CPNet: 3D Semantic Relation and Geometry Context Prior Network for Multi-organ Segmentation

Supplementary Document

Yuzhu Ji[1,*], Mingshan Sun[2,], Yiqun Zhang[1] and Haijun Zhang[3]**

[1]School of Computer Science and Technology, Guangdong University of Technology, Guangzhou, China
[2]CVTE Research, Guangzhou, China
[3]Harbin Institute of Technology, Shenzhen, China

A Normal Vectors Reversion

As mentioned in Section 3.4, since there is no access to information about the global geometric structure, the covariance analysis algorithm presented in Open3D [1] may yield two normal vectors with opposite directions as candidates, both of them are correct. Figure A.1 illustrates such cases with yellow and green arrows. To extract normal vectors, we reverse the confusing normal vectors, *i.e.*, yellow arrow cases, to make sure that the estimated normal vectors uniformly point to the outside of organs. To achieve this, we identify the semantic labels of a surface point and its nearest neighbour that the normal vector is towards to determine whether or not the normal vector on that point should reversed. Concretely, if the semantic labels are the same, *i.e.*, the yellow arrow cases, it indicates the direction of a normal vector points to the inside of the organ and the normal vector should be reversed.

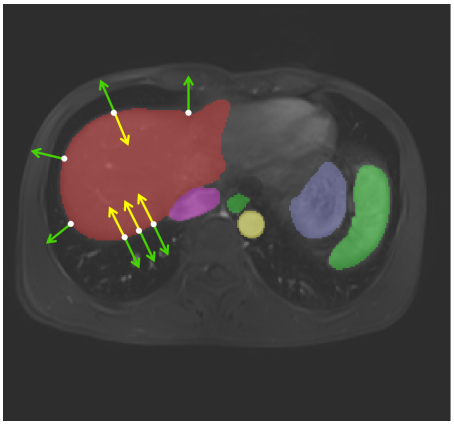

Figure A.1. Reversion of the direction of the normal vectors . Yellow arrows mean confusing normal vectors, and green arrows mean reversed normal vectors or normal vectors with correct directions.

* Corresponding Author. Email: yuzhu.ji@gdut.edu.cn
** Corresponding Author. Email: mingshine.sun@gmail.com

B Relation-balanced Cross-entropy Loss

To balance the loss weights for predicting different semantic relation cases, we categorize semantic relation priors into three types based on learning difficulties. Figure B.1 illustrates that three cases w.r.t. different types of difficulties. The first type is the simplest, where the surroundings are all background voxels. It means the voxel is not in any of the organs. The second type is slightly more difficult than the first type. It indicates the surrounding voxels are in the same category and located within an organ. The third type is the most challenging one, in which the central voxel is surrounded by voxels with multiple semantic categories. It shows the voxel is on the contact margins of multiple organs. Those complex and hard cases of semantic relation are difficult to predict. Therefore, during training, we empirically balance the `RCEloss` with loss weights 0.1, 0.3 and 0.6, respectively, based on the above analysis of three types of difficulties for semantic relation prediction.

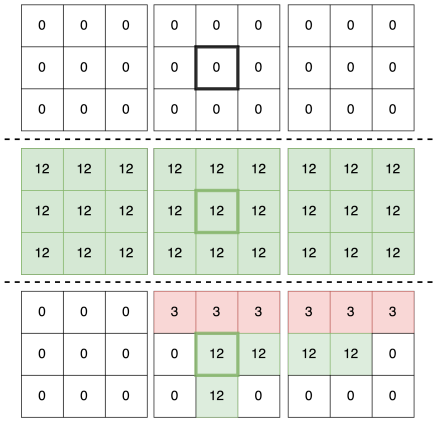

Figure B.1. Three types of semantic relation priors. The difficulty increases gradually from top to bottom to learn the surrounding semantics and capture the semantic relations of organs.

C Extra Experimental Results

Evaluation of other metrics. We calculate the Jaccard Index (JAC) and 95% Hausdorff distance (HD95) to further evaluate our

Table C.1. Organ-wise quantitative results in terms of Jaccard Index and HD95 on Abdomen CT dataset. The best result for each class is bolded.

	Jaccard Index (%)↑						HD95 (mm)↓					
	nnU-Net	SegResNet	UNETR	SwinUNETR	U-Mamba	Ours	nnUnet	SegResNet	UNETR	SwinUNETR	U-Mamba	Ours
liver	94.29	90.86	82.60	86.95	**94.42**	87.35	17.08	29.35	58.76	71.22	11.41	**8.09**
right kidney	77.59	70.87	55.10	60.51	76.26	**80.23**	18.90	64.94	95.00	57.82	**16.22**	18.72
spleen	84.54	76.15	64.87	72.44	87.92	**88.34**	30.88	23.86	97.18	63.42	14.57	**13.18**
pancreas	71.85	59.24	42.96	53.53	76.23	**76.35**	11.12	21.84	44.93	96.44	**8.16**	11.27
aorta	92.64	90.31	80.44	88.38	91.75	**94.42**	9.27	12.44	42.97	17.50	4.38	**3.20**
inferior vena cava	79.20	72.03	61.19	70.21	78.92	**80.93**	8.33	17.01	40.26	70.12	**7.62**	11.58
right adrenal gland	**69.94**	57.20	46.77	59.57	68.63	66.79	**3.71**	7.78	27.30	8.84	4.44	4.53
left adrenal gland	65.58	52.10	31.15	51.14	71.28	**69.13**	4.92	6.69	23.47	11.70	4.88	**4.07**
gallbladder	57.39	50.13	36.25	40.11	59.18	**65.74**	21.59	40.25	102.12	75.94	25.67	**20.27**
esophagus	**75.53**	65.66	53.01	64.65	74.26	70.33	**10.08**	11.23	21.00	18.64	10.46	11.26
stomach	80.21	69.59	55.38	62.02	80.68	**80.80**	25.65	34.53	81.30	98.04	14.83	**9.68**
duodenum	60.78	46.93	32.90	43.29	63.85	**64.49**	27.54	33.21	49.49	57.30	21.08	**16.94**
left kidney	**82.65**	65.96	50.65	58.32	81.21	80.60	32.26	75.29	86.95	97.07	24.01	**21.20**
Avgerage	76.32	66.69	53.33	62.39	77.28	**77.35**	17.03	29.11	59.29	57.24	12.90	**11.85**

Table C.2. Organ-wise quantitative results in terms of Jaccard Index and HD95 on Abdomen MRI dataset. The best result for each class is bolded.

	Jaccard Index (%)↑						HD95 (mm)↓					
	nnU-Net	SegResNet	UNETR	SwinUNETR	U-Mamba	Ours	nnUnet	SegResNet	UNETR	SwinUNETR	U-Mamba	Ours
liver	94.84	93.31	87.69	92.81	94.78	**95.12**	3.92	6.31	29.48	8.61	3.53	**2.24**
right kidney	**92.77**	88.50	70.44	87.88	92.20	92.23	5.71	2.64	44.03	16.08	2.07	**1.87**
spleen	84.01	82.20	75.70	84.06	**88.38**	86.92	7.67	6.30	53.78	12.54	**2.45**	2.70
pancreas	76.04	70.20	57.12	63.27	76.21	**77.75**	4.90	9.19	18.46	17.95	4.64	**3.50**
aorta	87.39	87.46	74.28	80.29	85.93	**90.51**	5.45	6.72	12.47	6.98	**4.85**	10.34
inferior vena cava	69.38	70.63	57.31	64.53	71.26	**71.73**	11.56	**6.03**	26.07	12.95	6.78	7.18
right adrenal gland	45.33	43.53	29.36	38.31	45.85	**46.99**	6.45	5.86	14.19	8.28	6.82	**5.80**
left adrenal gland	54.10	51.27	31.02	31.46	**56.56**	56.51	5.58	5.12	19.20	12.18	4.83	**4.56**
gallbladder	62.87	64.91	31.87	48.56	70.71	**74.34**	13.22	27.35	69.40	32.23	9.63	**7.92**
esophagus	59.86	58.83	39.42	47.90	66.04	**66.11**	4.80	5.16	14.99	8.60	5.32	**5.07**
stomach	70.36	61.41	50.34	56.23	70.31	**72.06**	17.65	25.14	35.93	27.07	19.49	**15.64**
duodenum	53.94	49.63	32.26	42.06	**57.59**	56.84	14.00	14.84	30.66	22.14	**12.63**	13.99
left kidney	92.99	92.47	75.78	86.32	93.22	**93.37**	2.90	3.20	45.39	8.14	2.66	**1.63**
Avgerage	72.61	70.34	54.81	63.36	74.54	**75.42**	7.98	9.53	31.85	14.90	6.59	**6.34**

CPNet. The quantitative results in terms of JAC and HD95 on the Abdomen CT and MRI datasets are listed in Table C.1 and Table C.2, respectively. Specifically, the average JAC of our model is 77.35% on the CT dataset. The JAC of SOTA models, *i.e*., nnU-Net, SegResNet, UNETR, SwinUNETR and U-Mamba are 76.32%, 66.69%, 53.33%, 62.39%, and 77.28%, respectively. For the MRI dataset, we get 75.42% of the average JAC. For SOTA models, the JACs are 72.61%, 70.34%, 54.81%, 63.36%, and 74.54%, respectively. Moreover, the average HD95(mm) of our CPNet is 11.85 on the CT dataset. For SOTA models, the results are 17.03, 29.11, 59.29, 57.24, and 12.90, respectively. For the MRI dataset, our CPNet achieves 6.34 HD95(mm). For SOTA models, the HD95(mm)s are 7.98, 9.53, 31.85, 14.90, and 6.59, respectively.

In addition, we have also performed another two runs of experiments on our CPNet. The mean and confidence interval (CI) of our trained models achieve 87.24%±0.81% DSC and 91.04%±0.90% NSD on the Abdomen CT dataset, and 85.07%±0.17% DSC and 91.92%±0.23% NSD on the Abdomen MRI dataset.

Runtime Efficiency. We further compared the inference time and the memory consumption with the SOTA methods. As shown in Table C.3, it takes about 0.11 seconds to infer a CT volume with a resolution of 40×224×192 on a machine with one NVIDIA GeForce RTX 4090 GPU (24G). It requires 4.077 GB of VRAM for testing. For SOTA methods, including nnU-Net, SegResNet, UNETR, SwinUNETR and U-Mamba, the time costs are 0.05, 0.08, 0.04, 0.29 and 0.16 seconds, respectively.

Table C.3. Quantitative results of time cost(s).

	Abdomen CT	Abdomen MRI
Ours	0.1123	0.1118
U-Mamba	0.1554	0.1533
UNETR	0.0363	0.0304
SwinUNETR	0.2894	0.2467
nnU-Net	0.0547	0.0538
SegResNet	0.0820	0.0813

D Disscusion on Limitations and Future work

Our proposed SRPP and MCPP modules are designed for segmenting multiple organs with regular shapes and anatomical structure priors. It may not be suitable for segmenting tissues or organs with irregular structures, such as coronary arteries, vessels, *etc*. Therefore, in future, it is worth further investigation into utilizing baseline models and their lightweight variations, leveraging domain adaptation techniques to improve the cross-modality generalization ability and exploring more anatomical constraints as extra regularizers for different medical image segmentation tasks.

For practical applications, our model can be used to identify and locate the organs, facilitate the analysis of the abnormalities of the organs, locate tumors and control the dose of the X-ray beam in radiotherapy. For model deployment, it is potential and feasible to integrate the trained model into some open-source software/system for biomedical image analysis, *e.g*., 3D Slicer, MITK, *etc*. Moreover, it is still an open question to clarify the interpretability of the proposed model, and so are other deep-learning-based models. We believe the deep-learning-based AI system still works in the assisted diagnosis mode, which requires the doctor to finalize the diagnosis.

References

[1] Q.-Y. Zhou, J. Park, and V. Koltun. Open3D: A modern library for 3D data processing. *arXiv:1801.09847*, 2018.