Superpixel-Guided Segment Anything Model for Liver Tumor Segmentation with Couinaud Segment Prompt

Fei Lyu[1], Jingwen Xu[1], Ye Zhu[1], Grace Lai-Hung Wong[2], and Pong C. Yuen[1]

[1] Department of Computer Science, Hong Kong Baptist University
{feilyu,csjwxu,csyzhu,pcyuen}@comp.hkbu.edu.hk
[2] Department of Medicine and Therapeutics, The Chinese University of Hong Kong

Abstract. The Segment Anything Model (SAM) is a powerful foundation model which has shown impressive performance for generic image segmentation. However, directly applying SAM to liver tumor segmentation presents challenges due to the domain gap between nature images and medical images, and the requirement of labor-intensive manual prompt generation. To address these challenges, we first investigate text promptable liver tumor segmentation by Couinaud segment, where Couinaud segment prompt can be automatically extracted from radiology reports to reduce massive manual efforts. Moreover, we propose a novel CouinaudSAM to adapt SAM for liver tumor segmentation. Specifically, we achieve this by: 1) a superpixel-guided prompt generation approach to effectively transform Couinaud segment prompt into SAM-acceptable point prompt; and 2) a difficulty-aware prompt sampling strategy to make model training more effective and efficient. Experimental results on the public liver tumor segmentation dataset demonstrate that our method outperforms the other state-of-the-art methods.

Keywords: Liver Tumor Segmentation · Couinaud Segment · SAM

1 Introduction

Liver cancer is one of the leading causes of death and has brought heavy global burden [10]. Computed tomography (CT) is a commonly used imaging tool for liver condition assessment, and liver tumor segmentation from CT images is a prerequisite for cancer diagnosis and treatment planning. Radiologists suffer from heavy workload to manually delineate liver tumors in clinical practice. Therefore, it is necessary to develop automatic methods for accurate liver tumor segmentation. Deep learning based methods have achieved great progress in liver tumor segmentation. First, many advanced convolutional neural network (CNN) architectures have reported impressive results, such as H-DenseUNet [19] and nnU-Net [15]. Then, the Transformer architecture brings substantial performance gains by introducing self-attention mechanism [18]. Recently, foundation models have caused a dramatic revolution to artificial intelligence (AI) model development, which are trained on massive datasets and adapted to diverse

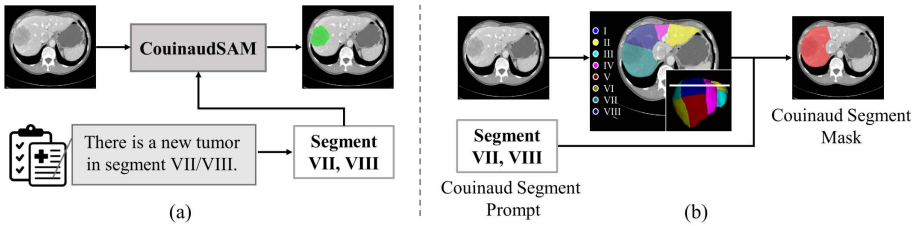

Fig. 1: (a) Text promptable segmentation by Couinaud segment: Couinaud segment prompt is a special type of text prompt, which can be automatically extracted from radiology reports with less manual effort. (b) Anatomical meaning of Couinaud segment: Liver can be divided into eight Couinaud segments based on hepatic veins, and each segment refers to a mask on the CT image.

downstream tasks [21]. The Segment Anything Model (SAM) is a pioneering foundation model for image segmentation [17], and SAM shows unprecedented performance by prompting it with points or bounding boxes. Given its merit, SAM has also been investigated to address the critical medical image segmentation tasks. However, there exists some challenges that hinder the application of SAM in liver tumor segmentation. First, SAM shows poor performance due to the significant domain gap between natural images and medical images. Second, SAM relies on manual point-or-box prompts to obtain acceptable performance, which is still labor-intensive. Therefore, how to adapt SAM to the medical domain without manually input prompts is an important question.

There is a growing interest to adapt SAM for medical image segmentation. To incorporate medical knowledge into SAM, many approaches propose to collect large datasets for fine-tuning SAM [20], [7]. Considering the large scale of SAM, parameter-efficient transfer learning (PETL) approaches are employed for efficient adaptation by optimizing a small amount of model parameters, such as LoRA used in SAMed [13], [28], FacT used in MA-SAM [16], [4], and Adapter used in MSA [12] , [27]. In the context of SAM prompts, some works have investigated the impact of different prompts, and demonstrated that suitable prompt is non-trivial for the segmentation results of SAM [6], [8], [9]. Moreover, there exists some attempts of auto-prompting to turn SAM into a fully automatic manner. However, it either requires additional localization frameworks to generate bounding box prompts [30], or ignores the prompt encoder, which substantially reduces the power of SAM [28].

In this paper, we investigate text promptable liver tumor segmentation by Couinaud segment, and propose a novel framework, namely CouinaudSAM, to address the above challenges. Liver can be divided into eight parts based on hepatic veins and each part is called a Couinaud segment. Couinaud segment is often used by radiologists to describe the localization of liver tumors in radiology reports, and it can be automatically extracted with existing text mining techniques. Couinaud segment prompt can be seen as a special type of text prompt, which describes the location of tumors based on anatomy knowledge.

Many methods have been developed to accurately segment the liver into Couinaud segments [29], [26], and it is straightforward to turn the Couinaud segment prompt into a Couinaud segment mask, as shown in Fig. 1. In CouinaudSAM, we first propose a superpixel-guided prompt generation module to obtain point prompts from the Couinaud segment mask. Superpixel is a group of pixels with similar characteristics, and the point prompt generation is on the superpixel level rather than on the pixel level. Consequently, it can help reduce the risk of missing small tumors and also benefit the model training with more informative point prompts. Then, we propose a difficulty-aware prompt sampling strategy for training CouinaudSAM. In contrast to random selection, our main idea is that the massive candidate points contain a large number of simple samples and a small number of hard samples, and sampling those hard ones can make model training more effective.

Our main contributions are summarized as follows: (1) We introduce a new type of text prompt using Couinaud segment, aiming to avoid the laborious manual prompt creation. (2) We propose a novel approach CouinaudSAM to take full advantage of the Couinaud segment prompt and make it adaptable to SAM. (3) Experimental results on the liver tumor segmentation dataset demonstrate that our method can achieve better performance than other state-of-the-art methods.

2 Method

Fig. 2 illustrates the proposed CouinaudSAM. SAM is a promptable segmentation model, which contains three main components: an image encoder based on pre-trained Vision Transformer (ViT) [11], a prompt encoder and a mask decoder. For fine-tuning the image encoder, we employ a state-of-the-art PETL technique of FacT [16]. We freeze the image encoder and only update the Fact layers, prompt encoder and mask decoder, to reduce the number of trainable parameters. SAM can accept sparse prompts including point, box or text. Couinaud segment prompt can be seen as a special type of text prompt, however, it is essentially a number from 1 to 8, which may not generate informative prompt embedding using general text encoder. Therefore, CouinaudSAM is proposed to make Couinaud segment prompt adaptable to SAM.

2.1 Superpixel-Guided Prompt Generation

Couinaud segment prompt cs represents a specific region on the CT image, and a Couinaud segment mask m can be obtained by adopting any existing Couinaud segmentation methods [29], [26]. The following question is how to generate SAM-acceptable prompts from the Couinaud segment mask. Considering that superpixel can divide a region to sub-regions based on connectivity and similarity, we can use superpixel to guide the transformation from mask to point or box. In our method, we use the popular Simple Linear Iterative Clustering (SLIC) [1] algorithm to generate the superpixels inside the liver region, and take the centroid of each superpixel as candidate points $\mathcal{SP} = \{sp_i\}_{i=1}^{N}$. For point

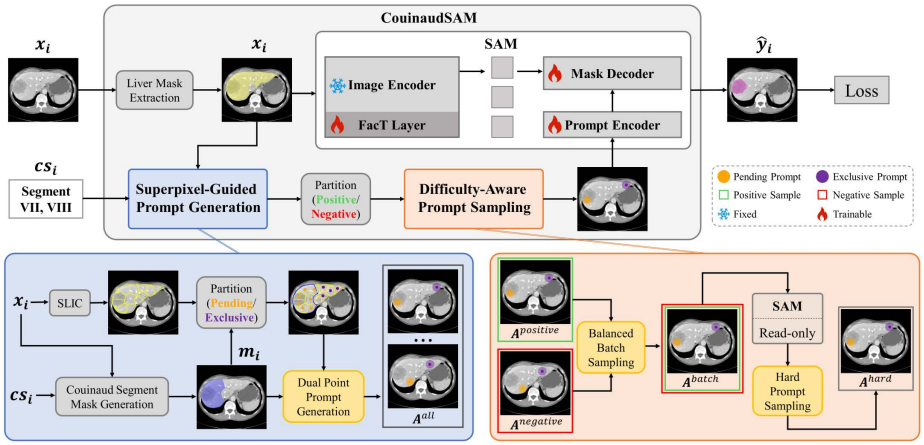

Fig. 2: Overview of the proposed CouinaudSAM. It contains two main parts: superpixel-guided prompt generation and difficulty-aware prompt sampling.

prompt in SAM, its label should be given to indicate whether it is foreground or background. However, the label information is not available without manual decision. One naive solution is to take all candidate points as foreground, and ensemble the results from them. To better utilize the candidate points, dual point prompt generation is introduced.

Dual Point Prompt Generation. Inspired by the observation that combining positive points with negative points shows substantial benefits [24], we can partition $\mathcal{SP}$ into two parts based on its overlap with the Couinaud segment mask: **Pending** points $\mathcal{SP}^p = \{sp_i\}_{i=1}^{N^p}$ within m which are possible to be tumor, and **Exclusive** points $\mathcal{SP}^e = \{sp_i\}_{i=1}^{N^e}$ outside m which are definitely not tumor. Unlike the original SAM, the label of a point is **pending** or **exclusive** instead of foreground or background in our method, since it can be directly obtained based on the Couinaud segment prompt. Consequently, we can generate a set of dual point prompts as: $A^{all} = \{sp_i^p, sp_i^e\}_{i=1}^{N^p}$, where $sp_i^p \in \mathcal{SP}^p, sp_i^e \in \mathcal{SP}^e$. For each pair, sp_i^p is sequentially sampled from $\mathcal{SP}^p$, while sp_i^e is randomly sampled from $\mathcal{SP}^e$ to form the pair.

2.2 Difficulty-Aware Prompt Sampling

Considering that random sampling from massive candidate points for each batch during training is sub-optimal, we propose a difficulty-aware prompt sampling strategy to make model training more effective and efficient.

Balanced Batch Sampling. Different from natural image segmentation, tumors normally occupy a small portion of the whole image. Therefore, most points

lie in the healthy liver region, and it is highly possible that no points lie in the tumor region using random sampling in a batch. To build a balanced batch, we first partition all dual point prompts A^{all} into two parts based on the overlap with the tumor mask, including a positive set A^{pos} and a negative set A^{neg}. Then we sample k^{batch} prompts from A^{pos} and A^{neg} with a balanced ratio to get A^{batch} for each batch during training. The sampling probability of each prompt in A^{all} is represented as follows:

$$p_i^{sample} = \begin{cases} \tau/size(A^{pos}), & \text{if } A_i^{all} \in A^{pos} \\ (1-\tau)/size(A^{neg}), & \text{if } A_i^{all} \in A^{neg} \end{cases}, \tag{1}$$

where τ is the balanced ratio for batch sampling.

Hard Prompt Sampling. After balancing the positive-negative ratio of each batch to shift the attention to the point prompts in the tumor region, we propose to further improve the training performance with hard prompt sampling. Inspired by online hard example mining [23], our proposed hard prompt sampling proceeds as follows. We first do a forward pass with the input image and all point prompts in A^{batch}. Prediction results can be obtained for each prompt, and we can perform evaluation by calculating the difference between the output and the ground truth. The evaluation results are used for ranking, and then the top $k^{hard}(k^{hard} < k^{batch})$ prompts can be selected to construct A^{hard} and used for backward pass. Hard prompts from A^{hard} force the model to pay more attention to false predictions, thus improving its discriminative ability. Moreover, since only a part of prompts in A^{batch} are selected for updating the model, the computation cost of the backward pass is less expensive than using all prompts.

2.3 Training and Inference

Training During training, for each input image x_i, its Couinaud segment prompt cs_i is provided. A set of hard prompts A^{hard} can be generated from cs_i using the above method. For each $prompt_i$ in A^{hard}, the predicted segmentation result $\hat{y}_i$ is obtained by sending x_i and $prompt_i$ to CouinaudSAM: $\hat{y}_i = CouinaudSAM(x_i, prompt_i)$. The ground truth y_i is used for supervision, and the training loss is a linear combination of binary cross-entropy ($\mathcal{L}_{BCE}$) and dice loss functions ($\mathcal{L}_{DICE}$), which is represented by:

$$\mathcal{L} = (1-\lambda) \cdot \mathcal{L}_{BCE}(\hat{y}_i,\ y_i) + \lambda \cdot \mathcal{L}_{DICE}(\hat{y}_i,\ y_i), \tag{2}$$

where λ represents the balanced weight for the loss items.

Inference During inference, given a test image x_i^t and its Couinaud segment prompt cs_i^t, N^t dual point prompts can be generated from cs_i^t with superpixel-guided prompt generation, and each prompt $prompt_{ij}$ can predict a segmentation result $\hat{y}_{ij}^t$. The final prediction result $\hat{y}_i^t$ is the ensemble of all prediction results,

and it can be further refined by removing the parts outside the Couinaud segment mask m_i:

$$\hat{y}_i^t = m_i \odot \sum_{j=1}^{N^t} \hat{y}_{ij}^t. \tag{3}$$

3 Experiments

3.1 Experimental Setup

Dataset and Metric We evaluate our method using the public liver tumor segmentation dataset from Medical Segmentation Decathlon Task08 (MSD08) [2]. Tumor masks of 303 CT volumes are provided, and 161 of them are manually annotated with pixel-wise Couinaud segment label (1∼8) from [25]. We first divide MSD08 into two subsets, *i.e.* MSD08−161 and MSD08−142. Couinaud segment labels for MSD08−142 are not available, and we use a pre-trained model based on [25] to generate the labels. Then, Couinaud segment prompts for each subset are obtained based on the overlap between tumor masks and Couinaud segment labels. MSD08−161 is divided into 113 training volumes, 16 validation volumes and 32 test volumes. MSD08−142 contains 142 CT volumes, which are all used for testing. Following [2], Dice similarity coefficient (DSC) and Normalized surface Dice (NSD) are used as the evaluation metrics.

Implementation Details During pre-processing, we first truncate the image intensity values of all CT slices to the range [-200, +250], and then normalized to [0, 255]. Liver mask is extracted using the pre-trained model provided in [15]. The implementation is based on PyTorch 1.10.0 using an NVIDIA A100 GPU. Similar to SAMed [28], we use AdamW optimizer with a warmup training strategy, where the initial learning rate is set to 0.001 and the warmup period if set to 250. The model is trained for 100 epochs with a batch size of 12. We adopt FacT to fine-tune the image encoder, and the rank is set to 32. The balanced sampling ratio is set to $\tau = 0.6$, the sampling numbers are set to $k^{batch} = 8$ and $k^{hard} = 4$ separately, and the balanced weight for the loss is set to $\lambda = 0.8$.

3.2 Self Analysis of CouinaudSAM

Evaluation of each component. We first evaluate CouinaudSAM with different variants using the ViT-B backbone on the MSD08−161 test set. Fig. 3 (a) shows that DSC is increased by 2.13% and NSD is increased by 3.19% when Couinaud segment prompt is added, where the false positives outside the corresponding Couinaud segment mask are removed. To better utilize the Couinaud segment prompt, our proposed CouinaudSAM with superpixel-guided prompt generation and difficulty-aware prompt sampling can further improve the performance, where DSC is increased to 72.98% and NSD is increased to 61.41%. The results demonstrate that Couinaud segment prompt can bring useful location information, and CouinaudSAM can make full use of this newly introduced prompt for producing better liver tumor segmentation results.

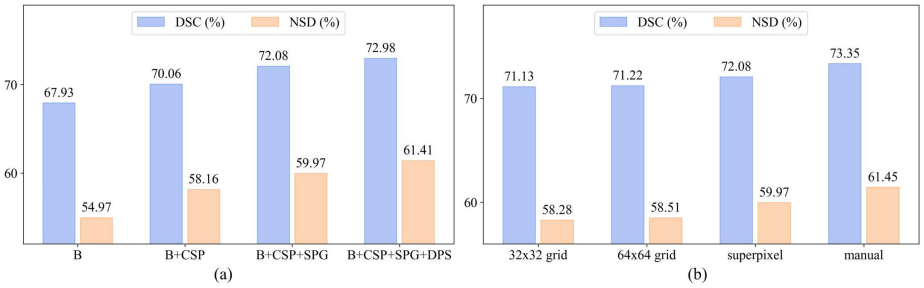

Fig. 3: Self analysis of CouinaudSAM. (a) Evaluation of each component. **B**: Baseline FacT fine-tuning without prompt; **CSP**: Couinaud segment prompt; **SPG**: Superpixel-guided prompt generation; **DPS**: Difficulty-aware prompt sampling. (b) Comparison of different prompt generation approaches.

Comparison of different prompt generation approaches. This part evaluates different prompt generation approaches. Original SAM [17] provides a solution for automatic mask generation, where point prompts are generated from a regular grid. Fig. 3 (b) shows that using a regular grid of 64x64 points does not show obvious difference compared to a 32x32 one. Note that the results are refined within the Couinaud segment mask. When applying our proposed superpixel-guided prompt generation approach, the performance is improved. After combining the difficulty-aware prompt sampling strategy, the segmentation performance is close to the one using manually input point prompts. The regular grid based point generation approach may miss small tumors since the grid divides the image equally. In comparison, our method divides the image into sub-regions by considering local characteristics, and tumors are regarded as different clusters from the background, which are less possible be ignored after iterating over all superpixels. As a result, we could achieve better performance by employing the superpixel-guided prompt generation approach. Moreover, when the resolution of the grid is higher, it takes longer inference time, brings higher false positive rate and benefits more from the refinement of the Couinaud segment mask. Compared to the manual approach, our method shows inferior results, but also greatly reduces the manual efforts of point prompt generation, which is the target of this work.

3.3 Comparison with State-of-the-Arts

Table 1 presents the quantitative comparison between CouinaudSAM and other state-of-the-art methods, and the qualitative comparison results are shown in Fig. 4. CNN based methods include Att-UNet [22], UNet++ [31] and nnU-Net [15], while Transformer based methods include TransUnet [5], MissFormer [14] and DAEFormer [3]. Though these methods do not accept prompts as input, the results are refined within the Couinaud segment mask derived from the prompt. For CouinaudSAM, we report the results with different backbones, *i.e.*,

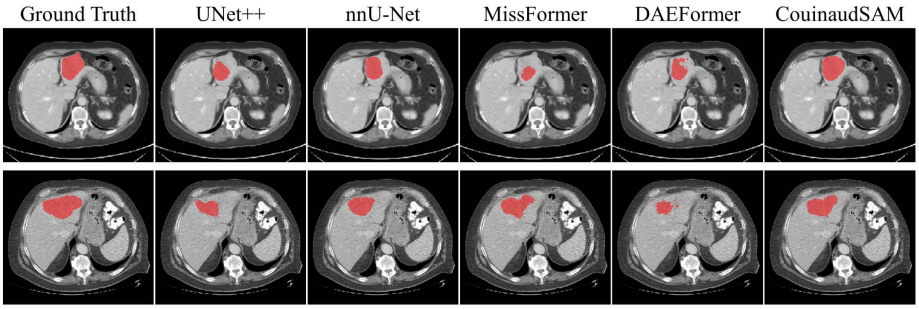

Fig. 4: Qualitative visualization of segmentation results generated from our proposed CouinaudSAM and other state-of-the-art methods.

Table 1: Comparison between CouinaudSAM and other state-of-the-art methods.

Methods	MSD08−161				MSD08−142			
	Slice Prompt		Volume Prompt		Slice Prompt		Volume Prompt	
	DSC(%)	NSD(%)	DSC(%)	NSD(%)	DSC(%)	NSD(%)	DSC(%)	NSD(%)
Att-UNet [22]	67.81	56.68	66.81	55.25	62.70	70.11	60.99	67.55
UNet++ [31]	68.64	56.23	67.83	55.01	64.50	71.19	62.82	68.70
nnU-Net [15]	68.19	56.03	67.26	54.89	64.78	71.24	63.66	69.81
TransUnet [5]	67.83	54.37	67.17	53.41	61.48	69.13	59.99	66.92
MissFormer [14]	69.85	56.95	68.62	55.19	65.27	72.14	63.20	68.76
DAEFormer [3]	70.49	58.06	69.40	56.47	66.26	73.83	64.03	70.15
CouinaudSAM^b	72.98	61.41	70.95	58.47	70.82	76.81	66.69	69.30
CouinaudSAM^l	72.98	62.99	71.21	60.04	72.30	78.99	68.74	71.23
CouinaudSAM^h	**73.52**	**63.02**	**71.64**	**60.12**	**72.50**	**79.40**	**69.75**	**74.17**

* CouinaudSAM^b: ViT-B version; CouinaudSAM^l: ViT-L version; CouinaudSAM^h: ViT-H version.

ViT-B, ViT-L and ViT-H. It can be observed that CouinaudSAM^b shows obvious superiority, and the performance of CouinaudSAM is further improved when the model size increases. Moreover, Couinaud segment prompt can be given in two formats, slice-level and volume-level. For volume-level prompt, all slices in one CT volume share the same prompt. For slice-level prompt, slices without tumor are given prompt '0' and slices with tumor are given their Couinaud segment location information. The performance of volume-level prompt is inferior while its cost is much lower compared to slice-level prompt. Moreover, we cannot directly compare pre-trained SAM variants, because the test dataset is already involved in their training, which may lead to unfair comparison. To summarize, CouinaudSAM can achieve superior performance compared to other state-of-the-art methods for liver tumor segmentation, and it has high potential for enhanced performance when more powerful foundation models are utilized as the backbone.

4 Conclusion

In this paper, we present CouinaudSAM, a novel approach to adapt SAM for text promptable liver tumor segmentation. Specifically, we first propose a new type of text prompt using Couinaud segment, and it can be automatically extracted from radiology reports to avoid the labor-intensive manual prompt creation process. Couinaud segment prompt is then transformed into SAM-acceptable prompts with the superpixel-guided prompt generation method. In addition, we propose a difficulty-aware prompt sampling strategy to make model training more effective and efficient. Extensive experimental results have verified the effectiveness of CouinaudSAM. More importantly, CouinaudSAM can also be extended to other organs that can be divided into sub-regions, such as lung with 5 lobes [26].

Acknowledgments. This work was supported by Hong Kong Research Grants Council General Research Fund under Grant RGC/HKBU12200122.

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

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence **34**(11), 2274–2282 (2012)
2. Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., Ronneberger, O., Summers, R.M., et al.: The medical segmentation decathlon. Nature communications **13**(1), 4128 (2022)
3. Azad, R., Arimond, R., Aghdam, E.K., Kazerouni, A., Merhof, D.: Dae-former: Dual attention-guided efficient transformer for medical image segmentation. In: International Workshop on PRedictive Intelligence In MEdicine. pp. 83–95. Springer (2023)
4. Chen, C., Miao, J., Wu, D., Yan, Z., Kim, S., Hu, J., Zhong, A., Liu, Z., Sun, L., Li, X., et al.: Ma-sam: Modality-agnostic sam adaptation for 3d medical image segmentation. arXiv preprint arXiv:2309.08842 (2023)
5. Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y.: Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)
6. Cheng, D., Qin, Z., Jiang, Z., Zhang, S., Lao, Q., Li, K.: Sam on medical images: A comprehensive study on three prompt modes. arXiv preprint arXiv:2305.00035 (2023)
7. Cheng, J., Ye, J., Deng, Z., Chen, J., Li, T., Wang, H., Su, Y., Huang, Z., Chen, J., Jiang, L., et al.: Sam-med2d. arXiv preprint arXiv:2308.16184 (2023)
8. Dai, H., Ma, C., Liu, Z., Li, Y., Shu, P., Wei, X., Zhao, L., Wu, Z., Zhu, D., Liu, W., et al.: Samaug: Point prompt augmentation for segment anything model. arXiv preprint arXiv:2307.01187 (2023)
9. Deng, G., Zou, K., Ren, K., Wang, M., Yuan, X., Ying, S., Fu, H.: Sam-u: Multi-box prompts triggered uncertainty estimation for reliable sam in medical image. arXiv preprint arXiv:2307.04973 (2023)

10. Devarbhavi, H., Asrani, S.K., Arab, J.P., Nartey, Y.A., Pose, E., Kamath, P.S.: Global burden of liver disease: 2023 update. Journal of Hepatology (2023)
11. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al.: An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations (2021)
12. Houlsby, N., Giurgiu, A., Jastrzebski, S., Morrone, B., De Laroussilhe, Q., Gesmundo, A., Attariyan, M., Gelly, S.: Parameter-efficient transfer learning for nlp. In: International Conference on Machine Learning. pp. 2790–2799. PMLR (2019)
13. 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. International Conference on Learning Representations (2022)
14. Huang, X., Deng, Z., Li, D., Yuan, X., Fu, Y.: Missformer: An effective transformer for 2d medical image segmentation. IEEE Transactions on Medical Imaging **42**(5), 1484–1494 (2023)
15. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)
16. Jie, S., Deng, Z.H.: Fact: Factor-tuning for lightweight adaptation on vision transformer. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 37, pp. 1060–1068 (2023)
17. 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. Proceedings of the IEEE International Conference on Computer Vision p. 4015–4026 (2023)
18. Li, R., Xu, L., Xie, K., Song, J., Ma, X., Chang, L., Yan, Q.: Dht-net: Dynamic hierarchical transformer network for liver and tumor segmentation. IEEE Journal of Biomedical and Health Informatics (2023)
19. Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A.: H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes. IEEE transactions on medical imaging **37**(12), 2663–2674 (2018)
20. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**(1), 654 (2024)
21. Moor, M., Banerjee, O., Abad, Z.S.H., Krumholz, H.M., Leskovec, J., Topol, E.J., Rajpurkar, P.: Foundation models for generalist medical artificial intelligence. Nature **616**(7956), 259–265 (2023)
22. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al.: Attention u-net: Learning where to look for the pancreas. Medical Imaging with Deep Learning (2018)
23. Shrivastava, A., Gupta, A., Girshick, R.: Training region-based object detectors with online hard example mining. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 761–769 (2016)
24. Stein, J., Di Folco, M., Schnabel, J.A.: Influence of prompting strategies on segment anything model (sam) for short-axis cardiac mri segmentation. arXiv preprint arXiv:2312.08932 (2023)
25. Tian, J., Liu, L., Shi, Z., Xu, F.: Automatic couinaud segmentation from ct volumes on liver using glc-unet. In: International Workshop on Machine Learning in Medical Imaging. pp. 274–282. Springer (2019)
26. Tian, Y., Qin, W., Xue, F., Lambo, R., Yue, M., Diao, S., Yu, L., Xie, Y., Cao, H., Li, S.: Arr-gcn: Anatomy-relation reasoning graph convolutional network for automatic fine-grained segmentation of organ's surgical anatomy. IEEE Journal of Biomedical and Health Informatics (2023)

27. 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)
28. Zhang, K., Liu, D.: Customized segment anything model for medical image segmentation. arXiv preprint arXiv:2304.13785 (2023)
29. Zhang, X., Liu, Y., Ali, S., Zhao, X., Sun, M., Han, M., Liu, T., Zhai, P., Cui, Z., Zhang, P., et al.: Anatomical-aware point-voxel network for couinaud segmentation in liver ct. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 465–474. Springer (2023)
30. Zhang, Y., Hu, S., Jiang, C., Cheng, Y., Qi, Y.: Segment anything model with uncertainty rectification for auto-prompting medical image segmentation. arXiv preprint arXiv:2311.10529 (2023)
31. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE transactions on medical imaging **39**(6), 1856–1867 (2019)