Segment Anything in Medical Images

Yu Jiening
Yu Jiening@umac.mo

arXiv preprint arXiv, 2023.7

Segment Anything in Medical Images

Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang

Abstract

Medical image segmentation is a critical component in clinical practice, facilitating accurate diagnosis, treatment planning, and disease monitoring. However, current methods predominantly rely on customized models, which exhibit limited generality across diverse tasks. In this study, we present MedSAM, the inaugural foundation model designed for universal medical image segmentation. Harnessing the power of a meticulously curated dataset comprising over one million images, MedSAM not only outperforms existing state-of-the-art segmentation foundation models, but also exhibits comparable or even superior performance to specialist models. Moreover, MedSAM enables the precise extraction of essential biomarkers for tumor burden quantification. By delivering accurate and efficient segmentation across a wide spectrum of tasks, MedSAM holds significant potential to expedite the evolution of diagnostic tools and the personalization of treatment plans.

INTRODUCTION

Segmentation is a fundamental task in medical imaging analysis, which involves identifying and delineating regions of interest (ROI) in various medical images, such as organs, lesions, and tissues. Accurate segmentation is essential for many clinical applications, including disease diagnosis, treatment planning, and monitoring of disease progression [1], [2]. Manual segmentation has long been the gold standard for delineating anatomical structures and pathological regions, but this process is time-consuming, labor-intensive, and often requires a high degree of expertise. Semi- or fully-automatic segmentation methods can significantly reduce the time and labor required, increase consistency, and enable the analysis of large-scale datasets.

Deep learning-based models have shown great promise in medical image segmentation due to their ability to learn intricate image features and deliver accurate segmentation results across a diverse range of tasks, from segmenting specific anatomical structures to identifying pathological regions [3]. However, a significant limitation of many current medical image segmentation models is their task-specific nature. These models are typically designed and trained for a specific segmentation task, and their performance can degrade significantly when applied to new tasks or different types of imaging data. This lack of generality poses a substantial obstacle to the wider application of these models in clinical practice. In contrast, recent advances in the field of natural image segmentation have witnessed the emergence of segmentation foundation models [4], [5], showcasing remarkable versatility and performance across various segmentation tasks. However, their application to medical image segmentation has been challenging due to the substantial domain gap [6] (Supplementary Related work).

Therefore, there is a growing demand for universal models in medical image segmentation: models that can be trained once and then applied to a wide range of segmentation tasks. Such models would not only exhibit heightened versatility in terms of model capacity, but also potentially lead to more consistent results across different tasks, benefiting from a shared underlying architecture and training process. Motivated by the remarkable generality of the Segment Anything Model (SAM) [4], we introduce MedSAM, the first foundation model for universal medical image segmentation. MedSAM is adapted from the SAM model on an unprecedented scale, with more than one million medical image-mask pairs. We thoroughly evaluate MedSAM through comprehensive experiments on over 70 internal validation tasks and 40 external validation tasks, spanning a variety of anatomical structures, pathological conditions, and medical imaging modalities. Experimental results demonstrate that MedSAM consistently outperforms the state-of-the-art (SOTA) segmentation foundation model, while achieving performance on par with, or even surpassing specialist models. These results highlight the

Catalogue

1. Introduction
2. Methods
3. Results
4. Discussion

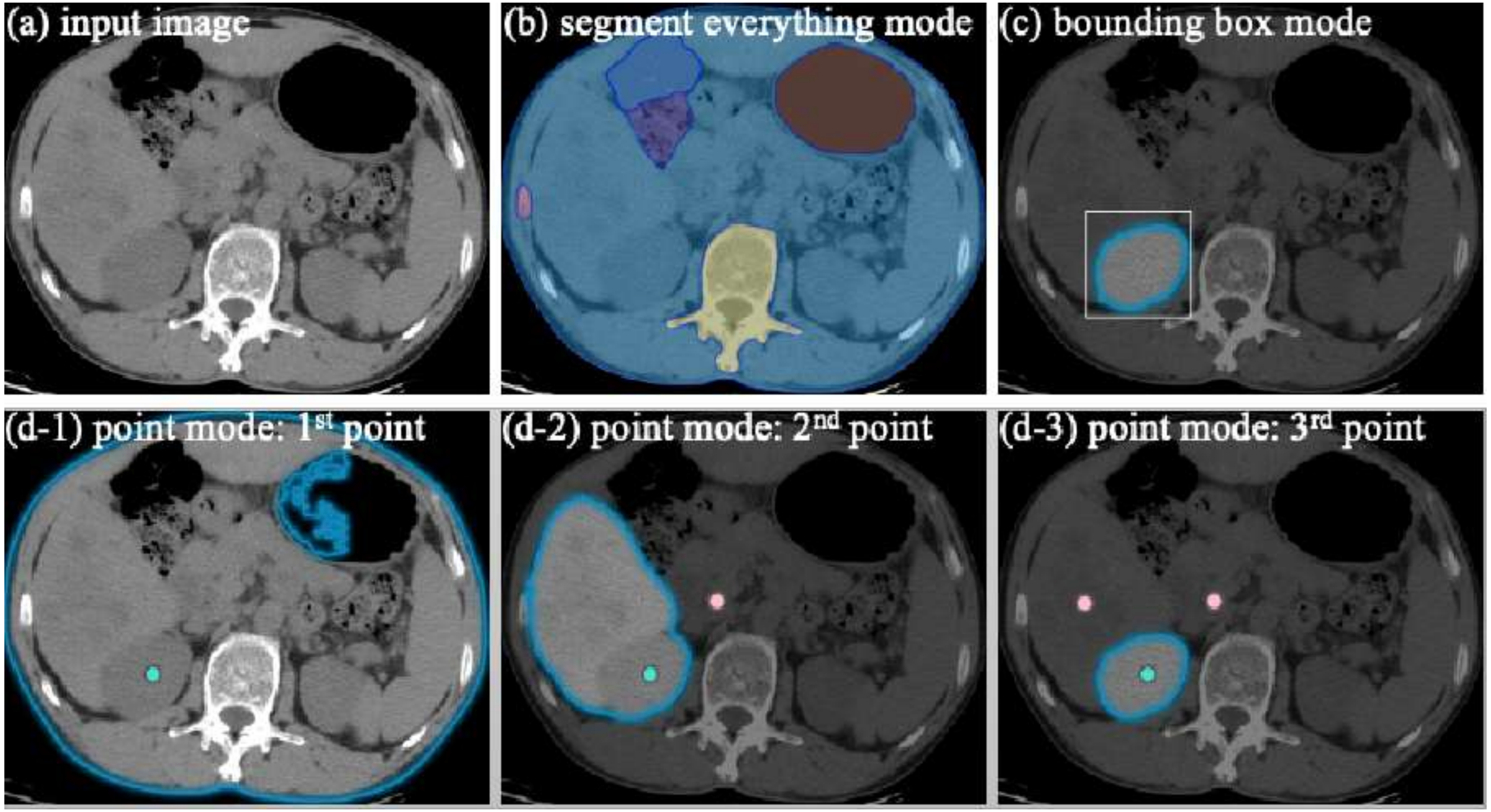
1.Introduction

- Segmentation is a fundamental task in medical image analysis
- Manual segmentation is time-consuming and labour-intensive, while semi-automatic or fully automatic segmentation methods can reduce the labour force.
- Medical image segmentation models can often only be used for specific segmentation tasks.
- **MedSAM** was introduced and evaluated through 70 internal and 40 external tasks.

2.Methods

2.1 Study design

- **Limitations** of existing medical image segmentation methods: **only for specific tasks**.
- Research objective of this paper: to develop a robust segmentation base model that can be widely applied to **different tasks**.

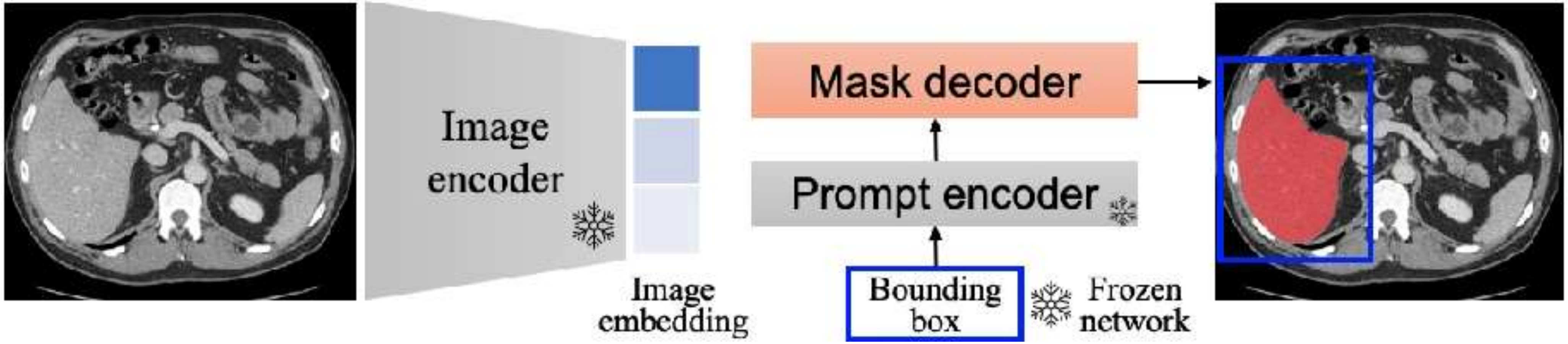


SAM supports three main segmentation modes: fully automatic segmentation mode, bounding box mode and point mode. The figure below shows the segmentation results of SAM under different Prompts for abdominal CT(腹部).

2.2 Dataset curation and pre-processing

- A large and diverse dataset containing 33 segmentation tasks was collated, including various segmentation targets.
- Perform **intensity normalisation** on all images.
- **For the CT images**, the intensity values were cropped to the range of [-500,1000].
- **For the other images**, the intensity values were clipped to the range between the 0.95th and 99.5th percentiles.
- Normalise all intensity values to the range [0,255] and resize the image to a uniform size of 256 x 256 x 3.

2.3 Network architecture



- Adoption of ViT model as an image encoder
- To reduce the computation, the image encoder and cue encoder were **frozen** and only the mask decoder was fine-tuned.

2.4 Training protocol and experimental setting

- Each dataset was randomly divided into 80%、 10%and 10% for training ,tuning and validation.
- Since SAM is designed for **2D** image segmentation, we divide 3D images (i.e., CT, MR, PET) into **2D slices** along out-of-plane dimensions.

2.5 Evaluation metrics

- DSC(Dice Similarity Coefficient) is a **region-based segmentation metric**

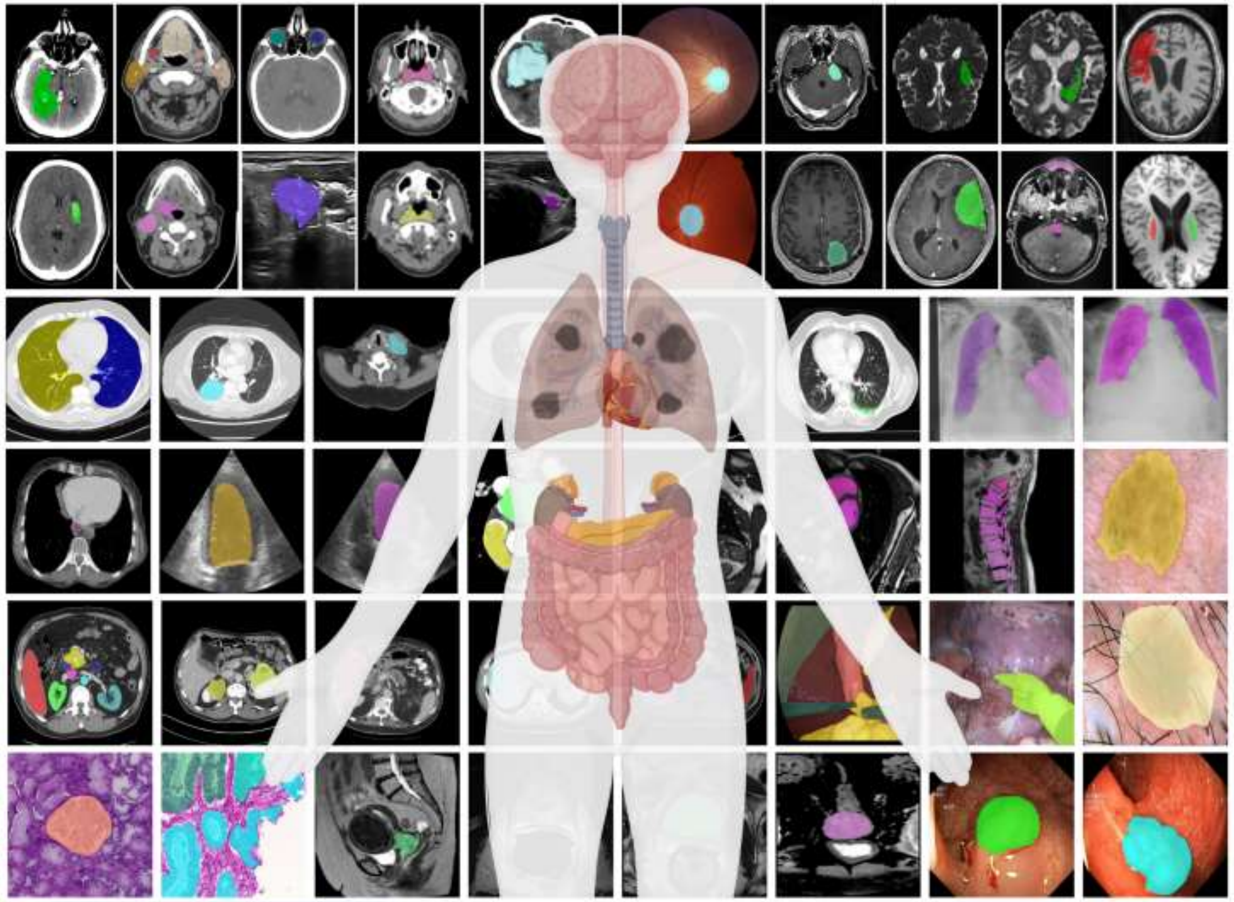
$$DSC(G,S) = \frac{2|G \cap S|}{|G| + |S|},$$

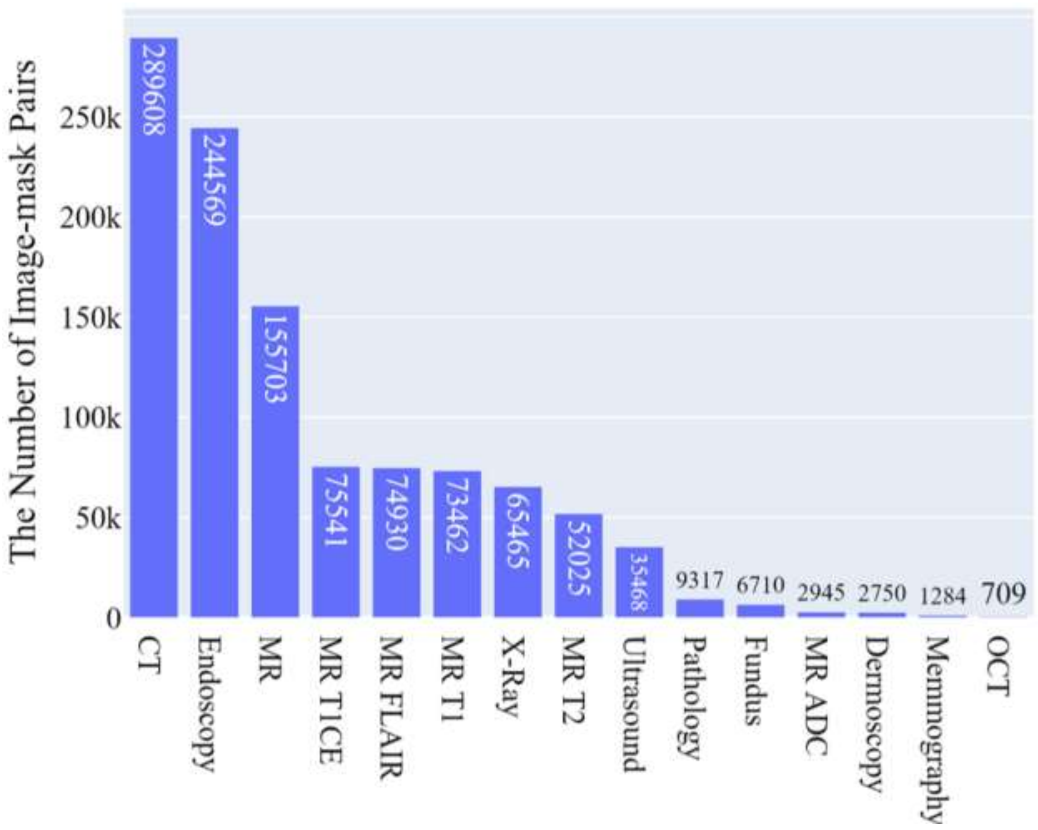
- NSD(Normalized Surface Distance) is a **boundary-based metric**

$$NSD(G,S) = \frac{|\partial G \cap B_{\partial S}^{(\tau)}| + |\partial S \cap B_{\partial G}^{(\tau)}|}{|\partial G| + |\partial S|},$$

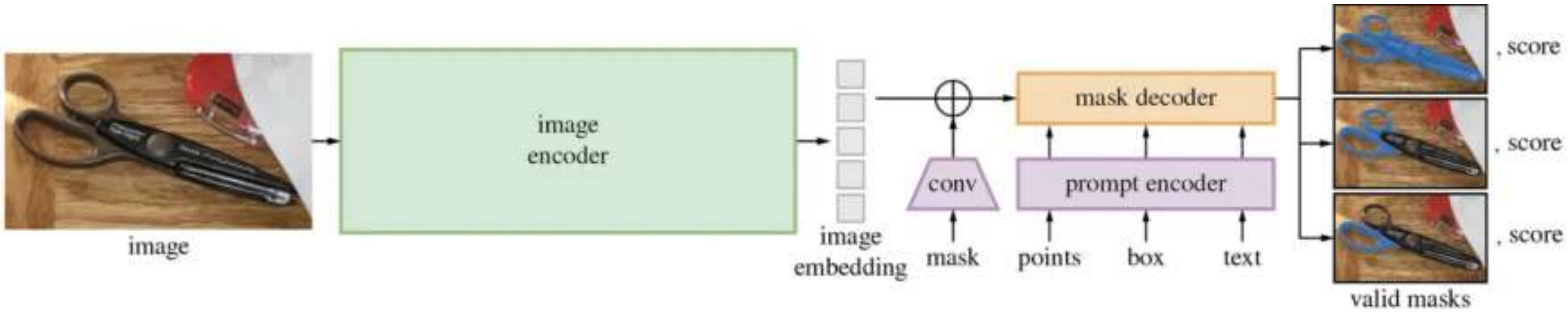
3.Results

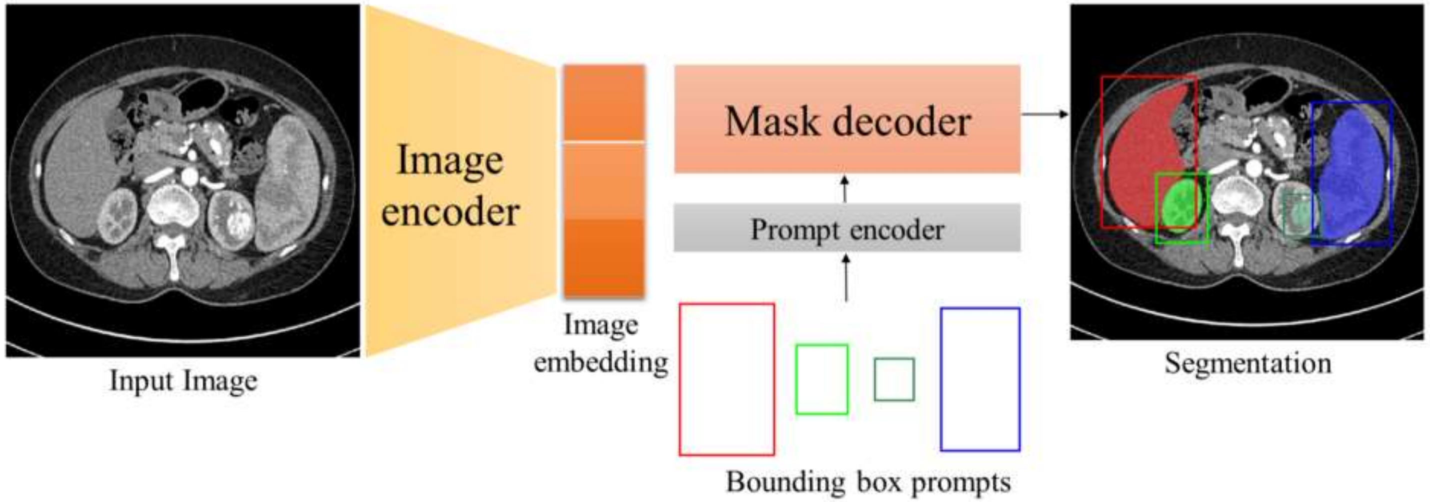
- Curated a diverse and large-scale medical image segmentation dataset with 1,090,486 medical image-mask pairs, covering 15 imaging modalities, over 30 cancer types, and a multitude of imaging protocols





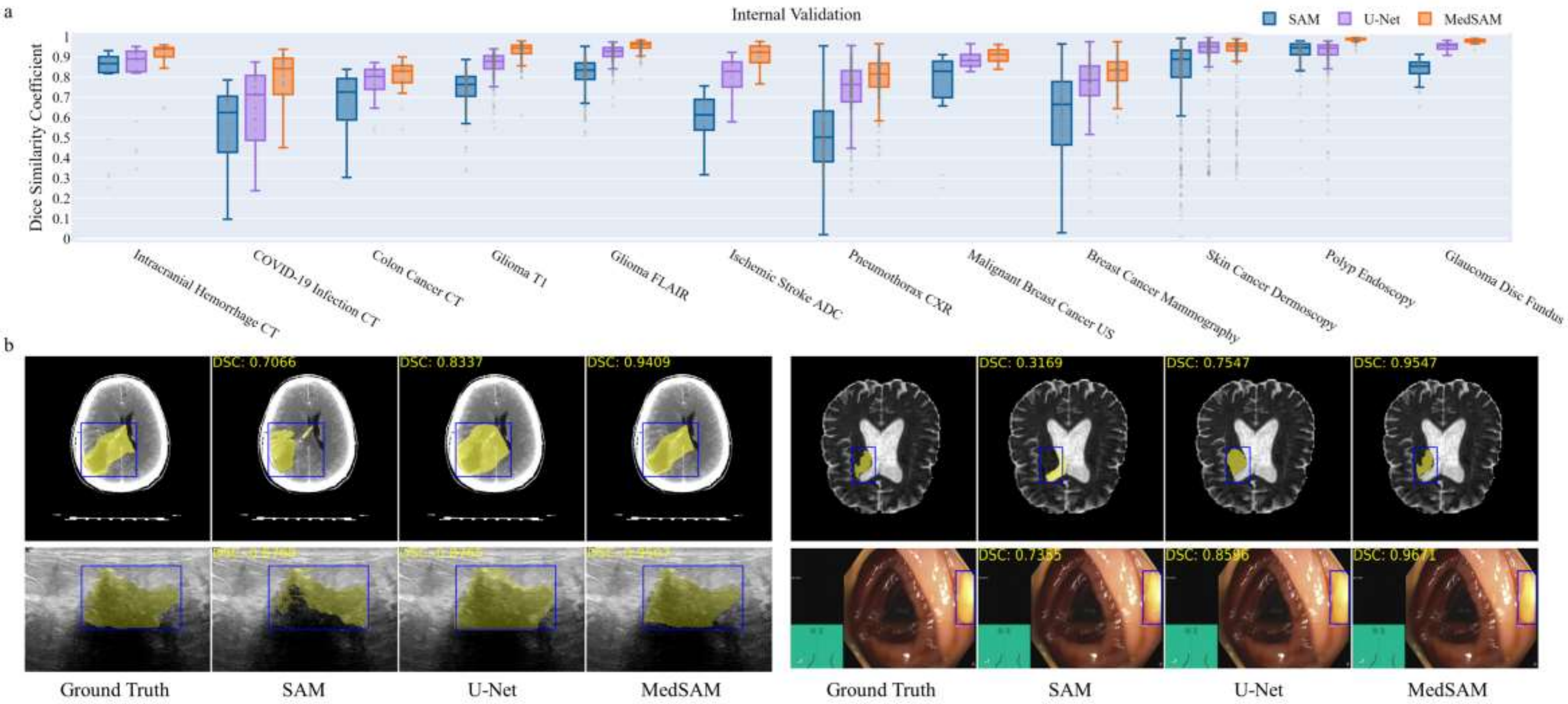
- Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and endoscopy(内窥镜检查) are the dominant modalities.





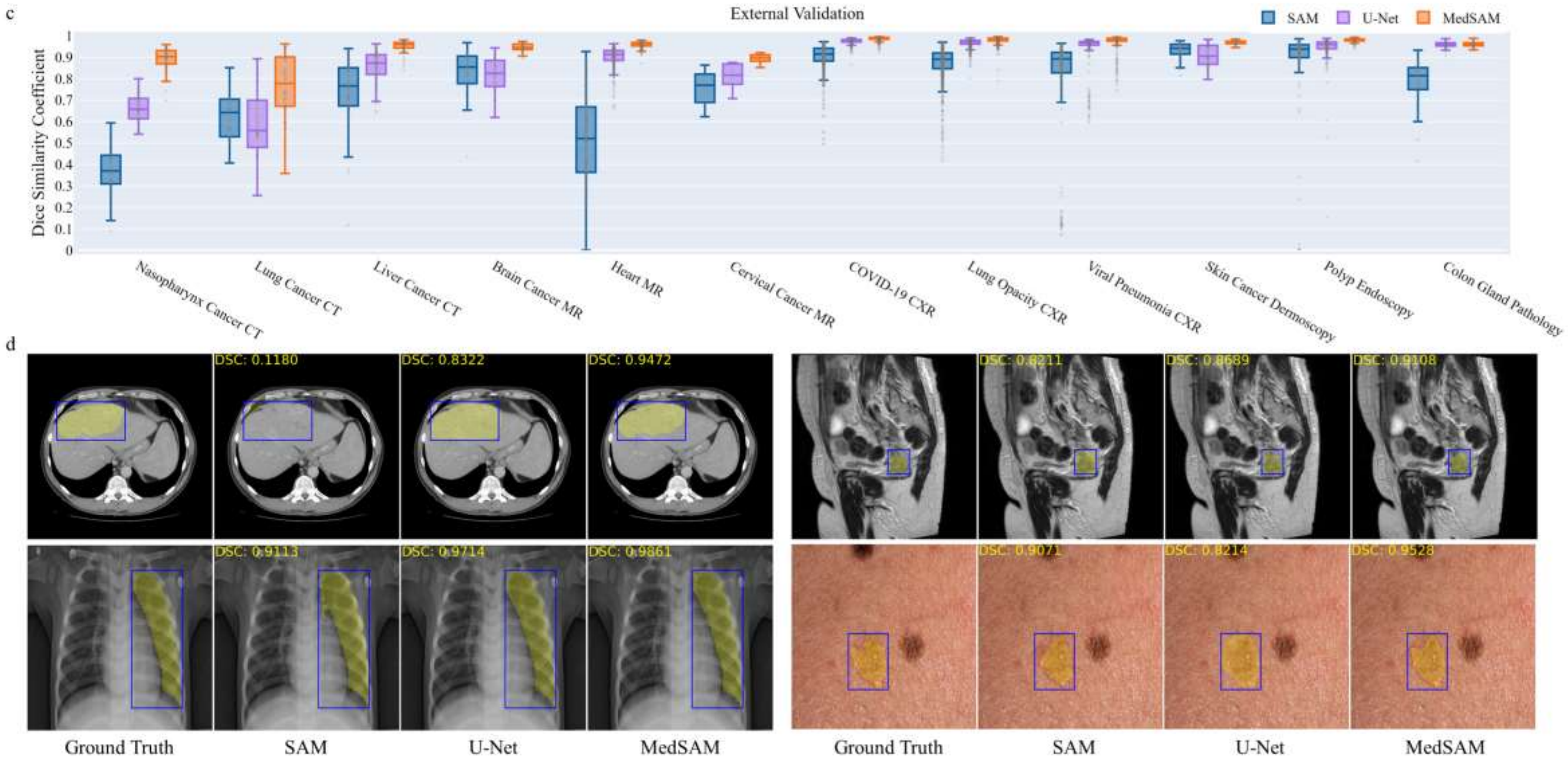
- follows the SAM network architecture, consisting of a image encoder, prompt encoder and mask decoder.

Internal validation results for 12 representative segmentation tasks

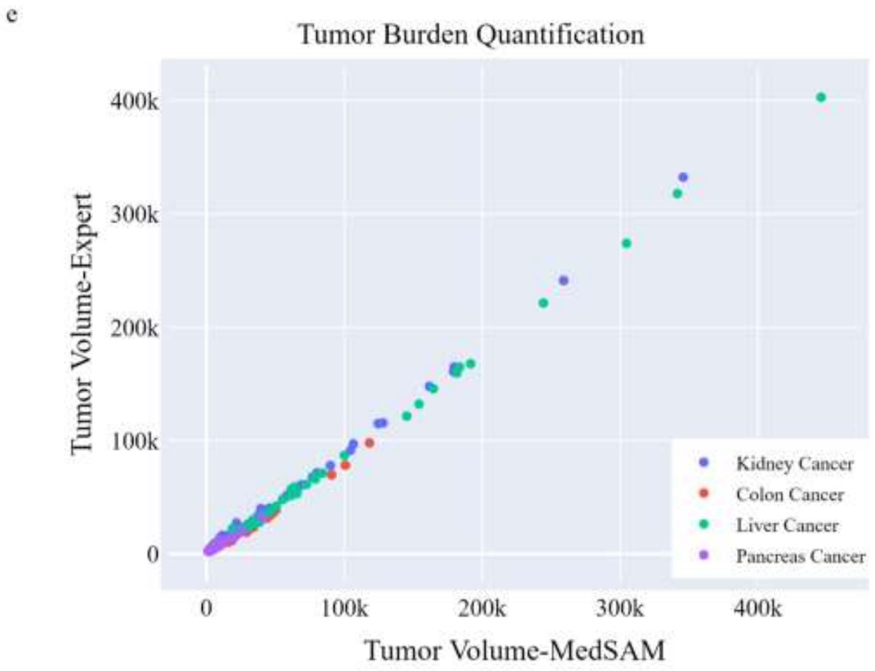


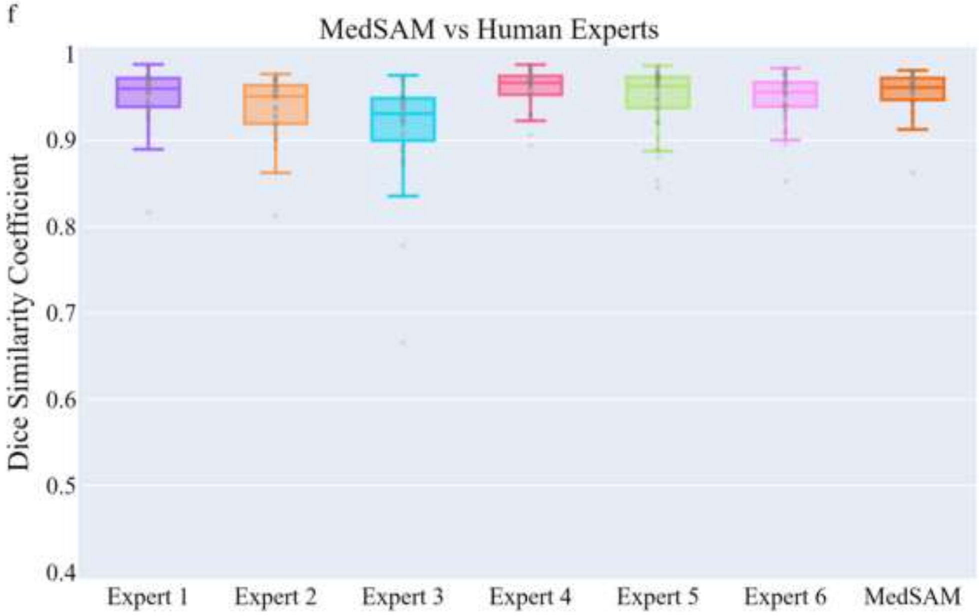
- SAM **performs poorly** on most CT, MR segmentation tasks
- MedSAM outperforms SAM and U-net

External validation results for 12 representative segmentation tasks



- MedSAM outperforms SAM and U-net
- MedSAM's performance on new datasets is highly generalisable





- MedSAM can be used for precise tumor burden quantification.
- MedSAM can obtain comparable or even better segmentation accuracy compared to human experts.

Performance comparison between MedSAM and SAM on 21 3D medicaimage segmentation tasks.

Segmentation Target	Modality	DSC (%)			NSD (%)		
		MedSAM	SAM	Improve	MedSAM	SAM	Improve
Brain Ventricles	MR-T1	74.82	41.96	**32.86**	78.17	31.26	**46.91**
Brain Ventricles	MR-T2	72.87	39.56	**33.31**	75.01	31.39	**43.62**
Brain Tumor	MR-FLAIR	89.15	74.00	**15.15**	76.13	38.27	**37.86**
Cerebellum	MR-T1	93.31	83.25	**10.06**	82.39	44.55	**37.84**
Cerebellum	MR-T2	90.78	81.88	**8.90**	70.01	38.84	**31.17**
Gallbladder	MR	77.78	61.97	**15.81**	76.36	36.34	**40.02**
Left Ventricle	MR	88.91	68.44	**20.48**	91.05	55.73	**35.32**
Right Ventricle	MR	85.92	72.11	**13.80**	88.85	68.98	**19.87**
Liver	MR	93.90	80.38	**13.52**	80.13	33.00	**47.13**
Pancreas	MR	80.07	51.11	**28.96**	79.42	32.81	**46.61**
Prostate	MR-ADC	92.25	79.61	**12.65**	92.72	60.12	**32.59**
Prostate	MR-T2	92.18	79.39	**12.78**	92.00	57.60	**34.40**
Abdomen Tumor	CT	65.54	42.86	**22.68**	64.99	34.49	**30.50**
Gallbladder	CT	84.36	47.28	**37.08**	87.07	30.48	**56.59**
Head-Neck Tumor	CT	68.29	23.87	**44.41**	47.36	23.88	**23.48**
Liver	CT	91.42	74.21	**17.21**	76.30	26.08	**50.21**
Lung Infections	CT	60.01	32.54	**27.47**	58.89	25.84	**33.04**
Pancreas	CT	76.76	43.53	**33.23**	81.09	32.38	**48.71**
Pleural Effusion	CT	59.46	9.52	**49.94**	75.77	11.19	**64.59**
Stomach	CT	82.66	68.95	**13.72**	70.93	35.49	**35.44**
Head-Neck Tumor	PET	81.17	72.45	**8.72**	62.76	38.57	**24.18**
Average		81.04	58.52	**22.51**	76.57	37.51	**39.05**

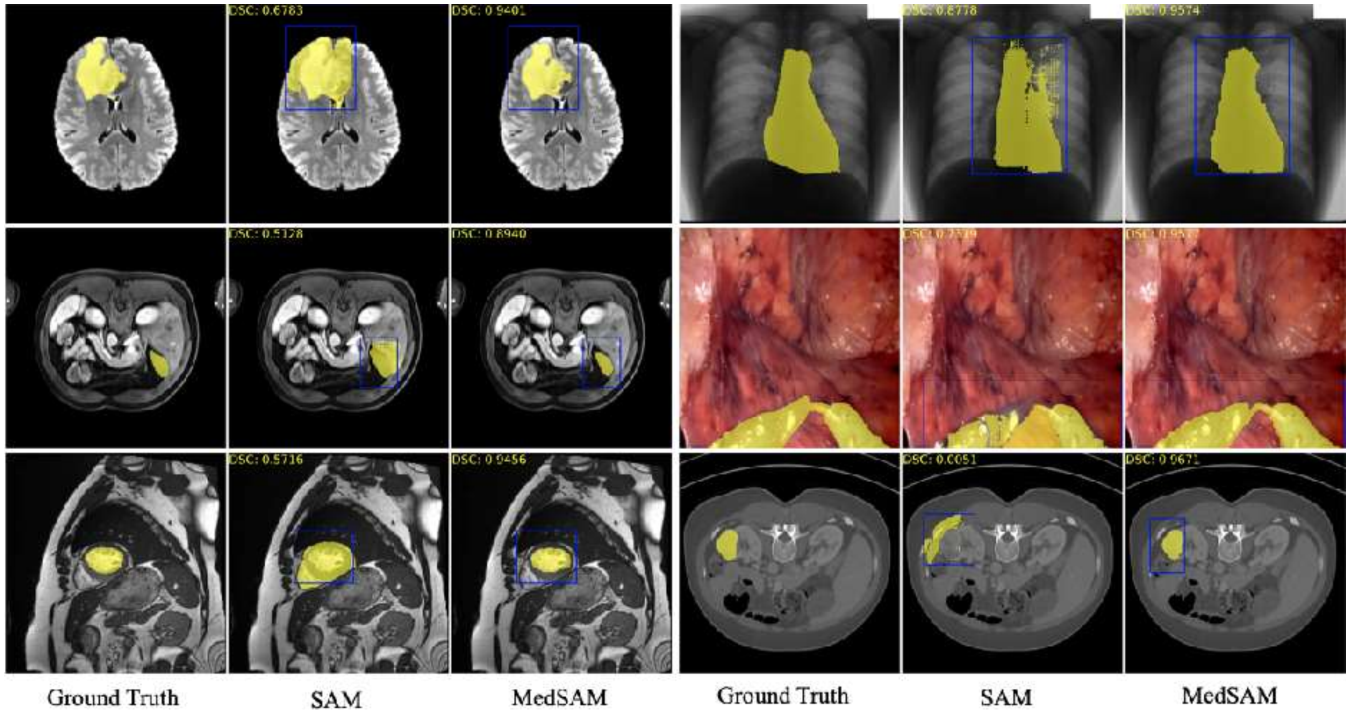
- MedSAM achieves signifcant and consistent improvementsacross all the tasks.

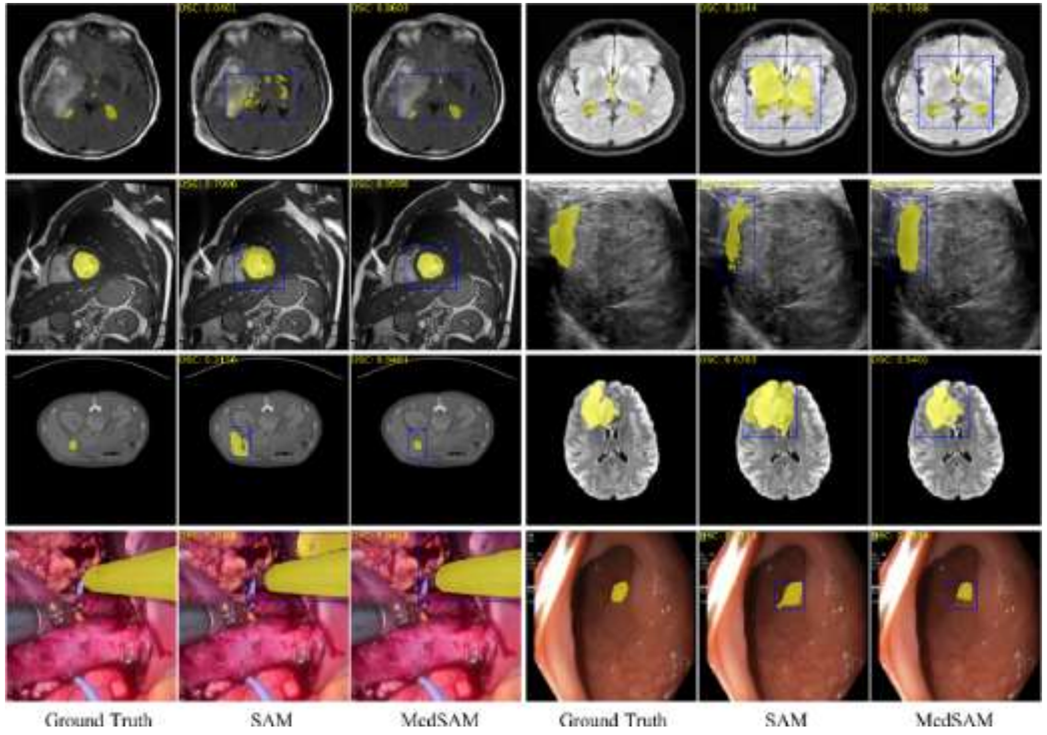
Performance comparison between MedSAM and SAM on nine 2D medicaimage segmentation tasks

Segmentation Target	Modality	DSC (%)			NSD (%)		
		MedSAM	SAM	Improve	MedSAM	SAM	Improve
Breast Tumor	Ultrasound	85.42	78.01	**7.41**	89.02	82.48	**6.54**
Liver	Ultrasound	74.36	67.81	**6.55**	79.07	72.07	**7.01**
Vessel	Ultrasound	70.88	57.60	**13.28**	78.41	65.10	**13.31**
Heart	X-Ray	91.19	79.28	**11.91**	94.10	83.85	**10.25**
Lungs	X-Ray	96.57	72.24	**24.33**	98.56	75.45	**23.11**
Polyp	Endoscope	86.90	81.60	**5.30**	90.91	85.93	**4.99**
Instrument	Endoscope	86.37	76.61	**9.76**	90.93	82.36	**8.57**
Retinal Vessel	Retinal Image	66.10	0.75	**65.35**	84.40	3.45	**80.95**
Gland	Pathology	37.23	22.63	**14.60**	43.09	27.75	**15.33**
Average		77.22	59.62	**17.61**	83.17	64.27	**18.89**

MedSAM achieves significant and consistent improvementsacross all the tasks.

More examples of the pre-trained SAM and MedSAM on different 3D and 2Lmedical image segmentation tasks.





- MedSAM is much better with the same prompts

4.Discussion

- MedSAM can be used as a tool for generic medical image segmentation.
- Data imbalance in the training set, which could potentially affect MedSAM.

Thank You!