

An Adapted UNet Convolutional Neural Network Architecture for Automatic Prostate MRI Segmentation

Nooshin Ghavami^{1*}, Yipeng Hu^{1,2}, Eli Gibson^{1,2}, Dean Barratt^{1,2}

¹ UCL Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering

² Wellcome/EPSRC Centre for Interventional and Surgical Sciences
University College London, London, UK

* nooshin.ghavami.15@ucl.ac.uk

June 2018

Abstract. We describe a deep learning method for automatic prostate MRI segmentation that uses convolutional neural networks (CNNs), based on a UNet architecture with the addition of residual shortcuts. The model was trained and tested on the data from the PROMISE12 challenge.

1 Introduction

Convolutional neural networks (CNNs) have yielded high accuracies for automatic prostate segmentations and cancer detection from MR images [1-3], using a variety of network architectures. In this paper, we evaluate a CNN architecture based on a UNet, proposed previously for the segmentation of transrectal ultrasound images of the prostate [4], for automatic segmentation of the prostate gland in MR images from the dataset in the PROMISE12 challenge [5]. We report the accuracy of the CNN-based method based on a 10-fold cross-validation using the publicly available training images and ground-truth segmentation from the challenge.

2 Methods

2.1 Data and Pre-processing

The PROMISE-12 challenge provides 50 T2-weighted MR prostate images with corresponding segmentations, as ground-truth, and an independent set of 30 MR images without ground-truth segmentations, for the purposes of training and evaluation of segmentation algorithms, respectively [5]. All images were normalised to have zero mean with unit variance intensity and resized to a volume size of $181 \times 181 \times 30$ voxels.

2.2 Network Architecture

The proposed ‘Adapted UNet’ algorithm uses a fully-convolutional 3D neural network based on a UNet architecture [6], with three modifications; First, convolutional layers are replaced by residual network blocks after each down- or up-sampling block. Second, concatenation is replaced by summation shortcuts before each down-sampling block to the output feature maps, and thirdly, incorporation of additive up-sampling shortcuts. The network takes a volume of size $S_0=181\times 181\times 30$ as input and propagates to feature maps of the same size with 8 initial channels using a convolution (Conv), a batch normalisation (BN) and a nonlinear rectified linear unit (ReLU). The resulting architecture is shown in Figure 1. All the convolution kernels are of size $3\times 3\times 3$. The feature maps are then down-sampled to $K=5$ different resolution levels by down-sampling blocks and each followed by a residual network unit (Resnet) block. For each level k , where $k=1, 2, \dots, K$, the number of channels is doubled and size is halved. Each down-sampling block consists of Conv, BN and ReLU, followed by a max-pooling layer with stride 2, whilst each Resnet block has two Conv layers with BN and ReLU, and an identity shortcut over these layers. Reverse Resnet blocks are also included with the addition of additive up-sampling shortcut layers over the transpose convolution (Deconv) layers [7].

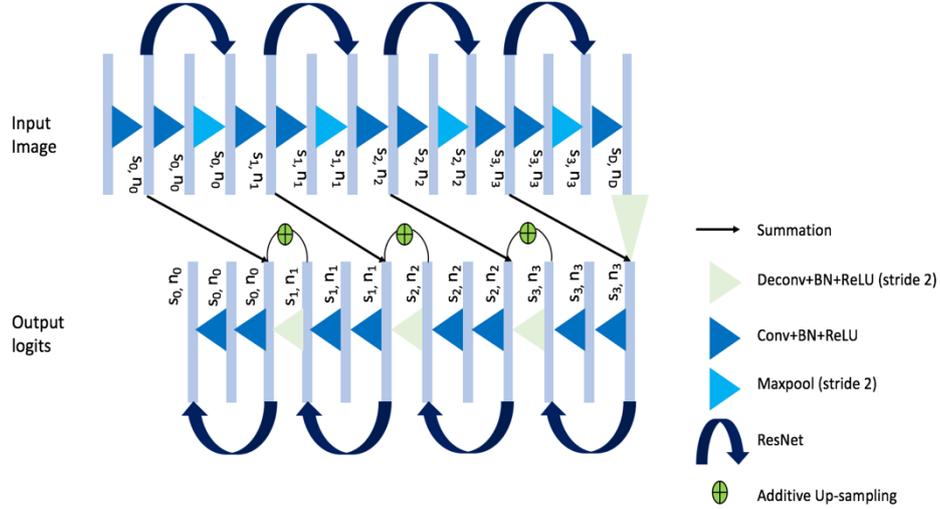


Figure 1: Architecture of the adapted UNet convolutional neural network. s_k and n_k denote different activation map sizes and number of channels.

2.3 Implementation Details and Training

The network described above was implemented in TensorFlowTM and trained on a 12GB NVIDIA[®] Titan GPU using a minibatch size of 4. The results presented in this

work were obtained by minimising a negative probabilistic Dice score. The Dice score is differentiable with an added L^2 -norm weight decay on the trainable parameters, with the weighting parameter being set to 1×10^{-3} . The network was trained for a total of 5000 iterations using the Adam optimiser with a learning rate of 0.01.

2.4 Performance Testing

To test the network performance during algorithm development, we used 10-fold cross-validation experiments with the 50 images that have available corresponding reference standard segmentations from the challenge training dataset. For each fold, 5 images were “left-out” and used for testing while the remaining 45 images were used to train the network. The Dice score and symmetric boundary distance were calculated as the measures of segmentation accuracy. The symmetric boundary distance is defined as the average of the mean absolute value of the distances between all the points from the automatically segmented boundary and the closest boundary points found on the left-out ground-truth, and vice versa.

3 Results

The mean (\pm standard deviation) values of the Dice score and boundary distance over all the folds were 0.84 ± 0.08 and 2.51 ± 1.23 mm, respectively. For illustration purpose, the automatic segmentations output by the network on the testing dataset for three arbitrarily chosen MRI slices, overlaid on the original images and compared with the ground-truth segmentations are shown in Figure 2. The submitted segmentations on the challenge test image data were obtained by one network trained using all 50 images and corresponding segmentations from the challenge training dataset.

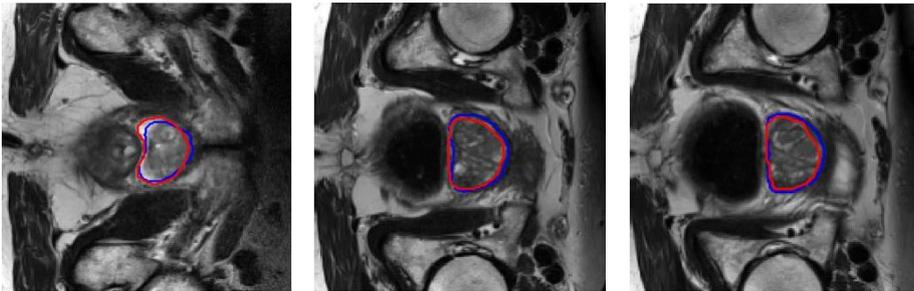


Figure 2: Automatic (blue) and manual segmentations (red) shown as overlays on three MRI slices.

Acknowledgements I would like to acknowledge the UCL EPSRC Centre for Doctor Training in Medical Imaging for funding this work (EPSRC Grant No. EP/L016478/1). Eli Gibson and Yipeng Hu are funded by Cancer Research UK

(CRUK/EPSRC Grant No. C28070/A19985; NS/A000030/1). Yipeng Hu also received funding from EPSRC grant EP/M020533/1.

References

1. Liu, S., Zheng, H., Li, W. (2017). Prostate cancer diagnosis using deep learning with 3D multiparametric MRI. *Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis*, 1013428.
2. Tsehay, Y., *et al.* (2017). Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images. *Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis*, 1013405
3. Liao, S., *et al.* (2013). Representation learning: a unified deep learning framework for automatic prostate MR segmentation. *MICCAI 2013, Part II. LNCS*, 8150, pp. 254–261
4. Ghavami, N., *et al.* (2018). Automatic slice segmentation of intraoperative transrectal ultrasound images using convolutional neural networks. *Proc. SPIE 10567, Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modelling*, 1057603 (12 March 2018); doi: 10.1117/12.2293300.
5. Litjens, G., *et al.* (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, pp.60-88.
6. Cicek, O., *et al.* (2016). 3D U-net: Learning dense volumetric segmentation from sparse annotation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*.
7. Wojna, Z., *et al.* (2017). The Devil is in the Decoder. *British Machine Vision Conference*.