Abstract ESTRO 40

Development and quantitative evaluation of AI-based pelvic MRI autocontouring for adaptive MRgRT

M.Nachbar[1], M.Lo Russo[2], S.Boeke[2], D.Wegener[2], J.Boldt[2], S.Butzer[2], T. Roque[3], A.Lombard[3], L. De Vitry[3], N.Paragios[3], D.Zips[2], D.Thorwarth[1]

[1] Section for Biomedical Physics. Department of Radiation Oncology. University Hospital and Medical Faculty. Eberhard Karls University Tübingen, Tübingen. Germany
[2] Department of Radiation Oncology. University Hospital and Medical Faculty. Eberhard Karls University Tübingen. Tübingen. Germany
[3] TheraPanacea, Paris, France

Purpose
In MR-guided radiotherapy (MRgRT) online plan adaptation can account for tumor shrinkage, interfraction motion and allows thus daily sparing of relevant OARs close to the target. Due to the high interfraction variability of bladder and rectum, patients with tumors in the pelvic region may strongly benefit from adaptive MRgRT. However, manual annotation of relevant structures limits the potential of online adaptive MRgRT, as time spent for MRI contouring results in intrafraction motion. Therefore a deep neural network was generated on MRIs of the 1.5 T MR-Linac and evaluated for the implementation in an online adaptive workflow.

Methods
From 47 patients, T2w MRI data acquired on a 1.5 T MR-Linac (Unity, Elekta) at five treatment fractions were contoured, identifying prostate, seminal vesicles, rectum, anal canal, bladder, penile bulb, and bony structures (cf. figure 1). This data was used for training in a three steps approach:: (i) a localization algorithm that maps data to different reference anatomies, (ii) automatic - multiple - delineation of the anatomical structures on the reference spaces, and (iii) a winner takes all approach among models predictions that enforces anatomical consistency and produces optimal outcomes at the original space. For quantitative evaluation, 10 Un-seen T2w-MRIs were contoured by a radiation oncologist (RO) and compared in MATLAB to the automatic contours (AI-C). For quantitative evaluation, dice similarity coefficients (DSC) and 95% Hausdorff distances (HD95) were calculated in a caudal-cranial window of ± 3 cm with respect to the PTV. For qualitative evaluation, three radiation oncologists scored the AI-C for possible usage in an online adaptive workflow as follows: (1) can be used as is; (2) slight modifications necessary; (3) large modifications necessary; (4) not usable.

Results
AI-C datasets were generated in a mean time of 52s. Qualitative evaluation revealed that 64% of structures could be used immediately, 27% needed minor, 7% large modifications and 2% would not be usable for online MRgRT (figure 1).
Quantitative comparison between RO and AI-C showed good agreement for all bony structures, except sacrum, with a median DSC >86%. Structure specific median DSCs varied between 65% for seminal vesicle and 95% for bladder. For all structures median HD95 were <5.8 mm. Structure specific results are depicted in figure 2.

Conclusion:

We show first results for the training and validation of AI-based autocontouring with the potential for real time annotation in online adaptive MRgRT. Physician based scoring shows good acceptance, corroborated by small HD95 in a quantitative evaluation (figure 1). However, prior to an online implementation the dosimetric effect of differences in DSC as well as more robust modeling for unusual anatomies must be investigated.

Funding: German Research Council (ZI 736/2-1, PAK 997/1).and European Union's Horizon 2020 research and innovation programme under grant agreement No. 880314

COI: Research agreement with Elekta, Therapanacea, Philips & PTW

Structure/Scoring	1	2	3	4
Anal canal	77%	23%	0%	0%
Bladder	40%	50%	0%	10%
Penile bulb	80%	20%	0%	0%
Prostate	13%	43%	40%	3%
Rectum	40%	47%	10%	3%
Seminal vesicle	33%	47%	17%	3%
Femur L	97%	3%	0%	0%
Femur R	100%	0%	0%	0%
Pelvis L	83%	17%	0%	0%
Pelvis R	80%	20%	0%	0%
Sacrum	57%	30%	13%	0%
Over all structures	**64%**	**27%**	**7%**	**2%**

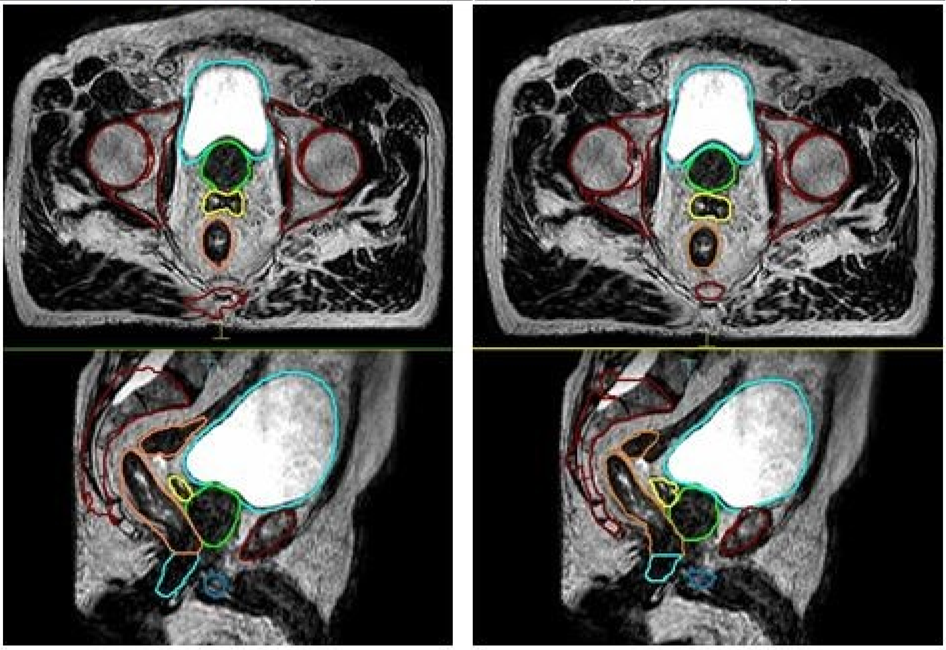

Figure 1: Physician based scoring (top) and corresponding visual comparison between AIC (bottom-left) and ROC (bottom-right)

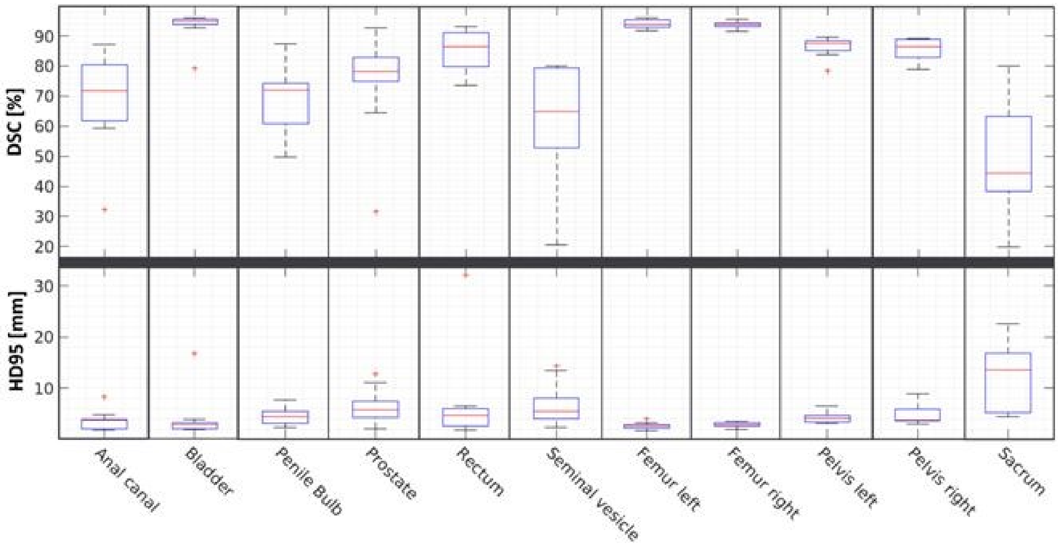

Figure 2 Structure specific quantitative comparison between manual and AI-based contours. Top row depicts the dice similarity coefficient, whereas the bottom row shows the 95% Hausdorff distance.