Optimizing Radiomics for Prostate Cancer Diagnosis: Feature Selection Strategies, Machine learning Classifiers, and MRI sequences

ELECTRONIC SUPPLEMENTARY MATERIAL

Supplementary Table S1. Configuration of PyRadiomics parameters used for radiomic feature extraction.

Pyradiomics Parameters	Parameters Values				
"imageType"	Original,				
	• Wavelet,				
	• Gradient,				
	• LoG with sigma: 2mm, 3mm, 4mm and 5mm.				
"featureClass"	• firstorder				
	• glcm				
	• glrlm				
	• glszm				
	• gldm				
	• shape				
"settings"	resampledPixelSpacing: [1, 1, 1]				
	normalize: True				
	normalizeScale: 100				
	padDistance: 5				
	binWidth:				
	ProstateNET:				
	o T2: 7				
	• ADC: 7				
	ProstateX2:				
	o T2:6				
Parameters other than those reporte	• ADC: 3				

Section	No	Item	Yes	No	n/a	Page
Title				•		
	1	Relevant title, specifying the radiomic methodology	V			1
Abstract						
	2	Structured summary with relevant information	\checkmark			1
Keywords			1		T	T
	3	Relevant keywords for radiomics	\checkmark			2
Introduction	T	1				1
	4	Scientific or clinical background	\checkmark			3
	5	Rationale for using a radiomic approach	\checkmark			3-4
	6	Study objective(s)	\checkmark			4
Method		1	1	1		1
Study Design	7	${\sf Adherencetoguidelinesorchecklists(e.g., CLEAR checklist)}$	\checkmark			9
	8	Ethical details (e.g., approval, consent, data protection)	V			Ethics declarations
	9	Sample size calculation		\checkmark		-
	10	Study nature (e.g., retrospective, prospective)	\checkmark			4
	11	Eligibility criteria	V			4-5
	12	Flowchart for technical pipeline	V			Fig. 1
Data	13	Data source (e.g., private, public)	V			4
	14	Data overlap	V			4-5
	15	Data split methodology	V			8
	16	Imaging protocol (i.e., image acquisition and processing)	V			5
	17	Definition of non-radiomic predictor variables		V		Beyond the intended focus of this study
	18	Definition of the reference standard (i.e., outcome variable)	V			5
Segmentatio n	19	Segmentation strategy	V			4-5
	20	Details of operators performing segmentation		V		Segmentatio ns performed by experts across 12 clinical centers.
Pre- processing	21	Image pre-processing details	V			5
-	22	Resampling method and its parameters	V			5
	23	Discretization method and its parameters	V			5

Supplementary Table S2. CheckList for EvaluAtion of Radiomics research (CLEAR checklist)

Insights Imaging (2024) Mylona E, Zaridis DI, Kalantzopoulos CN, et al.

	24	Image types (e.g., original, filtered, transformed)	\checkmark			5-6	
Feature extraction	25	Feature extraction method	V			5	
	26	Feature classes	V			5-6	
	27	Number of features	\checkmark			6	
	28	Default configuration statement for remaining parameters	V			Suppl. Table 1	
Data preparation	29	Handling of missing data			V	No missing data	
	30	Details of class imbalance	\checkmark			Table 1	
	31	Details of segmentation reliability analysis		V			
	32	Feature scaling details (e.g., normalization, standardization)	V			5	
	33	Dimension reduction details	\checkmark			6-8	
Modeling	34	Algorithm details	\checkmark			8	
	35	Training and tuning details	\checkmark			8	
	36	Handling of confounders	\checkmark			8	
	37	Model selection strategy	\checkmark			8	
Evaluation	38	Testing technique (e.g., internal, external)	\checkmark				
	39	Performance metrics and rationale for choosing	\checkmark			8-9	
	40	Uncertainty evaluation and measures (e.g., confidence intervals)	V			8	
	41	Statistical performance comparison (e.g., DeLong's test)	\checkmark			8	
	42	Comparison with non-radiomic and combined methods				Beyond the intended focus of this study	
	43	Interpretability and explainability methods		V		Beyond the intended focus of this study	
Results							
	44	Baseline demographic and clinical characteristics	\checkmark			Table 1	
	45	Flowchart for eligibility criteria			\checkmark		
	46	Feature statistics (e.g., reproducibility, feature selection)	V			Fig.6	
	47	Model performance evaluation	V			Tables 2-5, Fig. 2-5	
	48	Comparison with non-radiomic and combined approaches		V		Beyond the intended focus of this study	
Discussion		· · · · · · · · · · · · · · · · · · ·				· ·	
	49	Overview of important findings	V			12	
	50	Previous works with differences from the current study	\checkmark			13	
	51	Practical implications	\checkmark			14	

Insights Imaging (2024) Mylona E, Zaridis DI, Kalantzopoulos CN, et al.

	52	Strengths and limitations (e.g., bias and generalizability issues)	V			14
Open Science	e					
Data availability	53	Sharing images along with segmentation data [n/e]			V	Data privacy constraints
	54	Sharing radiomic feature data			V	Data privacy constraints
Code availability	55	Sharing pre-processing scripts or settings	V			Supp. Table 1
	56	Sharing source code for modeling		V		Available upon request
Model availability	57	Sharing final model files		V		Available upon request
	58	Sharing a ready-to-use system [n/e]		\checkmark		

Yes, details provided; No, details not provided; n/e, not essential; n/a, not applicable

Items/Condi			
tions	Definitions	Weights	Answers
Study Design			-
	Adherence to radiomics and/or machine learning-specific		
ltem#1	checklists or guidelines	0.0368	yes
	Eligibility criteria that describe a representative study		
Item#2	population	0.0735	yes
Item#3	High-quality reference standard with a clear definition	0.0919	yes
Imaging Data			
ltem#4	Multi-center	0.0438	yes
	Clinical translatability of the imaging data source for radiomics		
Item#5	analysis	0.0292	yes
ltem#6	Imaging protocol with acquisition parameters	0.0438	no
ltem#7	The interval between imaging used and reference standard	0.0292	yes
Segmentation			
Condition#1	Does the study include segmentation?		yes
Condition#2	Does the study include fully automated segmentation?		no
ltem#8	Transparent description of segmentation methodology	0.0337	no
Item#9	Formal evaluation of fully automated segmentation	0.0225	n/a
	Test set segmentation masks produced by a single reader or		
Item#10	automated tool	0.0112	no
Image Process	sing and Feature Extraction		
Condition#3	Does the study include hand-crafted feature extraction?		yes
	Appropriate use of image preprocessing techniques with		
ltem#11	transparent description	0.0622	yes
Item#12	Use of standardized feature extraction software	0.0311	yes
	Transparent reporting of feature extraction parameters,		
Item#13	otherwise providing a default configuration statement	0.0415	yes
Feature Proce	ssing		
Condition#4	Does the study include tabular data?		yes
Condition#5	Does the study include end-to-end deep learning?		no
Item#14	Removal of non-robust features	0.0200	yes
Item#15	Removal of redundant features	0.0200	yes
Item#16	Appropriateness of dimensionality compared to data size	0.0300	yes
Item#17	Robustness assessment of end-to-end deep learning pipelines	0.0200	n/a
Preparation fo		•	•
Item#18	Proper data partitioning process	0.0599	yes
Item#19	Handling of confounding factors	0.0300	yes
Metrics and C	omparison		
ltem#20	Use of appropriate performance evaluation metrics for task	0.0352	yes

Supplementary Table S3. METhodological RadiomICs Score (METRICS)

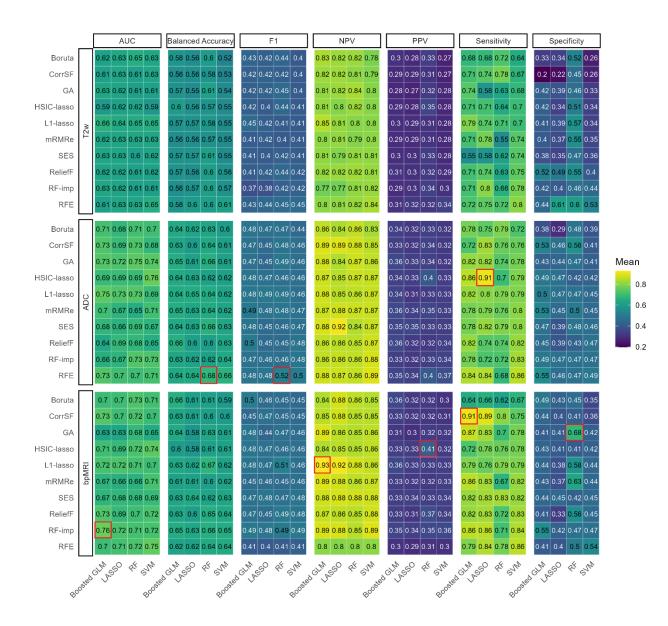
ltem#21	Consideration of uncertainty	0.0234	yes				
Item#22	Calibration assessment	0.0176	no				
Item#23	Use of uni-parametric imaging or proof of its inferiority	0.0117	yes				
	Comparison with a non-radiomic approach or proof of added						
Item#24	clinical value	0.0293	no				
Item#25	Comparison with simple or classical statistical models	0.0176	no				
Testing							
Item#26	Internal testing	0.0375	yes				
Item#27	External testing	0.0749	yes				
Open Science							
Item#28	Data availability	0.0075	no				
Item#29	Code availability	0.0075	no				
Item#30	Model availability	0.0075	no				
Total METRICS score: 81.7%							
Quality category *: Excellent							
* 0≤score<20%, "very low"; 20≤score<40%, "low"; 40≤score<60%, "moderate"; 60≤score<80%,							
"good"; 80≤score≤100%, "excellent" quality							

Supplementary Figure S1. Average models' performance in Setting 1 (across folds) for each feature selection method, ML classifier and MRI sequence. Red boxes indicate to the best performing combination of feature selection methods and ML classifiers for each metric.

		AUC	Balanced Accuracy	F1	NPV	PPV	Sensitivity	Specificity	
Boruta		0.69 0.67 0.66 0.69	0.64 0.62 0.62 0.62	0.75 0.75 0.73 0.73	0.4 0.39 0.38 0.37	0.82 0.81 0.82 0.81	0.69 0.69 0.66 0.66	0.58 0.56 0.59 0.58	
CorrSF		0.66 0.64 0.67 0.65	0.61 0.59 0.6 0.6	0.75 0.72 0.72 0.7	0.38 0.35 0.36 0.35	0.8 0.8 0.8 0.8	0.7 0.66 0.66 0.62	0.51 0.53 0.54 0.58	
GA		0.66 0.66 0.67 0.65	0.6 0.62 0.63 0.62	0.73 0.74 0.71 0.73	0.36 0.38 0.37 0.38	0.8 0.81 0.82 0.82	0.67 0.68 0.64 0.67	0.53 0.55 0.62 0.58	
HSIC-lasso		0.68 0.66 0.68 0.64	0.62 0.6 0.63 0.6	0.74 0.73 0.76 0.7	0.38 0.36 0.4 0.35	0.81 0.8 0.82 0.81	0.68 0.68 0.7 0.62	0.57 0.52 0.57 0.58	
L1-lasso	N	0.7 0.68 0.67 0.68	0.64 0.63 0.62 0.65	0.77 0.77 <mark>0.71</mark> 0.74	0.41 0.4 0.37 0.41	0.82 0.81 0.82 0.83	0.72 0.72 0.62 0.67	0.57 0.54 0.62 0.64	
mRMRe	T2w	0.61 0.62 0.64 0.63	0.57 0.58 0.6 0.58	0.64 0.63 0.69 0.65	0.31 0.32 0.34 0.32	0.79 0.8 0.8 0.8	0.54 0.52 0.61 0.55	0.61 0.64 0.58 0.61	
SES		0.68 0.67 0.66 0.67	0.62 0.62 0.63 0.63	0.73 0.75 0.74 0.74	0.37 0.38 0.39 0.39	0.81 0.81 0.82 0.82	0.67 0.7 0.67 0.68	0.57 0.54 0.6 0.58	
ReliefF		0.66 0.64 0.66 0.64	0.62 0.58 0.61 0.56	0.77 0.78 0.72 0.69	0.41 0.38 0.37 0.32	0.81 0.78 0.81 0.78	0.74 0.77 0.66 0.63	0.5 0.4 0.56 0.48	
RF-imp		0.69 0.68 0.68 0.69	0.62 0.62 0.63 0.66	0.75 0.75 <mark>0.72</mark> 0.75	0.38 0.38 0.38 0.41	0.81 0.81 <mark>0.83 0.83</mark>	0.69 0.7 0.64 0.69	0.55 0.54 0.63 0.62	
RFE		0.68 0.67 0.64 0.66	0.62 0.61 0.63 0.61	0.74 0.72 0.74 0.7	0.39 0.37 0.38 0.36	0.81 0.81 0.82 0.81	0.69 0.65 0.67 0.62	0.56 0.58 0.58 0.59	
Boruta		0.71 0.71 0.71 0.72	0.65 0.66 0.67 0.68	0.74 0.77 0.76 0.78	0.39 0.41 0.41 0.43	0.83 0.84 0.85 0.85	0.67 0.71 0.68 0.71	0.62 0.61 0.66 0.65	
CorrSF		0.73 0.71 0.7 0.74	0.67 0.66 0.65 0.68	0.75 0.75 0.74 <mark>0.78</mark>	0.41 0.41 0.39 0.44	0.85 0.84 0.84 0.85	0.67 0.68 0.67 0.72	0.66 0.63 0.63 0.65	
GA		0.74 0.71 0.68 0.71	0.68 0.66 0.64 0.69	0.78 0.76 0.75 0.77	0.45 0.41 0.39 0.44	0.85 0.84 0.83 0.86	0.71 0.69 0.68 0.7	0.65 0.64 0.6 0.68	Mean
HSIC-lasso		0.7 0.71 0.73 0.72	0.65 0.67 0.67 0.67	0.75 <mark>0.78</mark> 0.77 0.77	0.4 0.44 0.42 0.42	0.83 0.84 0.85 0.85	0.68 0.72 0.7 0.7	0.61 0.62 0.65 0.64	0.8
L1-lasso	S	0.7 0.72 0.7 0.72	0.66 0.65 0.66 0.65	0.77 <mark>0.78</mark> 0.76 0.76	0.42 0.43 0.41 0.4	0.84 0.83 0.84 0.83	0.72 0.73 0.7 0.7	0.6 0.58 0.62 0.6	0.7
mRMRe	AD	0.73 0.74 0.71 <mark>0.76</mark>	0.68 0.69 0.66 0.71	0.74 0.77 0.75 0.78	0.42 0.43 0.41 0.46	0.86 0.86 0.85 0.88	0.66 0.7 0.68 0.71	0.7 0.68 0.64 0.72	0.6
SES		0.7 0.7 0.71 0.7	0.67 0.67 0.66 0.66	0.77 0.78 0.75 0.78	0.43 0.43 0.4 0.43	0.84 0.84 0.84 0.83	0.72 0.73 0.68 0.74	0.62 0.6 0.63 0.58	0.5 0.4
ReliefF		0.73 0.7 0.7 0.69	0.68 0.64 0.65 0.66	0.78 0.77 <mark>0.73</mark> 0.76	0.44 0.41 0.39 0.41	0.85 0.82 0.84 0.84	0.72 0.73 0.65 0.69	0.64 0.54 0.66 0.63	0.4
RF-imp		0.74 0.73 0.7 0.71	0.67 0.68 0.65 0.68	0.78 0.78 0.77 0.78	0.44 0.44 0.42 0.44	0.85 0.85 0.84 0.85	0.72 0.72 0.72 0.72	0.63 0.64 0.59 0.64	
RFE		0.72 0.72 0.71 0.73	0.65 0.66 0.67 0.69	0.74 0.76 0.76 <mark>0.79</mark>	0.4 0.42 0.42 0.45	0.84 0.84 0.86 0.86	0.67 0.7 0.68 0.73	0.63 0.63 0.67 0.65	
Boruta		0.74 0.72 0.7 0.72	0.7 0.66 0.66 0.66	0.79 0.77 0.74 0.76	0.46 0.42 0.4 0.41	0.86 0.84 0.85 0.85	0.73 0.72 0.66 0.69	0.68 0.59 0.66 0.64	
CorrSF		0.72 0.71 0.7 0.71	0.66 0.65 0.63 0.64	0.73 <mark>0.76</mark> 0.73 0.74	0.4 0.41 0.38 0.39	0.85 0.83 0.83 0.84	0.65 0.7 0.66 0.66	0.66 0.6 0.59 0.63	
GA		0.66 0.68 0.66 0.67	0.62 0.64 0.62 0.62	0.74 0.75 0.74 0.73	0.38 0.39 0.37 0.37	0.82 0.84 0.82 0.82	0.68 0.68 0.67 0.66	0.57 0.61 0.57 0.58	
HSIC-lasso		0.71 0.71 0.68 0.7	0.67 0.66 0.66 0.65	0.75 <mark>0.78</mark> 0.75 0.74	0.41 0.42 0.4 0.39	0.85 0.84 0.84 0.84	0.67 0.72 0.67 0.66	0.67 0.61 0.64 0.65	
L1-lasso	ARI	0.68 0.68 0.69 0.7	0.67 0.65 0.66 0.64	0.78 0.78 0.75 0.76	0.44 0.43 0.41 0.4	0.84 0.83 0.85 0.83	0.74 0.74 0.68 0.7	0.6 0.57 0.64 0.58	
mRMRe	Vdq	0.73 0.73 <mark>0.68</mark> 0.72	0.68 0.67 0.63 0.67	0.73 0.74 0.74 <mark>0.76</mark>	0.4 0.4 0.39 0.42	0.87 0.86 0.82 0.85	0.64 0.65 0.67 0.69	0.72 0.69 0.59 0.66	
SES		0.69 0.68 0.67 0.69	0.65 0.65 0.65 0.64	0.76 <mark>0.78</mark> 0.76 0.76	0.4 0.42 0.4 0.4	0.83 0.82 0.84 0.83	0.71 0.75 0.69 0.7	0.59 0.55 0.61 0.58	
ReliefF		0.65 0.65 0.59 0.63	0.6 0.59 0.58 0.6	0.78 <mark>0.79 <mark>0.72</mark> 0.76</mark>	0.39 0.39 <mark>0.33</mark> 0.38	0.8 0.79 0.79 0.8	0.76 0.79 0.66 0.74	0.45 0.39 0.5 0.46	
RF-imp		0.72 0.73 0.69 0.71	0.66 0.69 0.64 0.66	0.74 0.8 0.73 0.77	0.4 0.46 0.39 0.42	0.84 0.85 0.84 0.84	0.67 0.75 0.66 0.72	0.64 0.63 0.62 0.6	
RFE		0.73 0.73 0.7 0.73	0.69 0.66 0.66 0.68	0.77 0.77 0.76 0.76	0.44 0.42 0.41 0.42	0.86 0.83 0.84 0.85	0.71 0.73 0.7 0.69	0.67 0.58 0.62 0.66	
	X	GLMASSO RY SUM	d GIM ASSO RY SUM	GIN ASSO RE SIM	dell'ASSO At SUM	GIN 550 Rt SVM	GLM 550 Rt 5VM	GLM SSO RE SUM	
\$°	oster	GIN 550 Rt SIM	dent as of the superior	chingso pt sun Booste	e v sooster	31, 12, 50 Rt 51, 10, 10, 10, 10, 10, 10, 10, 10, 10, 1	Booste	0.67 0.58 0.62 0.66	

Insights Imaging (2024) Mylona E, Zaridis DI, Kalantzopoulos CN, et al.

Supplementary Figure S2. Models' performance in Setting 2 for each feature selection method, ML classifier and MRI sequence. Red boxes indicate to the best performing combination of feature selection methods and ML classifiers for each metric.



Supplementary Figure S3. Results of Delong's test grouped based on (A) the MRI sequence, (B) the Feature Selection method, and (C) the ML classifier. Each colored point indicates how many times a specific model was found to have significantly higher ROC AUC than others. The superiority of certain groups can be determined by the larger number of points (models) and the larger number of counts per model. Specifically, significant differences were found mostly for models containing either bpMRI or solely ADC features, combined with RF-imp, L1-lasso and RFE feature selection methods. Models trained with RF, SVM or Boosted GLM had a marginal superiority to the ones based on LASSO classifier.

