Segment Any Cell: A SAM-based Auto-prompting Fine-tuning Framework for Nuclei Segmentation

Saiyang Na, Yuzhi Guo, Feng Jiang, Junzhou Huang, Hehuan Ma

University of Texas at Arlington

Contents

Introduction

Background – AI Research Evolution

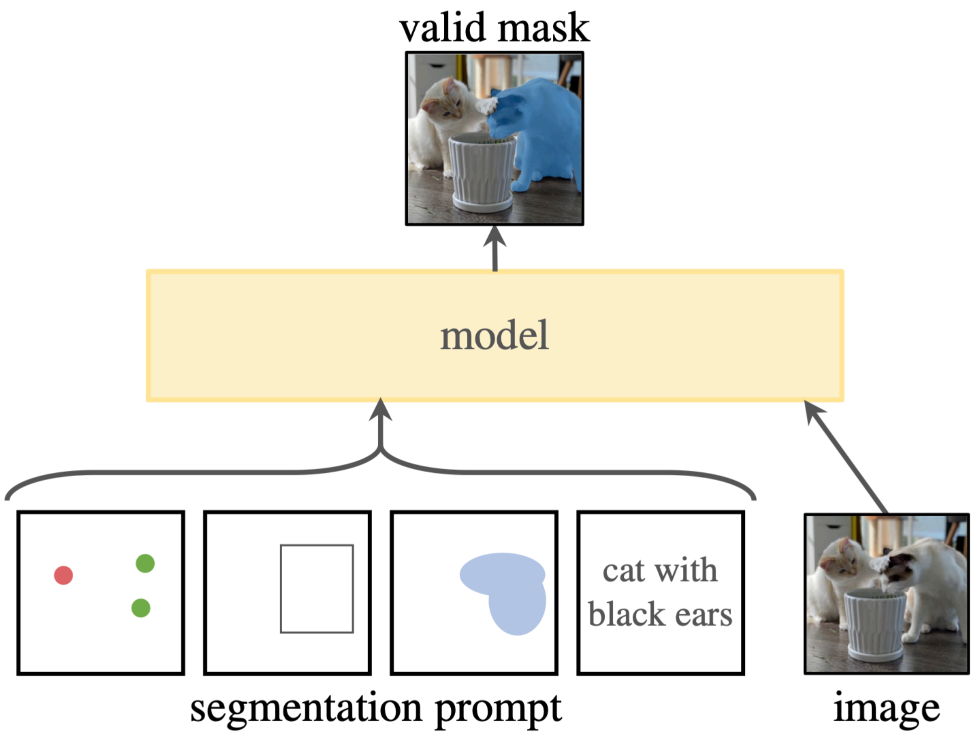
- Past: A well defined task, a well defined model — LeNet, AlexNet.
- Now:
  - Data-driven, large-scale model — BERT, GPT, ViT.
  - Those foundamental models can be adapted across a wide range of tasks and domains.

Introduction

SAM — Segment Anything Model

- **Purpose**: SAM is designed to perform image segmentation tasks with

a high degree of versatility, enabling it to understand and segment virtually any object or region within an image as instructed by user-generated textual prompts.



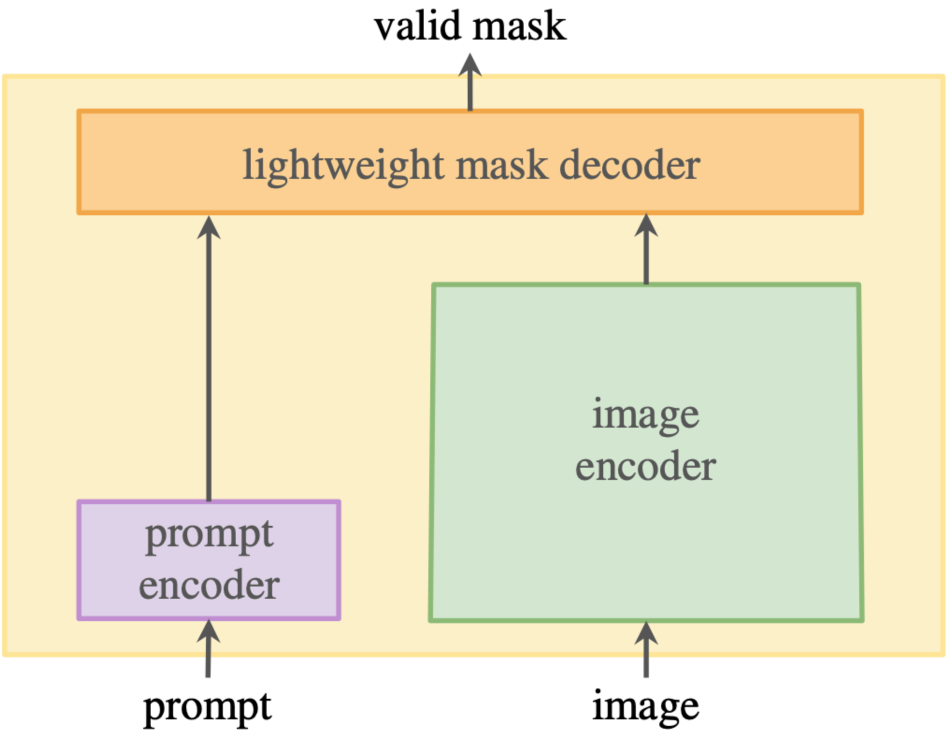
Promptable segmentation, figure from SAM.

Introduction

SAM — Segment Anything Model

- **Foundation**: At its core, SAM utilizes a transformer-based

architecture — VIT, similar to those found in state-of-the-art natural language processing (NLP) models. This design choice allows SAM to effectively process and interpret the complex relationships between visual elements and textual prompts.



Segment Anything Model, figure from SAM.

Introduction

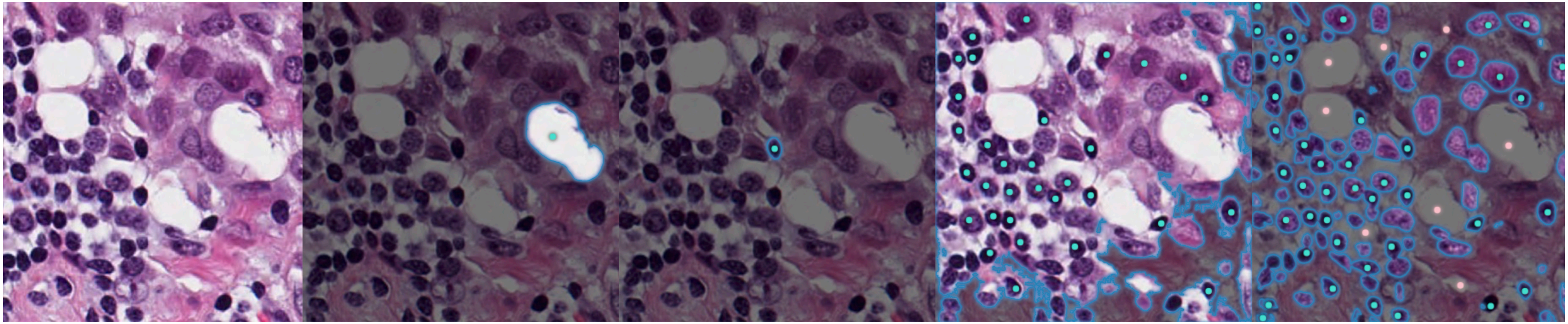
SAM Challenges in specialized areas like nuclei segmentation

Input image   Segmentation result

Natural Image Segmentation.

Introduction

SAM Challenges in specialized areas like nuclei segmentation



Input image | Unprofessional single prompt | Professional single prompt | Professional multiple prompts | Professional w/ negative prompts

Nuclei Segmentation.

Related works

1. Hörst, F. *et al.* CellViT: Vision Transformers for Precise Cell Segmentation and Classification. (2023)

- Which everages pre-trained ViT encoders, such as ViT256 and SAM,

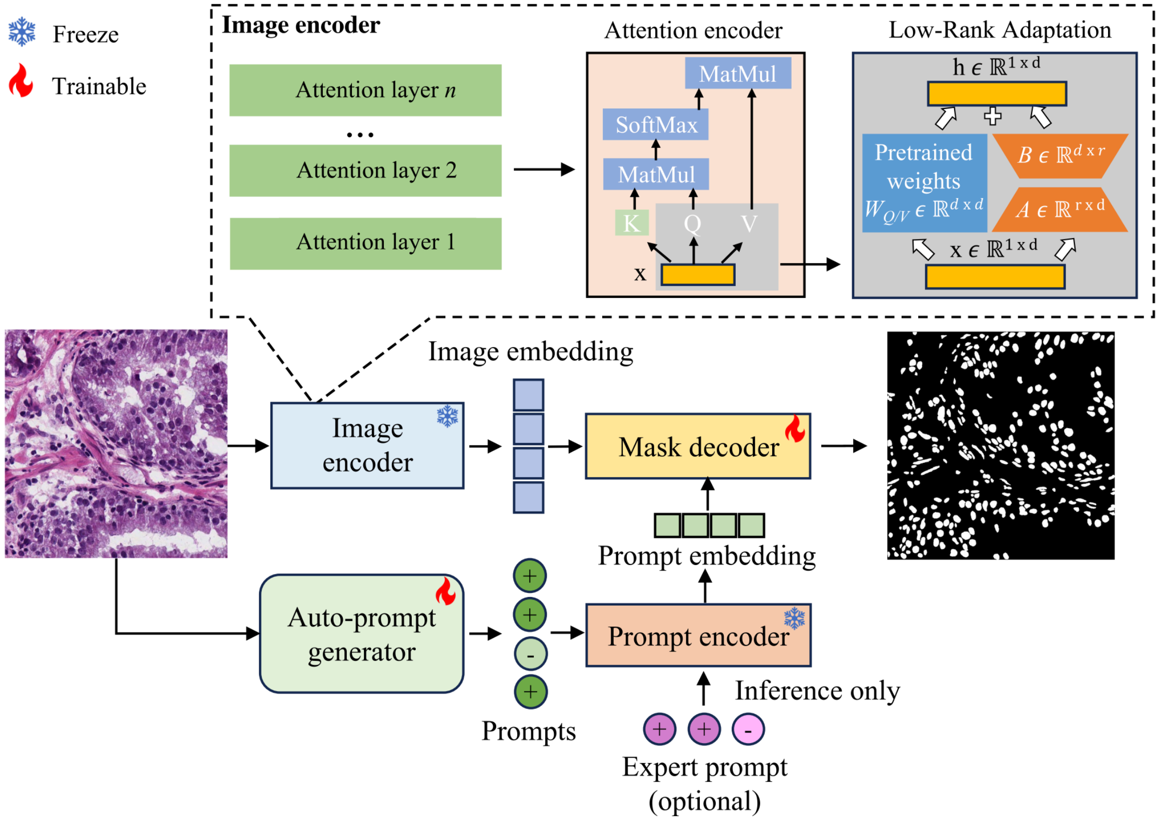
and its own decoder to do the cell segmentation.

- No prompt for the mask decoder.

2. Ma, J. *et al.* Segment Anything in Medical Images. *Nature Communications* **15,** 1–9 (2024)

- Which fine-tuning SAM with more than one million medical image-mask pairs.
- Still need professionals' to point out the nuclei in the prompt.

Methodology



Segment any cell framework.

Methodology

Auto-prompt generator

Methodology

Discriminating Strategy

Methodology

LoRA

Experiments

Dataset

- MoNuSeg
- 2018 Data Science Bowl (DSB)
- PanNuke

Experiments

MoNuSeg

Experiments

DSB

Experiments

PanNuke

Ablation Studies

Effectiveness of Prompts in SAM Fine-Tuning

Ablation Studies

Centroid-based prompt selection vs. Direct probability-based prompt selection

Ablation Studies

Efficiency Analysis

Ablation Studies

Enhancement Through Incremental Prompt Amplification

Ablation Studies

SAM VIT Backbone Comparison

Conclusion

- Segment Any Cell Framework
- Auto Prompting
- LoRA

Reference

1. Hörst, F. *et al.* CellViT: Vision Transformers for Precise Cell Segmentation and Classification. (2023)
2. Ma, J. *et al.* Segment Anything in Medical Images. *Nature Communications* **15,** 1–9 (2024)
3. Bommasani, R. *et al.* On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258* (2021)
4. Yu, J. *et al.* Coca: Contrastive captioners are image-text foundation models. *arXiv preprint arXiv:2205.01917* (2022)

5. Vaswani, A. *et al.* Attention is all you need. *Advances in neural information processing systems* **30,** (2017)

6. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018)

7. Brown, T. *et al.* Language models are few-shot learners. *Advances in neural information processing systems* **33,** 1877–1901 (2020)

8. Koubaa, A. GPT-4 vs. GPT-3.5: A concise showdown. (2023)

9. Hu, E. J. *et al.* Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685* (2021)

10. Zhang, R. *et al.* Llama-adapter: Efficient fine-tuning of language models with zero-init attention. *arXiv preprint arXiv:2303.16199* (2023)

11. Touvron, H. *et al.* Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288* (2023)

12. Kirillov, A. *et al.* Segment anything. *arXiv preprint arXiv:2304.02643* (2023)

13. Wu, J. *et al.* Medical sam adapter: Adapting segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.12620* (2023)

14. Chen, T. *et al.* SAM Fails to Segment Anything?–SAM-Adapter: Adapting SAM in Underperformed Scenes: Camouflage, Shadow, and More. *arXiv preprint arXiv:2304.09148* (2023)

15. Ma, J. & Wang, B. Segment anything in medical images. *arXiv preprint arXiv:2304.12306* (2023)

16. Sharma, Y., Syed, S. & Brown, D. E. Mani: Maximizing mutual information for nuclei cross-domain unsupervised segmentation. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* 345–355 (2022).

17. Yang, S., Zhang, J., Huang, J., Lovell, B. C. & Han, X. Minimizing labeling cost for nuclei instance segmentation and classification with cross-domain images and weak labels. in *Proceedings of the AAAI Conference on Artificial Intelligence* vol. 35 697–705 (2021).

18. Xu, Y. *et al.* Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC bioinformatics* **18,** 1–17 (2017)

19. Mahmood, F. *et al.* Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE transactions on medical imaging* **39,** 3257–3267 (2019)

20. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* 234–241 (2015).

21. He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask r-cnn. in *Proceedings of the IEEE international conference on computer vision* 2961–2969 (2017).

22. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 3431–3440 (2015).

23. Haq, M. M. & Huang, J. Self-supervised pre-training for nuclei segmentation. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* 303–313 (2022).

24. Lee, M. Y. *et al.* CellSeg: a robust, pre-trained nucleus segmentation and pixel quantification software for highly multiplexed fluorescence images. *BMC bioinformatics* **23,** 46–47 (2022)

25. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 770–778 (2016).

26. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014)
27. Dosovitskiy, A. *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929* (2020)
28. Deng, J. *et al.* Imagenet: A large-scale hierarchical image database. in *2009 IEEE conference on computer vision and pattern recognition* 248–255 (2009).
29. He, K. *et al.* Masked autoencoders are scalable vision learners. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* 16000–16009 (2022).
30. Kumar, N. *et al.* A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE transactions on medical imaging* **36,** 1550–1560 (2017)

31. Sirinukunwattana, K. *et al.* Gland segmentation in colon histology images: The glas challenge contest. *Medical image analysis* **35,** 489–502 (2017)

32. Caicedo, J. C. *et al.* Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nature methods* **16,** 1247–1253 (2019)

33. Wang, H., Cao, P., Wang, J. & Zaiane, O. R. Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. in *Proceedings of the AAAI conference on artificial intelligence* vol. 36 2441–2449 (2022).

34. Jha, D., Riegler, M. A., Johansen, D., Halvorsen, P. & Johansen, H. D. Doubleu-net: A deep convolutional neural network for medical image segmentation. in *2020 IEEE 33rd International symposium on computer-based medical systems (CBMS)* 558–564 (2020).

35. Valanarasu, J. M. J., Oza, P., Hacihaliloglu, I. & Patel, V. M. Medical transformer: Gated axial-attention for medical image segmentation. in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24* 36–46 (2021).

36. Wazir, S. & Fraz, M. M. HistoSeg: Quick attention with multi-loss function for multi-structure segmentation in digital histology images. in *2022 12th International Conference on Pattern Recognition Systems (ICPRS)* 1–7 (2022).

37. Tang, F. *et al.* DuAT: Dual-aggregation transformer network for medical image segmentation. *arXiv preprint arXiv:2212.11677* (2022)

38. Wang, J. *et al.* Stepwise feature fusion: Local guides global. in *International Conference on Medical Image Computing and Computer-Assisted Intervention* 110–120 (2022).

39. Srivastava, A. *et al.* MSRF-Net: a multi-scale residual fusion network for biomedical image segmentation. *IEEE Journal of Biomedical and Health Informatics* **26,** 2252–2263 (2021)

40. Tomar, N. K. *et al.* Fanet: A feedback attention network for improved biomedical image segmentation. *IEEE Transactions on Neural Networks and Learning Systems* (2022)

41. Pan, Z., Chen, J. & Shi, Y. Masked Diffusion as Self-supervised Representation Learner. *arXiv preprint arXiv:2308.05695* (2023)

42. Deng, R. *et al.* Segment anything model (sam) for digital pathology: Assess zero-shot segmentation on whole slide imaging. *arXiv preprint arXiv:2304.04155* (2023)

43. He, S., Bao, R., Li, J., Grant, P. E. & Ou, Y. Accuracy of segment-anything model (sam) in medical image segmentation tasks. *arXiv preprint arXiv:2304.09324* (2023)

44. Roy, S. *et al.* Sam. md: Zero-shot medical image segmentation capabilities of the segment anything model. *arXiv preprint arXiv:2304.05396* (2023)

45. Tang, L., Xiao, H. & Li, B. Can sam segment anything? when sam meets camouflaged object detection. *arXiv preprint arXiv:2304.04709* (2023)

46. Huang, Y. *et al.* Segment anything model for medical images?. *arXiv preprint arXiv:2304.14660* (2023)

47. Ji, G.-P. *et al.* SAM Struggles in Concealed Scenes–Empirical Study on" Segment Anything". *arXiv preprint arXiv:2304.06022* (2023)

48. Lee, P., Bubeck, S. & Petro, J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *New England Journal of Medicine* **388,** 1233–1239 (2023)

49. Nori, H., King, N., McKinney, S. M., Carignan, D. & Horvitz, E. Capabilities of gpt-4 on medical challenge problems. *arXiv preprint arXiv:2303.13375* (2023)

50. Raffel, C. *et al.* Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research* **21,** 5485–5551 (2020)

51. Good, I. J. Rational decisions. *Journal of the Royal Statistical Society: Series B (Methodological)* **14,** 107–114 (1952)

52. Lin, T.-Y., Goyal, P., Girshick, R., He, K. & Dollár, P. Focal loss for dense object detection. in *Proceedings of the IEEE international conference on computer vision* 2980–2988 (2017).

53. Milletari, F., Navab, N. & Ahmadi, S.-A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. in *2016 fourth international conference on 3D vision (3DV)* 565–571 (2016).

54. Gamper, J., Koohbanani, N. A., Benet, K., Khuram, A. & Rajpoot, N. PanNuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. in *European Congress on Digital Pathology* 11–19 (2019).

55. Gamper, J. *et al.* PanNuke Dataset Extension, Insights and Baselines. *arXiv preprint arXiv:2003.10778* (2020)

56. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. Cell Detection with Star-Convex Polygons. in *Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 - 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II* 265–273 (2018). doi:10.1007/978-3-030-00934-2_30

57. Weigert, M., Schmidt, U., Haase, R., Sugawara, K. & Myers, G. Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy. in *The IEEE Winter Conference on Applications of Computer Vision (WACV)* (2020). doi:10.1109/WACV45572.2020.9093435

58. Graham, S. *et al.* Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* **58,** 101563–101564 (2019)

59. Chen, S., Ding, C., Liu, M., Cheng, J. & Tao, D. CPP-Net: Context-Aware Polygon Proposal Network for Nucleus Segmentation. *IEEE Trans. Image Process.* **32,** 980–994 (2023)

60. Yao, K., Huang, K., Sun, J. & Hussain, A. PointNu-Net: Keypoint-Assisted Convolutional Neural Network for Simultaneous Multi-Tissue Histology Nuclei Segmentation and Classification. *IEEE Trans. Emerging Top. Comput. Intell.* **8,** 802–813 (2023)

61. Shui, Z. *et al.* Unleashing the Power of Prompt-driven Nucleus Instance Segmentation. *arXiv* (2023) doi:10.48550/arXiv.2311.15939