A deep learning-based framework for segmenting CTV with estimated uncertainties for post-operative prostate cancer radiotherapy

JULY 12–16 2020 VIRTUAL
JOINT AAPM | COMP MEETING
EASTERN TIME [GMT-4]

Anjali Balagopal, Dan Nguyen, Howard Morgan, Yaochung Weng, Michael Dohopolski, Mu-Han Lin, Azar Sadeghnejad Barkousaraie, Yesenia Gonzalez, Aurelie Garant, Neil Desai, Raquibul Hannan, Steve Jiang*
Medical Artificial Intelligence and Automation (MAIA) Laboratory, Department of Radiation Oncology, UT Southwestern Medical Center, Dallas, Texas
Steve.Jiang@utsouthwestern.edu

MAIA Laboratory
UT Southwestern Medical Center
Radiation Oncology

INTRODUCTION:

In post-operative radiotherapy for prostate cancer, precisely contouring the clinical target volume (CTV) to be irradiated is challenging, because the cancerous prostate gland has been surgically removed, so the CTV encompasses the microscopic spread of tumor cells, which cannot be visualized in clinical images like computed tomography or magnetic resonance imaging. In current clinical practice, physicians segment CTVs manually based on their relationship with nearby organs and other clinical information, but this allows large inter-physician variability. Automating post-operative prostate CTV segmentation with traditional image segmentation methods has yielded suboptimal results. We propose using deep learning to accurately segment post-operative prostate CTVs.

METHODS:

A three stage network is proposed. CTV and OAR volumes are localized and cropped from the original CT images through a 2D localization network; OARs are segmented individually by separate 3D segmentation networks; and CTV is segmented by a dedicated 3D segmentation network that takes the localized CTV volume and segmented bladder and rectum as inputs.

- The model proposed is trained using labels clinically approved and used for patient treatment.
- To segment the CTV, we segment nearby organs first, then use their relationship with the CTV to assist CTV segmentation.
- To make the DL model practical for clinical use, Monte Carlo dropout[1] is leveraged to give physicians the 95% confidence bounds together with the predicted mean contours.

RESULTS :

The model achieves an average Dice similarity coefficient (DSC) of 0.87 on a holdout test dataset, better than established methods, such as atlas-based methods (DSC<0.7). For each test patient, two out of five residents manually segmented the CTV with the assistance of pathology and MRI reports. We compared the DSC values between resident-drawn contours and clinical contours with the DSC values between the model contours and clinical contours for all test patients. The model performance was observed to be statistically superior. Experienced practicing physicians were presented, in a randomized and blinded way, with the model-predicted CTV contour and the clinical CTV contour, side by side.
These physicians, with the assistance of pathology reports, reviewed and scored the contours according to a 4-point grading system:

- 4 - acceptable without changes, 3 - acceptable with minor changes, 2 - acceptable with major changes, and 1 - completely unacceptable.
- Half of the patients were evaluated by their respective original treating physicians(same-observer evaluation), and the rest were evaluated by a physician who was not involved in the original treatment(Different-observer evaluation).

Structure	DSC(%)	ASD(mm)
	Mean ± SD	
CTV	86.8 ± 4.9	1.57 ± .46
Bladder	95.4 ± 6.8	1.02 ± .35
Rectum	90.2 ± 2.3	1.35 ± .56
Penile Bulb	77.4 ± 7.3	1.70 ± .67
Left Fem Head	96.0 ± 2.0	0.98 ± .24
Right Fem Head	95.8 ± 1.8	1.11 ± .23

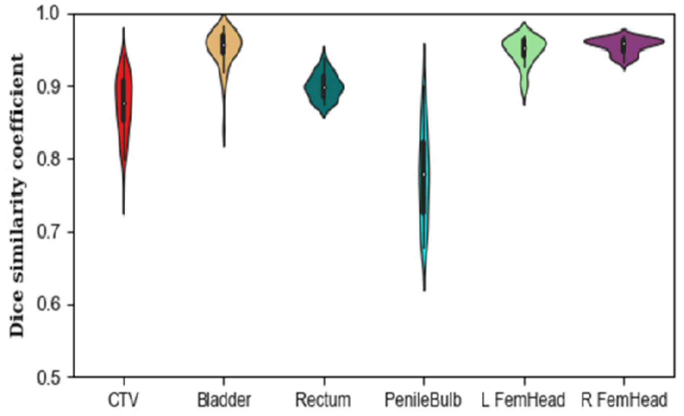

***Figure 1:** Quantitative evaluation of the predicted CTV and OAR contours against the clinical contours. **Top**, Mean values and standard deviations of Dice similarity coefficient (DSC) and average surface distance (ASD); **Bottom**, Violin plots of DSC values*

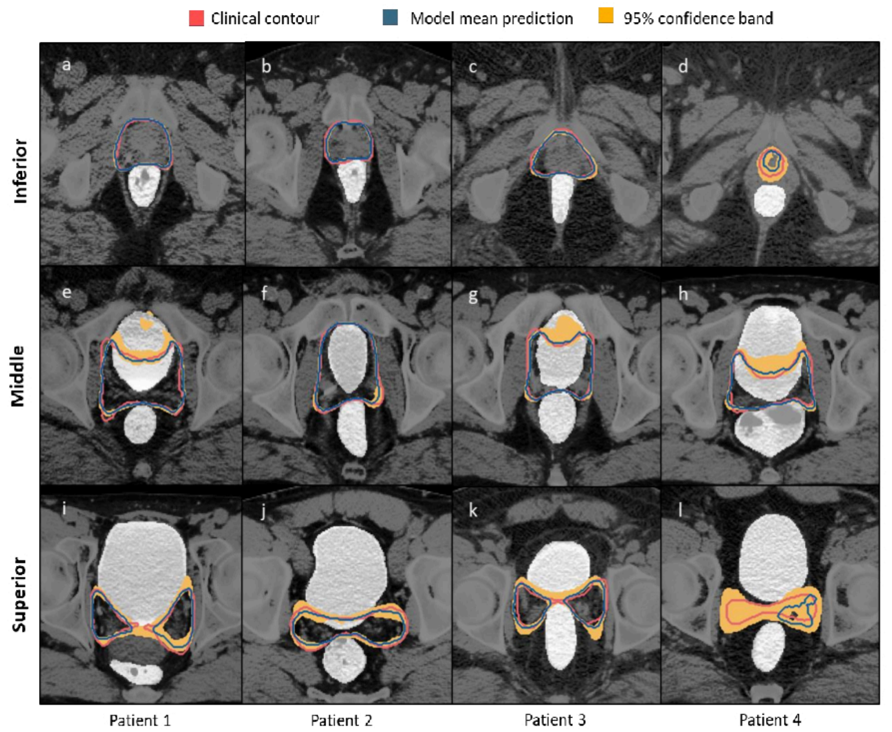

***Figure 2:** Visualization of the clinical CTV contours (red) and the predicted mean CTV contours (blue) with 95% confidence bounds (yellow) in axial CT images at three representative anatomical locations (top row - inferior, middle row – middle, and bottom row – superior) for four example testing patients (each column corresponding to one patient).*

	Mean	Variance	p-value
Model DSC	86.88%	0.17%	
Avg Resident DSC	79.56%	0.35%	1.21×10^{-12}
Max Resident DSC	81.75%	0.32%	1.82×10^{-6}

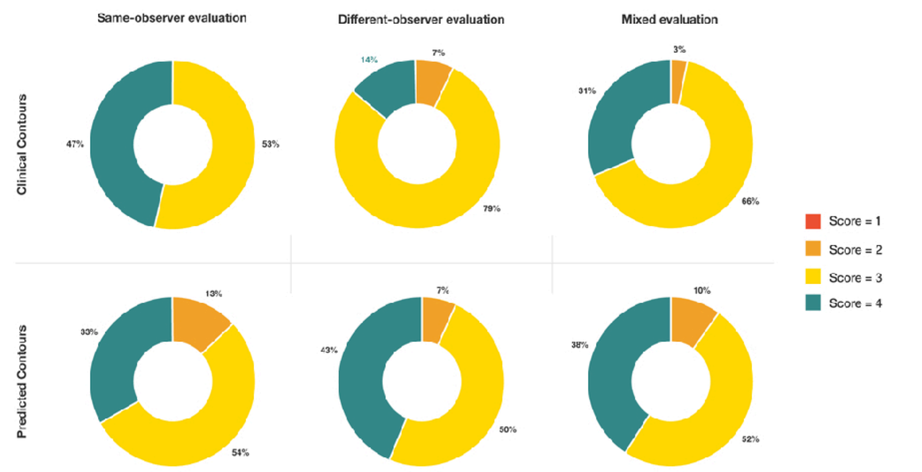

***Figure 3: Top.** The mean and variance of DSC values of the model-predicted CTV contours, the average of the two residents, and the better of the two residents for each patient, Bottom: Summary of the evaluation of clinical acceptability using a 4-point grading system. None of the contours received a score below 2.*

CONCLUSIONS :

In this work, we have designed and implemented a deep learning based framework that can segment invisible clinical tumor volumes on computed tomography images for post-operative prostate cancer radiation therapy. Uncertainty associated with the deep learning model is estimated , enabling the visualization of areas of large variability which would help physicians make informed and efficient corrections to the deep learning produced tumor segmentation.

REFERENCES

1. Y. Gal and Z. Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In international conference on machine learning, pages 1050– 1059, 2016.