Clinical implementation of deep learning: Automatic contouring via U-Net architecture

Matthew Cooper[1]
Simon Biggs[2]
Yu Sun[1]
Matthew Sobolewski[2]

[1]The University of Sydney (USyd). School of Physics. Institute of Medical Physics.

[2]Riverina Cancer Care Centre (RCCC). Cancer Care Associates.

Thesis: github.com/matthewdeancooper/masters_thesis
Video overview: docs.pymedphys/com/background/autocontouring

Current limitations

Variability

- Large intra and inter-observer variance (IOV).[1]
- AAPM TG275 risk assessment - multiple human-factor failure modes in RT.[2]

Time constraints

- Atlas methods $\implies$ significant correction times.[3]
- Barrier to future technologies that require fast contouring.[3]

Deep learning potential

- Shown to reduce IOV and contouring time.[3]
- Significant improvement cf. atlas methods (time & accuracy).[4]

---

[1] Dale Roach et al. "Multi-observer contouring of male pelvic anatomy: Highly variable agreement across conventional and emerging structures of interest". In: *Journal of Medical Imaging and Radiation Oncology* 63.2 (2019), pp. 264–271. DOI: 10.1111/1754-9485.12844

[2] Eric Ford et al. "Strategies for effective physics plan and chart review in radiation therapy: Report of AAPM Task Group 275". In: *Medical Physics* 47.6 (2020), e236–e272. DOI: https://doi.org/10.1002/mp.14030

[3] Shalini K Vinod et al. "A review of interventions to reduce inter-observer variability in volume delineation in radiation oncology". In: *Journal of Medical Imaging and Radiation Oncology* 60.3 (2016), pp. 393–406. DOI: 10.1111/1754-9485.12462

[4] Stanislav Nikolov et al. *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*. 2018. arXiv: 1809.04430 [cs.CV]

Research goals

Model 1: QA tool - Pelvic imaging for prostate cancer (Patient, bladder, rectum).

- Alert if prediction differs significantly from expert.
- Need for delineation to be part of regular QA.[4]

Model 2: Automatic contouring - Canine vacuum bag

- Manual vacuum bag contouring $\approx$ 30 min

Goal: Performance similar to human experts.

- Performance metric (sDSC) that takes into account expert IOV.[3]

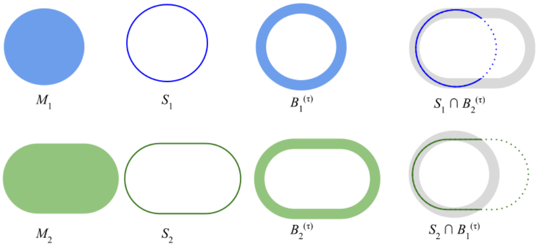
[3] Stanislav Nikolov et al. *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*. 2018. arXiv: 1809.04430 [cs.CV]

[4] Shalini K Vinod et al. "A review of interventions to reduce inter-observer variability in volume delineation in radiation oncology". In: *Journal of Medical Imaging and Radiation Oncology* 60.3 (2016), pp. 393–406. DOI: 10.1111/1754-9485.12462

Surface dice similarity coefficient (sDSC)

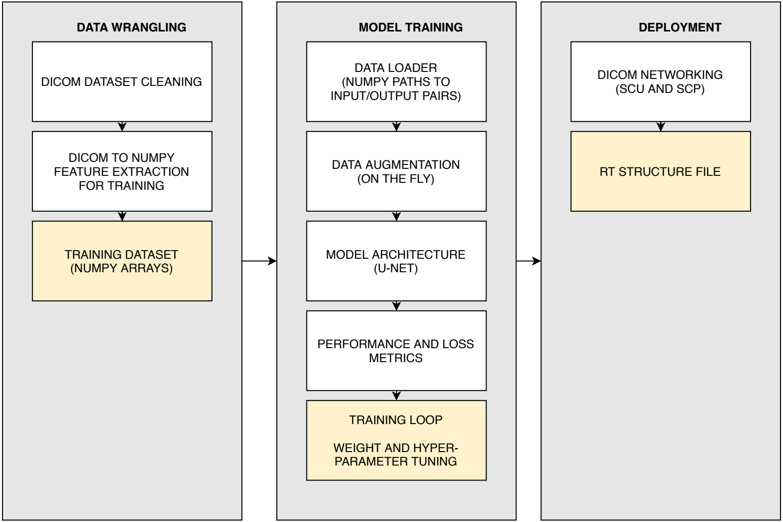
$$DSC_{1,2} = \frac{2|M_1 \cap M_2|}{|M_1| + |M_2|}$$

$$sDSC_{1,2}^{(\tau)} = \frac{|S_1 \cap B_2^{(\tau)}| + |S_2 \cap B_1^{(\tau)}|}{|S_1| + |S_2|}$$

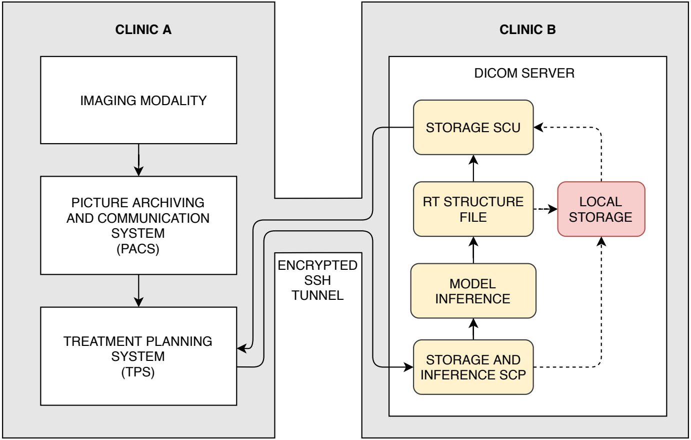


[3] Stanislav Nikolov et al. *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*. 2018. arXiv: 1809.04430 [cs.CV]

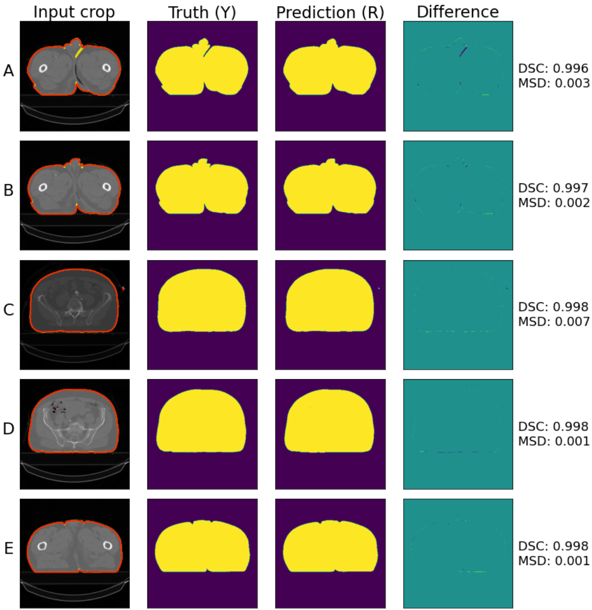
Modules - All happy models are alike...



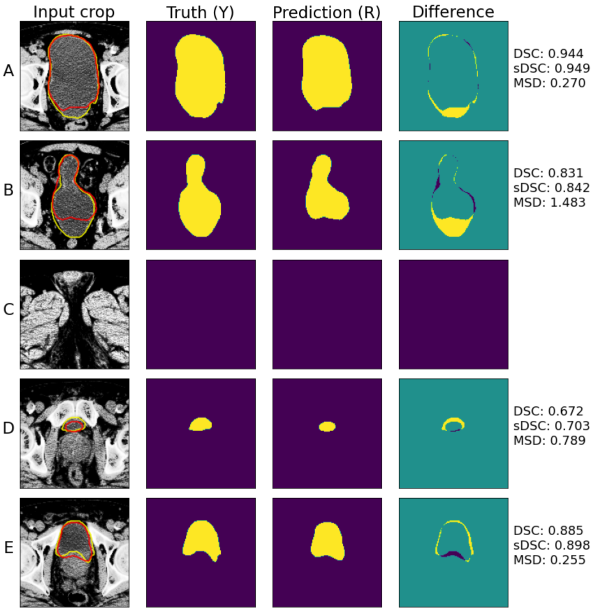
Deployment - DICOM networking



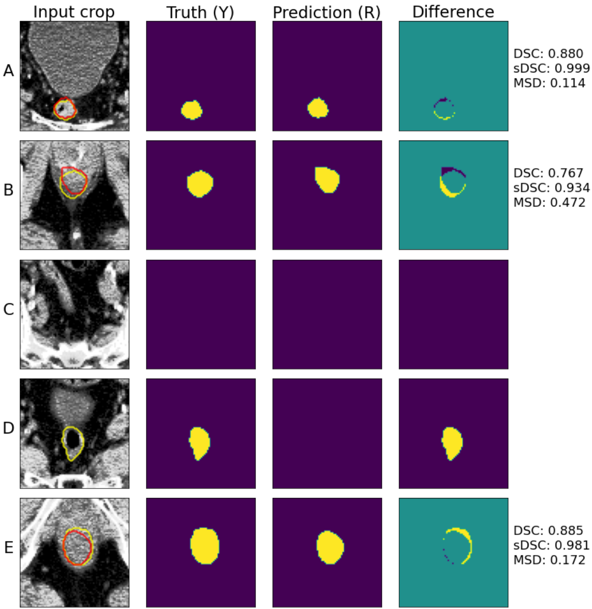
Pelvic imaging - Patient



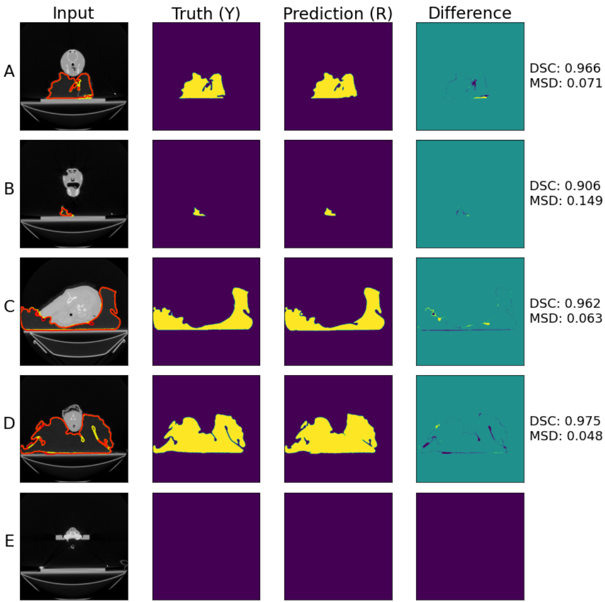
Pelvic imaging - Bladder



Pelvic imaging - Rectum



Canine imaging - Vacuum bag



Structure averaged metrics

	sDSC	DSC	MSD (mm)	Sensitivity	Specificity
Pelvic imaging					
Patient		0.998(0.001)	0.002(0.005)	0.99	0.99
Bladder (τ 1.46 mm)	0.9(0.2)	0.9(0.2)	1(3)	0.79	0.99
Rectum (τ 6.99 mm)	0.9(0.1)	0.7(0.1)	1(2)	0.62	0.99
Average		0.9(0.2)	0.6(2)	0.99	0.99
Canine imaging					
Vacbag		0.952(0.001)	0.2(0.3)	0.95	0.99

Cf. Expert IOV.[2]

- Clinically 'acceptable' bladder and rectum DSC ≥ 0.7
- Bladder: DSC 0.93 ± 0.03, MSD 0.9(0.3) mm.
- Rectum: DSC 0.81 ± 0.07, MSD 3(2) mm.

[2] Dale Roach et al. "Multi-observer contouring of male pelvic anatomy: Highly variable agreement across conventional and emerging structures of interest". In: *Journal of Medical Imaging and Radiation Oncology* 63.2 (2019), pp. 264–271. DOI: 10.1111/1754-9485.12844

[3] Stanislav Nikolov et al. *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*. 2018. arXiv: 1809.04430 [cs.CV]

Conclusion and future work

Pelvic imaging model:

- Patient contouring within tolerances
- Suspect more data will improve bladder and rectum segmentation.
- 3D architecture may identify gaseous rectal volumes.

Canine imaging model:

- Successfully deployed to clinic under a prototype warning
- Performance improvement of approximately 30 minutes per patient

Future

- Develop a soft surrogate for sDSC to optimise directly
- U-Net no longer S.O.T.A $\implies$ HR-Net.[16]

[16] Jingdong Wang et al. *Deep High-Resolution Representation Learning for Visual Recognition*. 2020. arXiv: 1908.07919 [cs.CV]

References I

Ford, Eric et al. "Strategies for effective physics plan and chart review in radiation therapy: Report of AAPM Task Group 275". In: *Medical Physics* 47.6 (2020), e236–e272. DOI: https://doi.org/10.1002/mp.14030.

Nikolov, Stanislav et al. *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*. 2018. arXiv: 1809.04430 [cs.CV].

Roach, Dale et al. "Multi-observer contouring of male pelvic anatomy: Highly variable agreement across conventional and emerging structures of interest". In: *Journal of Medical Imaging and Radiation Oncology* 63.2 (2019), pp. 264–271. DOI: 10.1111/1754-9485.12844.

Vinod, Shalini K et al. "A review of interventions to reduce inter-observer variability in volume delineation in radiation oncology". In: *Journal of Medical Imaging and Radiation Oncology* 60.3 (2016), pp. 393–406. DOI: 10.1111/1754-9485.12462.

Wang, Jingdong et al. *Deep High-Resolution Representation Learning for Visual Recognition*. 2020. arXiv: 1908.07919 [cs.CV].