

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

**Matthew Cooper**<sup>1</sup>  
Simon Biggs<sup>2</sup>  
Yu Sun<sup>1</sup>  
Matthew Sobolewski<sup>2</sup>

<sup>1</sup>Institute of Medical Physics, The University of Sydney

<sup>2</sup>Riverina Cancer Care Centre, Cancer Care Associates.

**Thesis:** [github.com/matthewdeancooper/masters\\_thesis](https://github.com/matthewdeancooper/masters_thesis)

**Video overview:** [docs.pymedphys.com/background/autocontouring](https://docs.pymedphys.com/background/autocontouring)



THE UNIVERSITY OF  
**SYDNEY**



**Riverina Cancer Care Centre**



# Contouring - Current limitations

## Variability

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

## Time constraints

- Atlas methods  $\implies$  significant correction times.<sup>3</sup>

## Deep learning potential

- Shown to reduce IOV and contouring time.<sup>3</sup>
- Significant improvement cf. atlas methods (time & accuracy).<sup>4</sup>

---

<sup>1</sup>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

<sup>2</sup>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>

<sup>3</sup>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

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

## Model 1: QA tool (prostate cancer).

- Compare contours to identify macro contouring errors
- Contoured patient, bladder, rectum volumes.

## Model 2: Canine vacuum bag.

- Automate time consuming aspect of canine RT.
- Previously, manual vacuum bag contouring  $\approx$  30 min.

## Goal: Performance similar to human experts.

- Performance metric (sDSC) that takes into account expert IOV.<sup>4</sup>
- Stronger correlation with correction time cf. DSC.<sup>5</sup>

---

<sup>3</sup>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

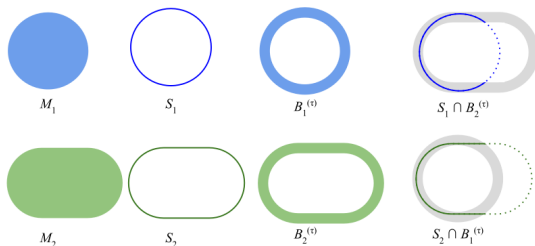
<sup>4</sup>Stanislav Nikolov et al. *Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy*. 2018. arXiv: 1809.04430 [cs.CV]

<sup>5</sup>Femke Vaassen et al. "Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy". In: *Physics and Imaging in Radiation Oncology* 13 (2020), 1–6. ISSN: 2405-6316. DOI: 10.1016/j.phro.2019.12.001

## Performance - Surface dice similarity coefficient (sDSC)

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

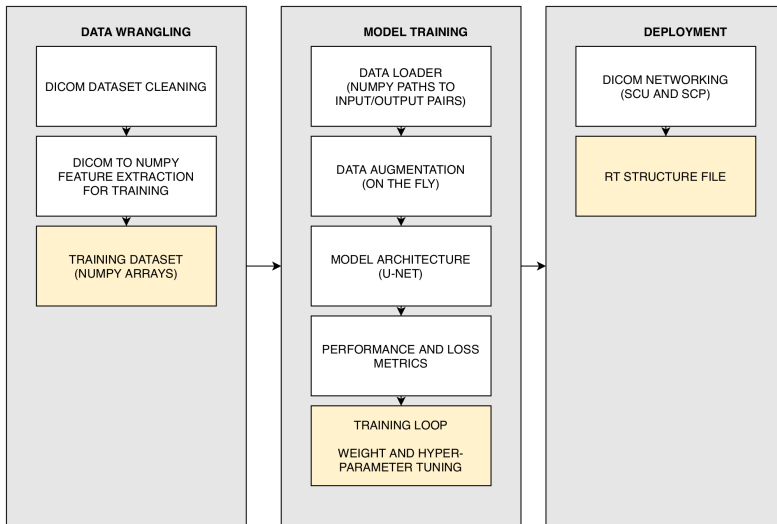
$$sDSC_{1,2}^{(\tau)} = \frac{|S_1 \cap B_2^{(\tau)}| + |S_2 \cap B_1^{(\tau)}|}{|S_1| + |S_2|}$$



**Figure:** DSC is a volumetric overlap score, sDSC is a surface overlap score - the percentage of surface contoured within an organ specific tolerance representative of expert IOV.<sup>3</sup>

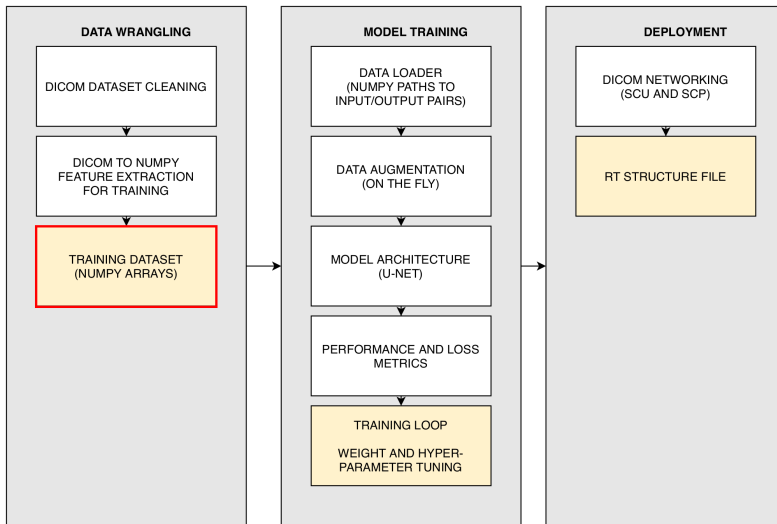
<sup>4</sup>Stanislav Nikolov et al. Deep learning to achieve clinically applicable segmentation of head and neck anatomy for radiotherapy. 2018. arXiv: 1809.04430 [cs.CV]

# All happy models are alike...



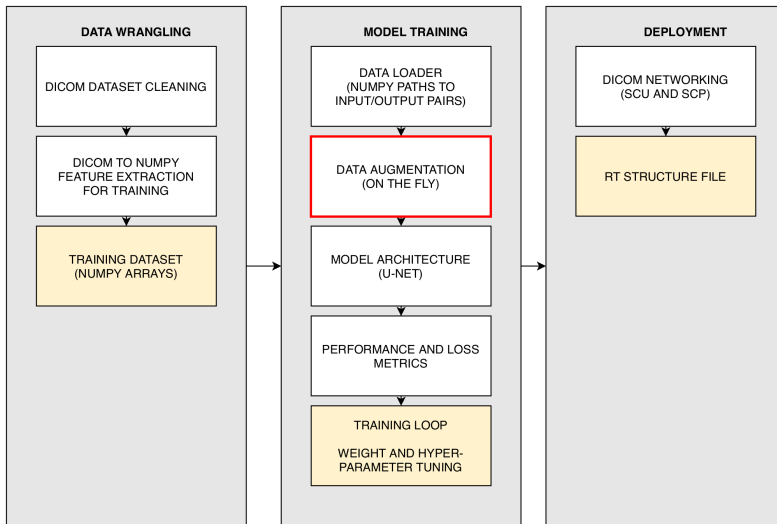
**Figure:** Modules required for end-to-end deep learning model deployment.

# All happy models are alike...



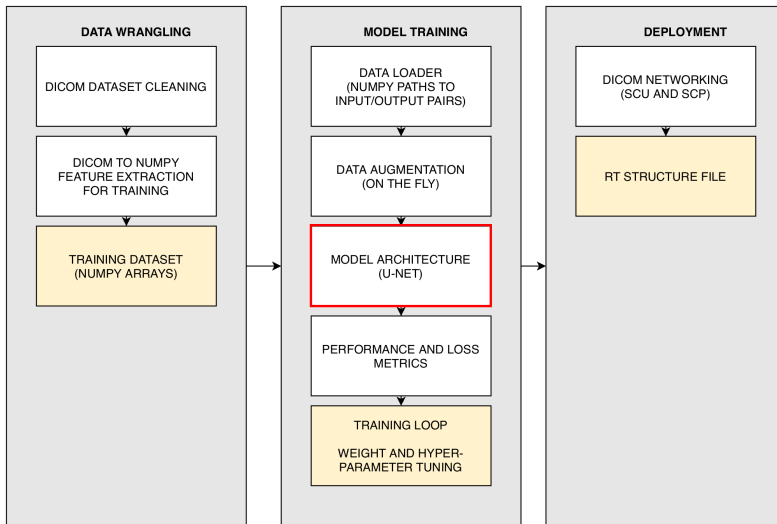
**Figure:** Modules required for end-to-end deep learning model deployment.

# All happy models are alike...



**Figure:** Modules required for end-to-end deep learning model deployment.

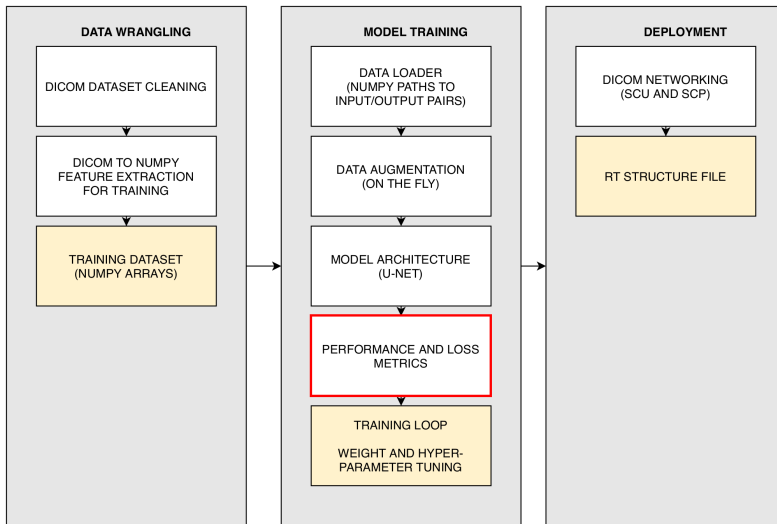
# All happy models are alike...



**Figure:** Modules required for end-to-end deep learning model deployment.

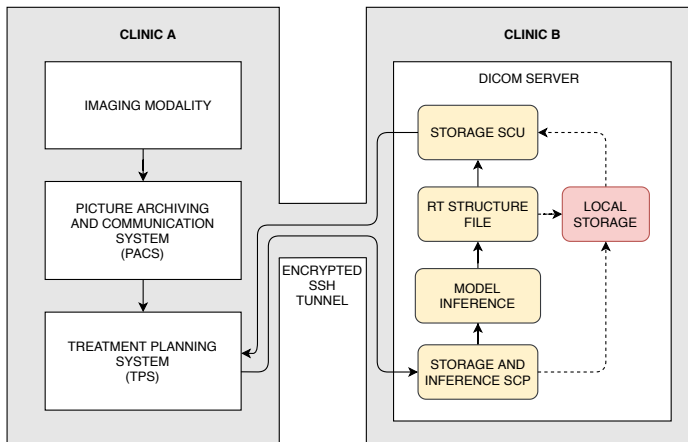


# All happy models are alike...

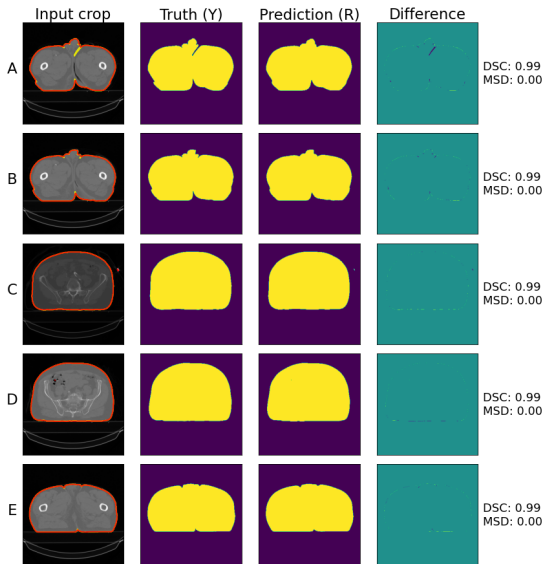


**Figure:** Modules required for end-to-end deep learning model deployment.

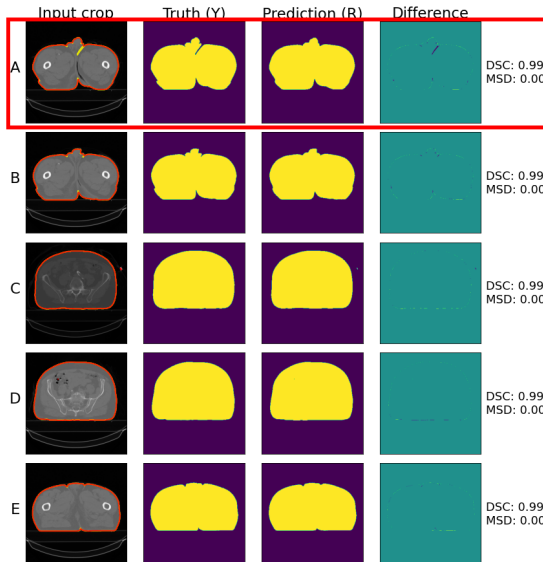
# Deployment - DICOM networking



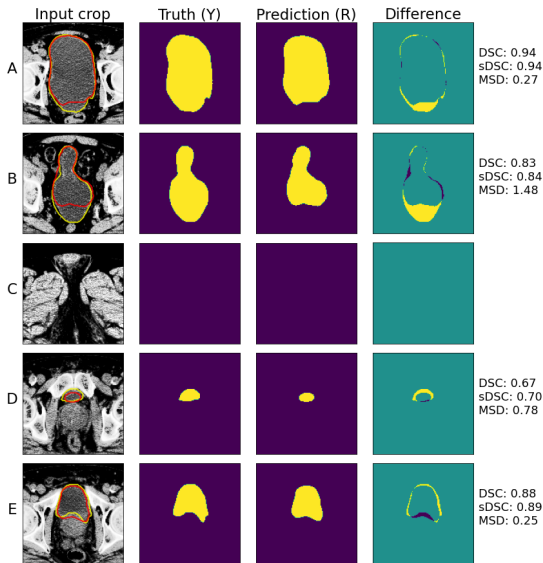
**Figure:** TPS exports to remote server via DICOM networking protocol.



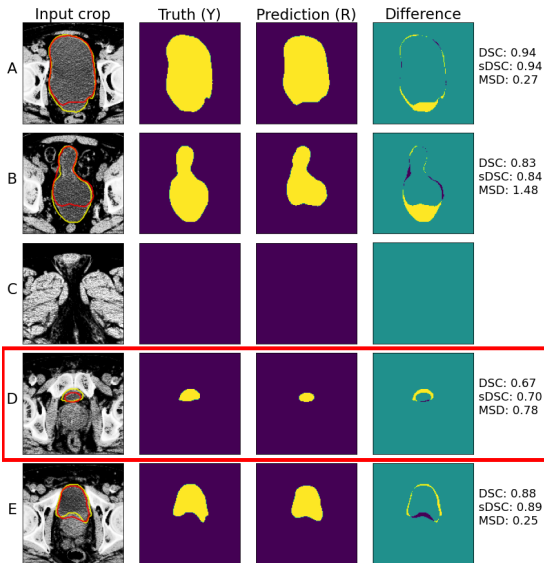
**Figure:** Representative output for **patient**. Truth contour (yellow), prediction contour (red). Metrics: Dice similarity coefficient (DSC), and mean surface distance (MSD) in mm.



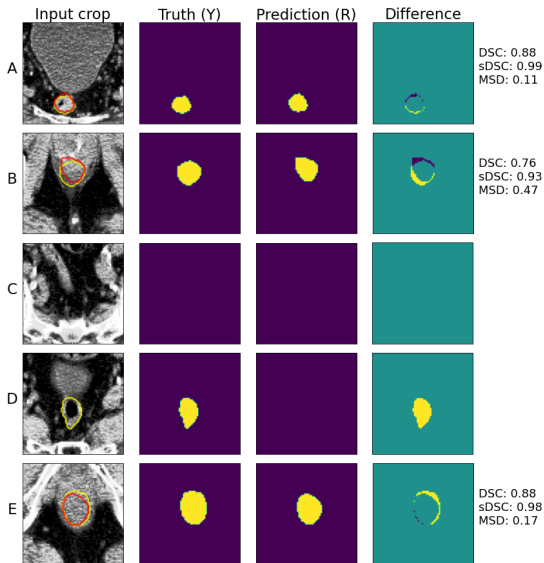
**Figure:** Representative output for **patient**. Truth contour (yellow), prediction contour (red). Metrics: Dice similarity coefficient (DSC), and mean surface distance (MSD) in mm.



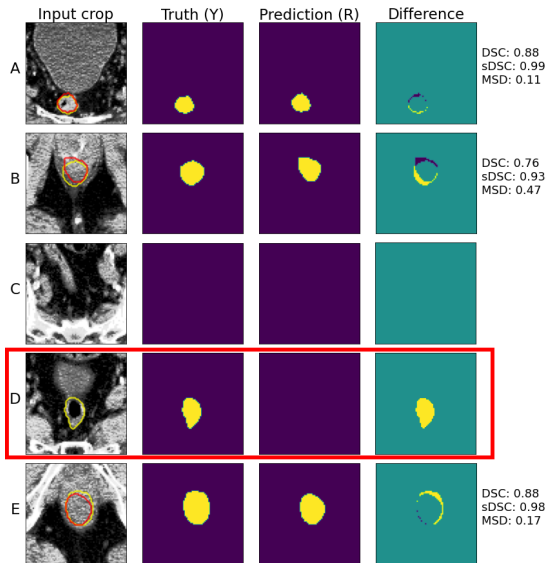
**Figure:** Representative output for **bladder**. Truth contour (yellow), prediction contour (red). Metrics: Dice coefficient (DSC), surface dice coefficient (sDSC), and mean surface distance (MSD) in mm.



**Figure:** Representative output for **bladder**. Truth contour (yellow), prediction contour (red). Metrics: Dice coefficient (DSC), surface dice coefficient (sDSC), and mean surface distance (MSD) in mm.

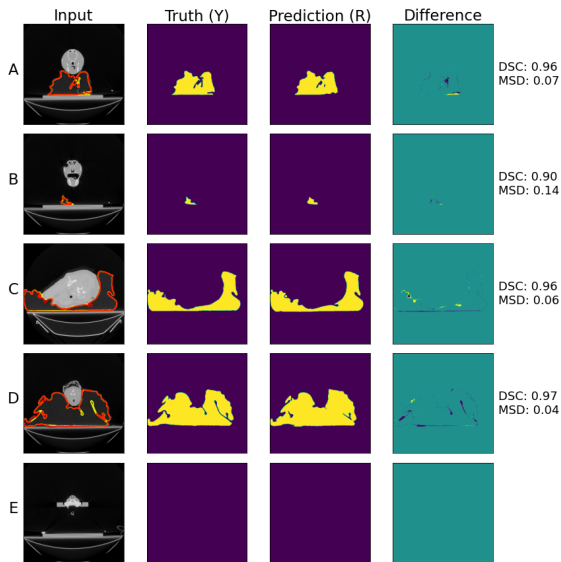


**Figure:** Representative output for **rectum**. Truth contour (yellow), prediction contour (red). Metrics: Dice coefficient (DSC), surface dice coefficient (sDSC), and mean surface distance (MSD) in mm.

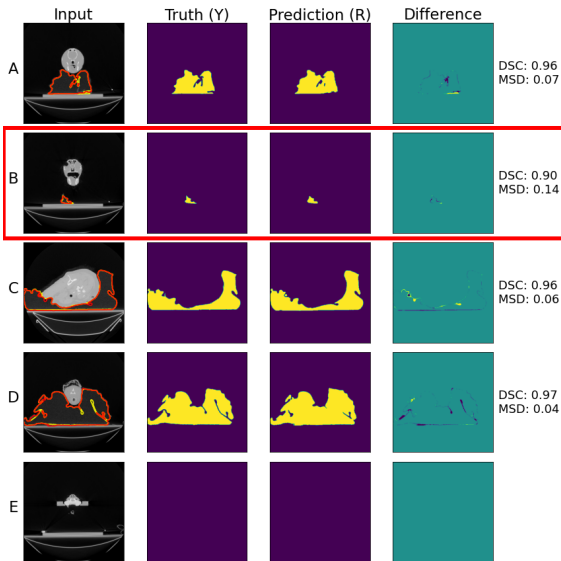


**Figure:** Representative output for **rectum**. Truth contour (yellow), prediction contour (red). Metrics: Dice coefficient (DSC), surface dice coefficient (sDSC), and mean surface distance (MSD) in mm.





**Figure:** Representative output for **vacuum bag**. Truth contour (yellow), prediction contour (red). Metrics: Dice similarity coefficient (DSC), and mean surface distance (MSD) in mm.



**Figure:** Representative output for **vacuum bag**. Truth contour (yellow), prediction contour (red). Metrics: Dice similarity coefficient (DSC), and mean surface distance (MSD) in mm.

**Table:** Organ specific evaluation on independent dataset.

	sDSC	DSC	MSD (mm)
<b>Pelvic imaging<sup>a</sup></b>			
Patient		0.99(1)	0.00(5)
Bladder ( $\tau$ 1.46 mm)	0.9(2)	0.9(2)	1(3)
Rectum ( $\tau$ 6.99 mm)	0.9(1)	0.7(1)	1(2)
Average		0.9(2)	0.6(2)
<b>Canine imaging</b>			
Vacbag		0.952(1)	0.2(3)

<sup>a</sup> Organ specific tolerance  $\tau$  = MSD<sub>95</sub> (Top 95% expert performance).

## Cf. expert IOV.<sup>1</sup>

- Clinically 'acceptable' bladder and rectum DSC  $\geq 0.7$
- Bladder: DSC 0.93(3), MSD 0.9(3) mm.
- Rectum: DSC 0.81(7), MSD 3(2) mm.

<sup>1</sup>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

### **Pelvic imaging model:**

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

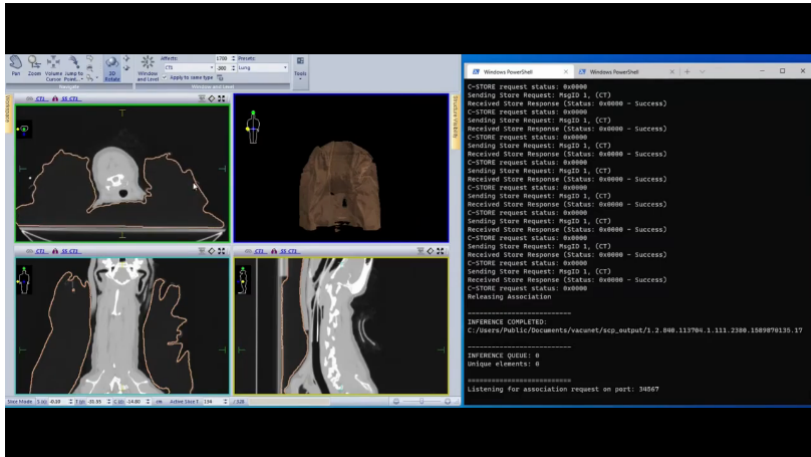
### **Canine imaging model:**

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

### **Future**

- Develop a continuous valued surrogate for sDSC to optimise directly.

## Questions



**Video overview:** [docs.pymedphys.com/background/autocontouring](https://docs.pymedphys.com/background/autocontouring)

# 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.
- Vaassen, Femke et al. "Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy". In: *Physics and Imaging in Radiation Oncology* 13 (2020), 1–6. ISSN: 2405-6316. DOI: 10.1016/j.phro.2019.12.001.
- 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.