World Wide Web (2026) 29:4
https://doi.org/10.1007/s11280-025-01392-6

RESEARCH

Exploring UNet-based models for prostate lesion segmentation from multi-sequence MRI (T2W, ADC, DWI)

Saman Fouladi[1] · Fatemeh Darvizeh[2] · Gabriele Gianini[3] · Rosario Di Meo[2] · Luca Di Palma[2] · Ernesto Damiani[1] · Alessandro Maiocchi[4] · Deborah Fazzini[2] · Marco Alì[2]

Received: 30 March 2025 / Revised: 20 August 2025 / Accepted: 28 November 2025

Abstract
Accurate delineation of lesions in prostate MRI is crucial for the diagnosis of prostate cancer. Manual segmentation is time-consuming, requires advanced medical expertise, and is subject to inter-operator variability. Automatic lesion segmentation therefore represents a valuable tool to support clinicians by reducing workload, minimizing observer bias, and enabling more consistent image analysis. In this work, we investigated the performance of four deep learning architectures for prostate lesion segmentation: nnU-Net, DenseUNet, SegResUNet, and U-Net. Unlike many existing studies that rely on publicly available data, we constructed a dedicated dataset to better capture real-world variability and challenges. The dataset, comprising T2-weighted (T2W), apparent diffusion coefficient (ADC), and diffusion-weighted imaging (DWI) sequences, was carefully annotated by medical experts to ensure high-quality labels. Training was performed using the full combination of these modalities. Two cohorts were considered based on lesion severity, as defined by PI-RADS (Prostate Imaging–Reporting and Data System) scores: one with only PI-RADS 4–5 lesions (151 patients), and another including PI-RADS 3 cases, totaling 209 patients. Evaluation was conducted both on a patient-by-patient basis and in a consolidated all-patient setting. In the patient-level analysis, nnU-Net achieved the highest Dice similarity coefficient (DSC) of 0.60 when trained on PI-RADS 4–5 lesions, while in the all-patient analysis, DenseUNet attained a DSC of 0.57 on the same dataset. These results are within the range reported in recent prostate lesion segmentation studies, and in some cases are comparable to or exceed those obtained with substantially larger datasets.

Keywords Automatic Prostate MRI segmentation · Prostate cancer · Lesion segmentation · Multimodal imaging · PI-RADS · U-Net · nnU-Net · DenseUNet · SegResUNet

Saman Fouladi and Fatemeh Darvizeh contributed equally.

Extended author information available on the last page of the article

Published online: 10 December 2025

1 Introduction

Prostate cancer (PCa) incidence has been increasing in recent years. In 2022, there were 1,467,854 new cases of prostate cancer worldwide [1]. In the United States, projections for 2025 estimate about 313,780 new cases and 35,770 deaths due to prostate cancer [2]. This upward trend underscores the critical importance of early detection and effective treatment strategies.

Magnetic resonance Imaging (MRI) is commonly used to locate tumors in the prostate because it does a better job of showing soft tissues than other medical images [3]. Multiparametric MRI (mpMRI) involves using various imaging sequences to emphasize tissue characteristics. The most common sequences include T2-weighted images (T2W), diffusion-weighted imaging (DWI), and dynamic contrast-enhanced MRI. The apparent diffusion coefficient (ADC) is calculated from DWI at different b-values. Biparametric MRI (bpMRI) simplifies this by focusing mainly on three sequences: T2W, DWI, and ADC [4, 5]. The European Society of Residents in Urology (ESRU) has developed a guideline for interpreting prostate cancer, the Prostate Imaging–Reporting and Data System (PI-RADS) version 2.1, which relies on T2W, DWI, and ADC to analyze malignant areas and assess tumor risk using PI-RADS scoring [6].

T2W plays a key role in prostate MRI, capturing high-resolution scans in axial, coronal, and sagittal planes with a focused field of view. It allows for detailed visualization of prostatic zonal anatomy and helps identify extra-prostatic extension [7]. DWI is a technique that generates image contrast based on differences in water molecule diffusion within tissues. In prostate cancer, tumor progression leads to increased cellularity, structural changes, and a reduced cytoplasm-to-nucleus ratio, all of which restrict water diffusion. This restriction becomes more pronounced with higher cancer grades. DWI analysis involves two key components: high b-value images and ADC maps. The b-value represents the strength of diffusion weighting, with higher values enhancing tissue contrast. ADC maps provide a quantitative measure of water diffusion, where areas with restricted diffusion appear bright on DWI and dark on ADC maps due to lower ADC values. These imaging techniques play a crucial role in evaluating prostate cancer by highlighting variations in tissue diffusion properties [8].

Accurate segmentation of the prostate anatomy and lesions is essential for the precise diagnosis and effective treatment planning of prostate cancer, enabling detailed localization and characterization of malignant tissues [9]. However, manually segmenting tumors is a time-intensive task, with the accuracy of results relying heavily on the radiologist's expertise. Consequently, significant efforts have been dedicated to developing automated methods for tumor segmentation to enhance efficiency and consistency [10]. The advancements in deep learning have notably enhanced the segmentation process in bpMRI, improving accuracy and efficiency in clinical applications. Convolutional neural networks (CNNs) play a crucial role in semantic segmentation, with U-Net [11] being a notable example, and have become essential in medical image segmentation. Various U-Net variants have been introduced for this purpose in the medical field. Using deep learning, automatic segmentation of the prostate and cancer lesions in MRI images has achieved promising results, benefiting from strong hierarchical feature representation and data-driven, end-to-end learning [12].

In this study, we introduce a multi-encoder U-Net–based framework specifically tailored for prostate lesion segmentation using multiparametric MRI. Unlike conventional single-

encoder U-Nets, our approach employs three parallel encoder streams, each dedicated to a distinct imaging modality (T2W, ADC, and DWI), to better capture complementary anatomical and functional information. We evaluated four architectures within this framework. For nnU-Net, one of the most widely recognized and benchmarked medical image segmentation frameworks, we applied it in a multi-input configuration to handle the three modalities simultaneously. For the other networks, U-Net, Dense-UNet, and SegRes-UNet, we redesigned the original architectures to incorporate the proposed multi-encoder structure, enabling parallel feature extraction from each modality before fusion in the decoding stage. This adaptation allows the networks to exploit the full richness of multiparametric MRI, which is often underutilized in conventional segmentation pipelines.

These networks were trained on a carefully assembled dataset collected at our clinical center, designed to reflect real-world clinical conditions. The dataset includes all three MRI sequences, along with corresponding lesion masks. To our knowledge, no publicly available dataset offers this complete combination of imaging modalities paired with lesion annotations linked to PI-RADS scores. Publicly available datasets typically include only one or two MRI sequences, most often T2W and/or ADC, and in many cases do not include DWI. Moreover, these datasets generally lack PI-RADS–specific lesion annotations. Such detailed labeling is essential for clinically meaningful segmentation, as it allows models to be developed and evaluated according to lesion severity and the likelihood of clinically significant prostate cancer.

Accurate segmentation of prostate lesions with PI-RADS scores of 4 and 5 is particularly critical, as these categories carry substantially higher risks of clinically significant cancer. For instance, in a prospective cohort of 454 men, the rates of clinically significant prostate cancer (csPCa) were 14% for PI-RADS 3, 37% for PI-RADS 4, and 77% for PI-RADS 5 lesions [13]. A systematic review further found that 52% of PI-RADS 4 and 89% of PI-RADS 5 lesions corresponded to csPCa [14]. These findings demonstrate that PI-RADS 4 and 5 lesions are more likely to be clinically significant compared to PI-RADS 3, and thus warrant focused architectural and methodological attention.

The remainder of this article is structured as follows: Sect. 2 details the Materials and Methods, including the preparation of our in-house dataset and the proposed approach, Sect. 3 presents the results, followed by a discussion in Sect. 4 and conclusions in Sect. 6.

2 Related work

Deep learning-based approaches have been widely explored for prostate lesion segmentation tasks in recent years. Various studies have employed neural network architectures, datasets, and evaluation metrics to enhance segmentation performance. These works primarily utilize publicly available and private medical imaging datasets, with the number of images varying significantly depending on the dataset. Additionally, different studies adopt architectures such as U-Net, U- Net++, Attention U-Net, and other advanced variants to improve accuracy.

Table 1 summarizes the relevant studies; their key findings are reported right after. To report about the performance of the different methods, we used the Dice Similarity Coefficient (DSC), for quantifying the predicted and ground truth segmentation masks. It is defined as:

Table 1 Comparison of the results different studies for prostate lesion segmentation. DSC: dice similarity coefficient; T2W: T2-weighted; ADC: apparent diffusion coefficients

Study	Dataset	Network	DSC
[12]	Prostate158 (70 T2W) + Prostate X (70 T2W)	Dual-pyramid network combining Convolutional Neural Networks (CNN) and tokenized Multi-Layer Perceptrons (MLP)	56.31%
[15]	Prostate158 (70 T2W) + Prostate X (70 T2W)	Adaptive Window Adjustment (AWA)	48.56%
	PI-CAI (220 T2W)		49.03%
[25]	Private dataset (390 ADC)	Integration of SE-Net, Cased Pyramid Convolution, and an Attention Gate	86.3%
[10]	Private dataset (355 T2W)	neighbor-aware multi-modal adaptive learning network (NaMa)	63.76%
	PI-CAI (215 T2W)		59.01%
[17]	PI-CAI (220 T2W)	A cascaded two-stage SegResNet	53.34%
	Prostate158 (70 T2W)		42.11%
	PSU-mpMRI (10 T2W)		35.88%
[18]	Prostate158 (70 T2W)	U-ResNets	45%
[19]	Private dataset (149 ADC)	U-Net	77%
[20]	I2CVB (151 T2W)	Cascaded pyramid convolution module	79%
[21]	I2CVB (151 T2W)	Deep Convolutional Encoder-Decoder	80.24%
[22]	Private dataset (77 T1W)	Cascaded scoring convolutional neural network	69%
[23]	Private dataset (120 T2W)	Mask-RCNN	62%
[24]	Private dataset (152 ADC)	Fully Convolutional Networks	41%
[25]	Private dataset (16 ADC) 16 patients with lesions in peripheral zone only	Support Vector Machines	58%
[26]	Private dataset (11 T2W)	Fuzzy Markov random fields	62%
[27]	Private dataset (64 T2W)	Mask R-CNN	60%
[28]	Prostate158 (70 T2W) + PI-CAI (220 T2W)	SwinUNETR	48%
[29]	I2CVB (151 T2W)	U-Net LSTM	67%
[30]	Private dataset (122 T2W)	U-Net	31%

$$DSC = \frac{2TP}{2TP + FP + FN}$$

(TP = true positives, FP = false positives, FN = false negatives). A higher DSC value indicates better segmentation performance, with a value of 1 representing perfect overlap between the predicted and ground truth masks. Another metric occasionally used in literature to this

purpose is the Intersection over Union (IoU) metric (a.k.a Jaccard index). While the two metrics are linked by a simple formula (IoU = DSC/(2-DSC)), they emphasize different aspects of the overlap. DSC is generally more sensitive to small regions, while IoU penalizes false positives more strongly. We reported our results in terms of DSC, since it is the most frequently used metrics, and for comparison converted the occasional literature results expressed as IoU into DSC values.

In [12], authors utilized 140 MRI cases, with 70 from the ProstateX dataset and 70 from Prostate158. The dataset included cases with and without lesions. This work introduces a dual-pyramid network combining CNN and tokenized MLP for prostate lesion segmentation in MRI. It employs a two-stage approach: prostate segmentation followed by lesion detection, integrating multi-scale feature extraction and an attention module for enhanced accuracy. The proposed network achieved an IoU of 44.82 ± 20.45 and a DSC of 56.31 ± 22.77 on T2W.

An Adaptive Window Adjustment (AWA) module proposed in [15] dynamically adjusts window settings for different image modalities and tasks. The study also evaluates loss functions for prostate segmentation, highlighting the superiority of Boundary Difference over Union (DoU) Loss, especially in 3D imaging. Additionally, a cascaded segmentation approach is introduced to enhance lesion segmentation using anatomical structure information. Using 5-fold cross-validation, the proposed method achieved a mean DSC of 48.56 for lesion segmentation on Prostate158 and ProstateX datasets and 49.03 on the PI-CAI dataset, demonstrating its effectiveness. The proposed multi-segmentation model utilizes T2WI and ADC to segment the prostate gland and tumor.

It integrates SE-Net, Cased Pyramid Convolution, and an Attention Gate for enhanced feature extraction, while multi-scale convolutions and a convolution-based edge detection block improve segmentation accuracy. The model was evaluated on a dataset of 390 MRIs from Guizhou Provincial People's Hospital, achieving a DSC of 86.3 ± 0.027 for lesion segmentation, demonstrating strong generalization in external validation [16].

A two-stage neighbor-aware multi-modal adaptive learning network (NaMa) was proposed for prostate tumor segmentation from multi-modal anisotropic MR images [10]. The method incorporates a Modality-Informative Adaptive Learning (MIAL) module for adaptive fusion of modality-specific information and an Activation Mapping Guidance (AMG) module to refine inter-slice feature correlations for consistent segmentation. A random mask strategy was also applied to enhance feature representation. The model achieved a DSC of 63.76% on their dataset and 59.01% on the PI-CAI dataset.

A cascaded two-stage SegResNet approach was proposed for prostate cancer lesion segmentation in bpMRI [17].The first model identifies the region of interest, while the second segments lesions within the prostate using T2W, DWI, and ADC sequences. Trained on the PI-CAI 2022 dataset, the model outperformed a single SegResNet baseline, achieving Dice scores of 0.5334 on PI-CAI 2022, 0.4221 on Prostate-158, and 0.3588 on PSU-mpMRI, demonstrating strong generalizability to external data.

The authors of [18] proposed a U-ResNets model to segment PCa prostate lesions employing the Prostate158 dataset. The network delineated the lesions achieving a DSC of 45%.

In [19] U-Net was designed for segmenting prostate lesions, achieving a DSC of 77% for accurate lesion segmentation. The dataset comprised 149 patients with a mean age of

69.2 years (ranging from 47 to 84 years), divided into training (80%) and internal validation (20%) sets.

The study [20] aimed to create an advanced segmentation network incorporating a cascading pyramid convolution module (CPCM) and a double-input channel attention module (DCAM) for accurate and automated segmentation of PCa lesions from mp-MRI scans. The model, evaluated on the I2CVB dataset, attained a DSC of 79%.

The research by [21] employed the I2CVB dataset to segment prostate cancer lesions. They utilized a sophisticated deep convolutional encoder-decoder model to segment not only the prostate and its anatomical features but also malignant lesions. For cancer segmentation, the system achieved an IoU score of 67%, i.e. a DSC of 80.24%.

A novel 3D cascaded scoring convolutional neural network was introduced for the automatic segmentation of both the prostate and lesions from MRI images [22]. This innovative network operates by sequentially identifying a region of interest (ROI), detecting the object within this ROI, and then defining the target area. The effectiveness of this method was assessed through a retrospective study involving 77 patients who had undergone MRI scans. Results showed that the proposed network attained a DSC of 69% for segmenting lesions.

A study was conducted to train and assess mask region-based convolutional neural networks for the segmentation of the prostate gland and lesions [23]. The study comprised two cohorts: 78 patients from public sources (cohort 1) and 42 patients from their institution (cohort 2). The model achieved a lesion segmentation DSC of 62%.

In a separate investigation [24], an FCN underwent training with 152 MRI scans to accomplish prostate lesion segmentation. The network achieved a DSC of 41%.

An experiment evaluated enhanced localization in multispectral MRI data using a semi-supervised method and an SVM classifier with varied feature vectors for segmentation, achieving a DSC of 58% for lesion segmentation [25].The study utilized a dataset of 16 MRIs, including T2 maps and ADC images, demonstrating promising results in improving localization accuracy.

A novel approach to segmenting prostate cancer in multispectral MR images is proposed by [26], employing fuzzy Markov random fields (fuzzy MRFs). The developed model attained a DSC of 62% on a dataset comprising 11 MRIs.

In [27], a non-local Mask R-CNN was enhanced through various training methodologies, showing improved performance in a retrospective analysis of 262 patients who underwent prostate bi-parametric MRI scans. The enhanced model was evaluated against baseline Mask R-CNN, 3D U-Net, and manual delineation by a radiologist using metrics such as detection rate, DSC, sensitivity, and Hausdorff Distance (HD). The dataset was divided into three cohorts: Cohort 1 (64 patients) used histopathology for ground truth, Cohort 2 (158 patients) relied on bp-MRI, and Cohort 3 (40 patients) was used for semi-supervised learning. The enhanced Mask R-CNN achieved the highest DSC of 60%.

In their research, the authors [28] introduced SAM-UNETR, a novel approach to delineating regions of clinically significant PCa in MRI images. SAM-UNETR combines a transformer-encoder derived from the Segment Anything Model (SAM), a highly adaptable segmentation model trained on an extensive dataset of 11 million images, with a residual-convolution decoder inspired by UNETR. Two distinct datasets were employed in the study: Prostate158 and the dataset from the PI-CAI Challenge, each offering unique perspectives and contributions. SAM-UNETR achieves a DSC of 48% for PCa prediction.

In their study, Gavade et al. [29] employed modern DL techniques to develop a framework aimed at segmenting and categorizing mp-MRI images. The DL methodology consists of two main stages: initially utilizing a U-Net architecture to segment ROI, followed by employing a long short-term memory (LSTM) network to classify these ROIs into cancerous or non-cancerous categories. Training of the DL models was conducted on the I2CVB dataset, and a comprehensive evaluation was performed through a detailed comparison with their experimental setup. The model achieved a significant DSC of 67% for lesion segmentation.

In [30], a transparent DL model is employed to decipher the forecasts of a CNN in prostate tumor segmentation. Utilizing a U-Net framework, the CNN was trained on multiparametric MRI data from 122 patients to autonomously delineate the prostate gland and tumor lesions. The CNN attained a mean DSC of 62% for the prostate gland and 31% for the tumor lesions segmentation.

3 Materials and methods

This section outlines the key components of our study methodology. We begin by describing the dataset used, including patient selection and imaging modalities. Next, we present the deep learning architectures employed for prostate lesion segmentation, detailing the four networks evaluated: U-Net, nnU-Net, Dense-UNet, and SegRes-UNet. We then explain our approach to combining segmentation masks derived from multi-sequence inputs. Finally, the section covers important implementation details, image pre-processing steps, network training strategies, and evaluation metrics used to assess model performance.

3.1 Dataset

In this study, we constructed our dataset at our research center by collecting and processing raw MRI scans. Specifically, we downloaded MRI images from 222 unique patients, which included 69 cases categorized as PI-RADS 3, 81 cases as PI-RADS 4, and 70 cases as PI-RADS 5. No follow-up or repeat scans from the same patients were included, ensuring that each scan corresponds to a distinct patient to prevent data leakage during training and evaluation.

These images were originally acquired at the CDI (Department of Diagnostic Imaging and Stereotactic Radiosurgery, Centro Diagnostico Italiano, Milan, Italy). To ensure patient confidentiality and compliance with ethical guidelines, we anonymized all MRI scans before further processing.

For all 222 cases, all three imaging sequences, T2W, ADC, and DWI, were available, as the foundation of our research relies on training networks using the complete set of sequences. we developed dedicated code using the SimpleITK library to extract and preprocess these sequences, ensuring transparency and reproducibility.

For DWI, we specifically selected images with a b-value of 1600, as this value is known to enhance lesion visibility. Our dataset originally included images with b-values of 800, 1600, and 2500. After consultation with two experienced radiologists, we chose b = 1600 because the image quality at b = 800 was generally lower, and many clinical centers do not have MRI machines capable of acquiring b = 2500 images. Therefore, b = 1600 represents

the best compromise between image quality and widespread availability across different hospitals, ensuring clinical relevance and consistency in our dataset.

For lesion annotation, a radiologist initially trained a PhD student in identifying and delineating lesions across all three sequences (T2W, ADC, and DWI). Following this preliminary annotation, two expert radiologists independently reviewed, corrected and finalized the segmentations to ensure accuracy and consistency. The finalized segmentations were then utilized as ground truth for lesion segmentation in our study.

3.2 Proposed neural networks

In our study, we employed four different deep learning architectures: U-Net, nnU-Net, SegResU-Net, and DenseNet, for prostate lesion segmentation. Each of these networks offers unique advantages in capturing spatial and contextual information, contributing to improved segmentation performance. We trained the networks using parallel T2W, ADC, and DWI sequences to capture the distinct features provided by each modality. As illustrated in Fig. 1, the image depicts the overall pipeline of the network training process, providing an overview of the data flow and model architecture. By using this U-shaped architecture, the network can effectively utilize both spatial and contextual information across the different MRI sequences (T2W, ADC, and DWI), contributing to enhanced performance in prostate lesion segmentation.

Our dataset provides separate masks for each image; however, for training, we combined these three masks into a single mask and trained the networks using this combined annotation. In the following sections, we provide a detailed description of each architecture, highlighting their structures and contributions to prostate lesion segmentation. Additionally, we describe the method used to combine the masks for training.

3.2.1 U-Net

The U-Net model is mainly composed of two parts: an encoder and a decoder. Its core idea is to reduce the size of an image to fit a display area, effectively creating smaller versions (thumbnails) that represent the original image. This reduction process helps in extracting deeper features from the image. To combine both high and low-level features, a skip-connection is added between the encoder and decoder. In the encoding phase, features

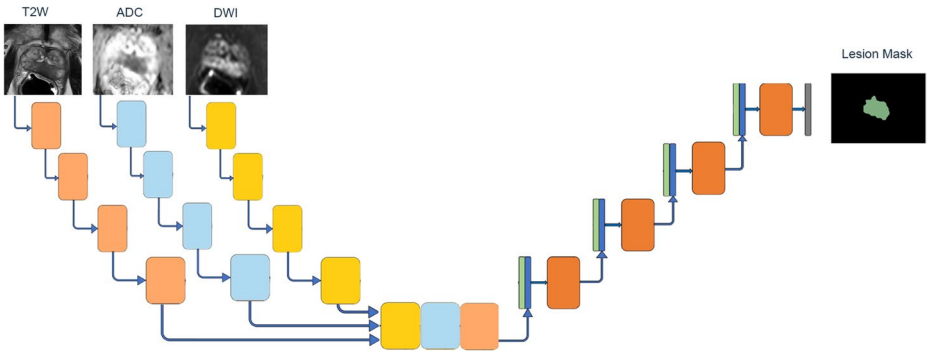


Fig. 1 The overall architecture of multi-encoder U-Net for prostate lesion segmentation

at different levels are extracted from the input image through convolution layers, with filters progressively learning these features. Then, the decoder uses up-sampling to create the segmentation mask. Feature detection in the segmented image depends on training the filters, down-sampling/up-sampling layers, and skip connections. Additionally, the backbone structure plays a key role in organizing the layers during the encoding and decoding stages [31].

3.2.2 nnU-Net

In this study, we used nnU-Net, a deep learning-based segmentation framework that automatically configures itself for new tasks, handling preprocessing, network architecture, training, and post-processing. Its design is based on a combination of fixed parameters, interdependent rules, and empirical decisions. Without requiring manual adjustments, nnU-Net outperforms many existing methods, including highly specialized models, across 23 public datasets from international biomedical segmentation challenges. It is publicly available as a ready-to-use tool, making state-of-the-art segmentation accessible. We leveraged its multi-input capability by feeding T2W, DWI, and ADC images along with a combined mask from these three sequences for prostate lesion segmentation [32].

3.2.3 Dense-UNet

Dense-UNet is formed by integrating U-Net with dense concatenation. It features a dense downsampling path and a dense upsampling path, which are symmetrically structured. Additionally, several skip connection channels are used to link these two paths. In the dense downsampling path, convolution operations are used to capture semantic contextual features across multiple scales. To overcome the depth limitation of the traditional U-Net architecture, we replaced standard pooling and convolution operations with dense_block and transition_block operations. This modification increases the network depth while maintaining efficient feature extraction, leading to what we refer to as the dense downsampling path. Each layer within a dense_block is directly connected to all its preceding layers in a feed-forward manner, promoting extensive feature reuse. Specifically, each layer receives the output feature maps from all previous layers, a process known as dense concatenation. To facilitate this, the feature maps within each layer must maintain a consistent size, ensuring seamless integration of dense blocks throughout the network [33].

3.2.4 SegRes-UNet

SegNet is a deep neural network (DNN) featuring an encoder-decoder structure with three layers. Its encoder is based on the convolutional layers of the VGG16 network. To generate the segmentation mask, the decoder reuses the pooling indices obtained from the max-pooling operation in the corresponding encoder layers. The fully connected layers were removed to simplify the model, reducing the number of parameters in the encoder section and improving efficiency [34]. Residual blocks are integral components in deep learning architectures, particularly within Residual Networks (ResNets). They address challenges associated with training deep neural networks, such as vanishing gradients, by introducing shortcut connections that allow the network to learn residual functions relative to the layer

inputs. This design facilitates the training of substantially deeper networks without degradation in performance [35].

In our approach, we have integrated residual blocks into the SegNet architecture. By incorporating these blocks, the network benefits from improved gradient flow and enhanced feature learning capabilities, leading to more accurate and efficient segmentation outcomes [36].

3.3 Combined mask

To generate a unified segmentation mask, we combined the individual masks from the T2W, ADC, and DWI sequences. Specifically, we applied a pixel-wise logical OR operation across the three masks, ensuring that a voxel was considered part of the segmented region if it was labeled in at least one of the modalities. This approach captures the union of detected regions across different MRI sequences, leveraging complementary information from each modality. The resulting mask was then converted to a float32 format for compatibility with the deep learning model. Figure 2 illustrates several sample masks of T2W, ADC, and DWI sequences, as well as their combined mask.

3.4 Implementation details

In this study, we employed four deep learning architectures to improve prostate lesion segmentation using multiparametric MRI (mpMRI) data: U-Net, SegResU-Net, DenseUNet, and nnU-Net. All four models are well-known for their effectiveness in medical image segmentation. Unlike the other three models, nnU-Net is an automated deep learning framework

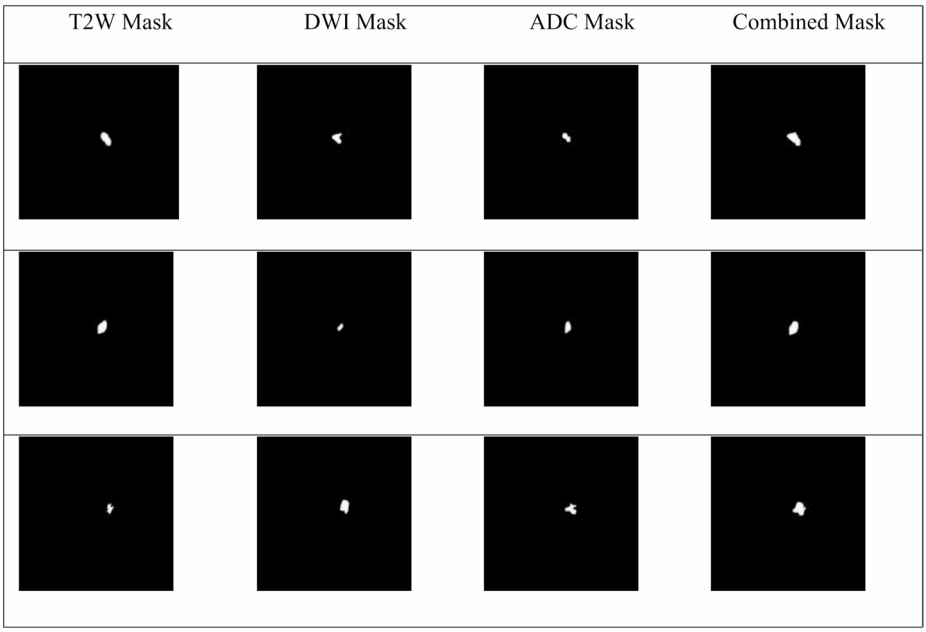


Fig. 2 Some samples of T2W, ADC, and DWI lesion ground truth and their combined mask

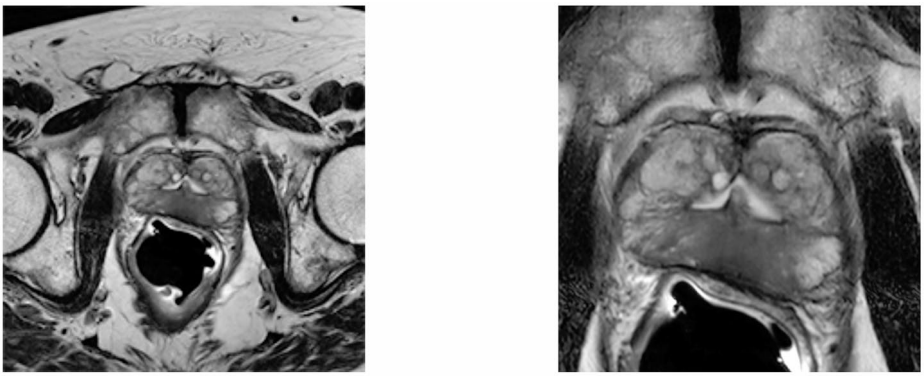
that dynamically optimizes its network architecture, training strategy, and preprocessing pipeline based on the specific properties of the dataset. For nnU-Net, we used the default configuration to train the network in our research.

Image pre-processing For preprocessing, images were resized to 256 × 256 pixels and center-cropped to 128 × 128 pixels to focus the models on the prostate region, where lesions are typically located. This step helped reduce the influence of background noise, as the prostate is generally located in the center of the MRI scans [37]. Figure 3 demonstrates the process of cropping the images on a sample of T2W. Data augmentation methods, including rotation, flipping, and zooming, were applied to further diversify the training data.

To align and register T2W, ADC, and DWI sequences, we employed the SimpleITK library in Python. Due to differences in resolution, field of view, and potential patient motion, these MRI sequences often require preprocessing and transformation to ensure proper alignment [38]. Initially, all images were preprocessed by converting them to a common format and standardizing their intensity ranges [39]. The registration process involved selecting one sequence, typically the T2W image, as the fixed reference and aligning the ADC and DWI images to it. We applied rigid and affine transformations using SimpleITK's registration framework, optimizing the transformation parameters through mutual information as the similarity metric. To refine the alignment, a multi-resolution approach was implemented, where registration was performed at different image scales to improve robustness [40]. Finally, the registered images were resampled to match the reference image's spacing and dimensions, ensuring spatial consistency. This automated approach effectively corrected misalignments, facilitating accurate image analysis and segmentation.

To account for intensity variations between the different MRI sequences, we used different normalization techniques: min-max normalization for T2W and ADC images to scale pixel intensities to a range of 0–1, and Z-score normalization for DWI images to standardize intensity values.

Networks training The dataset was divided based on PI-RADS scores, which categorize prostate lesions by their likelihood of malignancy. In one configuration, only PI-RADS 4



(a) (b)

Fig. 3 A T2W sample of the cropped image: (**a**) the original slice, and (**b**) cropped slices with focus on the prostate area

and 5 images, representing higher suspicion of cancer, were used for training and testing. Specifically, 127 images were used for training and 24 images for testing. In another configuration, images from PI-RADS 3, 4, and 5, representing a broader range of suspicion, were included. In this case, 177 images were used for training and 32 images for testing.

We ensured that all PI-RADS categories were represented in both the training and test sets. The test set consisted of unseen data that was never used during training. For training, we employed cross-validation, which internally manages the validation splits for hyperparameter tuning and early stopping. This approach allowed us to evaluate the models on both high- and low-suspicion lesions, improving the generalizability of the models for real-world clinical use.

This ensured that the models were not biased by differences in intensity distributions across sequences. The Focal Tversky loss was used for training, as it addresses class imbalance and enhances sensitivity to smaller lesions, which are critical for early prostate cancer detection [41].

The models were trained with a learning rate of 1e-4 for 100 epochs, using a batch size of 16 and early stopping to prevent overfitting. For early stopping, we applied the following criteria: the validation loss was monitored during training, with a patience of 5 epochs, starting from epoch 50. Training was halted if no improvement in validation loss was observed within these 5 epochs, and the model weights were restored to those corresponding to the best validation loss.

This strategy enabled the models to effectively capture the features essential for accurate prostate lesion segmentation.

4 Results

We trained all of the proposed networks using two different approaches.

In the first experiment, we used only MRI scans with PI-RADS 4 and 5, where 127 images were allocated for training and 24 for testing. In the second experiment, the dataset was expanded to include PI-RADS 3 cases in addition to PI-RADS 4 and 5, resulting in 177 images for training and 45 for testing. A 5-fold cross-validation strategy was employed consistently across both experiments to ensure reliable performance assessment. All networks were trained using the same hyperparameters in both experiments to maintain evaluation consistency.

The increase in training and testing sample sizes between the two experiments reflects the inclusion of PI-RADS 3 images, which increased the overall dataset size. The test sets were kept separate and mutually exclusive from training data in both experiments.

In addition, two proposed training setups, allowed us to analyze the impact of including lower PI-RADS scores on the performance of the segmentation model. To evaluate our segmentation networks, we employed two distinct approaches on the test set to assess their performance more thoroughly.

In the first approach, we calculated the DSC on a patient-by-patient basis. This involved evaluating each patient's segmentation independently, where the DSC was computed for each slice of the patient's images. We then reported the mean and median DSC values for all patients. In the second step, we performed predictions across all slices of all images in

Table 2 The results of lesion segmentation using proposed networks for PI-RADS 4 and 5

	Patient-by-Patient		Overall	
Neural Network	DSC (Median)	DSC (Mean)	DSC (Median)	DSC (Mean)
nnU-Net	**71%**	**60%**	**66%**	50%
Dense-Unet	51%	51%	49%	**57%**
U-Net	45%	44%	34%	34%
SegRes-UNet	43%	45%	45%	49%

the test set and calculated the mean and median DSC for the entire dataset. The patient-by-patient approach yielded lower results due to variability within each patient's dataset, such as differences in MRI scan quality, contrast, and lesion sizes. This caused the model to struggle with variations in individual scans, leading to lower DSC values. In contrast, the second approach, which aggregates predictions across all slices from all patients, benefits from the averaging effect. This smooths out discrepancies and allows the model's performance on higher-quality scans to compensate for lower-quality scans, resulting in better overall performance.

Table 2 presents the median and mean DSC for lesions classified as PI-RADS 4 and 5, evaluating the segmentation performance of the developed networks specifically for these higher-risk categories. Meanwhile, Table 3 expands the analysis by including PI-RADS 3, 4, and 5 lesions, reporting the median and mean DSC across all developed networks. This comparison highlights how the segmentation performance varies when incorporating lower-risk lesions (PI-RADS 3) alongside the more clinically significant PI-RADS 4 and 5 cases.

The results show that nnU-Net consistently outperformed the other models in both experiments and evaluation approaches, reflecting its adaptive architecture's ability to handle complex lesion features. Performance decreased across all models when PI-RADS 3 cases were included, likely due to the increased difficulty of segmenting smaller or less conspicuous lesions. Dense-Unet and SegRes-UNet showed moderate performance with some variability between patient-wise and overall metrics, while the standard U-Net lagged behind, indicating its relatively limited capacity for this task. The differences between patient-by-patient and overall DSC suggest that segmentation quality varies at both individual and aggregate slice levels, particularly for the more challenging dataset including PI-RADS 3.

In Fig. 4, we present the results of lesion segmentation using the nnU-Net neural network on a single slice from each of the 22 test images. Initially, we tested the model on 24 images; however, the network successfully generated segmentations for 22 of them, failing to predict segmentations for the remaining two.

In the visualized results, green represents the actual ground truth of the lesion, while red indicates the predicted lesion by the nnU-Net. The Fig. 4, provides an overview of the network's performance in identifying lesions across different test cases, illustrating its strengths and limitations in segmentation accuracy.

Table 3 The results of lesion segmentation using proposed networks for PI-RADS 3,4 and 5

	Patient-by-Patient		Overall	
Neural Network	DSC (Median)	DSC (Mean)	DSC (Median)	DSC (Mean)
nnU-Net	61%	50%	55%	50%
Dense-Unet	37%	39%	46%	46%
U-Net	30%	28%	27%	26%
SegRes-UNet	33%	35%	44%	44%

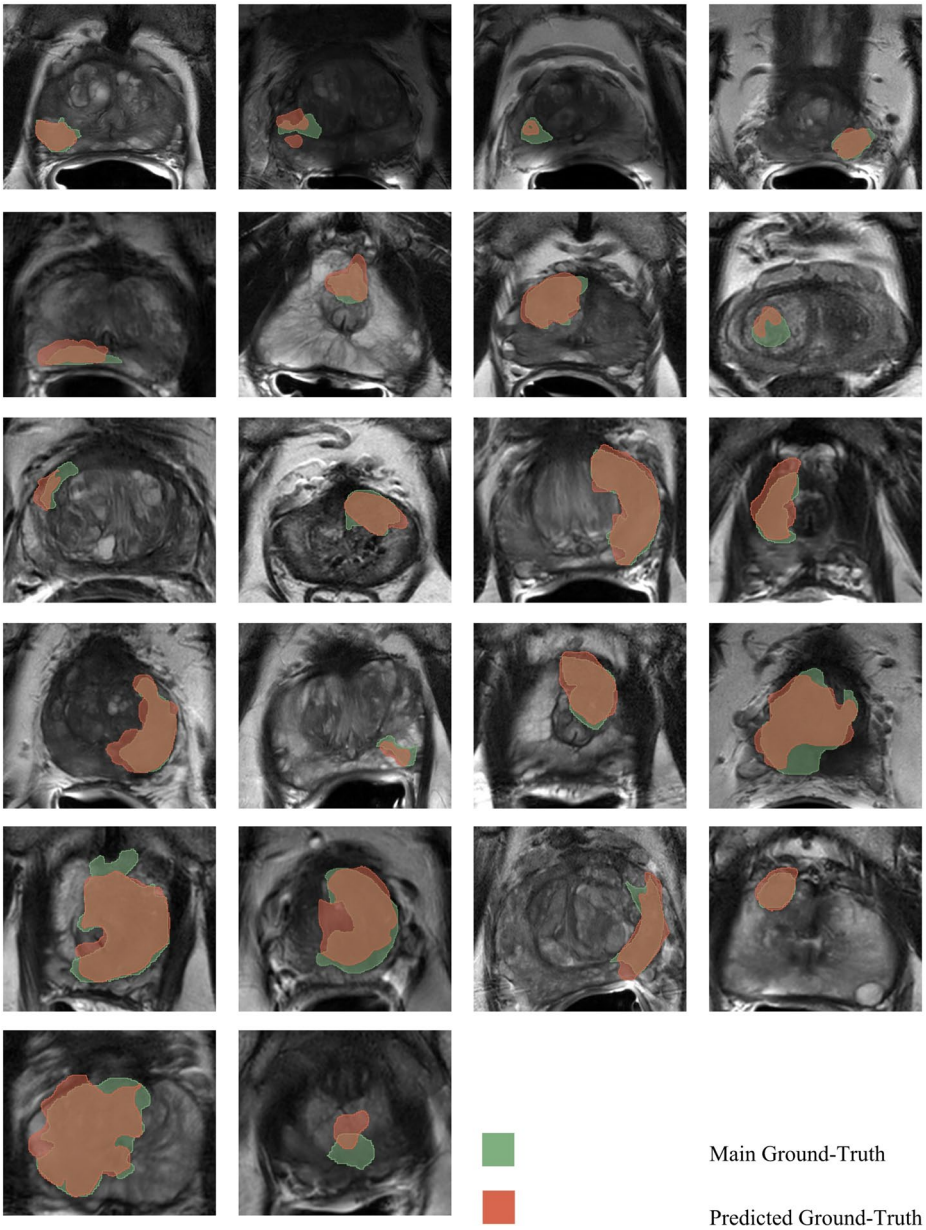


Fig. 4 Lesion segmentation results of nnU-Net on a single slice from each of the 22 test images

5 Discussion

Lesion segmentation in prostate MRI plays a crucial role in the early detection and diagnosis of clinically significant prostate cancer. Prostate lesion segmentation is considerably more challenging than prostate zonal segmentation [42, 43]. As a preliminary step to lesion

detection, zonal segmentation can achieve very high performance, even when using a single MRI sequence, typically T2W. Given the heterogeneity of prostate lesions, multi-parametric MRI (mpMRI) is widely used to enhance lesion characterization. In this study, we leveraged three essential MRI sequences: T2W, ADC, and DWI to improve segmentation accuracy. These sequences provide complementary anatomical and functional information, enhancing the differentiation between cancerous and non- cancerous tissues.

To our knowledge, there is no publicly available dataset that provides such a comprehensive combination of multiple MRI sequences alongside lesion annotations specifically associated with PI-RADS scores. Most public datasets tend to include only one or two imaging modalities, commonly T2W and/or ADC, and often omit DWI sequences. Additionally, these datasets frequently lack detailed lesion annotations categorized by PI-RADS, limiting their applicability for studies focused on standardized risk assessment.

The dataset was carefully annotated, ensuring high-quality labels for model training and evaluation. To assess the effectiveness of deep learning in this task, we employed four neural networks: nnU-Net, U-Net, SegResU-Net, and DenseU-Net. Each of these architectures has demonstrated strong performance in medical image segmentation tasks, with nnU-Net being an adaptive framework, U-Net being a widely used baseline, SegResU-Net incorporating residual connections for improved feature propagation, and DenseU-Net leveraging dense connections to enhance information flow.

To evaluate model performance, we trained all networks using two different approaches. In the first approach, the training set included only images from PI-RADS 4 and 5 lesions, representing.

clinically significant cases, with 127 images used for training and 24 images for testing. In the second approach, we expanded the training set to include PI-RADS 3, 4, and 5 lesions, introducing a broader range of lesion appearances, with 177 images for training and 32 images for testing. This comparative analysis allowed us to investigate the impact of including PI-RADS 3 cases, which often present greater diagnostic uncertainty. By evaluating the segmentation performance across both approaches, we aimed to determine the most effective strategy for training deep learning models in prostate lesion segmentation.

Unlike previous studies that often focus on single modalities or do not specify the PI-RADS classification, the present work provides a detailed analysis using multi-parametric MRI sequences (T2W, ADC, and DWI).

Our highest DSC of 60% (patient-by-patient, nnU-Net), achieved with a training set of 127 images, demonstrates the effectiveness of deep learning approaches for prostate lesion segmentation.

Compared to other studies, this work provides a detailed analysis using multi-parametric MRI sequences (T2W, ADC, and DWI), unlike previous studies that often focus on single modalities or do not specify the PI-RADS classification.

Our highest DSC of 60% (patient-by-patient, nnU-Net), achieved with a training set of 127 images, demonstrates the effectiveness of our approach for prostate lesion segmentation. A study by [16] reported a higher DSC of 86.3% using ADC images, it's important to note that they utilized a larger dataset of 390 MRIs. Unlike our approach of combining all sequences, they focused solely on ADC images and a "Simultaneous Estimation of Markov Random Field Parameters and Class" network, a different approach than our U-Net based architectures.

In comparison to our study, which achieved a DSC of 60% using the nnU-Net model, in [20] the authors employed a cascaded pyramid convolution module combined with the I2CVB method, training on a dataset of 151 MRIs, and reported a DSC of 79%. This indicates a more advanced architecture and potentially more effective feature extraction techniques in their approach.

Study [19] achieved a higher DSC of 77% using a U-Net on a private dataset of 149 ADC images, compared to our results. Our lower DSC is likely due to the use of multi-parametric MRI, while potentially offering richer clinical information and introducing complexity that can reduce segmentation accuracy compared to using ADC images alone and the smaller dataset size.

In comparison to the 69% DSC achieved by the authors in Article [22] using a dataset of 77 T1W MRIs and a cascaded scoring convolutional neural network, our study, which utilized all sequences (T2W, ADC, and DWI) with nnUNet, achieved a lower DSC of 60%.

Our study, using nnUNet on 127 MRIs with T2W, ADC, and DWI, achieved a DSC of 60%, lower than the 67% reported in Article [29]. They used U-Net LSTM on 151 MRIs of the I2CVB dataset with only T2W sequences. This difference may be due to dataset characteristics, the single T2W focus in their research, or model architecture efficacy.

Article [10] achieved a DSC of 63.76% on T2W images using the NaMa network with a dataset of 355 private MRIs. In contrast, our study utilized nnUNet and a smaller dataset of 127 MRIs, encompassing T2W, ADC, and DWI sequences, resulting in a DSC of 60%. Although our dataset was smaller, the difference in performance is minimal at 3.76%.

A limitation of our study is the exclusion of PI-RADS 2 lesions from the analysis, which was based on both clinical and technical reasons. PI-RADS 2 lesions are generally considered benign and are rarely reported by radiologists, as they are deemed to have a low likelihood of malignancy and typically do not require further intervention. Furthermore, PI-RADS 2 lesions are only visible on T2W MRI images and may not appear in other sequences, such as DWI and ADC maps. Since our study utilized a combination of these sequences, focusing on lesions that are visible across all modalities provided a more consistent dataset for segmentation. Additionally, the imbalance in the dataset resulting from this exclusion could influence the performance of machine learning models, suggesting that future work may need to address this imbalance through techniques like oversampling or under-sampling.

6 Conclusion

We studied the performance of some U-Net based models in prostate lesion segmentation using multi-sequence (T2W, DWI, ADC) MRI data. The results underscore the sensitivity of model performance to the composition of the training dataset, particularly regarding the inclusion of lower-grade lesions. While promising segmentation was achieved, further refinement is necessary to enhance accuracy and generalization capabilities. Future research should focus on expanding the dataset to encompass a wider spectrum of lesion grades and patient demographics. Moreover, exploring advanced training methodologies and network architectures, such as federated learning [44] holds promise for achieving improved segmentation performance across all lesion categories. These efforts will contribute to more

accurate diagnosis, treatment planning, and ultimately, improved patient outcomes in prostate cancer management.

Author contributions S.F., L.D.P.,F.D. and M.A: conceived the study, planned the experimentsS.F., L.D.P., and M.A: wrote the manuscriptS.F. and L.D.P.: implemented the code. S.F., F.D., R.D.M. and D .F.: contributed to data preparation. S.F., A.M., S.P., G.G., E.D., F.D. and M.A.: revised the initial draft and contributed to the final version of the submitted manuscriptF.D., G.G. and E.D.: contributed to supervision

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement. This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the CISC project (Marie Skłodowska-Curie grant agreement no. 955901 https://www.ciscproject.eu/, accessed on 18 March 2025). This work was partially supported by the MUSA (Multilayered Urban Sustainability Action) project, funded by the European UnionNextGenerationEU, under the Mission 4 Component 2 Investment Line of the National Recovery and Resilience Plan (NRRP) Mission 4 Component 2 Investment Line 1.5: Strengthening of research structures and creation of R\&D "innovation ecosystems", set up of "territorial leaders in R&D" (CUP G43C22001370007, Code ECS00000037); Program "piano sostegno alla ricerca" PSR and the PSR-GSA-Linea 6; Project ReGAInS (code 2023-NAZ-0207/DIP-ECC-DISCO23), funded by the Italian University and Research Ministry, within the Excellence Departments program 2023–2027 (law 232/2016); and FAIR-Future Artificial Intelligence Research-Spoke 4-PE00000013-D53C22002380006, funded by the European Union-Next Generation EU within the project NRPP M4C2, Investment 1.,3 DD. 341, 15 March 2022.

Data availability Data will be provided upon reasonable request.

Declarations

Ethics approval Approval for this study was obtained following a comprehensive review process, ensuring that all ethical considerations and privacy concerns were adequately addressed to protect participants' rights. Approval on 20/11/2024 by the Ethical Committee "Comitato Etico Territoriale Lombardia 3", Study ID: 5105, code "PI-RADSv2", title "Predizione della malignità delle lesioni prostatiche mediante analisi AI di immagini RM multiparametriche con mezzo di contrasto".

Human ethics and consent to participate Not applicable.

Clinical trial number Not applicable.

Competing interests The authors declare no competing interests.

References

1. World Cancer Research Fund International, Prostate Cancer Statistics: Available online: https://www.wcrf.org/preventing-cancer/cancer-statistics/prostate-cancer-statistics/
2. American Cancer Society: Key Statistics for Prostate Cancer Available online: https://www.cancer.org/cancer/types/prostate-cancer/about/key-statistics.html

3. Khosravi, P., Lysandrou, M., Eljalby, M., Li, Q., Kazemi, E., Zisimopoulos, P., Sigaras, A., Brendel, M., Barnes, J., Ricketts, C., et al.: A deep learning approach to diagnostic classification of prostate cancer using Pathology-Radiology fusion. J. Magn. Reson. Imaging. **54**, 462–471 (2021). https://doi.org/10.1002/jmri.27599
4. Greenberg, J.W., Koller, C.R., Casado, C., Triche, B.L., Krane, L.S.: A narrative review of biparametric MRI (BpMRI) implementation on Screening, Detection, and the overall accuracy for prostate cancer. Ther. Adv. Urol. **14**, 17562872221096376 (2022). https://doi.org/10.1177/17562872221096377
5. Tamada, T., Kido, A., Yamamoto, A., Takeuchi, M., Miyaji, Y., Moriya, T., Sone, T.: Comparison of biparametric and multiparametric MRI for clinically significant prostate cancer detection with PI-RADS version 2.1. J. Magn. Reson. Imaging. **53**, 283–291 (2021). https://doi.org/10.1002/jmri.27283
6. Steiger, P., Thoeny, H.C., Prostate, M.R.I.: Based on PI-RADS version 2: How we review and report. Cancer Imaging Off Publ Int. Cancer Imaging Soc. **16** (2016). https://doi.org/10.1186/s40644-016-0068-2
7. de Rooij, M., Hamoen, E.H.J., Witjes, J.A., Barentsz, J.O., Rovers, M.M.: Accuracy of magnetic resonance imaging for local staging of prostate cancer: A diagnostic Meta-Analysis. Eur. Urol. **70**, 233–245 (2016). https://doi.org/10.1016/j.eururo.2015.07.029
8. Cabarrus, M.C., Westphalen, A.C.: Multiparametric magnetic resonance imaging of the Prostate-a basic tutorial. Transl Androl. Urol. **6**, 376–386 (2017). https://doi.org/10.21037/tau.2017.01.06
9. 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. **138**, 109647 (2021). https://doi.org/10.1016/j.ejrad.2021.109647
10. Meng, R., Zhang, X., Huang, S., Gu, Y., Liu, G., Wu, G., Wang, N., Sun, K., Shen, D.: NaMa: Neighbor-Aware Multi-Modal Adaptive Learning for Prostate Tumor Segmentation on Anisotropic MR Images. In Proceedings of the Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence; AAAI Press, (2024)
11. Ronneberger, O., Fischer, P., Brox, T.U.: -Net: Convolutional Networks for Biomedical Image Segmentation BT - Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015.; Navab, N., Hornegger, J., Wells, W. M., Frangi, A. F. (eds.); Springer International Publishing: Cham, 234–241. (2015)
12. Wang, W., Pan, B., Ai, Y., Li, G., Fu, Y., Liu, Y., ParaCM-PNet:: A CNN-Tokenized MLP combined parallel dual pyramid network for prostate and prostate cancer segmentation in MRI. Comput. Biol. Med. **170**, 107999 (2024). https://doi.org/10.1016/j.compbiomed.2024.107999
13. Yilmaz, E.C., Shih, J.H., Belue, M.J., Harmon, S.A., Phelps, T.E., Garcia, C., Hazen, L.A., Toubaji, A., Merino, M.J., Gurram, S., et al.: Prospective evaluation of PI-RADS version 2.1 for prostate cancer detection and investigation of multiparametric MRI–Derived markers. Radiology. **307**, e221309 (2023). https://doi.org/10.1148/radiol.221309
14. Bonaffini, P.A., De Bernardi, E., Corsi, A., Franco, P.N., Nicoletta, D., Muglia, R., Perugini, G., Roscigno, M., Occhipinti, M., Da Pozzo, L.F., et al.: Towards the Definition of Radiomic Features and Clinical Indices to Enhance the Diagnosis of Clinically Significant Cancers in PI-RADS 4 and 5 Lesions. Cancers (Basel). **15** (2023)
15. Li, W., Zheng, B., Shen, Q., Shi, X., Luo, K., Yao, Y., Li, X., Lv, S., Tao, J., Wei, Q.: Adaptive window adjustment with boundary dou loss for cascade segmentation of anatomy and lesions in prostate cancer using BpMRI. Neural Netw. **181**, 106831 (2025). https://doi.org/10.1016/j.neunet.2024.106831
16. Yan, C., Liu, F., Peng, Y., Zhao, Y., He, J., Wang, R.: 3D convolutional network with edge detection for prostate gland and tumor segmentation on T2WI and ADC. Biomed. Signal. Process. Control. **90**, 105883 (2024). https://doi.org/10.1016/j.bspc.2023.105883
17. Thipkasorn, C., Chaichulee, S., Bejrananda, T., Tubtawee, T.: Cascaded Architecture for Segmenting Prostate Cancer Lesions in Biparametric MRI. In Proceedings of the 2024 21st International Joint Conference on Computer Science and Software Engineering (JCSSE). 167–173 (2024)
18. Adams, L.C., Makowski, M.R., Engel, G., Rattunde, M., Busch, F., Asbach, P., Niehues, S.M., Vinayahalingam, S., van Ginneken, B., Litjens, G., et al.: Prostate158 - An Expert-Annotated 3T MRI dataset and algorithm for prostate cancer detection. Comput. Biol. Med. **148**, 105817 (2022). https://doi.org/10.1016/j.compbiomed.2022.105817
19. Hong, S., Kim, S.H., Yoo, B., Kim, J.Y.: Deep Learning Algorithm for Tumor Segmentation and Discrimination of Clinically Significant Cancer in Patients with Prostate Cancer. Curr. Oncol. **30**, 7275–7285 (2023)
20. Liu, Y., Zhu, Y., Wang, W., Zheng, B., Qin, X., Wang, P.: Multi-Scale discriminative network for prostate cancer lesion segmentation in multiparametric MR images. Med. Phys. **49**, 7001–7015 (2022). https://doi.org/10.1002/mp.15861

21. Alkadi, R., Taher, F., El-Baz, A., Werghi, N.A.: Deep Learning-Based approach for the detection and localization of prostate cancer in T2 magnetic resonance images. J. Digit. Imaging. **32**, 793–807 (2019). https://doi.org/10.1007/s10278-018-0160-1
22. Eidex, Z.A., Wang, T., Lei, Y., Axente, M., Akin-Akintayo, O.O., Ojo, O.A.A., Akintayo, A.A., Roper, J., Bradley, J.D., Liu, T., et al.: MRI-Based prostate and dominant lesion segmentation using cascaded scoring convolutional neural network. Med. Phys. **49**, 5216–5224 (2022). https://doi.org/10.1002/mp.15687
23. Dai, Z., Carver, E., Liu, C., Lee, J., Feldman, A., Zong, W., Pantelic, M., Elshaikh, M., Wen, N.: Segmentation of the prostatic gland and the intraprostatic lesions on multiparametic magnetic resonance imaging using mask Region-Based convolutional neural networks. Adv. Radiat. Oncol. **5**, 473–481 (2020). https://doi.org/10.1016/j.adro.2020.01.005
24. Kohl, S., Bonekamp, D., Schlemmer, H.-P., Yaqubi, K., Hohenfellner, M., Hadaschik, B., Radtke, J.-P., Maier-Hein, K.H.: Adversarial Networks for the Detection of Aggressive Prostate Cancer. CoRR abs/1702.0. (2017)
25. Artan, Y., Haider, M.A., Langer, D.L., Yetik, I.S.: Semi-Supervised Prostate Cancer Segmentation with Multispectral MRI. In Proceedings of the 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. 648–651 (2010)
26. Liu, X., Langer, D.L., Haider, M.A., Yang, Y., Wernick, M.N., Yetik, I.S.: Prostate cancer segmentation with simultaneous Estimation of Markov random field parameters and class. IEEE Trans. Med. Imaging. **28**, 906–915 (2009). https://doi.org/10.1109/TMI.2009.2012888
27. Dai, Z., Jambor, I., Taimen, P., Pantelic, M., Elshaikh, M., Dabaja, A., Rogers, C., Ettala, O., Boström, P.J., Aronen, H.J., et al.: Prostate Cancer Detection and Segmentation on MRI Using Non-Local Mask R-CNN with Histopathological Ground Truth. Med. Phys. **50**, 7748–7763 (2023). https://doi.org/10.1002/mp.16557
28. Alzate-Grisales, J.A., Mora-Rubio, A., García-García, F., Tabares-Soto, R., Iglesia-Vayá, M.D.: La SAM-UNETR: Clinically significant prostate cancer segmentation using transfer learning from large model. IEEE Access. **11**, 118217–118228 (2023). https://doi.org/10.1109/ACCESS.2023.3326882
29. Gavade, A.B., Nerli, R., Kanwal, N., Gavade, P.A., Pol, S.S., Rizvi, S.T.: Automated Diagnosis of Prostate Cancer Using MpMRI Images: A Deep Learning Approach for Clinical Decision Support. Computers. **12** (2023)
30. Gunashekar, D.D., Bielak, L., Hägele, L., Oerther, B., Benndorf, M., Grosu, A.-L., Brox, T., Zamboglou, C., Bock, M.: Explainable AI for CNN-Based prostate tumor segmentation in Multi-Parametric MRI correlated to whole Mount histopathology. Radiat. Oncol. **17** (2022). https://doi.org/10.1186/s13014-022-02035-0
31. Rezaeijo, S.M., Jafarpoor Nesheli, S., Fatan Serj, M., Tahmasebi Birgani, M.J.: Segmentation of the Prostate, its Zones, anterior fibromuscular Stroma, and urethra on the MRIs and multimodality image fusion using U-Net model. Quant. Imaging Med. Surg. **12**, 4786–4804 (2022). https://doi.org/10.21037/qims-22-115
32. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: NnU-Net: A Self-Configuring method for deep Learning-Based biomedical image segmentation. Nat. Methods. **18**, 203–211 (2021). https://doi.org/10.1038/s41592-020-01008-z
33. Cai, S., Tian, Y., Lui, H., Zeng, H., Wu, Y., Chen, G.: Dense-UNet: A novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. Quant. Imaging Med. Surg. **10**, 1275–1285 (2020). https://doi.org/10.21037/qims-19-1090
34. Saood, A., Hatem, I.: COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. BMC Med. Imaging. **21**, 19 (2021). https://doi.org/10.1186/s12880-020-00529-5
35. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. 2016 IEEE Conf. Comput. Vis. Pattern Recognit. 770–778 (2015)
36. Weng, L., Xu, Y., Xia, M., Zhang, Y., Liu, J., Xu, Y.: Water Areas Segmentation from Remote Sensing Images Using a Separable Residual SegNet Network. ISPRS Int. J. Geo-Information **9** (2020)
37. Zaridis, D.G., Mylona, E., Marias, K., Papanikolaou, N., Tachos, N.S., Fotiadis, D.I.: A New Smart-Cropping Pipeline for Prostate Segmentation Using Deep Learning Networks. ArXiv abs/2107.0. (2021)
38. Yaniv, Z., Lowekamp, B.C., Johnson, H.J., Beare, R.: SimpleITK Image-Analysis notebooks: A collaborative environment for education and reproducible research. J. Digit. Imaging. **31**, 290–303 (2018). https://doi.org/10.1007/s10278-017-0037-8
39. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4ITK: Improved N3 bias correction. IEEE Trans. Med. Imaging. **29**, 1310–1320 (2010). https://doi.org/10.1109/TMI.2010.2046908

40. Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C.: A reproducible evaluation of ants similarity metric performance in brain image registration. Neuroimage. **54**, 2033–2044 (2011). https://doi.org/10.1016/j.neuroimage.2010.09.025
41. Abraham, N., Khan, N.: A Novel Focal Tversky Loss Function with Improved Attention U-Net for Lesion Segmentation. (2023)
42. Fouladi, S., Di Palma, L., Darvizeh, F., Fazzini, D., Maiocchi, A., Papa, S., Gianini, G., Alì, M.: Neural network models for prostate zones segmentation in magnetic resonance imaging. Information. **16**(3), 186 (2025)
43. Fouladi, S., Gianini, G., Fazzini, D., Maiocchi, A., Damiani, E., Papa, S., Ali, M.: Advanced Prostate MRI Analysis: UNET-Based Models for Zonal and Lesion Segmentation. In Proceedings of the 16th International Conference on Management of Digital Ecosystems 2024 (pp. 174–187). Cham: Springer Nature Switzerland
44. Bovio, A., Barile, M., Pallotta, F., Pede, L., Maiocchi, A., Ali, M., Darvizeh, F., Fazzini, D., Lacavalla, F., Banzi, M., Gianini, G., Mio, C., Berto, F., Bondaruc, R., Damiani, E., Fouladi, S.: A Federated Learning Architecture for Prostate MRI Image Segmentation. To appear in Proceedings of the 4th Italian Conference on Big Data and Data Science, Torino, Italy, September 2025 (ITADATA2025)

Authors and Affiliations

Saman Fouladi[1] · Fatemeh Darvizeh[2] · Gabriele Gianini[3] · Rosario Di Meo[2] · Luca Di Palma[2] · Ernesto Damiani[1] · Alessandro Maiocchi[4] · Deborah Fazzini[2] · Marco Alì[2]

✉ Gabriele Gianini
gabriele.gianini@unimib.it

[1] Department of Computer Science, University of Milano, via Celoria 18, Milan 20133, Italy

[2] CDI Centro Diagnostico Italiano S.p.A, via Saint Bon 20, Milan 20147, Italy

[3] Department of Informatics, Systems and Communication, University of Milano-Bicocca, viale Sarca 336, Milan 20126, Italy

[4] Bracco S.p.A, Via Egidio Folli, 50, Milan 20134, Italy