

Automated Radiomics Based Clinically Significant Prostate Cancer (csPCa) Grade Classification From Biparametric MRI

Md Rakibul Islam^{1*} Abdullah Nazib^{2*} Riad Hassan^{3,4} Abu Rumman Refat⁴
Kien Nguyen² Clinton Fookes² Md Zahidul Islam¹

¹Dept. of ICT, Islamic University, Kusthtia, Bangladesh.

²School of Electrical Engineering and Robotics, QUT, Brisbane, Australia.

³IICT, Bangladesh University of Engineering and Technology, Bangladesh.

⁴Dept. of CSE, Green University of Bangladesh.

{rakibuliuict, refatiuice, zahidimage}@gmail.com, {nazib, k.nguyenthanh, c.fookes}@qut.edu.au,
riad@cse.green.edu.bd

Abstract

Prostate cancer (PCa) is one of the most common cancers in males, and clinically significant PCa (csPCa) is one of the leading causes of death worldwide. Multi-parametric MRI (mp-MRI) has gained widespread acceptance as the primary diagnostic method for prostate cancer diagnosis. However, there are still difficulties with the widespread utilization of contrast agents based on gadolinium that brings on nephrogenic systemic fibrosis, increased complexity of interpretation, and lower productivity. To overcome the drawbacks of mp-MRI, Bi-parametric MRI (bp-MRI) is an emerging solution. Here, we propose a fully automated csPCa grade classification framework that utilizes bp-MRI for segmenting csPCa lesions and uses radiomics features. We use 124 different radiomics features from 8 feature categories and select statistically significant features to train three machine learning classifiers namely SVM, KNN, Bayesian, XGBoost and Random forest. Our framework with XGBoost classifier achieved a classification accuracy of 96% in distinguishing clinically significant PCa (ISUP 4+5) and non-significant PCa (ISUP 2+3) using the predicted maps drawn by our segmentation method.

Cancer Statistics in China [32] and the United States [9]. Accurately diagnosing PCa is challenging due to symptoms varying in a broad spectrum. The severity of prostate cancer varies greatly, with some lesions exhibiting minimal growth and posing a minimum risk, while others, categorized as csPCa, display aggressive behavior, spreading quickly and potentially leading to death.

Early detection and intervention are crucial for better prostate cancer treatment. However, this disease often lacks noticeable early symptoms [21]. Recently multi-parametric MRI (mpMRI) has gained worldwide acceptance as a primary method to identify potential lesions and screen for high-risk csPCa before biopsy [22]. mpMRI is formed with T2WI, DWI, and DCE sequences [29]. However, multiple concerns exist regarding DCE: 1) gadolinium-based contrast agents raise safety concerns about brain deposits and kidney damage [20] which is not suitable for routine screening, 2) radiologists less experienced with mpMRI may show lower consistency in diagnoses due to the contrast agent [23], and 3) mpMRI scans have a relatively long acquisition time, ranging from 30 to 45 minutes [14]. Bi-parametric MRI (bpMRI) in contrast consists of T2W and DWI sequences and requires much less time. To overcome the drawbacks of mpMRI that arise for DCE, bpMRI is now an emerging solution for prostate cancer routine screening [31].

1. Introduction

Prostate cancer (PCa) is a major contributor to cancer-related death rates for men and it has become a serious health issue that impacts millions. Global cancer statistics show that prostate cancer is the second most common cancer diagnosed in men, behind lung cancer, according to

Accurate diagnosis of prostate cancer grade is crucial for optimal treatment but challenging due to similar visual characteristics of low-risk and high-risk lesions on MRI. Radiomics, analyzing textural features, has emerged as a potential solution combined with bpMRI. This study proposes an end to end solution for automated segmentation of prostate cancer zones using deep learning, and classification of prostate tumor grades from the segmented zones using

*Equal contribution.

their radiomics features. Specifically, our contributions are:

- A fully automated pipeline for pre-processing, training and testing of bpMRI sequences for 3D semantic segmentation of cancer lesions. For training and validation semantic segmentation model, we use the publicly available PICA dataset.
- Unlike previous works [13, 35] where manual segmentation is used to explore radiomics features for csPCa classification, here, we automate the process by utilizing predicted segmentation maps for extracting 124 radiomics features and predicting csPCa grades of a given bpMRI sequence.

The remainder of the paper is organized as follows: Section 2 focuses on the related works. The experimental details along with required materials are presented in Section 3. The experimental result of our proposed method is compared with state-of-the-arts methods in Section 4. Finally, the overall conclusion is presented in Section 5.

2. Related Work

Prostate disorders may cause inconveniences in the patient's day-to-day lives. To enhance clinical judgments for prostate disorders, computer-aided diagnostics builds an automated system using deep learning techniques. Deep convolutional neural networks (CNN) have made it possible to create robust Computer Aided Diagnosis (CAD) systems that show significant advancements in image analysis, pattern recognition, classification, and segmentation [8, 17, 10]. Meanwhile [15] provides a completely automated deep learning-based prostate cancer detection system. It uses 525 patients for algorithm testing, validation, and training. The dice coefficients for UNet in the training and validation sets were 0.372 and 0.287 respectively. For csPCa detection, [34] presented a two-stage 2D UNet, in which an auxiliary second-stage module uses contextual information from nearby slices to reduce false positives. Meanwhile [26] provided a CNN model with two stages, in which the prostate gland is segmented and registered in the first stage, and the binary classification of csPCa is carried out by the second step of 2D U-Net (0.86 pixel-level AUROC on 20 cases not yet seen via histological verification). As an alternative, [5] proposed a two-stage CNN model that uses prostate MRI's whole-amount histopathology mapping in conjunction with associated cancer markers to extract data (0.86 pixel-level AUROC on 20 unseen prostatectomy cases). Since prostate MRI is a nonisotropic imaging protocol, most CAD solutions for csPCa only function in two dimensions due to computational constraints. The difficulty was addressed by [33] through the utilization of specialized 2D ResNets for every slice in a patient scan. All

slice-level predictions were then aggregated using a Random Forest classifier, which produced an AUROC of 0.84 on 108 unseen biopsy-confirmed cases. [1] suggested utilizing deep learning to automatically classify Gleason patterns for determining the prostate biopsies' Gleason grade group (GGG) and used an inception network to classify GGG in a prostate histopathology picture. Nevertheless, by using pictures of prostate histopathology, [7] developed a two-phase deep learning technique for the prostate cancer Gleason Score. The first step entailed segmenting the Gleason pattern using an image of the prostate histology. In contrast, step two used the features that were extracted from the model in stage one to classify Gleason scores. An automated methodology for the detection and lesion segmentation of prostate cancer using mpMRI images, together with the estimate of Gleason scores using the ProstateX and Valencia Oncology Institute Foundation (IVO) datasets, was presented by [19].

Radiomic uses certain data characterization methods to extract a significant quantity of quantitative information from medical images. This extracted information can be integrated with clinical information to create prediction models for prognostic assessment and treatment selection. The article [27] has mentioned that Radiomic models offer a non-invasive repeatable way to assess PCa aggressiveness. [3] has also mentioned that conventional classifier models can benefit from the use of imaging properties, which are extracted from the labelled region of mpMRI. [16] investigated that radiomic models must use feature selection techniques to identify the most informative characteristics. Despite their potential to classify csPCa, earlier efforts were hampered by the bottleneck of time-intensive manual segmentation [35, 30]. This study explores a fully automated approach to overcome this limitation.

3. Materials and Methods

The overall architecture of our proposed method is depicted in Fig. 1. The methodology consists of a bpMRI pre-processing step, prospective cancer lesion ROI localization or segmentation step, radiomics Feature extraction stage, and an csPCa grade classification stage based on ML algorithms. The details of our design are explained in the following subsections.

3.1. Dataset

For this study, we use a publicly available dataset for prostate disease medical images from PICA challenge-2022 [25], from where we received bpMRI 3D volumes. The dataset comprises 1500 cases, of which 328 are from the ProstateX Challenge. A total of 220 malignant cases are manually labeled. Among them, 1075 cases had benign tissue or indolent PCa labeled as entirely zero. This dataset

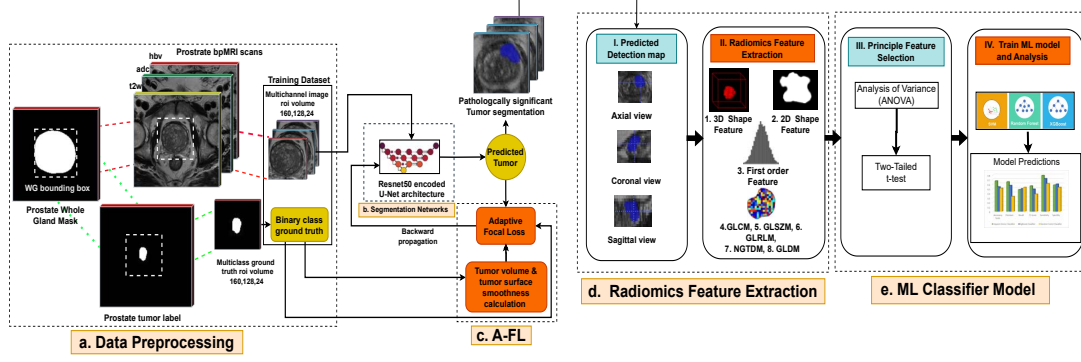


Figure 1. The overview of our working methodology consists of four main parts: a. Data pre-processing; b. Segmentation Network; c. Adaptive Focal Loss (A-FL) Function d. Radiomic Feature Extraction; e. ML classifier Model.

also provides clinical radiology information and prostate whole gland mask.

3.2. Data Pre-processing

The dataset consists of T2w images, ADC, and HBV maps for each patient. To ensure uniformity between these images and patients, the ADC and HBV maps are resampled using the maximum percentage voxel spacing 43% obtained from the axial T2W images observation and scans are resampled to $3.0 \times 0.5 \times 0.5$ mm/voxel. A two-step normalizing procedure is implemented, which includes min-max and z-score normalization. The formula for min-max and z score normalization is Equation 1 and 2 respectively where $X_{\text{normalized}}$ is the normalized value, X is the original value of the feature, and X_{\min} and X_{\max} is the minimum and maximum value of the feature in the dataset. For each image, μ represents the mean and σ the standard deviation.

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

$$Z_{\text{normalized}} = \frac{X - \mu}{\sigma} \quad (2)$$

After intensity normalization, all images, labels, and prostate whole gland masks are resized to $20 \times 256 \times 256$. We compute a bounding box around and subsequently extend these masks by 30 pixels in each direction. Using these extended bounding boxes, the prostate region is extracted from the T2W, ADC and HBV maps that belong to the same slice. All extracted regions are fixed to the size $16 \times 144 \times 128$ for segmentation model training. Finally by iteratively estimating and removing the low-frequency bias field, N4 bias field correction filter standardizes intensity values, improving the comparability of data from our multi institutional MRI images. The method involves the following key steps:

1. Log Transformation: Convert the intensity $I(\mathbf{x})$ at po-

sition \mathbf{x} into the logarithmic domain:

$$J(\mathbf{x}) = \log(I(\mathbf{x})) \quad (3)$$

2. B-Spline Approximation: Use B-splines to model the bias field $B(\mathbf{x})$. The bias field is parameterized by a set of control points \mathbf{c} , and the B-spline representation is:

$$B(\mathbf{x}) = \sum_k \mathbf{c}_k \phi_k(\mathbf{x}) \quad (4)$$

where $\phi_k(\mathbf{x})$ are the B-spline basis functions.

3. Optimization: Iteratively optimize the control points \mathbf{c} to maximize the entropy of the histogram of the corrected image. The optimization problem can be formulated as:

$$\mathbf{c}^* = \arg \max_{\mathbf{c}} H(J(\mathbf{x}) - B(\mathbf{x}; \mathbf{c})) \quad (5)$$

where H denotes the entropy function.

4. Bias Field Correction: Apply the optimized bias field to the original image:

$$I'(\mathbf{x}) = I(\mathbf{x}) \exp(-B(\mathbf{x}; \mathbf{c}^*)) \quad (6)$$

By iteratively estimating and removing the low-frequency bias field, N4 standardizes intensity values, improving the comparability of data from our multi institutional MRI images.

3.3. Segmentation Network

For segmentation, we compared the performance of our Resnet50 combined with UNet (shown in Figure 2) with 4 established models: RegUNet [18], AttUNet [24], UNet [6], and RRUNet [4]. The key challenge lies in accurately mapping the 3D volume data (T2W, ADC, and HBV channels) to a corresponding segmentation map with two channels representing foreground (tumor) and background. The PICA dataset consists of human-annotated labels with five

distinct prostate tumor grades based on the International Society of Urological Pathology (ISUP) grading system, ranging from ISUP grade 0 and 2 to 5. We convert the labels into a binary format where ISUP grade 0 tumors are considered background (0) and all other grades (ISUP 2-5) are grouped as foreground (1). We used mean Dice score and mean Intersection over Union (IoU) as a performance metric in the model validation stage. To make the models generalized we apply a combined augmentation strategy that includes random affine transformation, random flip, random scaling etc.

All segmentation models are trained with three-channel patches (T2w, ADC map, and High-b map) of size 160×128×24 as input. Adaptive Focal Loss (A-FL) [28] is used as a loss function. All models are trained with Adam optimizer with a learning rate of 0.0001, 300 epochs and batch size 1. The models are trained in a machine with an i7 CPU with 32 GB of RAM and a 12 GB NVIDIA Geforce RTX 3060 GPU.

3.4. Loss Function

The Adaptive Focal loss (A-FL) function [12] is adopted as a loss function in our research. The mathematical representation of A-FL is presented in equation 7.

$$\mathbf{A-FL}(P_t) = \alpha_{va} \cdot (1 - P_t)^{\gamma_{\text{adaptive}}} \cdot \log(P_t) \quad (7)$$

where γ_{adaptive} is the dynamically adaptive focusing parameter and α_{va} is a class balancing parameter.

3.4.1 Adaptive Focusing Parameter

The Adaptive Focusing Parameter (γ_{adaptive}) leverages tumor volume size and tumor surface smoothness information to improve Focal Loss in semantic segmentation. The equation for γ_{adaptive} is:

$$\gamma_{\text{adaptive}} = \gamma_{va} + \gamma_{mSa} \quad (8)$$

Here, the Volume Adaptive Parameter (γ_{va}) gives tumor size information to the focusing parameter. It is calculated as the ratio of foreground (tumor) pixels to the total pixels (foreground + background). The Mean Surface Smoothness Adaptive Parameter (γ_{mSa}) captures the smoothness of the tumor's surface and provides this information to the focusing parameter. It is computed by first calculating the image gradients along the x, y, and z axes within three dimension volume. The magnitude of these gradients is then averaged across the tumor to determine the mean gradient magnitude.

3.4.2 Adaptive Class Balancing Parameter

To address class imbalance and give more focus to small tumor cases, we dynamically calculate the Class Balancing Adaptive Parameter (α_{va}) using the ratios of cancerous and

non-cancerous pixels to the total pixels for each patient's tumor during training. The equation for α_{va} is given by:

$$\alpha_{va} = \frac{P_{bg}}{P_{fg} + P_{bg}} \quad (9)$$

Here, P_{fg} represents the count of foreground pixels (non-zero elements) and P_{bg} represents the count of background pixels (zero elements). This calculation ensures that the model pays appropriate attention to smaller tumors, which might otherwise be overshadowed by the larger number of non-cancerous pixels.

3.5. Radiomics Feature Extraction

We used Pyradiomics toolboxes for extracting radiomic features from images and predicted labels from the trained segmentation model. Radiomics features are extracted from the T2W modality. The extracted radiomics dataset consisted of 124 different features of 8 different feature groups calculated with PyRadiomics v3.7. The feature groups consisted of first-order features, 3D shape feature, 2D shape feature, and five texture-based feature groups: gray-level co-occurrence matrix (GLCM), gray-level dependence matrix (GLDM), gray-level run-length matrix (GLRLM), gray-level size-zone matrix (GLSZM), and the neighboring gray-tone difference matrix (NGTDM) features.

3.6. ML classifier Model Training

First, the z score normalization was used to independently normalize the radiomic features that were taken from the pyradiomics toolboxes. Second, to assess the impact of radiomic features on prostate cancer aggressiveness classification, we employed an analysis of variance (ANOVA) ($p < 0.05$). Following feature selection, we delved deeper with two-tailed t-tests to identify features significance associated with ISUP grade groups ($p < 0.05$). In this process, 20 principle radiomics features are selected. Finally, five machine learning classifiers are trained on selected radiomics features. The ML classifiers included XGBOOST, random forest (RF), support vector machine (SVM), K-Nearest Neighbor Algorithm (KNN), and Bayesian Classifier. We have used the ISUP grade from the clinical information datasheet as the label for the model training.

4. Results

For segmentation model training, 225 patients are used for the model training, validation and testing. Three cohort sets were created from the PICA dataset: a test cohort ($n = 22$, 10%), a validation cohort ($n = 45$, 20%), and a training cohort ($n = 158$, 70%). Table 1 displays the patient demographics following the separation of the data.

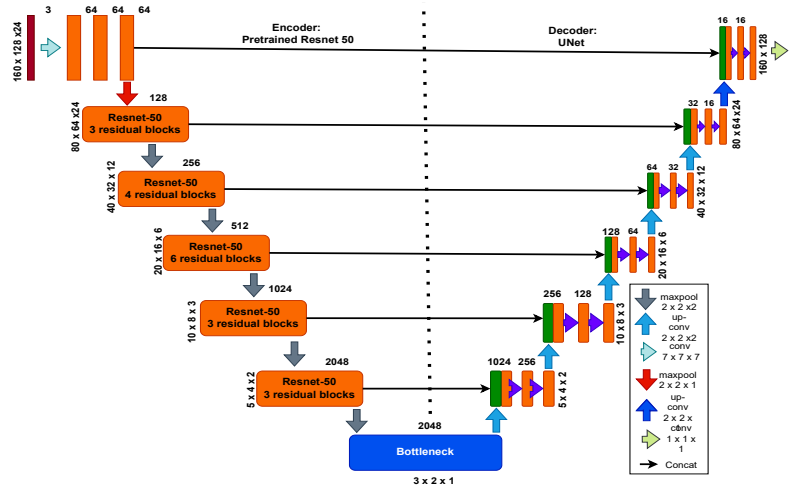


Figure 2. The U-Net network architecture uses pre-trained Resnet50 [11] as backbone. The residual blocks in ResNet50 enable the network to learn more complex features and deeper representations, which are crucial for accurate segmentation.

Table 1. csPCa grade distribution for the PICAI dataset before and after data splitting

Patient Characteristic Variable	Overall n=225	Training n=158 (70%)	Validation n=45 (20%)	Test, n=22 (10%)
Age (year)	66 (61-70)	67 (61-70)	64 (62-69)	65 (61-68)
PSA (ng/mL)	8.5 (6-13)	8.2 (6-12)	8.0 (7-12)	8.1 (6-13)
Volume (mL)	57 (40-80)	57 (40-80)	57 (42-76)	54 (40-75)
ISUP 2	74 (33%)	52	15	7
ISUP 3	59 (26%)	41	12	6
ISUP 4	45 (20%)	32	9	5
ISUP 5	47 (21%)	33	10	4

Table 2. Comparison of segmentation models, showing mean Dice and IoU scores on validation and test sets.

Split	Model	Dice Score	IoU Score
Valid set	RegUNet	0.635	0.573
	RRUNet	0.740	0.672
	AttUNet	0.750	0.679
	UNet	0.740	0.669
	ResNet50 using UNet	0.769	0.696
Test set	RegUNet	0.627	0.558
	RRUNet	0.732	0.661
	AttUNet	0.741	0.689
	UNet	0.729	0.672
	ResNet50 using UNet	0.758	0.683

4.1. Results on Segmentation

The Dice Score and IoU are metrics used to evaluate the performance of semantic segmentation models. Table 2 shows that the ResNet50 combined with a UNet architecture achieved the highest performance in our experiments. It achieves Dice Scores of 0.769 on the validation set and 0.758 on the test set. For IoU, it scores 0.696 on the validation set and 0.683 on the test set. Our experiment showed a 30.2% increase in Dice score over the method in [2] (DSC=0.467) and a 39.7% increase over the method in [15] (DSC=0.372), both of which were trained on the PICAI and ProstateX datasets respectively.

4.2. Results on csPCa Grade Classification

The analysis in Table 3 reveals 20 statistically significant radiomics features out of 124, using ANOVA and a two-tailed t-test with a threshold of $p < 0.05$. Highly significant features ($p < 0.001$), such as Skewness, GLRLM-Gray Level NonUniformity, GLSZM-Large Area High

Gray Level Emphasis, and Kurtosis, strongly differentiate clinically significant (ISUP grades 2 & 3) from non-clinically significant (ISUP grades 4 & 5) tumor grades. Moderately significant features ($0.001 \leq p < 0.01$), including Maximum 2D Diameter Column, Voxel Volume, Mask-original Voxel Number, MeshVolume, and gldm-Gray Level NonUniformity, also play a robust role in classification. Notably significant features ($0.01 \leq p < 0.05$), like MajorAxisLength, SurfaceArea, gldm-Dependence NonUniformity, Maximum3D Diameter, and others, contribute meaningfully as well. The inverse relationship between weight values and their informative value highlights that features with smaller weights provide more critical insights for tumor grading.

We evaluated classification models using Accuracy and Area Under the Curve(AUC). Table 4 shows the performance of five classifiers: SVM, KNN, Bayesian, RF, and XGBoost. These classifiers were chosen for their popularity and diverse approaches, aligning with prior studies

Table 3. Radiomic Features p-value for Tumor Grade Classification Task for Non-Clinical Significant (ISUP grade 2 & 3) vs. Clinical Significant (ISUP grade 4 & 5).

Radiomics Features	P-value	Radiomics Features	P-value
1. Skewness	0.000	8. MeshVolume	0.003
2. GLRLM-GrayLevel NonUniformity	0.000	9. gldm-GrayLevelNonUniformity	0.0066
3. GLSZM-Large Area High Gray-Level Emphasis	0.000	10. MajorAxis Length	0.014
4. Kurtosis	0.000	11. SurfaceArea	0.0051
5. Maximum 2D Diameter Column	0.001	12. gldm-Dependence NonUniformity	0.019
6. Voxel Volume	0.001	13. Maximum3D Diameter	0.019
7. Mask-original Voxel Number	0.003	14. MinorAxis Length	0.024
15. Maximum2D DiameterRow	0.024	16. TotalEnergy	0.038
17. Maximum2D DiameterSlice	0.030	18. LeastAxis Length	0.045
19. glrlm-RunLength NonUniformity	0.036	20. glcm-Cluster Shade	0.046

Table 4. The table presents the classification results for various classification models on prostate disease diagnosis based on principal radiomics features.

Framework	Classifier	Accuracy	AUC
[13]	SVM	0.79	0.94
	KNN	0.76	0.89
	RF	0.76	0.94
[35]	SVM	0.73	-
	KNN	0.64	-
	Bayesian	0.68	-
Our proposed	SVM	0.83	0.94
	KNN	0.78	0.90
	Bayesian	0.76	0.87
	RF	0.87	0.92
	XGBoost	0.96	0.95

[13, 35] on the same dataset. Among these classifiers, XGBoost shows an impressive accuracy of 96%, while Random Forest achieves the second highest accuracy with a score of 87%. XGBoost also excels in AUC with a score of 95%, indicating its effectiveness in distinguishing between classes. The SVM, KNN, and Bayesian models achieve accuracies of 83%, 78%, and 76%, respectively, with notable AUC values of 94%, 90%, and 87%, indicating strong discriminative power. The proposed framework with XGBoost classifier outperforms two previous radiomics-based classification studies by a significant margin in terms of accuracy. Overall, XGBoost accurately predicted almost 96% of all positive cases, demonstrating its excellent capability in both overall classification and capturing relevant positive instances.

5. Conclusion

In this study, we propose an end-to-end automated framework to classify clinically significant PCa (ISUP grade > 3) and non-significant PCa (ISUP grade ≤ 3). Instead of using human-annotated labels, our framework uses a predicted segmentation map from a deep-learning model to classify PCa grades and demonstrate comparable accuracy. Our high classification accuracy demonstrates the significance of an automated pipeline for optimizing treatment decisions for prostate cancer patients.

6. Acknowledgment

This research is funded by Grand NO:56.00.0000.052.33.001.23-59 from the **ICT Division, Bangladesh**. The authors also wish to thank Computer Vision & Intelligent Interfacing (CVIIL) Lab, Islamic University, Kushtia-7003, Bangladesh for accessing high-configuration computer set-up for performing experiments. The authors also wish to thank SAIVT, QUT for continuous support, study design and supervision of the project.

References

- [1] R. Alfano, G. S. Bauman, J. A. Gomez, M. Gaed, M. Moussa, J. Chin, S. Pautler, and A. D. Ward. Prostate cancer classification using radiomics and machine learning on mp-mri validated using co-registered histology. *European Journal of Radiology*, 156:110494, 2022. 2
- [2] J. A. Alzate-Grisales, A. Mora-Rubio, F. García-García, R. Tabares-Soto, and M. De La Iglesia-Vayá. Samunetr: Clinically significant prostate cancer segmentation using transfer learning from large model. *IEEE Access*, 11:118217–118228, 2023. 5
- [3] M. M. Badža and M. Č. Barjaktarović. Classification of brain tumors from mri images using a convolutional neural network. *Applied Sciences*, 10(6):1999, 2020. 2
- [4] X. Bi, Y. Wei, B. Xiao, and W. Li. Rru-net: The ringed residual u-net for image splicing forgery detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2019. 3
- [5] V. Brancato, G. Di Costanzo, L. Basso, L. Tramontano, M. Puglia, A. Ragozzino, and C. Cavaliere. Assessment of dce utility for pca diagnosis using pi-rads v2. 1: Effects on diagnostic accuracy and reproducibility. *Diagnostics*, 10(3):164, 2020. 2
- [6] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19*, pages 424–432. Springer, 2016. 3
- [7] J. I. Epstein. Prostate cancer grading: a decade after the 2005 modified system. *Modern Pathology*, 31:47–63, 2018. 2

- [8] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639):115–118, 2017. 2
- [9] J. Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, D. M. Parkin, M. Piñeros, A. Znaor, and F. Bray. Estimating the global cancer incidence and mortality in 2018: Globocan sources and methods. *International journal of cancer*, 144(8):1941–1953, 2019. 1
- [10] R. Hassan, M. R. H. Mondal, and S. I. Ahamed. Udbnnet: A novel uncertainty driven boundary refined network for organ at risk segmentation. *PLOS ONE*, 19(6):1–18, 06 2024. 2
- [11] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 5
- [12] M. R. Islam, R. Hassan, A. Nazib, K. Nguyen, C. Fookes, and M. Z. Islam. Enhancing semantic segmentation with adaptive focal loss: A novel approach, 2024. 4
- [13] H. Khanfari, S. Mehranfar, M. Cheki, M. Mohammadi Sadr, S. Moniri, S. Heydarheydari, and S. M. Rezaei. Exploring the efficacy of multi-flavored feature extraction with radiomics and deep features for prostate cancer grading on mpMRI. *BMC Medical Imaging*, 23(1):195, 2023. 2, 6
- [14] C. K. Kuhl, R. Bruhn, N. Krämer, S. Nebelung, A. Heidenreich, and S. Schrading. Abbreviated biparametric prostate mr imaging in men with elevated prostate-specific antigen. *Radiology*, 285(2):493–505, 2017. 1
- [15] S. Mehralivand, D. Yang, S. A. Harmon, D. Xu, Z. Xu, H. Roth, S. Masoudi, D. Kesani, N. Lay, M. J. Merino, et al. Deep learning-based artificial intelligence for prostate cancer detection at biparametric MRI. *Abdominal Radiology*, 47(4):1425–1434, 2022. 2, 5
- [16] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab. A high-bias, low-variance introduction to machine learning for physicists. *Physics reports*, 810:1–124, 2019. 2
- [17] A. Nazib, R. Hassan, Z. Islam, and C. Fookes. Uncertainty driven bottleneck attention u-net for organ at risk segmentation. In *2024 IEEE 21th International Symposium on Biomedical Imaging (ISBI)*, 2024. 2
- [18] G. D. Pais, S. Ramalingam, V. M. Govindu, J. C. Nascimento, R. Chellappa, and P. Miraldo. 3dregnet: A deep neural network for 3d point registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7193–7203, 2020. 3
- [19] O. J. Pellicer-Valero, J. L. Marenco Jimenez, V. Gonzalez-Perez, J. L. Casanova Ramon-Borja, I. Martin Garcia, M. Barrios Benito, P. Pelechano Gomez, J. Rubio-Briones, M. J. Rupérez, and J. D. Martín-Guerrero. Deep learning for fully automatic detection, segmentation, and gleason grade estimation of prostate cancer in multiparametric magnetic resonance images. *Scientific reports*, 12(1):2975, 2022. 2
- [20] M. Ramalho, J. Ramalho, L. M. Burke, and R. C. Semelka. Gadolinium retention and toxicity—an update. *Advances in Chronic Kidney Disease*, 24(3):138–146, 2017. 1
- [21] P. Rawla. Epidemiology of prostate cancer. *World journal of oncology*, 10(2):63, 2019. 1
- [22] J. Richenberg, V. Løgager, V. Panebianco, O. Rouviere, G. Villeirs, and I. G. Schoots. The primacy of multiparametric MRI in men with suspected prostate cancer. *European Radiology*, 29:6940–6952, 2019. 1
- [23] A. B. Rosenkrantz, L. A. Ginocchio, D. Cornfeld, A. T. Froemming, R. T. Gupta, B. Turkbey, A. C. Westphalen, J. S. Babb, and D. J. Margolis. Interobserver reproducibility of the PI-RADS version 2 lexicon: a multicenter study of six experienced prostate radiologists. *Radiology*, 280(3):793–804, 2016. 1
- [24] R. Roy, B. Annappa, and S. Dodia. 3d attu-net for brain tumor segmentation with a novel loss function. In *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, pages 1–8. IEEE, 2023. 3
- [25] A. Saha, J. S. Bosma, J. J. Twilt, B. van Ginneken, A. Bjartell, A. R. Padhani, D. Bonekamp, G. Villeirs, G. Salomon, G. Giannarini, J. Kalpathy-Cramer, J. Barentsz, K. H. Maier-Hein, M. Rusu, O. Rouvière, R. van den Bergh, V. Panebianco, V. Kasivisvanathan, N. A. Obuchowski, D. Yakar, M. Elschot, J. Veltman, J. J. Fütterer, M. de Rooij, H. Huisman, and the PI-CAI consortium. Artificial intelligence and radiologists in prostate cancer detection on MRI (PI-CAI): an international, paired, non-inferiority, confirmatory study. *The Lancet Oncology*, 25(7):879–887, 2024. 2
- [26] J. Sanyal, I. Banerjee, L. Hahn, and D. Rubin. An automated two-step pipeline for aggressive prostate lesion detection from multi-parametric MR sequence. *AMIA Summits on Translational Science Proceedings*, 2020:552, 2020. 2
- [27] P. Ström, K. Kartasalo, H. Olsson, L. Solorzano, B. Delahunt, D. M. Berney, D. G. Bostwick, A. J. Evans, D. J. Grignon, P. A. Humphrey, et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *The Lancet Oncology*, 21(2):222–232, 2020. 2
- [28] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*, pages 240–248. Springer, 2017. 4
- [29] B. Turkbey, A. M. Brown, S. Sankineni, B. J. Wood, P. A. Pinto, and P. L. Choyke. Multiparametric prostate magnetic resonance imaging in the evaluation of prostate cancer. *CA: a cancer journal for clinicians*, 66(4):326–336, 2016. 1
- [30] J. F. Veenland. Automated classification of significant prostate cancer on MRI: A systematic review on the performance of machine learning applications. 2
- [31] S. Woo, C. H. Suh, S. Y. Kim, J. Y. Cho, S. H. Kim, and M. H. Moon. Head-to-head comparison between biparametric and multiparametric MRI for the diagnosis of prostate cancer: a systematic review and meta-analysis. *American Journal of Roentgenology*, pages W226–W241, 2018. 1
- [32] C. Xia, X. Dong, H. Li, M. Cao, D. Sun, S. He, F. Yang, X. Yan, S. Zhang, N. Li, et al. Cancer statistics in China

- and united states, 2022: profiles, trends, and determinants. *Chinese medical journal*, 135(05):584–590, 2022. [1](#)
- [33] S. Yoo, I. Gujrathi, M. A. Haider, and F. Khalvati. Prostate cancer detection using deep convolutional neural networks. *Scientific reports*, 9(1):19518, 2019. [2](#)
- [34] X. Yu, B. Lou, B. Shi, D. Winkel, N. Arrahmane, M. Djalilo, T. Meng, H. von Busch, R. Grimm, B. Kiefer, et al. False positive reduction using multiscale contextual features for prostate cancer detection in multi-parametric mri scans. In *2020 IEEE 17th international symposium on biomedical imaging (ISBI)*, pages 1355–1359. IEEE, 2020. [2](#)
- [35] H. Zhuang, A. Chatterjee, X. Fan, S. Qi, W. Qian, and D. He. A radiomics based method for prediction of prostate cancer gleason score using enlarged region of interest. *BMC Medical Imaging*, 23(1):205, 2023. [2](#), [6](#)