

# CNN-based Prostate Zonal Segmentation on T2-weighted MR Images: A Cross-dataset Study

Leonardo Rundo<sup>1,2</sup>(✉), Changhee Han<sup>3</sup>, Jin Zhang<sup>3</sup>, Ryuichiro Hataya<sup>3</sup>,  
Yudai Nagano<sup>3</sup>, Carmelo Militello<sup>2</sup>, Claudio Ferretti<sup>1</sup>, Marco S. Nobile<sup>1</sup>,  
Andrea Tangherloni<sup>1</sup>, Maria Carla Gilardi<sup>2</sup>, Salvatore Vitabile<sup>4</sup>,  
Hideki Nakayama<sup>3</sup>, and Giancarlo Mauri<sup>1</sup>

<sup>1</sup> Department of Informatics, Systems and Communication,  
University of Milano-Bicocca, Milan, Italy

`leonardo.rundo@disco.unimib.it`,

<sup>2</sup> Institute of Molecular Bioimaging and Physiology (IBFM),  
Italian National Research Council (CNR), Cefalù (PA), Italy

<sup>3</sup> Graduate School of Information Science and Technology,  
The University of Tokyo, Tokyo, Japan

<sup>4</sup> Department of Biopathology and Medical Biotechnologies,  
University of Palermo, Palermo, Italy

**Abstract.** Prostate cancer is the most common cancer among US men. However, prostate imaging is still challenging despite the advances in multi-parametric Magnetic Resonance Imaging (MRI), which provides both morphologic and functional information pertaining to the pathological regions. Along with whole prostate gland segmentation, distinguishing between the Central Gland (CG) and Peripheral Zone (PZ) can guide towards differential diagnosis, since the frequency and severity of tumors differ in these regions; however, their boundary is often weak and fuzzy. This work presents a preliminary study on Deep Learning to automatically delineate the CG and PZ, aiming at evaluating the generalization ability of Convolutional Neural Networks (CNNs) on two multi-centric MRI prostate datasets. Especially, we compared three CNN-based architectures: SegNet, U-Net, and pix2pix. In such a context, the segmentation performances achieved with/without pre-training were compared in 4-fold cross-validation. In general, U-Net outperforms the other methods, especially when training and testing are performed on multiple datasets.

**Keywords:** Prostate zonal segmentation · Prostate cancer · Anatomical MRI · Deep convolutional neural networks · Cross-dataset generalization.

## 1 Introduction

Prostate cancer (PCa) is expected to be the most common cancer among US men during 2018 [1]. Several imaging modalities can aid PCa diagnosis—such as Transrectal Ultrasound (TRUS), Computed Tomography (CT), and Magnetic Resonance Imaging (MRI) [2]—according to the clinical context. Conventional

structural T1-weighted (T1w) and T2-weighted (T2w) MRI sequences can play an important role along with functional MRI, such as Dynamic Contrast Enhanced MRI (DCE-MRI), Diffusion Weighted Imaging (DWI), and Magnetic Resonance Spectroscopic Imaging (MRSI) [3]. Therefore, MRI conveys more information for PCa diagnosis than CT, revealing the internal prostatic anatomy, prostatic margins, and the extent of prostatic tumors [4].

The manual delineation of both prostate Whole Gland (WG) and PCa on MR images is a time-consuming and operator-dependent task, which relies on experienced physicians [5]. Besides WG segmentation, distinguishing between the Central Gland (CG) and Peripheral Zone (PZ) of the prostate can guide towards differential diagnosis, since the frequency and severity of tumors differ in these regions [6, 7]; the PZ harbors 70 – 80% of PCa and is a target for prostate biopsy [8]. Regarding this anatomic division of the prostate, the zonal compartment scheme proposed by McNeal is widely accepted [9]. In this context, T2w MRI is the *de facto* standard in the clinical routine of prostate imaging thanks to its high resolution, which allows for differentiating the hyper-intense PZ and hypo-intense CG in young male subjects [10]. However, the conventional clinical protocol for PCa based on Prostate-Specific Antigen (PSA) and systematic biopsy does not generally obtain reliable diagnostic outcomes; thus, the PZ volume ratio (i.e., the PZ volume divided by the WG volume) was recently integrated for PCa diagnostic refinement [11]; the CG volume ratio can be also useful for monitoring prostate hyperplasia [12]. Furthermore, for robust clinical applications, generalization—among different prostate MRI datasets from multiple institutions—is essential.

So, how can we extract the CG and PZ from the WG on different MRI datasets? In this work, we automatically segment the CG and PZ using Deep Learning to evaluate the generalization ability of Convolutional Neural Networks (CNNs) on two different MRI prostate datasets. However, this is challenging since multi-centric datasets are generally characterized by different contrast, visual consistencies, and image characteristics. Therefore, prostate zones on T2w MR images were manually annotated for supervised learning, and then automatically segmented using a mixed scheme by (*i*) training on either each individual dataset or both datasets and (*ii*) testing on both datasets, using CNN-based architectures: SegNet [13], U-Net [14], and pix2pix [15]. In such a context, we compared the segmentation performances achieved with/without pre-training [16].

The manuscript is structured as follows: Sect. 2 outlines the state-of-the-art about MRI prostate zonal segmentation methods; Sect. 3 describes the MRI datasets as well as the proposed CNN-based segmentation approach; Sect. 4 shows our experimental results; finally, some conclusive remarks and possible future developments are given in Sect. 5.

## 2 Background

In prostate MR image analysis, WG segmentation is essential, especially on T2w MR images [17] or both T2w and the corresponding T1w images [5, 18].

Towards it, literature works used atlas-based methods [19], deformable models, or statistical priors [20]. More recently, Deep Learning [21] has been successfully applied to this domain, combining deep feature learning with shape models [22] or using CNN-based segmentation approaches [23].

Among the studies on prostate segmentation, the most representative works are outlined hereafter. The authors of [24] used active appearance models combined with multiple level sets for simultaneously segmenting prostatic zones. Qiu *et al.* [25] proposed a zonal segmentation approach introducing a continuous max-flow model based on a convex-relaxed optimization problem with region consistency constraints. Unlike these methods that analyzed T2w images alone, Makni *et al.* [26] exploited the evidential C-means algorithm to partition the voxels into their respective zones by integrating the information conveyed by T2w, DWI, and CE T1w MRI sequences.

However, these methods do not evaluate the generalization ability on different MRI datasets from multiple institutions, making their clinical applicability difficult [27]; thus, we verify the cross-dataset generalization using three CNN-based architectures with/without pre-training. Moreover, differently from a recent CNN-based work on DWI data [28], to the best of our knowledge, this is the first CNN-based prostate zonal segmentation approach on T2w MRI alone.

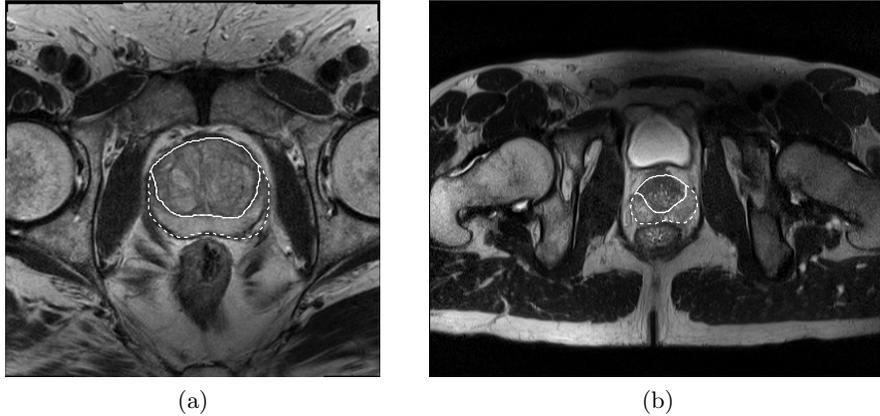
### 3 Materials and Methods

For clinical applications with better generalization ability, we evaluate prostate zonal segmentation performances of three CNN-based architectures: SegNet, U-Net, and pix2pix. We also compare the results in 4-fold cross-validation by (*i*) training on either each individual dataset or both datasets and (*ii*) testing on both datasets, with/without pre-training on a relatively large prostate dataset.

#### 3.1 MRI Datasets

We segment the CG and PZ from the WG using two completely different multi-parametric prostate MRI datasets, namely:

- #1 dataset containing 21 patients/193 MR slices with prostate, acquired with a whole body Philips Achieva 3T MRI scanner using a phased-array pelvic coil at the Cannizzaro Hospital (Catania, Italy) [5]. The MRI parameters: matrix size =  $288 \times 288$  pixels; slice thickness = 3.0 mm; inter-slice spacing = 4 mm; pixel spacing = 0.625 mm; number of slices = 18;
- #2 Initiative for Collaborative Computer Vision Benchmarking (I2CVB) dataset (19 patients/503 MR slices with prostate), acquired with a whole body Siemens TIM 3T MRI scanner using a body coil at the Hospital Center Regional University of Dijon-Bourgogne (France) [3]. The MRI parameters: matrix size  $\in \{308 \times 384, 336 \times 448, 360 \times 448, 368 \times 448\}$  pixels; slice thickness = 1.25 mm; inter-slice spacing = 1.0 mm; pixel spacing  $\in \{0.676, 0.721, 0.881, 0.789\}$  mm; number of slices = 64.



**Fig. 1.** Example input prostate T2w MR axial slices in their original image ratio: (a) dataset #1; (b) dataset #2. The CG and PZ are highlighted with solid and dashed white lines, respectively.

To make the proposed approach clinically feasible [10], we analyzed only T2w images—the most commonly used sequence for prostate zonal segmentation—among available sequences. Fig. 1 shows two example T2w MR images of the analyzed two datasets. We conducted the following three experiments using a 4-fold cross-validation scheme to confirm the generalization effect under different training/testing conditions in our multi-centric study:

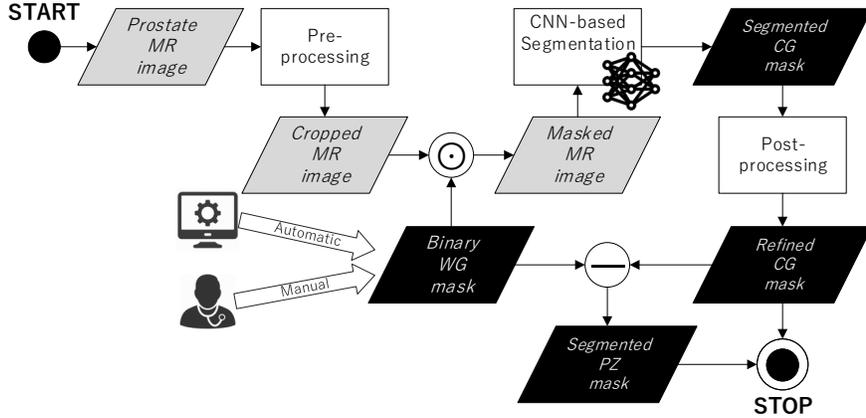
- Individual dataset #1: training on dataset #1 alone, and testing on the whole dataset #2 and the rest of dataset #1 separately for each round;
- Individual dataset #2: training on dataset #2 alone, and testing on the whole dataset #1 and the rest of dataset #2 separately for each round;
- Mixed dataset: training on both datasets #1 and #2, and testing on the rest of datasets #1 and #2 separately for each round.

For 4-fold cross-validation, we partitioned the datasets #1 and #2 using patient indices  $\{[1, \dots, 5], [6, \dots, 10], [11, \dots, 15], [16, \dots, 21]\}$  and  $\{[1, \dots, 5], [6, \dots, 10], [11, \dots, 15], [16, \dots, 19]\}$ , respectively. Finally, the results from the different cross-validation rounds were averaged.

### 3.2 CNN-based Prostate Zonal Segmentation

This work adopts a selective two-step delineation approach to focus on pathological regions in the CG and PZ denoted with  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{PZ}$ , respectively. Relying on [4, 25], the PZ was obtained by subtracting the CG from the WG ( $\mathcal{R}_{WG}$ ) meeting the constraints:  $\mathcal{R}_{WG} = \mathcal{R}_{CG} \cup \mathcal{R}_{PZ}$  and  $\mathcal{R}_{CG} \cap \mathcal{R}_{PZ} = \emptyset$ . The overall prostate zonal segmentation method is outlined in Fig. 2.

Starting from the whole prostate MR image, a pre-processing phase, comprising image cropping and resizing to deal with the different characteristics of



**Fig. 2.** Work-flow of our CNN-based prostate zonal segmentation approach. The gray and black data blocks denote gray-scale images and binary masks, respectively.

the two datasets, is performed (see Sect. 3.1). Afterwards, the resulting image is masked with the binary WG and fed onto the investigated CNN-based models for CG segmentation, which is then refined. Finally, the PZ delineation is obtained by subtracting the CG from the WG according to [25].

**Pre-processing** To fit the image resolution of the dataset #1, we center-cropped the images of the dataset #2 and resized them to  $288 \times 288$  pixels. Furthermore, the images of these datasets were masked using the corresponding prostate binary masks to omit the background and only focus on extracting the CG and PZ from the WG. This operation can be performed either by an automated method [5, 18] or previously provided manual WG segmentation [3]. For better training, we randomly cropped the input images from  $288 \times 288$  to  $256 \times 256$  pixels and horizontally flipped them.

**Investigated CNN-based Architectures** The following architectures were chosen in our comparative analysis since they cover several aspects regarding the CNN-based segmentation: SegNet [13] addresses semantic segmentation (i.e., assigning each pixel of the scene to an object class), U-Net [14] was successfully applied in biomedical image segmentation, while pix2pix [15] exploits an adversarial generative model to perform image-to-image translations.

During the training of all architectures, the  $\mathcal{L}_{DSC}$  loss function (i.e., a continuous version of the Dice Similarity Coefficient) was used [23]:

$$\mathcal{L}_{DSC} = -\frac{2 \sum_{i=1}^N s_i \cdot r_i}{\sum_{i=1}^N s_i + \sum_{i=1}^N r_i}, \quad (1)$$

where  $s_i$  and  $r_i$  represent the continuous values of the prediction map (i.e., the result of the final layer of the CNN) and the ground truth at the  $i$ -th pixel ( $N$  is the total number of pixels to be classified), respectively.

*SegNet* is a CNN architecture for semantic pixel-wise segmentation [13]. More specifically, it was designed for semantic segmentation of road and traffic scenes, wherein classes represent macro-objects, aiming at smooth segmentation results by preserving boundary information. This non-fully connected architecture, which allows for parameter-efficient implementations suitable for embedded systems, consists of an encoder-decoder network followed by a pixel-wise classification layer. Since our classification task involves only one class, the soft-max operation and Rectified Linear Unit (ReLU) activation function at the final layer were removed for stable training.

We implemented SegNet using PyTorch. During the training phase, we used the Stochastic Gradient Descent (SGD) [29] with a learning rate of 0.01, momentum of 0.9, weight decay of  $5 \times 10^{-4}$ , and batch size of 8. It was trained for 50 epochs and the learning rate was multiplied by 0.2 at the 20-th and 40-th epochs.

*U-Net* is a fully CNN capable of stable training with a reduced number of samples [14], combining pooling operators with up-sampling operations. The general architecture is an encoder-decoder with skip connections between mirrored layers in the encoder-decoder stacks. By so doing, high resolution features from the contracting path are combined with the up-sampled output for better localization. We utilized four scaling operations. U-Net achieved outstanding performance in biomedical benchmark problems [30] and has been also serving as an inspiration for novel Deep Learning models for image segmentation.

U-Net was implemented using Keras on top of TensorFlow. We used SGD with a learning rate of 0.01, momentum of 0.9, weight decay of  $5 \times 10^{-4}$ , and batch size of 4. Training was executed for 50 epochs, multiplying the learning rate by 0.2 at the 20-th, and 40-th epochs.

*pix2pix* is an image-to-image translation method coupled with conditional adversarial networks [15]. As a generator, U-Net is used to translate the original image into the segmented one [14], preserving the highest level of abstraction. The generator and discriminator include 8 and 5 scaling operations, respectively.

We implemented pix2pix on PyTorch. Adam [31] was used as an optimizer with a learning rate of 0.0002 and 0.01 for the discriminator and generator, respectively. The learning rate for generator was multiplied by 0.1 every 20 epochs. It was trained for 50 epochs with a batch size of 12.

**Post-processing** Two simple morphological steps were applied on the obtained CG binary masks to smooth boundaries and avoid disconnected regions:

- a hole filling algorithm on the segmented  $\mathcal{R}_{CG}$  to remove possible holes in the predicted map;

- a small area removal operation to delete connected components with area less than  $\lfloor |\mathcal{R}_{WG}|/8 \rfloor$  pixels, where  $|\mathcal{R}_{WG}|$  denotes the number of the pixels contained in the WG segmentation. This criterion effectively adapts according to the different dimensions of the  $\mathcal{R}_{WG}$ .

**Evaluation** The accuracy of the achieved segmentation results  $\mathcal{S}$  was quantitatively evaluated with respect to the real measurement (i.e., the gold standard  $\mathcal{G}$  obtained manually by experienced radiologists) using the *DSC*:

$$DSC = \frac{2|\mathcal{S} \cap \mathcal{G}|}{|\mathcal{S}| + |\mathcal{G}|} \times 100 (\%). \quad (2)$$

### 3.3 Influence of Pre-training

In medical imaging, due to the lack of training data, ensuring CNN’s proper training convergence is difficult from scratch. Therefore, pre-training models on a different application and then fine-tuning is common [16].

To evaluate cross-dataset generalization abilities *via* pre-training, we compared the performances of the three CNN-based architectures with/without pre-training on a similar application. We used a relatively large dataset of 50 manually segmented examples from the Prostate MR Image Segmentation 2012 (PROMISE12) challenge [32]. Since this competition focuses only on WG segmentation without providing prostate zonal labeling, we pre-trained the architectures on WG segmentation. To adjust this dataset to our experimental setup, the images of this dataset were resized from  $512 \times 512$  to  $288 \times 288$  pixels and randomly cropped to  $256 \times 256$  pixels; because our task only focuses on slices with prostate, we also omitted initial/final slices without prostate, so the number of slices for each sample was fixed to 25.

## 4 Results

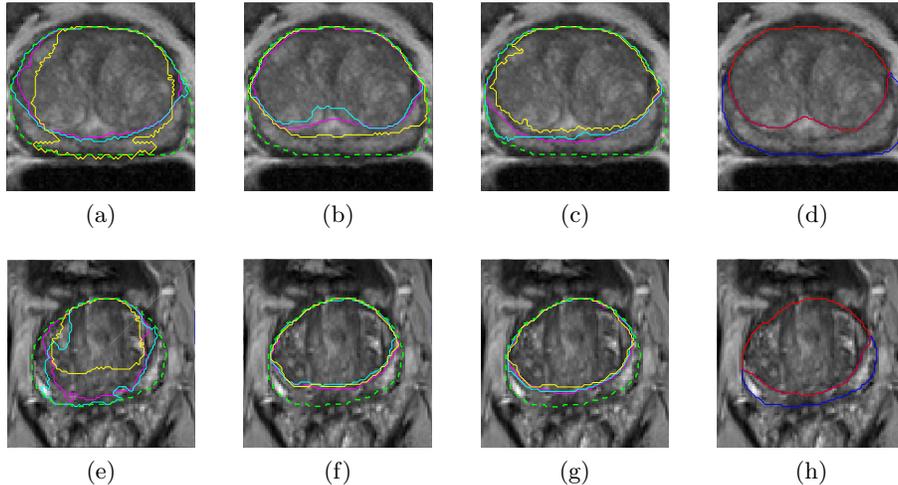
This section explains how the three CNN-based architectures segmented the prostate zones, evaluating their cross-dataset generalization ability.

Table 1 shows the 4-fold cross-validation results obtained in the different experimental conditions. When training and testing are both performed on the dataset #1, U-Net outperforms the other architectures on both CG and PZ segmentation; however, it experiences problems with testing on the dataset #2 due to the limited number of training images in the dataset #1. In such a case, pix2pix generalizes better thanks to its internal generative model. When trained on the dataset #2 alone, U-Net yields the most accurate results both in intra- and cross-dataset testing. This probably derives from the dataset #2’s relatively larger training data as well as U-Net’s good generalization ability when sufficient data are available. Moreover, SegNet reveals rather unstable results, especially when trained on a limited amount of data. Finally, when trained on the mixed

**Table 1.** Prostate zonal segmentation results of the three CNN-based architectures in 4-fold cross-validation assessed by the *DSC* (presented as the average and standard deviation). The experimental results are calculated on the different setups of (i) training on either each individual dataset or both datasets and (ii) testing on both datasets. Numbers in bold indicate the highest *DSC* values for each prostate region (i.e., CG and PZ) among all architectures with/without pre-training (PT).

	Architecture	Zone	Testing on Dataset #1		Testing on Dataset #2		
			Average	Std. Dev.	Average	Std. Dev.	
Training on Dataset #1	SegNet (w/o PT)	CG	80.20	3.28	74.48	5.82	
		PZ	80.66	11.51	59.57	12.68	
	SegNet (w/ PT)	CG	83.38	3.22	72.75	2.80	
		PZ	87.39	3.90	66.20	5.64	
	U-Net (w/o PT)	CG	84.33	2.37	74.18	3.77	
		PZ	88.98	2.98	66.63	1.93	
	U-Net (w/ PT)	CG	<b>86.88</b>	1.60	70.11	5.31	
		PZ	<b>90.38</b>	3.38	58.89	7.06	
	pix2pix (w/o PT)	CG	CG	82.35	2.09	<b>76.61</b>	2.17
			PZ	87.09	2.72	73.20	2.62
		CG	CG	80.38	2.81	76.19	5.77
			PZ	83.53	5.65	<b>73.73</b>	2.40
Training on Dataset #2	SegNet (w/o PT)	CG	76.04	2.05	87.07	2.41	
		PZ	77.25	3.09	82.45	1.77	
	SegNet (w/ PT)	CG	77.99	2.15	87.75	2.83	
		PZ	76.51	2.70	82.26	2.09	
	U-Net (w/o PT)	CG	78.88	0.88	88.21	2.10	
		PZ	74.52	1.85	<b>83.03</b>	2.46	
	U-Net (w/ PT)	CG	<b>79.82</b>	1.11	<b>88.66</b>	2.28	
		PZ	<b>74.56</b>	5.12	82.48	2.47	
	pix2pix (w/o PT)	CG	CG	77.90	0.73	86.95	2.93
			PZ	66.09	3.07	81.33	0.90
		CG	CG	77.21	1.02	85.94	4.31
			PZ	67.39	5.04	80.07	0.84
Training on Mixed Dataset	SegNet (w/o PT)	CG	84.28	3.12	87.92	2.80	
		PZ	87.74	1.66	82.21	0.79	
	SegNet (w/ PT)	CG	86.08	1.92	87.78	2.75	
		PZ	89.53	3.28	82.39	1.50	
	U-Net (w/o PT)	CG	<b>86.34</b>	2.10	<b>88.12</b>	2.34	
		PZ	90.74	2.40	<b>83.04</b>	2.30	
	U-Net (w/ PT)	CG	85.82	1.98	87.42	1.89	
		PZ	<b>91.44</b>	2.15	82.17	2.11	
	pix2pix (w/o PT)	CG	CG	83.07	3.39	86.39	3.16
			PZ	83.53	2.36	80.40	1.80
		CG	CG	82.08	4.37	85.96	5.40
			PZ	83.04	4.20	80.60	1.49

dataset, all three architectures—especially U-Net—achieve good results on both datasets without losing accuracy compared to training on the same dataset alone.



**Fig. 3.** Examples of prostate zonal segmentation in pre-training/fine-tuning. The first row concerns testing on dataset #1, trained on: (a) dataset #1; (b) dataset #2; (c) mixed dataset. The second row concerns testing on dataset #2, trained on: (e) dataset #1; (f) dataset #2; (g) mixed dataset. The  $\mathcal{R}_{CG}$  segmentation results are represented with magenta, cyan, and yellow solid contours for SegNet, U-Net, and pix2pix, respectively. The dashed green line denotes the  $\mathcal{R}_{WG}$  boundary. The last column (sub-figures (d) and (h)) shows the gold standard for  $\mathcal{R}_{CG}$  and  $\mathcal{R}_{PZ}$  with red and blue lines, respectively. The images are zoomed with a  $4\times$  factor.

Therefore, using mixed MRI datasets during training can considerably improve the performance in cross-dataset generalization towards other clinical applications. Comparing the CG and PZ segmentation, when tested on the dataset #1, the results on the PZ are generally more accurate, except when trained on the dataset #2 alone; however, for the dataset #2, segmentations on the CG are generally more accurate.

Fine-tuning after pre-training sometimes leads to slightly better results than training from scratch, when trained only on a single dataset. However, its influence is generally negligible or rather negative, when trained on the mixed dataset. This modest impact is probably due to the ineffective data size for pre-training.

For a visual assessment, two examples (one for each dataset) are shown in Fig. 3. Relying on the gold standards in Figs. 3d and 3h, it can be seen that U-Net generally achieves more accurate results compared with SegNet and pix2pix. This finding confirms the trend revealed by the  $DSC$  values in Table 1.

## 5 Discussion and Conclusions

Our preliminary results show that CNN-based architectures can segment prostate zones on two different MRI datasets to some extent, leading to valuable clinical insights; CNNs suffer when training and testing are performed on different

MRI datasets acquired by different devices and protocols, but this can be mitigated by training the CNNs on multiple datasets, even without pre-training. Generally, considering different experimental training and testing conditions, U-Net outperforms SegNet and pix2pix thanks to its good generalization ability. Furthermore, this study suggests that significant performance improvement *via* fine-tuning may require a remarkably large dataset for pre-training.

As future developments, we plan to improve the results by refining the predicted binary masks for better smoothness and continuity, avoiding disconnected segments; furthermore, we should enhance the output delineations considering the three-dimensional spatial information among slices. Furthermore, relying on the encouraging cross-dataset capability of U-Net, it is worth to devise and test new solutions aiming at improving the performance of the standard U-Net architecture [30]. Finally, for better cross-dataset generalization, additional prostate zonal datasets and domain adaptation using transfer learning with Generative Adversarial Networks (GANs) [33,34] and Variational Auto-Encoders (VAEs) [35] could be useful.

## Acknowledgment

This work was partially supported by the Graduate Program for Social ICT Global Creative Leaders of The University of Tokyo by JSPS.

## References

1. Siegel, R.L., Miller, K.D., Jemal, A.: Cancer statistics, 2018. *CA Cancer J. Clin.* **68**(1) (2018) 7–30
2. Rundo, L., Tangherloni, A., Nobile, M.S., Militello, C., Besozzi, D., Mauri, G., Cazzaniga, P.: MedGA: a novel evolutionary method for image enhancement in medical imaging systems. *Expert Syst. Appl.* **119** (2019) 387–399
3. Lemaître, G., Martí, R., Freixenet, J., Vilanova, J.C., Walker, P.M., Meriaudeau, F.: Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: A review. *Comput. Biol. Med.* **60** (2015) 8–31
4. Villeirs, G.M., De Meerleer, G.O.: Magnetic resonance imaging (MRI) anatomy of the prostate and application of MRI in radiotherapy planning. *Eur. J. Radiol.* **63**(3) (2007) 361–368
5. Rundo, L., Militello, C., Russo, G., Garufi, A., Vitabile, S., Gilardi, M.C., Mauri, G.: Automated prostate gland segmentation based on an unsupervised fuzzy c-means clustering technique using multispectral T1w and T2w MR imaging. *Information* **8**(2) (2017) 49
6. Choi, Y.J., Kim, J.K., Kim, N., Kim, K.W., Choi, E.K., Cho, K.S.: Functional MR imaging of prostate cancer. *Radiographics* **27**(1) (2007) 63–75
7. Niaf, E., Rouvière, O., Mège-Lechevallier, F., Bratan, F., Lartizien, C.: Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric MRI. *Phys. Med. Biol.* **57**(12) (2012) 3833
8. Haffner, J., Potiron, E., Bouyé, S., Puech, P., Leroy, X., Lemaitre, L., Villers, A.: Peripheral zone prostate cancers: location and intraprostatic patterns of spread at histopathology. *Prostate* **69**(3) (2009) 276–282

9. Selman, S.H.: The McNeal prostate: A review. *Urology* **78**(6) (2011) 1224–1228
10. Hoeks, C.M., Barentsz, J.O., Hambroek, T., Yakar, D., Somford, D.M., Heijmink, S.W., et al.: Prostate cancer: multiparametric MR imaging for detection, localization, and staging. *Radiology* **261**(1) (2011) 46–66
11. Chang, Y., Chen, R., Yang, Q., Gao, X., Xu, C., Lu, J., Sun, Y.: Peripheral zone volume ratio (PZ-ratio) is relevant with biopsy results and can increase the accuracy of current diagnostic modality. *Oncotarget* **8**(21) (2017) 34836
12. Kirby, R., Gilling, R.: *Fast Facts: Benign Prostatic Hyperplasia*. 7 edn. Health Press Limited, Abingdon, UK (2011)
13. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12) (2017) 2481–2495
14. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Volume 9351 of LNCS., Springer (2015) 234–241
15. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. arXiv preprint arXiv:1611.07004 (2016)
16. Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J.: Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans. Med. Imaging* **35**(5) (2016) 1299–1312
17. Ghose, S., Oliver, A., Martí, R., Lladó, X., Vilanova, J.C., Freixenet, J., et al.: A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images. *Comput. Meth. Prog. Biomed.* **108**(1) (2012) 262–287
18. Rundo, L., Militello, C., Russo, G., D’Urso, D., Valastro, L.M., Garufi, A., et al.: Fully automatic multispectral MR image segmentation of prostate gland based on the fuzzy c-means clustering algorithm. In: *Multidisciplinary Approaches to Neural Computing*. Volume 69 of Smart Innovation, Systems and Technologies. Springer (2018) 23–37
19. Klein, S., Van Der Heide, U.A., Lips, I.M., Van Vulpen, M., Staring, M., Pluim, J.P.: Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med. Phys.* **35**(4) (2008) 1407–1417
20. Martin, S., Troccaz, J., Daanen, V.: Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model. *Med. Phys.* **37**(4) (2010) 1579–1590
21. Bevilacqua, V., Brunetti, A., Guerriero, A., Trotta, G.F., Telegrafo, M., Moschetta, M.: A performance comparison between shallow and deeper neural networks supervised classification of tomosynthesis breast lesions images. *Cogn. Syst. Res.* **53** (2019) 3–19
22. Guo, Y., Gao, Y., Shen, D.: Deformable MR prostate segmentation via deep feature learning and sparse patch matching. *IEEE Trans. Med. Imaging* **35**(4) (2016) 1077–1089
23. Milletari, F., Navab, N., Ahmadi, S.A.: V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: *Proceedings of the 4th International Conference on 3D Vision (3DV)*, IEEE (2016) 565–571
24. Toth, R., Ribault, J., Gentile, J., Sperling, D., Madabhushi, A.: Simultaneous segmentation of prostatic zones using active appearance models with multiple coupled levelsets. *Comput. Vis. Image Underst.* **117**(9) (2013) 1051–1060
25. Qiu, W., Yuan, J., Ukwatta, E., Sun, Y., Rajchl, M., Fenster, A.: Dual optimization based prostate zonal segmentation in 3D MR images. *Med. Image Anal.* **18**(4) (2014) 660–673

26. Makni, N., Iancu, A., Colot, O., Puech, P., Mordon, S., Betrouni, N.: Zonal segmentation of prostate using multispectral magnetic resonance images. *Med. Phys.* **38**(11) (2011) 6093–6105
27. AlBadawy, E.A., Saha, A., Mazurowski, M.A.: Deep learning for segmentation of brain tumors: Impact of cross-institutional training and testing. *Med. Phys.* (2018)
28. Clark, T., Zhang, J., Baig, S., Wong, A., Haider, M.A., Khalvati, F.: Fully automated segmentation of prostate whole gland and transition zone in diffusion-weighted MRI using convolutional neural networks. *J. Med. Imaging* **4**(4) (2017) 041307
29. Bottou, L.: Large-scale machine learning with stochastic gradient descent. In: *Proceedings of COMPSTAT'2010*. Springer (2010) 177–186
30. Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., Böhm, A., Deubner, J., Jäckel, Z., Seiwald, K., et al.: U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**(1) (2019) 67
31. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
32. Litjens, G., Toth, R., van de Ven, W., Hoeks, C., Kerkstra, S., van Ginneken, B., et al.: Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. *Med. Image Anal.* **18**(2) (2014) 359–373
33. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al.: Generative adversarial nets. In: *Proceedings of Advances in Neural Information Processing Systems (NIPS)*. (2014) 2672–2680
34. Han, C., Hayashi, H., Rundo, L., Araki, R., Shimoda, W., Muramatsu, S., et al.: GAN-based synthetic brain MR image generation. In: *Proc. International Symposium on Biomedical Imaging (ISBI), IEEE* (2018) 734–738
35. Kingma, D., Welling, M.: Auto-encoding variational Bayes. In: *Proc. International Conference on Learning Representations (ICLR)*. (2014) *arXiv preprint arXiv:1312.6114*.