HHS Public Access

Author manuscript

Int J Comput Assist Radiol Surg. Author manuscript; available in PMC 2020 October 01.

Published in final edited form as:

Int J Comput Assist Radiol Surg. 2019 October ; 14(10): 1647–1650. doi:10.1007/s11548-019-01967-5.

Prostate Cancer Detection using Residual Networks

Helen Xu, PhD,
Ezra AI Canada, Unit 310, 545 King St. West, Toronto, Canada

John S.H. Baxter, PhD,
Université de Rennes 1, Rennes, France

Oguz Akin, MD [Professor of Radiology],
Memorial Sloan Kettering Cancer Center, New York, NY, USA

Diego Cantor-Rivera, PhD
Ezra AI Canada, Unit 310, 545 King St. West, Toronto, Canada

Abstract

Purpose: To automatically identify regions where prostate cancer is suspected on multi-parametric magnetic resonance images (mp-MRI).

Methods: A Residual Network was implemented based on segmentations from an expert radiologist on T2-weighted, apparent diffusion coefficient map, and high b-value diffusion-weighted images. Mp-MRIs from 346 patients were used in this study.

Results: The Residual Network achieved a Hit or Miss accuracy of 93% for lesion detection, with an average Jaccard score of 71% that compared the agreement between network and radiologist segmentations.

Conclusion: This paper demonstrated the ability for Residual Networks to learn features for prostate lesion segmentation.

Keywords

prostate cancer; multi-parametric MRI; lesion segmentation; deep learning

1 Introduction

Prostate cancer remains the most commonly diagnosed cancer and one of the leading causes of cancer death for men in developed countries. In the United States, it accounts for 19% of the newly diagnosed cases and 9% of cancer death in 2018 [6]. Magnetic resonance imaging (MRI) has emerged as an alternative to the clinical standard transrectal ultrasound (TRUS)

diego@ezra.ai.

for cancer detection, localization, staging, biopsy guidance, and focal therapy. More recently, multi-parametric (mp)-MRI has played an increasingly important role in prostate cancer assessment. In addition to the excellent soft tissue contrast, mp-MRI can provide metabolic, diffusion, and perfusion information of prostatic tissue that improves the identification of possible cancerous regions. As the use of mp-MRI increases in clinical practice, automatic lesion detection and segmentation as a part of the clinical decision support system can help radiologists interpret images faster and more accurately.

Recent advances in deep Convolutional Neural Networks (CNNs) have seen great success in natural image classification and medical image analysis. However, only a limited number of groups have conducted research in the use of CNN for the detection of prostate cancer. Inspired by the Visual Geometry Group (VGG) Net, Liu *et al.* [8] developed Xmas-Net for the ProstateX dataset. Data augmentation was used to address heterogeneity and small sample size. Song *et al.* [10] introduced a patch-based approach and augmented the test images to yield better prediction accuracy. Yang *et al.* [9] and Wang *et al.* [12] proposed similar CNNs based on GoogLeNet. Systematic 12-core TRUS-guided biopsy were used to train the model. Tsehay *et al.* [11] proposed a five layer CNN using ground truth obtained from MRI-ultrasound fusion biopsy. Finally, the use of adversarial networks to segment aggressive prostate cancer was explored by Kohl *et al.* [7].

Residual Networks (ResNets) are able to train deeper neural networks easier and faster [3]. This paper presents a deep learning framework using ResNets to identify suspicious lesions on prostate mp-MRI.

2 Methods

Mp-MR images of 346 subjects were downloaded from the Cancer Imaging Archive data portal [1]. For each subject, three radiologists evaluated axial T2-weighted (T2W), apparent diffusion coefficient (ADC) map, and high b-value (BVAL) diffusion-weighted images, and segmented lesions that were PI-RADS v2 category 3 or greater, using 3D Slicer. The segmentations generated by the most expert radiologist (genitourinary radiologist with 15 years of experience) were used as ground truth for network training and validation. Segmentations produced by the other two radiologists were used to establish a baseline comparison of the network performance.

For each subject, T2W, ADC, BVAL and segmentation images were resampled to a common volume. This volume was axially sliced to generate 2D images for T2W, ADC, BVAL and ground truth. A connected components algorithm was used to identify lesions in the ground truth image. Each lesion was then randomly sampled using disk point picking[1] centered in the lesion and with a radius equal to the maximum of the width and height of the lesion. Square patches of pixel size 75×75 centered at the sampled points were generated. Each patch contained three channels: T2W, ADC, and BVAL. Similarly, background patches were generated with the constraint that at least 85% of the patch contained non-lesion pixels. A total of 93K patches were generated. Subjects were divided into training (70%, 65K patches)

[1] http://mathworld.wolfram.com/DiskPointPicking.html

and validation (30%, 28k patches) sets, both containing an equal distribution of background and lesion patches.

A ResNet was trained using a class-weighted cross-entropy loss function with a weight of 2 for lesions and 1 for background classes. Training was performed using stochastic gradient descent with momentum. Batch normalization was employed following each convolution layer inside a residual block prior to the respective ReLU activation. The network was configured to train for a maximum of 50 epochs with an early-stop criteria. Online data augmentation was used on each epoch by applying a random 90 degree rotation or horizontal/vertical flip on each sample in the training batch. We implemented semantic segmentation [2] using a series of atrous convolutions to learn long-distance relationships without using pooling layers to retain high-frequency details [4]. A Softmax function was used to obtain pixel-wise probability maps for lesions. Since lesion identification is a binary problem, the probability maps were thresholded at 50% ($p \geq 0.5$) to obtain binarized predictions prior to evaluating the network performance.

Performance was measured using a Hit or Miss (HoM) accuracy metric comparing the ground truth (GT) and network prediction (PR) for a given patch. If the GT and PR intersected, the prediction was labelled a hit. A similar heuristic was applied for lesion absent patches. Mismatches between the GT and PR were labelled as a miss. We also evaluated the segmentation between GT and PR by using a Jaccard score. Both HoM and Jaccard scores were evaluated for the 103 subjects in the validation set, which were not available to the network during training.

3 Results

To avoid overfitting, training was stopped after the 12th epoch. This was the inflection point where improvement in validation loss was minimal compared to the training loss. All subsequent epochs showed a continuous improvement in training loss with no improvement of the respective validation loss. The ResNet detected lesions with an average HoM accuracy of 93% and achieved an average Jaccard score of 68%. A receiver operating curve (ROC) analysis demonstrated an area under the curve (AUC) of 97% (Figure 2).

Lesions in whole-axial slices were predicted by combining patch-based predictions (Figure 1). The accuracy for whole slice lesion detection remained the same (HoM = 93%), while the Jaccard score improved to 71%. A quantitative analysis comparing lesions outlined by the network against three radiologists showed a higher agreement with the segmentations of the most junior radiologist (average Jaccard score = 73%).

The distance between biopsy location and segmentation boundary were also calculated in order to compare the segmentation with histology results from MRI-guided biopsy. Considering the size of a clinically significant tumour [5] and the amount of segmentation inter-observer variability, a distance of 10 mm or less was considered as a match. Each segmentation was matched to at most one biopsy. The expert radiologist was able to detect 96.1% of the clinically significant cancers while the network detected 82.6%.

4 Discussion

This study has demonstrated that Residual Networks can be trained to learn features useful for segmenting prostate lesions suspicious for cancer. The results achieved are comparable to similar state-of-the-art studies: we obtained an AUC of 97%, whereas two of the leading groups [10] and [12] obtained 94% and 96%, respectively.

In standard clinical practice, a PI-RADS score is assigned to each lesion to indicate its probability of being malignant. A score of 1–5 is used with 1 being most likely to be benign and 5 being highly suspicious of malignancy. Incorporating this information along with the lesion segmentation provides a higher clinical value and is an element of future research for our group. Additional patient information such as age and medical history could also improve the network's performance. At this stage, we have obtained lesion segmentation results with 93% HoM accuracy compared to an expert radiologist. Our approach has the potential to reduce time radiologists spend on interpreting images, but histopathology data from biopsy is still needed for a definitive diagnosis of prostate cancer. In the future, this work will be extended to include segmentation of different prostate regions as this is highly relevant for the detection and localization of cancerous lesions.

Conflict of Interest

Dr. Oguz Akin receives support from Memorial Sloan Kettering Cancer Center Support Grant/Core Grant (P30 CA008748). Dr. Oguz Akin, Dr. Helen Xu and Dr. Diego Cantor-Rivera hold stock options and serve as scientific advisors for Ezra AI Inc., which is developing artificial intelligence algorithms related to the research being reported in this paper.

References

1. Litjens G, Debats O, Barentsz J, Karssemeijer N, and Huisman H. ProstateX Challenge data, The Cancer Imaging Archive (2017). 10.7937/K9TCIA.2017.MURS5CL.
2. Long J, Shelhamer E, and Darrel T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3431–3440, 2015.
3. He K, Zhang X, Ren S, and Sun J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
4. Chen L-C, Papandreou G, Schroff F, and Adam H. Rethinking atrous convolution for semantic image segmentation. CoRR, abs/1706.05587, 2017.
5. Guillaume Ploussard, Jonathan I Epstein, Rodolfo Montironi, Peter R Carroll, Manfred Wirth, Marc-Oliver Grimm, Anders S Bjartell, Francesco Montorsi, Stephen J Freedland, Andreas Erbersdobler, et al. The contemporary concept of significant versus insignificant prostate cancer. European urology, 60(2):291–303, 2011. [PubMed: 21601982]
6. Siegel RL, Miller KD, and Jemal A. Cancer statistics 2018. CA Cancer J Clin, 68:7–10, 2018. [PubMed: 29313949]
7. Kohl S, Bonekamp D, Schlemmer HP, Yaqubi K, Hohenfellner M, Hadaschik B, Radtke JP, and Maier-Hein K. Adversarial networks for the detection of aggressive prostate cancer. arXiv preprint: 1702.08014, 2017.
8. Liu S, Zheng H, Feng Y, and Li W. Prostate cancer diagnosis using deep learning with 3D multiparametric MRI. SPIE Medical Imaging: Computer-Aided Diagnosis, 10134:1013428, 2017.
9. Yang X, Wang Z, Lin C, Le HM, Chen J, Cheng KT, and Wang L. Joint detection of diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks. MICCAI 2017, Part III LNCS 10435:426–434, 2017.

10. Song Y, Zhang YD, Yan X, Liu H, Zhou M, Hu B, and Yang G. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. J Magn Reson Imaging, 48(6):1570–1577, 2018. [PubMed: 29659067]

11. Tsehay Y, Lay N, Wang X, Kwak JT, Turkbey B, Choyke P, Pinto P, Wood B, and Summers R. Biopsy-guided learning with deep convolutional neural networks for prostate cancer detection on multiparametric MRI. IEEE 14th International Symposium on Biomedical Imaging, pages 642–645, 2017.

12. Wang Z, Liu C, Cheng D, Wang L, Yang X, and Cheng KT. Automated detection of clinically significant prostate cancer in mp-MRI images based on an end-to-end deep neural network. IEEE transactions on medical imaging, 37(5):1127–1139, 2018. [PubMed: 29727276]

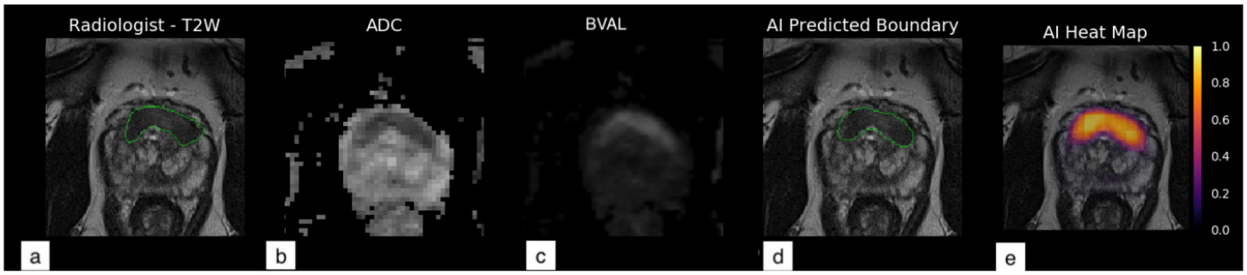

Fig. 1.

Images a-c show the type of images used for training: axial T2W, ADC map, and high b-value diffusion-weighted. The radiologist segmentation is shown on the T2W image. Two different visualizations of the lesion predicted by ResNet are provided: d shows the lesion as a boundary and e is the Softmax output.

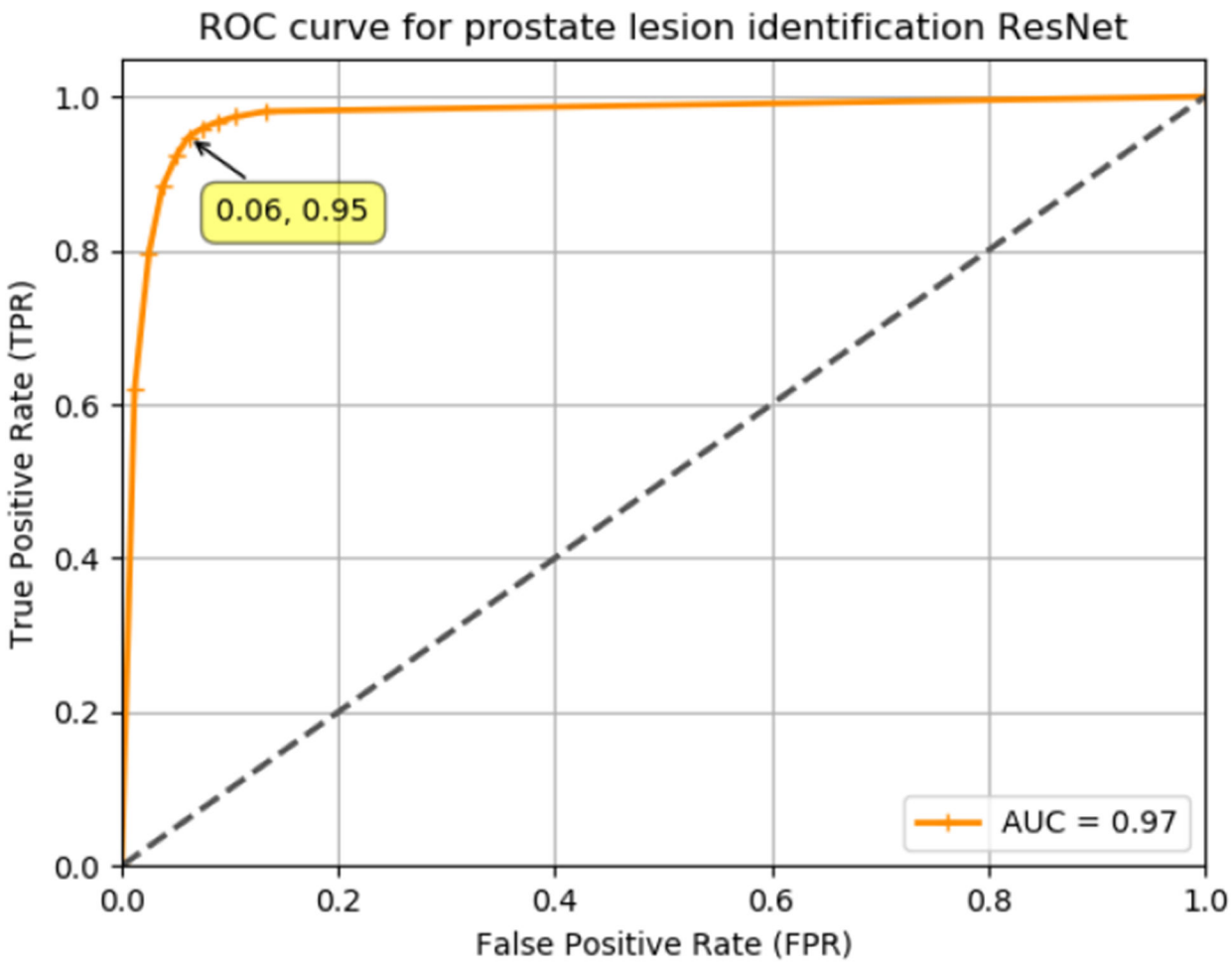

Fig. 2.

Receiver operating curve (ROC) for ResNet evaluated over 103 subjects constituting the validation set. The arrow indicates the False Positive Rate (FPR) and True Positive Rate (TPR) for a configuration where the probability map for lesions (Softmax output) is thresholded at 50%.

Table 1

List of acronyms

Acronym	Definition
ADC	Apparent Diffusion Coefficient
AUC	Area Under the Curve
BVAL	b-value diffusion-weighted images
CNN	Convolutional Neural Network
GT	Ground Truth
HoM	Hit or Miss
MRI	Magnetic Resonance Imaging
mp-MRI	Multi-parametric MRI
PI-RADS	Prostate Imaging - Reporting and Data System
PR	Prediction
ResNet	Residual Network
ROC	Receiver Operating Curve
VGG	Visual Geometry Group
TRUS	Transrectal Ultrasound