A Deep Learning-based cropping technique to improve segmentation of prostate’s peripheral zone

Dimitris Zaridis†, Eugenia Mylona†, Nikolaos Tachos†, Kostas Marias†, Manolis Tsiknakis‡, Dimitios I. Fotiadis†, Fellow IEEE

† Dept. of Biomedical Research, FORTH-IMBB, Ioannina, Greece
‡ Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering, University of Ioannina, Ioannina, Greece
§ Department of Electrical and Computer Engineering, Hellenic Mediterranean University, Heraklion, Crete, Greece

Abstract—Automatic segmentation of the prostate peripheral zone on Magnetic Resonance Images (MRI) is a necessary but challenging step for accurate prostate cancer diagnosis. Deep learning (DL) based methods, such as U-Net, have recently been developed to segment the prostate and its sub-regions. Nevertheless, the presence of class imbalance in the image labels, where the background pixels dominate over the region to be segmented, may severely hamper the segmentation performance. In the present work, we propose a DL-based preprocessing pipeline for segmenting the peripheral zone of the prostate by cropping unnecessary information without making a priori assumptions regarding the location of the region of interest. The effect of DL-cropping for improving the segmentation performance was compared to the standard center-cropping using three state-of-the-art DL networks, namely U-net, Bridged U-net and Dense U-net. The proposed method achieved an improvement of 24%, 12% and 15% for the U-net, Bridged U-net and Dense U-net, respectively, in terms of Dice score.

Index Terms—Image segmentation, deep-learning, image cropping, U-net, prostate

I. INTRODUCTION

Prostate segmentation consists a necessary step of medical image analysis. In prostate cancer applications, T2-weighted (T2w) magnetic resonance imaging (MRI) is the state-of-the-art imaging technique as it provides superior resolution and contrast in soft tissues [1].

The prostate gland can be divided to different zones, namely the peripheral zone (PZ), the transitional zone and the central zone with tumor characteristics having substantial differences depending on the zone they are identified [2]. Given that the majority of tumoral lesions are located in the PZ (70–80%), it is of paramount importance to accurate segment this region. Automatic segmentation of the prostatic gland and prostate zones from MR images may assist several diagnostic and therapeutic applications. Reliable segmentation of prostate PZ, in particular, can enable accurate tumor localization while alleviating the burden of manual annotation in clinical routine practice which is not only time-consuming but also highly variable based on reader’s expertise. Nevertheless, the wide range of inter-individual shape variation and the heterogeneous and inconsistent pixel representation surrounding the PZ boundary, make its automatic segmentation a challenging task [3].

Recent advances of machine learning (ML) techniques offer the potential to significantly improve segmentation accuracy and consistency for medical applications. To date, several deep learning (DL) approaches based on convolutional neural networks (CNN) have been proposed to segment automatically the prostatic gland and sub-regions of interest, with U-net being a major breakthrough [4]. Nevertheless, further improvement in automatic segmentation of the prostatic gland and its subdivisions is required before it can be applied in the clinical practice.

Advanced pre-processing techniques, prior to network training, are essential in order to obtain a robust network and improve segmentation accuracy. In this direction, a major challenge remains the class imbalance problem which is present in the image labels where the background pixels outnumber the pixels of the region of interest (ROI) [5]. Training a model with imbalanced data can result to an unstable segmentation network, which is biased towards the majority class (background pixels). With the assumption that the ROI is located at the center of the image, standard center-cropping prior to network training has been used to reduce the number of background pixels with respect to the ROI [6]. This may be efficient for stable and large ROIs, such as the thorax, but in the case of prostate and particularly in the PZ segmentation, where the MRI depicts the whole pelvic anatomy, center-cropping endangers leading to inaccurate segmentation results [7].

In the present work we propose a DL-based cropping technique that allows to accurately crop the prostate PZ on T2w MRI images in order to improve PZ segmentation accuracy. The effect of DL-cropping for improving segmentation performance using three state-of-the-art segmentation networks is compared to the standard center-cropping technique.

II. METHODOLOGY

A. Dataset

In this work, data from the publicly available database "The Cancer Imaging Archive (TCIA)" [8] were used.
Specifically, the Prostate-X dataset was utilized, containing 98 patients along with their annotations performed by experts. The number of annotated frames on the peripheral zone is 1319. The main MRI vendors for this study was Siemens, MAGNETON Trio and Skyra models with a magnetic field strength of 3 Tesla. For the 98 patients the slice thickness of the examination was 3.6mm and the number of slices were between 15-22 for each patient while the frames dimension were 384X384 pixels and resizing to 256X256 was applied to meet the models’ requirements. Data augmentation was first applied on the original data to increase model variability and generalizability and the methods was the following affine transformations: (i) image rotation in different fixed degrees (-20, -10, -5, 5, 10, 20), and (ii) image shifting upwards, downwards, left or right, by a factor of 0.5.

B. Deep Learning pipeline for DL cropping

In the present study we propose a deep learning cropping technique in order to reduce the class imbalance on the pixels of the peripheral zone compared to the background pixels. Our method acts as a preprocessing step in which a bounding box is created which encloses the original annotation of the peripheral zone on each frame. The ROI box is enlarged by 40 pixels both horizontally and vertically from the initial mask and the training images were cropped around the indicating by the bounding box area [10]. For this purpose, a U-net network [4] was trained to crop the images in the testing dataset. In our work, the network is trained to segment an area around the peripheral zone resulting in a more balanced distribution between foreground and background pixels.

The pipeline used to define the bounding box is presented in Fig. 1. In steps (i) & (ii), a sample from the center-cropped images and the corresponding annotation are shown. These images are used as input in the u-net model (step (iii)) which was previously trained with the enlarged bounding boxes in order to automatically recognize the area around the PZ. In step (iv), the predictions are extracted from the original images marking roughly the region of interest. Furthermore, in step (v) the bounding boxes are defined using the minimum and maximum coordinates on the x and y axis of the amorphous masks from step (iv). This process ensures that the original annotations are always included in the final cropped image. Resampling to 256X256 pixels is applied on the frames in step (vi) in order the input images to comply with the network’s requirements. Finally, in steps (vii) and (viii) the resulted cropped frames and annotations are shown and the data from the aforementioned steps can be used for the training of the networks to segment precisely the prostate’s peripheral zone.

C. Deep Learning segmentation networks

In the work presented, the original versions of three state-of-the-art segmentation networks were implemented for the evaluation of the DL-cropping approach with respect to the standard center-cropping method. The first one is the U-net model [4] which utilizes an encoder-decoder combination of layers and those layers are connected in serial and parallel
way with each other, increasing the ability of the network to learn spatial features. Furthermore, Dense U-net [11] is an encoder-decoder network where dense blocks [12] are used to pass the information from previous layers forward while the transitional blocks reduce the computational cost of the network flattening and keeping the most important features produced by dense blocks. Bridged U-net [9] consists of 2 stacked U-nets and apart from the connections between encoder and decoder paths, there are also inter-network connections between layers from the first U-net to the second one enabling the communication and collaborative feature extraction between the networks.

D. Network training

The network architectures were trained and tested using two pipelines. In the first, standard approach, the networks were trained using the images after conventional center-cropping (Fig.1, steps i, ii). For the second pipeline, the proposed DL cropping approach was applied to crop the images and the resulting images (Fig.1, steps vii, viii) were used to train the networks. To ensure model generalizability, training was performed using 5-fold cross validation. Each training and validation set included 78 patients while the remaining 20 patients were utilized for testing. The patients were partitioned in the folds in the same way for all the architectures and for both DL-crop and center-crop approaches. The overall performance of the trained networks was computed by averaging the 5-fold cross validation results over the test-sets. The segmentation performance was evaluated based on the Dice score coefficient [13].

The number of images (2D slices) in each training fold was approximately 891 before and 1692 after data augmentation with image size 256x256 pixels and around 152 images were used for validation and 276 images were included in each testing fold. The partition of the folds was made patient-wise due to the variation of prostate’s peripheral zone and the need for the stratified splitting between folds. In this way our work is unbiased. For the training process, training accuracy and binary cross-entropy loss have been utilized as cost functions. The optimization method used is the Adam [14] instead of the Stochastic gradient descend [15] because it converges faster. The model was trained for 120 epochs for almost all architectures. Checkpoint strategy and early stopping have been used to reduce the computation time and the tensorboard was also used for monitoring.

III. RESULTS

Table I shows the segmentation performance in terms of mean Dice Score for the three networks using standard center-cropping and the proposed DL-cropping. The p-values were computed using the Wilcoxon signed-rank test to compare pairwise the two methods. Overall, the proposed pipeline outperformed the center-cropping method and the improvement was significant for all architectures (p<0.0001). The highest improvement was achieved with the U-net architecture corresponding to 27% while for Dense U-net and Bridged U-net the improvement was 15% and 12%, respectively. The corresponding boxplots of Dice score for the three segmentation networks using center- and DL-cropping are shown in Fig. 2.

An example of the segmentation result is depicted in Fig. 3. Blue and red contours indicate the original and predicted PZ masks, respectively, for both center-crop and DL-crop.

![Fig. 2: Dice Score boxplots of the prostate PZ for three networks using center-cropping and DL-cropping.](image)

IV. DISCUSSION AND CONCLUSION

The present work describes a novel preprocessing step for improving the performance of state-of-the-art prostate segmentation networks for segmenting the peripheral zone of the prostate on T2w MR images. A DL-based smart cropping is, therefore, proposed in order to tackle the class imbalance problem in pixel distribution between foreground (prostate) and background pixels. As we demonstrate, for all the prostate segmentation networks used for evaluation, the proposed DL cropping technique outperformed the standard center cropping.

In a typical prostate MRI, the prostate gland and especially the PZ represent only a small proportion of the entire image. However, a well-known problem of machine- and deep-learning algorithms is their limited prediction accuracy when they are trained on imbalanced datasets. When the classes have unbalanced representation in the image, then the most prevalent class often dominates training [16]. An appropriate
loss function during model training is usually applied to counteract the class imbalance problem in an imaging dataset. Commonly, the weighted cross-entropy loss is used to assign weights to classes based on the inverse of their occurrences. This results in a higher model penalization for the most prevalent class. Nevertheless, the choice of weighting is not straightforward and it is application-dependent [17]. A comparison of the effect of the most common loss functions on PZ segmentation is provided in [18].

A limitation of our study is that the number of patients available, may not have been sufficient to generate the results provided in the literature. For example, the U-net and Dense U-net architectures have previously reached a Dice score of 75% and 78%, respectively, for segmenting the prostate's PZ on a training population of 141 patients [18]. In our training population of 60 patients, the corresponding performance was no better than 69%. Clearly, more data are needed to establish the superiority of the proposed framework. To summarize, we developed a preprocessing method that effectively balances prostate MR images and significantly improves segmentation accuracy of PZ compared to the conventional center cropping. The generalizability of our method needs to be established through external validation on independent populations including images taken from different MRI vendors.

V. ACKNOWLEDGEMENTS

This work is supported by the ProCancer-I project, funded by the European Union’s Horizon 2020 research and innovation program under grant agreement No 952159. It reflects only the author’s view. The Commission is not responsible for any use that may be made of the information it contains.

REFERENCES

[1] de Rooij M, Israel B, Tummers M et al., “ESUR/ESUI consensus statements on multi-parametric MRI for the detection of clinically significant prostate cancer: quality requirements for image acquisition, interpretation and radiologists’ training,” European Radiology, vol. 30, no.10, pp. 5404-5416, 2020.

[2] Viswanath S, Bloch N, Chappelow J et al., “Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 tesla endorectal, in vivo T2-weighted MR imagery,” Journal of Magnetic Resonance Imaging, vol. 36, no. 1, pp. 213–224, 2012.

[3] Hassanzadeh T, Hamy L and Ho-Shon K, “Convolutional Neural Networks for Prostate Magnetic Resonance Image Segmentation,” IEEE Access, vol. 7, pp. 36748–36760, 2019.

[4] Ronneberger O, Fischer P and Brox T, “U-net: Convolutional networks for biomedical image segmentation,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9351, pp. 234–241, 2015.

[5] Gocer E, “Challenges and Recent Solutions for Image Segmentation in the Era of Deep Learning,” 2019 9th International Conference on Image Processing Theory, Tools and Applications, IPTA 2019, 2019.

[6] Feng X, Kun Qing K, Tustison NJ et al., “Deep convolutional neural network for segmentation of thoracic organs-at-risk using cropped 3D images,” Medical Physics, vol. 46, no. 5, pp. 2169–2180, 2019.

[7] To M, Vu D, Turkbey B et al., “Deep dense multi-path neural network for prostate segmentation in magnetic resonance imaging,” International Journal of Computer Assisted Radiology and Surgery, vol. 13, no. 11, pp. 1687–1696, 2018.

[8] Clark K, Vendt B, Smith K et al., “The cancer imaging archive (TCIA): Maintaining and operating a public information repository,” Journal of Digital Imaging, vol. 26, no. 6, pp. 1045–1057, 2013.

[9] Chen W, Zhang Y, He J et al., “Prostate segmentation using 2D bridged U-net,” arXiv, 2018.

[10] Zaridis DG, Mylona E, Tachos NS et al., “A new smart-cropping pipeline for prostate segmentation using deep learning networks,” arXiv, 2018.

[11] Guan S, Khan I, Sikdar S et al., “Fully dense U-Net for 2D sparse photoacoustic tomography artifact removal,” arXiv, 2018.

[12] Huang G, Liu Z, Van Der Maaten L et al., “Densely connected convolutional networks,” Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, vol. 2017-Jama, pp. 2261–2269, 2017.

[13] Zou K, Wurfield S, Bharadra A et al., “Statistical validation of image segmentation quality based on a spatial overlap index,” Academic radiologist, vol.11, no.2, pp. 178–189, 2004.

[14] Kingma D and Ba JI, “Adam: A method for stochastic optimization,” 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, pp. 1–15, 2015.

[15] Bottou L et al., “Large-scale machine learning with stochastic gradient descent,” in COMPSTAT, 2010.

[16] Taghanaki A, Abhishek S, Cohen K et al., “Deep semantic segmentation of natural and medical images: a review,” Artificial Intelligence Review, vol. 54, no. 1, pp. 137–178, 2021.

[17] Long J, Shelhamer E and Darrell T, “Fully Convolutional Networks for Semantic Segmentation,” arXiv, 2015.

[18] Aldo N, Biavati F, Michalek F et al., “Automatic prostate and prostate zones segmentation of magnetic resonance images using DenseNet-like U-net,” Scientific Reports, vol. 10, 2020.