

Supplementary Material – Continual atlas-based segmentation of prostate MRI

Amin Ranem¹ Camila González² Daniel Pinto dos Santos³ Andreas M. Bucher⁴
 Ahmed E. Othman⁵ Anirban Mukhopadhyay¹

¹ Technical University of Darmstadt, ² Stanford University, ³ University of Cologne,
⁴ University of Frankfurt, ⁵ University Medical Center Mainz
 amin.ranem@gris.tu-darmstadt.de

1. End-to-end vs. atlas-based segmentation

Table 1 provides the Dice scores with standard deviation for every trained baseline evaluated across all tasks which were used to create Figure 9 from the main manuscript.

Baselines		Dice $\uparrow \pm \sigma \downarrow$ [%]						
		\mathcal{T}_1	\mathcal{T}_2	\mathcal{T}_3	\mathcal{T}_4	\mathcal{T}_5	\mathcal{T}_6	\mathcal{T}_7
U-Net	\mathcal{T}_1	75.38 \pm 5.96	37.82 \pm 15.47	17.04 \pm 4.83	42.02 \pm 23.75	6.66 \pm 6.86	50.86 \pm 1.78	76.14 \pm 5.66
	\mathcal{T}_2	36.36 \pm 19.55	70.25 \pm 6.30	25.81 \pm 9.26	52.21 \pm 15.53	50.30 \pm 8.01	36.78 \pm 7.98	37.01 \pm 10.70
	\mathcal{T}_3	21.12 \pm 17.64	53.55 \pm 6.23	78.50 \pm 5.04	55.44 \pm 2.65	37.54 \pm 6.98	50.12 \pm 5.39	18.94 \pm 9.90
	\mathcal{T}_4	46.48 \pm 17.13	65.69 \pm 6.29	60.23 \pm 2.08	66.44 \pm 11.89	49.19 \pm 13.37	65.73 \pm 10.23	44.92 \pm 11.52
	\mathcal{T}_5	9.61 \pm 7.19	48.44 \pm 8.02	29.82 \pm 5.24	40.26 \pm 18.72	59.55 \pm 4.97	38.21 \pm 4.85	7.87 \pm 7.35
	\mathcal{T}_6	22.05 \pm 19.25	54.02 \pm 10.05	58.41 \pm 4.21	59.37 \pm 4.41	52.10 \pm 1.99	77.42 \pm 3.27	22.13 \pm 17.11
	\mathcal{T}_7	73.52 \pm 7.57	37.38 \pm 9.17	20.28 \pm 3.27	32.83 \pm 22.86	10.75 \pm 7.59	45.11 \pm 3.84	76.12 \pm 3.83
Atlas Replay	\mathcal{T}_1	70.50 \pm 4.63	56.02 \pm 4.78	72.87 \pm 6.30	73.79 \pm 1.55	34.81 \pm 22.70	63.69 \pm 9.80	73.02 \pm 3.63
	\mathcal{T}_2	44.27 \pm 18.47	72.30 \pm 6.05	73.92 \pm 3.50	75.16 \pm 0.91	53.73 \pm 13.33	66.62 \pm 7.50	43.99 \pm 9.74
	\mathcal{T}_3	34.30 \pm 7.69	34.69 \pm 6.41	73.11 \pm 4.42	71.33 \pm 2.28	23.81 \pm 11.07	44.28 \pm 8.05	31.97 \pm 7.77
	\mathcal{T}_4	23.20 \pm 9.93	23.61 \pm 7.13	63.36 \pm 7.56	78.40 \pm 0.72	20.57 \pm 7.71	37.11 \pm 4.82	19.55 \pm 5.34
	\mathcal{T}_5	32.79 \pm 15.13	53.30 \pm 2.42	72.87 \pm 5.67	74.55 \pm 2.59	60.31 \pm 3.02	58.50 \pm 9.19	30.31 \pm 8.77
	\mathcal{T}_6	34.66 \pm 15.19	44.05 \pm 6.22	68.91 \pm 3.17	71.48 \pm 4.08	39.16 \pm 15.10	64.97 \pm 13.40	33.47 \pm 8.32
	\mathcal{T}_7	69.02 \pm 5.74	53.63 \pm 8.64	74.94 \pm 5.43	73.69 \pm 1.35	36.15 \pm 16.09	59.41 \pm 10.48	67.73 \pm 5.20

Table 1. Results for all baseline networks trained on every task individually and evaluated across all tasks; Bold values indicate the performance of the baseline on the validation set of the task it has been trained on.