

## Article

# Computational Detection of Extraprostatic Extension of Prostate Cancer on Multiparametric MRI Using Deep Learning

Ștefania L. Moroianu <sup>1,\*</sup>, Indrani Bhattacharya <sup>2,3</sup>, Arun Seetharaman <sup>4</sup>, Wei Shao <sup>2</sup>, Christian A. Kunder <sup>5</sup>, Avishkar Sharma <sup>2</sup>, Pejman Ghanouni <sup>2,3</sup>, Richard E. Fan <sup>3</sup>, Geoffrey A. Sonn <sup>2,3,†</sup>, Mirabela Rusu <sup>2,\*,†</sup>

<sup>1</sup> Department of Applied Physics, Stanford University, Stanford, CA 94305, USA

<sup>2</sup> Department of Radiology, Stanford University School of Medicine, Stanford, CA 94305, USA; ibhatt@stanford.edu (I.B.); weishao@stanford.edu (W.S.); avsharma@stanford.edu (A.S.); ghanouni@stanford.edu (P.G.); gsonn@stanford.edu (G.A.S.)

<sup>3</sup> Department of Urology, Stanford University School of Medicine, Stanford, CA 94305, USA; refan@stanford.edu

<sup>4</sup> Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA; arun.s@att.net

<sup>5</sup> Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA; ckunder@stanford.edu

\* Correspondence: slmoro@stanford.edu (Ș.L.M.); mirabela.rusu@stanford.edu (M.R.)

† These authors contributed equally to this work.

**Simple Summary:** In approximately 50% of prostate cancer patients undergoing surgical treatment, cancer has extended beyond the prostate boundary (i.e., extraprostatic extension). The aim of our study was to expand artificial intelligence (AI) models that identify cancer in the prostate to also identify the cancer that spreads outside the boundary of the prostate. By combining past models with image post-processing steps and clinical decision rules, we built an autonomous approach to detect the extension of the cancer beyond the prostate boundary using prostate MRI. Our study included 123 prostate cancer patients (38 with extraprostatic extension), and our proposed method can detect cancer outside the prostate boundary in more cases than radiologists.

**Abstract:** The localization of extraprostatic extension (EPE), i.e., local spread of prostate cancer beyond the prostate capsular boundary, is important for risk stratification and surgical planning. However, the sensitivity of EPE detection by radiologists on MRI is low (57% on average). In this paper, we propose a method for computational detection of EPE on multiparametric MRI using deep learning. Ground truth labels of cancers and EPE were obtained in 123 patients (38 with EPE) by registering pre-surgical MRI with whole-mount digital histopathology images from radical prostatectomy. Our approach has two stages. First, we trained deep learning models using the MRI as input to generate cancer probability maps both inside and outside the prostate. Second, we built an image post-processing pipeline that generates predictions for EPE location based on the cancer probability maps and clinical knowledge. We used five-fold cross-validation to train our approach using data from 74 patients and tested it using data from an independent set of 49 patients. We compared two deep learning models for cancer detection: (i) UNet and (ii) the Correlated Signature Network for Indolent and Aggressive prostate cancer detection (CorrSigNIA). The best end-to-end model for EPE detection, which we call EPENet, was based on the CorrSigNIA cancer detection model. EPENet was successful at detecting cancers with extraprostatic extension, achieving a mean area under the receiver operator characteristic curve of 0.72 at the patient-level. On the test set, EPENet had 80.0% sensitivity and 28.2% specificity at the patient-level compared to 50.0% sensitivity and 76.9% specificity for the radiologists. To account for spatial location of predictions during evaluation, we also computed results at the sextant-level, where the prostate was divided into sextants according to standard systematic 12-core biopsy procedure. At the sextant-level, EPENet achieved mean sensitivity 61.1% and mean specificity 58.3%. Our approach has the potential to provide the location of extraprostatic extension using MRI alone, thus serving as an independent diagnostic aid to radiologists and facilitating treatment planning.

**Keywords:** extraprostatic extension; computer-aided diagnosis; deep learning



**Citation:** Moroianu, Ș.L.;

Bhattacharya, I.; Seetharaman, A.; Shao, W.; Kunder, C.A.; Sharma, A.; Ghanouni, P.; Fan, R.E.; Sonn, G.A.; Rusu, M. Computational Detection of Extraprostatic Extension of Prostate Cancer on Multiparametric MRI Using Deep Learning. *Cancers* **2022**, *14*, 2821. <https://doi.org/10.3390/cancers14122821>

Academic Editors: Yoganand Balagurunathan, Kenneth Gage and Lubomir Hadjiiski

Received: 30 April 2022

Accepted: 3 June 2022

Published: 7 June 2022

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



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Prostate cancer is the second most frequent cancer and the fifth leading cause of cancer death among men worldwide [1]. It is estimated that in 2022 prostate cancer will account for 268,490 new cases and 34,500 deaths in the United States alone [2]. Radical prostatectomy is a common treatment for localized prostate cancer, yet 20% to 40% of patients experience biochemical recurrence after surgery [3]. These patients are much more likely to develop metastases [4]. Positive surgical margins caused by the extension of cancer beyond the prostate gland [5], i.e., extraprostatic extension (EPE), represent a major risk factor for biochemical recurrence and reduced cancer-specific survival after radical prostatectomy [6]. Cancers with extraprostatic extension are common [7–9], occurring in 50% of patients undergoing radical prostatectomy [10].

Detection of extraprostatic extension may improve selection of treatment (i.e., surgery or radiation) and help with surgical planning for patients who choose surgery (radical prostatectomy). If extraprostatic extension is suspected pre-operatively, surgeons may need to resect more tissue to reduce the likelihood of positive margins. Conversely, if extraprostatic extension can be confidently ruled out, surgeons will perform nerve-sparing radical prostatectomies to reduce the risk of post-operative erectile dysfunction and urinary incontinence [11]. Current guidelines from the American Urological Association recommend multiparametric MRI imaging studies for the evaluation of extraprostatic extension [12].

The ability of radiologists to detect extraprostatic extension on MRI varies greatly (sensitivity 12–83% and specificity 63–92%) [13]. One meta-analysis concluded that, with an average sensitivity of 57% [14], MRI is not sensitive enough to reliably find tumors with extraprostatic extension. One challenge is that grading systems for reporting extraprostatic extension on MRI lack standardization [15,16]. Quantitative assessment criteria such as tumor–capsule contact line length [5,12] have been shown to be independent predictors of extraprostatic extension.

Automated approaches have the potential to improve the accuracy and reliability of extraprostatic extension detection on MRI. Prior work [17–20] primarily used radiomics to predict the presence of extraprostatic extension, with one recent study employing deep learning [7] (Table 1). These prior studies have a common set of limitations: (1) They lack spatially accurate ground truth labels, since the histopathology images and the MRI are cognitively registered, a process prone to error; (2) they solve a binary classification problem i.e., predict the presence or absence of extraprostatic extension at the index lesion level, without attempting to spatially localize extraprostatic extension; (3) none of the approaches are fully automatic, as they require radiologist input usually in the form of manual lesion segmentations (see column 3 in Table 1).

In recent years, many studies have shown promising results in automatically detecting prostate cancer on MRI using convolutional neural networks [21–26]. We argue that tumor detection inside the prostate can be viewed as a sub-task for extraprostatic extension prediction, and we propose a fully automated extraprostatic extension detection framework using these existing cancer models as building blocks. The goal of our study was not to build a better deep learning model for prostate cancer, but to evaluate the utility of existing models in detecting cancer that extends beyond the capsule.

In this paper, we propose a two-step fully automated workflow for prediction and localization of extraprostatic extension on multiparametric MRI. First, we used pre-trained deep learning models to generate cancer probability maps both inside and outside the prostate gland. Second, we defined a series of post-processing steps and clinically inspired decision rules to predict the presence or absence of extraprostatic extension for each lesion.

**Table 1.** Summary of previous studies predicting presence of extraprostatic extension on multiparametric MRI. Abbreviations: EPE = extraprostatic extension; w/ = with; CNN = convolutional neural network; SVM = support vector machine; TCL = tumor–capsule contact line length (also known as capsular contact length); PI-RADS = Prostate Imaging–Reporting and Data System; ESUR = European Society of Urogenital Radiology; ADC = apparent diffusion coefficient.

First Author (Year)	Evaluation Granularity	Radiologist Input Required	Patient Number	Method	AUC
Hou (2021) [7]	Per index lesion	Tumor segmentation	849	Radiologists CNN	0.63–0.74 0.73–0.81
Cuocolo (2021) [17]	Per index lesion	Tumor segmentation	193	Radiologists Radiomics + SVM	81–83% acc 0.73–0.80 74–79% acc
Eurboonyanun (2021) [27]	Per index lesion	Measure TCL	95	Logistic regression w/ absolute TCL (euclidean) actual TCL (curvilinear)	0.80 0.74
Losnegard (2020) [28]	Per index lesion	Tumor segmentation	228	Radiologists Radiomics + Random forest	0.75 0.74
Park (2020) [29]	Per patient	Measure TCL	301	Radiologists using MRI-based EPE grade, ESUR score, Likert scale, TCL	0.77–0.81 0.79–0.81 0.78–0.79 0.78–0.85
Xu (2020) [19]	Per lesion (all those MRI visible)	Tumor segmentation	95	Radiomics + Regression algorithm	0.87
Shiradkar (2020) [18]	Per index lesion	Tumor and periprostatic fat segmentation	45	Radiomics + SVM	0.88
Mehralivand (2019) [30]	Per index lesion	Measure TCL	553	Logistic regression w/ MRI-based EPE grade + clinical features	0.77 0.81
Ma (2019) [31]	Per index lesion	Tumor segmentation	210	Radiologists Radiomics + Regression algorithm	0.60–0.70 0.88
Stanzione (2019) [20]	Per index lesion	Tumor segmentation	39	Radiomics + Bayesian Network	0.88
Krishna (2017) [32]	Per lesion (all those MRI visible)	Tumor segmentation	149	Radiologists Logistic regression w/ PI-RADS scores, tumor size, TCL, ADC entropy	0.61–0.67, 0.61–0.72, 0.73, 0.69, 0.76

## 2. Materials and Methods

### 2.1. Dataset

#### 2.1.1. Population Characteristics

This retrospective chart review study was approved by the Institutional Review Board of Stanford University. As a chart review of previously collected data, patient consent was waived. The study focused on a cohort of 123 patients with confirmed prostate cancer who had a pre-operative MRI and underwent radical prostatectomy at our institution. Patients were randomly split into a training and validation set ( $n = 74$ ) and a held-out testing set ( $n = 49$ ) (Table 2). Extraprostatic extension was present in 28 of the 74 training set patients, and 10 out of the 49 testing set patients.

**Table 2.** Details of the train and test cohorts. Lesion volumes are presented as median (IQR), where IQR = interquartile range. Lesion volumes were computed based on the histopathology labels, assuming continuity between slices and using the MRI slice thickness.

Cohort	Train	Test
Patient number	74	49
Lesion count	90	58
Indolent	9	10
Aggressive	81	48
EPE (pathologically proven)	29	10
Lesion volume (mm <sup>3</sup> )	1541.6 (714.7, 3418.6)	1099.1 (743.2, 2544.7)
EPE volume (where applicable)	8.6 (3.6, 44.6)	10.6 (5.6, 36.3)

### 2.1.2. Image Acquisition

Magnetic resonance images (MRI) were acquired using 3 Tesla GE scanners (GE Healthcare, Waukesha, WI, USA) with external 32-channel body array coils without endorectal coils. The imaging protocol included T2-weighted MRI (T2w), diffusion weighted imaging (DWI), derived Apparent Diffusion Coefficient (ADC) maps and dynamic contrast-enhanced imaging sequences. Axial T2w MRI (acquired using a 2D Spin Echo protocol) and ADC maps were used in this study (see characteristics in Table 3).

**Table 3.** Radical prostatectomy cohort data characteristics. Abbreviations: T2w—T2-weighted MRI; ADC—Apparent Diffusion Coefficient map. [*a*, *b*] indicates range between *a* and *b*.

Number of Patients	123
T2w	
Repetition time (TR, range) (s)	[3.9, 6.3]
Echo time (TE, range) (ms)	[122, 130]
Pixel size (range) (mm)	[0.27, 0.94]
Distance between slices (mm)	[3, 5.2]
Matrix size	[256, 512]
Number of slices	[20, 44]
ADC	
b-values (s/mm <sup>2</sup> )	0, 50, 800, 1000, 1200
Pixel size (range) (mm)	[0.78, 1.50]
Distance between slices (mm)	[3, 4.5]
Matrix size	[50, 256]
Number of slices	[15, 40]

The excised prostates were fixed in formalin and embedded in paraffin and then serially sectioned using customized 3D-printed molds with slice orientation and thickness matching that of T2w images, followed by staining with hematoxylin and eosin. All stained slices were scanned at 20× magnification (pixel size 0.5 μm) to generate digitized whole-mount histopathology images.

### 2.1.3. Labels

All radical prostatectomy specimens were reviewed by a genitourinary pathologist (C.A.K.) with 10 years of experience. The expert pathologist outlined cancer, including extraprostatic extension, on all whole-mount digital histopathology slices, generating pixel-level cancer and extraprostatic extension labels. By computationally registering surgical histopathology images and the corresponding MRI, we mapped the extent of extraprostatic extension and cancer onto MRI creating pixel-level labels for both classes (details in Section 2.2.1). Prostate segmentations were available on all T2w MRI and digital histopathology slices for all patients. The prostate was segmented on T2w images by expert



technologists (mean experience = 9 years) and adjusted as necessary by our expert team (C.A.K, G.S.—a urologic oncologist with 13 years of experience; P.G.—a body MR imaging radiologist with 14 years of experience; M.R.—an image analytics expert with 10 years of experience working on prostate cancer).

Additionally, all MRI in the test set were reviewed retrospectively by a radiologist (P.G. or A.S.—a body MR imaging radiologist with 5 years of experience) with the knowledge that patients had undergone prostatectomy for cancer, who indicated whether extraprostatic extension was present. These labels were used to compare our model's performance with that of radiologists in detecting the presence of extraprostatic extension on multiparametric MRI.

## 2.2. Data Pre-Processing

### 2.2.1. Histopathology Pre-Processing

Several pre-processing steps were applied to the digitized histopathology images:

1. **Registration:** Each digital histopathology image was aligned with its corresponding T2w MR image using the automated affine and deformable registration method RAPSODI [33]. This enabled accurate mapping of pixel-level cancer and extraprostatic extension labels from digital histopathology images onto MRI. For details on this process, refer to [26,33].
2. **Smoothing:** Images were smoothed with a Gaussian filter with  $\sigma = 0.25$  mm to avoid downsampling artifacts.
3. **Resampling:** The Gaussian smoothed images were downsampled to an X-Y size of  $224 \times 224$  pixels, resulting in an in-plane pixel size of  $0.29 \times 0.29$  mm<sup>2</sup>.
4. **Intensity normalization:** Each RGB channel of the resulting digital histopathology images was Z-score normalized.

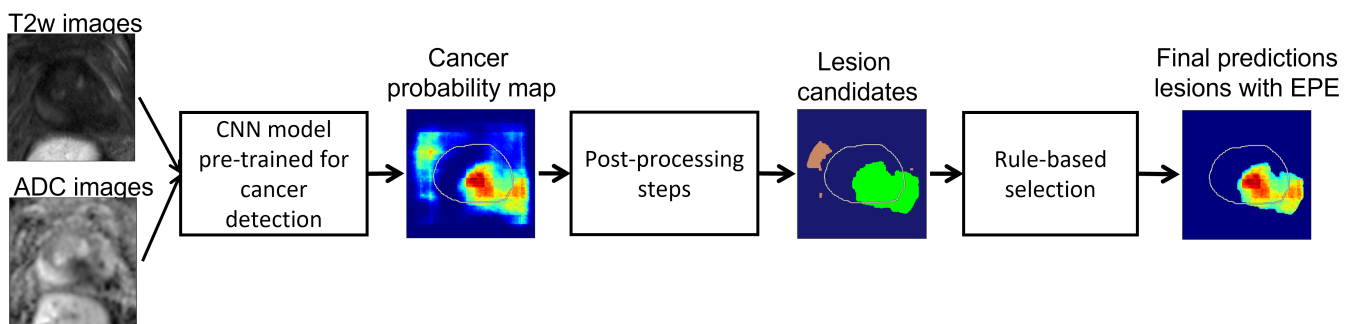
### 2.2.2. MRI Pre-Processing

Several pre-processing steps were applied to the MR images, following the procedure in [25,26]:

1. **Affine Registration:** The T2w images and ADC images were manually registered using an affine transformation driven by the prostate segmentations on both modalities.
2. **Resampling:** The T2w images, ADC images, prostate masks and cancer labels were projected and resampled on the corresponding histopathology images, resulting in images of  $224 \times 224$  pixels, with pixel size of  $0.29 \times 0.29$  mm<sup>2</sup>.
3. **Intensity standardization:** We followed the procedure by Nyul et al. [34]. Using the training dataset, we learned a set of intensity histogram landmarks for T2w and ADC sequences independently. Then, we transformed the image histograms to align with the learned mean histogram of each MRI sequence. The histogram average learned in the training set was also used to align the cases in the test set. This histogram alignment intensity standardization method helps ensure similar MRI intensity distribution for all patients irrespective of scanners and scanning protocols.
4. **Intensity normalization:** Finally, Z-score normalization was applied to the prostate regions of T2w and ADC images.

## 2.3. Proposed Approach

Our approach consists of two stages (Figure 1). First, we used deep learning models pre-trained for cancer detection to generate cancer probability maps over the entire image. Second, we defined a set of post-processing steps and heuristic rules based on clinical knowledge in order to generate localized predictions for lesions with extraprostatic extension from the cancer probability maps.



**Figure 1.** Flowchart describing our method for generating spatial predictions of cancer lesions with extraprostatic extension. A convolutional neural network takes in T2w and ADC images and outputs a probability map for cancer. We apply post-processing steps of masking, thresholding and connected components to the probability map to obtain a set of lesion candidates. We check each lesion candidate against a set of heuristic rules and determine which are the lesions with suspected extraprostatic extension.

### 2.3.1. Step 1: Deep Learning Models for Cancer Detection

Many studies focused on cancer detection inside the prostate using convolutional neural networks. The goal of our study was to evaluate the utility of these existing models in detecting cancer that has spread beyond the prostate capsule. Specifically, we implemented two networks: UNet [35] and the state-of-the-art Correlated Signature Network for Indolent and Aggressive prostate cancer detection (CorrSigNIA) [26]. Both models were pre-trained for cancer detection inside the prostate gland using a five-fold cross-validation scheme. Details on the choice and training of these models can be found in Appendix A.

At inference time, given input images of T2w and ADC axial slices  $\mathcal{I}_{T2}(x, y)$ ,  $\mathcal{I}_{ADC}(x, y)$  (Figure 2a,b), the pre-trained model outputs a cancer probability map for the 2D slice,  $p_{Ca}(x, y)$  (Figure 2d):

$$p_{Ca}(x, y) = f_{\theta}(\mathcal{I}_{T2}, \mathcal{I}_{ADC}), \quad (1)$$

where  $x$  and  $y$  are the pixel coordinates in the left–right and posterior–anterior directions, respectively; function  $f_{\theta}$  is the convolutional neural network with learned weights  $\theta$ .

### 2.3.2. Step 2: Post-Processing Pipeline

We apply several post-processing steps to the cancer probability maps output by the deep learning models in order to generate predictions for lesions with extraprostatic extension (Figure 2):

1. **Dilated prostate mask.** The deep learning cancer predictions become less reliable the further we look outside the prostate, since other anatomical features may drive false positives. To prevent this, we applied a dilated prostate mask to the cancer probability map. Based on the diameter of the largest extraprostatic extension lesions in our cohort, we chose to dilate the original prostate mask using kernels of size  $64 \times 64$  pixels (corresponding to  $1.86 \text{ cm} \times 1.86 \text{ cm}$ ):

$$\mathcal{M}_{pr}(x, y) = \text{dilate}(\mathcal{L}_{pr}(x, y)), \quad (2)$$

where  $\mathcal{L}_{pr}(x, y)$  represents the prostate segmentation mask (pixels have value 1 inside the prostate, value 0 outside) and  $\mathcal{M}_{pr}(x, y)$  is the resulting dilated prostate mask. We set all values in the probability map outside this region to zero:

$$p(x, y) = p_{Ca}(x, y) * \mathcal{M}_{pr}(x, y), \quad (3)$$

where  $p_{Ca}(x, y)$  is the cancer probability map output by the model (Equation (1)) (Figure 2d);  $*$  denotes element-wise multiplication; the resulting  $p(x, y)$  is the 2D masked cancer probability map for a given slice (Figure 2e).

2. **Binary threshold.** All pixels in the prediction map with probability  $p(x, y)$  greater than a fixed threshold,  $\alpha$ , were considered to be cancer, and the rest were set to zero;  $\alpha$  is a hyperparameter:

$$p_{\alpha}(x, y) = \begin{cases} p(x, y) & \text{if } p(x, y) > \alpha \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

and this results in a thresholded cancer probability map  $p_{\alpha}(x, y)$  for the slice. We computed  $p_{\alpha}(x, y)$  for all slices in a case, resulting in a volume  $P_{\alpha}(x, y, z)$  for the patient, where  $z$  is the slice index.

3. **Connected components.** Next, we computed all 3D connected components in the  $P_{\alpha}(x, y, z)$  volume with connectivity value 26 using the python cc3d library [36]:

$$\{\mathcal{C}_{\alpha}^{(i)}(x, y, z)\} = \text{cc3d.connected\_components}(P_{\alpha}(x, y, z)). \quad (5)$$

Each component  $\mathcal{C}_{\alpha}^{(i)}(x, y, z)$  is a lesion candidate (Figure 2f). Note that the connected components function returns binary mask objects, i.e., each pixel in  $\mathcal{C}_{\alpha}^{(i)}(x, y, z)$  is either 0 or 1, and all the pixels with value 1 are connected.

4. **Logical rules:** We used logical rules to prune these components and determine the final predictions for extraprostatic extension:
  - **Rule I:** Component must predict cancer both inside:

$$\mathcal{C}_{\alpha}^{(i)} \cap \mathcal{L}_{pr} \neq 0, \quad (6)$$

and outside the prostate capsule

$$\mathcal{C}_{\alpha}^{(i)} \setminus \mathcal{L}_{pr} \neq 0, \quad (7)$$

where  $\mathcal{L}_{pr}$  is the prostate binary mask, i.e.,  $\mathcal{L}_{pr}(x, y, z) = 1$  for pixels inside the prostate and  $\mathcal{L}_{pr}(x, y, z) = 0$  for pixels outside the prostate. If a component is either fully inside ( $\mathcal{C}_{\alpha}^{(i)} \setminus \mathcal{L}_{pr} = 0$ ) or fully outside the prostate ( $\mathcal{C}_{\alpha}^{(i)} \cap \mathcal{L}_{pr} = 0$ ), it is not a viable candidate for extraprostatic extension (Figure 2g, and brown components were discarded because they are fully outside the prostate and the green component was accepted since it crosses the prostate border).

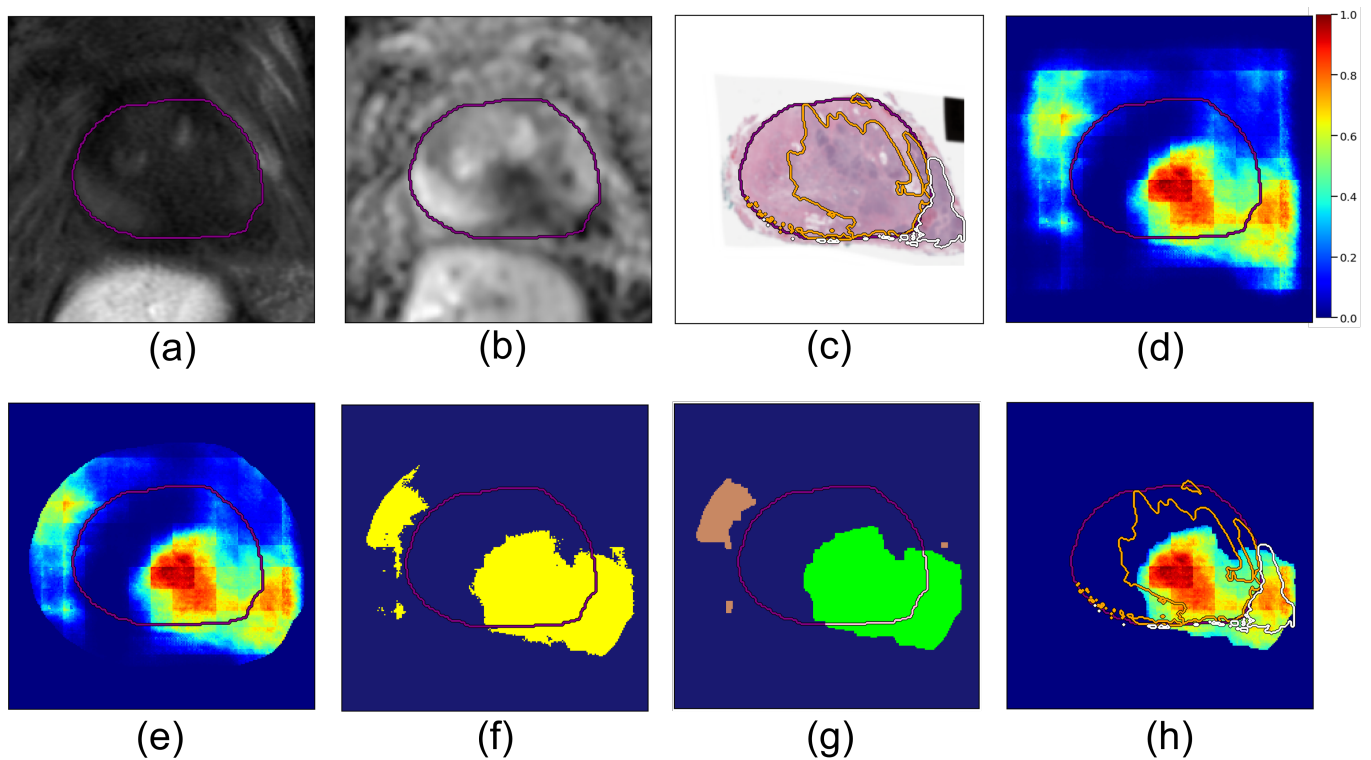
- **Rule II:** For each viable lesion candidate, compute tumor–capsule contact line length (TCL) and compare with threshold  $th_{TCL}$ . The overlap between a candidate  $\mathcal{C}_{\alpha}^{(i)}$  and the prostate boundary  $\mathcal{L}_{pr}$  defines a curvilinear segment (shown in pink in Figure 2g), and  $l_{TCL}^{(i)}$  is the length of this segment.

$$l_{TCL}^{(i)} = \text{tumor\_capsule\_contact\_line\_length}(\mathcal{C}_{\alpha}^{(i)}, \mathcal{L}_{pr}). \quad (8)$$

Lesion candidates with  $l_{TCL}^{(i)} < th_{TCL}$  are discarded. Candidates with  $l_{TCL}^{(i)} \geq th_{TCL}$  constitute our final predictions for cancer lesions with extraprostatic extension. Each final candidate  $\mathcal{C}_{\alpha}^{(i)}$  is a binary mask; multiplying it element-wise with the probability map  $P_{\alpha}$  gives the probability map for cancer with extraprostatic extension (Figure 2h).

$$q^{(i)}(x, y, z) = \mathcal{C}_{\alpha}^{(i)}(x, y, z) * P_{\alpha}(x, y, z). \quad (9)$$

We denote the final extraprostatic extension probability map for the entire case volume  $\mathcal{Q}(x, y, z)$ .



**Figure 2. Row 1:** Registered MRI and histopathology slices, along with output from cancer detection deep learning model in a patient with bulky extraprostatic extension. (a) T2w image, (b) ADC image, (c) Histopathology image overlaid with ground truth cancer labels and prostate segmentation mask. Purple contour shows prostate gland segmentation, orange contour is cancer inside the prostate, white contour is extraprostatic extension (EPE). (d) Cancer probability map output by the pre-trained deep learning cancer detection model. **Row 2:** Post-processing steps to generate final predictions for cancer with extraprostatic extension. (e) Cancer probability map after applying dilated prostate mask. (f) Candidate lesions (shown in yellow) were obtained by applying binary threshold and detecting connected components. (g) Candidate lesions were pruned based on their location relative to the prostate; components fully inside or fully outside the prostate were rejected (brown), those crossing the border were accepted (green). For the green lesion candidate, the tumor-capsule contact line (TCL) is displayed in pink. (h) Final prediction map for lesion with EPE.

Algorithm 1 presents all these steps in pseudo-code.

We refer to our proposed end-to-end model that includes pre-trained CorrSigNIA + post-processing steps + decision rules as EPENet, and we refer to the baseline end-to-end model pre-trained UNet + post-processing steps + decision rules as UNet\_EPE. Both models take as input the T2w and ADC images and output the extraprostatic extension probability map volume  $Q(x, y, z)$  for the patient.

**Algorithm 1** Steps for predicting lesions with extraprostatic extension**Require:**  $\mathcal{I}_{T2}, \mathcal{I}_{ADC}, \mathcal{L}_{pr}$ **Require:** user defined parameters  $\alpha$ ,  $TCL_{th}$ **Ensure:** load pre-trained cancer model  $f_{\theta}$ 

```

 $p_{Ca} \leftarrow f_{\theta}(\mathcal{I}_{T2}, \mathcal{I}_{ADC})$  ▷ deep learning model inference, Equation (1)

 $\mathcal{M}_{pr} \leftarrow \text{dilate}(\mathcal{L}_{pr})$  ▷ Equation (2)
 $p \leftarrow p_{Ca} * \mathcal{M}_{pr}$  ▷ apply dilated prostate mask, Equation (3)
 $p_{\alpha} \leftarrow \text{binary threshold}(p; \alpha)$  ▷ Equation (4)
 $\mathcal{C}_{\alpha} \leftarrow \text{connected components}(p_{\alpha})$  ▷ Equation (5)
 $n \leftarrow \text{len}(\mathcal{C}_{\alpha})$ 

 $q \leftarrow \text{empty list}$  ▷ where final EPE lesion predictions will be appended
for  $i$  in  $0, 1, \dots, n - 1$  do ▷ iterate over all candidates  $\mathcal{C}_{\alpha}^{(i)}$ 
  if  $\mathcal{C}_{\alpha}^{(i)} \cap \mathcal{L}_{pr} \neq 0$  and  $\mathcal{C}_{\alpha}^{(i)} \setminus \mathcal{L}_{pr} \neq 0$  then ▷ Equations (6) and (7)
     $l_{TCL}^{(i)} \leftarrow \text{compute tumor-capsule contact line length}(\mathcal{C}_{\alpha}^{(i)}, \mathcal{L}_{pr})$  ▷ Equation (8)
    if  $l_{TCL}^{(i)} > TCL_{th}$  then
       $q \leftarrow \text{append}(\mathcal{C}_{\alpha}^{(i)} * p_{\alpha})$  ▷ component  $\mathcal{C}_{\alpha}^{(i)}$  satisfied all criteria, Equation (9)
      ▷ append to list of final EPE predictions
    end if
  else continue
  end if
end for
 $Q \leftarrow \text{aggregate all final predictions for the case, } q$ 

```

## 2.4. Evaluation

### 2.4.1. Patient-Level Evaluation

If there are extraprostatic extension labels on any slice in a case, the patient is labeled as ground truth EPE positive; otherwise, the patient is ground truth EPE negative. If our approach predicts extraprostatic extension on any of the slices ( $Q \neq 0$ ), then the patient is predicted EPE positive. If  $Q(x, y, z) = 0$  on all slices, the patient is predicted EPE negative. The standard definitions for true positive, false positive, true negative and false negative predictions apply.

### 2.4.2. Sextant-Level Evaluation

Following standard clinical practice for systematic 12-core biopsies [37,38], we computationally divide the prostate volume into six regions, i.e., sextants, corresponding to left and right side and the apex, midgland and base regions. For details on this, refer to Seetharaman et al. [25]. For each sextant, we assign ground truth labels as follows: if the sextant contains any extraprostatic extension labels, then it is ground truth EPE positive; otherwise, it is ground truth EPE negative. Consequently, ground truth EPE negative patients will have six ground truth negative sextants, and ground truth EPE positive patients will have one or more ground truth positive sextants.

If our approach predicts extraprostatic extension anywhere within a sextant ( $Q \neq 0$ ), then that sextant is predicted EPE positive. If  $Q(x, y, z) = 0$  throughout a sextant, that sextant is predicted EPE negative. The standard definitions are used to count true positive, false positive, true negative and false negative predictions.

### 2.4.3. ROC Analysis

To generate receiver operating characteristic (ROC) curves, we varied the binary threshold parameter  $\alpha$  uniformly between  $[0, 1]$  in steps of 0.05. For each value of  $\alpha$  we performed the post-processing steps in Algorithm 1 and the evaluation steps at the patient

and sextant levels, recording true positive and false positive rates. We then generated two ROC curves for each model: a patient-level ROC and a sextant-level ROC, each with its corresponding area under the curve (AUC) score.

### 2.5. Experimental Design

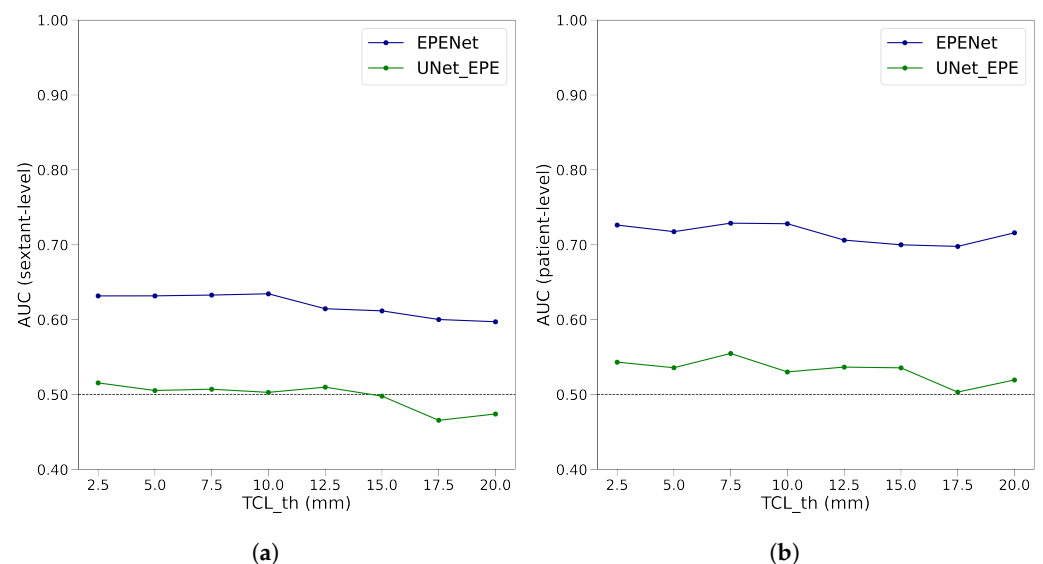
#### Hyper-Parameter Optimization

The method described in Algorithm 1 has two user-specified parameters: the binary threshold parameter  $\alpha$ , and the tumor-capsule contact line length threshold  $TCL_{th}$ .

Parameter  $\alpha$  specifies the threshold applied on the cancer probability map to create a binary mask (Equation (4)), from which connected components are computed and lesion candidates determined. By varying  $\alpha$  we generate receiver operating characteristic curves (Section 2.4.3).

The tumor-capsule contact line length threshold  $TCL_{th}$  is the length of the contact line between a lesion and the prostate boundary, above which the algorithm predicts extraprostatic extension for the lesion (Equation (8)). There is no consensus in the literature on the best tumor-capsule contact line length value to use in clinical nomograms, with proposed values ranging from 10 mm to 25 mm [39]. To find the optimal value for the  $TCL_{th}$  parameter, we ran a grid search on the training data, in which we evaluated a range of thresholds and selected the one with the best performance. Grid search is a tuning technique that attempts to compute the optimum values of hyperparameters. It is an exhaustive search that is performed on a specific parameter.

For the proposed EPENet and the baseline UNet\_EPE, we generated extraprostatic extension predictions at different values for  $TCL_{th} \in \{2.5 \text{ mm}, 5.0 \text{ mm}, 7.5 \text{ mm}, 10.0 \text{ mm}, 12.5 \text{ mm}, 15.0 \text{ mm}, 17.5 \text{ mm}, 20.0 \text{ mm}\}$ . We performed five-fold cross-validation with respect to the training dataset of the deep learning models. We constructed the receiver operating characteristic curves and computed AUC at per-patient and per-sextant level, as described in Section 2.4. Mean AUC was computed over the validation sets of the five cross-validation folds, both at sextant-level and patient-level. Figure 3 shows the results of the grid search over parameter  $TCL_{th}$  for EPENet and UNet\_EPE.



**Figure 3.** Grid search over the tumor-capsule contact line length threshold parameter ( $TCL_{th}$ ) for EPENet and UNet\_EPE. The two panels correspond to the evaluation metric used (displayed on the  $y$  axis): sextant-level AUC in panel (a) and patient-level AUC in panel (b). The  $x$  axis shows the values of  $TCL_{th}$  in millimeters. Results are averages computed over the validation sets of the five cross-validation data folds.



Based on the graphs in Figure 3, we chose  $TCL_{th} = 10$  mm, since it provided the best balance between good performance at sextant-level and patient-level. All experiments in the following sections were conducted with the tumor-capsule contact line length threshold set to  $TCL_{th} = 10$  mm.

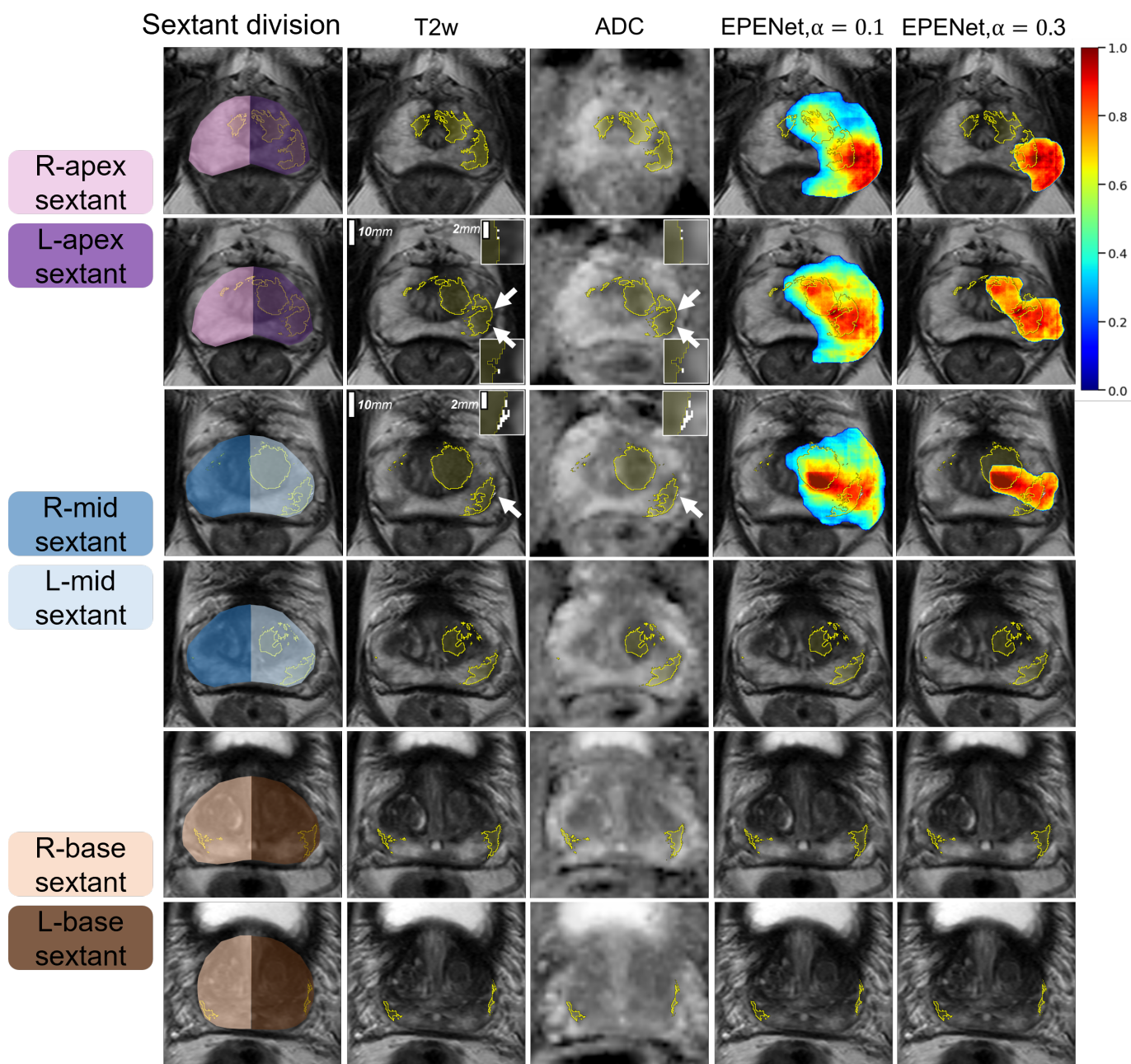
### 3. Results

#### 3.1. Qualitative Results

We used 3DSlicer [40] to visualize EPENet predictions overlaid on top of MR images. EPENet successfully detected extraprostatic extension (representative case shown in Figure 4). Note that EPENet final outputs indicate only lesions with a predicted extraprostatic extension component. Slices on which the model predicts cancer that is fully contained within the prostate capsule will not show any EPENet final predictions. For selective identification of cancer inside the prostate gland, the CorrSigNIA model should be applied directly. EPENet is designed as a tool to specifically highlight areas at risk for extraprostatic extension and should be used accordingly.

Figure 4 shows predictions computed with tumor–capsule contact line length threshold  $TCL_{th} = 10$  mm and two different binary thresholds  $\alpha = 0.1$  and  $\alpha = 0.3$ . For both values of  $\alpha$ , EPENet generated true positive predictions at the patient-level. However, varying  $\alpha$  changes the model performance at the sextant-level (which takes into account the spatial location of predictions). Column 1 of Figure 4 schematically illustrates the division of prostate into sextants corresponding to the apex, mid, base, right- and left-hand side regions of the gland. This case has two ground truth EPE positive sextants (spanned by the extraprostatic extension on the left side of the patient) and four ground truth EPE negative sextants. The EPENet model with  $\alpha = 0.3$  was 100% sensitive and 100% specific, since the final prediction was confined to the two ground-truth positive sextants (apex and mid, patient left-hand side). The EPENet model with  $\alpha = 0.1$  was 100% sensitive but only 50% specific, because the EPE prediction extended into the apex and mid sextants on the patient right-hand side, thus accounting for two false positive sextant predictions. To ensure spatially accurate predictions, it is important to select an appropriate operating point where the model balances sensitivity and specificity at the sextant-level. We address this in the next subsection.

Figure A1 in Appendix B shows qualitative results for additional cases in the test set.



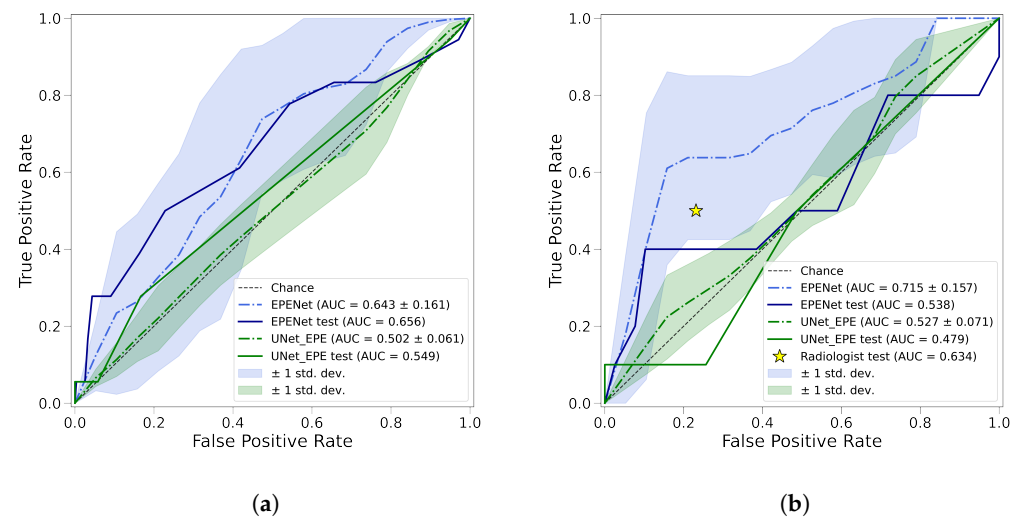
**Figure 4.** EPENet predictions for cancers with extraprostatic extension in one example case from the test set, from apex to base. EPENet does not show any predictions for cancer lesions fully contained within the prostate capsule. Column (1) Schematic illustration of the prostate division into sextant regions. Column (2) Input T2w images. Column (3) Input ADC images; yellow labels outline cancer within the prostate, white labels outline extraprostatic extension (EPE); white arrows point to the small EPE regions; the panels in the corners show zoomed-in 5 mm  $\times$  5 mm areas containing EPE. This patient shows cancer on all slices, but EPE is present only in rows 2 and 3. Column (4) and (5) show EPENet predictions for different values of the  $\alpha$  parameter. Predictions are probability maps displayed in a cold to hot color scheme (dark blue–0, dark red–1).

### 3.2. Quantitative Results

Table 4 shows sensitivity and specificity results for EPENet operating at various binary thresholds  $\alpha$ . We observed that for  $\alpha < 0.30$  specificity at the sextant level dropped below 50% in the cross-validation set. Conversely, for  $\alpha > 0.35$  sensitivity at the sextant level dropped below 50% in the cross-validation set. We recommend an optimal oper-

ating point with  $0.30 \leq \alpha \leq 0.35$  to achieve the best balance between sensitivity and specificity at patient-level and sextant-level. Within this range, we selected  $\alpha = 0.30$  to maximize sensitivity.

Figure 5 displays the receiver operating characteristic curves (ROC) for EPENet and UNet\_EPE in the cross-validation folds and test set. Over the five cross-validation folds, at the patient level, EPENet had mean  $AUC = 0.72 \pm 0.16$ , while UNet\_EPE performed close to chance with mean  $AUC = 0.53 \pm 0.07$ . At the sextant-level, EPENet (mean  $AUC = 0.64 \pm 0.16$ ) also outperformed the baseline UNet\_EPE (mean  $AUC = 0.50 \pm 0.06$ ).



**Figure 5.** Receiver operating characteristic curves based on performance metrics at the sextant-level (a) and patient-level (b). Dashed lines are the mean performance over the five cross-validation folds and the shaded regions represent one standard deviation around the average. Solid lines are the ROC results in the test set. EPENet model is shown in blue, UNet\_EPE model is green. In panel (b), the star marker represents radiologists performance on the test set.

On the test set, the radiologists assessment of EPE was 50.0% sensitive and 76.9% specific. By comparison, EPENet was 80.0% sensitive and 28.2% specific in the test set. While our proposed method had more false positives (and hence lower specificity), it also detected extraprostatic extension missed by the radiologists. The test set included 10 patients with extraprostatic extension: Radiologists correctly identified five cases, while EPENet correctly identified eight cases. The two cases missed by EPENet had focal extraprostatic extensions associated with smaller cancer lesions inside the prostate (index lesion volumes of  $640 \text{ mm}^3$  and  $470 \text{ mm}^3$  compared to test cohort median  $1100 \text{ mm}^3$ —see Table 2). The radiologists also missed one of these extraprostatic extensions. Out of the 10 patients with extraprostatic extension in the test set, 5 presented focal extraprostatic extension and 5 showed multifocal extraprostatic extension. EPENet correctly predicted all multifocal cases, whereas the radiologists only correctly predicted two of the multifocal cases.

**Table 4.** Sensitivity and specificity results in the cross-validation and test sets for the EPENet model operating at various binary thresholds  $\alpha$ . We recommend  $0.30 \leq \alpha \leq 0.35$  as the optimal operating range for this model. Results within the recommended operating range are shown in bold.

Mode	Threshold $\alpha$	Cohort	Sensitivity (Patient) %	Specificity (Patient) %	Sensitivity (Sextant) %	Specificity (Sextant) %
EPENet	0.10	cross-val	100.0 $\pm$ 0.0	1.7 $\pm$ 3.3	96.5 $\pm$ 7.1	17.9 $\pm$ 8.3
		test	90.0	0.0	88.9	13.0
EPENet	0.15	cross-val	100.0 $\pm$ 0.0	11.1 $\pm$ 6.7	95.3 $\pm$ 9.4	29.2 $\pm$ 9.5
		test	90.0	0.0	83.3	23.9
EPENet	0.20	cross-val	100.0 $\pm$ 0.0	14.3 $\pm$ 9.6	88.8 $\pm$ 13.0	36.8 $\pm$ 11.6
		test	80.0	5.1	83.3	34.4
EPENet	0.25	cross-val	97.5 $\pm$ 5.0	19.2 $\pm$ 13.4	71.8 $\pm$ 19.1	45.7 $\pm$ 11.5
		test	80.0	12.8	77.8	45.7
EPENet	<b>0.30</b>	cross-val	<b>95.0 <math>\pm</math> 10.0</b>	<b>26.8 <math>\pm</math> 8.8</b>	<b>64.4 <math>\pm</math> 21.6</b>	<b>54.6 <math>\pm</math> 8.1</b>
		test	<b>80.0</b>	<b>28.2</b>	<b>61.1</b>	<b>58.3</b>
EPENet	<b>0.35</b>	cross-val	<b>85.0 <math>\pm</math> 20.0</b>	<b>33.2 <math>\pm</math> 13.8</b>	<b>59.3 <math>\pm</math> 21.3</b>	<b>63.2 <math>\pm</math> 7.7</b>
		test	<b>50.0</b>	<b>41.0</b>	<b>55.6</b>	<b>67.8</b>
EPENet	0.40	cross-val	81.0 $\pm$ 18.5	39.9 $\pm$ 15.5	43.2 $\pm$ 31.7	71.2 $\pm$ 8.8
		test	50.0	51.3	50.0	77.2
EPENet	0.45	cross-val	74.5 $\pm$ 21.2	51.5 $\pm$ 23.3	31.4 $\pm$ 24.6	79.0 $\pm$ 8.7
		test	40.0	61.5	38.9	83.7
EPENet	0.50	cross-val	68.7 $\pm$ 19.8	54.3 $\pm$ 22.4	25.7 $\pm$ 25.0	82.6 $\pm$ 6.8
		test	40.0	74.4	27.8	90.9
EPENet	0.55	cross-val	56.6 $\pm$ 16.2	65.0 $\pm$ 21.6	19.7 $\pm$ 22.7	87.6 $\pm$ 6.0
		test	40.0	87.2	27.8	95.7
EPENet	0.60	cross-val	53.8 $\pm$ 11.3	79.6 $\pm$ 14.0	17.9 $\pm$ 17.2	92.1 $\pm$ 3.9
		test	40.0	89.7	16.7	96.7
Radiologists	—	test	<b>50.0</b>	<b>76.9</b>	—	—

#### 4. Discussion

We developed and validated a fully automated pipeline for the prediction and localization of extraprostatic extension in patients with prostate cancer. We used pre-trained deep learning models to generate cancer probability maps and applied post-processing steps and clinical decision rules to arrive at final predictions for lesions with extraprostatic extension. The end-to-end EPENet model operates without any radiologist input and only requires T2w and ADC images to make a final prediction. We recommend  $0.30 \leq \alpha \leq 0.35$  as the optimal operating range for this model and tumor-capsule contact line length threshold  $TCL_{th} = 10$  mm. For  $TCL_{th} = 10$  mm and  $\alpha = 0.30$ , EPENet achieved sensitivity 80.0% and specificity 28.2% at the patient-level, compared to the radiologists 50.0% sensitivity and 76.9% specificity on test set data. When accounting for the spatial location of predictions, at the sextant-level, EPENet model had a balanced 61.1% sensitivity and 58.3% specificity. High sensitivity shows that our method can reliably detect the presence of extraprostatic extension. As such, EPENet can be useful in highlighting regions suspicious for extraprostatic extension that might otherwise be difficult to detect by the human eye. Given the low specificity at the patient-level, the intended use for EPENet would be as a detection aid tool for radiologists rather than a standalone diagnosis software. Future work will assess performance of radiologists with and without our EPENet system and investigate the best strategies to integrate the proposed application to guide radiologists during prostate MRI interpretation in clinical practice.

We note that local artifacts or other anatomical features outside the prostate capsule that appear dark on MRI can mislead deep learning models into predicting extraprostatic extension (see examples in Figure A1 rows 2 and 3). None of the prior works on EPE

detection were confronted with this challenge since they used radiologist segmented lesions as the starting point for extraprostatic extension prediction. This diminishes the risk of false positives, but increases the risk of false negatives, since any lesion missed by the radiologist will be, by default, also missed by the EPE model.

Previous studies showed that the radiologists AUC for predicting extraprostatic extension ranged from 0.60 to 0.75. This is similar to our radiologists performance (AUC = 0.63). Park et al. [29] showed that by using additional metrics such as tumor–capsule contact line length or the EPE grading system proposed by Mehralivand et al. [30], radiologists performance improved, with AUC ranging from 0.77 to 0.85. Radiomics methods reported AUCs between 0.73 and 0.88, while the convolutional neural network PAGNet [7] had an AUC in the range 0.73 to 0.81. Since all the aforementioned methods are trained using radiologist cancer labels as inputs together with MR images, we cannot directly compare the reported AUC performance with that of EPENet ( $0.72 \pm 0.16$  at patient-level). Both the radiomics methods and PAGNet required the radiologists to manually segment the index lesions on all slices, making them time-consuming methods. Furthermore, they all solve a binary classification problem at the patient-level rather than a spatial detection problem and will omit any lesion missed by the radiologists. While EPENet generates a localized prediction for extraprostatic extension, it does not predict the size of the EPE lesion. Future research will focus on distinguishing between predictions for macroscopic EPE and microscopic EPE.

The automatic detection of extraprostatic extension is a challenging task for two key reasons. First, extraprostatic extension regions are small (see Table 2, cancer lesion volumes range between 0.7–3.4 mL, while median extraprostatic extension volume is  $<0.01$  mL) and it can even be debated whether MRI might not even capture these small regions (sometimes only representing a handful of pixels). Second, MRI intensities vary greatly across patients, making it difficult to learn features that capture very small cancers, especially since both the prostate boundary and cancer are hypointense on T2w images. Appendix C provides further details and causes for the large number of false positives.

Our study is the first to explore deep learning-based strategies that simultaneously detect tumors and assess the presence of extraprostatic extension. EPENet had an AUC of  $0.64 \pm 0.16$  at the sextant-level (0.66 in the test set) and  $0.72 \pm 0.16$  at the patient level (0.54 in the test set). The main strength of our method is that it operates without any human input. This has two advantages. First, EPENet brings added value by being able to highlight regions suspicious for extraprostatic extension that might otherwise be missed by the radiologists. Second, it eliminates the need for expensive and time-consuming manual annotations. Furthermore, the framework presented in this paper is fully independent of the underlying deep learning model used and, thus, can be applied out-of-the-box to any convolutional neural network trained for prostate cancer detection. As we develop better deep learning models, or ensemble models, the proposed framework can be applied to further improve extraprostatic extension detection, at minimal additional computational cost. An optimized system for EPE detection could have high impact in treatment planning, since it would enable the surgeon to consider nerve-sparing radical prostatectomy, which has been shown to significantly reduce the risk of post-operative erectile dysfunction and urinary incontinence [11,41].

Our study had two noteworthy limitations. First, it included a relatively small dataset from a single institution, with all MRIs acquired on scanners from a single manufacturer. This limits the generalization power of our model. There are also limitations in our dataset labels. Extraprostatic extension labels were defined by mapping from histology. Regions with missing histology (such as the extreme apex and extreme base of the prostate) may have missing EPE labels. We also disregarded seminal vesicle invasion. Second, we had several extraprostatic extension false positives driven by anatomical features outside the prostate that appear naturally dark on MRI. Future work will focus on increasing the cohort size and validate our framework on data from external institutions, as well as improved strategies for reducing false positives.



## 5. Conclusions

We introduced an approach to detect cancer lesions with extraprostatic extension on prostate multiparametric MRI. We used existing deep learning architectures and defined a framework to identify which candidate lesions might have extraprostatic extension. An optimized approach for EPE detection can serve to assist radiologists during their interpretation of MRI, providing an independent assessment of the presence of extraprostatic extension. Our approach may facilitate the planning of treatments, e.g., image-guided focal therapy or surgical removal of the prostate.

**Author Contributions:** Conceptualization, Ş.L.M. and M.R.; data curation, I.B., W.S., C.A.K., A.S. (Avishkar Sharma), P.G., R.E.F. and G.A.S.; formal analysis, Ş.L.M. and I.B.; funding acquisition, G.A.S. and M.R.; investigation, Ş.L.M.; methodology, Ş.L.M.; project administration, M.R.; resources, R.E.F., G.A.S. and M.R.; software, Ş.L.M., I.B. and A.S. (Arun Seetharaman); supervision, G.A.S. and M.R.; validation, Ş.L.M.; visualization, Ş.L.M.; Writing—original draft, Ş.L.M.; writing—review and editing, G.A.S. and M.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** We acknowledge the following funding sources: Departments of Radiology and Urology, Stanford University, National Institutes of Health, National Cancer Institute (U01CA196387 to JDB) and the generous philanthropic support of our patients (GS). Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R37CA260346. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Stanford University Medical Center (Protocol ID: IRB-44998, Approval Date: 22 October 2021).

**Informed Consent Statement:** Patient consent was waived since this study was a chart review of previously collected data.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** Mirabela Rusu is a consultant for Roche and has research grants from GE Healthcare and Philips Healthcare.

## Abbreviations

The following abbreviations are used in this manuscript:

ADC	Apparent Diffusion Coefficient
AUC	Area Under the Curve
CNN	Convolutional Neural Network
EPE	Extraprostatic Extension
MRI	Magnetic Resonance Imaging
PI-RADS	Prostate Imaging-Reporting and Data System
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
TCL	Tumor–capsule Contact Line Length (also known as Capsular Contact Length, CCL)

## Appendix A. Details of Deep Learning Models for Cancer Detection

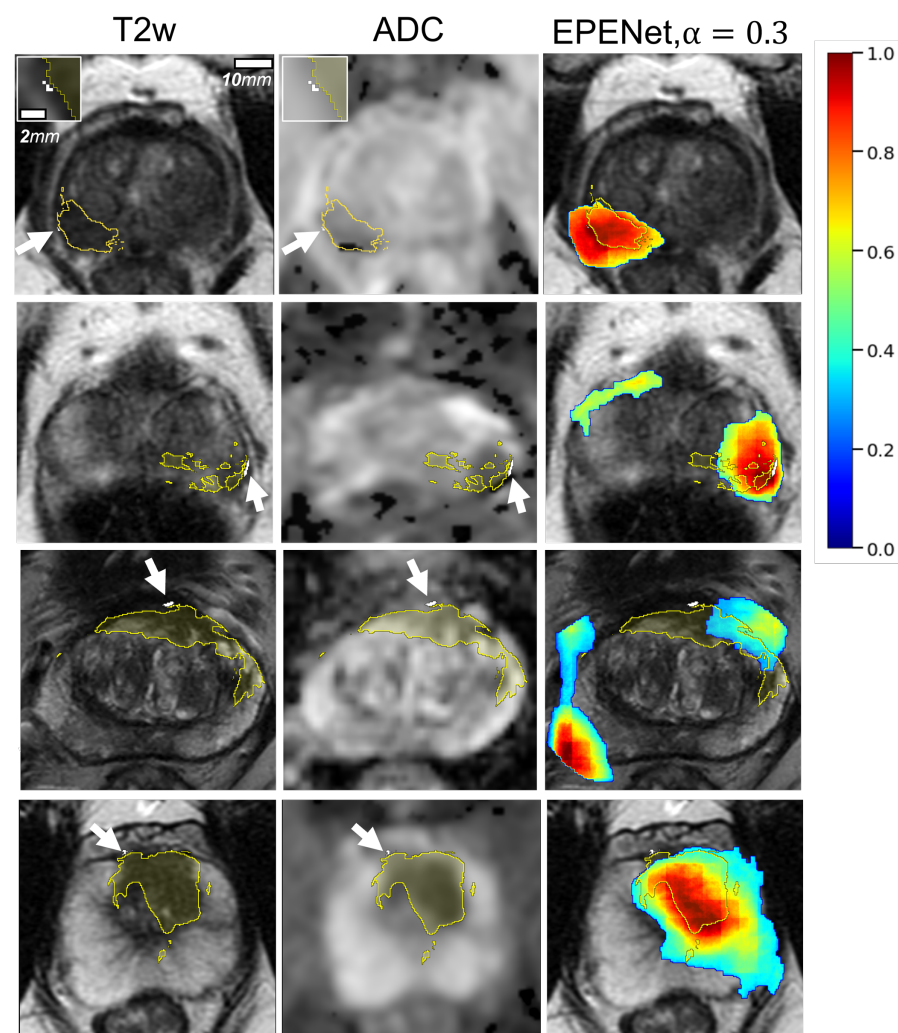
We implemented two networks: UNet [35] and the state-of-the-art Correlated Signature Network for Indolent and Aggressive prostate cancer detection (CorrSigNIA) [26]. U-Net showed promising results in several prostate cancer detection studies [23,42,43], and it was used as a baseline reference in others [25,26]. The UNet model takes three adjacent slices of T2w and ADC images as input, and outputs a cancer probability map for the middle slice. Inspired by the Holistically Nested Edge Detector [44], CorrSigNIA uses a three-branch architecture to optimize independent kernels for T2w MRI, ADC maps and CorrNet features, obtained via Common Representation Learning. The latter approach identifies combinations of MRI features that correlate with histopathology image



features [26,45]. The CorrSigNIA network takes an input of three consecutive slices in each modality, and predicts cancer probabilities for the middle slice [26].

Both models were pre-trained for the prostate cancer detection task within the boundaries of the prostate gland. We used a five-fold cross-validation scheme. We used Adam optimizer with a batch size of 32, initial learning rate  $\eta = 10^{-3}$  and weight decay  $\alpha = 0.1$ . We trained for a maximum of 100 epochs, with early stopping on a patience of 20 epochs when there was no improvement in validation loss. A patience of 10 epochs was used to reduce learning rate on plateau. We augmented training images with left to right flipping and random rotation with angles ranging between  $-15$  and  $15$  degrees.

## Appendix B. Additional Visual Results



**Figure A1.** Visualization of predictions for extraprostatic extension in four example cases from the test set; each row shows the leading slice from a different case. Column (1) input T2w images; Column (2) input ADC images; yellow labels outline cancer within the prostate, white labels outline extraprostatic extension (EPE); white arrows point to the small EPE regions; the panels in the corners show zoomed-in  $5\text{ mm} \times 5\text{ mm}$  areas containing EPE. Column (3) EPENet predictions. Predictions are probability maps displayed in a cold to hot color scheme (dark blue=0, dark red=1). Note that EPENet predictions highlight lesions with extraprostatic extension. EPENet does not show any predictions for cancer lesions fully contained within the prostate capsule.

### Appendix C. Analysis of False Positive Predictions

To better understand the high number of false positives, we looked at the size of cancer and extraprostatic extension lesions (Figure A2) the size and position relative to prostate boundary of EPENet predictions (Figure A3). Figure A2 shows scatter plots where each point represents a different case, and the x and y coordinates correspond to the volumes of cancer and extraprostatic extension as computed from ground truth labels. The color of each point shows how the case was classified by EPENet (TP-true positive, FP-false positive, FN-false negative, TN-true negative). The first observation we make is that there is no clear correlation between cancer lesion volume and EPE volume. Both in the train set (Figure A2a) and test set (Figure A2c) we see that cancer volumes vary widely (from less than 1% of the prostate volume up to 30% or more), but EPE volumes don't. This means that even with ground truth knowledge about cancer extent inside the prostate it is not straightforward to predict extraprostatic extension, re-emphasizing the difficulty level of the task. The second observation is that most false positives correspond to cases where cancer is 5–15% of the prostate volume, and the majority of model true negatives correspond to cases with small cancers (<5% of prostate volume). There are few false negative predictions.

Figure A3 takes a closer look at the true positive and false positive cases, this time only plotting information that is available to the model. The x coordinate shows the volume of EPENet predictions in a case (i.e., entire predicted cancer, both inside and outside the prostate), and the y coordinate shows how much of the EPENet prediction is outside the prostate capsule. What this scatter plot shows is that the two classes are not linearly separable given the information available to the model. Because of this, our current method does not use cut-off values neither for the volume of EPENet predictions nor the fraction of prediction volume outside the prostate. For example, the dotted lines in Figure A3 show that if we tried to find cut-off values on the training set and apply them to the test set, the reduction in true positives would be bigger than the reduction of false positives in the test set. Overall, this shows the scenarios in which our current decision mechanism reaches its limitations, and helps us gain insight for future work to improve the classifier.

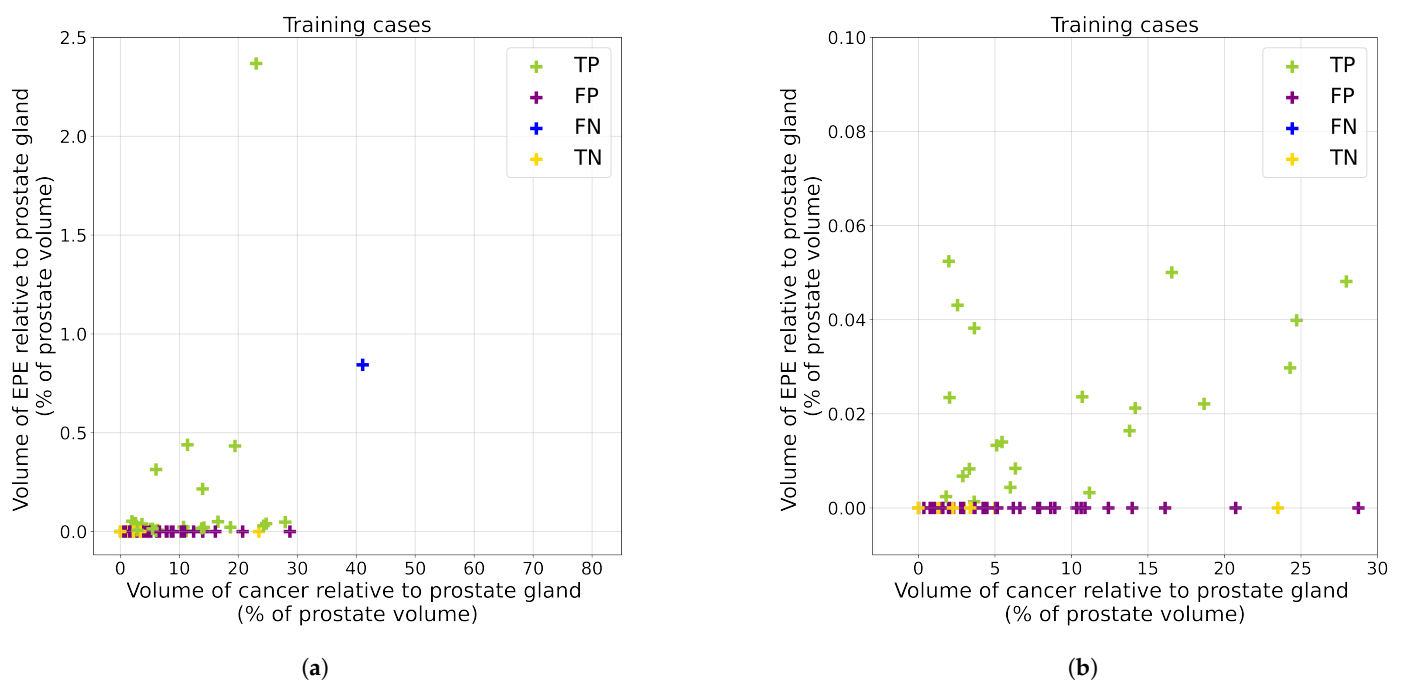
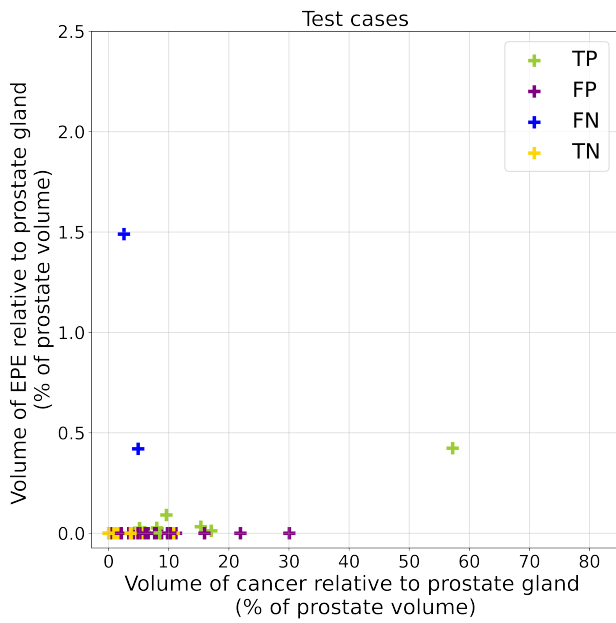
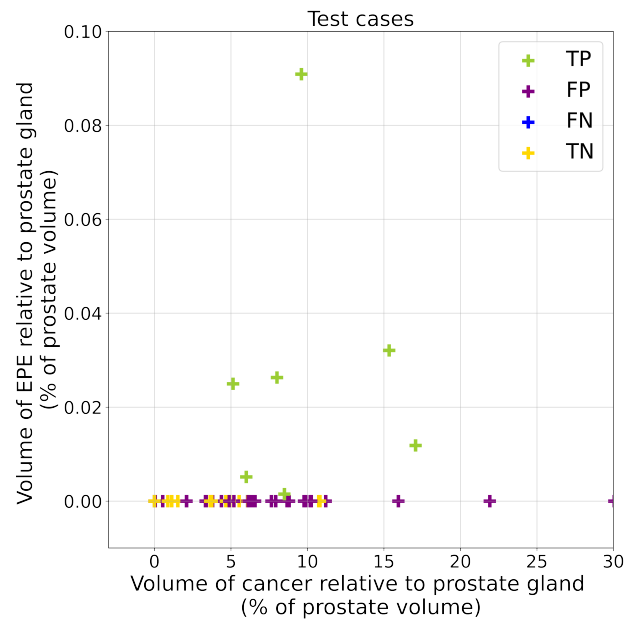


Figure A2. Cont.

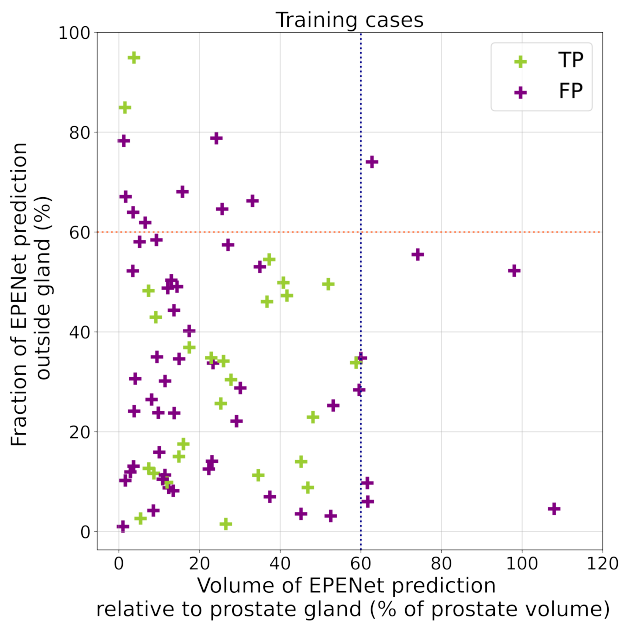


(c)

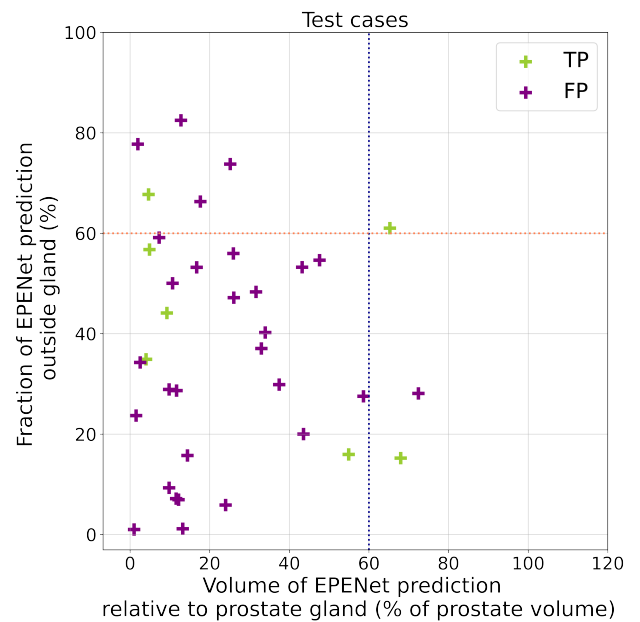


(d)

**Figure A2.** True and false predictions, with ground truth information about cancer and EPE lesion volumes. The *x*-axis shows the volume of cancer inside the prostate as a percentage of the total prostate gland volume. The *y*-axis shows the volume of extraprostatic extension as a percentage of the total prostate gland volume. Cases which are ground truth negative for EPE therefore have coordinate  $y = 0$ . (a) All training set cases. (b) Training set, zoomed-in view on the cases with small cancer and EPE volumes. (c) All test set cases. (d) Test set cases, zoomed-in view on the cases with small cancer and EPE volumes.



(a)



(b)

**Figure A3.** True positive (TP) and false positive (FP) predictions, with information about the EPENet prediction volumes. The *x*-axis shows the total volume of EPENet prediction as a percentage of the prostate gland volume. The *y*-axis shows the percentage of the prediction volume that is outside the

prostate. The TP and FP classes are not linearly separable based on size and position of prediction alone. (a) An example of cut-offs we might select in the training set. (b) Effect of applying training set cut-offs on the test set. (a) Training set cases. The dotted lines show an example of cut-off values we could apply to eliminate some false positives (i.e., declare predictions that have either  $x > 60\%$  or  $y > 60\%$  negative for EPE). This would eliminate 14 false positives and create 2 false negatives (15% reduction in FP and 7% reduction in TP). (b) Test set cases, with the example cut-off values from the training set. We observe that applying the cut-offs would eliminate five false positives, but would also lead to three false negatives (14% reduction in FP and 25% reduction in TP).

## References

- Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]
- Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* **2022**, *72*, 7–33. [[CrossRef](#)]
- Tourinho-Barbosa, R.; Srougi, V.; Nunes-Silva, I.; Baghdadi, M.; Rembeyo, G.; Eiffel, S.S.; Barret, E.; Rozet, F.; Galiano, M.; Cathelineau, X.; et al. Biochemical recurrence after radical prostatectomy: What does it mean? *Int. Braz. J.* **2018**, *44*, 14–21. [[CrossRef](#)] [[PubMed](#)]
- Lin, X.; Kapoor, A.; Gu, Y.; Chow, M.; Xu, H.; Major, P.; Tang, D. Assessment of biochemical recurrence of prostate cancer (Review). *Int. J. Oncol.* **2019**, *55*, 1194–1212. [[CrossRef](#)]
- Park, C.K.; Chung, Y.S.; Choi, Y.D.; Ham, W.S.; Jang, W.S.; Cho, N.H. Revisiting extraprostatic extension based on invasion depth and number for new algorithm for substaging of pT3a prostate cancer. *Sci. Rep.* **2021**, *11*, 13952. [[CrossRef](#)] [[PubMed](#)]
- Cheng, L.; MacLennan, G.T.; Bostwick, D.G. *Urologic Surgical Pathology*; Saunders: Philadelphia, PA, USA, 2019.
- Hou, Y.; Zhang, Y.H.; Bao, J.; Bao, M.L.; Yang, G.; Shi, H.B.; Song, Y.; Zhang, Y.D. Artificial intelligence is a promising prospect for the detection of prostate cancer extracapsular extension with mpMRI: A two-center comparative study. *Eur. J. Nucl. Med. Mol. Imaging* **2021**, *48*, 3805–3816. [[CrossRef](#)]
- Gandaglia, G.; Ploussard, G.; Valerio, M.; Mattei, A.; Fiori, C.; Roumiguié, M.; Fossati, N.; Stabile, A.; Beauval, J.B.; Malavaud, B.; et al. The Key Combined Value of Multiparametric Magnetic Resonance Imaging, and Magnetic Resonance Imaging–targeted and Concomitant Systematic Biopsies for the Prediction of Adverse Pathological Features in Prostate Cancer Patients Undergoing Radical Prostatectomy. *Eur. Urol.* **2020**, *77*, 733–741. [[CrossRef](#)]
- Diamand, R.; Ploussard, G.; Roumiguié, M.; Oderda, M.; Benamran, D.; Fiard, G.; Quackels, T.; Assenmacher, G.; Simone, G.; Van Damme, J.; et al. External Validation of a Multiparametric Magnetic Resonance Imaging–based Nomogram for the Prediction of Extracapsular Extension and Seminal Vesicle Invasion in Prostate Cancer Patients Undergoing Radical Prostatectomy. *Eur. Urol.* **2021**, *79*, 180–185. [[CrossRef](#)] [[PubMed](#)]
- Somford, D.; Hamoen, E.; Fütterer, J.; van Basten, J.; van de Kaa, C.H.; Vreuls, W.; van Oort, I.; Vergunst, H.; Kiemeny, L.; Barentsz, J.; et al. The Predictive Value of Endorectal 3 Tesla Multiparametric Magnetic Resonance Imaging for Extraprostatic Extension in Patients with Low, Intermediate and High Risk Prostate Cancer. *J. Urol.* **2013**, *190*, 1728–1734. [[CrossRef](#)] [[PubMed](#)]
- Wang, X.; Wu, Y.; Guo, J.; Chen, H.; Weng, X.; Liu, X. Intrafascial nerve-sparing radical prostatectomy improves patients' postoperative continence recovery and erectile function. *Medicine* **2018**, *97*, e11297. [[CrossRef](#)]
- Shieh, A.C.; Guler, E.; Ojili, V.; Paspulati, R.M.; Elliott, R.; Ramaiya, N.H.; Tirumani, S.H. Extraprostatic extension in prostate cancer: primer for radiologists. *Abdom. Radiol.* **2020**, *45*, 4040–4051. [[CrossRef](#)]
- Falagario, U.G.; Jambor, I.; Ratnani, P.; Martini, A.; Treacy, P.J.; Wajswol, E.; Lantz, A.; Papastefanou, G.; Weil, R.; Phillip, D.; et al. Performance of prostate multiparametric MRI for prediction of prostate cancer extraprostatic extension according to NCCN risk categories: Implication for surgical planning. *Ital. J. Urol. Nephrol.* **2020**, *72*, 746–754.
- De Rooij, M.; Hamoen, E.H.; 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.* **2016**, *70*, 233–245. [[CrossRef](#)]
- Eberhardt, S.C. Local Staging of Prostate Cancer with MRI: A Need for Standardization. *Radiology* **2019**, *290*, 720–721. [[CrossRef](#)] [[PubMed](#)]
- Choyke, P.L. A Grading System for Extraprostatic Extension of Prostate Cancer That We Can All Agree Upon? *Radiol. Imaging Cancer* **2020**, *2*, e190088. [[CrossRef](#)] [[PubMed](#)]
- Cuocolo, R.; Stanzione, A.; Faletti, R.; Gatti, M.; Callaris, G.; Fornari, A.; Gentile, F.; Motta, A.; Dell'Aversana, S.; Creta, M.; et al. MRI index lesion radiomics and machine learning for detection of extraprostatic extension of disease: A multicenter study. *Eur. Radiol.* **2021**, *31*, 7575–7583. [[CrossRef](#)]
- Shiradkar, R.; Zuo, R.; Mahran, A.; Ponsky, L.; Tirumani, S.H.; Madabhushi, A. Radiomic features derived from periprostatic fat on pre-surgical T2w MRI predict extraprostatic extension of prostate cancer identified on post-surgical pathology: Preliminary results. In Proceedings of the Medical Imaging 2020: Computer-Aided Diagnosis, Houston, TX, USA, 15–20 February 2020; Hahn, H.K., Mazurowski, M.A., Eds.; SPIE: Bellingham, WA, USA, 2020. [[CrossRef](#)]
- Xu, L.; Zhang, G.; Zhao, L.; Mao, L.; Li, X.; Yan, W.; Xiao, Y.; Lei, J.; Sun, H.; Jin, Z. Radiomics Based on Multiparametric Magnetic Resonance Imaging to Predict Extraprostatic Extension of Prostate Cancer. *Front. Oncol.* **2020**, *10*, 940. [[CrossRef](#)]

20. Stanzione, A.; Cuocolo, R.; Coccozza, S.; Romeo, V.; Persico, F.; Fusco, F.; Longo, N.; Brunetti, A.; Imbriaco, M. Detection of Extraprostatic Extension of Cancer on Biparametric MRI Combining Texture Analysis and Machine Learning: Preliminary Results. *Acad. Radiol.* **2019**, *26*, 1338–1344. [CrossRef]
21. Sumathipala, Y.; Lay, N.; Turkbey, B. Prostate cancer detection from multi-institution multiparametric MRIs using deep convolutional neural networks. *J. Med. Imaging* **2018**, *5*, 044507. [CrossRef]
22. Cao, R.; Bajgirani, A.M.; Mirak, S.A.; Shakeri, S.; Zhong, X.; Enzmann, D.; Raman, S.; Sung, K. Joint Prostate Cancer Detection and Gleason Score Prediction in mp-MRI via FocalNet. *IEEE Trans. Med. Imaging* **2019**, *38*, 2496–2506. [CrossRef]
23. Sanyal, J.; Banerjee, I.; Hahn, L.; Rubin, D. An Automated Two-step Pipeline for Aggressive Prostate Lesion Detection from Multi-parametric MR Sequence. *AMIA Jt. Summits Transl. Sci. Proc.* **2020**, *2020*, 552–560.
24. Wildeboer, R.R.; van Sloun, R.J.; Wijkstra, H.; Mischi, M. Artificial intelligence in multiparametric prostate cancer imaging with focus on deep-learning methods. *Comput. Methods Programs Biomed.* **2020**, *189*, 105316. [CrossRef]
25. Seetharaman, A.; Bhattacharya, I.; Chen, L.C.; Kunder, C.A.; Shao, W.; Soerensen, S.J.C.; Wang, J.B.; Teslovich, N.C.; Fan, R.E.; Ghanouni, P.; et al. Automated detection of aggressive and indolent prostate cancer on magnetic resonance imaging. *Med. Phys.* **2021**, *48*, 2960–2972. [CrossRef]
26. Bhattacharya, I.; Seetharaman, A.; Kunder, C.; Shao, W.; Chen, L.C.; Soerensen, S.J.; Wang, J.B.; Teslovich, N.C.; Fan, R.E.; Ghanouni, P.; et al. Selective identification and localization of indolent and aggressive prostate cancers via CorrSigNIA: An MRI-pathology correlation and deep learning framework. *Med. Image Anal.* **2022**, *75*, 102288. [CrossRef]
27. Eurboonyanun, K.; Pisuchpen, N.; O’Shea, A.; Lahoud, R.M.; Atre, I.D.; Harisinghani, M. The absolute tumor-capsule contact length in the diagnosis of extraprostatic extension of prostate cancer. *Abdom. Radiol.* **2021**, *46*, 4014–4024. [CrossRef]
28. Losnegård, A.; Reisæter, L.A.R.; Halvorsen, O.J.; Jurek, J.; Assmus, J.; Arnes, J.B.; Honoré, A.; Monssen, J.A.; Andersen, E.; Haldorsen, I.S.; et al. Magnetic resonance radiomics for prediction of extraprostatic extension in non-favorable intermediate- and high-risk prostate cancer patients. *Acta Radiol.* **2020**, *61*, 1570–1579. [CrossRef] [PubMed]
29. Park, K.J.; hyun Kim, M.; Kim, J.K. Extraprostatic Tumor Extension: Comparison of Preoperative Multiparametric MRI Criteria and Histopathologic Correlation after Radical Prostatectomy. *Radiology* **2020**, *296*, 87–95. [CrossRef] [PubMed]
30. Mehralivand, S.; Shih, J.H.; Harmon, S.; Smith, C.; Bloom, J.; Czarniecki, M.; Gold, S.; Hale, G.; Rayn, K.; Merino, M.J.; et al. A Grading System for the Assessment of Risk of Extraprostatic Extension of Prostate Cancer at Multiparametric MRI. *Radiology* **2019**, *290*, 709–719. [CrossRef]
31. Ma, S.; Xie, H.; Wang, H.; Han, C.; Yang, J.; Lin, Z.; Li, Y.; He, Q.; Wang, R.; Cui, Y.; et al. MRI-Based Radiomics Signature for the Preoperative Prediction of Extracapsular Extension of Prostate Cancer. *J. Magn. Reson. Imaging* **2019**, *50*, 1914–1925. [CrossRef] [PubMed]
32. Krishna, S.; Lim, C.S.; McInnes, M.D.; Flood, T.A.; Shabana, W.M.; Lim, R.S.; Schieda, N. Evaluation of MRI for diagnosis of extraprostatic extension in prostate cancer. *J. Magn. Reson. Imaging* **2017**, *47*, 176–185. [CrossRef]
33. Rusu, M.; Shao, W.; Kunder, C.A.; Wang, J.B.; Soerensen, S.J.C.; Teslovich, N.C.; Sood, R.R.; Chen, L.C.; Fan, R.E.; Ghanouni, P.; et al. Registration of presurgical MRI and histopathology images from radical prostatectomy via RAPSODI. *Med. Phys.* **2020**, *47*, 4177–4188. [CrossRef] [PubMed]
34. Nyul, L.; Udupa, J.; Zhang, X. New variants of a method of MRI scale standardization. *IEEE Trans. Med. Imaging* **2000**, *19*, 143–150. [CrossRef] [PubMed]
35. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Lecture Notes in Computer Science*; Springer International Publishing: Berlin/Heidelberg, Germany, 2015; pp. 234–241. [CrossRef]
36. Silversmith, W. cc3d: Connected Components on Multilabel 3D & 2D Images. 2021. Available online: <https://zenodo.org/record/5719536#.YpWNCv5BxPY> (accessed on 1 May 2022).
37. Elliott, S.P.; Shinohara, K.; Logan, S.L.; Carroll, P.R. Sextant Prostate Biopsies Predict Side and Sextant Site of Extracapsular Extension of Prostate Cancer. *J. Urol.* **2002**, *168*, 105–109. [CrossRef]
38. Mayes, J.M.; Mouraviev, V.; Sun, L.; Madden, J.F.; Polascik, T.J. Can the conventional sextant prostate biopsy reliably diagnose unilateral prostate cancer in low-risk, localized, prostate cancer? *Prostate Cancer Prostatic Dis.* **2008**, [CrossRef]
39. Kim, T.H.; Woo, S.; Han, S.; Suh, C.H.; Ghafoor, S.; Hricak, H.; Vargas, H.A. The Diagnostic Performance of the Length of Tumor Capsular Contact on MRI for Detecting Prostate Cancer Extraprostatic Extension: A Systematic Review and Meta-Analysis. *Korean J. Radiol.* **2020**, *21*, 684. [CrossRef]
40. Kikinis, R.; Pieper, S.D.; Vosburgh, K.G. 3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support. In *Intraoperative Imaging and Image-Guided Therapy*; Springer: New York, NY, USA, 2013; pp. 277–289. [CrossRef]
41. Kumar, A.; Patel, V.R.; Panaiyadiyan, S.; Bhat, K.R.S.; Moschovas, M.C.; Nayak, B. Nerve-sparing robot-assisted radical prostatectomy: Current perspectives. *Asian J. Urol.* **2021**, *8*, 2–13. [CrossRef]
42. Schelb, P.; Kohl, S.; Radtke, J.P.; Wiesenfarth, M.; Kickingereder, P.; Bickelhaupt, S.; Kuder, T.A.; Stenzinger, A.; Hohenfellner, M.; Schlemmer, H.P.; et al. Classification of Cancer at Prostate MRI: Deep Learning versus Clinical PI-RADS Assessment. *Radiology* **2019**, *293*, 607–617. [CrossRef]
43. de Vente, C.; Vos, P.; Hosseinzadeh, M.; Pluim, J.; Veta, M. Deep Learning Regression for Prostate Cancer Detection and Grading in Bi-Parametric MRI. *IEEE Trans. Biomed. Eng.* **2021**, *68*, 374–383. [CrossRef] [PubMed]

44. Xie, S.; Tu, Z. Holistically-Nested Edge Detection. *arXiv* **2015**, arXiv:1504.06375.
45. Bhattacharya, I.; Seetharaman, A.; Shao, W.; Sood, R.; Kunder, C.A.; Fan, R.E.; Soerensen, S.J.C.; Wang, J.B.; Ghanouni, P.; Teslovich, N.C.; et al. CorrSigNet: Learning CORRelated Prostate Cancer SIGNatures from Radiology and Pathology Images for Improved Computer Aided Diagnosis. In Proceedings of the MICCAI 2020 Medical Image Computing and Computer Assisted Intervention, Lima, Peru, 4–8 October 2020; Springer International Publishing: Cham, Switzerland, 2020; pp. 315–325. [[CrossRef](#)]