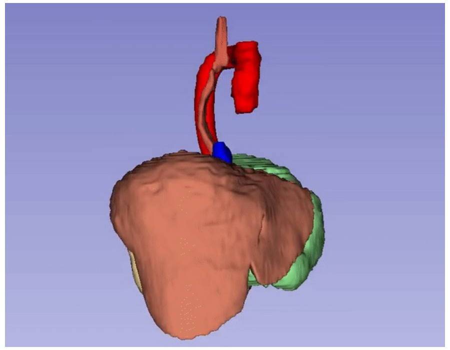
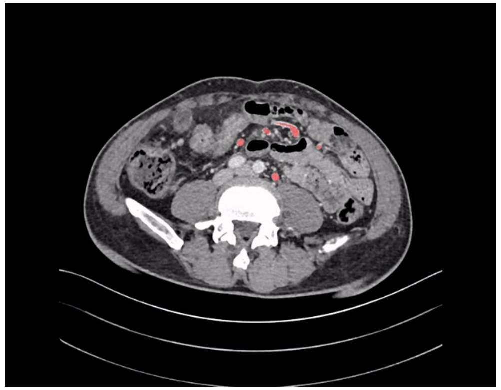
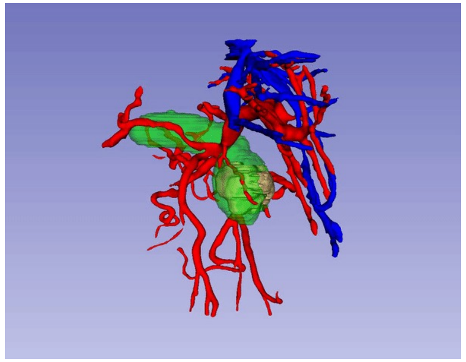
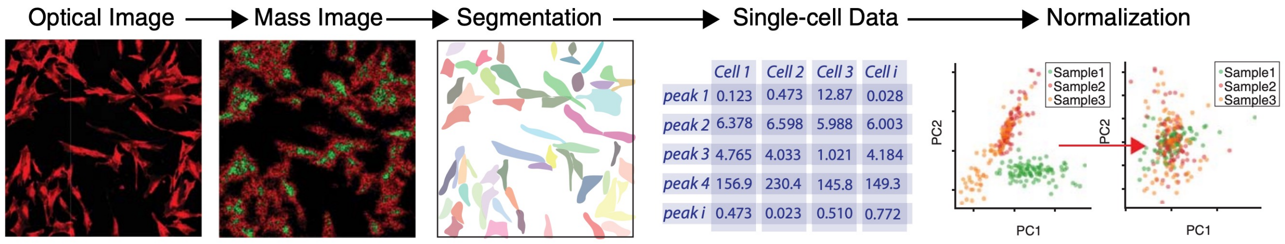
Segment Anything In Medical Images

Jun Ma (Bo WangLab)
University of Toronto
University Health Network
Vector Institute

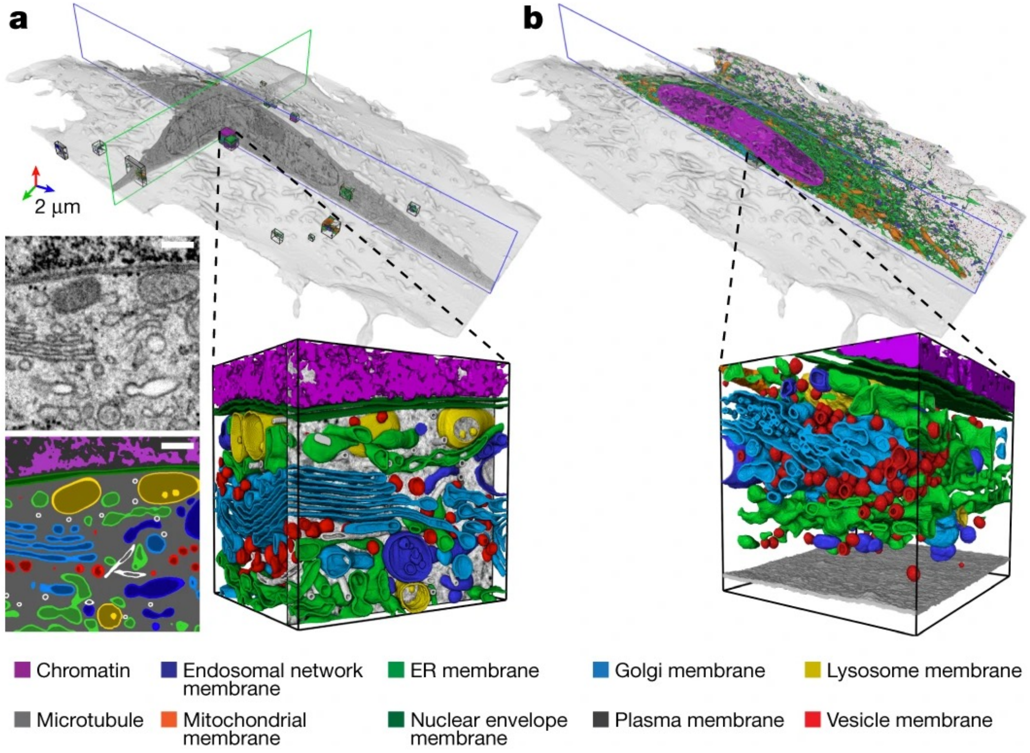
Biomedical Image Segmentation: What and Why

- ➢ Quantification of anatomical structures and disease progression
- ➢ Cancer microenvironment analysis



Capolupo, L et al. Science, 2022

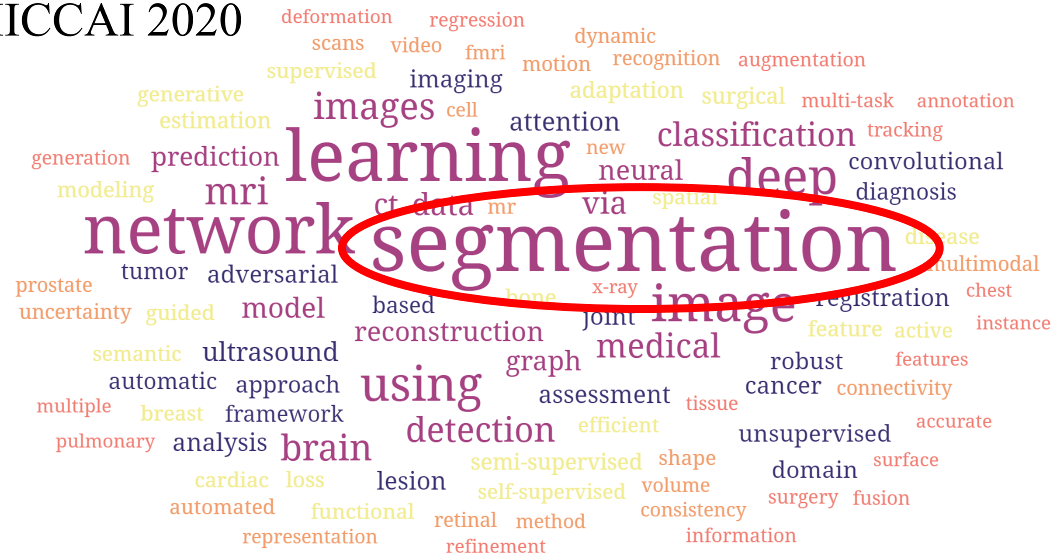


Heinrich, L. et al. Nature, 2021

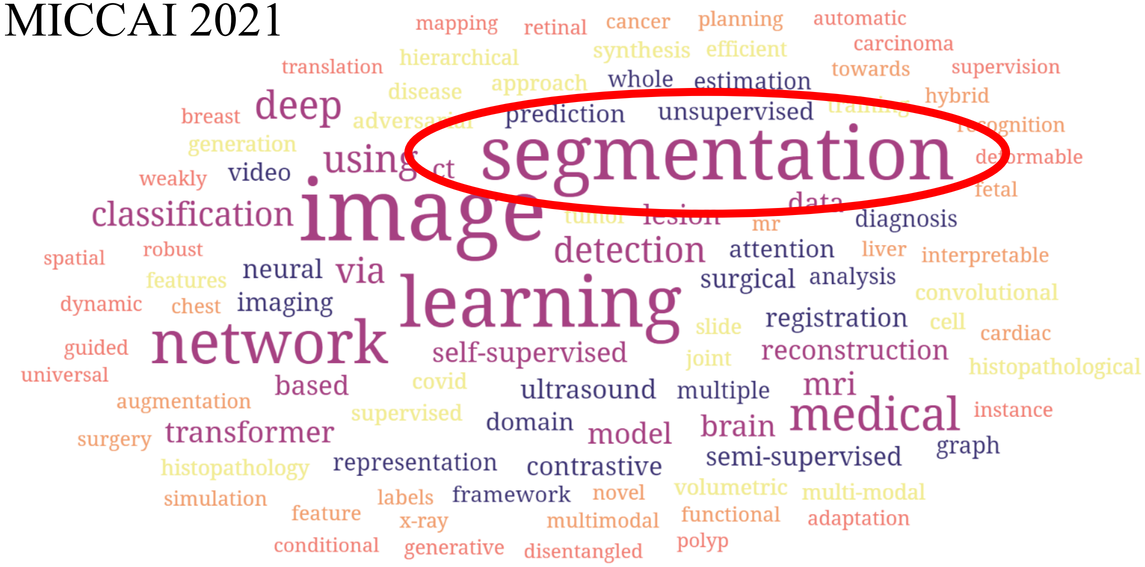
Segmentation is the core technology towards precise biomedical image analysis!

Biomedical Image Segmentation is Still an Active Research Field!

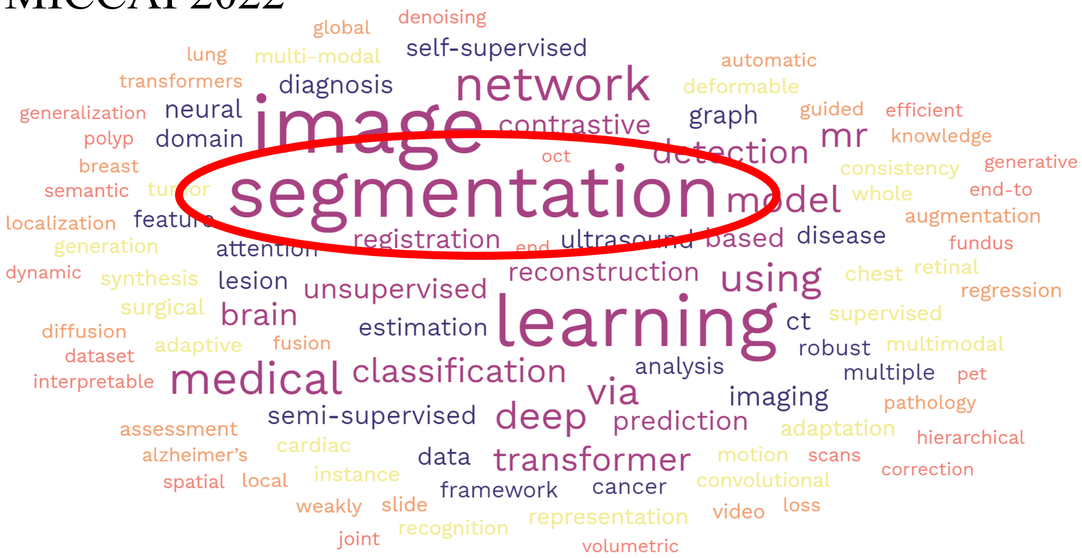
MICCAI 2020



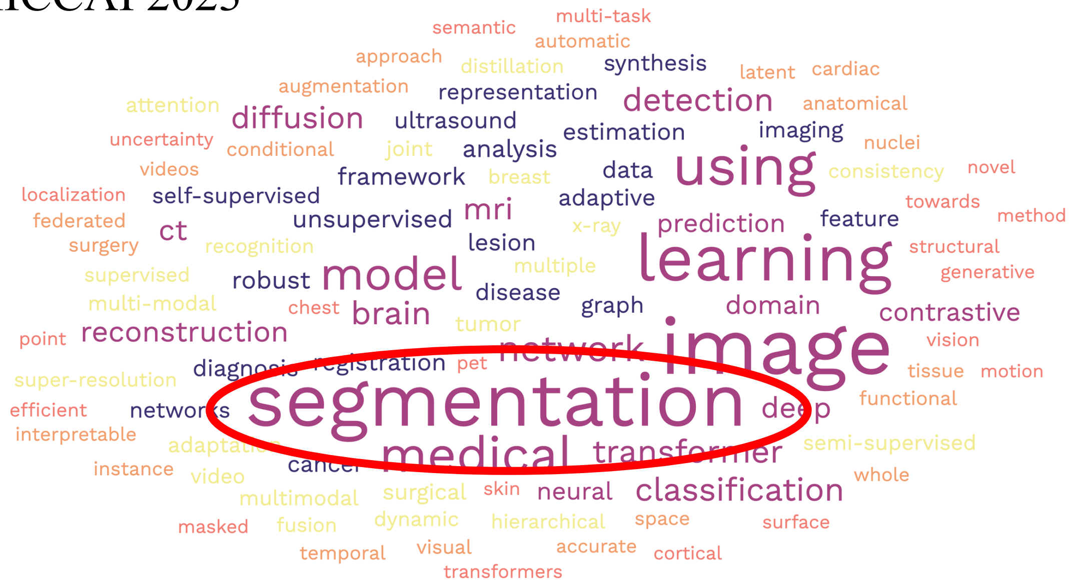
MICCAI 2021



MICCAI 2022



MICCAI 2023



Word clouds of paper titles in MICCAI 2020-2023 accepted papers.

Segmentation Paradigm in the Last half a Century

Heuristic methods

Optimization/model-based methods

1980s

Thresholding
Region growing
Watershed
…

Variational Models
(e.g., Snakes, GAC,…)

$$S^* = arg \min_S E(S)$$

Ma et al. ICTM-GAC, SIAM-IS, 2021

Mathematical Model-based Segmentation: An Example

Proposed Model $\min_{u(x)\in\{0,1\}} \sqrt{\frac{\pi}{\tau}} \int_\Omega \sqrt{g} u G_\tau * \left(\sqrt{g}(1-u)\right) + \lambda g u dx$

Relaxation $\min_{u(x)\in[0,1]} \sqrt{\frac{\pi}{\tau}} \int_\Omega \sqrt{g} u G_\tau * (\sqrt{g}(1-u)) + \lambda g u dx$

Ma, J. et al. A characteristic function-based algorithm for geodesic active contours, SIAM Journal on Imaging Sciences, 2021

Linearization $u^{k+1} = \arg\min_{u(x)\in[0,1]} \sqrt{\frac{\pi}{\tau}} \int_\Omega u\varphi(u^k) dx$

$$\varphi = \sqrt{g} G * \left(\sqrt{g}\left(1-2u^k\right)\right) + \lambda g$$

Algorithm 1 The iterative convolution-thresholding method (ICTM) for geodesic active contours

Require: Image edge indicator function, $\tau > 0$ and initialization $u^0 \in B$.

Ensure: Segmentation results $u^* \in B$;

while not converged **do**

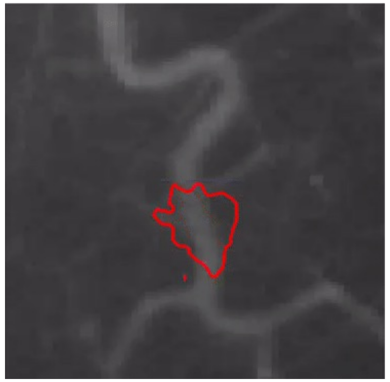
(1) Convolution. Fix u^k, compute

$$\varphi^k(x) = \sqrt{g} G_\tau * (\sqrt{g}(1-2u^k)) + \lambda g$$

(2) Thresholding. Set

$$u^{k+1}(\mathbf{x}) = \begin{cases} 1 & \text{if } \varphi(\mathbf{x}) \leq 0 \\ 0 & \text{otherwise} \end{cases}.$$

end while



Theorem (stability): $E^\tau(u^{k+1}) \leq E^\tau(u^k)$

😊 Explainable and transparent;

😊 Do not need large training set;

😭 Sensitive to initializations;

😭 Many hyper-parameter tunings;

Segmentation Paradigm in the Last half a Century

Heuristic methods

Optimization/model -based methods

DL-based tailored models

1980s

2015-2022

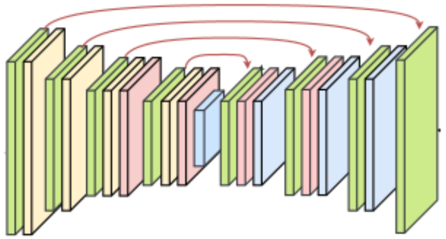
Thresholding
Region growing
Watershed
…

Variational Models
(e.g., Snakes, GAC,…)

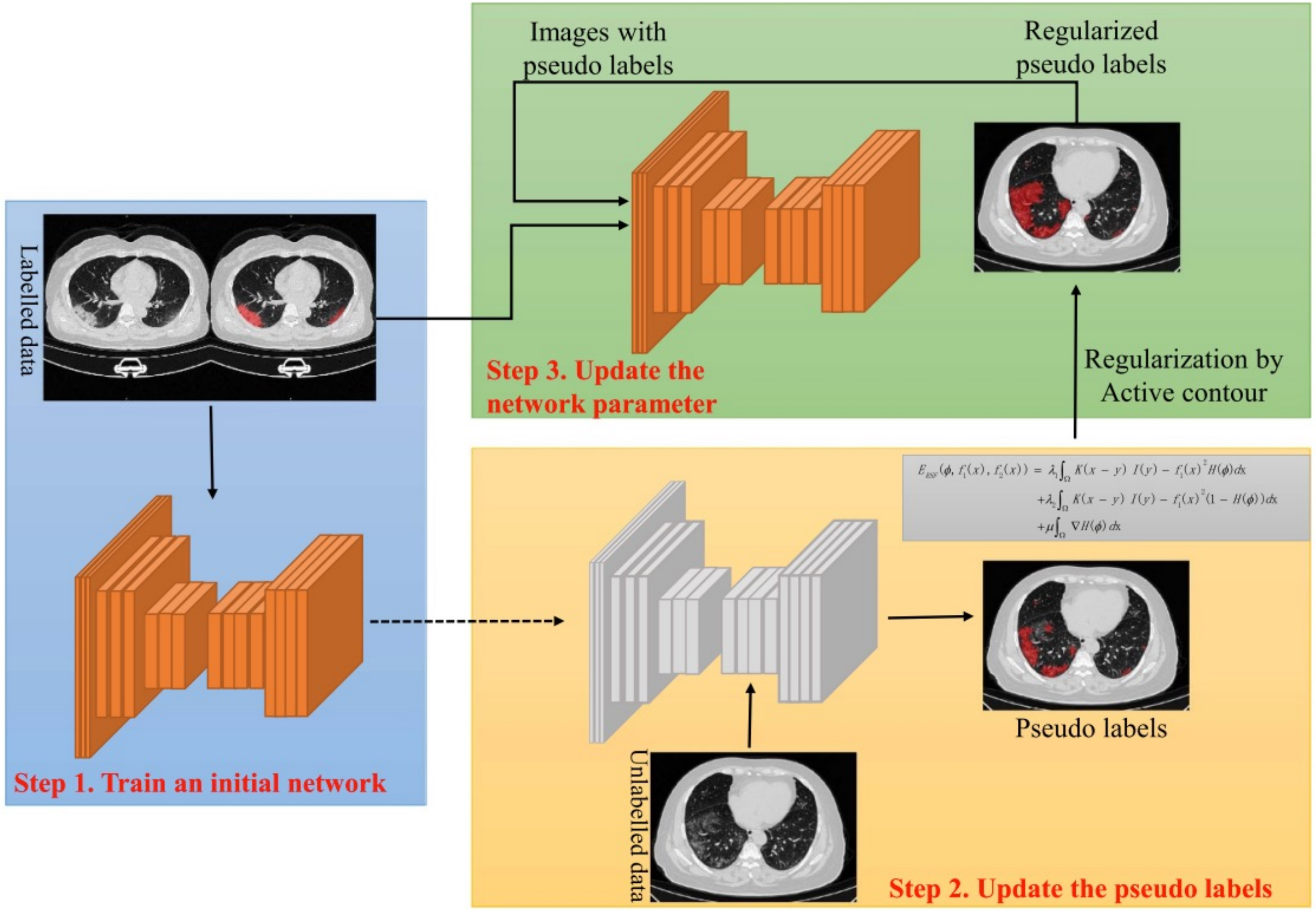
$$S^* = arg\min_S E(S)$$

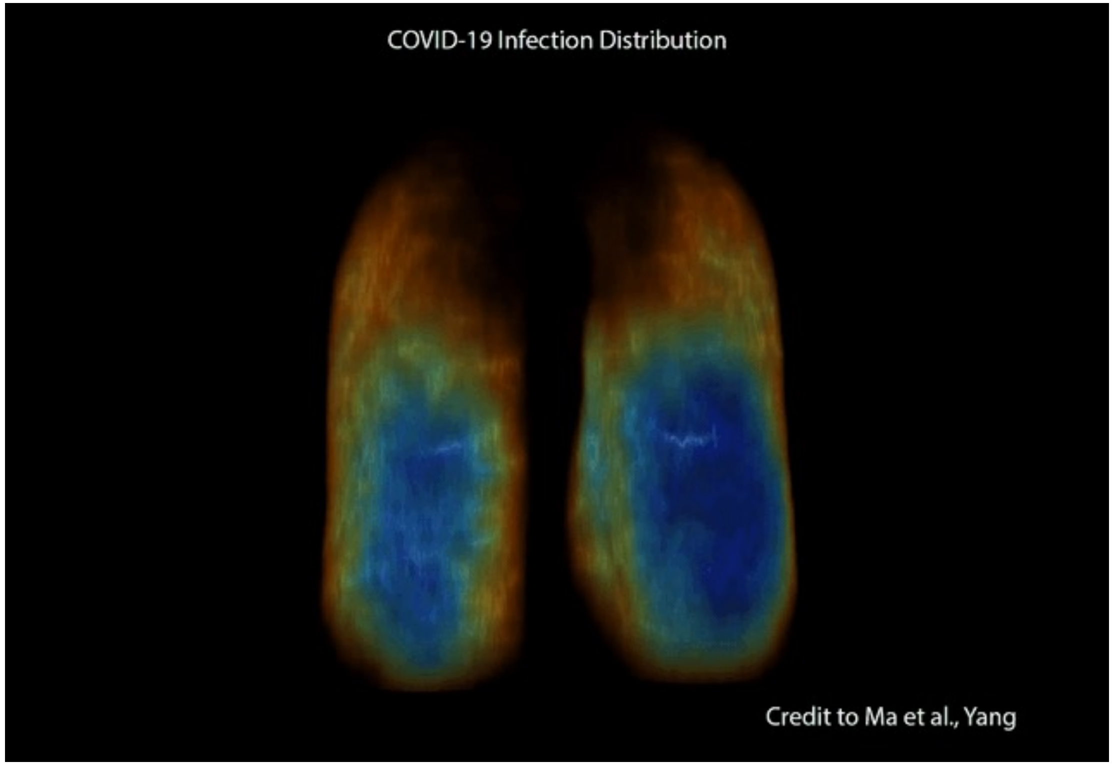
Ma et al. ICTM-GAC, SIAM-IS, 2021

FCN, U-Net, DeepLab, V-Net…



CNN-based Tailored Segmentation Model: An Example





COVID-19 infection spatial distribution map

😊 Automatic extract features;
😊 Automatic inference process;

😭 Task-specific adjustment;
😭 Limited generalizability and adaptability;

Ma, J. et al. “Active contour regularized semi-supervised learning for COVID-19 CT infection segmentation with limited annotations.” Physics in Medicine & Biology, 2020

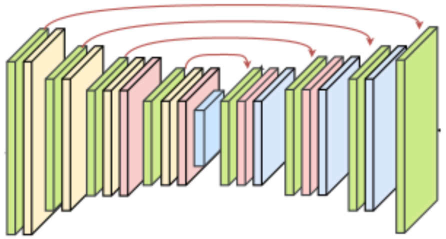
Ma, J. et al. "Toward data-efficient learning: A benchmark for COVID-19 CT lung and infection segmentation." Medical physics, 2021 (ESI highly cited paper)

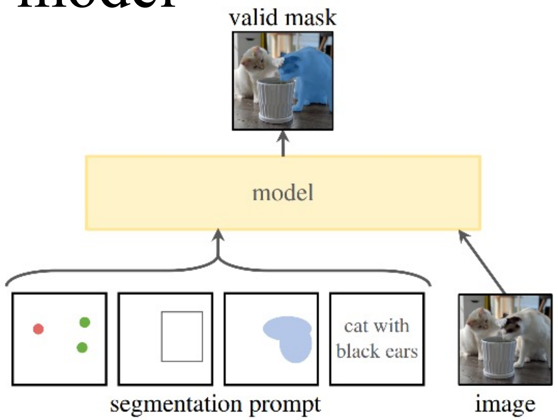
Segmentation Paradigm in the Last half a Century

Heuristic methods	Optimization/model -based methods	DL-based tailored models	Foundation models
	1980s	2015-2022	2023
Thresholding Region growing Watershed …	Variational Models (e.g., Snakes, GAC,…)	FCN, U-Net, DeepLab, V-Net…	Segment anything model

$$S^* = arg\min_{S} E(S)$$

Ma et al.
ICTM-GAC,
SIAM-IS,
2021

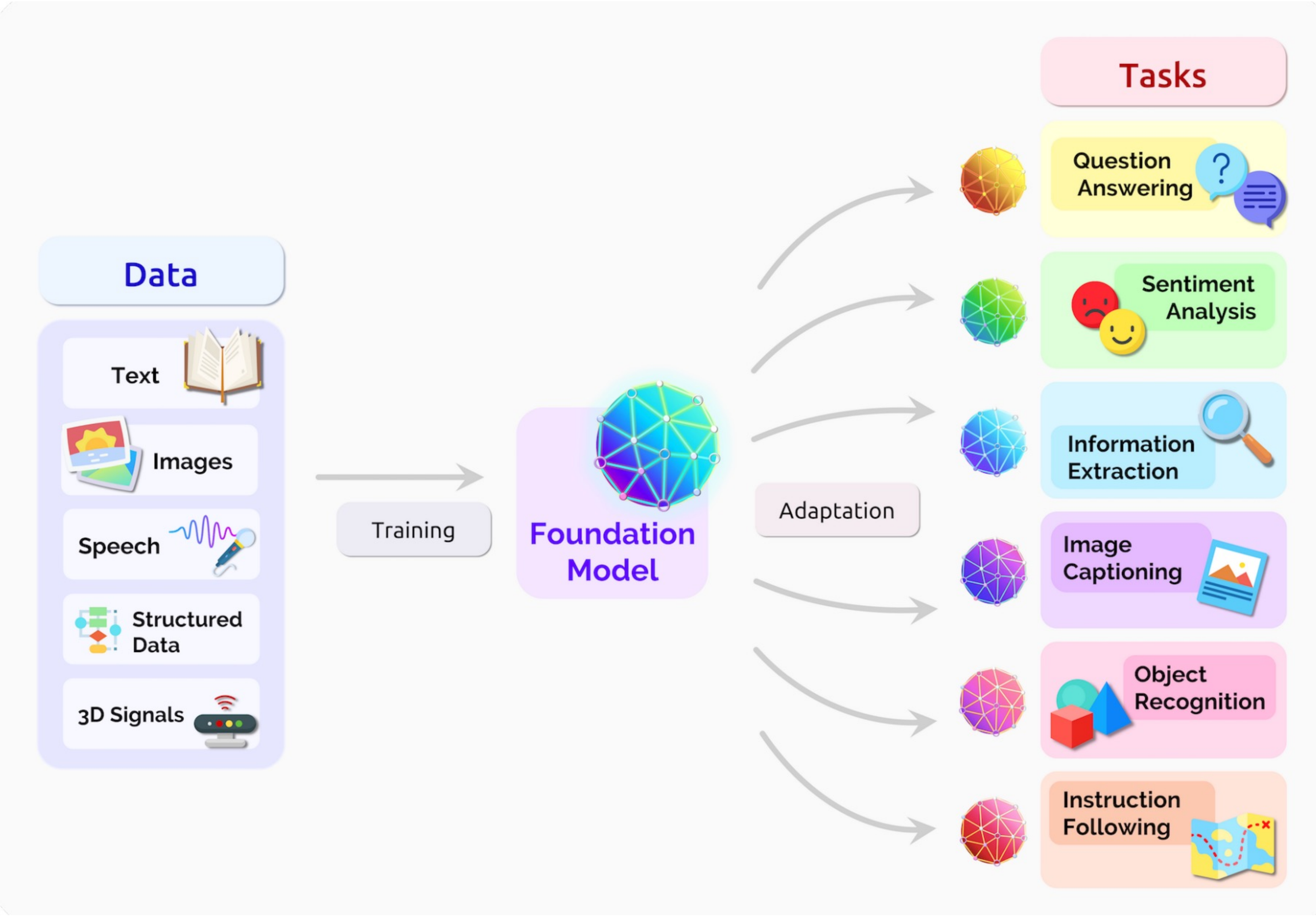




More data

More powerful computing

Foundation Models



Models trained on broad data that can be adapted to a wide range of downstream tasks.

Strong generalizability and adaptability

Bommasani, Rishi, et al. "On the opportunities and risks of foundation models." *arXiv preprint arXiv:2108.07258* (2021).

Segment Anything Model

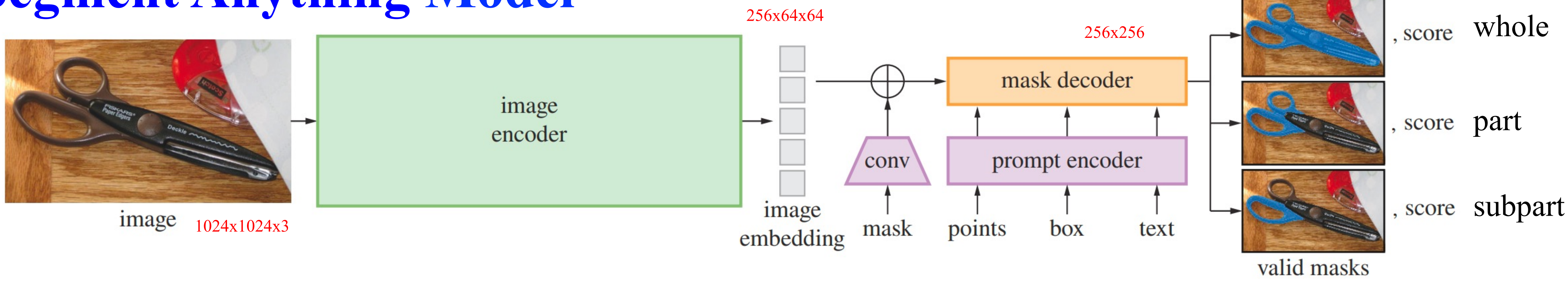


Figure 4: Segment Anything Model (SAM) overview. A heavyweight image encoder outputs an image embedding that can then be efficiently queried by a variety of input prompts to produce object masks at amortized real-time speed. For ambiguous prompts corresponding to more than one object, SAM can output multiple valid masks and associated confidence scores.

Image encoder

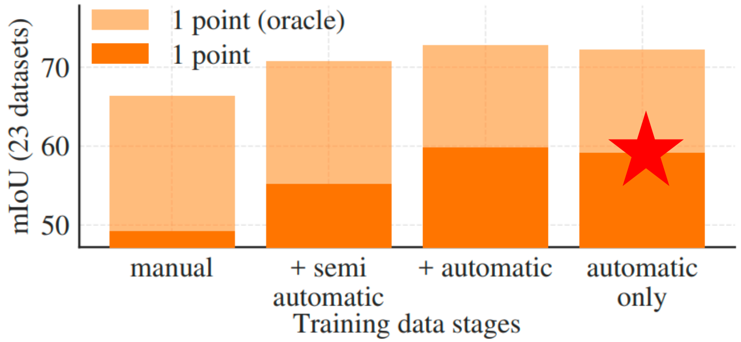
- Masked Auto-Encoder (MAE) pretrained Vision Transformer;
- Runs once per image
- Input: 1024x1024x3; Output embedding: 256x64x64 (16x downsacled)

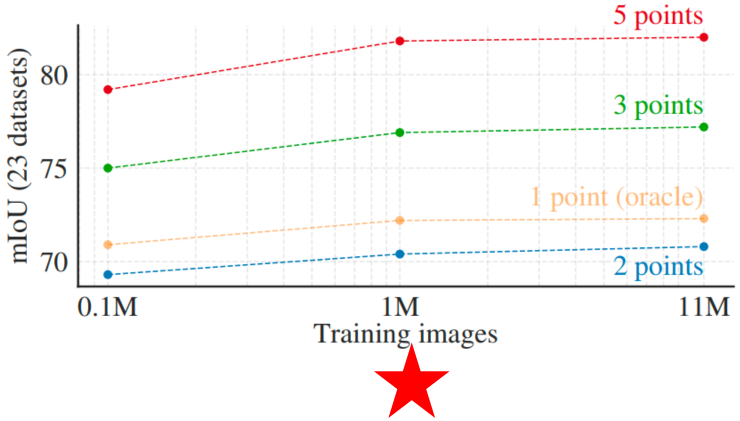
Mask decoder

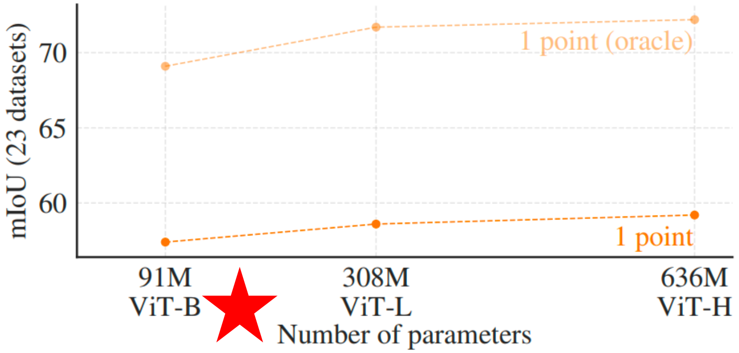
- Two-layer transformer decoder
- Inputs: image embedding + prompt embeddings
- Outputs: masks + IoU scores

SAM Performance









- Training with only the automatic data yields similar results to using all the data (manually labeled + model automatically generated data).
- SAM trained with ~10% of SA-1B (**1M images**) and full SA-1B is comparable.
- Scaling SAM's image encoder shows meaningful, yet saturating gains.

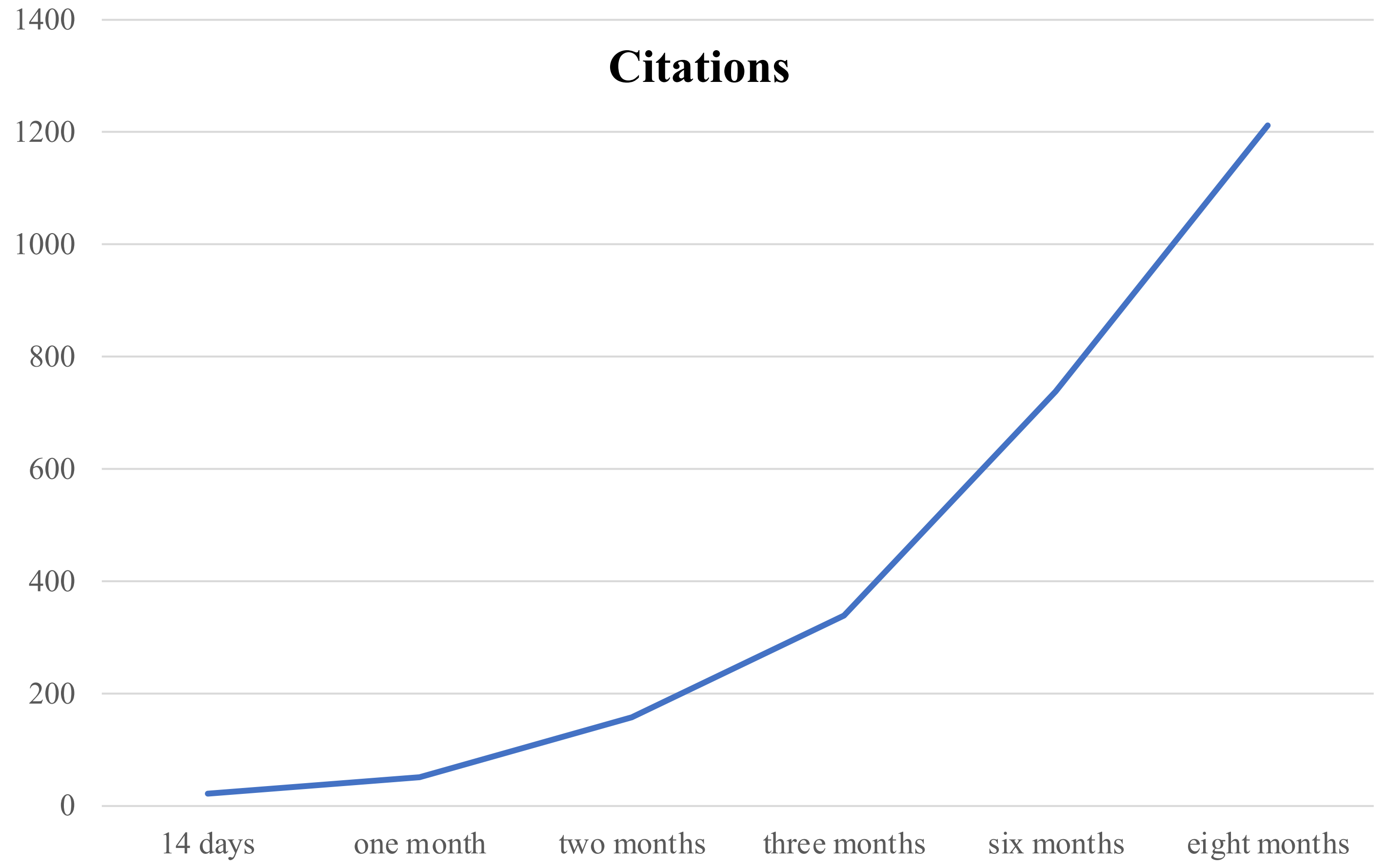
SAM's Friends

[Submitted on 5 Apr 2023]

Segment Anything

Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, Ross Girshick

We introduce the Segment Anything (SA) project: a new task, model, and dataset for image segmentation. Using our efficient model in a data collection loop, we built the largest segmentation dataset to date (by far), with over 1 billion masks on 11M licensed and privacy respecting images. The model is designed and trained to be promptable, so it can transfer zero-shot to new image distributions and tasks. We evaluate its capabilities on numerous tasks and find that its zero-shot performance is impressive -- often competitive with or even superior to prior fully supervised results. We are releasing the Segment Anything Model (SAM) and corresponding dataset (SA-1B) of 1B masks and 11M images at this https URL to foster research into foundation models for computer vision.

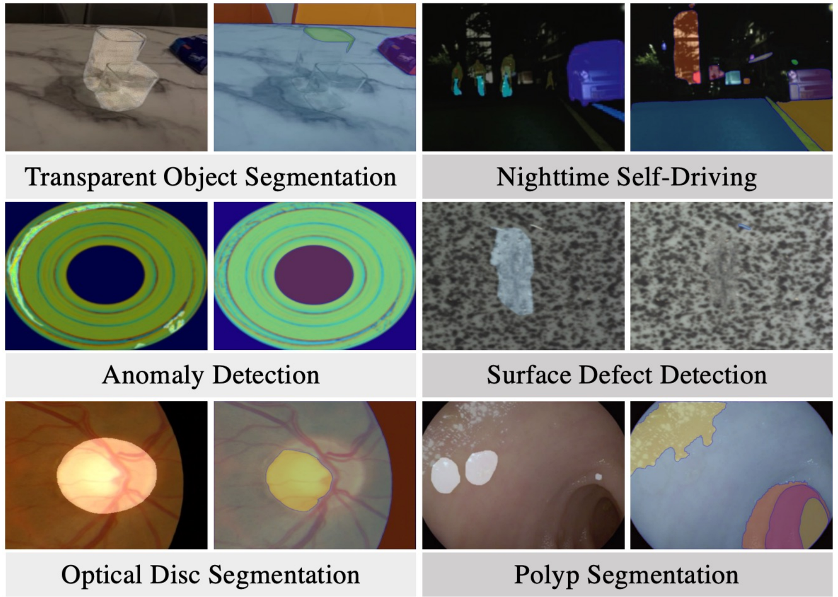


SAM was released on April 5th 2023

SAM's Friends

➢ **Benchmark SAM on new datasets**

- Title 1: Can SAM Segment A, B, C…
- Title 2: SAM Struggles in A, B, C
- Title 3: SAM meets A, B, C

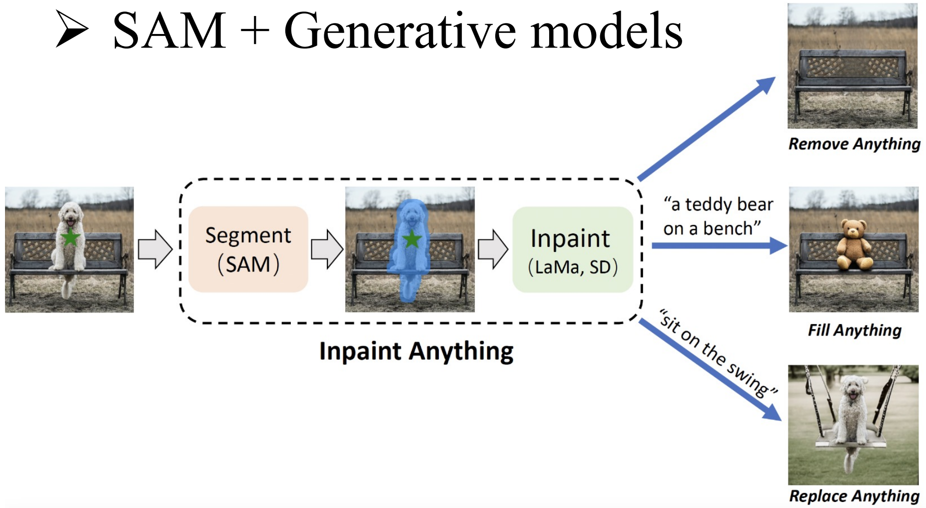


Ji et al. Segment Anything Is Not Always Perfect, 2023

➢ **Combine SAM with other tools**

➢ SAM + Detection/semantic segmentation/tracking models that can provide category information

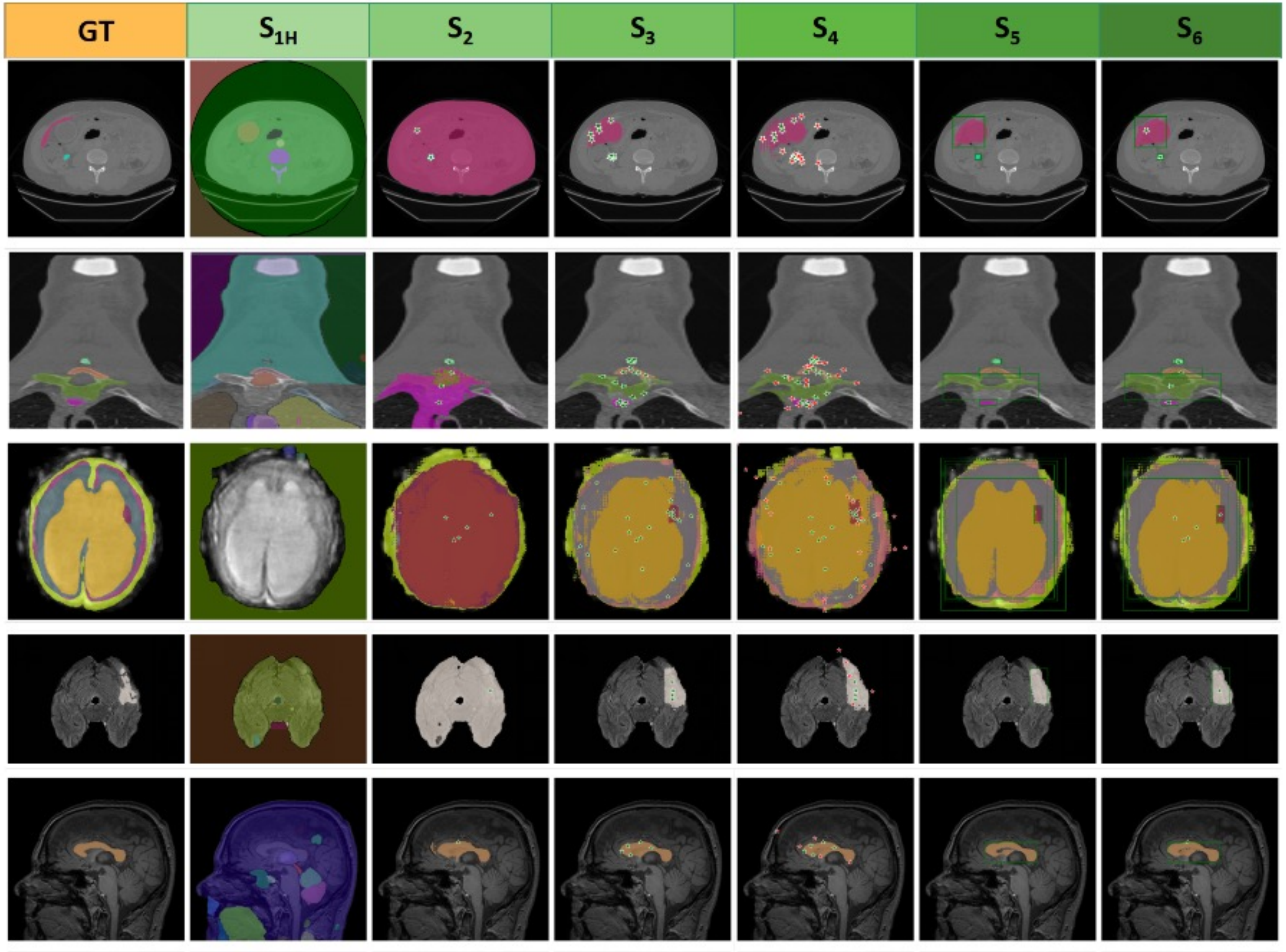
➢ SAM + Generative models



https://github.com/IDEA-Research/Grounded-Segment-Anything

https://github.com/gaomingqi/Track-Anything

SAM Fails to Segment Medical Images



Six different prompt modes

S_{1H}: Segment everything mode

S_2: One positive point

S_3: Five positive points

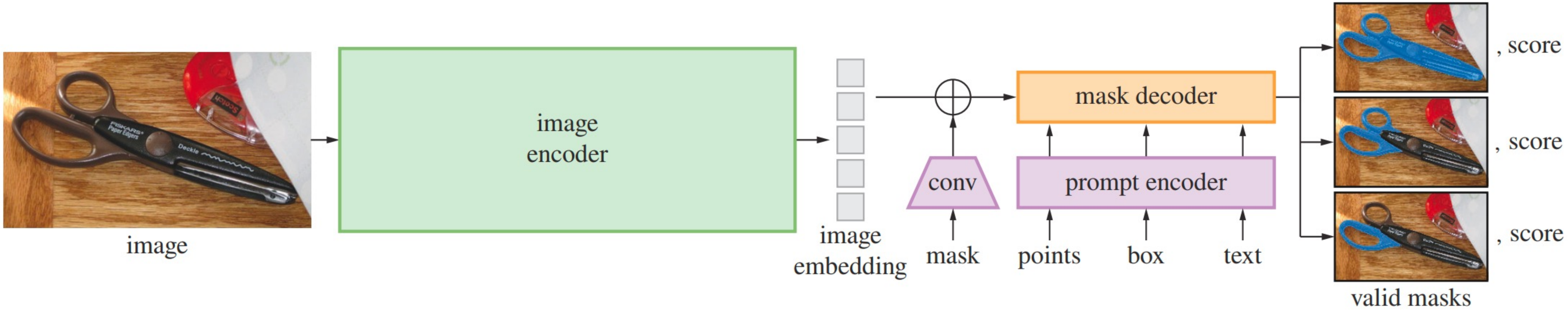
S_4: Five positive points and five negative points

S_5: One bounding box

S_6: One bounding box and one positive point

Huang et al. Segment Anything Model for Medical Images? 2023

MedSAM: Towards Medical Image Segmentation Foundation Model



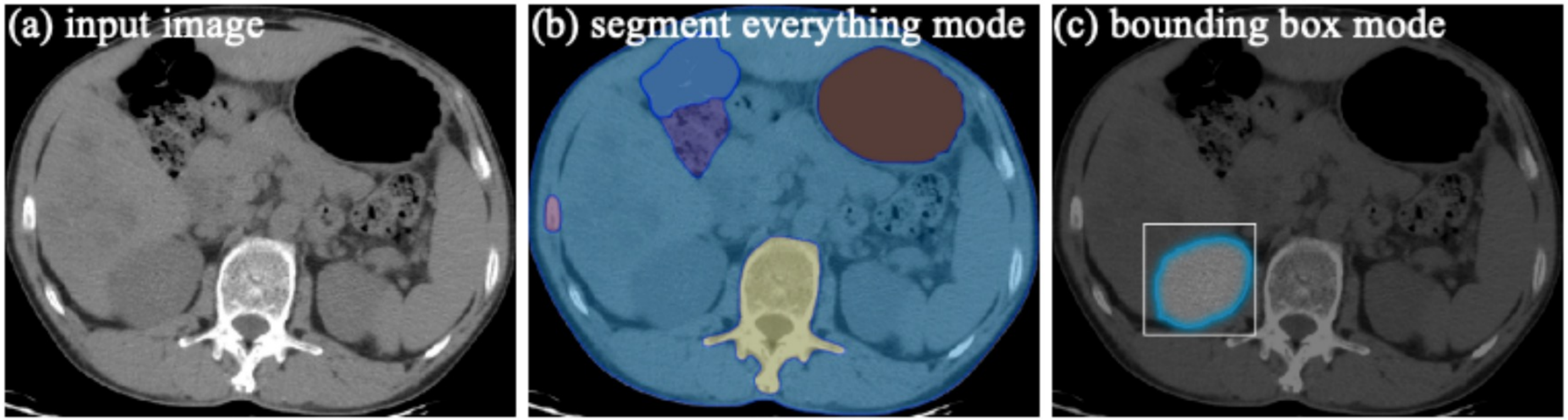
Key questions when adapting SAM to MedSAM

Q1. What can (not) SAM do in medical images?

Q2. How to choose the prompt?

Understand SAM's utility from medical perspectives

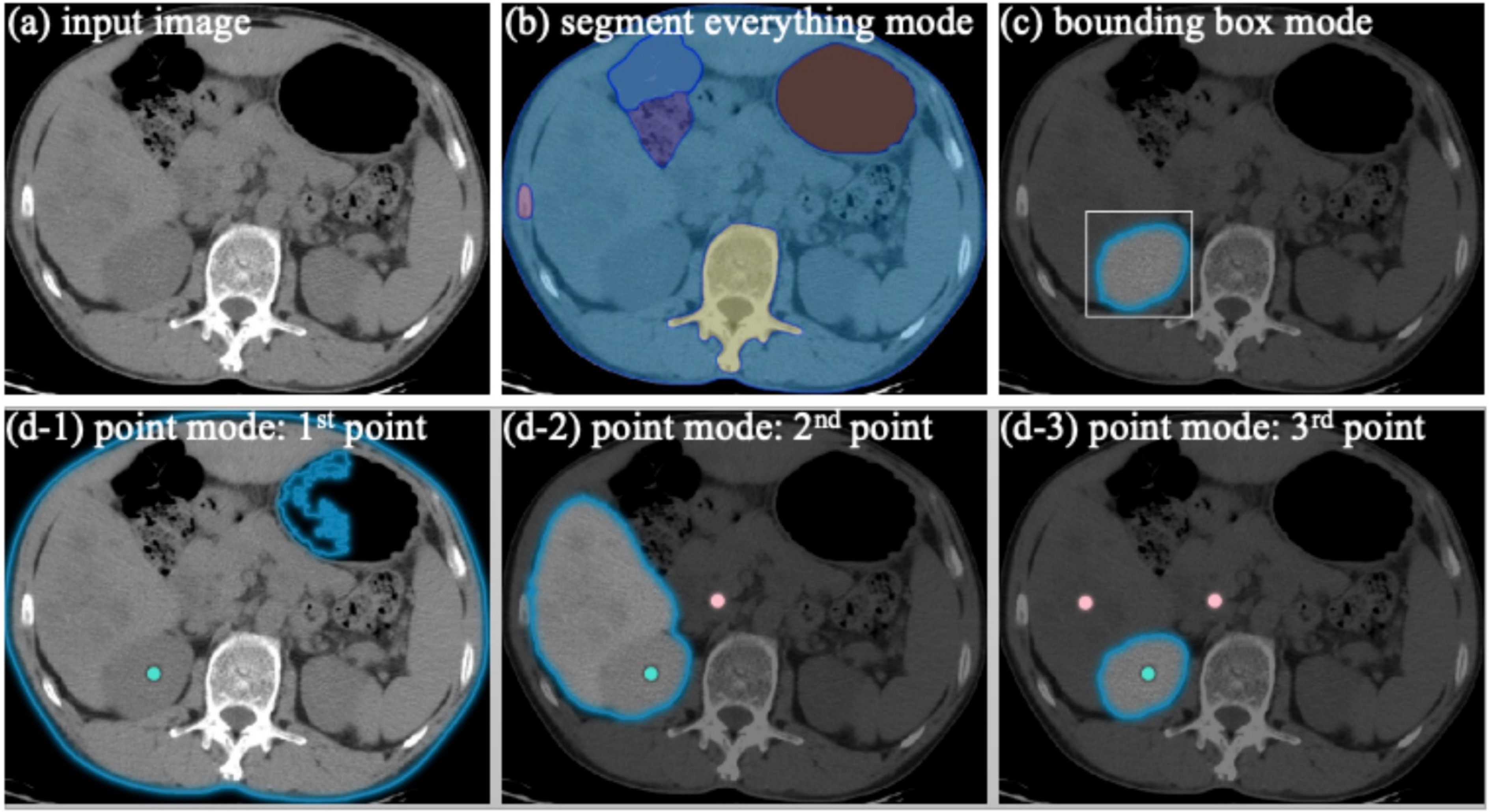
SAM supports three segmentation modes: segment everything, bounding box, and points



Segmentation performance under different modes

- Segment everything mode: prone to generate useless region partitions
- Bounding box mode: clearly specify the ROI and obtain reasonable segmentation results by just clicking two bounding points

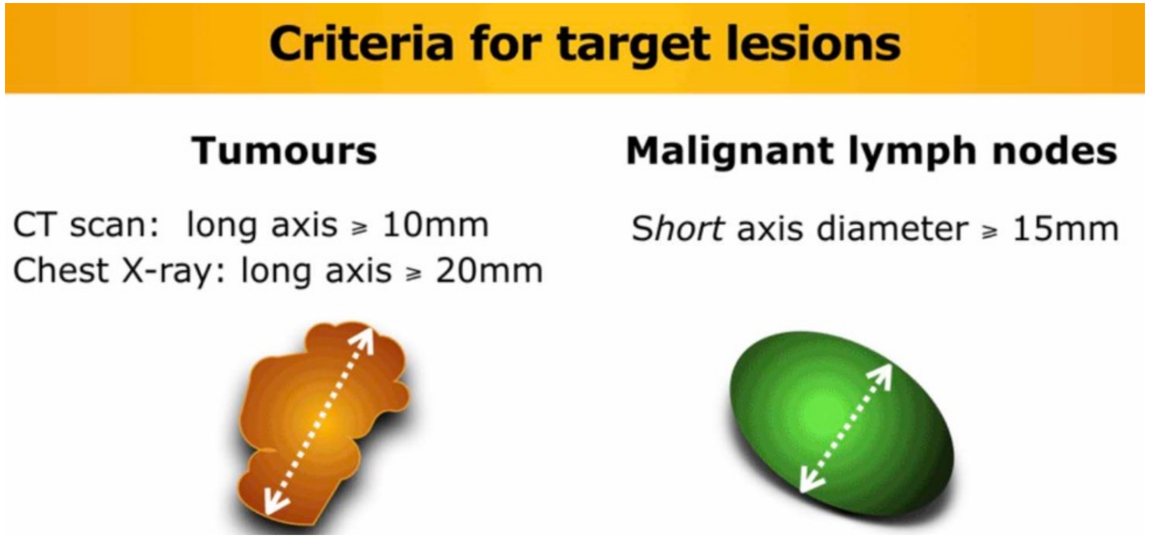
Understand SAM's utility from medical perspectives

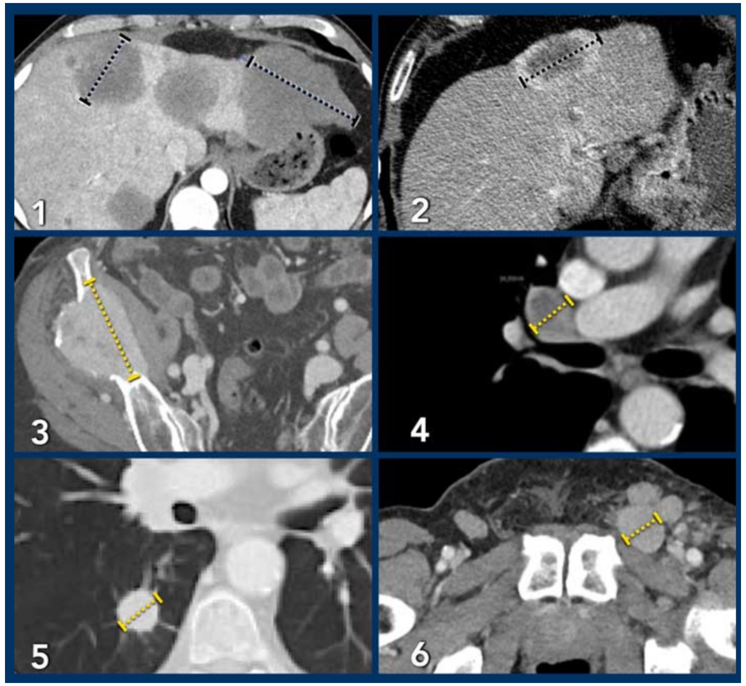


- Point mode: ambiguous and requires multiple prediction-correction iterations

Understand SAM's utility from medical perspectives

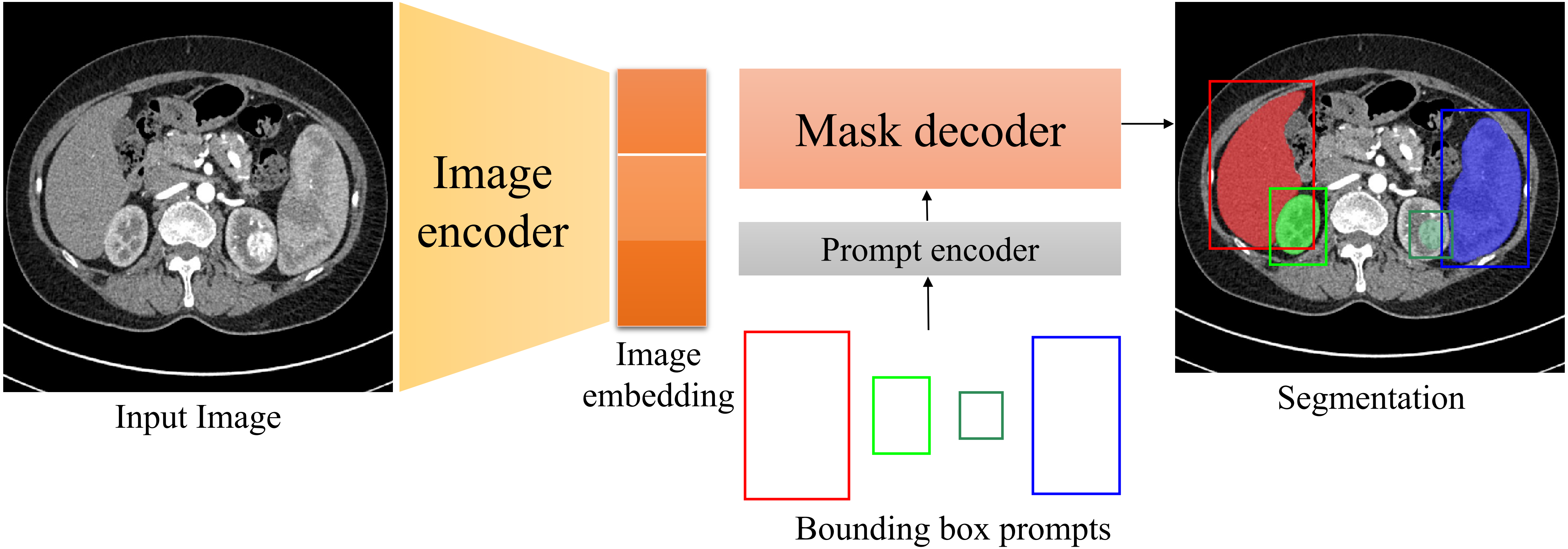
- SAM is essentially a point/bounding box-based segmentation method
- When applying SAM to medical image segmentation, bounding box is a better prompt.
    - It has less ambiguity
    - It doesn't require trial and errors
    - It fits typical clinical practice (e.g., RECIST) since we can simulate a bounding box from the long-short axis tumor annotation



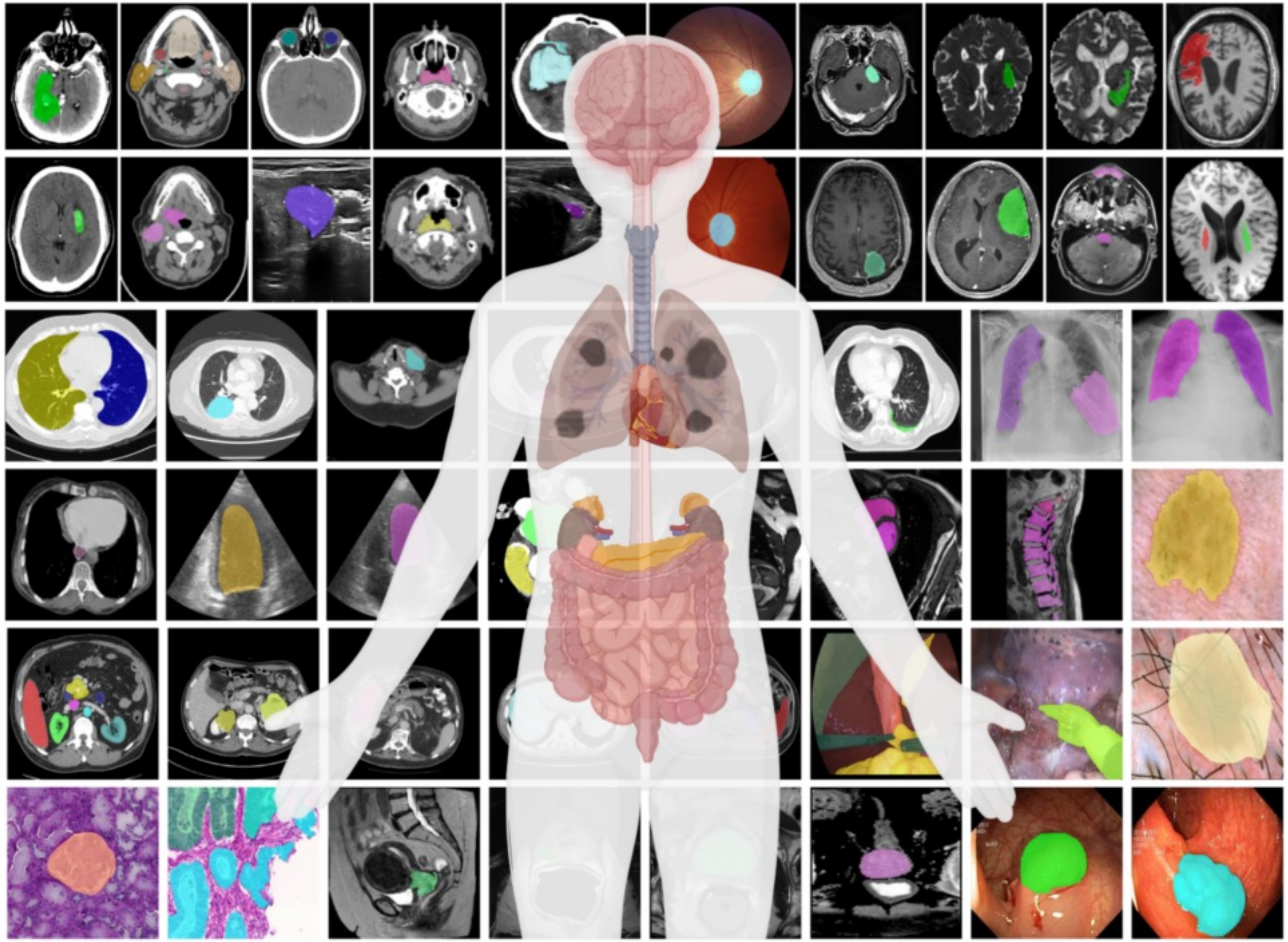


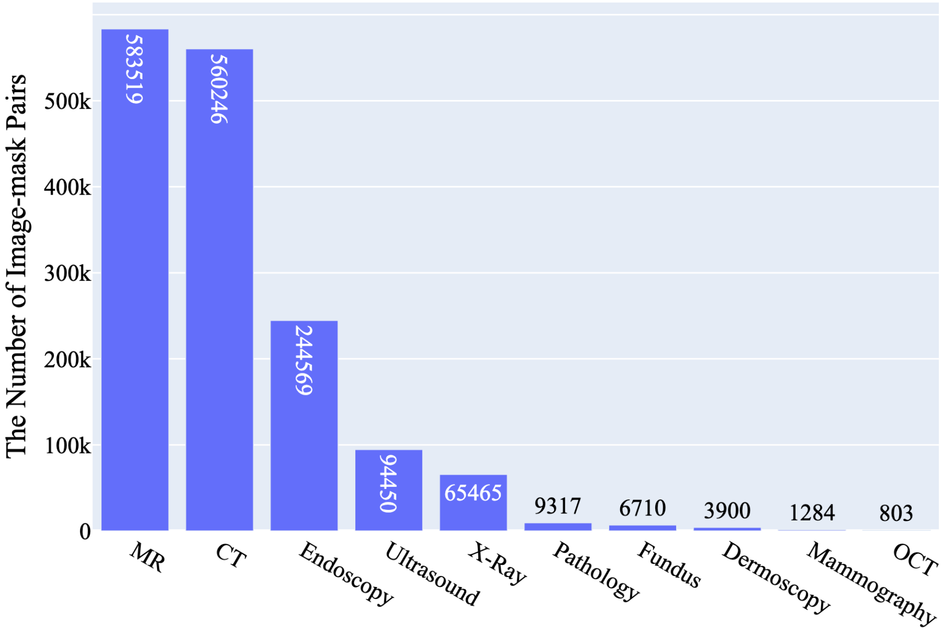
MedSAM: Pipeline

The first foundation model for promptable medical image segmentation



MedSAM: 1M+ Image-mask Pairs

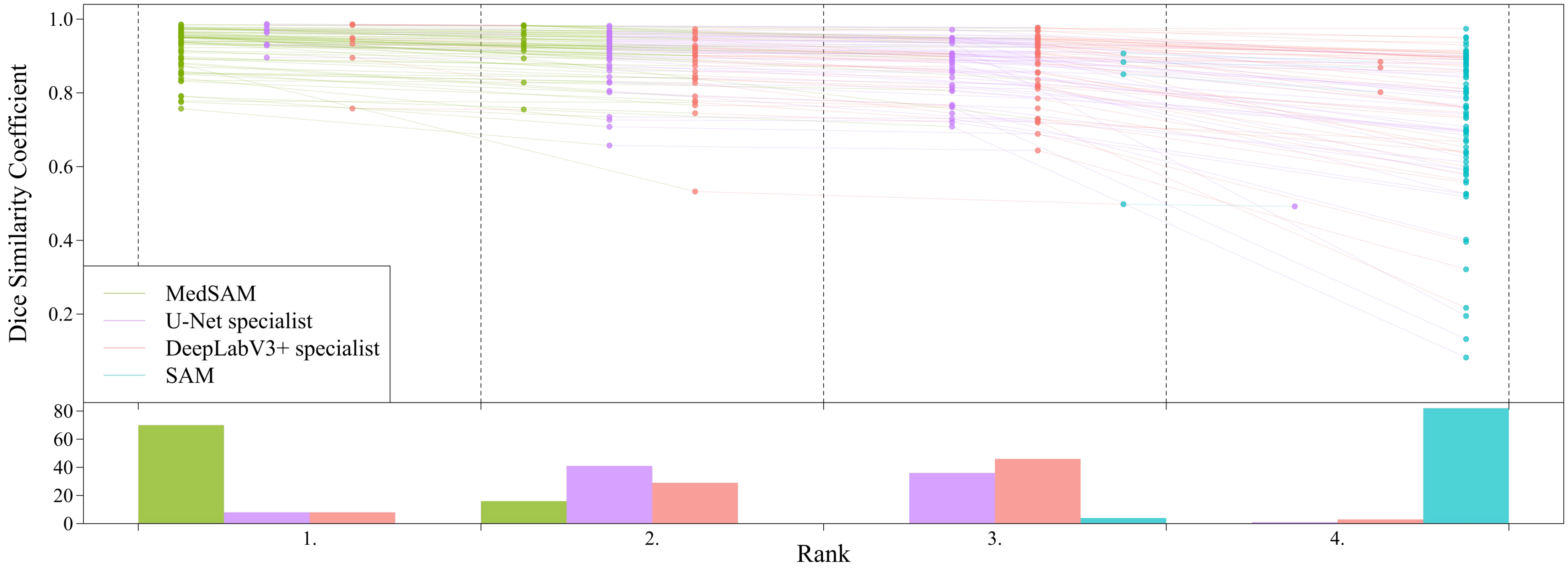


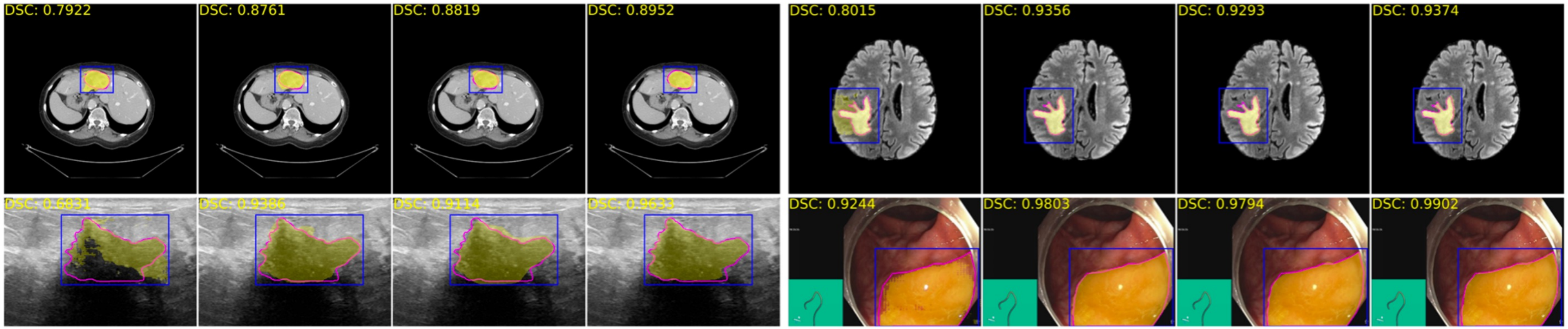


Experimental Settings

- 86 internal validation tasks
- 60 external validation tasks
- Compared to specialist U-Nets and DeepLabV3+ that are trained on each modality

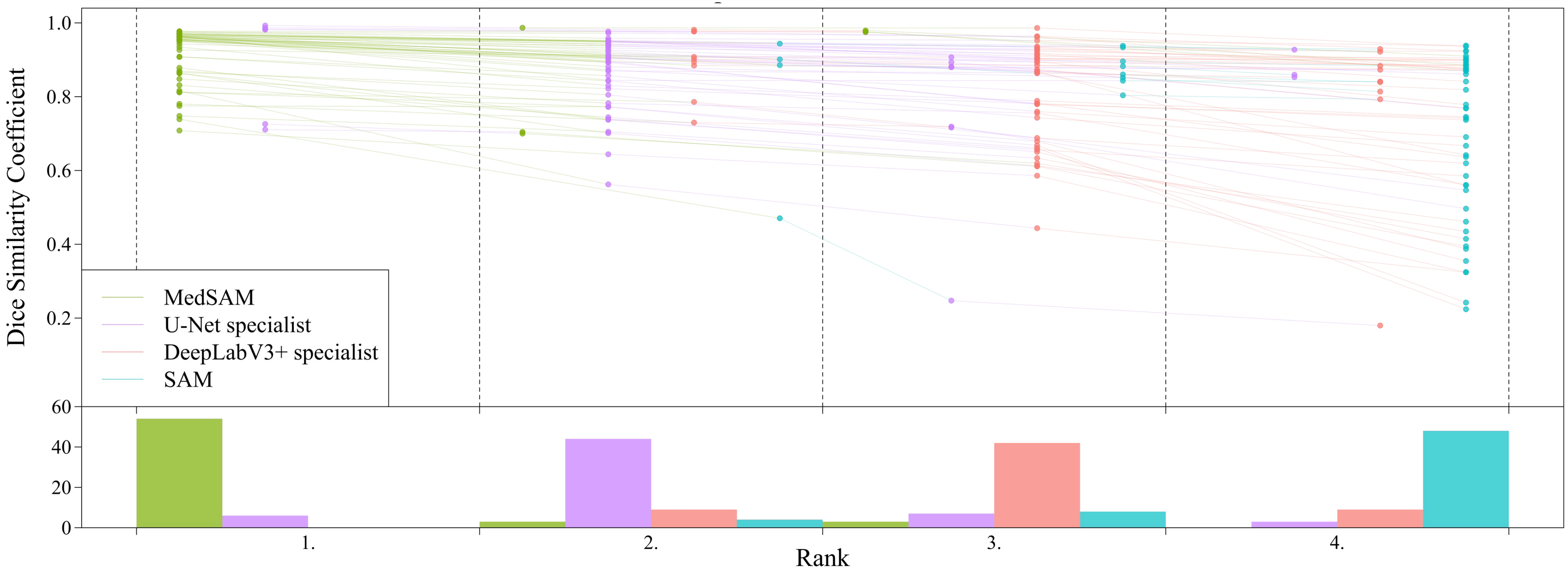
MedSAM: Internal Validation Results

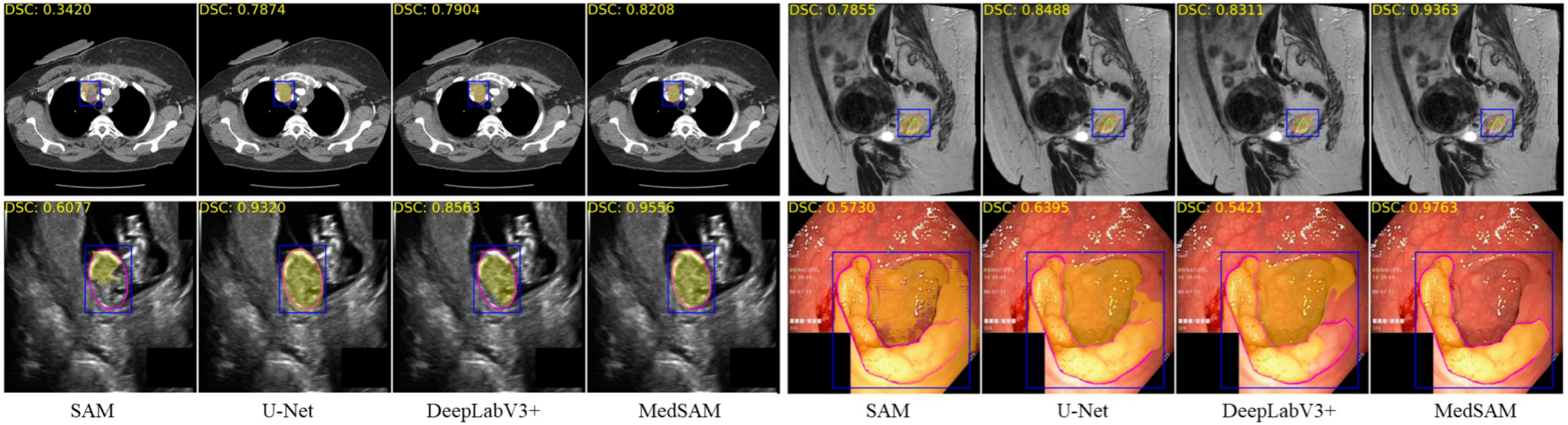




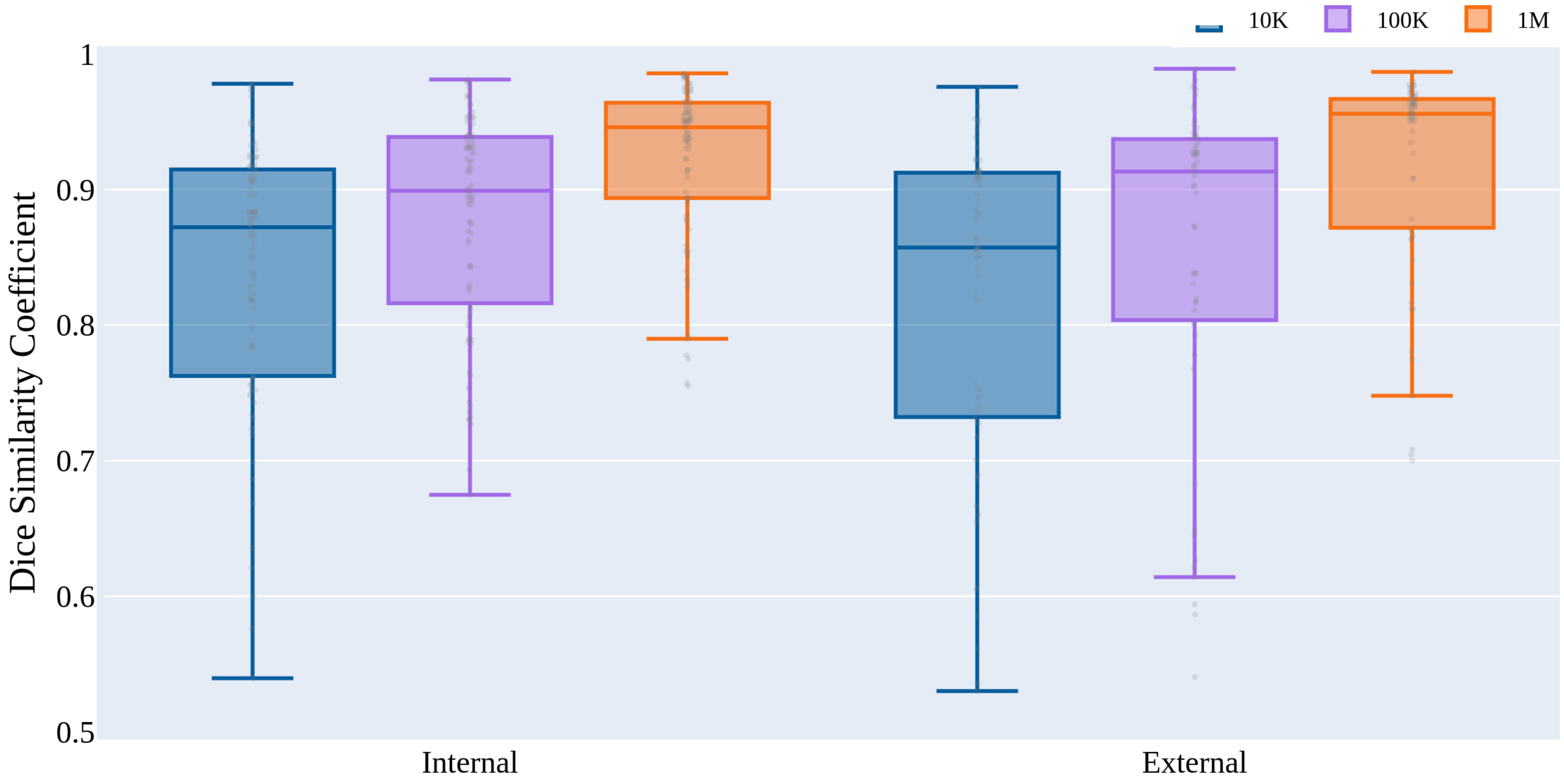
SAM U-Net DeepLabV3+ MedSAM SAM U-Net DeepLabV3+ MedSAM

MedSAM: External Validation Results

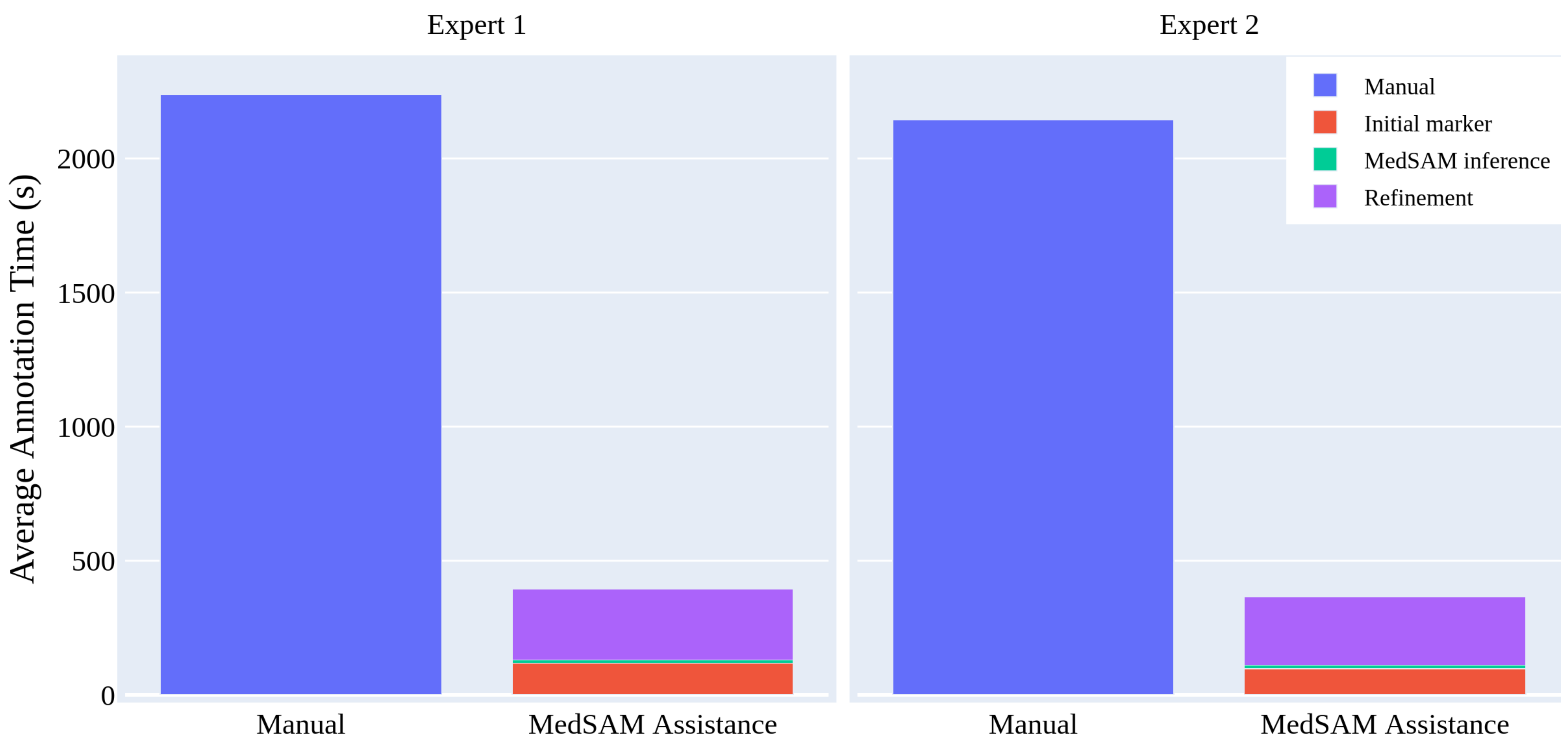




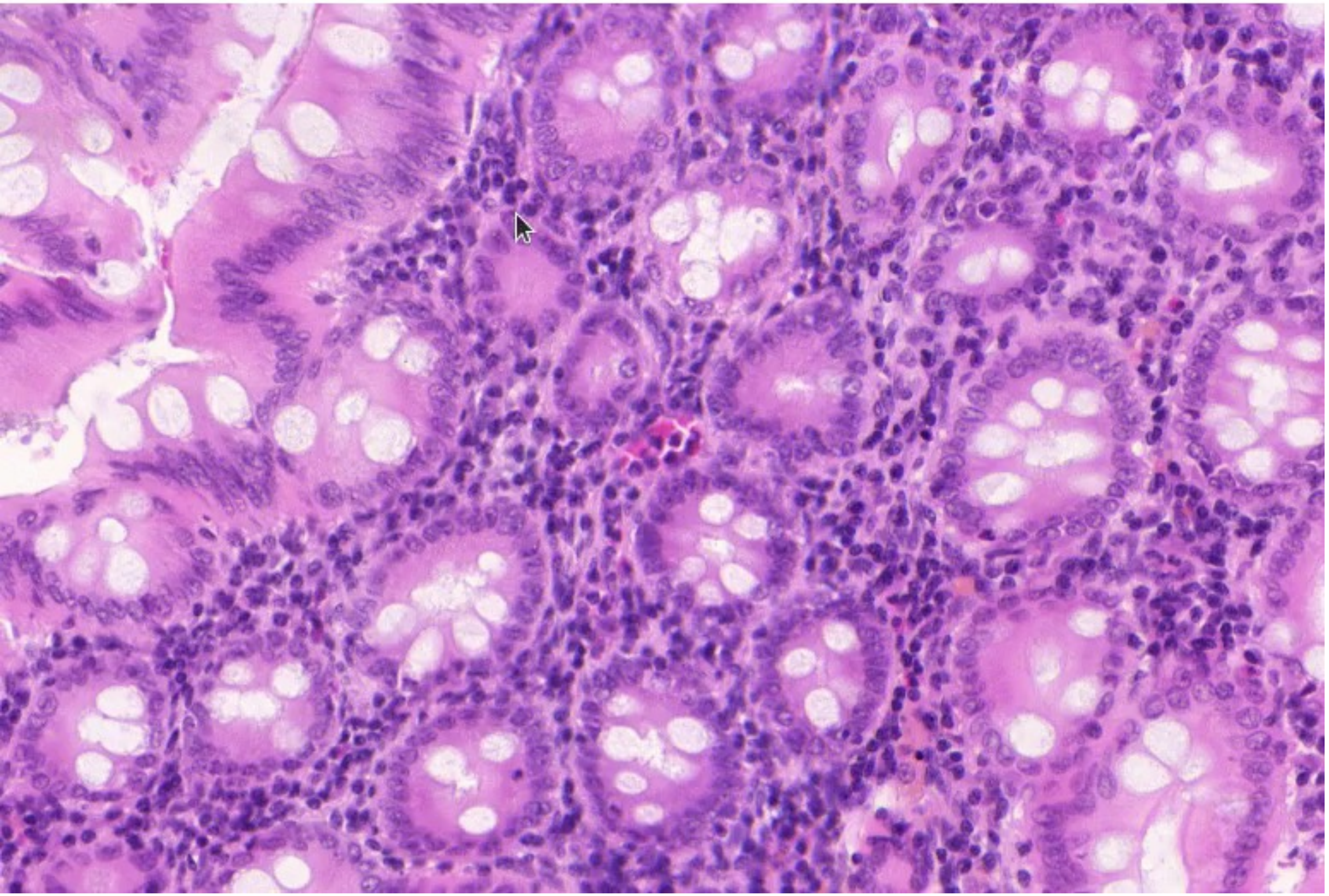
Performance Change of Varying Training Data Size



Human Annotation Study



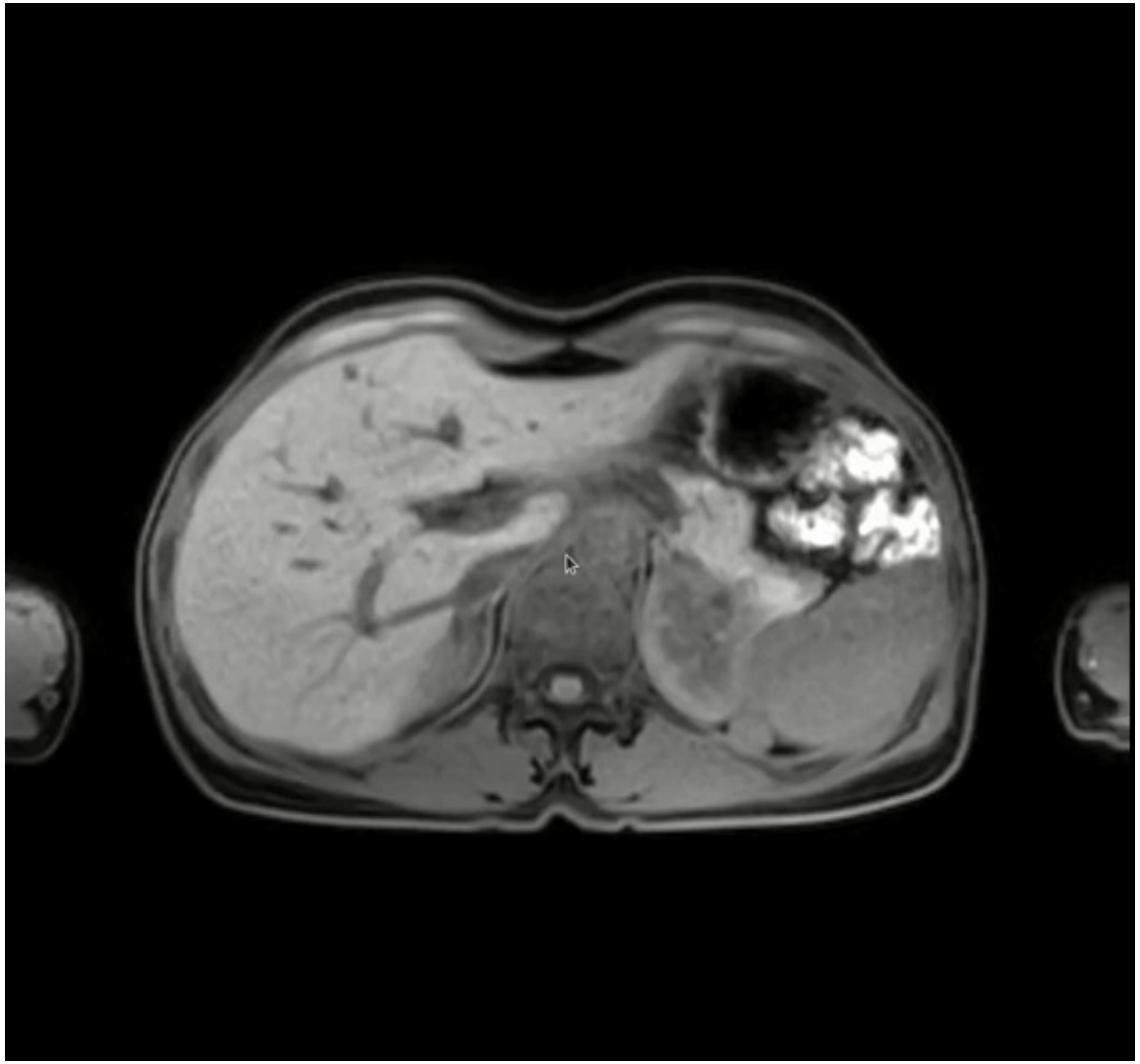
MedSAM: Demo



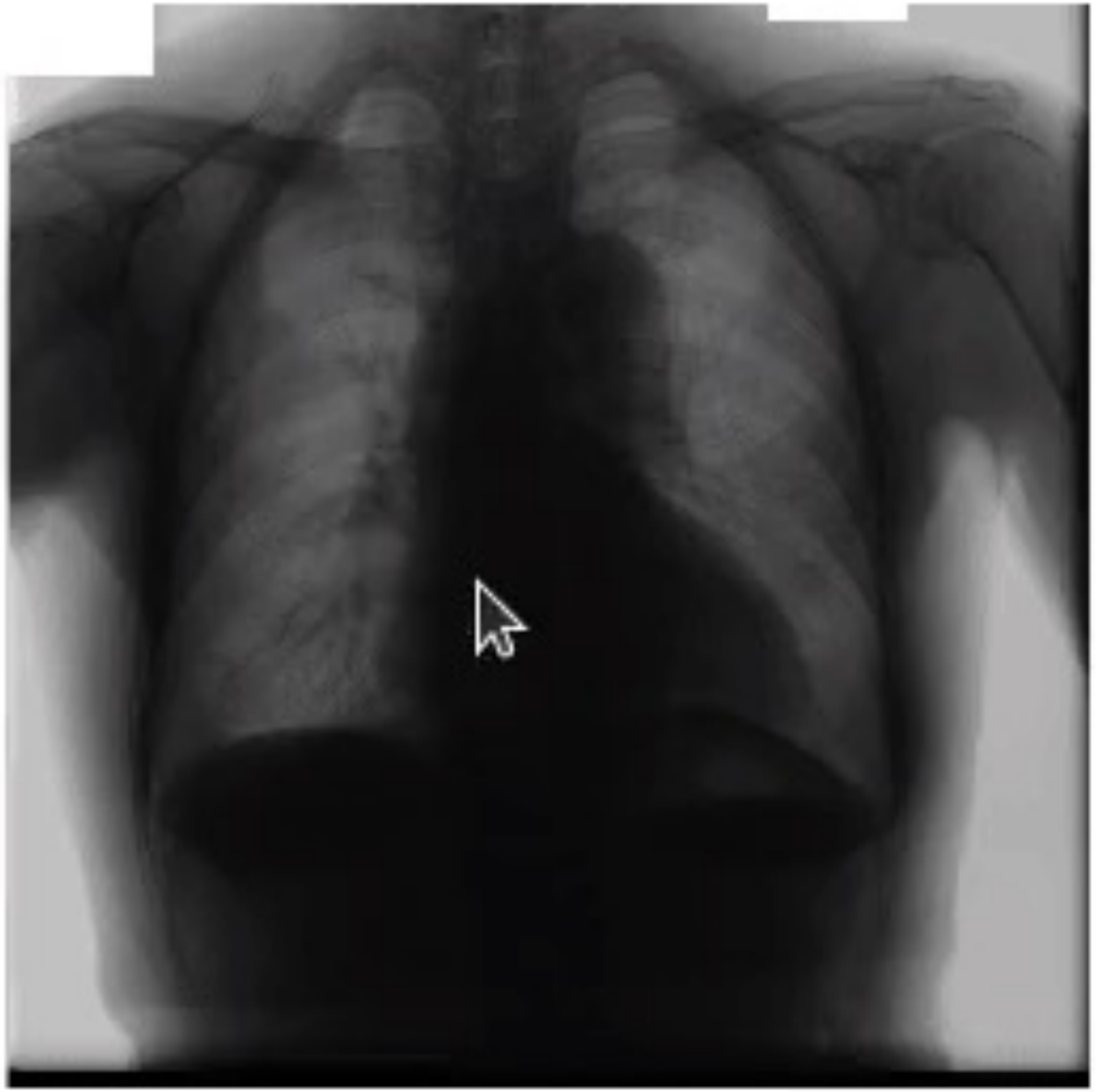
Gland Segmentation in Pathology Images

Liver and Tumor Segmentation in CT

MedSAM: Demo



Abdominal Organ Segmentation in MR



Lungs and Heart Segmentation in X-Ray

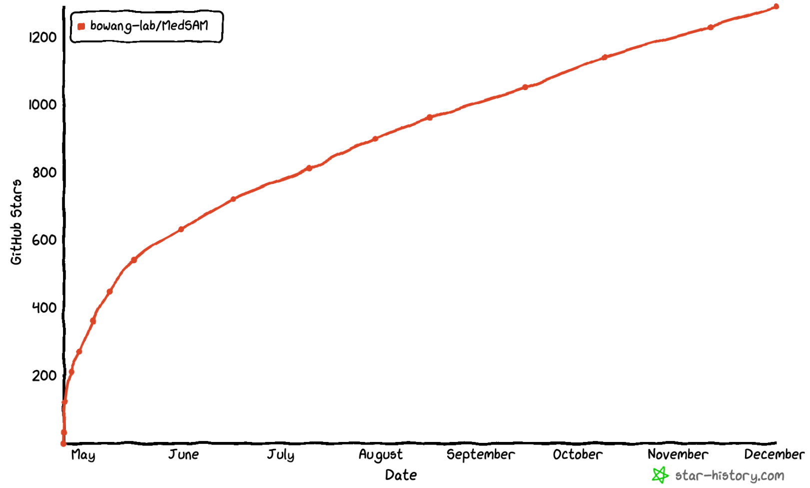
MedSAM in Community

Google Scholar (150+ citations in half a year)

Segment anything in medical images

Authors	Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, Bo Wang
Publication date	2023/4/24
Journal	arXiv preprint arXiv:2304.12306
Description	Segment anything model (SAM) has revolutionized natural image segmentation, but its performance on medical images is limited. This work presents MedSAM, the first attempt at extending the success of SAM to medical images, with the goal of creating a universal tool for the segmentation of various medical targets. Specifically, we first curate a large-scale medical image dataset, encompassing over 200,000 masks across 11 different modalities. Then, we develop a simple fine-tuning method to adapt SAM to general medical image segmentation. Comprehensive experiments on 21 3D segmentation tasks and 9 2D segmentation tasks demonstrate that MedSAM outperforms the default SAM model with an average Dice Similarity Coefficient (DSC) of 22.5% and 17.6% on 3D and 2D segmentation tasks, respectively. The code and trained model are publicly available at \url{https://github.com/bowang-lab/MedSAM}.

GitHub Stars (1.2K+)



https://github.com/bowang-lab/MedSAM

MedSAM in HuggingFace

Segment medical images with MedSAM

In this notebook, we're going to perform inference with MedSAM, a fine-tuned version of the SAM (segment-anything model) by Meta AI on the medical domain (thereby greatly improving its performance).

- Original repo
- Hugging Face docs.

https://github.com/NielsRogge/Transformers-Tutorials/blob/master/SAM/Run_inference_with_MedSAM_using_HuggingFace_Transformers.ipynb

MedSAM in napari

I integrated MedSam into napari FYI #36

Closed Karol-G opened this issue on May 5 · 2 comments

Karol-G commented on May 5

Hey,

I just wanted to let you know that I integrated MedSam already into my Napari SAM plugin: https://github.com/MIC-DKFZ/napari-sam

So you can check the mark on "3D slicer and napari support" on your todo list if you want ;)

Best,
Karol

👍 1

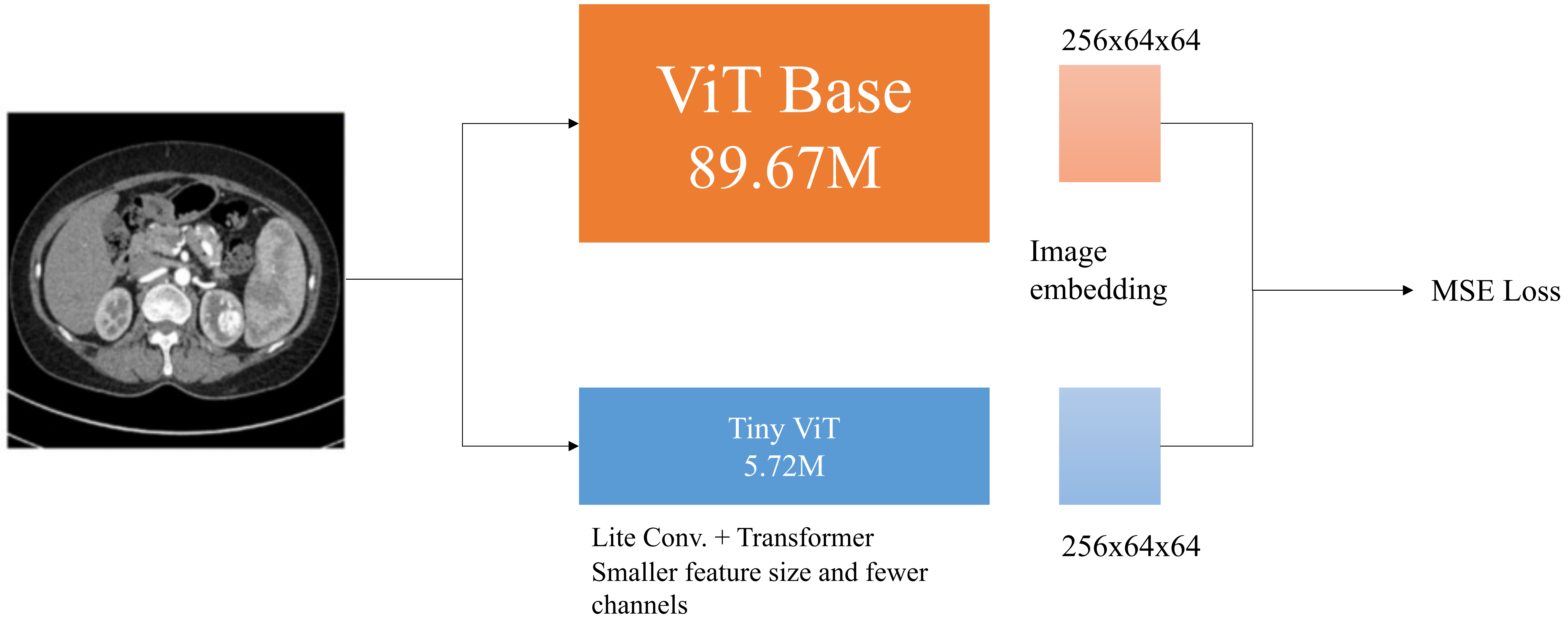
https://github.com/MIC-DKFZ/napari-sam

MedSAM is a useful segmentation tool,

but how to incorporate it into the clinical practice?

Lite MedSAM: 10× Faster

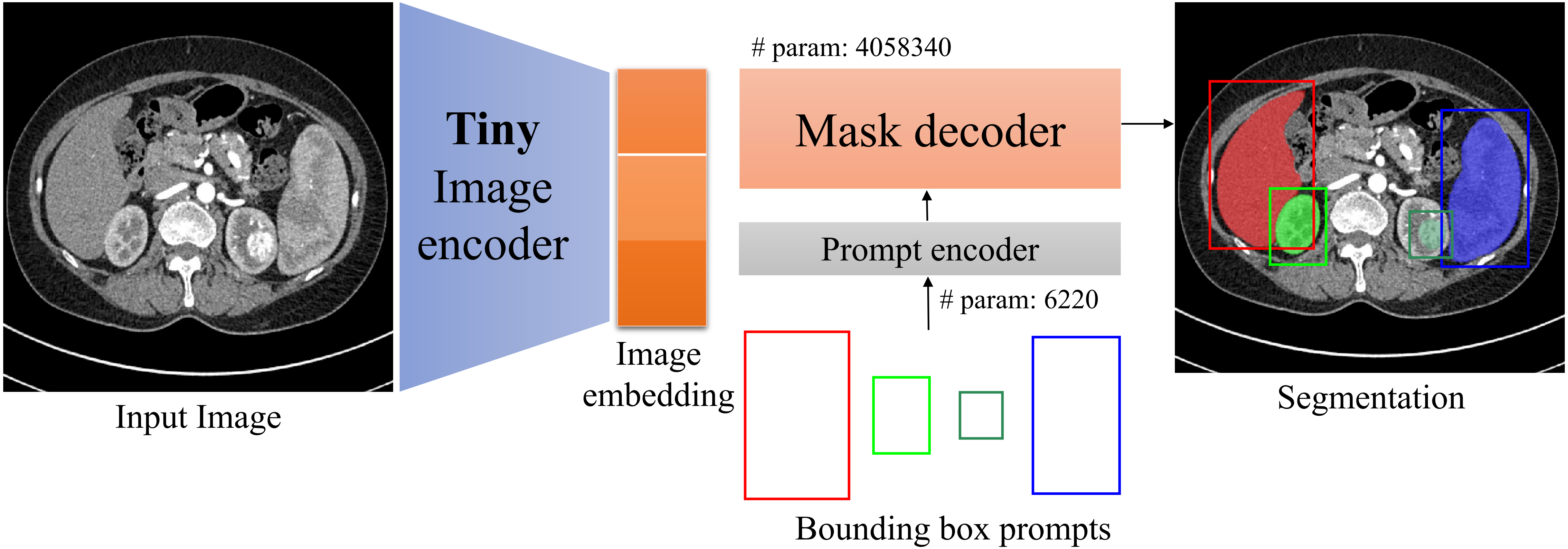
Stage 1. Distillation a small image encoder



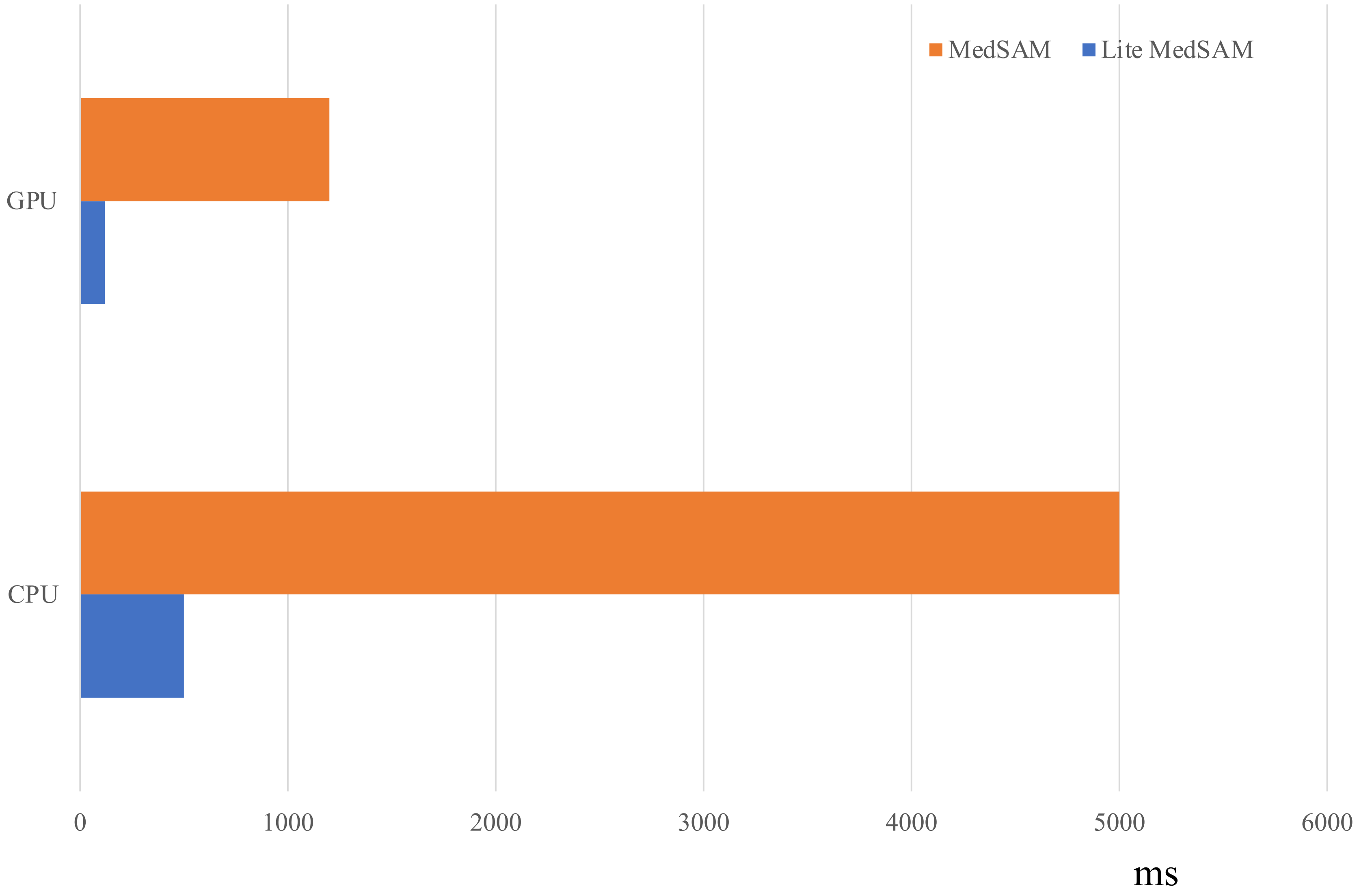
Wu, Kan, et al. "Tinyvit: Fast pretraining distillation for small vision transformers." ECCV, 2022.
Zhang, Chaoning, et al. "Faster Segment Anything: Towards Lightweight SAM for Mobile Applications." *arXiv preprint arXiv:2306.14289* (2023).
Zhao, Xu, et al. "Fast Segment Anything." arXiv preprint arXiv:2306.12156 (2023).

Lite MedSAM: 10× Faster

Stage 2. Fine-tune the whole model

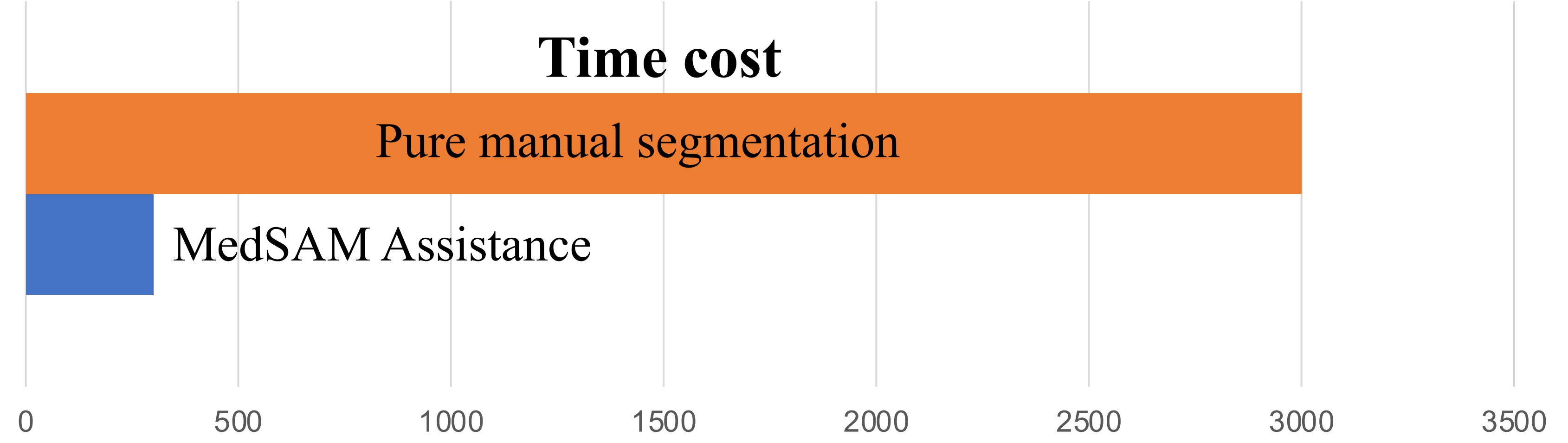


Lite MedSAM: 10× Faster



Lite MedSAM Adaptability: A case study

- Generating initial (~30) 2D masks with Lite MedSAM (draw bounding box) and manually refining.
- Fine-tune Lite MedSAM based on the labeled data
- Label another 30 images
- Train a fully automatic model
- Repeat this process (select hard cases in each stage)



3D Slicer Integration

3D Slicer image computing platform

Download | Documentation | Developers
Training | Forum | Twitter

3D Slicer is a **free**, **open source** software for visualization, processing, segmentation, registration, and analysis of medical, biomedical, and other 3D images and meshes; and planning and navigating image-guided procedures.

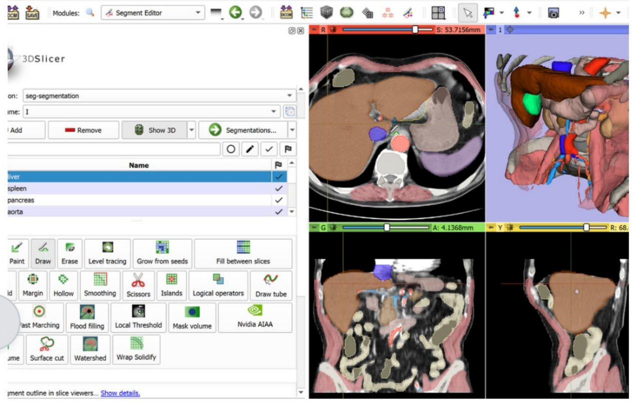
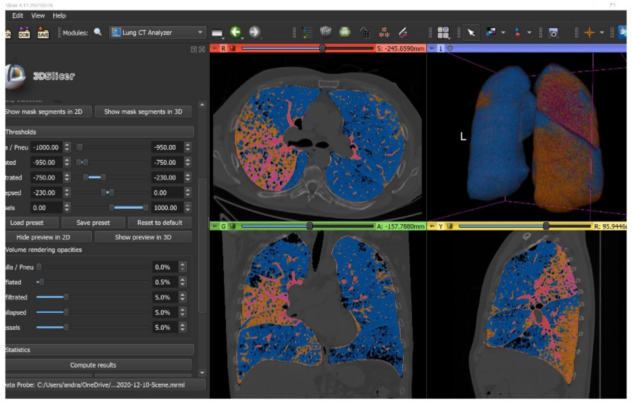


Image segmentation

Create surgical plans, create high-quality atlases, or training data sets for deep learning using the Segment Editor module. learn more >



Lung CT analysis for COVID-19

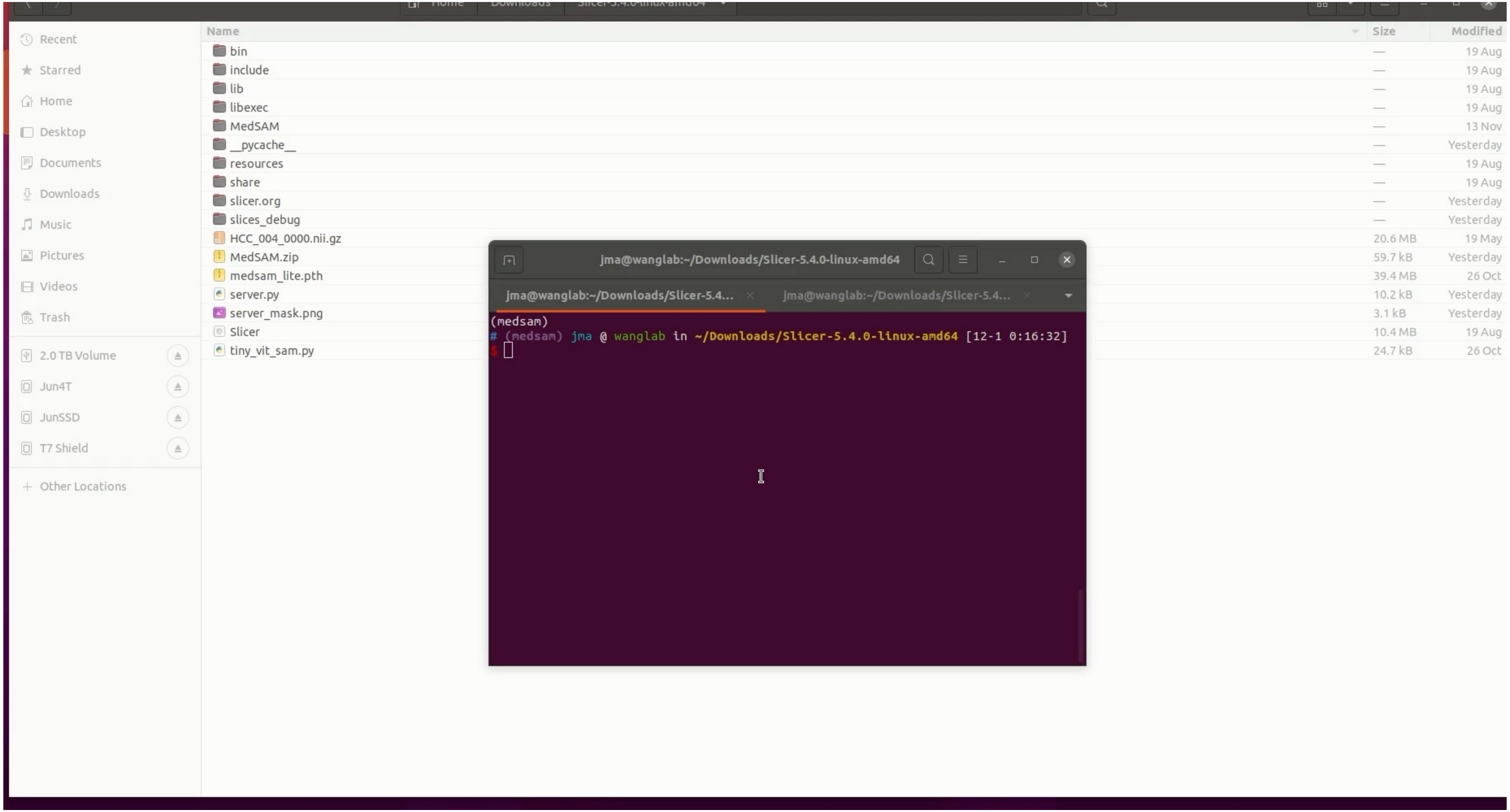
LungCTAnalyzer extension offers automated lung segmentation and quantative analysis for COVID-19 cases. video > learn more >

Surgical navigation

3D Slicer is used in real-time navigation of breast cancer surgery. video > journal article > learn more >

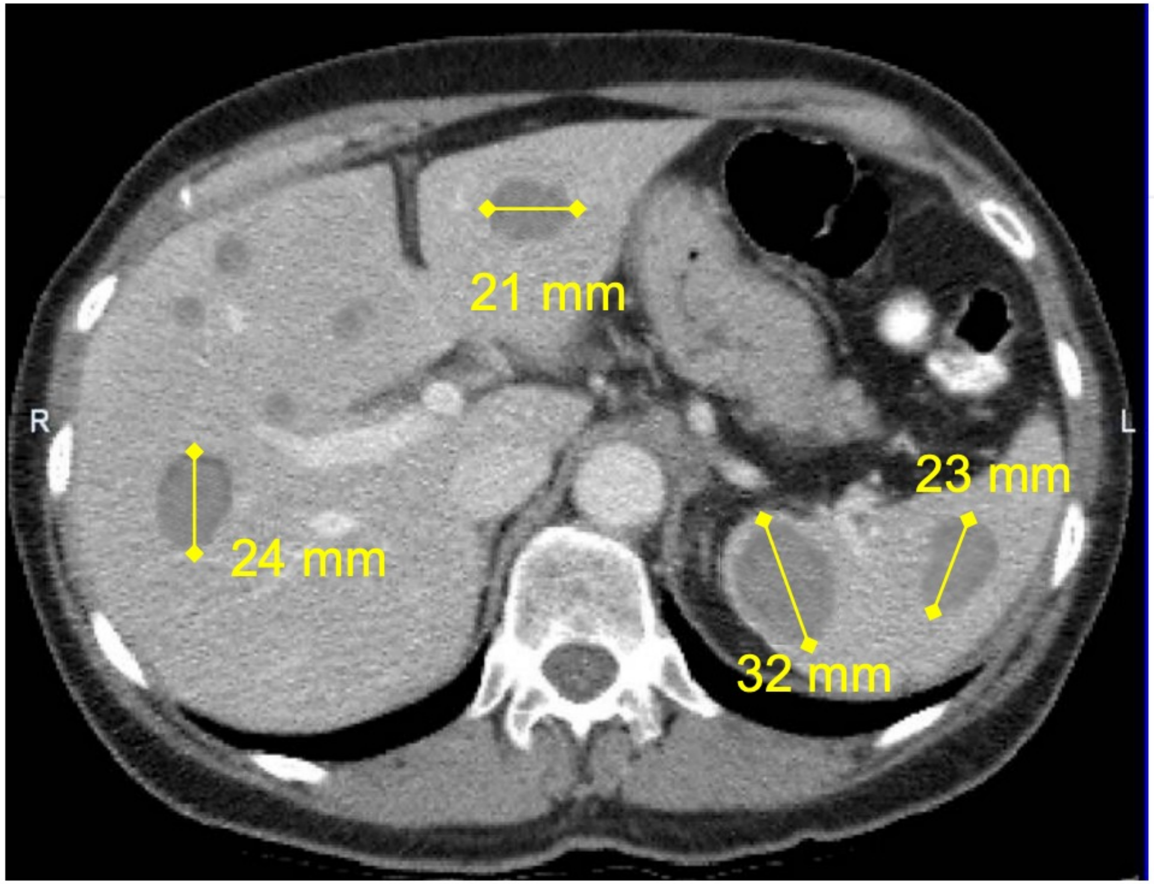
https://www.slicer.org /

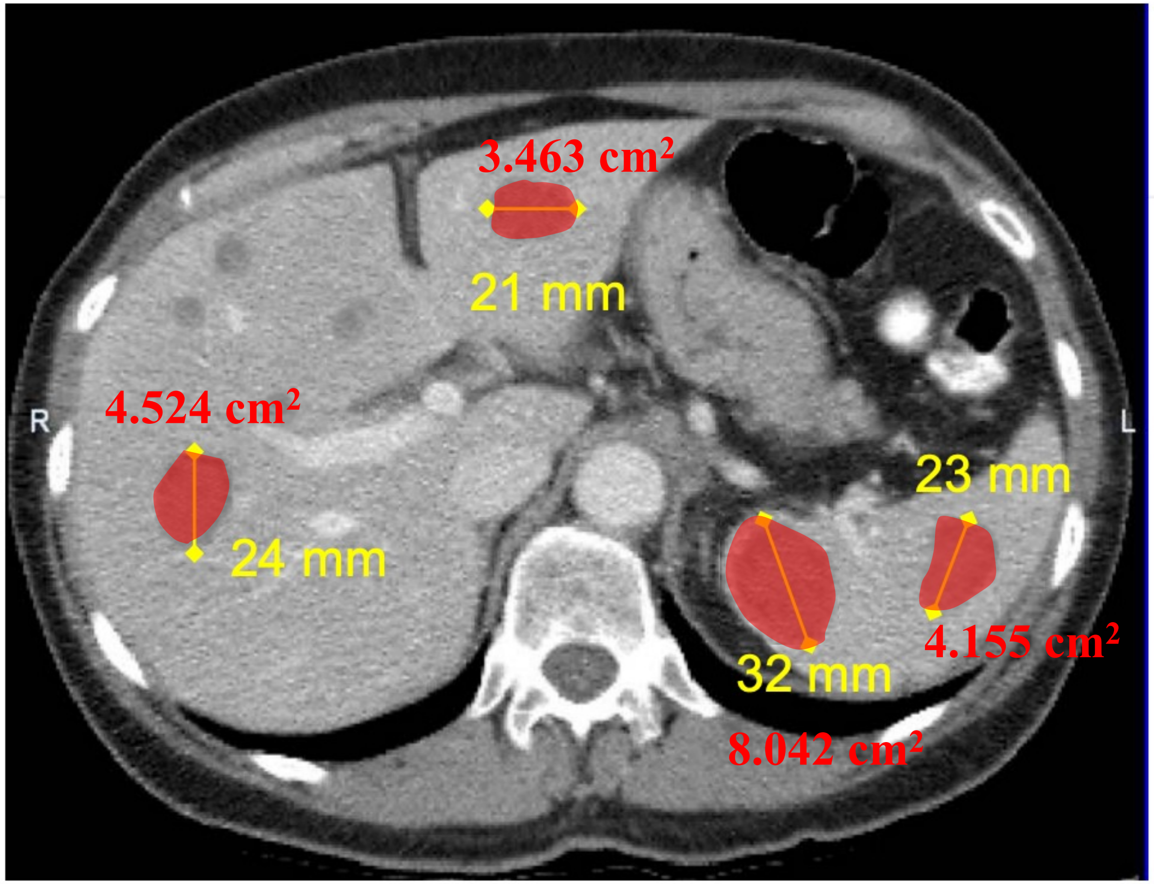
3D Slicer Integration



Next: Towards Next-generation Tumor Response Evaluation

Response Evaluation Criteria in Solid Tumors (RECIST)



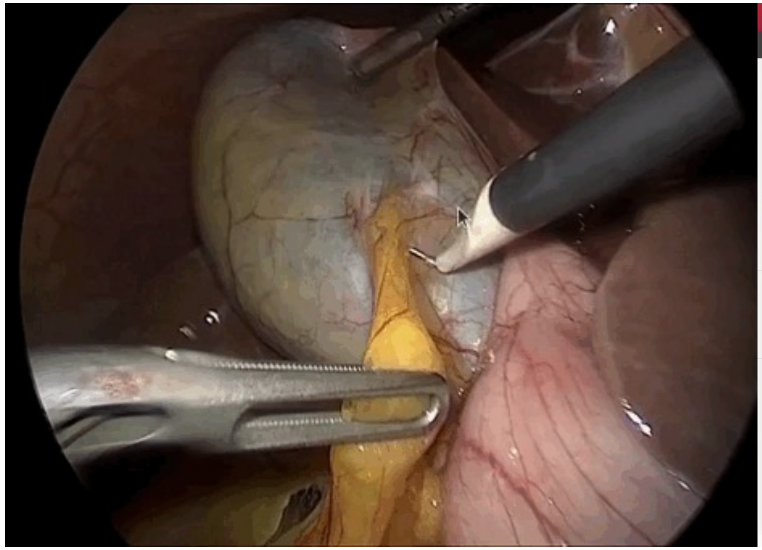


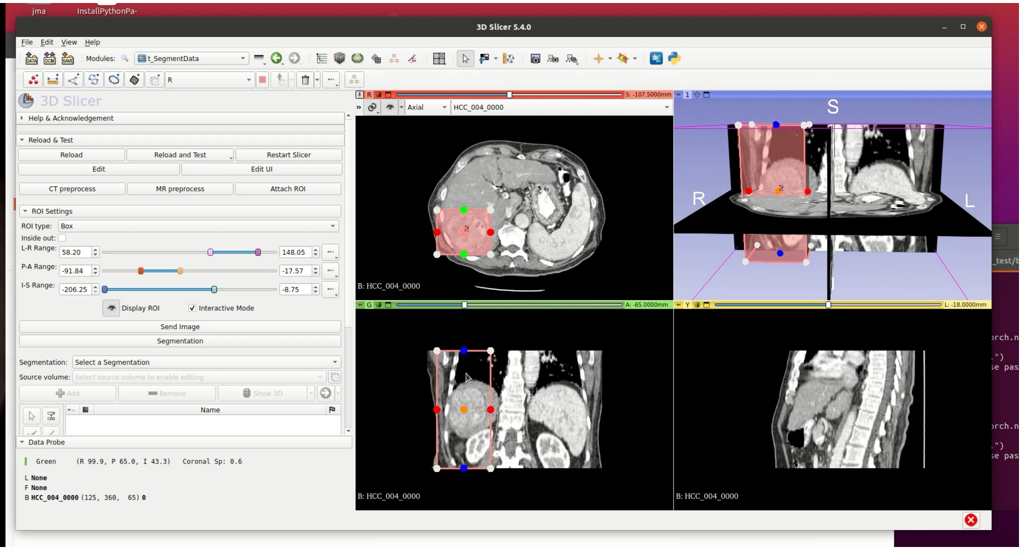
- Inaccurate: using a single diameter to measure the tumor
- Low reproducibility: impacted by reader experience, choice of target lesions, and lesion characteristics.

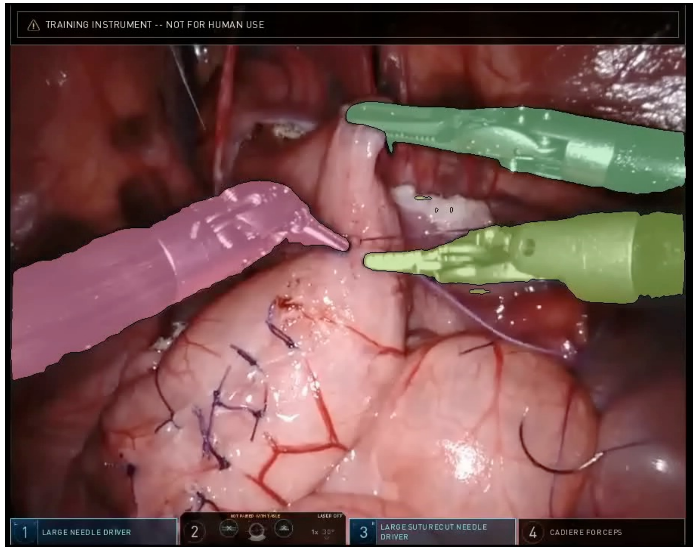
- A more authentic representation of tumor morphology.
- Enabling extraction of not just the largest diameter, but also a number of other tumor image biomarkers for holistic tumor quantifications.

Conclusion

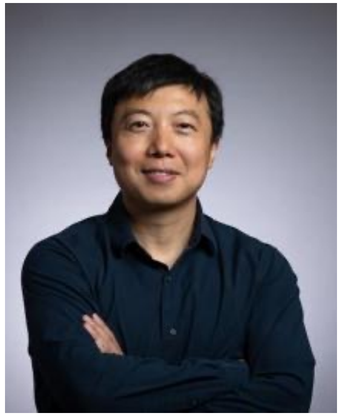
- ➢ MedSAM: The first foundation model for promptable medical image segmentation
- ➢ Lite MedSAM: from bench to bedside
- ➢ The trend in segmentation models is shifting towards greater flexibility: from closed-set to open-vocabulary, from isolated target to referring segmentation; from one modality to multi-







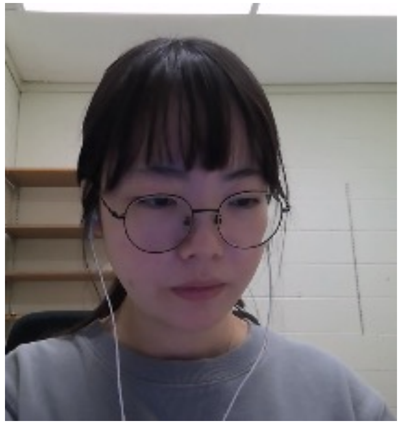
Acknowledgments



Prof. Bo Wang

Feifei Li
MedSAM,
Lite MedSAM

Beatrice Chen
MedDet



Sumin Kim
Image-based
Referring
Segmentation

Reza Asakereh

Andrew Qiao

Lite MedSAM plugin in Slicer

UNIVERSITY OF TORONTO

UHN Peter Munk Cardiac Centre AI Team

VECTOR INSTITUTE

Thanks for Listening!