Mammo-SAM2 Framework for improving breast cancer diagnosis with enhanced medical image segmentation

zheng zhou
2036873

1. Abstract:

Breast mass segmentation will play a crucial role in assisting radiologists in the diagnosis of breast cancer. Although Segment Anything Model 2 (SAM2) has achieved remarkable success in natural image segmentation, its zero-shot performance on medical images, such as mammograms, remains suboptimal, falling short of clinical application requirements. Therefore, we aim to fine-tune SAM2 to enhance its performance in breast mass segmentation tasks. This study will propose a parameter-efficient fine-tuning framework, named Mammo-SAM2. This framework will leverage multi-scale feature adapters and a redesigned decoder to significantly improve SAM2's performance in mass segmentation tasks. We plan to conduct experiments on public medical datasets such as CBIS-DDSM and INbreast to validate the effectiveness of Mammo-SAM2. The expected experimental results will demonstrate that Mammo-SAM2 can surpass existing mass segmentation methods, achieving a new state-of-the-art level, providing more accurate and efficient tools for breast cancer diagnosis in medical image analysis.

2. Research Background:

Breast cancer is one of the leading causes of cancer-related deaths among women worldwide[7], and early detection and treatment are crucial for improving survival rates[8]. While mammography screening can detect breast cancer at early stages, manual screening is prone to errors and heavily relies on the radiologist's experience[4]. Automated breast mass segmentation[6, 3, 5] can help reduce diagnostic errors and improve diagnostic efficiency. However, the high resolution and complexity of mammogram images, as well as the subtlety of breast masses within normal glandular tissue, make automated segmentation tasks highly challenging.

Huang, Yuhao, et al.[2] demonstrate that while the zero-shot segmentation performance of SAM on medical images is somewhat limited compared to natural images, it still exhibits strong foundational segmentation capabilities under certain conditions. The research further reveals that by fine-tuning SAM for medical imaging tasks and incorporating manual prompts, the model's performance in handling complex medical image segmentation can be significantly improved. In addition, Huang, Yuhao, et al.[2] show some good cases of SAM to deal with medical images in the Fig. 1. These methods not only enhance the model's feature extraction capabilities but also effectively address the challenges posed by complex structures and diverse characteristics of medical images.

Subsequent studies have introduced methods such as Mammo-SAM[9], SAM Adapter[1], and SAMed[11], which focus on different fine-tuning strategies and structural optimizations, significantly enhancing SAM's performance in breast mass segmentation tasks. Mammo-SAM, for example, improves segmentation accuracy and efficiency by incorporating multi-scale feature adapters and multi-level decoders; SAM Adapter enhances the model's segmentation capability in complex scenarios by inserting convolutional adapter modules; SAMed uses low-rank adaptation techniques to enable efficient fine-tuning at a lower computational cost.

Building on these successful strategies, the development of SAM2[10] aims to further optimize model architecture and fine-tuning strategies to improve its adaptability and performance in medical image segmentation. By adopting more powerful adapter and decoder designs, flexible fine-tuning strategies, and enhanced cross-domain generalization capabilities, SAM2 is expected to achieve higher accuracy and efficiency in challenging tasks such as breast mass segmentation. These advancements demonstrate that foundational models, through careful design and optimization, can play an increasingly pivotal role in medical AI, providing stronger support for future clinical applications and research.

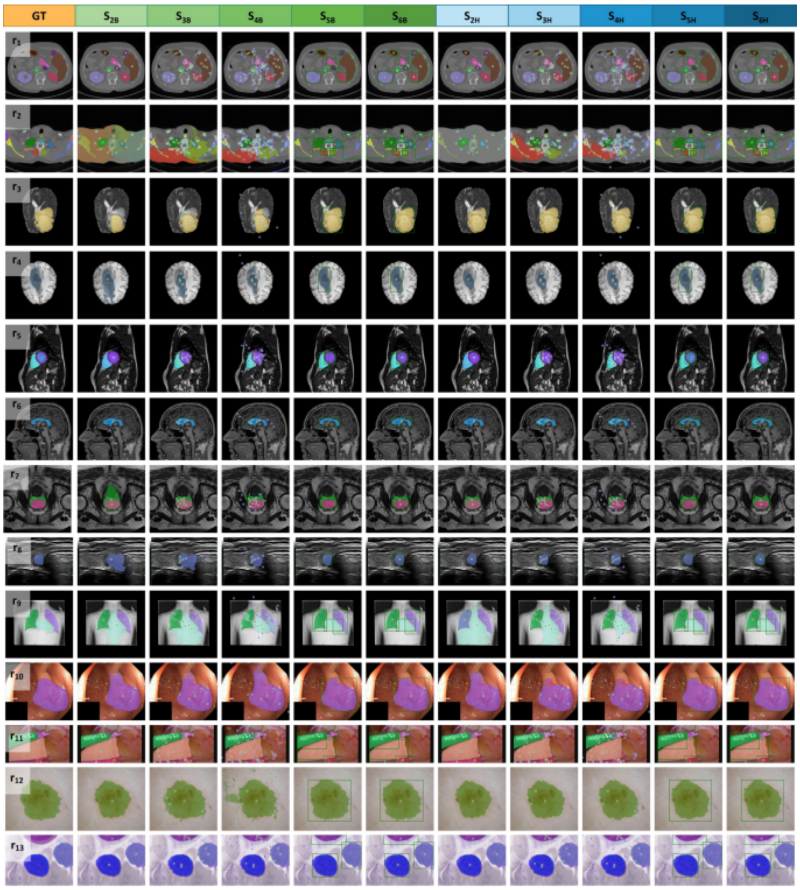

Figure 1: Typical good cases of SAM

3.Objectives and Research Questions:

The goal of this project is to develop a powerful segmentation framework that fully leverages the capabilities of the SAM2 model. This framework should address the unique challenges presented by breast cancer images, such as varying tumor shapes, sizes, and densities. We will attempt to replace the Hiera backbone network in SAM2 with a domain-specific pre-trained model from the medical field. This will enhance the model's ability to understand the unique characteristics of medical images and improve its performance in breast cancer image segmentation tasks.

Additionally, we will explore the integration of various adapters into the U-Net architecture, employing dimensionality reduction and expansion strategies to boost the model's efficiency and accuracy in handling medical images. Furthermore, we will investigate different fine-tuning methods, such as transfer learning and continual learning, to optimize the model's performance for specific breast cancer segmentation tasks. To address the challenge of limited labeled data, we will utilize data augmentation techniques to diversify the training set and employ generative models to create supplementary training data.

To reach our goal, several research questions will be answered in this project.

How can the integration of a domain-specific pre-trained model improve the SAM2-based framework for breast cancer image segmentation?

The integration of a domain-specific pre-trained model into the SAM2-based framework can significantly enhance its performance in breast cancer image segmentation tasks. A domain-specific model that has been pre-trained on a large dataset of medical images is already familiar with the unique characteristics of medical data, such as the textures, shapes, and contrast patterns typical of mammograms. By replacing the Hiera backbone in SAM2 with such a specialized model, the framework gains a better ability to discern subtle differences between healthy and cancerous tissues, which is crucial for accurate segmentation of breast tumors. This tailored approach improves the model's ability to handle the complex variability present in breast cancer images, such as diverse tumor shapes, sizes, and densities, leading to more precise and reliable segmentation outcomes.

What impact do various adapter designs and integration strategies have on the segmentation performance of the U-Net architecture within the SAM2 framework?

The design and integration of adapters within the U-Net architecture of the SAM2 framework can significantly influence its segmentation performance, particularly in handling medical images. Adapters serve as intermediate layers that can modify and optimize feature representations by applying dimensionality reduction and expansion strategies. By carefully designing these adapters, the model can efficiently manage the information flow between different network layers, enhancing its capacity to learn from complex and high-dimensional medical data. This approach not only improves the model's ability to capture fine-grained details in breast cancer images but also optimizes computational efficiency, enabling the model to perform well even in resource-constrained environments. Different adapter designs, such as those incorporating convolutional layers or attention mechanisms, can further refine feature extraction and boost overall segmentation accuracy, particularly when dealing with the nuanced characteristics of breast tumors in mammographic images.

4. Research Methods and Approach:

In this project, we will use a multi-level research approach to optimize the SAM2-UNet[10] model to make it more suitable for breast mass segmentation tasks. The research method mainly focuses on the following aspects:

Model fine-tuning and adaptation

As shown in Fig. 2, the core component of the SAM2 model is its Encoder Block, which utilizes multiple Hiera Blocks. These blocks combine normalization layers, attention layers, and multi-layer perceptrons (MLPs) to progressively extract multi-scale features from the images. Our research approach will focus on replacing the Hiera backbone network in SAM2 with a domain-specific pre-trained model to enhance its ability to understand and process medical image features more effectively.

The adapter module in SAM2-UNet optimizes feature representation through dimensionality reduction, activation functions, and dimensionality expansion, enhancing computational efficiency and non-linear expressiveness. To better adapt to breast mass segmentation tasks, we can introduce more complex non-linear transformations, multi-scale adapters, adaptive strategies, and regularization techniques to improve the model's understanding and segmentation accuracy for complex medical image features.

Weak Supervision and Data Augmentation

To address the challenge of limited labeled data, we will employ weak supervision and data augmentation techniques. By using image-level labels and partially annotated data, combined with weak supervision strategies, we aim to reduce annotation costs while maximizing the use of available data resources to enhance the model's learning efficiency. Additionally, we will apply various data augmentation techniques, such as rotation, flipping, scaling, and contrast adjustment, to increase the diversity of the training set.

Model evaluation and optimization

To ensure the optimal performance of the model, we will conduct extensive evaluation and optimization. Performance evaluation will be carried out on multiple public breast cancer datasets, such as CBIS-DDSM and INbreast, using standard segmentation metrics like Dice coefficient, IoU, precision, and recall to measure the model's segmentation performance. Hyperparameter optimization and experimental design will be performed using techniques such as grid search or Bayesian optimization to determine the best training parameter configurations. Additionally, we will conduct ablation studies to analyze the impact of different model components and design choices on the final performance.

Cross-dataset validation and generalization ability testing

We will further validate the generalization ability of the fine-tuned SAM2-UNet model across different datasets. Cross-dataset validation will involve testing the model on multiple mammography datasets with varying image quality, imaging conditions, and patient populations to evaluate its robustness and adaptability in different environments. Model robustness analysis will be conducted by assessing the model's performance when dealing with different types of image artifacts and abnormalities, such as noise, blurring, and low contrast.

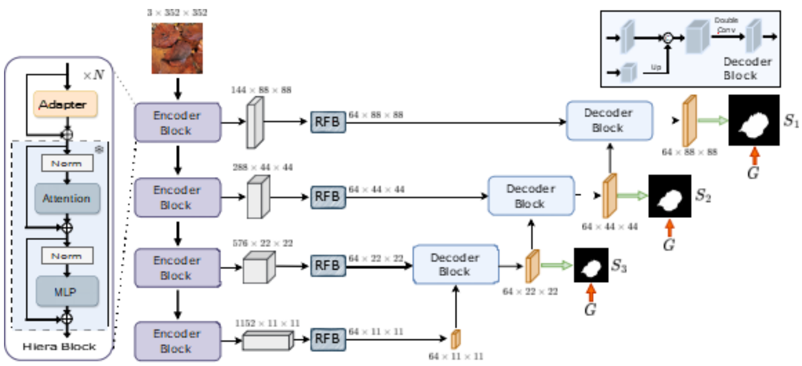

Figure 2: Overview of the proposed SAM2-UNet.

5.Timeline

Based on a 12-week preparation period, the following timeline outlines the key activities and milestones for the project:

- **Weeks 1-2: Preliminary Research and Planning**
  - Conduct a comprehensive literature review on breast mass segmentation and SAM2 model adaptation.
  - Define the project scope, objectives, and research questions.
  - Set up the development environment and acquire necessary datasets (e.g., CBIS-DDSM, INbreast).
- **Weeks 3-4: Data Preparation and Initial Model Setup**
  - Preprocess the datasets (e.g., data cleaning, normalization, augmentation).
  - Prepare training and validation sets.
  - Set up the baseline SAM2-UNet model with the original Hiera backbone.
- **Weeks 5-6: Model Adaptation and Integration**
  - Replace the Hiera backbone with a domain-specific pre-trained model.

- Integrate the new model into the SAM2-UNet architecture.
- Implement and integrate adapter modules with the U-Net architecture.

- **Weeks 7-8: Training and Hyperparameter Optimization**
  - Train the adapted SAM2-UNet model on the prepared datasets.
  - Perform hyperparameter optimization using grid search or Bayesian optimization to find the best configurations.
- **Weeks 9-10: Model Evaluation and Analysis**
  - Evaluate the model's performance on multiple public datasets using segmentation metrics (Dice coefficient, IoU, precision, recall).
  - Conduct ablation studies to analyze the impact of different components and configurations.
- **Week 11: Cross-Validation and Robustness Testing**
  - Perform cross-dataset validation to assess model generalization across different datasets.
  - Test model robustness against various image artifacts (noise, blurring, low contrast).
- **Week 12: Final Review and Reporting**
  - Compile results and analyze findings.
  - Prepare a final report summarizing the methodology, results, and potential future work.
  - Submit the final report and prepare for any presentations or follow-up discussions.

References

[1] Tianrun Chen, Lanyun Zhu, Chaotao Ding, Runlong Cao, Yan Wang, Zejian Li, Lingyun Sun, Papa Mao, and Ying Zang. Sam fails to segment anything?–sam-adapter: Adapting sam in underperformed scenes: Camouflage, shadow, medical image segmentation, and more. *arXiv preprint arXiv:2304.09148*, 2023.

[2] Yuhao Huang, Xin Yang, Lian Liu, Han Zhou, Ao Chang, Xinrui Zhou, Rusi Chen, Junxuan Yu, Jiongquan Chen, Chaoyu Chen, et al. Segment anything model for medical images? *Medical Image Analysis*, 92:103061, 2024.

[3] Qingqiu Li, Jilan Xu, Runtian Yuan, Yuejie Zhang, and Rui Feng. Scsgnet: Spatial-correlated and shape-guided network for breast mass segmentation. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2023.

[4] Michael G Marmot, DG Altman, DA Cameron, JA Dewar, SG Thompson, and Maggie Wilcox. The benefits and harms of breast cancer screening: an independent review. *British journal of cancer*, 108(11):2205–2240, 2013.

[5] Jiande Pi, Yunliang Qi, Meng Lou, Xiaorong Li, Yiming Wang, Chunbo Xu, and Yide Ma. Fs-unet: Mass segmentation in mammograms using an encoder-decoder architecture with feature strengthening. *Computers in Biology and Medicine*, 137:104800, 2021.

[6] Hui Sun, Cheng Li, Boqiang Liu, Zaiyi Liu, Meiyun Wang, Hairong Zheng, David Dagan Feng, and Shanshan Wang. Aunet: attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms. *Physics in Medicine & Biology*, 65(5):055005, 2020.

[7] Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3):209–249, 2021.

[8] László Tabár, Bedrich Vitak, Tony Hsiu-Hsi Chen, Amy Ming-Fang Yen, Anders Cohen, Tibor Tot, Sherry Yueh-Hsia Chiu, Sam Li-Sheng Chen, Jean Ching-Yuan Fann, Johan Rosell, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*, 260(3):658–663, 2011.

[9] Xinyu Xiong, Churan Wang, Wenxue Li, and Guanbin Li. Mammo-sam: Adapting foundation segment anything model for automatic breast mass segmentation in whole mammograms. In *International Workshop on Machine Learning in Medical Imaging*, pages 176–185. Springer, 2023.

[10] Xinyu Xiong, Zihuang Wu, Shuangyi Tan, Wenxue Li, Feilong Tang, Ying Chen, Siying Li, Jie Ma, and Guanbin Li. Sam2-unet: Segment anything 2 makes strong encoder for natural and medical image segmentation. *arXiv preprint arXiv:2408.08870*, 2024.

[11] Kaidong Zhang and Dong Liu. Customized segment anything model for medical image segmentation. *arXiv preprint arXiv:2304.13785*, 2023.