Automated Intraoperative Lumpectomy Margin Detection using SAM-Incorporated Forward-Forward Contrastive Learning

Tyler Ward[1] tyler.ward@uky.edu
Braxton McFarland[1] braxtonjmcfarland@gmail.com
Sahar Nozad[1] sahar.nozad@uky.edu
Talal Arshad[1] t.arshad@uky.edu
Hafsa Nebbache[1] hafsa.nebbache@uky.edu
Jin Chen[2] jinchen@uab.edu
Xiaoqin Wang[1] xiaoqin.wang@uky.edu
Abdullah Imran[1] aimran@uky.edu
[1] *University of Kentucky, Lexington, KY 40506, USA*
[2] *University of Alabama at Birmingham, AL 35233, USA*

Editors: Under Review for MIDL 2025

Abstract

Complete removal of cancerous tumors with a negative specimen margin during lumpectomy is essential to reduce breast cancer recurrence. However, interpretation of 2D specimen radiography (SR) by radiologists, the current method used to assess intraoperative margin status, has limited accuracy. This study aims to improve positive margin detection performance on SRs by leveraging cutting-edge deep learning models. We developed a novel lumpectomy margin assessment method using an innovative pre-training technique, Forward-Forward Contrastive Learning (FFCL), followed by few-shot segmentation training leveraging the Segment Anything Model 2 (SAM 2). Experimental results on independent annotated breast SRs demonstrate the effectiveness of the proposed FFCL-SAM method in classifying and segmenting positive margins in SRs.

Keywords: Breast, Lumpectomy, Specimen radiography, Positive margin, Forward-Forward Contrastive learning

1. Introduction

Accurate intraoperative detection of positive margins during breast-conserving surgery (BCS) is critical to minimizing cancer recurrence and avoiding repeat surgeries (Siegel et al., 2025). Radiography-based inspection, the standard practice during BCS for margin assessment, offers limited sensitivity and often fails to reliably differentiate between tumor tissue and surrounding structures (Scimone et al., 2021). This leads to roughly 20–25% of patients requiring re-excision (Scimone et al., 2021).

Recent advancements in deep learning have shown promise in medical imaging, but most margin assessment approaches rely on supervised learning and require extensive annotation, which is labor-intensive and difficult to scale (To et al., 2023). Moreover, foundation models such as the Segment Anything Model (SAM) (Kirillov et al., 2023; Ravi et al., 2024), are not tailored to specimen radiographs (SRs) and typically require manual prompting.

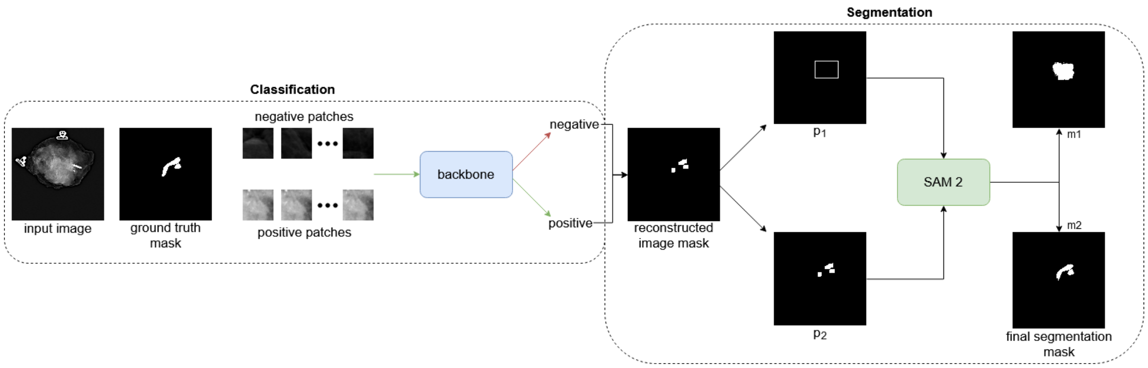

Figure 1: Illustration of the breast cancer margin detection pipeline consisting of both classification and segmentation tasks.

In this work, we propose FFCL-SAM, a novel method that combines self-supervised Forward-Forward Contrastive Learning (FFCL) (Ahamed et al., 2023) with SAM for automated, accurate lumpectomy margin detection. Our method first classifies radiograph patches using an FFCL-pretrained classification model and then uses the predictions to reconstruct masks from the patches and train SAM under a few-shot setting for refined pixel-level segmentation. This integrated approach improves detection performance and may reduce re-excision rates, with fast inference suitable for intraoperative use.

2. Methods

FFCL employs the Forward-Forward Algorithm (FFA) (Hinton, 2022) and performs two-stage pre-training (local followed by global contrastive learning) without requiring any back-propogation. To generate segmentation masks from FFCL-based classification, we reconstruct images from their patches in the form of a binary mask, where an activation in the mask indicates a positive prediction, and no activation indicates the opposite. To improve the accuracy of these coarse reconstructed masks, we refine them using SAM 2, which we fine-tune using a small labeled set of SRs to aid better approximation of the ground truth margin annotations. Once fine-tuned, SAM 2 receives two prompts based on the class predictions from the FFCL pipeline: a bounding box encompassing the predicted region and the coarse binary mask itself, allowing the model to leverage both spatial and form information when making its prediction.

3. Experiments and Results

In this HIPPA (Health Insurance Portability and Accountability Act)-compliant retrospective study, intraoperative SRs were collected from 215 lumpectomy procedures performed between 2019 and 2022 at the University of Kentucky and the Markey Cancer Center Comprehensive Breast Care Center. Institutional Review Board (IRB) approval was obtained, and all images were de-identified prior to use. Samples were excluded if they lacked anatomic orientation, intraoperative labeling, or pathology-confirmed malignancy. Tumor localization

Table 1: Classification (top) and segmentation (bottom) performance. Best scores are **bolded** and second best are underlined.

Classification Results					
Model	Accuracy	F1	Precision	Recall	AUC
ConvNeXt w/ FFCL	0.7510 ± 0.0029	0.6396 ± 0.0100	0.6542 ± 0.0186	<u>0.6322 ± 0.0072</u>	0.7751 ± 0.0002
ViT w/ FFCL	**0.7816 ± 0.0316**	<u>0.6898 ± 0.0373</u>	**0.7882 ± 0.0582**	0.6006 ± 0.0387	<u>0.8067 ± 0.0530</u>
ResNet-18 w/ FFCL	<u>0.7815 ± 0.0174</u>	**0.6914 ± 0.0238**	<u>0.7478 ± 0.0656</u>	**0.6900 ± 0.0142**	**0.8455 ± 0.0152**

Segmentation Results									
Mask	Positive Margin			Negative Margin			Overall		
	DSC	HD	Accuracy	DSC	HD	Accuracy	DSC	HD	Accuracy
Reconstructed Mask	0.6313	85.3282	0.8157	0.6861	179.5214	0.8431	0.6587	132.4248	0.8294
SAM-zero shot	0.5206	87.4305	0.7603	0.7631	307.6746	0.8816	0.6419	197.5535	0.8210
SAM2-zero-shot	0.8861	6.8267	0.9431	0.9220	9.3258	0.9610	0.9041	8.0763	0.9521
SAM2-few-shot	**0.9053**	**6.6736**	**0.9527**	**0.9443**	**9.0721**	**0.9722**	**0.9248**	**7.8729**	**0.9625**

and 3D orientation were confirmed using diagnostic mammograms and post-biopsy imaging. Each SR was annotated with three region labels: nonmalignant tissue, tumor presence, and pathology-confirmed positive margin. Positive margin labels were determined following guidelines from the Society of Surgical Oncology, American Society of Clinical Oncology, and American Society of Radiation Oncology. The final dataset included 92 SR images from 46 patients, which were split into training, validation, and test sets (38/2/6).

For the classification task, each SR was dividing into overlapping 64x64 patches (stride=3). To address class imbalance, fewer patches were extracted from negative regions, and for validation and testing, only patches from positive margin regions were used to better assess generalization. For the segmentation task, five SRs with corresponding ground-truth masks were used to fine-tune SAM 2 in a few-shot setting. The coarse masks reconstructed from patch-level predictions were refined using SAM 2, which received both bounding box and mask prompts derived from the FFCL predictions.

Table 1 summarizes the performance of the classification and segmentation approaches. Among the FFCL-pretrained backbones, ResNet-18 achieved the best classification performance, with top scores in F1, recall, and AUC, and second-best accuracy and precision. For segmentation, zero-shot SAM 2 already showed strong performance over the coarse baseline, but few-shot fine-tuning yielded the best results overall on Dice similarity coefficient (DSC), Hausdorff distance (HD), and accuracy.

4. Conclusions

FFCL-SAM achieves high classification performance for detecting positive margins in SRs. Evaluation of the subsequent segmentation task demonstrates that SAM 2 under a few-shot training setting shows promise in refining segmentation masks, improving in performance compared to the zero-shot settings. The findings from this study indicate the potential to reduce the rate of lumpectomy re-excision and improve patient outcomes through the use of deep learning methods, particularly the integration of domain-specific pre-training with a generic foundation model.

References

Md Atik Ahamed, Jin Chen, and Abdullah-Al-Zubaer Imran. Forward-forward contrastive learning. *arXiv preprint arXiv:2305.02927*, 2023.

Geoffrey Hinton. The forward-forward algorithm: Some preliminary investigations. *arXiv preprint arXiv:2212.13345*, 2(3):5, 2022.

Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4015–4026, 2023.

Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, et al. SAM 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*, 2024.

Mark T Scimone, Savitri Krishnamurthy, Gopi Maguluri, Dorin Preda, Jesung Park, John Grimble, Min Song, Kechen Ban, and Nicusor Iftimia. Assessment of breast cancer surgical margins with multimodal optical microscopy: A feasibility clinical study. *PLoS One*, 16(2):e0245334, 2021.

Rebecca L Siegel, Tyler B Kratzer, Angela N Giaquinto, Hyuna Sung, and Ahmedin Jemal. Cancer statistics, 2025. *CA*, 75(1):10, 2025.

Tyrell To, Tongtong Lu, Julie M Jorns, Mollie Patton, Taly Gilat Schmidt, Tina Yen, Bing Yu, and Dong Hye Ye. Deep learning classification of deep ultraviolet fluorescence images toward intra-operative margin assessment in breast cancer. *Frontiers in Oncology*, 13: 1179025, 2023.