**Abstract**

We propose a novel automatic method for accurate segmentation of the prostate in T2-weighted magnetic resonance imaging (MRI).

Our method is based on convolutional neural networks based autoencoders. Because of the large variability in the shape, size, and appearance of the prostate and the scarcity of annotated training data, we suggest training four separate Deep Learning Models. To fully exploit the training data, we synthesize additional data by deforming the training images and segmentations using a learned shape model. 4 separate Auto Encoder models based on CNN are trained separately on each type of Image – KTrans, T2, ADC and MHD. The outputs from the model can then be averaged to produce the final output or the result of prostate cancer

We apply the proposed method on the CancerX2 Challenge dataset and achieve state of the art results. Our proposed method generates accurate, smooth, and artefact-free segmentations.

**Introduction**

According to a current report of the American Cancer Society (ACS) for 2019, the number of deaths from prostate cancer and new cases of prostate cancer will be approximately 31,620 and 174,650, respectively. In the U.S, prostate cancer is the second leading cause of death, behind lung cancer [1]. This high disease burden represents a tremendous cost associated with the diagnosis of prostate cancer. Early stage detection is amenable to a greater range of treatment options and could lead to better outcomes [2].

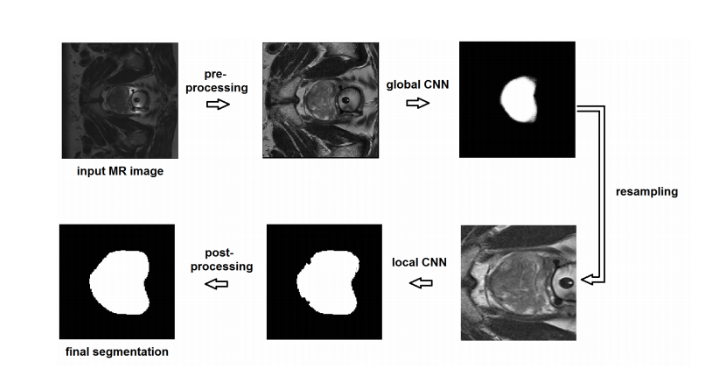
The two most common methods for prostate cancer screening are through the detection of prostate specific antigen (PSA) in blood test and this is usually followed by the transrectal ultrasound (TRUS) biopsy for further diagnosis. One of the critics of the PSA blood test is its relatively high false positive rates that leads to unnecessary TRUS biopsy which is invasive and may sometimes lead to complications. MP-MRI scanning is a non-invasive way to detect early stage prostate cancer. It was also observed that more than 20% of the prostate cancers are missed or under sampled during the first biopsy session [5]. Thus, the use of other non-invasive imaging techniques such as Magnetic Resonance Imaging (MRI) are increasingly being preferred as a prostate cancer computer aided diagnostic (CAD) tool.

**Existing system & Related Works**

A few previous papers have been published in developing an automatic Gleason grading system for prostate cancer diagnosis. A commonly used approach is to extract tissue features and apply classifiers upon the selected features. Stotzka et al. [8] extracted statistical and structural features from the spatial distribution of epithelial nuclei over the image area. They used a hybrid neural network/Gaussian statistical classifier to distinguish moderately and poorly differentiated histological samples. Smith et al. [9] used the power spectrum of tissue images to represent their texture characteristics. They used a nearest neighbor classifier to assign the input image to Gleason grades 1 through 3 and the combined grades of 4 and 5. Wetzel et al. [10] proposed the use of features derived from spanning trees connecting cell nuclei across the tumor image to represent tissue images belonging to each grade. Jafari-Khouzani and Soltanian-Zadeh [11] used features based on co-occurrence matrices, wavelet packets, and multi-wavelets combined with a k -nearest neighbor (k NN) classifier to classify each image into grades 2 through 5. Farjam et al. [12] proposed a multistage classifier based on morphometric and texture features for Gleason grading. First, gland units are identified using texture features. Then, morphometric and texture features obtained from gland units are used in a series of classification stages to classify the image into grades 1 through 5. Tabesh et al. [13] aggregated color, texture, and morphometric cues at the global and histological object levels for classification and compared Gaussian, k -nearest neighbor, and support vector machine classifiers along with the sequential forward feature selection algorithm. Nguyen et al. [14] used structural features of prostate glands to classify pre-extracted regions of interest (ROIs) into benign, G3, and G4. Gorelick et al. [15] proposed a two stage Adaboost model to classify around 991 sub-images extracted from 50 whole-mount sections of 15 patients.

Though most of these papers achieved good results on their datasets due to heavy reliance on feature extraction, the systems described above are prone to subjectivity and limited intra- and inter-system reproducibility. Moreover, all the systems require accurate localization of the small image area (region of interest, RoI) to extract features from, which is a non-trivial problem [16].

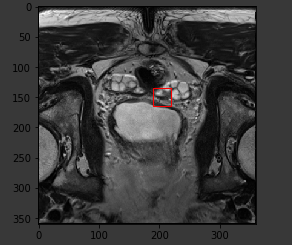
There were also few Ensemble CNN Models which use global and local CNN models which picks up features at various levels. A global CNN picks up high level features. Local CNN Picks the low-level features to classify as Cancer / non-Cancer patient. This is one of the state-of-art models besides U-Net.



**Fig 1 – State of art model 1**

**Overview of Proposed System**

* Data for 4 Channels – ADC, T2, KTrans, Bval are retained and processed separately without combining them. Hence, Different models for different channels
* Based on Gleason grade, Cancer Images are marked as Clinically Significant or not
* Clinically Significant Cancers are Marked



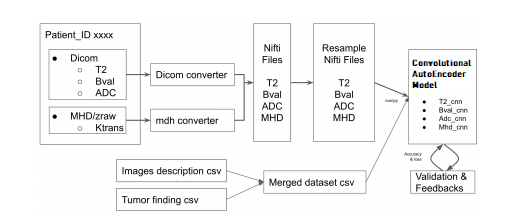
**Fig 2 – Raw Image**

**Dataset**

The dataset used is taken from PROSTATEx-2 — SPIEAAPM-NCI Prostate MR Gleason Grade Group Challenge [22]. The dataset consists of 162 patients’ MRI scans, 99 patients as train cases and 63 patients as test cases. Each patient has five sets of MRI sequence: two sets of T2-weighted images (transverse and sagittal; DICOM format), one set of apparent diffusion coefficient (ADC) images (DICOM format), one set of diffusion-weighted (DWI) images (DICOM format), and one set of Ktrans images (MHD format). All the sequences were used except for the T2-weighted sagittal sequence. Each patient may have more than 1 lesion findings in his or her MRI sequences. Each lesion finding is labelled a Gleason Grade Group (GGG) of 1-5, 1 being the most favorable and 5 being the least favorable, and the X, Y, and Z coordinates of each lesion finding was also given in the accompanying csv file.

**Pre-processing**

Pre-processing steps are as shown in the diagram

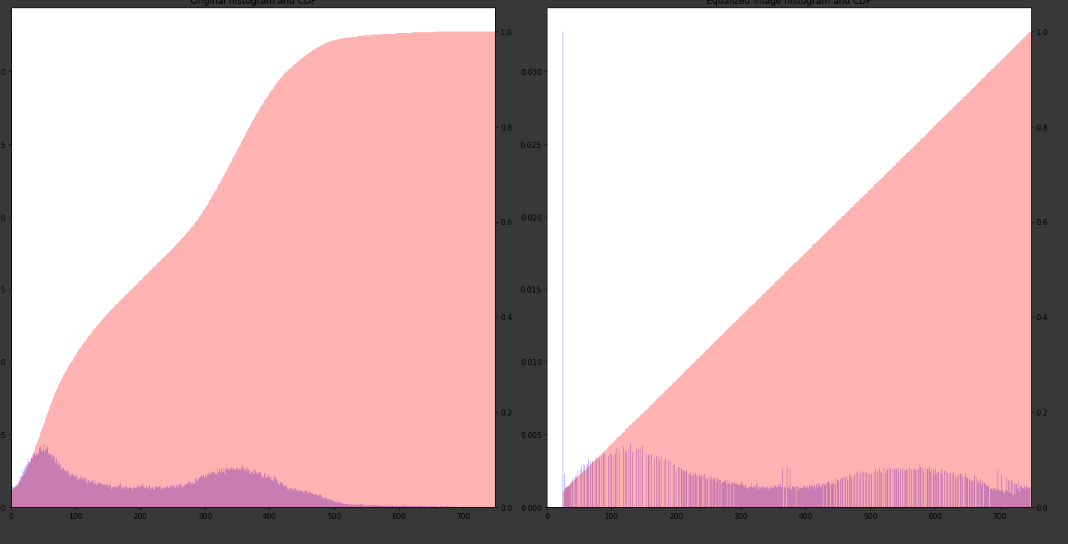


**Conversion to NIFTI file format**

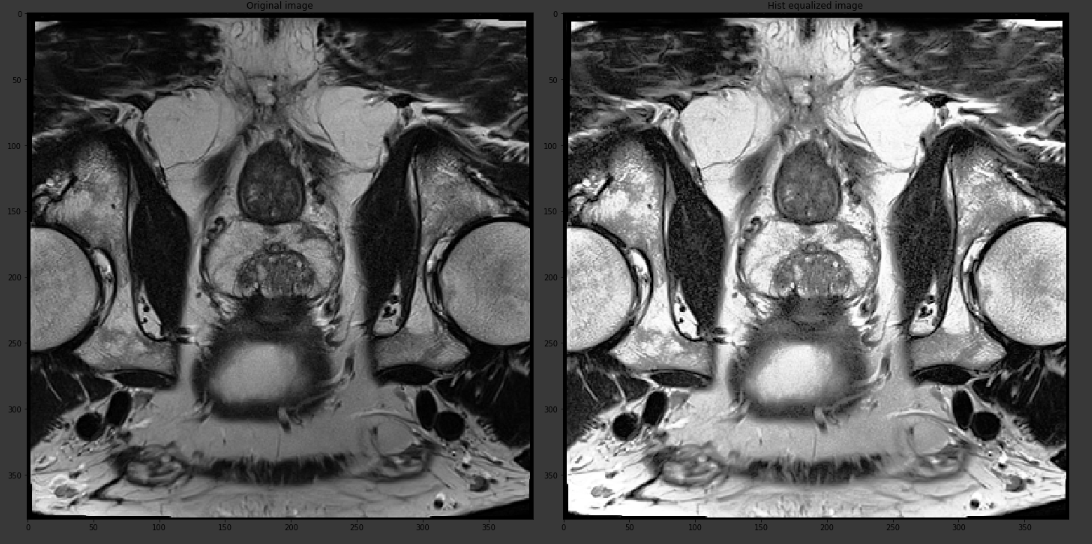
All the DICOM files and MHD files were first converted to NIFTI files using SimpleITK that use to convert DICOM to NIFTI

**Data cleansing** consists of images resampling, histogram equalization, and class balance in between benign and malignant [10]. Before extracting tumour patch from images, the raw images were converted into NIFTI format with resampling image size to fit into the NN model. Histogram equalization was used to improve the global contrast of each format images. The following sample images were processed with reference to cumulative density function

From the equalized images comparison above, the contrast intensity value was dramatically improved.



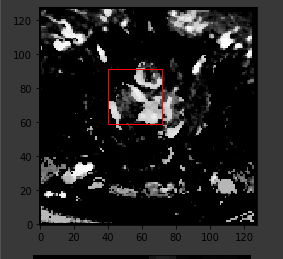
Histogram Equalization – Before / After



Cancer Image – Before / After Histogram equalization

And the next step of data cleansing, **data balancing between benign/malignant classes** balance for the dataset was processed. Initially, the class balance was approximately 3.3 to 1 as benign to malignant. **Hence, those cancer samples were oversampled by 4 rotations (45, 90, 180, 270)** for adding to the final dataset, an example of T2w image format is displayed.

**Patch size of tumor extraction** was based on performance evaluation of a prototype NN model with original image size. Particularly, for T2w, the original image size is 384 × 384 while for ADC, Bval, and Ktrans, the original image sizes are 128 × 84, 128 × 84, and 128 × 128, respectively. Therefore, the patch matrix sizes for T2w, ADC, Bval and Ktrans were designed as 32 × 32, 8 × 8, 8 × 8 and 8 × 8, respectively for data input of individual NN model.



**Labels Extraction**

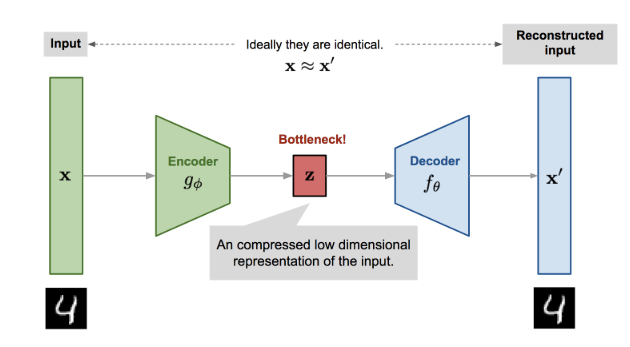
The labels given are the GGG (1-5) of each lesion findings. The lesion findings can be further grouped to be of no clinical significance when the GGG is 1 and of clinical significance when the GGG is greater than 1. This can be seen when comparing these findings of PROSTATEx2 to the findings of PROSTATEx as they share the same cases. Each lesion case is then re-labelled to **True or False** depending on whether they are of clinical significance based on the GGG findings. With binary labels, the task ahead is binary classifications on whether the lesion is of clinical significance.

**Train-Test**

Split The dataset consists of 99 patients’ MRI sequences, but each patient may have more than 1 lesion finding in his or her MRI sequences. A total of 112 findings was present and extracted as described in the pre-processing steps. The 112 findings were split to 78 findings for training and 34 findings for validation.

**Proposed System**

Convolutional Autoencoder are autoencoders that use CNNs in their encoder/decoder parts.

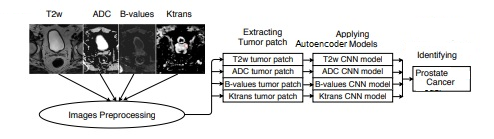


**MODEL DESCRIPTION**

1. Convolutional Autoencoder is an autoencoder, a network that tries to encode its input into another space (usually a smaller space) and then decode it to its original value.
2. Based on the type of input, you can use different types of networks for encoder and decoder. For example, if your input is Bag Of Word Vector, you’ll probably use a Fully Connected network; if your input is a time series, you’ll probably use a recurrent neural network (RNN); and if your input is image, you’ll probably use a Convolutional Neural Network (CNN).
3. In this case it is an encoder decoder network with CNN + RELU

**Implementation details**

The proposed network is trained in an end-to-end manner from scratch.

****

All weights are initialized from the zero-mean.

All four Models are trained independently based on T2W, ADC, B-Val and KTrans dataset.

All four Convolutional Autoencoder models Our model is implemented in keras.

The results of the models are averaged out to conclude whether the patient has cancer or not.

**Training, Testing & Results**

**Training Time:**

|  |  |
| --- | --- |
| Method | Training Time / Epoch after tuing the hyper parameters and cudnn parameters |
| Base CNN | 2 hrs |
| Convolutional Auto Encoders | 3 hrs |

Base CNN Model with K-Fold cross validation (5-Fold):

|  |  |
| --- | --- |
| **Channel** | **Accuracy** |
| T2 | 78.85 |
| ADC | 79.96 |
| B-Val | 78.87 |
| K-Trans | 73.62 |

Convolutional Autoencoder Model with K-Fold cross validation (5-Fold):

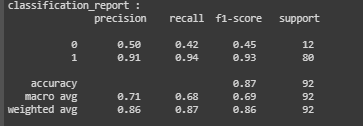
|  |  |
| --- | --- |
| **Channel** | **Accuracy** |
| T2 | 80.18 |
| ADC | 78.88 |
| B-Val | 80.39 |
| K-Trans | 79.52 |

As we can see, the average accuracy of the models has improved compared to the baseline model.

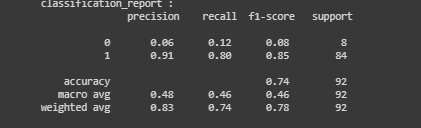
Let’s look at the confusion matrix and roc curves to examine the model and address any issues with the model.

|  |  |
| --- | --- |
| Type | AutoEncoder |
| T2 | A picture containing screenshot  Description automatically generated |
| ADC | A picture containing screenshot  Description automatically generated |
| B-Val | A picture containing screenshot  Description automatically generated |
| K-Trans | A picture containing screenshot  Description automatically generated |

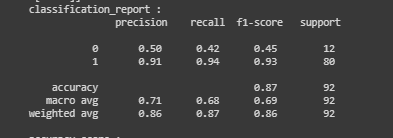
T2:



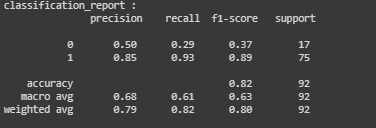
ADC:



KTrans:



BVal:



**Conclusion and Future Work**

Binary classification of significance of GGG score was done. Previously, using a separate dataset which has a mask provided – Cancer areas were segmented. Future work could explore obtaining a dataset with both tumour masks and GGG score provided. A more balanced dataset could also aid multi-label classification of GGG as opposed to binary classification of significance or not. With this dataset, an end-to-end prostate segmentation and tumour recognition system can be developed to segment the prostate from MRI scans, cropped sequences at the prostate region and classify each cropped sequence to be of a GGG group, thus identifying the location of tumour and its seriousness.

**References**

1. R. L. Siegel, K. D. Miller, A. Jemal, "Cancer statistics 2016", *CA Cancer J. Clin.*, vol. 66, no. 1, pp. 7-30, 2016.
2. D. F. Gleason, "Histologic grading of prostate cancer: A perspective", *Hum. Pathol.*, vol. 23, no. 3, pp. 273-279, 1992.
3. J. I. Epstein et al., "The 2005 international society of urological pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma", *Amer. J. Surg. Pathol.*, vol. 29, no. 9, pp. 1228-1242, 2005.
4. H. J. Lavery, M. J. Droller, "Do Gleason patterns 3 and 4 prostate cancer represent separate disease states?", *J. Urol.*, vol. 188, no. 5, pp. 1667-1675, 2012.
5. C. C. Huang et al., "Gleason score 3+4=7 prostate cancer with minimal quantity of Gleason pattern 4 on needle biopsy is associated with low-risk tumor in radical prostatectomy specimen", *Amer. J. Surg. Pathol.*, vol. 38, no. 8, pp. 1096-1101, 2014.
6. P. A. Humphrey, "Gleason grading and prognostic factors in carcinoma of the prostate", *Mod. Pathol.*, vol. 17, no. 3, pp. 292, 2004.
7. K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask R-CNN", *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, pp. 2980-2988, Oct. 2017.
8. R. Stotzka, R. Männer, P. H. Bartels, D. Thompson, "A hybrid neural and statistical classifier system for histopathologic grading of prostatic lesions", *Anal. Quant. Cytol. Histol.*, vol. 17, no. 3, pp. 204-218, 1995.
9. Y. Smith, G. Zajicek, M. Werman, G. Pizov, Y. Sherman, "Similarity measurement method for the classification of architecturally differentiated images", *Comput. Biomed. Res.*, vol. 32, no. 1, pp. 1-12, 1999.
10. A. W. Wetzel et al., "Evaluation of prostate tumor grades by content-based image retrieval", *Proc. 27th AIPR Workshop Adv. Comput.-Assist. Recognit.*, vol. 3584, pp. 244-253, 1999.
11. K. Jafari-Khouzani, H. Soltanian-Zadeh, "Multiwavelet grading of pathological images of prostate", *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 697-704, Jun. 2003
12. R. Farjam, H. Soltanian-Zadeh, R. A. Zoroofi, K. Jafari-Khouzani, "Tree-structured grading of pathological images of prostate", *Proc. SPIE*, vol. 5747, pp. 840-852, Apr. 2005.
13. A. Tabesh et al., "Multifeature prostate cancer diagnosis and Gleason grading of histological images", *IEEE Trans. Med. Imag.*, vol. 26, no. 10, pp. 1366-1378, Oct. 2007.
14. K. Nguyen, B. Sabata, A. K. Jain, "Prostate cancer grading: Gland segmentation and structural features", *Pattern Recognit. Lett.*, vol. 33, no. 7, pp. 951-961, 2012.
15. L. Gorelick et al., "Prostate histopathology: Learning tissue component histograms for cancer detection and classification", *IEEE Trans. Med. Imag.*, vol. 32, no. 10, pp. 1804-1818, Oct. 2013
16. A deep learning-based method for prostate segmentation in T2-weighted magnetic resonance imaging - Davood Karimi, Golnoosh Samei, Yanan Shao, Tim Salcudean
17. Ensemble of Convolutional Neural Networks for the Detection of Prostate Cancer in Multi-Parametric MRI Scans - Quang H. Nguyen, Mengnan Gong, Tao Liu, Ou Yang Youheng, Binh P. Nguyen and Matthew Chin Heng Chua
18. PROPOSAL FOR 2-STAGE CONVOLUTIONAL PROSTATE SEGMENTATION AND TUMOUR CLASSIFICATION - Chian Yan Tao Eugene, Han Yuen Kwang Andy
19. <https://medium.com/@SeoJaeDuk/achieved-post-a-novel-deep-learning-based-method-for-prostate-segmentation-in-t2-weighted-229cec073758>
20. <https://towardsdatascience.com/autoencoders-are-essential-in-deep-neural-nets-f0365b2d1d7c>
21. <https://arxiv.org/abs/1901.09462>
22. <https://openreview.net/pdf?id=Syoj0k2iG>
23. <https://www.sciencedirect.com/science/article/pii/S2352914819302588>
24. <https://github.com/alexhamiltonRN/ProstateX>
25. [https://wiki.cancerimagingarchive.net/display/Public/SPIE-AAPM-NCI+PROSTATEx+Challenges#715aea9571044e3e8577f712808173e8](https://wiki.cancerimagingarchive.net/display/Public/SPIE-AAPM-NCI+PROSTATEx+Challenges)
26. <https://arxiv.org/ftp/arxiv/papers/1904/1904.02575.pdf>
27. <https://meetshah1995.github.io/semantic-segmentation/deep-learning/pytorch/visdom/2017/06/01/semantic-segmentation-over-the-years.html>
28. <https://towardsdatascience.com/semantic-segmentation-popular-architectures-dff0a75f39d0>
29. <https://towardsdatascience.com/semantic-segmentation-of-aerial-images-using-deep-learning-90fdf4ad780>
30. <https://towardsdatascience.com/u-net-b229b32b4a71>
31. de Rooij M, Hamoen EH, Fütterer JJ, Barentsz JO, Rovers MM (2014) Accuracy of multiparametric MRI for prostate cancer detec-tion: a meta-analysis. Am J Roentgenol 202(2):343–351
32. Fütterer JJ, Briganti A, De Visschere P, Emberton M, Giannarini G, Kirkham A, Taneja SS, Thoeny H, Villeirs G, Villers A (2015) Can clinically significant prostate cancer be detected with multi-parametric magnetic resonance imaging? a systematic review of the literature. Eur Urol 68(6):1045–1053
33. Garvey B, Türkbey B, Truong H, Bernardo M, Periaswamy S, Choyke PL (2014) Clinical value of prostate segmentation and vol-ume determination on MRI in benign prostatic hyperplasia. Diagn Interv Radiol 20(3):229
34. Valerio M, Donaldson I, Emberton M, Ehdaie B, Hadaschik BA, Marks LS, Mozer P, Rastinehad AR, Ahmed HU (2015) Detection of clinically significant prostate cancer using magnetic resonance imaging—ultrasound fusion targeted biopsy: a systematic review. Eur Urol 68(1):8–19
35. Muller BG, Fütterer JJ, Gupta RT, Katz A, Kirkham A, Kurhanewicz J, Moul JW, Pinto PA, Rastinehad AR, Robertson C (2014) The role of magnetic resonance imaging (MRI) in focal therapy for prostate cancer: recommendations from a consensus panel. BJU Int 113(2):218–227
36. Ghai S, Louis AS, Van Vliet M, Lindner U, Haider MA, Hlasny E, Spensieri P, Van Der Kwast TH, McCluskey SA, Kucharczyk W (2015) Real-time MRI-guided focused ultrasound for focal therapy of locally confined low-risk prostate cancer: feasibility and preliminary outcomes. Am J Roentgenol 205(2):W177–W184
37. Zhu Y, Williams S, Zwiggelaar R (2004) Segmentation of volu-metric prostate MRI data using hybrid 2D + 3D shape modeling. In: Proceeding of medical image understanding and analysis, pp 61–64
38. Allen PD, Graham J, Williamson DC, Hutchinson CE (2006) Differential segmentation of the prostate in MR images using com-bined 3D shape modelling and voxel classification. In: 3rd IEEE international symposium on biomedical imaging: nano to macro. IEEE, pp 410–413
39. Toth R, Madabhushi A (2012) Multifeature landmark-free active appearance models: application to prostate MRI segmentation. IEEE Trans Med Imaging 31(8):1638–1650
40. Zwiggelaar R, Zhu Y, Williams S (2003) Semi-automatic seg-mentation of the prostate. In: Perales FJ, Campilho AJC, de la Blanca NP, Sanfeliu A (eds) Pattern recognition and image analy-sis. Springer, Berlin, pp 1108–1116
41. El Naqa I, Yang D, Apte A, Khullar D, Mutic S, Zheng J, Bradley JD, Grigsby P, Deasy JO (2007) Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. Med Phys 34(12):4738–4749
42. Yin Y, Fotin SV, Periaswamy S, Kunz J, Haldankar H, Muradyan N, Cornud F, Turkbey B, Choyke P (2012) Fully automated prostate segmentation in 3D MR based on normalized gradient fields cross-correlation initialization and LOGISMOS refinement. In: Medical imaging 2012: image processing. International Society for Optics and Photonics, p 831406
43. Zhu Q, Du B, Turkbey B, Choyke PL, Yan P (2017) Deeply-supervised CNN for prostate segmentation. In: 2017 international joint conference on neural networks (IJCNN). IEEE, pp 178–184
44. Yu L, Yang X, Chen H, Qin J, Heng P-A (2017) Volumetric Con-vNets with mixed residual connections for automated prostate segmentation from 3D MR images. In: AAAI, pp 66–72