Supplementary material: Progressively refined joint registration-segmentation (ProRSeg) of gastrointestinal organs at risk: Application to MRI and cone-beam CT

Jue Jiang[1,†], **Jun Hong**[1,⋆], **Kathryn Tringale**[2], **Marsha Reyngold**[2], **Christopher Crane**[2], **Neelam Tyagi**[1], **Harini Veeraraghavan**[1,†]

[1] Department of Medical Physics, Memorial Sloan Kettering Cancer Center, 1250 First Ave, New York, NY 10065, USA

[2] Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, 1250 First Ave, New York, NY 10065, USA

[†] These authors contributed equally

⋆: work performed when author was at MSKCC

Supplementary Table 1. Recurrent registration network (RRN) architecture used for registration. We use the following abbreviation for ease of presentation: N=number of features; K=Kernel size; S=Stride size; CLSTM=convolutional Long short-term memory; VecInc=diffeomorphic intergration layer.

Layers	Registration net G	Concatenation
1	CLSTM-(N16,K3,S2) LeakyReLu	1
2	CLSTM-(N16,K3,S2) LeakyReLu	2
3	CLSTM-(N16,K3,S2) LeakyReLu	3
5	CLSTM-(N16,K3,S2) LeakyReLu	4
6	CLSTM-(N16,K3,S1)	-
7	CONV-(N32,K3,S1) LeakyReLu	4
8	CONV-(N32,K3,S1) LeakyReLu	3
9	CONV-(N32,K3,S1) LeakyReLu	2
10	CONV-(N32,K3,S1) LeakyReLu	1
11	CONV-(N16,K3,S1) LeakyReLu	-
12	CONV-(N3,K3,S1) LeakyReLu	-
13	VecInC	-

@articleliang2020auto, title=Auto-segmentation of pancreatic tumor in multi-parametric MRI using deep convolutional neural networks, author=Liang, Ying and Schott, Diane and Zhang, Ying and Wang, Zhiwu and Nasief, Haidy and Paulson, Eric and Hall, William and Knechtges, Paul and Erickson, Beth and Li, X Allen, journal=Radiotherapy and Oncology, volume=145, pages=193–200, year=2020, publisher=Elsevier

Supplementary Table 2. RSN architecture used for segmentation. We use the following abbreviation for ease of presentation: N=number of features; K=Kernel size; S=Stride size;CLSTM=convolutional Long short-term memory;

Layers	Unet	Concatenation
1	CONV-(N32,K3,S1), ReLu	-
2	CLSTM-(N32,K3,S1), ReLu	1
3	Max-Pooling (S2)	-
4	CONV-(N64,K3,S1), ReLu	-
5	CLSTM-(N64,K3,S1), ReLu	2
6	Max-Pooling (S2)	-
7	CONV-(N128,K3,S1), ReLu	-
8	CLSTM-(N128,K3,S1), ReLu	3
9	Max-Pooling (S2)	-
10	CONV-(N256,K3,S1), ReLu	-
11	CLSTM-(N256,K3,S1), ReLu	4
12	Max-Pooling (S2)	-
13	CONV-(N512,K3,S1), ReLu	-
14	CLSTM-(N512,K3,S1), ReLu	-
15	UP-Pooling (S2)	4
16	CONV-(256,K3,S1), ReLu	-
17	CONV-(N256,K3,S1), ReLu	-
18	UP-Pooling (S2)	3
19	CONV-(N128,K3,S1), ReLu	-
20	CONV-(N128,K3,S1), ReLu	-
21	UP-Pooling (S2)	2
22	CONV-(N64,K3,S1), ReLu	-
23	CONV-(N64,K3,S1), ReLu	-
24	UP-Pooling (S2)	1
25	CONV-(N32,K3,S1), ReLu	-
26	CONV-(N2,K1,S1), Softmax	-

Supplementary Table 3. P-values measuring differences in segmentation accuracy produced by analyzed methods with respect to ProRSeg. DSC metric was used for comparison using two-sided, paired Wilcoxon signed rank test at 95% signficance level. LG bowel: Large bowel, SM bowel: small bowel, Sto-Duo: stomach-duodenum. †

Method	DSC				HD95			
	Liver	LG Bowel	SM Bowel	Sto-Duo	Liver	LG Bowel	SM Bowel	Sto-Duo
SyN	8.10E-10	8.10E-10	1.10E-07	2.10E-08	3.60E-05	1.80E-07	4.40E-09	6.10E-06
Voxmorph	1.20E-04	9.00E-08	7.40E-07	7.00E-09	3.30E-05	8.70E-10	3.90E-07	3.90E-05
Unet3D	4.40E-07	0.00018	4.50E-08	3.50E-08	1.80E-05	8.70E-09	1.80E-06	1.40E-05
nnUnet	0.006323	0.00118	0.00017	4.90E-08	4.90E-05	9.80E-06	0.00019	9.30E-06
UResNet	0.00012	9.00E-08	7.40E-07	0.00123	0.00798	2.90E-08	2.70E-05	3.90E-05

Supplementary Table 4. Organ-specific coefficient of variation computed using DSC metric (CV_{DSC} per patient. CV_{DSC} was computed using the ratio of mean DSC to standard deviation of DSC was computed per patient based on all possible MR-MR combinations from the different treatment fractions to evaluate variability due to anatomic prior.

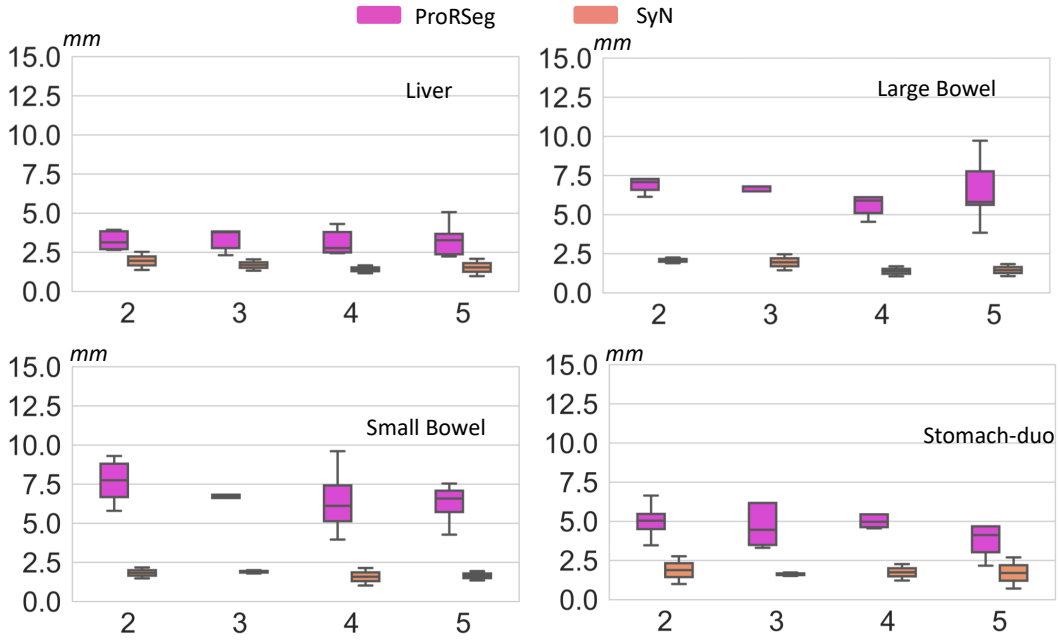
Method	CV_{DSC} %				CV_{DSC}^{Mean} %
	Liver	LG Bowel	SM Bowel	Sto-Duo	
P1	0.137	0.977	5.36	1.29	1.94
P2	0.681	0.885	2.39	0.654	1.15
P3	2.10	1.82	4.60	2.81	2.83
P4	0.425	0.165	0.719	6.89	2.05
P5	0.277	9.79	5.33	11.02	6.60
P6	0.164	2.65	3.89	3.05	2.44
P7	0.526	5.03	5.05	6.59	4.29
P8	0.669	0.524	6.54	5.03	3.19
P9	0.452	0.238	4.49	2.69	1.97
P10	0.452	2.66	3.73	7.49	3.58

Supplementary Table 5. Mean displacement measured in all three directions using SyN, Voxelmorph, and ProRSeg of the individual organs. Stomach-Duo: Stomach duodenum

Metric	Liver (mm)			Large bowel (mm)			Small bowel (mm)			Stomach-Duo (mm)		
	x	y	z	x	y	z	x	y	z	x	y	z
SyN	0.12	0.13	0.17	0.12	0.17	0.25	0.15	0.12	0.15	0.11	0.11	0.16
Voxmorph	0.91	0.80	1.10	0.91	0.89	1.14	0.76	0.83	0.97	0.75	0.74	0.85
ProRSeg	2.60	2.71	1.88	2.60	3.33	3.81	4.92	3.21	4.06	3.54	2.95	3.48

Supplementary Table 6. Accumulated dose and toxicity for analyzed subset of 5 patients

Metric	Stomach-Duo		Small bowel		Large Bowel		Acute toxicity	Late toxicity
	$0.035cm^3$	$5cm^3$	$0.035cm^3$	$5cm^3$	$0.035cm^3$	$5cm^3$		
P1	32.92	22.79	33.74	22.39	24.70	20.26	-	-
P2	41.33	28.93	29.14	20.90	27.03	22.00	Grade I abdominal pain	Grade I abdominal pain
P3	33.03	24.41	40.38	31.07	17.41	15.07	-	-
P4	40.23	27.82	34.39	21.06	23.76	19.50	-	-
P5	26.04	20.89	25.45	21.41	26.14	20.48	-	-



Supplementary Figure 1. Mean displacement computed for each organ at different treatment fractions shown for ProRSeg and SyN methods.

Supplementary Table 7. Signficance test results comparing ProRSeg++ with other baseline methods with Bonferroni correction.

Method	DSC				HD95 mm			
	Liver	LG Bowel	SM Bowel	Sto-Duo	Liver	LG Bowel	SM Bowel	Sto-Duo
SyN	0.0004	0.01	0.0030	0.045	0.048	0.12	0.028	0.45
Voxelmorph	0.0003	0.012	0.014	0.024	0.085	0.004	0.183	0.45
Unet3D	0.0002	0.044	0.018	0.0001	0.075	0.012	0.015	4e-4
nnUnet	0.0011	0.044	0.014	0.024	0.060	0.086	0.016	0.032
UResNet	0.00000001	0.010	0.0046	1e-4	0.075	0.12	0.018	0.3

Supplementary Table 8. Signficance test results comparing ProRSeg with other baseline methods with Bonferroni correction on CBCT dataset.

Method	DSC		HD95 mm	
	LG Bowel	SM Bowel	LG Bowel	SM Bowel
SyN	5.80E-10	1.90E-07	2.80E-05	1.00E-06
Voxelmorph	4.40E-07	7.60E-05	1.89E-02	0.03116
nnUnet	5.10E-13	2.00E-10	7.90E-12	4.10E-11
UResNet	1.20E-06	8.70E-04	1.42E-01	1.85E-01