2D Semantic Segmentation of the Prostate Gland in Magnetic Resonance Images using Convolutional Neural Networks

Silvia P. Vacacela and Marco E. Benalcázar

Artificial Intelligence and Computer Vision Research Lab
Department of Computer Science and Informatics
Escuela Politécnica Nacional, Quito, Ecuador
e-mails: {silvia.vacacela, marco.benalcazar}@epn.edu.ec

Abstract: Convolutional Neural Networks is one of the most commonly used methods for automatic prostate segmentation. However, few studies focus on the segmentation of the two main zones of the prostate: the central gland and the peripheral zone. This work proposes and evaluates two models for 2D semantic segmentation of these two zones of the prostate. The first model (Model-A) uses an encoder-decoder architecture based on the global U-net and the local U-net architectures. The global U-net segments the whole prostate, whereas the local U-net segments the central gland. The peripheral zone is obtained by subtracting the central gland from the whole prostate. On the other hand, the second model (Model-B) uses an encoder-classifier architecture based on the VGG16 network. Model-B performs segmentation by classifying each pixel of a Magnetic Resonance Image (MRI) into three categories: background, central gland, and peripheral zone. Both models are tested using MRIs from the dataset NCI-ISBI 2013 Challenge. The experimental results show a superior segmentation performance for Model-A, encoder-decoder architecture, ($DSC = 96.79\% \pm 0.15\%$ and $IoU = 93.79\% \pm 0.29\%$) compared to Model-B, encoder-classifier architecture, ($DSC = 92.50\% \pm 1.19\%$ and $IoU = 86.13\% \pm 2.02\%$).

Keywords: Convolutional Neural Networks, Prostate Segmentation, Central Gland, Peripheral Zone, MRIs, Encoder-Decoder, U-net, Encoder-Classifier, VGG16, NCI-ISBI 2013.

1. INTRODUCTION

Prostate cancer is one of the public health problems that affect men. The Global Cancer Observatory GLOBOCAN (2020) showed that prostate cancer was the second frequently diagnosed cancer and the sixth leading cause of death among men. Diagnosis of prostate cancer performs by visual analysis of tissue samples obtained through biopsy from the patient Heidenreich et al. (2008); Lomas and Ahmed (2020). An alternative is to use multiparametric MRIs Villers et al. (2012), which helps to check for the presence or absence of prostate cancer without the need for biopsy sampling. However, this approach requires the urologist to outline or segment the prostate border and target Lomas and Ahmed (2020) which is challenging due to: a) noise present in medical images; b) voxel intensities variation; c) finite image resolution; and d) anatomy variation of prostate among different individuals. Therefore, automatic segmentation would address the problems of manual segmentation and can improve the speed and accuracy of segmentation.

In automatic segmentation, there are some traditional methods in literature such as Atlas-based models, deformable models, threshold-based models, and region-based models Yan et al. (2019). However, recent works for prostate segmentation using Artificial Neural Networks have shown remarkable results. Therefore, we carried out a Systematic Literature Review (SLR): "Prostate Segmentation in Magnetic Resonance Image using Artificial Neural Networks". This SLR was analyzed 65 studies of interest between 2014 and 2020. The results of the SLR shown that 81.53% (53 studies) used convolutional neural networks (CNNs), whereas 15.38% (10 studies) combined CNNs with other segmentation techniques, and 3.08% (2 studies) used feed-forward neural networks. In addition, the results of the SLR shown most studies focus on segmenting the whole prostate leaving aside the main zones of the prostate. There were only 12 studies that segmented the prostate gland in its two main zones: the central gland and the peripheral zone. However, 7 studies P1 proposed by Gelder and Huisman (2018), P2 proposed by Rundo et al. (2020), P3 proposed by Jensen et al. (2019), P4 proposed by Zabihollahy et al. (2019), P5 proposed by Cao et al. (2020), P6 proposed by Khan et al. (2020), and P7 proposed by Umapathy et al. (2020) used private datasets. Whereas the remaining 5 studies P8 proposed by Benalcázar et al. (2015), P9 proposed by khan et al. (2019), P10 proposed by Liu et al. (2019), P11 proposed by Meyer et al. (2019), and P12 proposed by Aldoj et al. (2020) used public datasets. Therefore, there is still a gap for improvement in the prostate segmentation in its main parts. Also, the results of the SLR shown that the type of segmentation commonly used was 2-dimensional 2D. Taking into consideration the results of the SLR, we

⋆ Authors acknowledge the financial support provided by Escuela Politécnica Nacional.

propose this research work, and their main contributions are:

(1) Two models for 2D semantic segmentation of the prostate gland in its two main zones in MRIs using CNNs developed and compared. One of the models, Model-A was based on an encoder-decoder architecture. The other model, Model-B was based on an encoder-classifier architecture.
(2) The images used in this research are the MRIs without endorectal coil, which is more challenging for 2D semantic segmentation because it does not provide a reference point. This MRIs belong to the dataset NCI-ISBI 2013, which has segmentation labels for the main areas of the prostate.
(3) Comparison of the performance achieved by the proposed models with respect to models of SLR that use the NCI-ISBI 2013 dataset.

This article is organized as follows: The materials and method are presented in Section 2. Next, the result and discussion presents in Section 3. Finally, in Section 4 presents the conclusions of this work.

2. MATERIALS AND METHOD

This section describes the materials and methodology used to develop this work.

2.1 Prostate MRI Dataset

The National Cancer Institute (NCI) in collaboration with the International Society for Biomedical Imaging (ISBI) created the NCI-ISBI2013 public dataset, which was used to segmentation the prostate in its two main zones. The used 3T MRIs collection contained 30 transverse volumes accompanied by their respective manual segmentation or gold standard volumes. The number of slices per volume varied between 15 and 24. Also, each slice contained T2-weighted images of 320×320 pixels.

2.2 Evaluation Metrics

The metrics for evaluation the models were Dice Similarity Coefficient (DSC) and Intersection over Union (IoU), which were shown in equations (1), and (2).

$$DSC = \frac{2|X \cap Y|}{|X \cup Y|} = \frac{2TP}{2TP + FP + FN} \tag{1}$$

$$IoU = \frac{|X \cap Y|}{|X \cup Y|} = \frac{TP}{TP + FP + FN} \tag{2}$$

Where X *and* Y denoted the region of the predicted segmentation and reference segmentation, respectively in (1), and (2). The modulus sign '| |' defines the cardinal of the corresponding sets.

Also, other metrics were used, such as precision, recall and accuracy, that were obtained from the confusion matrix.

2.3 Model-A Architecture

Model-A was based on the U-net network Ronneberger et al. (2015), which had an encoder-decoder architecture. This model implemented two U-net networks in cascade: global U-Net and local U-Net. The Model-A architecture was showed in Fig. 1, where first the global U-net was used to segment the whole prostate gland. Then, the local U-net was used to segment the central gland. Subsequently, the peripheral zone is obtained by subtracting the central gland from the whole prostate. Also, the global U-Net and local U-Net architectures contained the pre-processing, data augmentation, feature extraction and segmentation, and post-processing blocks.

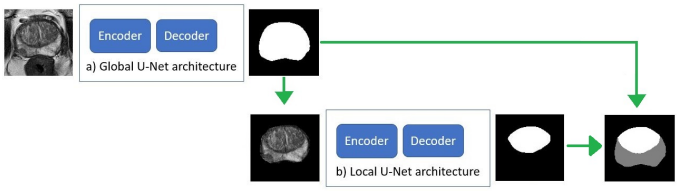

Fig. 1. Model-A based on the Encoder-Decoder Architecture

Pre-processing: regions of interest were extracted as input of the local and global U-net from 2D MRIs.

- For the global U-net input was extracted regions of interest with three different sizes (151×151 pixels, $M \times N$ pixels where $M \leq 151$ and $N \leq 151$ and mask's size) of which the region of interest with size $M \times N$ obtained superior results.
- For the U-net local input was extracted the region of interest with size $M \times N$. Additionally, we were considered the dependence or independence of results of global and local U-net, where the dependence of results of the global and local U-net obtained superior results.

Data Augmentation The dataset contained 578 2D MRIs (30 MRIs volumes). However, only 421 2D MRI contained the main zones of the prostate. To increase the number of samples of MRIs were used data augmentation techniques in global and local U-net. Among augmentation techniques were used: flipping, random rotation between $[-45°, +45°]$, zoom, and zoom with random rotation. Finally, the dataset increased in 2105 2D MRIs.

Feature Extraction and Segmentation: In the global and local U-net architecture, the encoder block extracted features using the repeated application of 3×3 convolutions, followed by dropout layer, a 3×3 convolutions, and a 2×2 max-pooling layer for downsampling. Each downsampling doubled the number of feature channels. On the other hand, the decoder block segmented the regions of interest, which consisted in the repeated application of a 2×2 transposed convolution that halves the number of feature channels then the concatenation of the corresponding cropped feature map from the encoder block, followed by one 3×3 convolutions, a dropout layer, and a 3×3 convolutions. The final layer used a 1×1 convolution that assigned the classes. Also, the convolution block used the ELU activation function in U-net. Fig. 2 shown the local U-net architecture. The global U-net architecture employed the same architecture that local U-net with different inputs and outputs. Finally, the parameters trained in global or local U-net consisted of a total of 1940817 parameters.

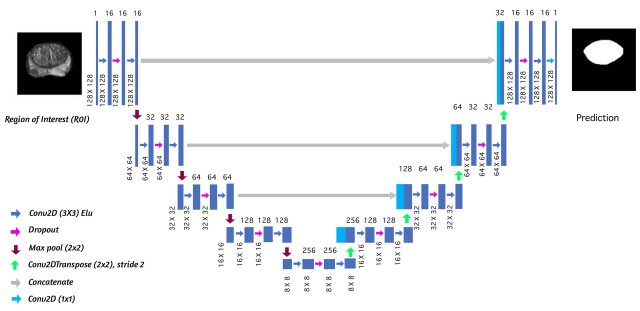

Fig. 2. Local U-Net architecture

Post-processing: helped to improve the prediction of segmentation of both the global and local U-net. The prediction result was binarized using a 0.5 threshold. Then, an opening morphological operation was applied using a disk structuring element with radius 5. Next, small areas were removed leaving the result the segmentation mask.

2.4 Model-B Architecture

The Model-B was based on VGG16 pre-training model Simonyan and Zisserman (2015), which has an encoder-classifier architecture used in classification tasks. Model-B used a sliding window to predict the class label of each pixel by providing a patch (local region). The blocks defined to create Model-B were: pre-processing, data augmentation, feature extraction and segmentation, and post-processing blocks.

Pre-processing: In this step were extracted regions of interest with size $M \times N$ from 421 2D MRIs that contained the two main zones of prostate described above. Also, patches of 15×15 pixels were extracted using a sliding window from the regions of interest. These patches were labeled depending on the central pixel as the background, the central gland, and the peripheral zone. Finally, these patches were employed as the input of the Model-B.

Data Augmentation: In the patches extraction, we noted the imbalanced data. Therefore the patches for each of the classes was manipulated as follows: a) the patches for the background was reduced a 2/3 parts, b) the patches for the central gland was maintained, and c) the patches for the peripheral zone was increased 3 *times* using data augmentation techniques such as flipping, and random rotation between $[-45°, +45°]$. The obtained patches were 161607 patches for the background, 152042 patches for the central gland, and 155040 patches for the peripheral zone.

Feature Extraction and Segmentation: Model-B used the VGG16 pre-trained model. Also, the training for Model-B consisted of a fine-tuning keeping the weights of the 15 first layers and adjusting weights of the last layers for the classification of 3 classes: central gland, peripheral zone, and background (see Fig. 3). Therefore, Model-B had a total of 14846787 parameters, where the 7211523 parameters were trained and the7635264 parameters were re-used

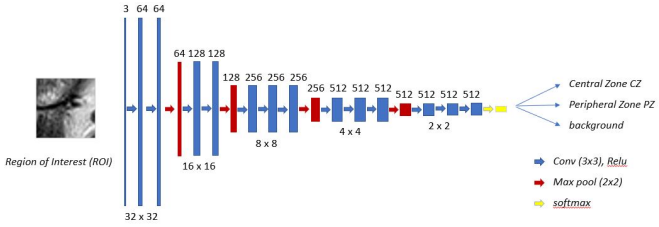

Fig. 3. Model-B based on the Encoder-Classifier Architecture

Post-processing: In the segmentation result of the MRI was removed the small areas keeping the largest areas of the prediction of the two main zones.

2.5 Implementation Details

Training and testing of the models were implemented on Tensorflow v13.1 as the machine learning framework. Also, Training and testing were run on an Nvidia Tesla K80, with 12 GB of memory and Cuda edition of 10.1.

- The training parameters for Model-A was detailed below: a) loss function: binary cross-entropy, b) optimizer: ADAM, c) batch size: 16, d) total number of epochs: 200, e) early stopping patience: 100, and f) the class weight was calculated to balance the classes of the region of interest.
- The training parameters for Model-B was detailed below: a) loss function: categorical cross-entropy, b) optimizer: SGD, c) learning rate: 0.001 with decay of 0,000001 and momentum of 0.9, d) batch size: 64, e) number of epochs: 10, and f) early stopping patience: 20.

3. RESULT AND DISCUSSION

This section presents the results of Model-A and Model-B and a comparison with others proposed in SLR. The estimation of the Model-A and Model-B results in DSC and IoU metric consisted of the 10-time repetition of the cross-validation process with K-fold equal to 5. In other words, the cross-validation process consisted of the 5-time repetition dividing the 30 MRIs volume into five random groups, of which four groups (24 MRIs volume) were used for training and validation, and one group (6 MRIs volume) was used for testing. Also, to establish a comparison of Model-A and Model-B with other proposed are got the confusion matrix using the 16-30 MRI volume, where the number of processed voxels was 29237248.

3.1 Results of Model-A

This section presented the results achieved by Model-A. The results of training, validate and testing of Model-A were presented in table 1. Notice in this table, the testing results reached 96.79% and 93.79% in the DSC and IoU metrics, respectively. In addition, the results showed that there was not over-fitting.

Further, in Table 2 was shown the confusion matrix and the precision, recall and accuracy metrics. The confusion matrix shown a true positive rate of 95.81% for the

Table 1. Result of Model A

	DSC [%]	**IoU** [%]
training	96.59 ± 0.23	93.42 ± 0.43
valid	94.09 ± 0.07	88.86 ± 0.13
test	96.79 ± 0.15	93.79 ± 0.29

background, 0.55% for the peripheral zone, and 3.20% for the central gland. Notice that peripheral zone and central gland pixels occupied $\leq$ 4.19% compared to background pixels. Furthermore, the precision and recall results for the peripheral zone were 77.44% and 70.65%. These results were due peripheral zone occupied 0.78% of the total number of pixels.

Table 2. Confusion matrix of Model-A

		Actual labels (%)			
		Background	Peripheral zone	Central gland	Precision
Predicted labels (%)	Background	**28013085** *95.81%*	**22445** *0.08%*	**11147** *0.04%*	99.88% *0.12%*
	Peripheral zone	**7004** *0.02%*	**159555** *0.55%*	**39484** *0.14%*	77.44% *22.56%*
	Central gland	**6193** 0.02%	**43766** 0.15%	**934569** 3.20%	94.93% 5.07%
	Recall	99.95% *0.05%*	70.67% *29.33%*	94.86% *5.14%*	99.56% *0.44%*

3.2 Results of Model-B

The training, validate and testing results of Model-B using DSC and IoU metric were presented in Table 3. The Model-B obtained a testing results of 92.50% and 86.13% in the DSC and IoU metrics, respectively. Notice that the testing result of Model-B was lower than Model-A because the result of Model-B did not take adjacent pixels into account.

Table 3. Result of Model-B

	DSC [%]	**IoU** [%]
training	94.27 ± 1.33	89.27 ± 2.36
valid	76.48 ± 0.36	62.25 ± 0.47
test	92.50 ± 1.19	86.13 ± 2.02

Table 4 presented the confusion matrix, the precision, recall and accuracy metrics. Also, the confusion matrix showed a true positive rate of 95.93% for the background, 0.33% for the peripheral zone, and 2.27% for the central gland. The precision and recall of the central gland and the peripheral zone were less than the background class, due to the peripheral zone and central gland patches occupied $\leq$ 4.07% compared to background patches.

Table 4. Confusion matrix of Model-B

		Actual labels (%)			
		Background	Peripheral zone	Central gland	Precision
Predicted labels (%)	Background	**28047402** *95.93%*	**69069** *0.24%*	**199736** *0.68%*	99.05% *0.95%*
	Peripheral zone	**38330** *0.13%*	**97275** *0.33%*	**21284** *0.07%*	62.00% *38.00%*
	Central gland	**66231** 0.23%	**32994** 0.11%	**664927** 2.27%	87.02% 12.98%
	Recall	99.63% *0.37%*	48.80% *51.20%*	75.05% *24.95%*	98.54% *1.46%*

3.3 Comparison of Proposed Models

The comparison of Model-A and Model-B consisted in compare the testing results (see Table 1 and table 3), which were shown in Fig 4. The results of Model-A and Model-B did not show overlap. Also, the result of Model-A overcame Model-B in DSC metric with 4.29%, and in IoU metric with 7.66%. The results of Model-A were due to the encoder-decoder architecture of the U-net network, which took as input the region of interest to predict results based on the information in context. The input and output of the U-net network used the same sizes. Also, Model-A used the region of interest of $M \times N$ size, which reduced 52% of the original MRI size. In addition, Model-A implemented the two U-net networks in a cascade.

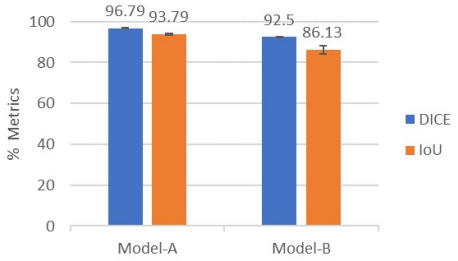

Fig. 4. Comparison of Model-A and Model-B

3.4 Comparison of Proposed Models with State of the Art

Comparison of models proposed with models that used the same dataset obtained the Model-P8. The results comparison of these model consisted in compare the precision, recall and accuracy results of proposed models (see tables 2 and 4). The results of this comparison were shown in the table 5. Notice that the precision, recall, and accuracy of Model-A achieved high results compared to Model-B and Model-P8. Moreover, Model-B outperformed Model-P8 in the precision and recall metrics of the whole prostate and central gland. However, Model-P8 outperformed Model-B in the precision and recall metrics of the peripheral zone.

Table 5. Comparison of Models A and B with the State of the Art

Metric	**Model-A**	**Model-B**	**Model-P8**
Precision (Predicted labels — Actual labels)			
Prostate — Prostate	98.8915	88.6480	77.7071
Central gland — Central gland	94.9256	87.0174	72.0175
Peripheral zone — Peripheral zone	77.4377	62.0015	76.3224
Recall (Actual labels — Predicted labels)			
Prostate — Prostate	97.2260	75.2310	62.0712
Central gland — Central gland	94.8608	75.0512	56.7926
Peripheral zone — Peripheral zone	70.6727	48.8045	63.0397
Accuracy	99.56	98.54	96.84

The prostate segmentation in their two main zones was evaluated qualitatively by visual comparison with manual ground truth and results of Model-A, Model-B, and Model-P8 (see Fig. 5). Although peripheral zone and central gland occupied a small area of pixels in the MRIs, the Model-A got a similar result to ground truth. Additionally, the Model-B occupied second place in the segmentation

of the two zones. Finally, the Model-P8 contained fewer true positives. Note that Model-B and Model-P8 detected more spurious objects of the central gland or peripheral zone because these models segment the two zones of the prostate by classifying each pixel of its 2D slices as either central gland, peripheral zone, or background.

4. CONCLUSION

In this work, we have presented two models for the automatic segmentation of the two main zones of the prostate: the central gland and the peripheral zone. The first model (Model-A) uses an encoder-decoder architecture, which takes as input the 2D MRIs, and returns the segmentation of the two zones of the prostate. On the other hand, the second model (Model-B) uses an encoder-classifier architecture. The Model-B takes as input a patch of a 2D MRI and returns the classification of that patch into three classes: background, central gland, or peripheral zone. The classification of all the patches of an MRI gives as a result of the segmentation of the two main zones of the prostate.

Model-A is based on the U-net network architecture. This model is composed of a cascade of two networks: global U-net and local U-net. The global U-net is used to segment the whole prostate. This network takes as input 2D MRIs with a region of interest of $M \times N$ pixels, where $M \leq 151$ and $N \leq 151$. This region of interest of $M \times N$ pixels contains an area of 48% of the original MRI. The local U-net is used to segment the central gland. The input of this local U-net is the region of interest of $M \times N$ pixels together with its segmentation predicted by the global U-net. The peripheral zone is obtained by subtracting the central gland from the whole prostate. The global and local U-nets share the same architecture.

Model-B is based on a pre-trained VGG16 network. This model takes as input patches of 15×15 pixels extracted from the region of interest of $M \times N$ pixels obtained from 2D MRIs using the process described above for Model-A. Each patch of 15×15 pixels was labeled with the value of its central pixel in the ground truth image. The three classes that were used to define the labels for these patches are central gland, peripheral zone, and background. To train this encoder-classifier model, we used a balanced dataset composed of 468689 pairs of labels and patches of 15×15 pixels: 161607 patches for the background, 152042 patches for the central gland, and 155040 patches for the peripheral zone. The training of Model-B consisted of a fine-tuning of the weights of its last three layers using this dataset of 468689 patches and labels.

The two models proposed in this work were tested using the MRI images from the public dataset NCI-ISBI 2013 Challenge. The Model-A, with an encoder-decoder architecture, achieved a performance of $DSC = 96.79\% \pm 0.15\%$ and $IoU = 93.79\% \pm 0.29\%$; whereas the Model-B, with an encoder-classifier architecture, achieved a performance of $DSC = 92.50\% \pm 1.19\%$ and $IoU = 86.13\% \pm 2.02\%$. The comparison between the results of these models indicates that Model-A performs better than Model-B, with 4.29% and 7.66% for the DSC and IoU metric, respectively. The superior performance of Model-A compared to Model-B occurs because Model-A exploits the 2D geometrical structure in both the input and ground truth images, whereas the Model-B exploits the 2D geometrical structure of the input images only.

Future work includes the development and evaluation of models that exploit the 3D structure of the MRIs for the segmentation of the main parts of the prostate.

ACKNOWLEDGEMENTS

Authors would like thank the computing facilities and advisory services provided by the Scientific Computing Laboratory of the Research Center on Mathematical Modeling (MODEMAT) of the Escuela Politécnica Nacional. Additionally, Silvia Vacacela thanks the academic and financial support provided by the Escuela Politécnica Nacional to conclude successfully her Master on Computer Science.

REFERENCES

Aldoj, N., Biavati, F., Michallek, F., Stober, S., and Dewey, M. (2020). Automatic prostate and prostate zones segmentation of magnetic resonance images using densenet-like u-net. *Scientific Reports*, 10. doi: 10.1038/s41598-020-71080-0.

Benalcázar, M., Brun, M., and Ballarin, V. (2015). Automatic design of window operators for the segmentation of the prostate gland in magnetic resonance images. In A. Braidot and A. Hadad (eds.), *VI Latin American Congress on Biomedical Engineering CLAIB 2014, Paramá, Argentina 29, 30 & 31 October 2014*, 417–420. Springer International Publishing, Cham.

Cao, P., Noworolski, S.M., Starobinets, O., Korn, N., Kramer, S.P., Westphalen, A.C., Leynes, A.P., Pedoia, V., and Larson, P. (2020). Development of conditional random field insert for unet-based zonal prostate segmentation on t2-weighted mri.

Gelder, A. and Huisman, H. (2018). Autoencoders for multi-label prostate mr segmentation.

GLOBOCAN (2020). Estimated age-standardized incidence and mortality rates (world) in 2020, worldwide, males, all ages. URL `https://gco.iarc.fr/`.

Heidenreich, A., Aus, G., Bolla, M., Joniau, S., Matveev, V.B., Schmid, H.P., and Zattoni, F. (2008). Eau guidelines on prostate cancer. *European urology*, 53(1), 68–80.

Jensen, C., Sørensen, K., Jørgensen, C., Nielsen, C., Høy, P., Langkilde, N., and Østergaard, L. (2019). Prostate zonal segmentation in 1.5t and 3t t2w mri using a convolutional neural network. *Journal of medical imaging (Bellingham, Wash.)*, 6, 8. doi:10.3233/XST-190524.

Khan, Z., Yahya, N., Alsaih, K., Azhar, S., and Meriaudeau, F. (2020). Evaluation of deep neural networks for semantic segmentation of prostate in t2w mri. *Sensors (Basel, Switzerland)*, 20. doi:10.3390/s20113183.

khan, Z., Yahya, N., Alsaih, K., and Meriaudeau, F. (2019). Zonal segmentation of prostate t2w-mri using atrous convolutional neural network. In *2019 IEEE Student Conference on Research and Development (SCOReD)*, 95–99. doi: 10.1109/SCORED.2019.8896248.

Liu, Y., Sung, K., Yang, G., Afshari Mirak, S., Hosseiny, M., Azadikhah, A., Zhong, X., Reiter, R., Lee, Y., and Raman, S. (2019). Automatic prostate zonal segmentation using fully convolutional network with feature pyramid attention. *IEEE Access*, 7, 1–1. doi: 10.1109/ACCESS.2019.2952534.

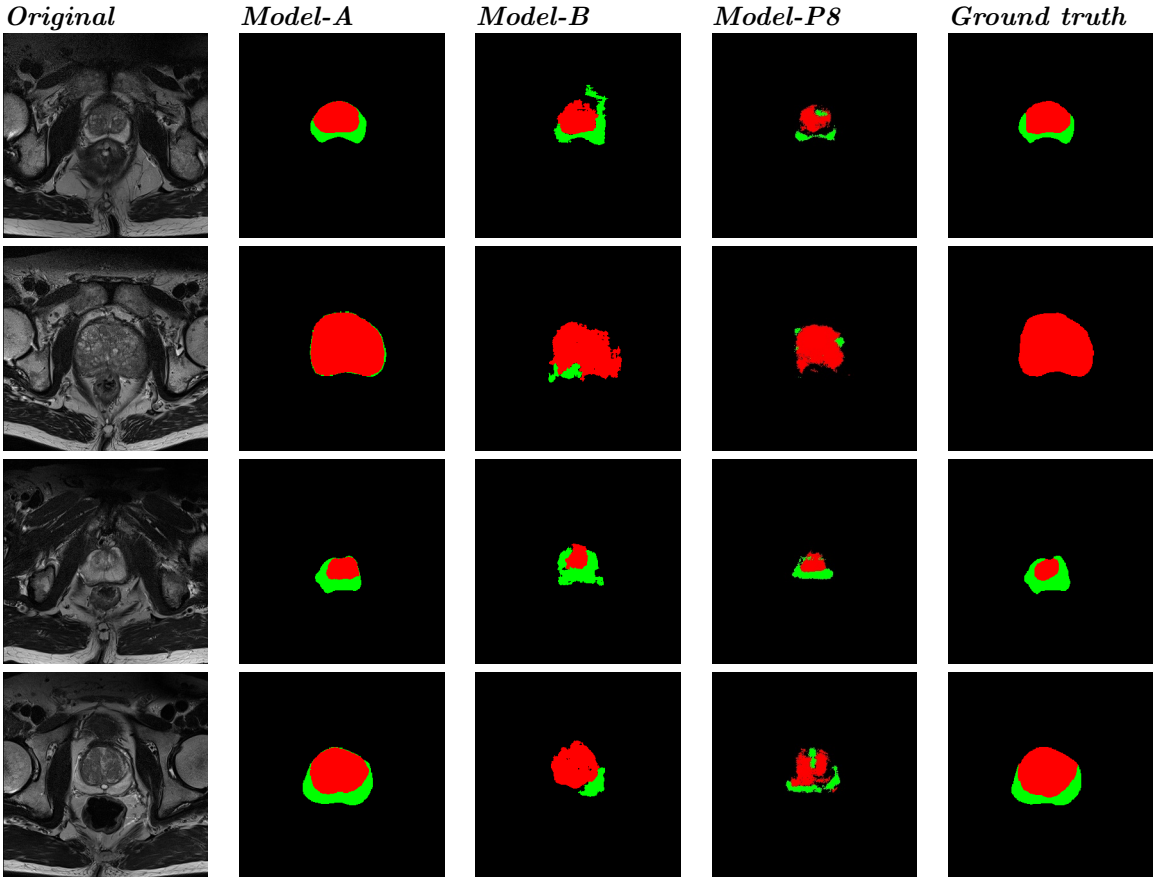

Fig. 5. Examples of original, results of Model-A, results of Model-B, results of other model proposed in SLR, and ground truth slices. Green = peripheral zone; Red = central gland

Lomas, D. and Ahmed, H. (2020). All change in the prostate cancer diagnostic pathway. *Nature Reviews Clinical Oncology*, 17, 1–10. doi:10.1038/s41571-020-0332-z.

Meyer, A., Rakr, M., Schindele, D., Blaschke, S., Schostak, M., Fedorov, A., and Hansen, C. (2019). Towards patient-individual pi-rads v2 sector map: Cnn for automatic segmentation of prostatic zones from t2-weighted mri. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 696–700. doi:10.1109/ISBI.2019.8759572.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. doi:https://arxiv.org/pdf/1505.04597.pdf.

Rundo, L., Han, C., Zhang, J., Hataya, R., Nagano, Y., Militello, C., Ferretti, C., Nobile, M.S., Tangherloni, A., Gilardi, M.C., Vitabile, S., Nakayama, H., and Mauri, G. (2020). *CNN-Based Prostate Zonal Segmentation on T2-Weighted MR Images: A Cross-Dataset Study*, 269–280. Springer, Singapore. doi:10.1007/978-981-13-8950-4_25.

Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. doi:https://arxiv.org/pdf/1409.1556.pdf.

Umapathy, L., Unger, W., Shareef, F., Arif, H., Martin, D., Altbach, M., and Bilgin, A. (2020). A cascaded residual unet for fully automated segmentation of prostate and peripheral zone in t2-weighted 3d fast spin echo images.

Villers, A., Marliere, F., Ouzzane, A., Puech, P., and Lemaître, L. (2012). Mri in addition to or as a substitute for prostate biopsy: The clinician's point of view. *Diagnostic and interventional imaging*, 93(4), 262–267. doi:10.1016/j.diii.2012.01.018.

Yan, K., Wang, X., Kim, J., Khadra, M., Fulham, M., and Feng, D.D.F. (2019). A propagation-dnn: Deep combination learning of multi-level features for mr prostate segmentation. *Computer Methods and Programs in Biomedicine*, 170, 11–21. doi:10.1016/j.cmpb.2018.12.031.

Zabihollahy, F., Scheida, N., Krishna, S., and Ukwatta, E. (2019). Automated segmentation of prostate zonal anatomy on t2-weighted (t2w) and apparent diffusion coefficient (adc) map mr images using cascaded u-nets. *Medical Physics*, 46. doi:10.1002/mp.13550.