SemiGDA: Generative Dual-distribution Alignment for Semi-Supervised Medical Image Segmentation

Supplementary Material

A. More Experiment

A.1. Experiments on Colonoscopy Datasets

Dataset description: In the endoscopic dataset, two additional unseen datasets are included, namely CVC-ColonDB [14] and ETIS [13]. Note that, consistent with the previous settings [6], the weights used for testing here are those trained on the two seen datasets (CVC-ClinicDB [1] and Kvasir [7]) in the main text. The number of test samples in these two datasets is 380 and 196, respectively.

Quantitative and Qualitative Results: Tab. 1 reports the quantitative results on two unseen endoscopic datasets, CVC-ColonDB and ETIS. As observed, with only 10% of annotations, our method achieves increases of 10.18% and 7.65% in the Dice and IoU metrics, respectively, compared to the second-best method (UnCo). Similarly, on the ETIS dataset, the advantage of our method becomes even more pronounced with limited annotations. This demonstrates that generative-based methods outperform discriminative segmentation methods under low-annotation settings. Additionally, although this advantage decreases at the 30% annotation level, our method still achieves the best performance across all metrics. As shown in Fig. 2, our method effectively captures the full shape of the target, particularly in small targets and along contour boundaries.

A.2. Experiments on Thyroid Datasets

Dataset description: The thyroid dataset is composed of three subsets: TN3K [4], TG3K [17], and DDTI [12]. The training and validation sets are derived from $2,879$ images of TN3K and $3,585$ images of TG3K, comprising $6,142$ images for training and 322 images for validation. The test dataset is constructed using images from TN3K and DDTI, totaling 614 images.

Comparison with State-of-the-Art Methods: Tab. 2 presents the comparative results on the ultrasound task with other semi-supervised medical image segmentation methods, where DDTI serves as the unseen test set. It can be seen that our model achieves state-of-the-art performance across all metrics. Notably, on the TN3K test set, our method exhibits significantly superior performance compared to others. For instance, with 30% labeled data on the TN3k dataset, our method improves the Dice coefficient from 80.73% to 83.57% and the IoU from 71.41% to 72.21%, compared to the second-best method. Moreover, Fig. 2 shows visual results of different semi-supervised medical image segmentation methods. It can be observed that our model accurately identifies and segments the nodule regions, while other methods frequently struggle with inaccuracies in segmentation fragments.

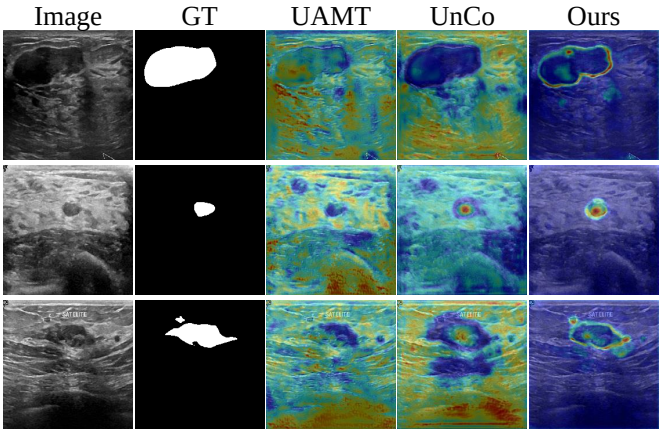

Figure 1. Heatmap comparison with other semi-supervised methods on the BUSI dataset.

B. More Model Analysis

As shown in the heatmap in Fig. 1, our method demonstrates clear advantages in target region focusing and boundary accuracy across large, small, and irregular targets. This confirms that our generative-based semi-supervised framework effectively captures fine-grained boundary features, crucial for medical image segmentation and clinical diagnosis. Additionally, its ability to reduce background interference and learn feature distributions from unlabeled data further supports the effectiveness of our method in feature alignment and consistency constraint design.

C. More Implementation Details

Our method employs a two-stage training paradigm. The first stage is specifically designed to warm up the Dual-distribution Alignment Module (DAM). Given that DAM involves complex distribution-matching operations, direct end-to-end training is prone to optimization oscillations due to unstable initial parameters. In the second stage, pre-trained weights are loaded for fine-tuning, which effectively enhances the robustness of the generative model. For the Consistency-Driven Skip Adapter strategy, the maximum number of N_f is set to 5, considering the original model architecture and avoiding excessive increases in the number of learnable parameters. We adopt the PolyLR learning rate scheduling strategy.

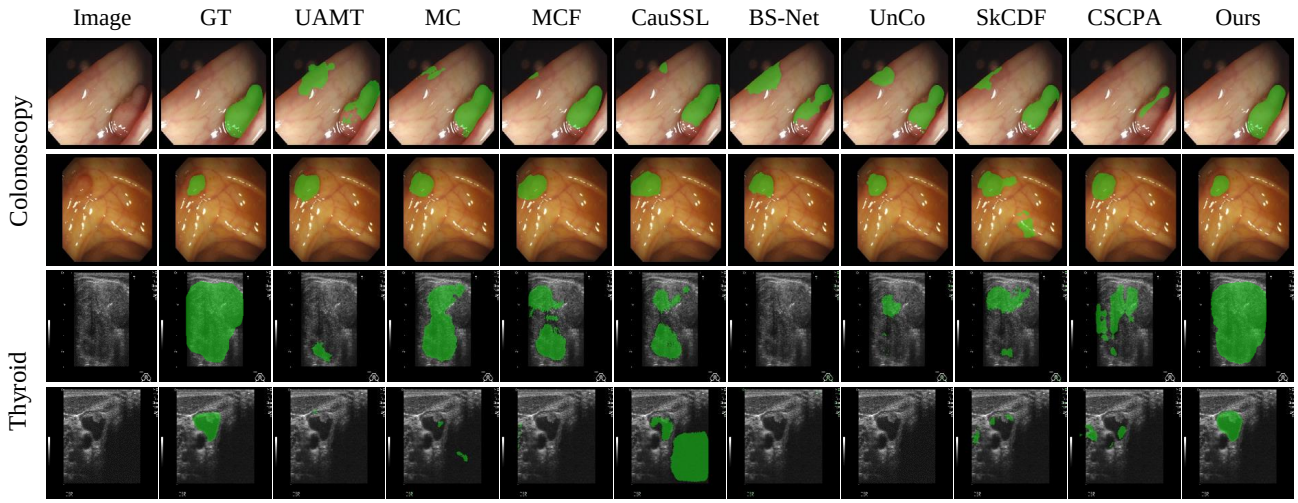

Figure 2. Visual comparisons of state-of-the-art medical segmentation methods on the colonoscopy and thyroid datasets.

Table 1. Quantitative results on the ETLS and CVC-ColonDB datasets.

Methods		CVC-ColonDB						ETIS					
		10% labeled			30% labeled			10% labeled			30% labeled		
		Dice (%)	IoU (%)	95HD	Dice (%)	IoU (%)	95HD	Dice (%)	IoU (%)	95HD	Dice (%)	IoU (%)	95HD
UA-MT [18]	[MICCAI'19]	41.34	33.49	5.72	59.43	51.30	4.74	31.84	25.59	4.96	41.18	34.80	4.64
DTC [9]	[AAAI'21]	35.35	27.17	6.10	54.75	46.18	5.05	26.74	20.48	6.05	44.83	37.72	4.40
MC-Net [16]	[MICCAI'21]	55.69	46.53	5.24	63.65	55.38	4.71	39.88	33.15	4.81	49.51	41.69	4.74
URPC [10]	[MIA'22]	46.84	38.98	5.55	60.00	51.89	4.78	38.98	32.92	4.78	45.00	36.75	4.89
MCF [15]	[CVPR'23]	48.72	40.60	5.25	63.24	55.05	4.64	31.13	25.81	4.84	48.46	41.28	4.33
CauSSL [11]	[ICCV'23]	48.51	40.89	5.24	57.84	49.36	4.95	30.93	24.96	5.18	32.97	27.41	5.42
CDMA [21]	[MICCAI'23]	44.88	35.72	5.77	59.39	49.10	5.20	37.23	28.85	4.87	38.29	29.37	6.07
BS-Net [5]	[TMI'24]	43.88	36.57	5.43	55.89	48.15	4.87	31.77	26.26	5.28	38.08	31.65	4.91
PMT [3]	[ECCV'24]	47.21	39.47	4.74	58.63	51.12	4.53	29.09	23.90	4.90	39.34	32.90	4.67
VCLIPSeg [8]	[MICCAI'24]	50.19	42.19	4.79	62.02	54.34	4.67	37.92	31.14	4.83	45.89	38.99	4.46
UnCo [19]	[TMI'25]	57.62	50.64	4.64	62.18	55.04	4.45	37.73	32.92	4.61	42.04	36.34	4.27
SKCDF [20]	[CVPR'25]	45.82	37.51	4.55	58.02	47.58	4.79	35.02	28.34	4.82	36.50	30.14	5.15
CSCPA [2]	[CVPR'25]	55.66	46.80	4.50	68.70	60.22	4.06	38.86	32.08	4.68	52.58	44.72	4.25
Ours		**67.80**	**58.29**	**4.47**	**71.81**	**63.38**	**4.02**	**56.77**	**48.26**	**4.52**	**59.86**	**53.10**	**4.11**

Table 2. Quantitative results on the thyroid datasets.

Methods		DDTI						TN3K					
		10% labeled			30% labeled			10% labeled			30% labeled		
		Dice (%)	IoU (%)	95HD	Dice (%)	IoU (%)	95HD	Dice (%)	IoU (%)	95HD	Dice (%)	IoU (%)	95HD
UA-MT [18]	[MICCAI'19]	34.96	27.31	6.73	49.76	39.85	6.07	74.73	64.23	4.83	78.42	69.16	4.53
DTC [9]	[AAAI'21]	27.13	19.67	7.35	38.53	28.05	6.87	62.67	50.50	5.61	64.48	51.66	5.74
MC-Net [16]	[MICCAI'21]	29.42	23.12	7.15	45.36	36.60	6.20	71.45	60.72	5.28	78.04	68.72	4.65
URPC [10]	[MIA'22]	31.37	24.23	6.88	56.01	45.26	5.92	74.42	64.17	4.88	79.19	70.00	4.56
MCF [15]	[CVPR'23]	**57.69**	**46.84**	**5.71**	55.09	44.08	5.97	74.50	64.46	4.79	78.04	68.79	4.43
CauSSL [11]	[ICCV'23]	41.88	32.76	6.52	46.12	36.51	6.22	70.89	60.22	5.17	73.93	63.63	4.97
CDMA [21]	[MICCAI'23]	46.42	36.17	6.28	54.68	42.48	6.34	72.97	61.54	5.12	76.24	65.60	4.81
BS-Net [5]	[TMI'24]	31.41	24.82	6.74	40.55	32.42	6.33	73.70	64.22	4.77	79.25	70.60	4.39
PMT [3]	[ICCV'24]	33.84	26.83	5.35	34.68	27.60	5.90	71.81	62.12	4.54	71.56	62.06	4.41
VCLIPSeg [8]	[MICCAI'24]	37.80	29.99	6.25	49.74	40.77	5.92	71.31	61.56	4.48	77.65	68.24	4.31
UnCo [19]	[TMI'25]	26.04	20.42	7.02	43.23	34.82	6.24	72.36	63.03	4.58	74.93	65.64	4.32
SKCDF [20]	[CVPR'25]	53.66	42.08	6.22	52.37	40.89	6.58	72.44	62.02	4.82	75.89	65.75	4.74
CSCPA [2]	[CVPR'25]	52.70	42.73	6.21	57.92	47.61	5.83	78.28	68.81	4.45	80.73	71.41	4.30
Ours		56.07	45.37	6.20	**61.49**	**50.85**	**5.71**	**80.25**	**68.94**	**4.44**	**83.57**	**72.21**	**4.18**

CVPR #34844

D. Model Complexity

Tab. 3 details the efficiency comparison. The parameter increase is inherent to the generative paradigm. While heavier than BSNet (86.49M vs 23.6M), SemiGDA gains +12.56% Dice (75.57% vs 63.01%). Crucially, the 7.90 GB footprint enables effortless single-GPU deployment, while 32 FPS satisfies real-time requirements.

Table 3. Efficiency comparison on BUSI at 224 × 224 resolution.

Method	Params (M)	Memory (GB)	FPS	Dice (10%)
BSNet	23.6	4.2	164	63.01
CSCPA	46.5	4.7	121	65.16
Ours	86.4	7.9	32	**75.57**

References

[1] Jorge Bernal, F Javier Sánchez, Gloria Fernández-Esparrach, Debora Gil, Cristina Rodríguez, and Fernando Vilariño. WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Computerized Medical Imaging and Graphics*, 43:99–111, 2015. 1

[2] Zhenhui Ding, Guilian Chen, Qin Zhang, Huisi Wu, and Jing Qin. Csc-pa: Cross-image semantic correlation via prototype attentions for single-network semi-supervised breast tumor segmentation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 15632–15641, 2025. 2

[3] Ning Gao, Sanping Zhou, Le Wang, and Nanning Zheng. PMT: Progressive mean teacher via exploring temporal consistency for semi-supervised medical image segmentation. *European Conference on Computer Vision*, 2024. 2

[4] Haifan Gong, Jiaxin Chen, Guanqi Chen, Haofeng Li, Guanbin Li, and Fei Chen. Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules. *Computers in Biology and Medicine*, 155:106389, 2023. 1

[5] Along He, Tao Li, Juncheng Yan, Kai Wang, and Huazhu Fu. Bilateral supervision network for semi-supervised medical image segmentation. *IEEE Transactions on Medical Imaging*, 43(5):1715–1726, 2024. 2

[6] Kaiwen Huang, Tao Zhou, Huazhu Fu, Yizhe Zhang, Yi Zhou, Chen Gong, and Dong Liang. Learnable prompting sam-induced knowledge distillation for semi-supervised medical image segmentation. *IEEE Transactions on Medical Imaging*, 44(5):2295–2306, 2025. 1

[7] Debesh Jha, Pia H Smedsrud, Michael A Riegler, Pål Halvorsen, Thomas De Lange, Dag Johansen, and Håvard D Johansen. Kvasir-seg: A segmented polyp dataset. In *MMM*, pages 451–462. Springer, 2020. 1

[8] Lei Li, Sheng Lian, Zhiming Luo, Beizhan Wang, and Shaozi Li. Vclipseg: Voxel-wise clip-enhanced model for semi-supervised medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 692–701. Springer, 2024. 2

[9] Xiangde Luo, Jieneng Chen, Tao Song, and Guotai Wang. Semi-supervised medical image segmentation through dual-task consistency. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 8801–8809, 2021. 2

[10] Xiangde Luo, Guotai Wang, Wenjun Liao, Jieneng Chen, Tao Song, Yinan Chen, Shichuan Zhang, Dimitris N Metaxas, and Shaoting Zhang. Semi-supervised medical image segmentation via uncertainty rectified pyramid consistency. *Medical Image Analysis*, 80:102517, 2022. 2

[11] Juzheng Miao, Cheng Chen, Furui Liu, Hao Wei, and Pheng-Ann Heng. Caussl: Causality-inspired semi-supervised learning for medical image segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21426–21437, 2023. 2

[12] Lina Pedraza, Carlos Vargas, Fabián Narváez, Oscar Durán, Emma Muñoz, and Eduardo Romero. An open access thyroid ultrasound image database. In *International Symposium on Medical Information Processing and Analysis*, pages 188–193. SPIE, 2015. 1

[13] Juan Silva, Aymeric Histace, Olivier Romain, Xavier Dray, and Bertrand Granado. Toward embedded detection of polyps in wce images for early diagnosis of colorectal cancer. *International Journal of Computer Assisted Radiology and Surgery*, 9:283–293, 2014. 1

[14] Nima Tajbakhsh, Suryakanth R Gurudu, and Jianming Liang. Automated polyp detection in colonoscopy videos using shape and context information. *IEEE Transactions on Medical Imaging*, 35(2):630–644, 2015. 1

[15] Yongchao Wang, Bin Xiao, Xiuli Bi, Weisheng Li, and Xinbo Gao. Mcf: Mutual correction framework for semi-supervised medical image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 15651–15660, 2023. 2

[16] Yicheng Wu, Minfeng Xu, Zongyuan Ge, Jianfei Cai, and Lei Zhang. Semi-supervised left atrium segmentation with mutual consistency training. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 297–306. Springer, 2021. 2

[17] Tom Wunderling, B Golla, Prabal Poudel, Christoph Arens, Michael Friebe, and Christian Hansen. Comparison of thyroid segmentation techniques for 3D ultrasound. In *Medical Imaging: Image Processing*, pages 346–352. SPIE, 2017. 1

[18] Lequan Yu, Shujun Wang, Xiaomeng Li, Chi-Wing Fu, and Pheng-Ann Heng. Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 605–613. Springer, 2019. 2

[19] Xiang Zeng, Shengwu Xiong, Jinming Xu, Guangxing Du, and Yi Rong. Uncertainty co-estimator for improving semi-supervised medical image segmentation. *IEEE Transactions on Medical Imaging*, 44(9):3870–3881, 2025. 2

[20] Zheng Zhang, Guanchun Yin, et al. A semantic knowledge complementarity based decoupling framework for semi-supervised class-imbalanced medical image segmentation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 25940–25949, 2025. 2

[21] Lanfeng Zhong, Xin Liao, Shaoting Zhang, and Guotai Wang. Semi-supervised pathological image segmentation

via cross distillation of multiple attentions. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 570–579. Springer, 2023. 2