Spatial and Planar Consistency for Semi-Supervised Volumetric Medical Image Segmentation

Yanfeng Zhou[1,2], Yiming Huang[1,2], Ge Yang[1,2]
[1]School of Artificial Intelligence, University of Chinese Academy of Sciences
[2]Institute of Automation, Chinese Academy of Sciences

1. Background

Semantic Segmentation *(Fundamental Task)* → Fully-Supervised *(Large-Scale Labeled Images)* → Semi-Supervised *(Labeled + Unlabeled)*

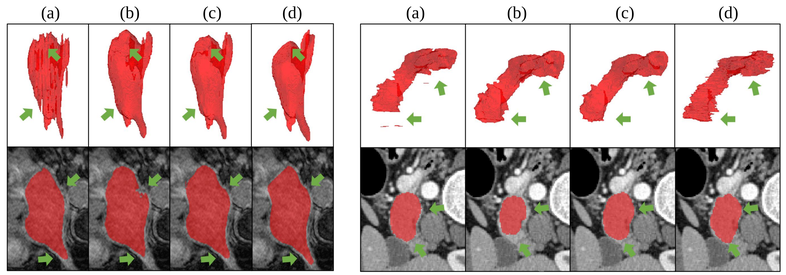
Semi-Supervised Volumetric Segmentation *(DTC, MC-Net, CC-Net, etc.)* → Consistency Regularization *(Noise, Distance Map, Multiple Outputs, etc.)*

Few Studies Consider Exploiting Potential Differences of 2D and 3D Models for Consistency Regularization

We Propose A Spatial and Planar Complementary Consistency (SPC) Strategy for Volumetric Medical Image Segmentation

2. Motivation

2D Network (↓ *Spatial Structure*, ↑ *Planar Detail*) + 3D Network (↑ *Spatial Structure*, ↓ *Planar Detail*) = SPC (2D + 3D) (↑ *Spatial Structure*, ↑ *Planar Detail*)

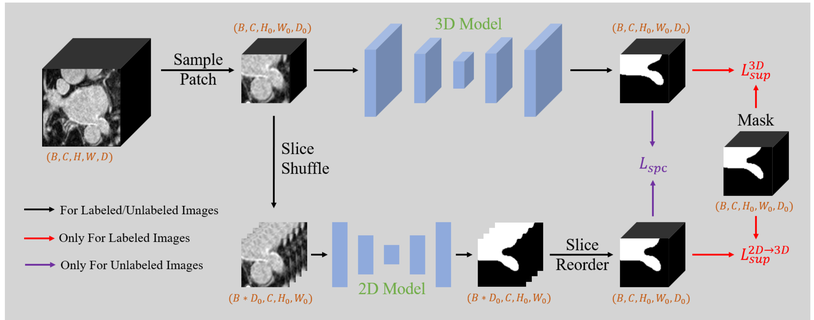


LA (right) and P-CT (left). (a) 2D. (b) 3D. (c) SPC (Combining 2D and 3D). (d) Ground truth.

3. Contribution

- A spatial and planar consistency strategy
- Focus on both spatial structure and planar detail
- State-of-the-art on LA and P-CT

4. Method



Architecture

3D Branch → Spatial Structure
2D Branch → Planar Detail
3D and 2D Model Consistency

Loss

$$L_{total} = L_{sup} + \lambda L_{spc} \qquad \lambda = \lambda_{max} * \frac{epoch}{\max\_epoch}$$

$$L_{sup} = L_{sup}^{3D}(p_{3D}, y_{3D}) + L_{sup}^{2D\to 3D}(p_{2D\to 3D}, y_{3D})$$

$$L_{spc} = L_{spc}^{3D}(p_{3D}, \hat{p}_{2D\to 3D}) + L_{spc}^{2D\to 3D}(p_{2D\to 3D}, \hat{p}_{3D})$$

5. Experiments

Comparison with State-of-the-arts

Dataset	Method	# Labeled	# Unlabeled	Dice ↑	Jaccard ↑	ASD ↓	95HD ↓
LA	MT	16	64	88.23	79.29	2.73	10.64
	SASSNet	16	64	89.17	80.69	2.86	8.57
	DTC	16	64	89.43	81.00	2.12	7.39
	MC-Net	16	64	90.12	82.12	1.99	8.07
	MC-Net+	16	64	91.05	83.64	1.69	5.81
	TraCoCo	16	64	90.94	83.47	1.79	**5.49**
	CC-Net	16	64	**91.27**	**84.02**	**1.54**	5.75
	SPC (Ours)	16	64	**92.52**	**86.08**	**1.40**	**4.59**
P-CT	MT	12	50	76.79	62.33	2.94	10.97
	EM	12	50	75.98	61.26	3.77	12.80
	UAMT	12	50	77.14	62.79	3.85	14.91
	SASSNet	12	50	77.81	63.67	3.06	9.15
	DTC	12	50	78.25	64.26	2.14	7.17
	MC-Net	12	50	77.71	63.54	2.74	9.02
	MC-Net+	12	50	**78.87**	**65.11**	**1.89**	8.15
	SPC (Ours)	12	50	**79.82**	**66.42**	**1.83**	**6.68**

Ablation Studies

Different λ_{max}

Dataset	λ_{max}	Dice ↑	Jaccard ↑	ASD ↓	95HD ↓
LA	1	91.49	84.32	1.63	5.35
	3	91.89	85.00	1.54	5.16
	5	**92.52**	**86.08**	**1.40**	**4.59**
	7	92.00	85.18	1.51	5.13
P-CT	0.5	76.53	61.99	2.79	11.63
	1	**79.82**	**66.42**	**1.83**	**6.68**
	3	79.46	65.92	1.85	7.41

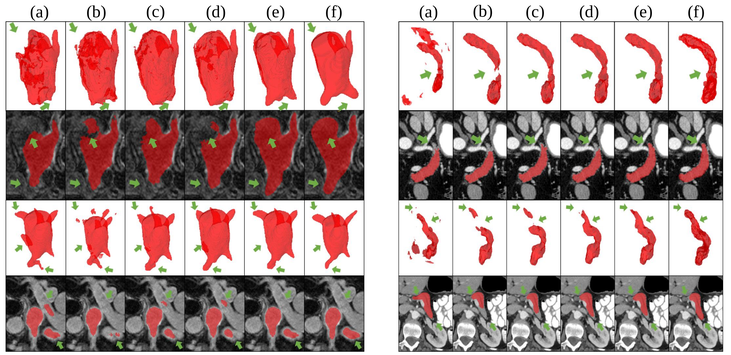
Loss CE vs Dice

Dataset	L_{sup}	L_{spc}	Dice ↑	Jaccard ↑	ASD ↓	95HD ↓
LA	Dice	Dice	92.10	85.36	1.45	**4.52**
	Dice	CE	91.79	84.82	1.52	4.89
	CE	Dice	91.41	84.18	1.66	5.71
	CE	CE	**92.52**	**86.08**	**1.40**	4.59
P-CT	Dice	Dice	79.61	66.13	**1.81**	7.13
	Dice	CE	79.09	65.41	1.95	6.71
	CE	Dice	78.18	64.18	1.97	7.58
	CE	CE	**79.82**	**66.42**	1.83	**6.68**

Effectiveness of Combining 2D and 3D

Dataset	2D	3D	Dice ↑	Jaccard ↑	ASD ↓	95HD ↓
LA	✓		85.31	74.39	2.36	8.17
		✓	89.79	81.48	2.03	6.86
	✓	✓	**92.52**	**86.08**	**1.40**	**4.59**
P-CT	✓		75.30	60.39	3.45	17.33
		✓	78.70	64.88	2.03	8.76
	✓	✓	**79.82**	**66.42**	**1.83**	**6.68**

Qualitative Results



LA (right) and P-CT (left). (a) MT. (b) DTC. (c) MC-Net. (d)MC-Net+. (e) SPC (Ours). (f) Ground truth.

Acknowledgements. *This work was supported in part by the Natural Science Foundation of China (grants 31971289, 91954201) and the Strategic Priority Research Program of the Chinese Academy of Sciences (grant XDB37040402).*