



A Two-Stage Approach for Automated Prostate Lesion Detection and Classification with Mask R-CNN and Weakly Supervised Deep Neural Network

Zhiyu Liu¹, Wenhao Jiang¹, Kit-Hang Lee¹, Yat-Long Lo¹,
Yui-Lun Ng¹, Qi Dou², Varut Vardhanabhuti³,
and Ka-Wai Kwok¹✉

¹ Department of Mechanical Engineering, The University of Hong Kong,
Pok Fu Lam, Hong Kong
kwokkw@hku.hk

² Department of Computing, Imperial College London, London, UK

³ Department of Diagnostic Radiology, Li Ka Shing Faculty of Medicine,
The University of Hong Kong, Pok Fu Lam, Hong Kong

Abstract. Early diagnosis of prostate cancer is very crucial to reduce the mortality rate. Multi-parametric magnetic resonance imaging (MRI) can provide detailed visualization of prostate tissues and lesions. Their malignancy can be diagnosed before any necessary invasive approaches, such as needle biopsy, at the risk of damage to or inflammation of the periprostatic nerves, prostate and bladder neck. However, the prostate tissue malignancy on magnetic resonance (MR) images can also be difficult to determine, with often inconclusive results among the clinicians. With the progress in artificial intelligence (AI), research on MR image-based lesion classification with AI tools are being explored increasingly. So far, existing classification approaches heavily rely on manually labelling of lesion areas, which is a labor-intensive and time-consuming process. In this paper, we present a *novel* two-stage method for *fully-automated* prostate lesion detection and classification, using input sequences of T2-weighted images, apparent diffusion coefficient (ADC) maps and high b-value diffusion-weighted images. In the first stage, a Mask R-CNN model is trained to automatically segment prostate structures. In the second stage, a weakly supervised deep neural network is developed to detect and classify lesions in a *single* run. To validate the accuracy of our system, we tested our method on two datasets, one from the PROSTATEx Challenge and the other from our local cohort. Our method can achieve average area-under-the-curve (AUC) of 0.912 and 0.882 on the two datasets respectively. The proposed approach present a promising tool for radiologists in their clinical practices.

Keywords: Prostate cancer · MR images classification · Weakly supervised learning

1 Introduction

Prostate cancer is the second leading cancer-related cause of death in the male population [1], with the estimate that 1 in 9 men in the US will be diagnosed with prostate cancer in their lifetime [2]. The prostate-specific antigen (PSA) test is mainly used as an upfront screening test for men without symptoms. Men with PSA level >4 would have an increasing chance of prostate cancer, and will be followed-up by further tests, such as MRI. mpMRI is commonly employed before tissue biopsy. Diagnosis with MRI would reduce unnecessary biopsy by 70%. Daniel *et al.* [3] reported 486% increased use of mpMRI from Oct 2013 to Dec 2015. Thus, a lot of prostate magnetic resonance (MR) images require radiologists to interpret in current clinical routine. Recently, the Prostate Imaging Reporting and Data System [4] guidelines were established to standardize and minimize the variation in interpreting prostate MRI. Geoffrey *et al.* [5] observed variability in cancer yield across radiologists, with 24% of males who were assigned a score of having a benign lesion turned out to have clinically significant prostate cancer on biopsy. This false negative rate could vary from 13% to 60% among radiologists. With the advances of deep learning in medical imaging, our goal is to automate the diagnosis process by detecting and classifying prostate lesions with high accuracy in a single framework.

There have been several works on prostate lesion classification with multi-parametric MRI. Karimi *et al.* [6] proposed to combine hand-crafted features and learned features for classifying malignancy of prostate lesions mpMRI images, which led to an area-under-the-curve (AUC) of 0.87. Liu *et al.* [7] proposed XmasNet that reformulates 3D mpMRI images as a 2D problem. To incorporate 3D information when learning from 2D slices, they employed data augmentation through 3D rotation and slicing. The approach attained a performance of 0.84 AUC.

However, these methods require labeling of lesion centroid, which is necessarily done by clinicians. In a related research on prostate lesion detection and classification in MRI, Kiraly *et al.* [8] used a deep convolutional encoder-decoder architecture to simultaneously detect and classify prostate cancer, and they reached an average classification performance of 0.834 AUC. However, they applied a region of interest (ROI) of roughly the same size to ensure only the prostate and its surrounding areas were considered. Such simplification may lead to imprecision.

In this paper, we propose a novel two-stage framework for *fully*-automated prostate lesion detection and diagnosis. In the *first* stage, prostate zone/contour in MR images is automatically segmented. *Secondly*, an analytical framework is implemented to detect and classify malignancy of prostate lesions. The major work contributions are: (i) A Mask R-CNN [9] model is trained to segment/pinpoint a smaller ROI that contains all the lesion candidates. (ii) A weakly supervised deep neural network is developed to process the detection/classification in a *single* run. (iii) Detailed validation is conducted on two datasets, namely PROSTATEx Challenge [10] dataset and our local cohort. Classification performance is also compared with the state-of-the-art approaches.

2 Methods

Figure 1 illustrates our proposed two-stage automatic framework. We applied an object detection method, Mask R-CNN, to train a prostate segmentation model for T2-weighted images that show tissues in high contrast and brightness. The model outputs a smaller size image that contains the prostate ROI. With coordinate transformation, prostate regions in ADC maps and high b -value diffusion-weighted images become available. Next, the prostate ROI will act as input to a deep neural network (DNN), outputting lesion areas and classification results in one single feedforward pass. Malignancy of the classified lesions will then be determined through the ensemble learning of images of different input sequences.

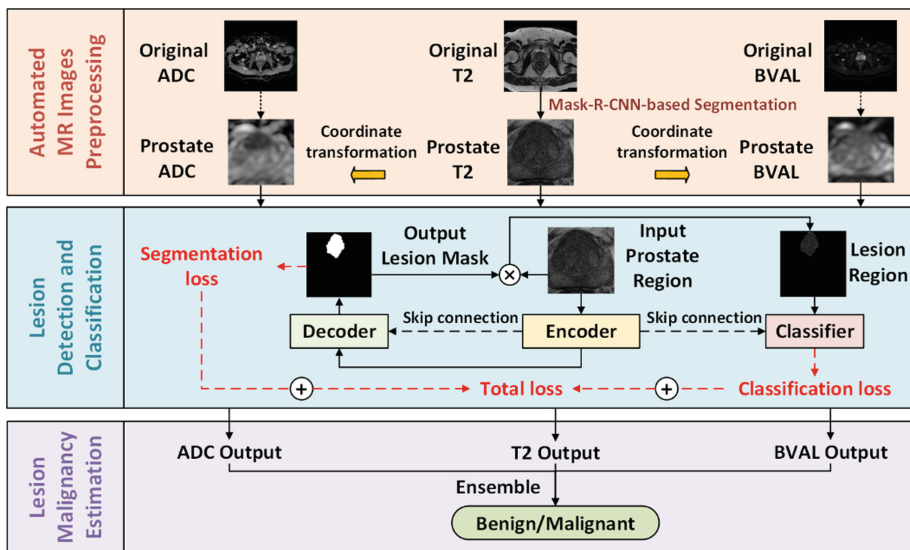


Fig. 1. Schematics of the proposed automated prostate lesion detection/classification.

2.1 Mask R-CNN for Automated Prostate Structures Segmentation

Prostate lesions detection can be difficult, as the lesions are small relative to the entire image size. Thus, identifying the prostate ROI is crucial to enable more accurate and effective lesion detection. Mask R-CNN object detection approach relies on generated region proposals, each of which outputs a class, bounding box, and a mask. Multi-scale features are extracted from various layers, which provide more powerful information to segment specific objects at various scales. These are particularly useful for our case, as the prostate sizes among patients can vary. To train this Mask R-CNN model, we employ an online prostate segmentation dataset, along with well labelled masks, which were released by the Initiative for Collaborative Computer Vision Benchmarking (I2CVB) [11]. The trained model can narrow the entire MR images down to the prostate ROI only, thus facilitating the subsequent detection and classification process.

2.2 Weakly Supervised Deep Neural Network for Prostate Lesion Detection and Classification

In the prostate classification datasets used by existing approaches, only lesion centroids are located, but not their exact outlines. Previous work [8] attempted to resolve this issue by applying lesion labels with identical gaussian distribution at each lesion point. However, this approach could not account for variations in lesion sizes. To this end, we employ a distance regularized level set evolution [12] to generate weak lesion labels. Provided with the lesion centroids, this method can be applied to edge-based active contour model for weak lesion labels.

Given the prostate region identified at stage one, we employed our novel weakly supervised DNN for simultaneous lesion detection and classification based on the weak masks and classification labels. The network comprises *three* key components, namely encoder, decoder and classifier. The encoder consists of five groups of convolutional layers and max-pooling layers while the decoder contains five upsampling layers. The encoder and decoder are linked with skip connections between their corresponding layers. The decoder is trained to produce lesion masks, indicating the location of predicted lesions. The predicted weak lesion mask, together with prostate region and features extracted from the encoder are then reused by the classifier. We hypothesize that the additional lesion and prostate structure masks can act as a form of attention map to guide the classifier, leading to the improved classification performance. To predict the malignancy of lesions, the classifier itself also contains five groups of convolutional and max-pooling layers with two fully connected layers. A composite loss function is applied, which combines the lesion segmentation loss and classification loss, to train the entire network. For each sample, the classification loss L_C can be designed as:

$$L_C = -y \log c(x) - (1 - y) \log[1 - c(x)] \quad (1)$$

where $x \in \mathbb{R}^{w \times h}$, $y \in \mathbb{R}^1$ and $c(x) \in \mathbb{R}^1$, respectively, denotes input, label and output of the data sample; w and h are the width and height of input images. Lesion segmentation loss L_S can be described as:

$$L_S = 1 - \frac{2 \sum_i^w \sum_j^h [M_{ij} S_{ij}(x)] + \varepsilon}{\sum_i^w \sum_j^h M_{ij} + \sum_i^w \sum_j^h S_{ij}(x) + \varepsilon} \quad (2)$$

where, $M \in \mathbb{R}^{w \times h}$ and $S(x) \in \mathbb{R}^{w \times h}$, respectively, denote mask label and output mask of the lesion; M_{ij} and $S_{ij}(x)$ represent pixel values of their corresponding masks in the i_{th} column and j_{th} row of matrix. The parameter ε is a numerical constant to avoid division of zero and ensure numerical stability. The total loss function L_T is also weighted with two coefficients, λ_C and λ_S as below. Note that the Adam [13] optimizer is applied to train the network with a learning rate of 10^{-6} .

$$L_T = \lambda_C L_C + \lambda_S L_S \quad (3)$$

3 Experimental Results

3.1 Prostate Structures Segmentation

In the *first* stage, 646 slices of T2-weighted images from 21 patients scanned with 1.5-T (GE) scanner and 15 patients scanned with 3.0-T (Siemens) scanner were extracted to train and validate our prostate segmentation model. We used a 7:2:1 split ratio for the slices extracted from the I2CVB dataset for training, validation and testing. Data augmentation was performed on the training sets by random rotation of $[\pm 30^\circ, \pm 60^\circ, \pm 90^\circ]$. All the inputs were resized to 512×512 pixels before feeding in the Mask R-CNN model. Learning rate is set to 10^{-3} . We trained on two GPUs with a batch size of 4 for 200 epochs. The model with the best dice coefficient on the validation set is chosen as final model, with results on the test split reported. Figure 2a shows a sample of segmentation result. As shown in Table 1, we have compared our results with other methods using mean intersection over union (IoU). These show that Mask R-CNN can achieve higher IoU in segmenting prostate and central gland. Our segmentation model can obtain prostate regions on T2 sequences of the two datasets (Fig. 2b and c).

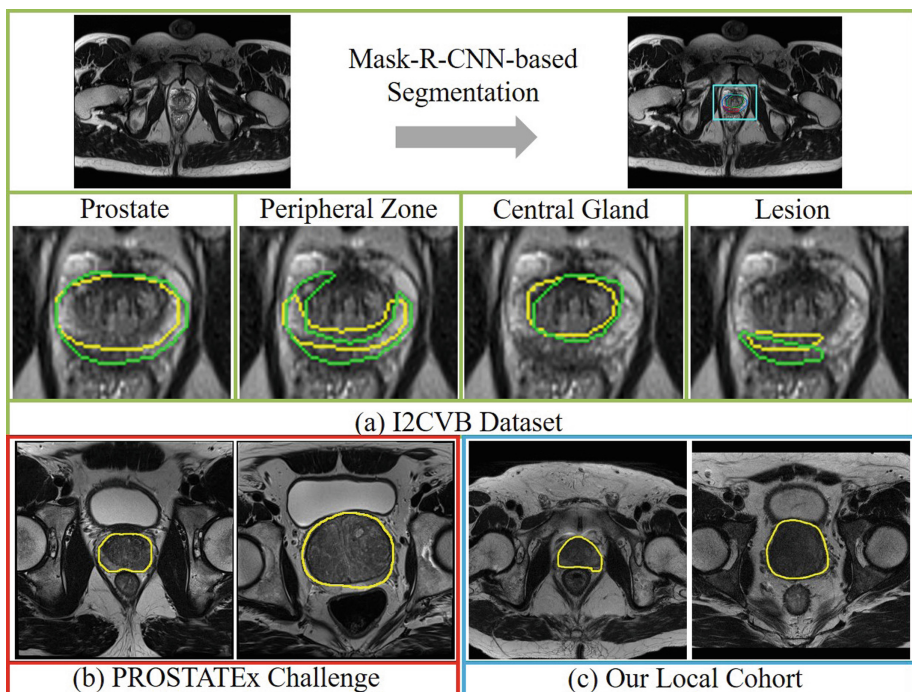


Fig. 2. Samples of automated prostate segmentation with Mask R-CNN. (a) Results for test split of I2CVB dataset. The ground truth is indicated in green, and the predicted area is in yellow. Predicted prostate area of PROSTATEx Challenge dataset (b) and our local cohort (c). (Color figure online)

3.2 Prostate Lesion Detection and Classification

In the *second* stage, our proposed weakly supervised network is trained and validated on PROSTATEx Challenge dataset (330 lesion samples, 76 malignant and 254 benign) and our local cohort (74 lesion samples, 51 malignant lesions and 23 benign lesions). We used GLEASON score (malignant for values ≥ 7) to label lesion malignancy on both datasets. We trained two sets of models on these two datasets. Each set of models consists of models trained on three sequences, namely T2-weighted images, ADC maps and high b -value diffusion-weighted images. All the segmented prostate regions were prepared for both datasets in stage one. Data augmentation was applied to the training data with random rotation of $[\pm 3^\circ, \pm 6^\circ, \pm 9^\circ, \pm 12^\circ, \pm 15^\circ]$. The image sizes were then scaled to 224×224 pixels for training. We conducted 5-fold cross validation experiments on both datasets. Through repeated experimental trials, parameters weights λ_C and λ_S were both set to 1, with ε set to 10 -5 . The final classification results and AUC were obtained through ensemble learning from all three sequences.

Table 1. Comparison of Mean IoU on I2CVB dataset with other methods.

Methods	Prostate	Central gland	Peripheral zone	Lesion
Ruba Alkadi - M1 [14]	–	0.673	0.563	0.677
Ruba Alkadi - M2 [14]	–	0.657	0.599	0.679
Mask R-CNN (ours)	0.843	0.781	0.516	0.405

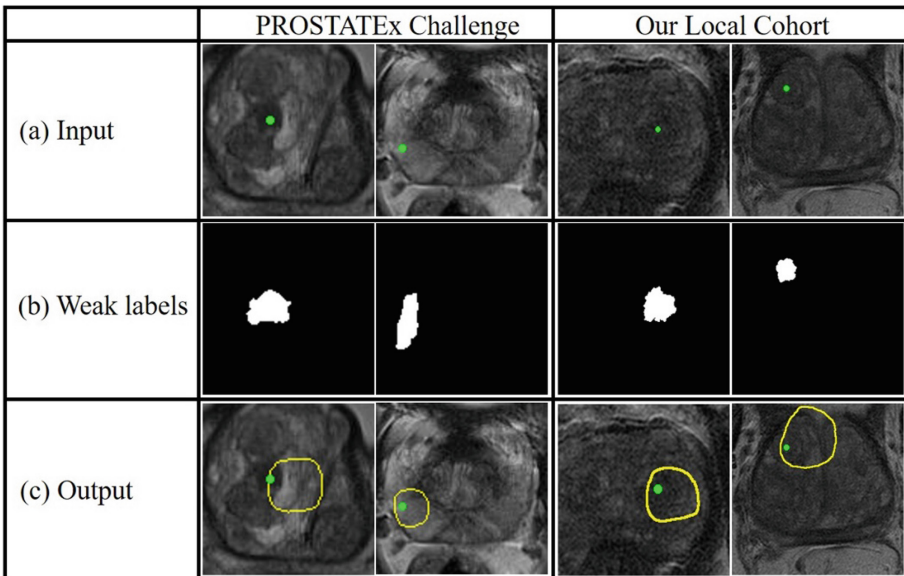


Fig. 3. Samples of result for lesion detection of T2-weighted images. Green spots indicate the given lesions centroid. (a) Input of raw prostate region. (b) Weak lesion mask obtained from level set method. (c) Predicted lesion area contoured by yellow. (Color figure online)

Figure 3 shows the results of lesion detection in T2 sequences. Figure 4a and b illustrate the final average receiver operating characteristic curve (ROC curve) on the two prostate lesion datasets with our proposed approach. For PROSTATEx Challenge dataset, our model outputs the average AUC of 0.912 with ensemble input sequences of ADC maps and T2-weighted images, outperforming the other two existing methods that individually combine hand-craft features and learned features (AUC of 0.870) [6], and encoder-decoder architecture (AUC of 0.834) [8]. For our local cohort, the average AUC of 0.882 is obtained with ensemble learning over input sequences. Moreover, we compare the results *without* pre-segmentation. Figure 4c and d indicate significantly *lower* AUC on the two datasets with the non-segmented/cropped image as input, showing the crucial role of pre-segmentation in this detection/classification framework.

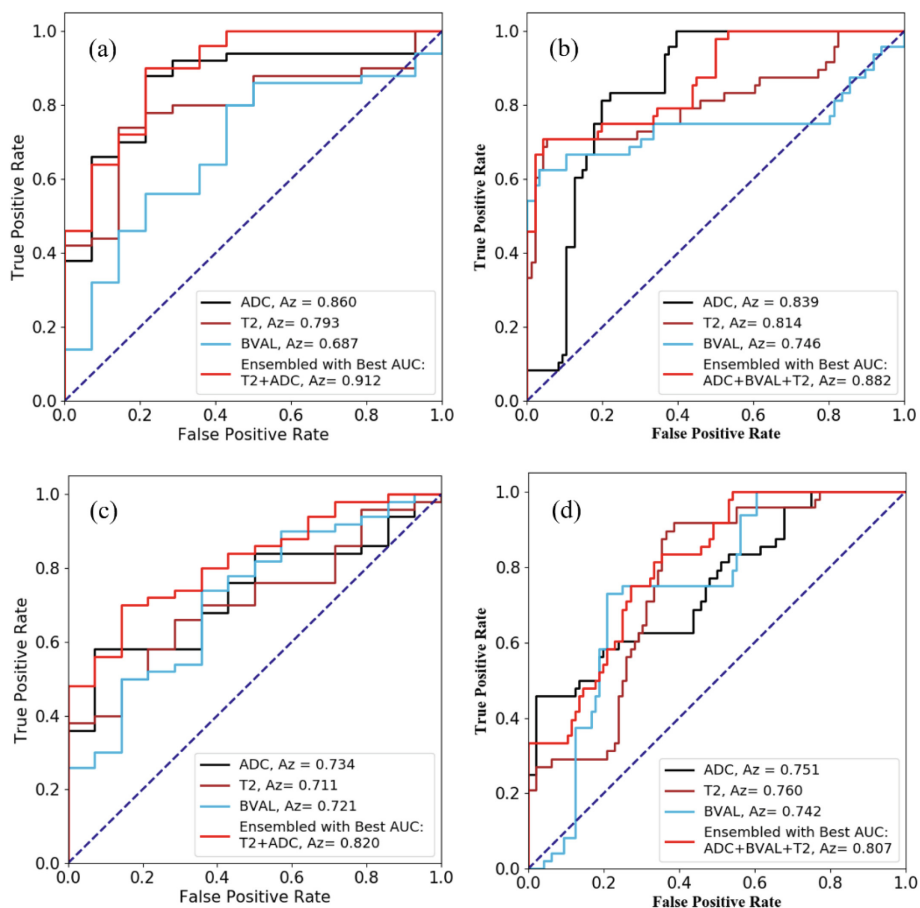


Fig. 4. Average AUC results on the two datasets: PROSTATEx Challenge (Left column) and our local cohort (Right column). AUC results (a) and (b) are input *with* prostate region segmented by Mask R-CNN models. Lower AUC can be observed in (c) and (d) *without* the segmentation.

4 Conclusion

This paper proposes a novel framework for *fully*-automated prostate lesion detection and diagnosis in MR images. Experiments on PROSTATEx Challenge dataset and our local cohort achieve a promising average AUC of 0.912 and 0.882 on their validation set respectively. The resultant efficacy is comparable to the first (champion) and second highest in AUC in PROSTATEx challenge [7], which achieved 0.87 and 0.84 on the test set, respectively. Our proposed method is extensible to other structures demanding for similar lesion diagnosis using MRI. For future work, we will extend the framework 3D MRI. To improve robustness, we will also attempt to consider using specific regions, such as segmented central gland and peripheral prostate zones for classification.

Acknowledgements. This work is supported by the Research Grants Council (RGC) of HK (Ref. No.: 17202317, 17227616, 17206818, 27209515), and Innovation and Technology Commission (Ref. No.: UIM/353).

References

1. Prostate Cancer International Inc. <https://prostatecancerinfolink.net/2018/01/09/prostate-cancer-projections-for-2018/>. Accessed 02 Apr 2019
2. American Cancer Society Inc. <https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html>. Accessed 02 Apr 2019
3. Oberlin, D.T., Casalino, D.D., Miller, F.H., Meeks, J.J.: Dramatic increase in the utilization of multiparametric magnetic resonance imaging for detection and management of prostate cancer. *Abdom. Radiol.* **42**(4), 1255–1258 (2017)
4. Hassanzadeh, E., Glazer, D.I., Dunne, R.M., Fennessy, F.M., Harisinghani, M.G., Tempany, C.M.: Prostate imaging reporting and data system version 2 (PI-RADS v2): a pictorial review. *Abdom. Radiol.* **42**(1), 278–289 (2017)
5. Sonn, G.A., Fan, R.E., Ghanouni, P., et al.: Prostate magnetic resonance imaging interpretation varies substantially across radiologists. *Eur. Urol. Focus* **5**(4), 592–599 (2017)
6. Karimi, D., Ruan, D.: Synergistic combination of learned and hand-crafted features for prostate lesion classification in multiparametric magnetic resonance imaging. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 391–398. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_45
7. Liu, S., Zheng, H., Feng, Y., Li, W.: Prostate cancer diagnosis using deep learning with 3D multiparametric MRI. In: Samuel, G.A., Nicholas, A.P. (eds.) *Medical Imaging 2017: Computer-Aided Diagnosis*. International Society for Optics and Photonics, vol. 10134, p. 1013428 (2017). <https://doi.org/10.1117/12.2277121>
8. Kiraly, A.P., et al.: Deep convolutional encoder-decoders for prostate cancer detection and classification. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 489–497. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_56
9. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961–2969 (2017)
10. Litjens, G., Debats, O., Barentsz, J., Karssemeijer, N., et al.: Computer-aided detection of prostate cancer in MRI. *IEEE Trans. Med. Imaging* **33**(5), 1083–1092 (2014)

11. Lemaitre, G., Marti, R., Freixenet, J., Vilanova, J.C., Walker, P.M., et al.: Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: a review. *Comput. Biol. Med.* **60**, 8–31 (2015)
12. Li, C., Xu, C., Gui, C., et al.: Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. Image Process.* **19**(12), 3243–3254 (2010)
13. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014)
14. 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*, 1–15 (2018)