Supplementary Material for 'I^{2}Net: Exploiting Misaligned Contexts Orthogonally with Implicit-Parameterized Implicit Functions for Medical Image Segmentation'

Anonymous[1]

Anonymous anonymous@anonymous

Table 1. The impact of K on model performance on Glas test and Synapse val. When $K > 4$, the performance on Synapse begins to decline, whose reason may be that Synapse data have fewer discrimination patterns than Glas data.

K	Glas				Synapse											
	#Params	GFLOPs	DSC(%)	HD95(mm)	#Params	GFLOPs	DSC(%)	HD95(mm)	Aorta	Gallbladder	Kidney(L)	Kidney(R)	Liver	Pancreas	Spleen	Stomach
2	11.35M	11.61	93.47	7.15	17.40M	32.07	78.94	29.08	88.47	65.70	84.60	80.84	94.22	56.98	86.45	74.26
3	11.40M	12.30	93.91	4.18	17.43M	32.53	**79.59**	**25.99**	88.56	66.55	83.99	**81.17**	**94.24**	**59.10**	87.17	**75.91**
4	11.44M	12.99	93.84	**3.55**	17.47M	32.99	78.99	28.70	**89.44**	**69.89**	83.98	79.90	92.98	55.78	86.71	73.27
5	11.48M	13.68	93.95	3.97	17.51M	33.45	78.87	28.62	88.77	66.54	**85.05**	81.15	93.69	55.05	86.35	74.38
6	11.53M	14.37	**94.00**	6.27	17.54M	33.91	78.78	33.55	88.63	69.95	83.39	79.59	93.44	53.62	**87.62**	73.99

Table 2. Comparison with different aligning methods on Cityscapes `val`. 'SS' and 'MS' denote single-scale and multi-scale inference respectively.

Method	#Params	GFLOPs	mIoU(%)(SS/MS)	
Bilinear Up-sampling	27.7M	183.4	74.47	76.46
Nearest Neighbor	27.7M	183.4	75.83	77.03
Deconvolution	29.5M	304.4	71.03	72.70
Semantic-FPN	31.0M	219.1	78.00	79.14
CARAFE++	30.3M	211.5	77.50	78.62
SFNet	42.9M	327.3	79.14	79.67
AlignSeg	49.7M	348.6	78.72	79.92
IFA	27.8M	186.9	78.04	79.35
I^{2}Net ($K = 4$)	30.3M	280.5	**79.67**	**80.26**

Table 3. Comparison with state-of-the-art methods on Cityscapes `val`.

Method	Backbone	mIoU(%) (SS/MS)		#Params	GFLOPs
PSPNet	DResNet-101	78.34	79.74	68M	1104
PSANet	DResNet-101	79.69	80.89	89M	1206
DeepLabV3+	DResNet-101	79.46	80.50	63M	2032
CCNet	DResNet-101	79.45	80.66	70M	1190
DANet	DResNet-101	79.88	81.47	70M	1336
ANNet	DResNet-101	79.32	80.94	67M	1121
GCNet	DResNet-101	79.18	80.71	-	354
Mask2Former	DResNet-101	80.10	81.90	63M	-
UPerNet	ResNet-101	79.03	80.77	86M	1029
PointRend	ResNet-101	78.30	79.97	79M	300
SFNet	ResNet-101	79.80	81.80	55M	460
IFA	ResNet-101	79.92	81.20	64M	281
OCRNet	HRNetV2p-W48	80.58	81.79	71M	1297
SETR	ViT-L	79.21	81.02	318M	2352
Segmenter	ViT-L	79.10	81.30	337M	-
DDRNet	DDRNet23	79.50	-	20M	143
I^{2}Net ($K = 4$)	ResNet-101	**81.07**	**82.03**	66M	417

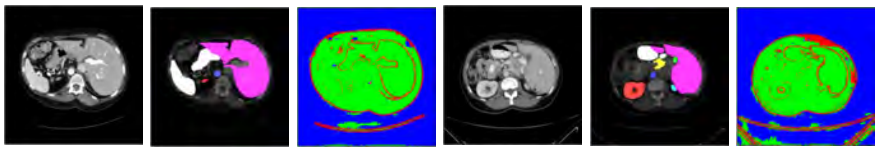

Fig. 1. Visualization of 'hard' gate maps from I^{2}Net ($K = 3$) trained on Synapse. We observe that the red PL tends to capture patterns along the edges, the green learns patterns within the body, and the blue handles regions outside the body.