Pixel Perfect: Exploring the Segment Anything Model

Isioma Emordi   Arnav Jain   Rami Kamel   Kashir Khan   Andy Vayalali

Ermina Ashraf   Dr. Ovidiu Daescu

Introduction

Detecting and segmenting pneumonia in chest X-rays is a complex challenge in medical imaging due to the subtle and often ambiguous nature of the disease's manifestations. While models like the Segment Anything Model (SAM) perform well on general segmentation tasks, their application to medical images, such as chest X-rays, frequently shows limitations.

This ablation study examines the effectiveness of each preprocessing method individually and in combination, providing insights into how such techniques influence SAM's segmentation accuracy for pneumonia detection in a clinical imaging context. Additionally, we introduce custom evaluation metrics, such as segmentation consistency checks without ground truth, to assess the comparative impact of these methods. By isolating and analyzing each preprocessing technique, this work aims to contribute a deeper understanding of SAM's adaptability to medical imaging .

Preprocessing Methods

Our first approach for image pre-processing was Canny Edge Detection (CED) and Contrast-Limited Adaptive Histogram Equalization (CLAHE). CED identifies clear, continuous edges by smoothing the image and filtering pixels based on thresholding and edge connectivity. CLAHE enhances image contrast by dividing an image into small regions, applying histogram equalization to each, and limiting contrast to reduce noise. These techniques were useful in making the pneumonia-affected regions more easily identifiable, without being overly sensitive to the noisy character of pneumonia-affected chest X-rays. We later added Gaussian blur, entropy, and SSIM optimization. Gaussian blur reduces noise by averaging pixel values with a function gives more weight to central pixels in a specified neighborhood. The entropy and SSIM optimization were used to maximize information content by adjusting image features to enhance clarity and still preserve each image's structural information.

Our Dataset

The dataset we used was the RSNA Pneumonia Detection Challenge dataset on Kaggle. It contains over 29,000 images total, consisting of around 6000 images of chest X-rays that displayed likely cases of pneumonia and a CSV file that marked the location of box annotations surrounding the pneumonia affected regions.

SAM Model

The Segment Anything Model (SAM) by Meta AI redefines object segmentation with its "prompt-anything" capability, using minimal input, such as bounding boxes or points, to identify and segment objects. SAM's architecture, combining an image encoder, prompt encoder, and mask decoder, enables it to perform well even on unfamiliar objects. However, SAM's ability to handle complex segmentation in medical imaging remains uncertain.

In our work, we applied SAM on pneumonia chest X-rays, using bounding box annotations to generate masks of affected regions. To improve segmentation accuracy, we converted these masks into contours, allowing detailed focus on boundaries around pneumonia areas. By isolating edges, we refined the segmentation to capture subtle details in affected regions. This contour-based approach helps analyze SAM's capability to maintain edge precision, a crucial factor for accurate medical segmentation.
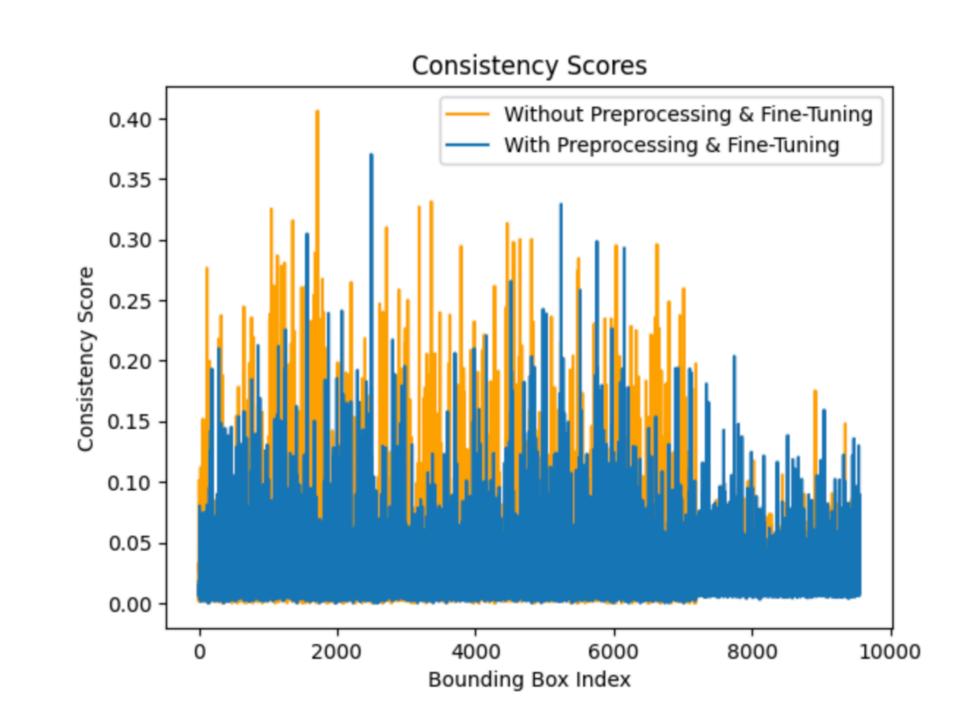
Results

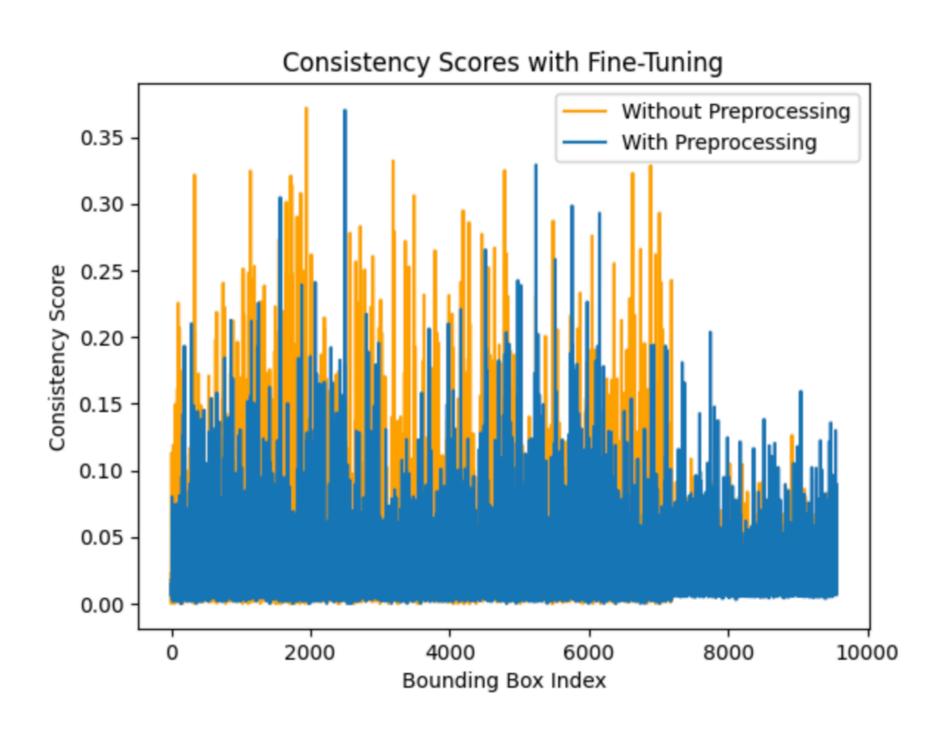Figure 1. Consistency scores with image preprocessing & SAM fine-tuning
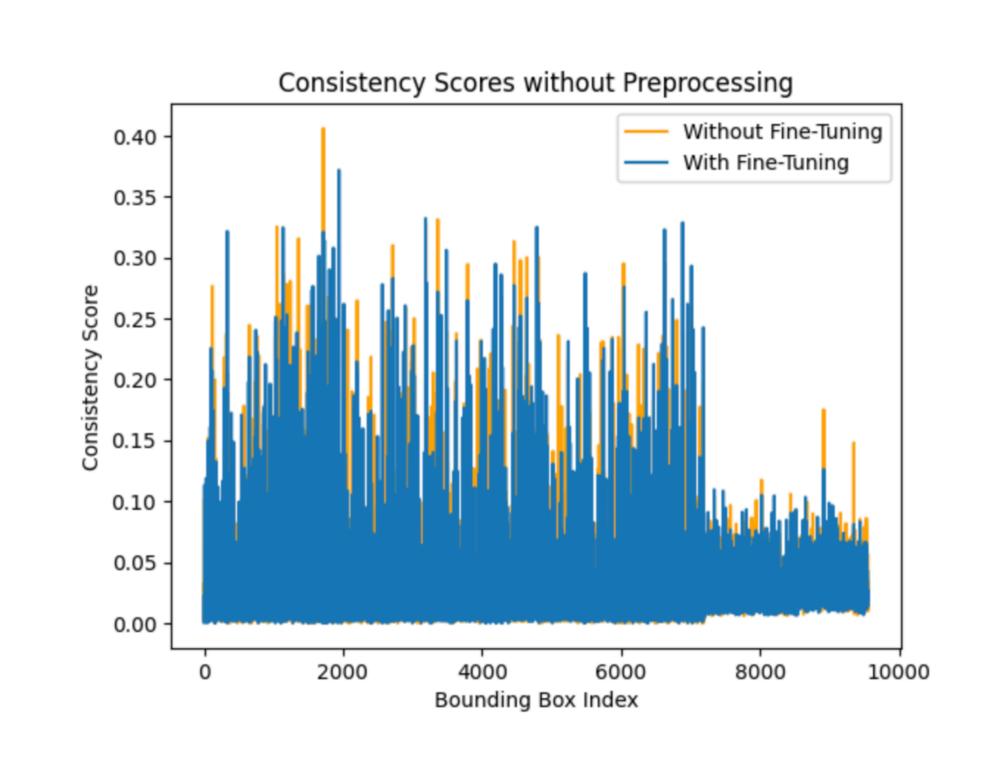

Figure 2. Consistency scores without image processing
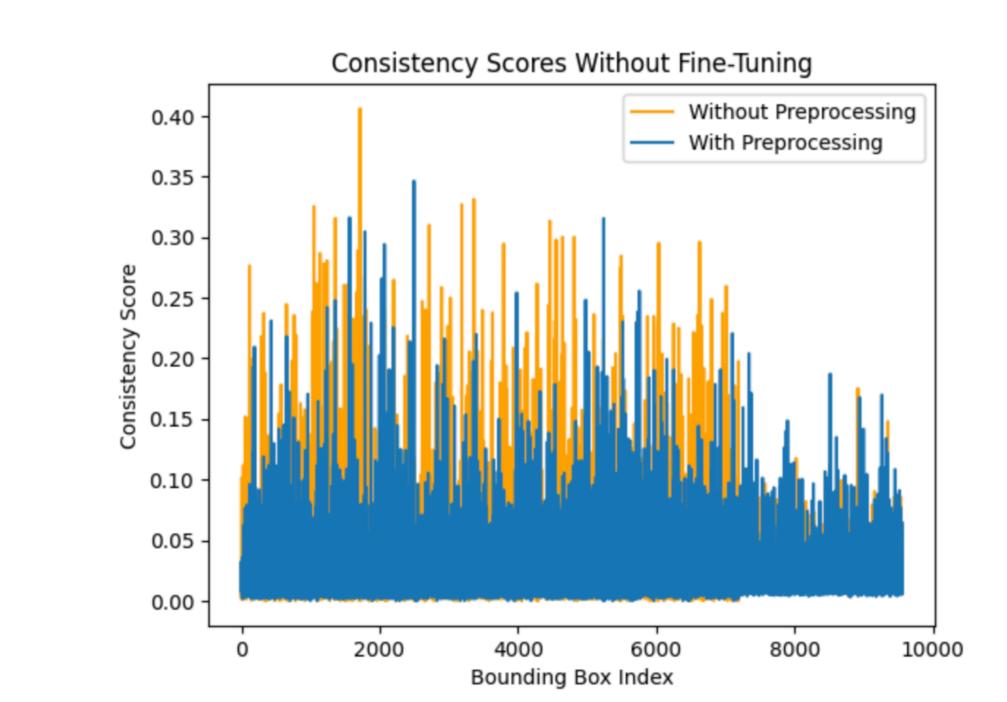

Figure 3. Consistency scores with SAM fine-tuning

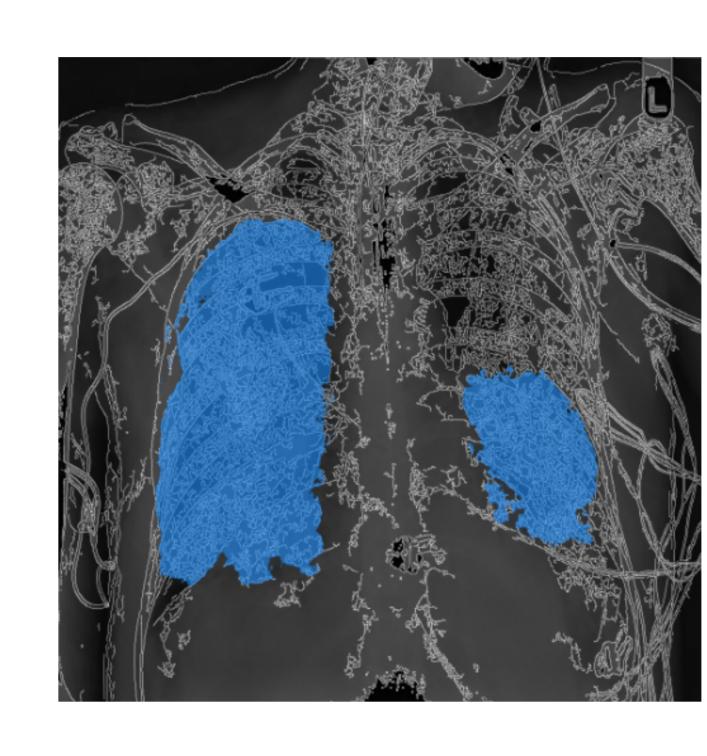Figure 4. Consistency scores without SAM fine-tuning

Analysis and Comparison
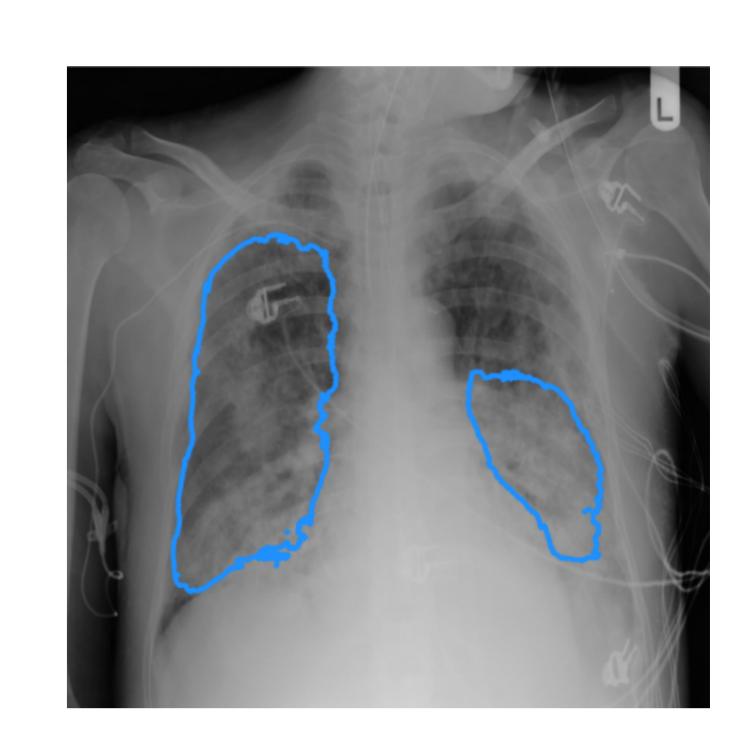

Figure 5. Masked image after preprocessing

Figure 6. SAM contoured image

To assess SAM's transferability, we used a linear regression model to predict the consistency of SAM scores across various bounding box perturbations as a measure of segmentation stability. After applying preprocessing, we observed a significant decrease in consistency score values for images masked by the SAM model, indicating increased consistency and thereby improved accuracy. However, applying additional training methods to the model had minimal impact on the consistency of the generated masks.

Conclusion

We found that, by employing image preprocessing and model training techniques, we can improve SAM's accuracy and consistency, thus enhancing its ability to be effectively applied to medical imaging contexts.We found that that by employing image preprocessing and model training techniques. However, the presence of numerous outliers suggests that further fine-tuning is necessary. The results demonstrate the potential for optimized preprocessing and training protocols to improve the model's performance in clinical settings.

References

Kohlberger, T., Singh, V., Alvino, C., Bahlmann, C., Grady, L. (2012). Evaluating segmentation error without ground truth. Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, 15(Pt 1), 528–536.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, A., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P. (2023). Segment Anything.

Nneji, G. U., Cai, J., Deng, J., Monday, H. N., James, E. C., Ukwuoma, C. C. (2022). Multi-Channel Based Image Processing Scheme for Pneumonia Identification. Diagnostics (Basel, Switzerland), 12(2), 325.