GA-SAM: Geometry-Aware SAM Adaptation with Sparse Annotation-Driven Point Cloud Completion

Shumeng Li[1], Jian Zhang[2,1], Lei Qi[3], and Yinghuan Shi[1,⋆]

[1] State Key Laboratory of Novel Software Technology, National Institute of Healthcare Data Science, Nanjing University, China
lism@smail.nju.edu.cn, syh@nju.edu.cn
[2] School of Intelligence Science and Technology, Nanjing University, China
[3] School of Computer Science and Engineering, Key Lab of Computer Network and Information Integration (Ministry of Education), Southeast University, China

Abstract. Segment Anything Model (SAM) adaptation has shown remarkable performance in medical image segmentation, but typically relies on large and precisely annotated datasets. However, acquiring such dense annotation is a labor-intensive and time-consuming task that requires significant expertise. An effective direction is to focus on sparse annotation, where only a few slices are annotated. However, sparse annotations are insufficient for capturing the complete 3D anatomical structure. To address this limitation, we innovatively leverage point cloud completion to generate robust volumetric shape from sparse annotation, offering a promising solution to this challenge. In this paper, we propose a novel Geometry-Aware SAM adaptation framework (namely **GA-SAM**) that integrates point cloud shape generation module with cross-view segmentation supervision mechanism. Specifically, we train a point cloud completion network to infer the 3D structure of the target anatomy. The generated point cloud shapes are then used to produce pseudo-labels, guiding the adaptation of SAM via a geometry-aware shape constraints. Furthermore, we incorporate a cross-view supervision mechanism, leveraging multi-view consistency to ensure reliable segmentation across different planes. We demonstrate the effectiveness of our method on Pancreas-CT dataset, surpassing the state-of-the-art SAM adaptation method by a Dice score of 15.25% and significantly improving segmentation robustness. Our code is available at https://github.com/ShumengLI/GA-SAM.

Keywords: Medical Image Segmentation · Sparse Annotation · Foundation Model · Point Cloud Completion.

1 Introduction

Recently, foundation models [9, 10, 18, 22, 30] for visual segmentation have gained significant attention in medical imaging due to their strong generalization and

⋆ Corresponding author.

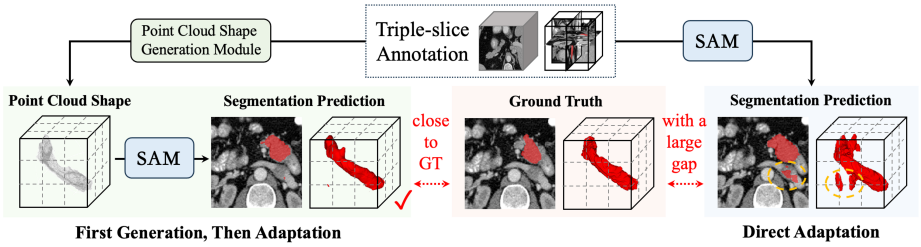

Fig. 1. An example of triple-slice annotation and comparison of segmentation strategies. It demonstrates the direct adaptation (right) exhibits a noticeable discrepancy from the ground truth. In contrast, the point cloud generation followed by SAM adaptation (left) produces a closer result to the ground truth, highlighting the effectiveness of point cloud completion in enhancing anatomical shape representation.

segmentation performance. As one of the most popular universal image segmentation models, Segment Anything Model (SAM) [10] has been employed to medical image segmentation by adaptation techniques under fully supervised [3, 5, 7, 14, 16, 17, 23, 28] or semi-supervised [4, 15, 17, 29] way. However, it remains heavily dependent on precisely dense annotations from experienced radiologists, making the annotation process expensive and time-consuming. For instance, with the Pancreas-CT dataset [19], annotating just 10% of the samples for SAM adaptation still takes over 400 hours (assuming an experienced radiologist spends 10 minutes annotating each 512×512 slice). Therefore, we wonder, *whether a foundation model can be adapted with only few annotations while still achieving comparable performance to that with extensive annotations?*

Compared to simply reducing the number of full annotated volumes, which often results in inferior results [1, 26], sparse annotation [1, 2, 13, 21, 25, 26] is a more effective strategy within the same annotation budget. With sparse annotation, fewer slices are labeled in each volume, but information is provided from more volumes. Previous studies [1] have revealed that annotations from different planes provide varied structural information. Building on this insight, we introduce Triple-Slice Annotation, where a single annotated slice is selected from each of the three planes (*e.g.*, transverse, coronal, and sagittal), enabling the model to integrate multi-view structural information.

However, sparse annotations alone may not capture the full 3D anatomical structure. For example, when only a few slices are labeled per volume, the limited supervision is insufficient to establish global shape constraints, leading to incomplete or imprecise object representations [25], as shown in Fig. 1. To address this, it becomes essential to *establish reliable shape constraints from sparse annotations* that guide the model to infer full structure information.

With the aforementioned goal, it is crucial to incorporate a representation that can capture the full 3D anatomy. We recognize that point clouds provide a natural and efficient way to represent 3D structures. Point cloud-based shape modeling [27] enables the estimation of the complete geometry of objects from

partial observations. PCN [27] pioneered learning-based point completion by directly performing reconstruction on the input points, while FSC [24] further investigated the point cloud completion with only a few input points. In medical image segmentation, many targets, such as organs, share similar anatomical shapes. This inherent regularity allows point cloud completion, guided by learned shape priors, to infer the global 3D structure of an organ from sparse annotations. Consequently, it could serve as a powerful mechanism for shape generation, which offers an effective solution to the challenges of incomplete structural representation caused by sparse annotations.

To this end, we propose a novel Geometry-Aware SAM adaptation framework (namely **GA-SAM**) that effectively integrates point cloud shape modeling with the SAM adaptation. Specifically, we introduce a point cloud completion network trained on a medical shape dataset, MedShapeNet [11], to infer the complete 3D structure of the target anatomy using only triple-slice annotation. The predicted point cloud shapes are then employed to generate pseudo-labels through voxel reconstruction, providing a structural guide for segmentation. We introduce geometry-aware shape guidance, where the pseudo-labels serve as a soft shape prior to enhance anatomical consistency. Furthermore, to effectively leverage multi-view information, we introduce three independent branches for three planes and establish a cross-view supervision mechanism. We use the consistently predicted foreground regions predicted by two models to supervise the third model to ensure the reliability of supervision. This cross-view supervision strengthens the structural consistency across different views. To sum up, our main contributions are illustrated as follows:

- A novel paradigm for triple-slice annotation, leveraging point cloud completion to generate robust volumetric shapes.
- A new geometry-aware guidance strategy, incorporating soft shape priors to enhance shape awareness in segmentation.
- A cross-view supervision mechanism, leveraging multi-view consistency to ensure reliable segmentation and enhance robustness.

We evaluate our framework on the Pancreas-CT dataset [19] and demonstrate that point cloud shape generation provides an anatomically reliable and structurally robust solution for triple-slice annotated medical image segmentation. Our method achieves significant performance improvements, with a Dice score of 67.12%, surpassing state-of-the-art approaches under sparse annotation.

2 Method

2.1 Overview

The architecture overview of our GA-SAM framework is illustrated in Fig. 2. The framework consists of two key components: (1) point cloud shape generation module and (2) cross-view foundational model adaptation module. First, we generate a point cloud representation from three mutually orthogonal annotated slices and reconstruct their corresponding complete 3D shapes by a point

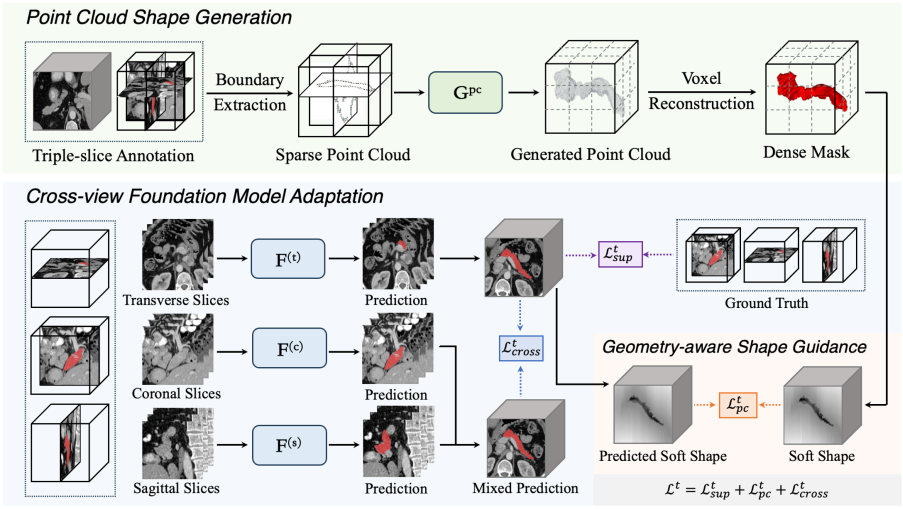

Fig. 2. Overview of our proposed GA-SAM. The segmentation models $\mathtt{F}^{(t)}$, $\mathtt{F}^{(s)}$, and $\mathtt{F}^{(c)}$ are trained alternately. In the example shown, $\mathtt{F}^{(t)}$ acts as the student model, learning from the supervision provided by the other two views.

cloud completion network. Next, we serve the generated 3D shapes as volumetric pseudo-labels to enhance the spatial awareness of the segmentation model. To fully exploit the triple-slice annotations, the segmentation model performs fine-grained predictions from three views and we employ a cross-view adaptation strategy to process transverse, sagittal, and coronal planes collaboratively.

2.2 Triple-Slice Annotation

Acquiring dense annotations for 3D volumetric data is costly and time-consuming. To alleviate this burden, we introduce triple-slice annotation, a sparse yet effective annotation approach. Instead of annotating full 3D volumes, we annotate only one slice with visible targets for each of the transverse, sagittal, and coronal planes. This approach provides information about the target structure from different perspectives. Formally, let the training set consist of N volumetric scans $\{X_1, X_2, ..., X_N\}$, where $X_i \in \mathbb{R}^{H \times W \times D}$. For each volume $X_i (1 \leq i \leq N)$, we annotate one slice per orthogonal plane, denoted as $X_i^{(t)}$, $X_i^{(s)}$, $X_i^{(c)}$, where t, s, c correspond to the transverse, sagittal, and coronal planes, respectively. Their corresponding segmentation masks are given by $Y_i^{(t)}$, $Y_i^{(s)}$, $Y_i^{(c)}$.

2.3 Point Cloud Shape Generation

To generate 3D anatomical shapes from sparsely annotated slices, we introduce a pre-trained point cloud completion network to reconstruct their corresponding 3D shapes. By leveraging it trained on external anatomical datasets (*e.g.*, MedShapeNet [11]), we generate a high-quality shape representation.

Shape Generation from Triple-slice Annotations Given a volumetric scan X_i with triple-slice annotation, we first extract a set of foreground boundary points from the annotated masks $\{Y_i^{(t)}, Y_i^{(s)}, Y_i^{(c)}\}$ to capture the anatomical contour. The resulting boundary point set then serves as an initial point cloud representation $P_i^{\text{sparse}} = \bigcup_{j\in\{t,s,c\}} \mathcal{B}(Y_i^{(j)}) \subset \mathbb{R}^3$, where $\mathcal{B}(\cdot)$ denotes the boundary extraction operation, which identifies the edge points of the annotated mask. The set of P_i^{sparse} can be denoted as $\{P_{i,j}^{\text{sparse}}\}_{j=1}^{K}$, where K is the total number of points, and each point $P_{i,j}^{\text{sparse}}$ is a 3D coordinate. Next, we normalize the point cloud by aligning its centroid μ_i to the origin and rescaling it based on the maximum Euclidean distance d_i from the centroid, ensuring that the normalized coordinates lie within a consistent scale for point cloud completion. Specifically, the normalized point is $\bar{P}_{i,j}^{\text{sparse}} = \frac{P_{i,j}^{\text{sparse}} - \mu_i}{d_i}$ for each point, and the normalized point cloud is obtained as $\bar{P}_i^{\text{sparse}} = \{\bar{P}_{i,j}^{\text{sparse}}\}_{j=1}^{K}$.

Enlightened by the shape completion scheme presented in [24, 27], we leverage a point completion network to predict the missing regions and generate a dense point cloud representation. The point completion network adopts an encoder-decoder structure [24]. To achieve this, we train the shape completion model on a large-scale external dataset, MedShapeNet [11], and simulate to uniformly sample points from three slices. It enables capturing organ morphology patterns and referring plausible 3D structures from triple-slice annotated inputs. As with [24], we train the point completion network consisting of a feature extractor and a decoder. The feature extractor is designed to separately capture the extensive and salient feature representation, while the decoder refines and expands this representation to generate a dense and coherent point cloud $P_i^{\text{dense}} = G^{pc}(\bar{P}_i^{\text{sparse}})$, where $G^{pc}(\cdot)$ represents the trained point completion network, which refines and densifies the sparse input. The resulting P_i^{dense} provides a structurally consistent and anatomically plausible 3D shape, which serves as the foundation for generating shape-aware pseudo-labels in our segmentation framework. Since the point cloud are normalized before, we rescal the generated point cloud P_i^{dense} back to its original space by applying the inverse of the normalization process to each point $P_{i,j}^{\text{dense}} = P_{i,j}^{\text{dense}} \cdot d_i + \mu_i$ in P_i^{dense}. This ensures that the generated shape remains anatomically plausible and aligned with the input data.

Voxel Reconstruction and Mask Generation Once the generated point cloud P_i^{dense} is obtained, we convert it into a voxel representation that aligns with the original volume space. Specifically, we discretize the point cloud into a binary voxel grid V_i using a point-to-voxel mapping function:

$$V_i(x, y, z) = \begin{cases} 1, & \text{if } (x, y, z) \in P_i^{\text{dense}}, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Since direct voxelization may introduce discrete voxels and sparse coverage, we apply a morphological refinement operation $\mathcal{R}(\cdot)$, including hole filling and smoothing, to generate the final binary mask $M_i = \mathcal{F}(V_i)$. The reconstructed

mask M_i serves as a shape-aware pseudo-label, integrating global anatomical priors into our segmentation network.

2.4 Cross-View Foundation Model Adaptation

For adaptation, we freeze the image encoder of SAM [10], adopt LoRA adaptation [8] by adding a bypass, and fine-tune the mask decoder.

Geometry-aware Shape Guidance The generated point cloud provides a plausible estimation of the target shape, but it lacks voxel-level correspondence with the original image due to the limited supervision. To incorporate global shape constraints into the segmentation process, we encode M_i into a signed distance field (SDF) [6, 12] with a soft shape prior, enabling geometry awareness rather than voxel-wise supervision. We compute the SDF S_i of M_i, where each voxel's value represents the signed distance to the nearest boundary of the target object. The SDF is normalized to the range $[-1, 1]$. To encourage the segmentation model to align with the shape prior, we develop a geometry-aware loss:

$$\mathcal{L}_{\text{pc}} = \|S_i - \hat{S}_i\|^2 + \lambda_{\text{topo}}\mathcal{L}_{\text{topo}}(\hat{S}_i, S_i), \tag{2}$$

where S_i is the SDF computed from M_i and $\hat{S}_i$ is the predicted SDF from the model's segmentation output. The topology consistency loss $\mathcal{L}_{\text{topo}}$ denote the discrepancies of the number of connected components. λ_{topo} is a weighting factor balancing the two terms and we set it as 0.01 empirically.

Cross-View Supervision To effectively leverage multi-view information, we establish three independent branches, each responsible for processing slices along one of the three axes. Specifically, the segmentation models $\mathtt{F}^{(t)}$, $\mathtt{F}^{(s)}$, and $\mathtt{F}^{(c)}$ are trained on the transverse, sagittal, and coronal views, respectively. Take the example of model $\mathtt{F}^{(t)}$, we denote the supervised loss $\mathcal{L}_{sup}$:

$$\mathcal{L}_{\text{sup}} = \mathcal{L}_{\text{CE}}(\mathtt{F}^{(t)}(X_i), Y_i^{(t)}) + \mathcal{L}_{\text{Dice}}(\mathtt{F}^{(t)}(X_i), Y_i^{(t)}), \tag{3}$$

where $\mathcal{L}_{\text{CE}}$ is the cross-entropy loss and $\mathcal{L}_{\text{Dice}}$ is the Dice loss.

To enhance the 3D structural understanding of the model, we introduce a cross-view supervision mechanism, where predictions from two views provide guidance for training the third. At each training step, one of the three models is randomly selected as the student model to receive supervision from the other two models. Take the example of $\mathtt{F}^{(t)}$ as the student model. The regions where both models consistently predict the presence of the target structure are identified as the foreground, to supervise the student model's segmentation. The prediction fusion of the two models is $\tilde{Y}_i = \mathbb{I}\left(\mathtt{F}^{(s)}(X_i) = 1 \wedge \mathtt{F}^{(c)}(X_i) = 1\right)$, where $\mathbb{I}(\cdot)$ is an indicator function. The cross-view consistency loss is denoted as:

$$\mathcal{L}_{\text{cross}} = \|\mathtt{F}^{(t)}(X_i) - \tilde{Y}_i\|^2. \tag{4}$$

The total loss is $\mathcal{L} = \mathcal{L}_{\text{sup}} + \mathcal{L}_{\text{pc}} + \lambda_{\text{cross}}\mathcal{L}_{\text{cross}}$, where λ_{cross} acts as a time-dependent Gaussian weighting function.

Table 1. Comparison Result on Pancreas-CT dataset.

Method	Venue	Metrics			
		Dice (%) ↑	Jaccard (%) ↑	HD (voxel) ↓	ASD (voxel) ↓
3D2DCT [2]	MICCAI'23	45.34±17.71	31.00±14.83	30.47±12.68	8.87±6.04
SAMed [28]	Arxiv'23	45.53±11.08	30.13±9.19	39.34±9.74	15.23±4.34
SemiSAM [29]	Arxiv'23	40.52±8.69	25.77±6.61	47.61±8.68	18.77±3.47
H-SAM [5]	CVPR'24	47.64±12.61	32.15±10.65	38.66±11.49	14.07±5.18
CPC-SAM [17]	MICCAI'24	51.87±12.18	35.91±10.95	39.11±9.51	14.76±4.21
GA-SAM (Ours)	this paper	**67.12±9.74**	**51.28±10.56**	**27.51±12.76**	**8.47±4.70**

3 Experiments

3.1 Dataset and Implementation Details

Pancreas-CT Dataset We conduct our experiments on the NIH Pancreas-CT dataset [19], which consists of 82 contrast-enhanced 3D abdominal CT scans. Following previous work [20], we apply the same data splits and pre-processing strategies. All volumes are reshaped to [128, 128, 128] with linear interpolation.

Implementation Details All the experiments are implemented in PyTorch on the Tesla V100 and NVIDIA RTX A6000 GPUs. For point cloud completion, we train the point completion network with the pancreas data from MedShapeNet [11], which contains 745 pancreas shapes. It is trained by the Adam optimizer with an initial learning rate of 0.0001. For segmentation, we conduct all the experiments based on the "ViT-B" version of SAM, resize the positional embeddings [28] to [256, 256] and interpolate each slice as the same size. We adopt LoRA adaptation [8] and the rank of LoRA is set to 4 for efficiency and performance optimization. Following [28], we use the AdamW optimizer with weight decay set to 0.9, 0.999, and 0.1. The data augmentations include random flip, rotation, and pixel-wise transforms (Gaussian blur, brightness, contrast, gamma). The point completion and cross-view supervision are only in the training phase. For inference, we retain the segmentation model trained on transverse plane, which is often used in clinical diagnostics, and point completion network and segmentation models from other views are eventually discarded.

3.2 Comparison with SOTA methods

We compare our approach with several state-of-the-art methods on the Pancreas-CT dataset in Table 1. We employ Dice, Jaccard, the 95% Hausdorff Distance (HD), and the Average Surface distance (ASD) measurements to quantitatively evaluate the performance. First, we evaluate our method against the three-plane version of 3D2DCT [2], making it compatible with our Triple-Slice Annotation. We also compare against models fine-tuned from SAM, including SAMed [28]

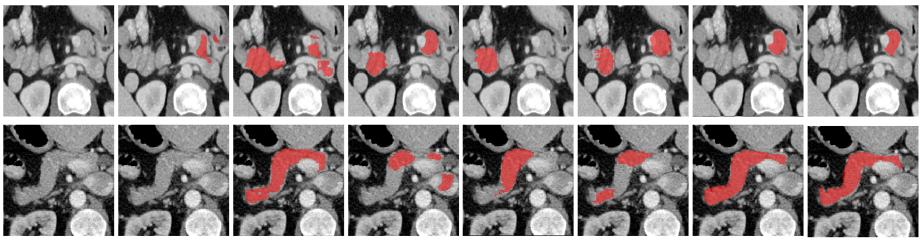

Fig. 3. Visual examples of segmentation results on Pancreas-CT dataset.

Table 2. Effectiveness of Each Component on Pancreas-CT dataset.

Method	Component		Metrics			
	$\mathcal{L}_{\text{cross}}$	$\mathcal{L}_{\text{pc}}$	Dice (%) ↑	Jaccard (%) ↑	HD (voxel) ↓	ASD (voxel) ↓
Baseline	✘	✘	45.53±11.08	30.13±9.19	39.34±9.74	15.23±4.34
w/o CV	✘	✔	51.80±10.56	35.62±9.45	40.10±10.05	14.46±4.31
w/o PC	✔	✘	59.56±10.25	43.12±9.85	36.65±13.45	11.94±5.01
Ours	✔	✔	**67.12±9.74**	**51.28±10.56**	**27.51±12.76**	**8.47±4.70**

and H-SAM [5]. Additionally, we evaluate SAM-based semi-supervised models, SemiSAM [29] and CPC-SAM [17], with the labeled and unlabeled slices as labeled and unlabeled samples, respectively. Our GA-SAM outperforms these SAM-based methods with significant improvements, demonstrating the benefits of our geometry-aware shape guidance techniques. The visual examples on the Pancreas-CT dataset are shown in Fig. 3, which demonstrates that our approach effectively maintains segmentation reliability under sparse annotation conditions. Compared to other methods, our framework produces more accurate boundaries and better capturing the anatomical shape and structure.

3.3 Ablation Study

To evaluate the effectiveness of different components of our GA-SAM framework, we conduct ablation experiments on removing key modules. Specifically, we investigate the impact of point cloud geometry-aware shape guidance $\mathcal{L}_{\text{pc}}$ and cross-view supervision $\mathcal{L}_{\text{cross}}$ by comparing three variants: the baseline without PC or CV, removing point cloud guidance (*i.e.*, w/o PC), and removing cross-view supervision (*i.e.*, w/o CV). As shown in Table 2, the results demonstrate that both PC and CV significantly contribute to performance improvements. The cross-view supervision bolsters the performance from 45.53% to 59%, demonstrating that it effectively leverages multi-view information. Furthermore, we observe that the removal of point cloud guidance leads to a significant drop in performance, highlighting the importance of shape priors.

4 Conclusion

In this work, we proposed a novel geometry-aware SAM adaptation framework GA-SAM that integrates point cloud completion to generate anatomical shapes under sparse annotations. The generated point cloud transformed into voxel representations, serves as a soft shape prior to guiding the adaptation of SAM through geometry-aware shape guidance. Additionally, we incorporated a cross-view supervision mechanism, ensuring segmentation reliability by enforcing multi-view consistency. The result of the Pancreas-CT dataset validates the effectiveness of our framework.

Acknowledgments. This work is supported by NSFC Project (62222604, 62206052), China Postdoctoral Science Foundation (2024M750424), Fundamental Research Funds for the Central Universities (020214380120, 020214380128), State Key Laboratory Fund (ZZKT2024A14, ZZKT2025B05), Postdoctoral Fellowship Program of CPSF (GZC20240252), Jiangsu Funding Program for Excellent Postdoctoral Talent (2024ZB 242) and Jiangsu Science and Technology Major Project (BG2024031).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Cai, H., Li, S., Qi, L., Yu, Q., Shi, Y., Gao, Y.: Orthogonal annotation benefits barely-supervised medical image segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3302–3311 (2023)
2. Cai, H., Qi, L., Yu, Q., Shi, Y., Gao, Y.: 3d medical image segmentation with sparse annotation via cross-teaching between 3d and 2d networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 614–624. Springer (2023)
3. Chen, C., Miao, J., Wu, D., Zhong, A., Yan, Z., Kim, S., Hu, J., Liu, Z., Sun, L., Li, X., et al.: Ma-sam: Modality-agnostic sam adaptation for 3d medical image segmentation. Medical Image Analysis **98**, 103310 (2024)
4. Chen, S., Lin, L., Cheng, P., Tang, X.: Aslseg: Adapting sam in the loop for semi-supervised liver tumor segmentation. In: IEEE International Symposium on Biomedical Imaging. pp. 1–5. IEEE (2024)
5. Cheng, Z., Wei, Q., Zhu, H., Wang, Y., Qu, L., Shao, W., Zhou, Y.: Unleashing the potential of sam for medical adaptation via hierarchical decoding. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3511–3522 (2024)
6. Dangi, S., Linte, C.A., Yaniv, Z.: A distance map regularized cnn for cardiac cine mr image segmentation. Medical physics **46**(12), 5637–5651 (2019)
7. Gowda, S.N., Clifton, D.A.: Cc-sam: Sam with cross-feature attention and context for ultrasound image segmentation. In: Proceedings of the European Conference on Computer Vision. Springer (2024)
8. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W.: LoRA: Low-rank adaptation of large language models. In: International Conference on Learning Representations (2022)

9. Khani, A., Asgari, S., Sanghi, A., Amiri, A.M., Hamarneh, G.: SLiMe: Segment like me. In: International Conference on Learning Representations (2024)
10. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4015–4026 (2023)
11. Li, J., Zhou, Z., Yang, J., Pepe, A., Gsaxner, C., Luijten, G., Qu, C., Zhang, T., Chen, X., Li, W., et al.: Medshapenet–a large-scale dataset of 3d medical shapes for computer vision. Biomedical Engineering/Biomedizinische Technik **70**(1), 71–90 (2025)
12. Li, S., Zhang, C., He, X.: Shape-aware semi-supervised 3d semantic segmentation for medical images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 552–561. Springer (2020)
13. Li, S., Cai, H., Qi, L., Yu, Q., Shi, Y., Gao, Y.: Pln: Parasitic-like network for barely supervised medical image segmentation. IEEE Transactions on Medical Imaging **42**(3), 582–593 (2022)
14. Lin, X., Xiang, Y., Zhang, L., Yang, X., Yan, Z., Yu, L.: SAMUS: Adapting segment anything model for clinically-friendly and generalizable ultrasound image segmentation. arXiv preprint arXiv:2309.06824 (2023)
15. Lu, W., Hong, Y., Yang, Y.: Up-sam: Uncertainty-informed adaptation of segment anything model for semi-supervised medical image segmentation. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine. pp. 2256–2261. IEEE (2024)
16. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**(1), 654 (2024)
17. Miao, J., Chen, C., Zhang, K., Chuai, J., Li, Q., Heng, P.A.: Cross prompting consistency with segment anything model for semi-supervised medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (2024)
18. Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., et al.: Sam 2: Segment anything in images and videos. arXiv preprint arXiv:2408.00714 (2024)
19. Roth, H.R., Lu, L., Farag, A., Shin, H.C., Liu, J., Turkbey, E.B., Summers, R.M.: Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 556–564. Springer (2015)
20. Shi, Y., Zhang, J., Ling, T., Lu, J., Zheng, Y., Yu, Q., Qi, L., Gao, Y.: Inconsistency-aware uncertainty estimation for semi-supervised medical image segmentation. IEEE Transactions on Medical Imaging **41**(3), 608–620 (2021)
21. Su, J., Shen, Z., Cao, P., Yang, J., Zaiane, O.R.: Self-paced sample selection for barely-supervised medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 582–592. Springer (2024)
22. Wang, X., Zhang, X., Cao, Y., Wang, W., Shen, C., Huang, T.: Seggpt: Towards segmenting everything in context. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1130–1140 (2023)
23. Wu, J., Fu, R., Fang, H., Liu, Y., Wang, Z., Xu, Y., Jin, Y., Arbel, T.: Medical SAM Adapter: Adapting segment anything model for medical image segmentation. arXiv preprint arXiv:2304.12620 (2023)

24. Wu, X., Wu, X., Luan, T., Bai, Y., Lai, Z., Yuan, J.: Fsc: Few-point shape completion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 26077–26087 (2024)
25. Wu, X., Xu, Z., Tong, R.K.y.: Few slices suffice: Multi-faceted consistency learning with active cross-annotation for barely-supervised 3d medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 286–296. Springer (2024)
26. Xu, Z., Chen, C., Lu, D., Sun, J., Wei, D., Zheng, Y., Li, Q., Tong, R.K.y.: Fmabs: Promptable foundation model drives active barely supervised learning for 3d medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 294–304. Springer (2024)
27. Yuan, W., Khot, T., Held, D., Mertz, C., Hebert, M.: Pcn: Point completion network. In: International Conference on 3D Vision. pp. 728–737. IEEE (2018)
28. Zhang, K., Liu, D.: Customized segment anything model for medical image segmentation. arXiv preprint arXiv:2304.13785 (2023)
29. Zhang, Y., Cheng, Y., Qi, Y.: Semisam: Exploring sam for enhancing semi-supervised medical image segmentation with extremely limited annotations. arXiv preprint arXiv:2312.06316 (2023)
30. Zou, X., Yang, J., Zhang, H., Li, F., Li, L., Wang, J., Wang, L., Gao, J., Lee, Y.J.: Segment everything everywhere all at once. Advances in Neural Information Processing Systems **36** (2023)