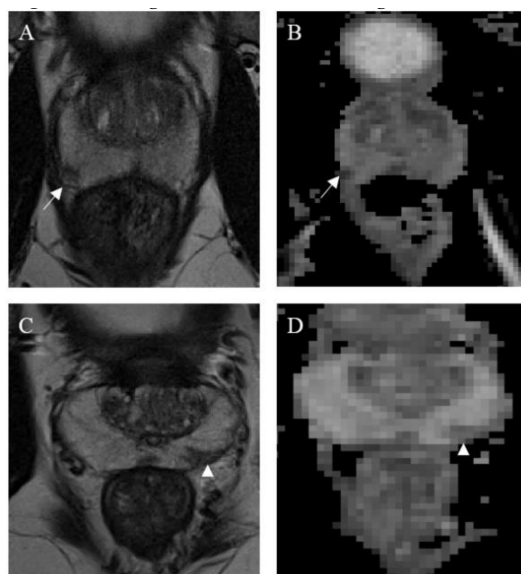


Figure 2. Magnetic resonance images from a clinically significant (A,B) and a non-clinically significant (C,D) prostate cancer lesion. Both are located in the peripheral zone; the clinically significant lesion, found in the right middle-posterior location (white arrows), appears as a nodular and well-defined area of low signal on both T2w (A) and ADC map (B). Conversely, the non-clinically significant lesion, found in the left middle-posterior location (white arrow heads), appears wedge-shaped, with no bulging on T2w (C) and only is mildly hypointense on the ADC map (D). Both lesions were deemed worthy of targeted biopsy.

2.2. Radiomics Features Extraction

Images underwent a preprocessing stage before feature extraction, including resampling to isotropic voxel, the normalization of pixel intensity values and discretization [20]. A freely accessible software (PyRadiomics, v 3.0) was used for image pre-processing and feature extraction [21]. Z-score normalization was paired with scaling by a factor of 100 and a grey level value shift of +300, resulting in a final expected intensity range of 0–600. Discretization prior to first-order feature extraction was implemented using a fixed bin width of 5. Laplacian Gaussian filtering (sigma values= 1, 2, 3, 4, 5) and wavelet decomposition (all high- and low-pass filter combinations along the three axes) were applied, in addition to the original images. These settings were based on recommendations from the software developers and previous experiences in the literature [22]. Feature stability was tested for multiple segmentations on a random sample of 30 lesions (in total, masks from three operators were used), by calculating intraclass correlation coefficient and using a cut-off of 0.75. Low variance features were then excluded using a variance threshold of 0.01. Highly intercorrelated features (Pearson pairwise correlation > 0.8) were discarded, leaving a final number of 60 stable, informative features. Radiomic features are subsequently extracted to the prostate segmentation to simplify the detection, similarly to a previous published study [23]. A detailed description of the extracted radiomic features is available in the official PyRadiomics documentation (<https://pyradiomics.readthedocs.io/en/latest/features.html>, accessed on 1 July 2020).

2.3. Statistical Analysis

An inferential statistical analysis was performed by means of Levene's test to assess the equality of variances for each feature of the two classes. Moreover, an unpaired *t*-test was carried out to assess the differences in the mean values for each feature between the two classes. Both statistic tests were implemented assuming a two-tailed distribution and a confidence level equal to 95% (definition of statistical significance: *p*-value < 0.05). The main purpose of this analysis was to understand whether the radiomics features extracted from the images could distinguish the significance of the lesion.

SPSS Software for Statistics v. 25 was used to perform the statistical analysis.

2.4. Machine Learning

Afterward, a ML analysis was conducted to evaluate the predictive power of the extracted features in classifying significant and non-significant lesions.

The following ML algorithms were implemented.

Decision Tree (DT) is based on an ordinary tree structure, which is made-up of a root, nodes, branches and leaves [24]. A DT starts from the root, then moves downward. The node from which the tree starts is named the root node, while the node where the chain ends is named the leaf node. Two or more branches can be extended from each internal node; in this case, it is not a leaf node. A node represents a certain feature while the branches represent a range of values [8]. J48 DT, which uses the C4.5 algorithm [25], was considered in the present work.

Random Forest (RF) [26], considered a classification task, is an ensemble of unpruned classification trees generated from the random selection of training set instances. Random features are selected in the induction process. A prediction is made by aggregating the ensemble predictions using the majority vote strategy. The Information Gain Ratio was used as a split criterion.

Gradient Boosted Tree (GBT) builds one DT at a time to fit the residual of the trees that precede it [27]. In the case of a binary classification, as in this study, a scalar score function is formed to distinguish the two classes. Given the training data and the classes related to each training instances, the goal of GBT is to choose a classification function that minimizes the aggregation of some specified loss function [27].

Ada Boost (ADA-B) is part of the boosting algorithms, in which several individual classifiers, DT in the case under study, are produced iteratively, and each classifier tries to accurately classify the training data [28]. The classifier uses an adaptive resampling strategy to choose the training samples. Each iteration assigns a weight to the dataset so that the next integration concentrates on reweighted datasets that were previously misclassified. The final classifier is a weighted sum of the ensemble predictions [29]. The advantage of the ADA-B algorithm is significant for solving several issues, including two-class problems, as in the case under study.

Naïve Bayes (NB) is based on the assumption that features are independent within a class in order to simplify the learning process [30]. Although this is an unrealistic assumption, NB competes well with more sophisticated classifiers [31], finding concrete applications in several scenarios including medical diagnosis [32].

K Nearest Neighbor (KNN) requires, in addition to training data, a fixed k value to search the k-nearest data based on distance computation. If the k found instances of different class labels, the classifier predicts that the class of the unknown example would be the same as the majority class [33]. Different distance metrics have been proposed in the scientific literature; for our purpose, we considered the Euclidian distance.

Two workflows of analyses were carried out using two different validation strategies for all the ML algorithms.

The first analysis used a 10-fold cross-validation to validate the predictive models by including all 60 radiomics features [34].

The second analysis used a hold-out validation; the dataset was divided into two non-overlapping parts and these two parts were used for training (70%) and testing (30%), respectively. This validation allows to avoid the problem of overfitting that is present in a re-substitution validation to be removed [35]. This analysis was performed using a feature selection method by means of a wrapper method based on backward feature elimination [36]. The usefulness of this method relies on the elimination of useless features and the building of a more reliable model based on a reduced set of features.

The main difference between the models was the presence of a feature selection step.

The performance of the proposed predictive models was evaluated through the following evaluation metrics: accuracy, sensitivity, specificity, area under the receiver operating

characteristic curve (AUC-ROC) [37] and accuracy max, computed as the maximum value among the accuracies obtained in the ten cycles of 10-fold cross-validation.

ML algorithms were implemented through the artificial intelligence platform Knime Analytics Platform (version 3.7.1), which is increasingly diffused in the scientific literature [38–40] and has achieved an interesting performance when compared with other platforms and programming languages.

3. Results

The following subsections show the univariate statistical analysis results for the radiomics features and the ML analyses.

Altogether, 466 out of the 2576 features were considered stable after the inter-observer intra-class correlation analysis. An additional reduction was performed by removing zero variant features (n = 54 removed). Then, 352 out of the remaining 412 were excluded due to their high pairwise correlation, leaving 60 radiomic parameters in the dataset.

3.1. Statistical Analysis

Levene’s test was employed to verify the equality of variances, and then the univariate statistical analysis was performed through a *t*-test. Table 1 shows the descriptive statistics and the *p*-value of the *t*-test for all the radiomic features.

Table 1. Descriptive statistics for the 60 radiomic variables and comparison between significant (0) and non-significant (1) lesions. Feature names are structured as follows: “MRI sequence”, “image type (original or filtered)”, “feature class”, “feature name”, each separated by an underscore.

Variable	Class	Mean	±	std. dev.	<i>t</i> -Test <i>p</i> -Value
t2_original_shape_MeshVolume	0	0.11	±	0.13	0.003 **
	1	0.18	±	0.19	
t2_original_firstorder_10Percentile	0	0.47	±	0.20	0.002 **
	1	0.40	±	0.15	
t2_original_glcM_Imc2	0	0.65	±	0.26	0.007 **
	1	0.56	±	0.26	
t2_original_glcM_Idm	0	0.59	±	0.16	0.098
	1	0.62	±	0.15	
t2_logsigma10mm3D_firstorder_90Percentile	0	0.24	±	0.15	0.119
	1	0.21	±	0.13	
t2_logsigma10mm3D_ngtdm_Busyness	0	0.15	±	0.15	0.042 *
	1	0.20	±	0.19	
t2_logsigma10mm3D_gldm_DependenceVariance	0	0.37	±	0.21	0.001 ***
	1	0.46	±	0.21	
t2_logsigma20mm3D_firstorder_90Percentile	0	0.44	±	0.19	0.874
	1	0.44	±	0.15	
t2_logsigma20mm3D_glcM_DifferenceVariance	0	0.19	±	0.16	0.506
	1	0.17	±	0.15	
t2_logsigma20mm3D_glszm_LargeAreaLowGrayLevelEmphasis	0	4.76	±	1.17	0.345
	1	6.16	±	9.36	
t2_logsigma30mm3D_glcM_Contrast	0	0.15	±	0.13	0.112
	1	0.12	±	0.11	
t2_logsigma30mm3D_glrIm_LongRunEmphasis	0	0.25	±	0.17	0.015 *
	1	0.31	±	0.19	
t2_logsigma30mm3D_ngtdm_Busyness	0	0.15	±	0.15	0.181
	1	0.17	±	0.13	
t2_logsigma40mm3D_firstorder_10Percentile	0	0.63	±	0.17	0.072
	1	0.66	±	0.12	
t2_logsigma40mm3D_firstorder_90Percentile	0	0.44	±	0.17	0.033 *
	1	0.49	±	0.14	
t2_logsigma40mm3D_firstorder_InterquartileRange	0	0.38	±	0.17	0.571
	1	0.39	±	0.19	
t2_logsigma40mm3D_glcM_Idm	0	0.50	±	0.17	0.054
	1	0.54	±	0.16	
t2_logsigma40mm3D_glcM_InverseVariance	0	0.85	±	0.11	0.325
	1	0.86	±	0.08	
t2_logsigma40mm3D_glszm_SizeZoneNonUniformity	0	0.07	±	0.10	0.064
	1	0.11	±	0.14	

Table 1. Cont.

Variable	Class	Mean	±	std. dev.	t-Test p-Value
t2_logsigma50mm3D_firstorder_Minimum	0	0.58	±	0.17	0.708
	1	0.57	±	0.14	
t2_logsigma50mm3D_firstorder_Variance	0	0.17	±	0.15	0.026 *
	1	0.22	±	0.17	
t2_logsigma50mm3D_glcmm_Autocorrelation	0	0.16	±	0.15	0.006 **
	1	0.22	±	0.16	
t2_logsigma50mm3D_glcmm_Contrast	0	0.11	±	0.11	0.465
	1	0.10	±	0.82	
t2_logsigma50mm3D_glrmm_LongRunEmphasis	0	0.26	±	0.16	0.413
	1	0.28	±	0.18	
t2_logsigma50mm3D_glszm_LargeAreaEmphasis	0	5.40	±	1.16	0.172
	1	7.50	±	1.13	
t2_logsigma50mm3D_gldm_LargeDependenceHighGrayLevelEmphasis	0	0.09	±	0.10	0.003 **
	1	0.13	±	0.11	
t2_waveletLLH_glcmm_JointEnergy	0	0.28	±	0.19	0.927
	1	0.28	±	0.15	
t2_waveletLHL_firstorder_90Percentile	0	0.21	±	0.15	0.213
	1	0.19	±	0.12	
t2_waveletLHH_glcmm_JointEnergy	0	0.45	±	0.21	0.305
	1	0.47	±	0.17	
t2_waveletHLL_glrmm_LongRunEmphasis	0	0.36	±	0.16	0.217
	1	0.38	±	0.17	
t2_waveletHLL_ngtdm_Busyness	0	0.14	±	0.14	0.059
	1	0.17	±	0.17	
t2_waveletHHL_firstorder_Variance	0	0.24	±	0.14	0.244
	1	0.22	±	0.11	
t2_waveletHHL_glszm_LargeAreaLowGrayLevelEmphasis	0	0.40	±	0.08	0.386
	1	0.06	±	0.15	
t2_waveletHHL_ngtdm_Busyness	0	0.17	±	0.17	0.047 *
	1	0.22	±	0.19	
t2_waveletLLL_firstorder_Energy	0	0.10	±	0.14	0.131
	1	0.13	±	0.15	
adc_original_firstorder_10Percentile	0	0.58	±	0.16	0.001 ***
	1	0.47	±	0.18	
adc_original_glrmm_LongRunEmphasis	0	0.31	±	0.18	0.612
	1	0.32	±	0.17	
adc_logsigma10mm3D_glcmm_Contrast	0	0.10	±	0.11	0.718
	1	0.09	±	0.09	
adc_logsigma10mm3D_glcmm_Idm	0	0.54	±	0.21	0.265
	1	0.57	±	0.19	
adc_logsigma10mm3D_ngtdm_Strength	0	0.11	±	0.15	0.070
	1	0.09	±	0.09	
adc_logsigma30mm3D_firstorder_90Percentile	0	0.11	±	0.14	0.001 ***
	1	0.09	±	0.09	
adc_logsigma30mm3D_glcmm_DifferenceAverage	0	0.35	±	0.19	0.163
	1	0.38	±	0.16	
adc_logsigma30mm3D_glrmm_LongRunEmphasis	0	0.26	±	0.16	0.140
	1	0.23	±	0.12	
adc_logsigma30mm3D_glszm_GrayLevelNonUniformity	0	0.16	±	0.13	0.001 ***
	1	0.25	±	0.21	
adc_logsigma40mm3D_glcmm_InverseVariance	0	0.66	±	0.16	0.694
	1	0.65	±	0.16	
adc_logsigma40mm3D_glszm_LargeAreaHighGrayLevelEmphasis	0	4.28×10^2	±	1.22×10^2	0.472
	1	5.39×10^2	±	9.64×10^2	
adc_logsigma50mm3D_firstorder_10Percentile	0	0.50	±	0.18	0.100
	1	0.54	±	0.19	
adc_logsigma50mm3D_glrmm_RunPercentage	0	0.58	±	0.15	0.051
	1	0.62	±	0.12	
adc_logsigma50mm3D_glszm_ZoneVariance	0	6.50×10^2	±	1.39×10^1	0.720
	1	5.99×10^2	±	8.49×10^2	
adc_waveletLLH_glcmm_JointEnergy	0	0.27	±	0.16	0.572
	1	0.29	±	0.18	
adc_waveletLLH_glrmm_LongRunEmphasis	0	0.19	±	0.12	0.015 *
	1	0.24	±	0.16	
adc_waveletLHL_firstorder_90Percentile	0	0.36	±	0.14	0.926
	1	0.36	±	0.13	
adc_waveletLHL_firstorder_Kurtosis	0	0.10	±	0.10	0.004 **
	1	0.15	±	0.14	

Table 1. Cont.

Variable	Class	Mean	±	std. dev.	t-Test p-Value
adc_waveletHLL_firstorder_90Percentile	0	0.22	±	0.12	0.060
	1	0.25	±	0.12	
adc_waveletHLL_glcm_Imc2	0	0.46	±	0.23	0.984
	1	0.46	±	0.20	
adc_waveletHLL_glcm_Idm	0	0.53	±	0.16	0.743
	1	0.53	±	0.14	
adc_waveletHLL_glrmlm_RunVariance	0	0.28	±	0.15	0.05 *
	1	0.32	±	0.18	
adc_waveletHHL_glcm_Contrast	0	0.07	±	0.12	0.834
	1	0.08	±	0.10	
adc_waveletHHL_glszm_LargeAreaEmphasis	0	4.70×10^2	±	8.20×10^2	0.064
	1	8.26×10^2	±	1.58×10^1	
adc_waveletLLL_glcm_Imc2	0	0.88	±	0.17	0.477
	1	0.89	±	0.12	

* = $0.01 < p < 0.05$; ** = $0.001 < p < 0.01$; *** = $p < 0.001$.

The statistical analysis showed that 16 features, out of a total of 60, were useful to distinguish a significant from a non-significant lesion (p -value < 0.05).

3.2. Machine Learning Analysis

The ML analysis was performed twice.

First, all 60 radiomics features were given as input to the six algorithms and the 10-fold cross-validation was employed to compute the evaluation metrics; the results of this analysis are shown in Table 2.

Table 2. Evaluation metrics (%) of the models computed on all 60 features with the 10-fold cross-validation.

Algorithms	Accuracy	Accuracy Max	Sensitivity	Specificity	AUCROC
J48	74.2	83.3	35.5	87.4	0.567
ADA-B	74.6	86.7	42.1	85.7	0.720
RF	77.9	83.3	48.7	87.9	0.713
GBT	74.9	86.2	34.2	88.8	0.682
NB	68.9	80.0	56.6	73.1	0.650
KNN	73.2	76.7	18.4	91.9	0.643

From this, the following results can be seen: the best algorithms were RF, according to their accuracy (77.9%), NB, which achieved the highest sensitivity (56.6%), KNN, with the highest specificity (91.9%), and ADA-B, which obtained the best AUCROC (0.720) and the highest accuracy max (86.7%).

Then, the dataset was divided, with 70% in the training set, and 30% in the test set, as per hold-out cross-validation. The training set was used to perform backward feature elimination, starting from all 60 features, and a set of variables was chosen for each algorithm. Finally, the evaluation metrics were computed on the test set for each algorithm by implementing 10-fold cross-validation. The results are shown in Table 3.

From this, the following results can be seen: the best algorithms were RF, again according to accuracy (82.1%) and also regarding specificity (91.0%), J48, which achieved the highest sensitivity (56.5%), and GBT, which obtained the best AUCROC (0.774). GBT and J48 achieved the highest accuracy max during the 10-fold cross-validation (100%). The application of backward feature elimination on the best algorithm, RF, made the algorithm select 39 features, which are shown in the Appendix A.

Table 3. Evaluation metrics (%) of the models computed using hold-out, backward feature elimination, 10-fold cross validation.

Algorithms	Number of Features	Accuracy	Accuracy Max	Sensitivity	Specificity	AUCROC
J48	16	82.2	100	56.5	91.0	0.635
ADA-B	14	81.1	88.9	52.2	91.0	0.708
RF	39	82.2	88.9	39.1	97.0	0.730
GBT	50	76.7	100	43.5	88.1	0.774
NB	15	70.0	88.9	21.7	86.6	0.546
KNN	25	74.4	88.9	30.4	89.6	0.676

4. Discussion and Conclusions

The present study describes 60 stable, uncorrelated and non-invariant radiomics features, extracted from MRI images, which previously underwent a quality assessment [16], and used to distinguish significant from non-significant prostate cancer lesions through an ML approach. Firstly, a univariate statistical analysis was performed to prove that these 60 features were useful in distinguishing the lesions by themselves (16 of them were revealed to be statistically significant). Secondly, J48, ADA-B, RF, GBT, NB and KNN were implemented twice: i) they were applied with a 10-fold cross-validation on all 60 features; ii) a different ML workflow was employed, including a backward feature elimination strategy to identify the best subset of features, maximizing the evaluation metrics (i.e., accuracy).

Several studies have used a similar approach, combining radiomics and ML, for the diagnosis and characterization of prostatic lesions, aiming to differentiate clinically significant from non-significant lesions, and thus to stratify patient's risk [13]. This differentiation is considered crucial in the management of prostate cancer patients for different causes: i) a growing number of prostate lesions, discovered through prostate-specific antigen (PSA) screening, are often clinically insignificant [41]; ii) in cases of clinically non-significant prostate cancer, the method of choice is active surveillance, whereas clinically significant lesions undergo surgical and medical treatment [42]; iii) then, the definition of clinically significant cancer becomes even more urgent [43].

A further analysis by subgroups showed that eight groups had used an ML approach while four used deep learning [44–47]. In the latter case, the used algorithms were convolutional neural network, artificial neural network and transfer deep learning with a pooled AUC-ROC of 0.78. Instead, in the former case, the ML algorithms used were NB, linear regression, RF, logistic regression, and support vector machine, with a pooled AUC-ROC of 0.90.

Another interesting finding is the variability of the sequences used; three studies [48–50] employed a similar approach to ours, relying on T2 and ADC acquisitions with a pooled AUC-ROC of 0.90. Abraham et al. [44] and Bonekamp et al [51] also associated DWI with a pooled AUC-ROC of 0.81, which presented a lower stability in the extracted radiomic features [52]. Similarly, the use of automated analysis on T1- and T2-w sequences—without the need for gadolinium-based contrast medium—was also recently described [53,54].

Dynamic contrast-enhanced sequences were combined with baseline T2 and ADC sequences in two studies [55,56], resulting in a pooled AUC-ROC of 0.85. In addition, other studies extracted radiomic features from images of advanced MRI sequences that are not normally used in prostate MRI protocols, limiting the resulting algorithm's clinical applicability [16].

Moreover, among the studies analyzed by Cuocolo et al., only five, like ours, started from a public archive, the Cancer Imaging Archive (<https://www.cancerimagingarchive.net/>, accessed on 1 July 2020) [44–48,57]. The others were based on data from single institutions, thus limiting the reproducibility and standardization of the algorithms used.

Of note, Papa et al. proposed a deep neural network architecture for classifying clinically significant prostate lesions of non-contrast-enhanced MRI images using Conditional Random Fields as a Recurrent Neural Network to enhance the classification

performance [58]; although high evaluation metrics were achieved in this research, the proposed scores were affected by a high level of variability.

However, the present investigation evaluated a public dataset for improving the consistency of our technique, whereas most of the published studies are based on data from a single institution [13].

In addition, combining ML algorithms and radiomics has several advantages and potentialities. Since conventional image interpretation is based on radiologists' experience, this technique could decrease inter-individual variability, as well as reporting time, leading to a potential benefit for less-experienced radiologists [59].

Moreover, the present paper demonstrated the usefulness of a ML and radiomics approach to images, which presents advantages, e.g., the non-necessity of a contrast medium. Therefore, the MRI acquisition protocol could be faster (by selecting only the most useful sequences) and cheaper, limiting the risk of possible side effects [60]. Indeed, the absence of a gadolinium-based contrast medium does not expose patients to different types of toxicities, such as nephrogenic systemic fibrosis, gadolinium brain accumulation and the invasiveness of intravenous access [61,62]. Moreover, an easier and faster protocol could also be more reproducible, allowing a better quality of images to be acquired in both local databases and public archives, in turn facilitating and implementing radiomic feature extraction and ML application. The present study has some limitations. Firstly, we did not demonstrate a potential association between the Gleason grading with clinical outcomes. Secondly, the performance of the used technique needs to be confirmed with further investigations. Thirdly, we cannot consider the histopathological variants.

In conclusion, the ML and radiomics approach, based on a public dataset, demonstrated a successfully discriminating, clinically significant prostate cancer. In the future, this radiomic signature could be interpreted as a "virtual biopsy", which could potentially help to reduce the number of invasive procedures that are currently performed, and also guide the management of patients.

Author Contributions: Conceptualization, G.S. and C.R.; methodology, L.D., G.C. and C.R.; software, L.D., G.C. and C.R.; validation, A.C., D.R.D.L., F.N. and G.S.; formal analysis, L.D., G.C. and C.R.; investigation, A.C., D.R.D.L., F.N. and G.S.; writing—original draft preparation, all authors; writing—review and editing, all authors; supervision, G.S. and C.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable for a study on a public dataset.

Informed Consent Statement: Not applicable for a study on a public dataset.

Data Availability Statement: Data for this study can be found at https://github.com/rcuocolo/PROSTATEx_masks and <https://wiki.cancerimagingarchive.net/display/Public/SPIE-AAPM-NCI+PROSTATEx+Challenges> (accessed on 1 July 2020).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Features selected by Random Forests through backward feature elimination:

1. t2_log-sigma-3-0-mm-3D_glcm_Contrast
2. t2_log-sigma-3-0-mm-3D_ngtdm_Busyness
3. t2_log-sigma-4-0-mm-3D_firstorder_10Percentile
4. t2_log-sigma-4-0-mm-3D_firstorder_90Percentile
5. t2_log-sigma-4-0-mm-3D_firstorder_InterquartileRange
6. t2_log-sigma-4-0-mm-3D_glcm_Idm
7. t2_log-sigma-4-0-mm-3D_glcm_InverseVariance
8. t2_log-sigma-5-0-mm-3D_firstorder_Minimum
9. t2_log-sigma-5-0-mm-3D_glcm_Contrast

10. t2_log-sigma-5-0-mm-3D_glszm_LargeAreaEmphasis
11. t2_log-sigma-5-0-mm-3D_gldm_LargeDependenceHighGrayLevelEmphasis
12. t2_wavelet-LLH_glcm_JointEnergy
13. t2_wavelet-LHL_firstorder_90Percentile
14. t2_wavelet-LHH_glcm_JointEnergy
15. t2_wavelet-HLL_glrlm_LongRunEmphasis
16. t2_wavelet-HHL_firstorder_Variance
17. t2_wavelet-HHL_glszm_LargeAreaLowGrayLevelEmphasis
18. t2_wavelet-HHL_ngtdm_Busyness
19. t2_wavelet-LLL_firstorder_Energy
20. adc_original_firstorder_10Percentile
21. adc_original_glrlm_LongRunEmphasis
22. adc_original_glszm_LargeAreaEmphasis
23. adc_log-sigma-1-0-mm-3D_glcm_Contrast
24. adc_log-sigma-1-0-mm-3D_glcm_Idm
25. adc_log-sigma-3-0-mm-3D_firstorder_90Percentile
26. adc_log-sigma-3-0-mm-3D_glrlm_LongRunEmphasis
27. adc_log-sigma-3-0-mm-3D_glszm_GrayLevelNonUniformity
28. adc_log-sigma-4-0-mm-3D_glcm_InverseVariance
29. adc_log-sigma-4-0-mm-3D_glszm_LargeAreaHighGrayLevelEmphasis
30. adc_log-sigma-5-0-mm-3D_glrlm_RunPercentage
31. adc_log-sigma-5-0-mm-3D_glszm_ZoneVariance
32. adc_wavelet-LHL_firstorder_Kurtosis
33. adc_wavelet-HLL_firstorder_90Percentile
34. adc_wavelet-HLL_glcm_Imc2
35. adc_wavelet-HLL_glcm_Idm
36. adc_wavelet-HLL_glrlm_RunVariance
37. adc_wavelet-HHL_glcm_Contrast
38. adc_wavelet-HHL_glszm_LargeAreaEmphasis
39. adc_wavelet-LLL_glcm_Imc2

References

1. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics. *CA Cancer J. Clin.* **2021**, *71*, 7–33. [\[CrossRef\]](#)
2. Matoso, A.; Epstein, J.I. Defining clinically significant prostate cancer on the basis of pathological findings. *Histopathology* **2019**, *74*, 135–145. [\[CrossRef\]](#)
3. Mottet, N.; Van den Bergh, R.C.N.; Briers, E.; Van den Broeck, T.; Cumberbatch, M.G.; De Santis, M.; Fanti, S.; Fossati, N.; Gandaglia, G.; Gillessen, S.; et al. EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer-2020 Update. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. *Eur. Urol.* **2021**, *79*, 243–262. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Gupta, R.T.; Mehta, K.A.; Turkbey, B.; Verma, S. PI-RADS: Past, present, and future. *J. Magn. Reson. Imaging* **2020**, *52*, 33–53. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Del Monte, M.; Leonardo, C.; Salvo, V.; Grompone, M.D.; Pecoraro, M.; Stanzione, A.; Campa, R.; Vullo, F.; Sciarra, A.; Catalano, V.; et al. MRI/US fusion-guided biopsy: Performing exclusively targeted biopsies for the early detection of prostate cancer. *Radiol. Med.* **2018**, *123*, 227–234. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Wei, C.G.; Zhang, Y.Y.; Pan, P.; Chen, T.; Yu, H.C.; Dai, G.C.; Tu, J.; Yang, S.; Zhao, W.L.; Shen, J.K. Diagnostic accuracy and interobserver agreement of PI-RADS version 2 and version 2.1 for the detection of transition zone prostate cancers. *Am. J. Roentgenol.* **2021**, *216*, 1247–1256. [\[CrossRef\]](#)
7. Sosnowski, R.; Zagrodzka, M.; Borkowski, T. The limitations of multiparametric magnetic resonance imaging also must be borne in mind. *Cent. Eur. J. Urol.* **2016**, *69*, 22–23. [\[CrossRef\]](#)
8. Stanzione, A.; Ricciardi, C.; Cuocolo, R.; Romeo, V.; Petrone, J.; Sarnataro, M.; Mainenti, P.P.; Improta, G.; De Rosa, F.; Insabato, L.; et al. MRI Radiomics for the Prediction of Fuhrman Grade in Clear Cell Renal Cell Carcinoma: A Machine Learning Exploratory Study. *J. Digit. Imaging* **2020**, *33*, 879–887. [\[CrossRef\]](#)
9. Carlo, R.; Renato, C.; Giuseppe, C.; Lorenzo, U.; Giovanni, I.; Domenico, S.; Valeria, R.; Elia, G.; Maria, C.L.; Mario, C. Distinguishing Functional from Non-functional Pituitary Macroadenomas with a Machine Learning Analysis. In Proceedings of the XV Mediterranean Conference on Medical and Biological Engineering and Computing—MEDICON 2019, IFMBE Proceedings, Coimbra, Portugal, 26–28 September 2019; Henriques, J., Neves, N., de Carvalho, P., Eds.; Springer: Cham, Switzerland, 2020; Volume 76. [\[CrossRef\]](#)

10. Romeo, V.; Cuocolo, R.; Ricciardi, C.; Ugga, L.; Coccozza, S.; Verde, F.; Stanzione, A.; Napolitano, V.; Russo, D.; Improta, G.; et al. Prediction of tumor grade and nodal status in oropharyngeal and oral cavity squamous-cell carcinoma using a radiomic approach. *Anticancer. Res.* **2020**, *40*, 271–280. [[CrossRef](#)]
11. Chaddad, A.; Kucharczyk, M.J.; Cheddad, A.; Clarke, S.E.; Hassan, L.; Ding, S.; Rathore, S.; Zhang, M.; Katib, Y.; Bahoric, B.; et al. Magnetic resonance imaging based radiomic models of prostate cancer: A narrative review. *Cancers* **2021**, *13*, 552. [[CrossRef](#)]
12. Stanzione, A.; Gambardella, M.; Cuocolo, R.; Ponsiglione, A.; Romeo, V.; Imbriaco, M. Prostate MRI radiomics: A systematic review and radiomic quality score assessment. *Eur. J. Radiol.* **2020**, *129*, 109095. [[CrossRef](#)]
13. Cuocolo, R.; Cipullo, M.B.; Stanzione, A.; Romeo, V.; Green, R.; Cantoni, V.; Ponsiglione, A.; Ugga, L.; Imbriaco, M. Machine learning for the identification of clinically significant prostate cancer on MRI: A meta-analysis. *Eur. Radiol.* **2020**, *30*, 6877–6887. [[CrossRef](#)] [[PubMed](#)]
14. Cutaia, G.; La Tona, G.; Comelli, A.; Vernuccio, F.; Agnello, F.; Gagliardo, C.; Salvaggio, L.; Quartuccio, N.; Sturiale, L.; Stefano, A.; et al. Radiomics and Prostate MRI: Current Role and Future Applications. *J. Imaging* **2021**, *7*, 34. [[CrossRef](#)] [[PubMed](#)]
15. Spadarella, G.; Calareso, G.; Garanzini, E.; Ugga, L.; Cuocolo, A.; Cuocolo, R. MRI based radiomics in nasopharyngeal cancer: Systematic review and perspectives using radiomic quality score (RQS) assessment. *Eur. J. Radiol.* **2021**, *140*, 109744. [[CrossRef](#)] [[PubMed](#)]
16. Cuocolo, R.; Stanzione, A.; Castaldo, A.; De Lucia, D.R.; Imbriaco, M. Quality control and whole-gland, zonal and lesion annotations for the PROSTATEx challenge public dataset. *Eur. J. Radiol.* **2021**, *138*, 109647. [[CrossRef](#)]
17. Litjens, G.; Debats, O.; Barentsz, J.; Karssemeijer, N.; Huisman, H. Computer-aided detection of prostate cancer in MRI. *IEEE Trans. Med Imaging* **2014**, *33*, 1083–1092. [[CrossRef](#)] [[PubMed](#)]
18. Steiger, P.; Thoeny, H.C. Prostate MRI based on PI-RADS version 2: How we review and report. *Cancer Imaging* **2016**, *16*, 9. [[CrossRef](#)] [[PubMed](#)]
19. Litjens, G.; Debats, O.; van de Ven, W.; Karssemeijer, N.; Huisman, H. A pattern recognition approach to zonal segmentation of the prostate on MRI. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Nice, France, 1–5 October 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 413–420.
20. Stanzione, A.; Cuocolo, R.; Del Grosso, R.; Nardiello, A.; Romeo, V.; Travaglino, A.; Raffone, A.; Bifulco, G.; Zullo, F.; Insabato, L.; et al. Deep myometrial infiltration of endometrial cancer on MRI: A radiomics-powered machine learning pilot study. *Acad. Radiol.* **2021**, *28*, 737–744. [[CrossRef](#)]
21. Van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.; Fillion-Robin, J.C.; Pieper, S.; Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)]
22. Cuocolo, R.; Stanzione, A.; Faletti, R.; Gatti, M.; Callaris, G.; Fornari, A.; Gentile, F.; Motta, A.; Dell’Aversana, S.; Creta, M.; et al. MRI index lesion radiomics and machine learning for detection of extraprostatic extension of disease: A multicenter study. *Eur. Radiol.* **2021**, *31*, 7575–7583. [[CrossRef](#)]
23. Klein, S.; Van Der Heide, U.A.; Lips, I.M.; Van Vulpen, M.; Staring, M.; Pluim, J.P. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med. Phys.* **2008**, *35*, 1407–1417. [[CrossRef](#)]
24. Ali, J.; Khan, R.; Ahmad, N.; Maqsood, I. Random forests and decision trees. *Int. J. Comput. Sci. Issues (IJCSI)* **2012**, *9*, 272.
25. Quinlan, J.R. *C4. 5: Programs for Machine Learning*; Elsevier: Amsterdam, The Netherlands, 1994.
26. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
27. Si, S.; Zhang, H.; Keerthi, S.S.; Mahajan, D.; Dhillon, I.S.; Hsieh, C.J. Gradient boosted decision trees for high dimensional sparse output. In Proceedings of the 34th International Conference on Machine Learning, PMLR, Sydney, Australia, 6–11 August 2017; pp. 3182–3190.
28. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [[CrossRef](#)]
29. Hong, H.; Liu, J.; Bui, D.T.; Pradhan, B.; Acharya, T.D.; Pham, B.T.; Zhu, A.X.; Chen, W.; Ahmad, B.B. Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). *Catena* **2018**, *163*, 399–413. [[CrossRef](#)]
30. Rish, I. An empirical study of the naive Bayes classifier. In *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*; IBM: New York, NY, USA, 2001; Volume 3, pp. 41–46.
31. Langley, P.; Iba, W.; Thompson, K. An analysis of Bayesian classifiers. In Proceedings of the Tenth National Conference on Artificial Intelligence, AAAI’92, San Jose, CA, USA, 12–16 July 1992; AAAI: Menlo Park, CA, USA, 1992; Volume 90, pp. 223–228.
32. Mitchell, T.M. *Machine Learning*; McGraw-Hill Education: New York, NY, USA, 1997.
33. Chomboon, K.; Chujai, P.; Teerassamee, P.; Kerdprasop, K.; Kerdprasop, N. An empirical study of distance metrics for k-nearest neighbor algorithm. In Proceedings of the 3rd International Conference on Industrial Application Engineering, Kitakyushu, Japan, 28–31 March 2015; pp. 280–285.
34. Anguita, D.; Ghelardoni, L.; Ghio, A.; Oneto, L.; Ridella, S. The ‘K’ in K-fold Cross Validation. In Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), Bruges, Belgium, 25–27 April 2012; pp. 441–446.
35. Yadav, S.; Shukla, S. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In Proceedings of the 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 27–28 February 2016; pp. 78–83.

36. Kostrzewa, D.; Brzeski, R. The data dimensionality reduction in the classification process through greedy backward feature elimination. In Proceedings of the International Conference on Man–Machine Interactions, Kraków, Poland, 3–6 October 2017; Springer: Cham, Switzerland, 2017; pp. 397–407.
37. Hossin, M.; Sulaiman, M.N. A review on evaluation metrics for data classification evaluations. *Int. J. Data Min. Knowl. Manag. Process.* **2015**, *5*, 1–11.
38. Scrutinio, D.; Ricciardi, C.; Donisi, L.; Losavio, E.; Battista, P.; Guida, P.; Cesarelli, M.; Pagano, G.; D’Addio, G. Machine learning to predict mortality after rehabilitation among patients with severe stroke. *Sci. Rep.* **2020**, *10*, 20127. [[CrossRef](#)]
39. O’Hagan, S.; Kell, D.B. Software review: The KNIME workflow environment and its applications in Genetic Programming and machine learning. *Genet. Program. Evolvable Mach.* **2015**, *16*, 387–391. [[CrossRef](#)]
40. Donisi, L.; Cesarelli, G.; Coccia, A.; Panigazzi, M.; Capodaglio, E.M.; D’Addio, G. Work-Related Risk Assessment According to the Revised NIOSH Lifting Equation: A Preliminary Study Using a Wearable Inertial Sensor and Machine Learning. *Sensors* **2021**, *21*, 2593. [[CrossRef](#)] [[PubMed](#)]
41. Crawford, E.D.; Grubb, R., 3rd; Black, A.; Andriole, G.L., Jr.; Chen, M.H.; Izmirlian, G.; Berg, C.D.; D’Amico, A.V. Comorbidity and mortality results from a randomized prostate cancer screening trial. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2011**, *29*, 355–361. [[CrossRef](#)] [[PubMed](#)]
42. Tosoian, J.J.; Carter, H.B.; Lepor, A.; Loeb, S. Active surveillance for prostate cancer: Current evidence and contemporary state of practice. *Nat. Rev. Urol.* **2016**, *13*, 205–215. [[CrossRef](#)]
43. Edmund, L.; Rotker, K.L.; Lakis, N.S.; Brito, J.M., III; Lepe, M.; Lombardo, K.A.; Renzulli, J.F., II; Matoso, A. Upgrading and upstaging at radical prostatectomy in the post-prostate-specific antigen screening era: An effect of delayed diagnosis or a shift in patient selection? *Hum. Pathol.* **2017**, *59*, 87–93. [[CrossRef](#)]
44. Abraham, B.; Nair, M.S. Computer-aided grading of prostate cancer from MRI images using convolutional neural networks. *J. Intell. Fuzzy Syst.* **2019**, *36*, 2015–2024. [[CrossRef](#)]
45. Le, M.H.; Chen, J.; Wang, L.; Wang, Z.; Liu, W.; Cheng, K.T.T.; Yang, X. Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks. *Phys. Med. Biol.* **2017**, *62*, 6497–6514. [[CrossRef](#)]
46. Sobiecki, P.; Życka-Malesa, D.; Mykhalevych, I.; Sklinda, K.; Przelaskowski, A. MRI imaging texture features in prostate lesions classification. In *EMBECE & NBC 2017. EMBEC 2017, NBC 2017, IFMBE Proceedings*; Eskola, H., Väisänen, O., Viik, J., Hyttinen, J., Eds.; Springer: Singapore, 2018; Volume 65, pp. 827–830.
47. Zhong, X.; Cao, R.; Shakeri, S.; Scalzo, F.; Lee, Y.; Enzmann, D.R.; Wu, H.H.; Raman, S.S.; Sung, K. Deep transfer learning-based prostate cancer classification using 3 Tesla multi-parametric MRI. *Abdom. Radiol.* **2019**, *44*, 2030–2039. [[CrossRef](#)]
48. Chaddad, A.; Niazi, T.; Probst, S.; Bladou, F.; Anidjar, M.; Bahoric, B. Predicting Gleason score of prostate cancer patients using radiomic analysis. *Front. Oncol.* **2018**, *8*, 630. [[CrossRef](#)] [[PubMed](#)]
49. Chen, T.; Li, M.; Gu, Y.; Zhang, Y.; Yang, S.; Wei, C.; Wu, J.; Li, X.; Zhao, W.; Shen, J. Prostate cancer differentiation and aggressiveness: Assessment with a radiomic-based model vs. PI-RADS v2. *J. Magn. Reson. Imaging* **2019**, *49*, 875–884. [[CrossRef](#)] [[PubMed](#)]
50. Fehr, D.; Veeraraghavan, H.; Wibmer, A.; Gondo, T.; Matsumoto, K.; Vargas, H.A.; Sala, E.; Hricak, H.; Deasy, J.O. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E6265–E6273. [[CrossRef](#)]
51. Bonekamp, D.; Kohl, S.; Wiesenfarth, M.; Schelb, P.; Radtke, J.P.; Götz, M.; Kickingereder, P.; Yaqubi, K.; Hitthaler, B.; Gähler, N.; et al. Radiomic machine learning for characterization of prostate lesions with MRI: Comparison to ADC values. *Radiology* **2018**, *289*, 128–137. [[CrossRef](#)] [[PubMed](#)]
52. Peerlings, J.; Woodruff, H.C.; Winfield, J.M.; Ibrahim, A.; Van Beers, B.E.; Heerschap, A.; Jackson, A.; Wildberger, J.E.; Mottaghy, F.M.; DeSouza, N.M.; et al. Stability of radiomics features in apparent diffusion coefficient maps from a multi-centre test-retest trial. *Sci. Rep.* **2019**, *9*, 4800. [[CrossRef](#)]
53. Rundo, L.; Militello, C.; Russo, G.; Garufi, A.; Vitabile, S.; Gilardi, M.C.; Mauri, G. Automated prostate gland segmentation based on an unsupervised fuzzy C-means clustering technique using multispectral T1w and T2w MR imaging. *Information* **2017**, *8*, 49. [[CrossRef](#)]
54. Wang, Z.; Liu, C.; Cheng, D.; Wang, L.; Yang, X.; Cheng, K.T. Automated detection of clinically significant prostate cancer in mp-MRI images based on an end-to-end deep neural network. *IEEE Trans. Med. Imaging* **2018**, *37*, 1127–1139. [[CrossRef](#)]
55. Antonelli, M.; Johnston, E.W.; Dikaios, N.; Cheung, K.K.; Sidhu, H.S.; Appayya, M.B.; Giganti, F.; Simmons, L.A.; Freeman, A.; Allen, C.; et al. Machine learning classifiers can predict Gleason pattern 4 prostate cancer with greater accuracy than experienced radiologists. *Eur. Radiol.* **2019**, *29*, 4754–4764. [[CrossRef](#)]
56. Dikaios, N.; Alkalbani, J.; Abd-Alazeez, M.; Sidhu, H.S.; Kirkham, A.; Ahmed, H.U.; Emberton, M.; Freeman, A.; Halligan, S.; Taylor, S.; et al. Zone-specific logistic regression models improve classification of prostate cancer on multi-parametric MRI. *Eur. Radiol.* **2015**, *25*, 2727–2737. [[CrossRef](#)]
57. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The cancer imaging archive (TCIA): Maintaining and operating a public information repository. *J. Digit. Imaging* **2013**, *26*, 1045–1057. [[CrossRef](#)] [[PubMed](#)]
58. Lapa, P.; Castelli, M.; Gonçalves, I.; Sala, E.; Rundo, L. A Hybrid End-to-End Approach Integrating Conditional Random Fields into CNNs for Prostate Cancer Detection on MRI. *Appl. Sci.* **2020**, *10*, 338. [[CrossRef](#)]

59. Kang, H.C.; Jo, N.; Bamashmos, A.S.; Ahmed, M.; Sun, J.; Ward, J.F.; Choi, H. Accuracy of Prostate Magnetic Resonance Imaging: Reader Experience Matters. *Eur. Urol. Open Sci.* **2021**, *27*, 53–60. [[CrossRef](#)] [[PubMed](#)]
60. Cuocolo, R.; Verde, F.; Ponsiglione, A.; Romeo, V.; Petretta, M.; Imbriaco, M.; Stanzione, A. Clinically significant prostate cancer detection with biparametric MRI: A systematic review and meta-analysis. *Am. J. Roentgenol.* **2021**, *216*, 608–621. [[CrossRef](#)] [[PubMed](#)]
61. Pasquini, L.; Napolitano, A.; Visconti, E.; Longo, D.; Romano, A.; Tomà, P.; Espagnet, M.C.R. Gadolinium-based contrast agent-related toxicities. *CNS Drugs* **2018**, *32*, 229–240. [[CrossRef](#)]
62. Wallström, J.; Geterud, K.; Kohestani, K.; Maier, S.E.; Månsson, M.; Pihl, C.G.; Socratous, A.; Godtman, R.A.; Hellström, M.; Hugosson, J. Bi- or multiparametric MRI in a sequential screening program for prostate cancer with PSA followed by MRI? Results from the Göteborg prostate cancer screening 2 trial. *Eur. Radiol.* **2021**, 1–11. [[CrossRef](#)]