

# Cascading Classifiers Improve Prostate Segmentation

Ronald J Nowling<sup>1</sup>, John Bukowy<sup>2</sup>, Sean D McGarry<sup>3</sup>, Andrew S Nencka<sup>2</sup>, Jay Urbain<sup>1,4</sup>, Allison Lowman<sup>2</sup>, Alexander Barrington<sup>5</sup>, Mark Hohenwalter<sup>2</sup>, Anjishnu Banerjee<sup>6</sup>, Kenneth A Iczkowski<sup>7</sup>, Peter S LaViolette<sup>2,5</sup>

<sup>1</sup>Electrical Engineering and Computer Science, Milwaukee School of Engineering, Milwaukee, WI

<sup>2</sup>Radiology, Medical College of Wisconsin, Milwaukee, WI

<sup>3</sup>Biophysics, Medical College of Wisconsin, Milwaukee, WI

<sup>4</sup>Clinical and Translational Science Institute, Medical College of Wisconsin, Milwaukee, WI

<sup>5</sup>Biomedical Engineering, Medical College of Wisconsin, Milwaukee, WI

<sup>6</sup>Biostatistics, Medical College of Wisconsin, Milwaukee, WI

<sup>7</sup>Pathology, Medical College of Wisconsin, Milwaukee, WI

## Synopsis (100 words)

We evaluated the U-Net segmentation model on prostate segmentation using data from 39 patients, achieving a Dice score of 73.9%. We improved segmentation performance by applying a convolutional neural network (CNN) to determine whether slices have prostates. Images with prostates are then forwarded to a U-Net model for segmentation. Our two-phase approach achieves a higher Dice score of 85.2%.

## INTRODUCTION

Prostate segmentation is a necessary pre-processing step for computer-aided detection and diagnosis algorithms for prostate disorders and associated cancers<sup>7</sup>.

Convolutional neural networks (CNN) are a popular class of deep learning models<sup>3-6</sup> that have enabled significant advances in image-based machine learning tasks. The U-Net<sup>1</sup> and V-Net<sup>2</sup> models, which combine CNNs with variational autoencoders (VAEs), have recently been proposed for biomedical image segmentation and received significant interest. We evaluated the U-Net model on a prostate segmentation task using T2-weighted images from 39 patients (Dice score of 73.9%). Using a cascading classifiers<sup>8</sup> approach (classification of slices followed by segmentation of slices), we were able to increase the overall Dice score to 85.2%.

## METHODS

### Data Set

The data set consisted of 39 patients with prostate cancer (mean age 60 years | mean PSA 8.2 ng/mL). Only T2-weighted images were considered in pipeline for localization. All images were collected on a 3-T MRI scanner (GE) using an endorectal coil. Datasets were stratified randomly by patient - training (75%) | testing (25%). Ground truth prostate masks were drawn by a single, trained observer.

### Segmentation Model

Images were segmented with the U-Net model<sup>1</sup>. The U-Net model combines convolutional neural networks (CNNs) with variational autoencoders (VAEs). Training was performed with a cross-entry loss function. For the segmentation of prostate-only slices, slices without prostates were removed from the

training set and slices were augmented with rotation and translation (using periodic boundaries). Performance was evaluated using Dice score calculated over all images and on individual images.

### **Classification Model**

The goal of the classification task was to classify slices as having or not having prostates. Our classification model is a convolutional neural network (CNN) based on the forward part of the U-Net model. Labels were generated by identifying empty masks. Interfacial slices were defined as those whose immediate slices above and below did not both have or not have a prostate.

### **Integrated Pipeline**

For the integrated pipeline, the two models were trained using the same training set. For the classification model, the training set slices were used as-is. For training the segmentation model, empty slices were removed, and the remaining slices were augmented. For testing, classification model was applied to the testing set slices. Slices predicted to have prostates were passed to the segmentation model. The Dice score was calculated over all images passed to the segmentation phase.

## **RESULTS**

We trained and tested the U-Net model on prostate segmentation of T2-weighted images. The model achieved a maximum overall Dice score of 73.9% with 100 training epochs. Attempts to improve the results with additional training epochs resulted in the model overfitting by outputting only empty masks (see Figure 1). When trained and tested only on images with prostates, augmented with rotations and translations, the model achieved an improved Dice score of 87.8% after 750 training epochs. Our results suggest that the U-net is more effective at segmentation when images without prostates are filtered out before segmentation.

Based on our results with prostate-only images, we designed a pipeline of cascading classifiers. Images are first classified as having prostates or not using a separate classification model; images predicted to have prostates are then segmented using the U-Net model.

Our classification model is based on the forward Convolutional Neural Network (CNN) of the U-Net model. In a 4-fold cross-fold validation, our classification model achieved a per-slice accuracy of 89.1%. We analyzed the misclassified slices and found that 46.0% of the misclassified slices were interfacial. When the interfacial slices were excluded from the accuracy calculation, our classification accuracy went up to 93.5%.

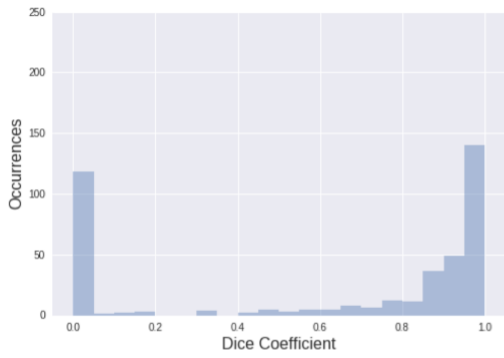
The complete cascading classifiers pipeline achieved a Dice score of 85.2%, close to the Dice score achieved by the segmentation model on images filtered by their ground truth labels.

## **DISCUSSION AND CONCLUSION**

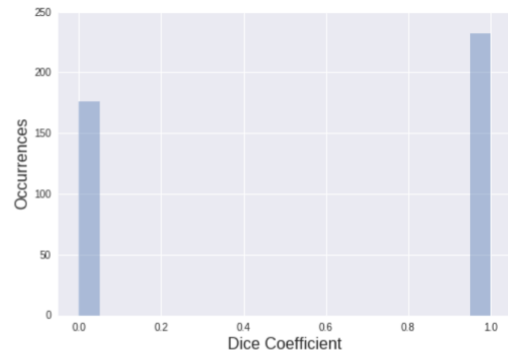
The U-Net and V-Net models have recently been proposed for biomedical image segmentation and received significant interest. We evaluated the U-Net model on our data set and found that it generated a large number of false predictions. When applied only to slices with prostates, the Dice score of the U-Net model increased from 73.9% to 87.8%. In response, we proposed dividing the segmentation problem into two sub-tasks: classification followed by segmentation. The classification model is able to accurately classify 89.1% of the slices. A pipeline integrating the two models achieved a Dice score of 85.2%. Our approach represents a simple but practical way to improve the segmentation performance of the U-Net model. Future work will focus on improving the accuracy of the classification model.

## REFERENCES

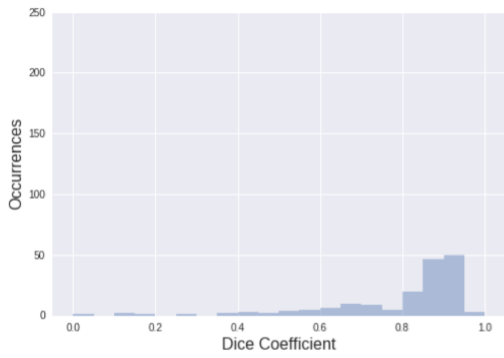
1. Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham
2. Milletari, F., N. Navab, and S. Ahmadi. 2016. "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation." In *2016 Fourth International Conference on 3D Vision (3DV)*, 565–71.
3. Liao S., Gao Y., Oto A., Shen D. (2013) Representation Learning: A Unified Deep Learning Framework for Automatic Prostate MR Segmentation. In: Mori K., Sakuma I., Sato Y., Barillot C., Navab N. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013. MICCAI 2013. Lecture Notes in Computer Science, vol 8150. Springer, Berlin, Heidelberg
4. Guo, Yanrong, Yaozong Gao, and Dinggang Shen. 2016. "Deformable MR Prostate Segmentation via Deep Feature Learning and Sparse Patch Matching." *IEEE Transactions on Medical Imaging* 35 (4): 1077–89.
5. Ruida Cheng, Holger R. Roth, Le Lu, Shijun Wang, Baris Turkbey, William Gandler, Evan S. McCreedy, Harsh K. Agarwal, Peter Choyke, Ronald M. Summers, Matthew J. McAuliffe, "Active appearance model and deep learning for more accurate prostate segmentation on MRI," Proc. SPIE 9784, Medical Imaging 2016: Image Processing, 97842I (21 March 2016);
6. Saifeng Liu, Huaixiu Zheng, Yesu Feng, Wei Li, "Prostate cancer diagnosis using deep learning with 3D multiparametric MRI," Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis, 1013428 (3 March 2017);
7. Litjens, Geert, Robert Toth, Wendy van de Ven, Caroline Hoeks, Sjoerd Kerkstra, Bram van Ginneken, Graham Vincent, et al. 2014. "Evaluation of Prostate Segmentation Algorithms for MRI: The PROMISE12 Challenge." *Medical Image Analysis* 18 (2): 359–73.
8. Gama, João, and Pavel Brazdil. 2000. "Cascade Generalization." *Machine Learning* 41 (3): 315–43.



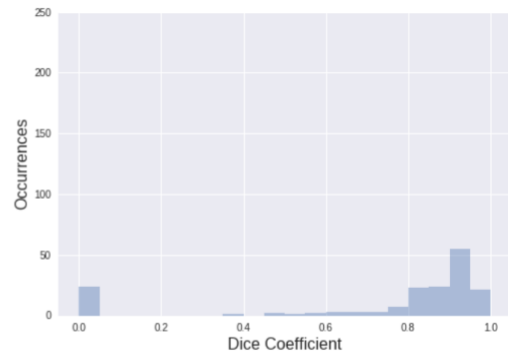
**(a) All Slices, 100 epochs**



**(b) All Slices, 200 epochs**

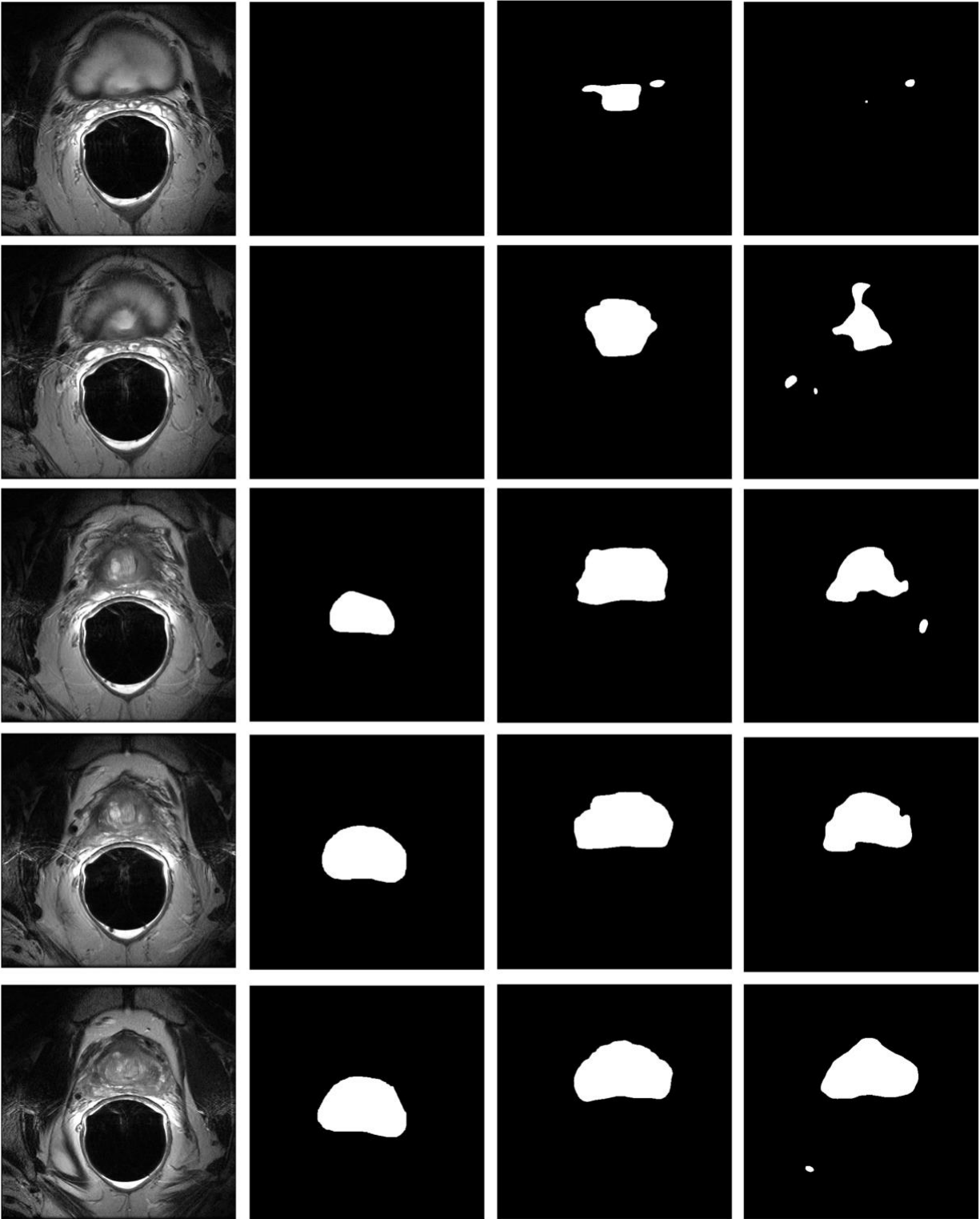


**(c) Prostate-only Slices, 100 epochs**



**(d) Predicted Prostate Slices, 750 epochs**

Figure 1 Histograms of Per-Image Dice Scores



(a) Image

(b) Manual Segmentation

(c) U-Net All Images

(d) Cascading U-Net

Figure 2 Image and Masks for a Single Patient