SAM-DA: Decoder Adapter for Efficient Medical Domain Adaptation Supplementary Material

Javier Gamazo Tejero[1], Moritz Schmid[1], Pablo Márquez Neila[1]
Martin S. Zinkernagel[2], Sebastian Wolf[2], Raphael Sznitman[1]
[1]University of Bern, [2]Inselspital Bern, Switzerland
{javier.gamazo-tejero, moritz.schmid, raphael.sznitman, pablo.marquez}@unibe.ch
{martin.zinkernagel, sebastian.wolf}@insel.ch

1. Implementation details

Figure 1 shows the architecture of our adapter (A_ℓ) as a neural diagram (introduced in [1]). It is combined with the embeddings T_ℓ in an attention module in which T_ℓ act as queries and the adapter weights A_ℓ act as keys and values.

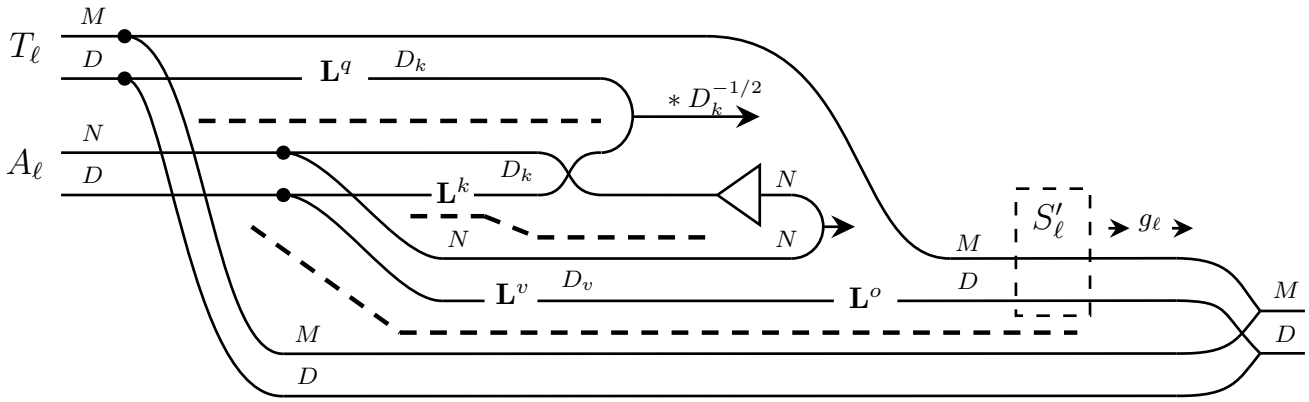

Figure 1. Neural circuit diagram for the proposed SAM-Decoder-Adapter

2. Statistical Analysis

To evaluate the statistical significance of performance differences between Med-SA [7] and SAM-DA on the fully supervised task, we conducted a paired t-test on image-wise mIoU scores obtained from both methods on the Retouch-Spectralis dataset. The test was conducted under the null hypothesis that Med-SA and SAM-DA achieve the same mean mIoU across images. A p-value of $p < 0.01$ was obtained, providing sufficient evidence to reject the null hypothesis and indicating a significant performance difference between the methods.

As each image may contain a unique subset of classes, image-wise mIoU scores are not directly comparable across images, and the application of this paired t-test is not completely justified from a theoretical standpoint. However, this approach was selected as the most feasible option among available alternatives, despite its limitations. Alternative paired t-test methods were considered but ultimately dismissed as unworkable. For instance, conducting a t-test over multiple random seeds per model would have required more than 100 seeds to reach a statistical power of 0.9, which was impractical. A pixel-level paired t-test was also ruled out due to the high correlation between pixels within images, which would likely yield an artificially low and unreliable p-value.

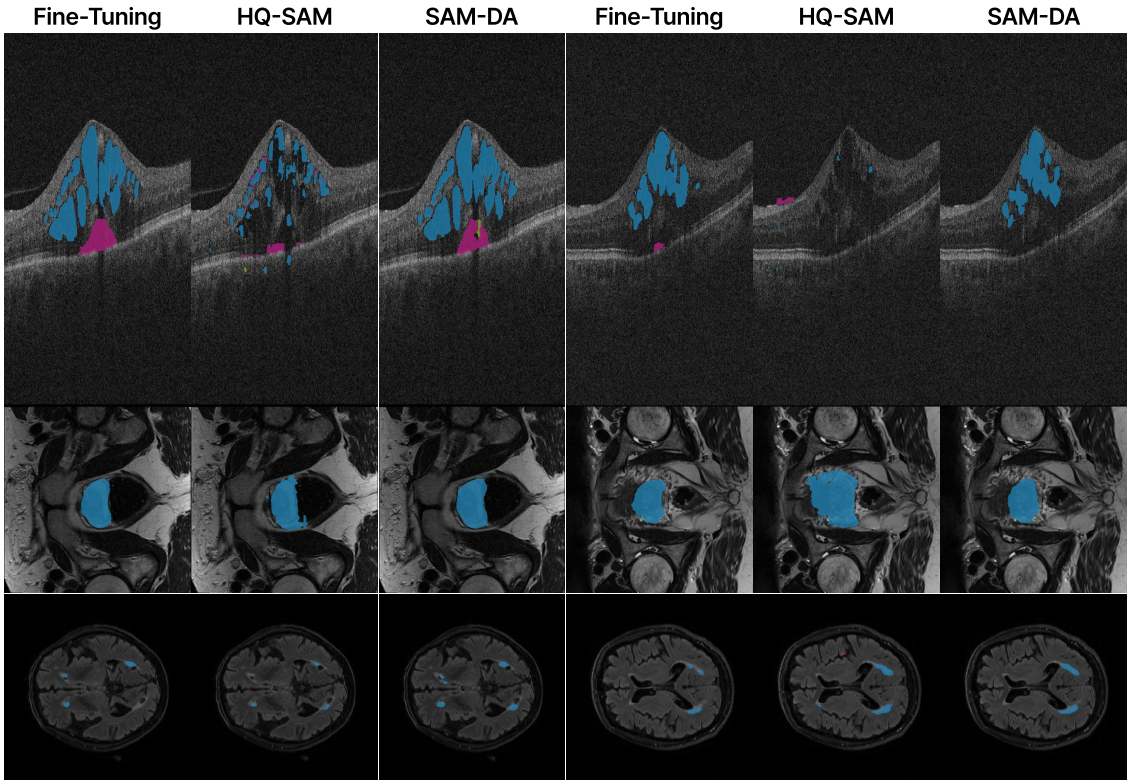

Figure 2. Qualitative results on three randomly selected test samples from domain generalization subsets

3. Further results

Table 1 shows the IoU of all the methods on the four different subsets that compose HQSeg-44K [3]. Encoder Adapter and Decoder Adapter represent our proposed methods with two different placements of the adapter (encoder and decoder, respectively). Due to the high number of images, we see that adaption methods with a higher number of parameters, such as LoRA, outperform smaller ones, such as HQ-SAM or our SAM Adapter.

	COIFT	HRSOD	ThinObject5k	DIS5K
Fine-Tuning	$82.07_{\pm 0.58}$	$79.44_{\pm 0.14}$	$79.32_{\pm 0.33}$	$63.33_{\pm 0.34}$
Decoder FT	$84.91_{\pm 0.58}$	$82.46_{\pm 0.14}$	$84.54_{\pm 0.33}$	$71.61_{\pm 0.34}$
LoRA	$86.01_{\pm 0.19}$	$84.50_{\pm 0.18}$	$87.86_{\pm 0.26}$	$74.22_{\pm 0.52}$
Med-SA	$85.74_{\pm 0.15}$	$83.94_{\pm 0.73}$	$89.83_{\pm 0.31}$	$75.69_{\pm 0.28}$
HQ-SAM	$84.17_{\pm 0.20}$	$81.41_{\pm 0.26}$	$81.41_{\pm 0.12}$	$69.88_{\pm 0.12}$
Encoder Adapter	$84.82_{\pm 0.26}$	$82.41_{\pm 0.30}$	$84.61_{\pm 0.15}$	$71.41_{\pm 0.42}$
Decoder Adapter	$84.61_{\pm 0.28}$	$81.81_{\pm 0.50}$	$82.73_{\pm 0.10}$	$69.24_{\pm 0.56}$

Table 1. IoU of all the methods on the different subsets that compose HQSeg-44K. Variance has been obtained over four trained models on the validation set

Figure 2 shows qualitative results on three test samples selected randomly from the untrained domains (Cirrus for Retouch [2], UCL for MRI [5, 6], and NUHS Singapore for WMH [4]).

Tables 2 and 3 show the impact of the dimension of the adaption prompt A_ℓ on the performance.

	Retouch - Spectralis	MRI - BMC	WMH - Utrecht	HQ-Seg
512	$75.4_{\pm 0.6}$	$86.2_{\pm 1.5}$	$44.2_{\pm 0.3}$	$79.6_{\pm 0.4}$
1024	$75.8_{\pm 0.8}$	$85.6_{\pm 1.6}$	$40.5_{\pm 3.8}$	$71.6_{\pm 2.8}$
2048	$76.3_{\pm 0.8}$	$85.2_{\pm 1.7}$	$39.0_{\pm 4.7}$	$70.2_{\pm 2.3}$

Table 2. Ablation study for the size of the adapter embeddings. IoU scores for full supervision. Variances are computed over four trained models tested on the testing set.

	Retouch - Cirrus	MRI - UCL	WMH - Singapore
512	$70.2_{\pm 3.1}$	$80.6_{\pm 1.0}$	$39.6_{\pm 0.7}$
1024	$69.4_{\pm 1.3}$	$79.8_{\pm 1.3}$	$38.7_{\pm 3.1}$
2048	$69.2_{\pm 1.2}$	$80.2_{\pm 1.1}$	$35.5_{\pm 3.4}$

Table 3. Ablation study for the size of the adapter embeddings. Application of trained models to zero-shot domain generalization. The variance was obtained over four trained models tested on the testing set.

References

[1] Vincent Abbott. Neural circuit diagrams: Robust diagrams for the communication, implementation, and analysis of deep learning architectures. *arXiv preprint arXiv:2402.05424*, 2024.

[2] Hrvoje Bogunovic, Freerk Venhuizen, Sophie Klimscha, et al. Retouch: The retinal oct fluid detection and segmentation benchmark and challenge. *IEEE Transactions on Medical Imaging*, 38:1858–1874, 8 2019.

[3] Lei Ke, Mingqiao Ye, Martin Danelljan, Yu-Wing Tai, Chi-Keung Tang, , et al. Segment anything in high quality. *Advances in Neural Information Processing Systems*, 36, 2024.

[4] Hugo J. Kuijf, Adria Casamitjana, D. Louis Collins, Mahsa Dadar, Achilleas Georgiou, et al. Standardized assessment of automatic segmentation of white matter hyperintensities and results of the wmh segmentation challenge. *IEEE Transactions on Medical Imaging*, 38:2556–2568, 11 2019.

[5] Quande Liu, Qi Dou, and Pheng-Ann Heng. Shape-aware meta-learning for generalizing prostate mri segmentation to unseen domains. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part II 23*, pages 475–485. Springer, 2020.

[6] Quande Liu, Qi Dou, Lequan Yu, and Pheng Ann Heng. Ms-net: Multi-site network for improving prostate segmentation with heterogeneous mri data. *IEEE Transactions on Medical Imaging*, 39:2713–2724, 9 2020.

[7] Junde Wu, Rao Fu, Huihui Fang, et al. Medical sam adapter: Adapting segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.12620*, 2023.