Conference on Artificial Intelligence in Medical Imaging
SIIM CAIMI26
OCT 26-27 | University of Pennsylvania, Philadelphia, PA

Fully Automated Left Ventricular Segmentation with LoRA-Adapted UltraSAM and Learned Prompt Generation

Ryan Chang-Wu, Belmont Hill School; Kaan Sel, PhD

Introduction/Background

Accurate ejection-fraction estimation depends on reliable left-ventricular (LV) segmentation at end-diastole (ED) and end-systole (ES). EchoNet-Dynamic established an effective automated baseline using DeepLabv3, but performance can be further improved, where small contour errors can disproportionately affect ventricular volume and ejection-fraction estimates. UltraSAM offers ultrasound-specific foundation-model representations that can further boost the ED/ES segmentation performance; however, it requires an input prompt and has not been optimized for cardiac segmentation.

Methods/Intervention

We developed a two-stage adaptation pipeline using the EchoNet-Dynamic training (N=7,465) and validation cohort sets (N=1,288). First, UltraSAM was fine-tuned for LV segmentation using rank-8 Low-Rank Adaptation (LoRA) applied to the attention layers of its vision-transformer backbone. Second, to eliminate manual prompting, we trained a lightweight 33,000-parameter multilayer perceptron using UltraSAM backbone features to estimate the LV centroid and generate an automated point prompt. We compared the original EchoNet-Dynamic DeepLabv3 model, LoRA-UltraSAM with an annotation-derived box prompt, and LoRA-UltraSAM with the learned point prompt on the EchoNet-Dynamic test set. https://github.com/rycw22/auto-ultrasam-lora/

Results/Outcome

Tested on the EchoNet-Dynamic test cohort (N=1,276), baseline DeepLabv3 model achieved ED/ES Dice scores of 0.927/0.903. LoRA-UltraSAM with an user annotation-derived box prompt achieved 0.955/0.953, corresponding to significant relative improvements of 3.0% at ED and 5.5% at ES and establishing an upper bound for prompted performance. Replacing the annotation-derived box with the fully automated learned point prompt achieved 0.926/0.920 Dice score for ED/ES, maintaining ED performance while improving ES Dice by 1.9% relative to DeepLabv3.

Conclusion

Parameter-efficient fine-tuning substantially improved UltraSAM when provided a high-quality prompt. Automated prompt generation eliminated human interaction while preserving baseline ED performance and improving ES segmentation. The performance gap between box- and point-prompted results suggests that prompt localization remains the primary opportunity for further improvement.

Statement of Impact

This approach converts a prompt-dependent ultrasound foundation model into a compact, fully automated LV segmentation pipeline that may reduce the need for manual clinical image segmentation and support scalable, reproducible ejection-fraction assessment and down-stream clinical information extraction.

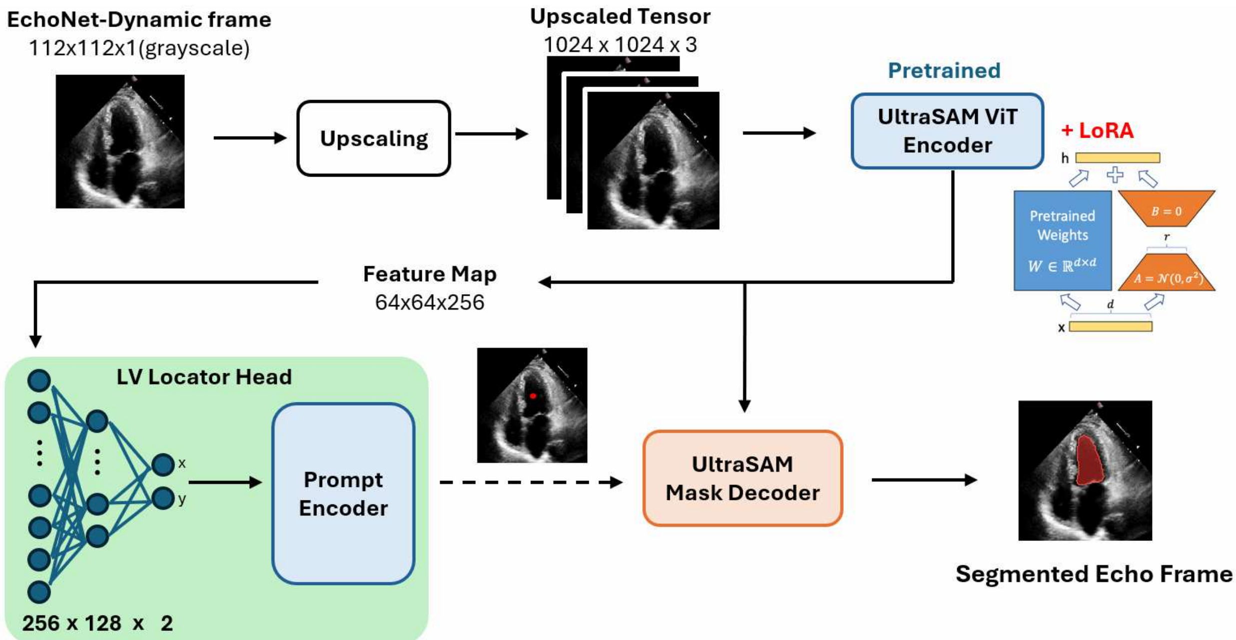

Fig. 1. Model Architecture. A grayscale echocardiogram frame is upscaled to a 1024x1024x3 tensor and encoded by the rank-LoRA encoded frozen UltraSAM ViT encoder. The encoder produces a 64x64x256 feature map that is fed to the 33k-parameter MLP that regresses the prompt LV centroid and the UltraSAM mask decoder which outputs the final segmented echocardiogram frame.

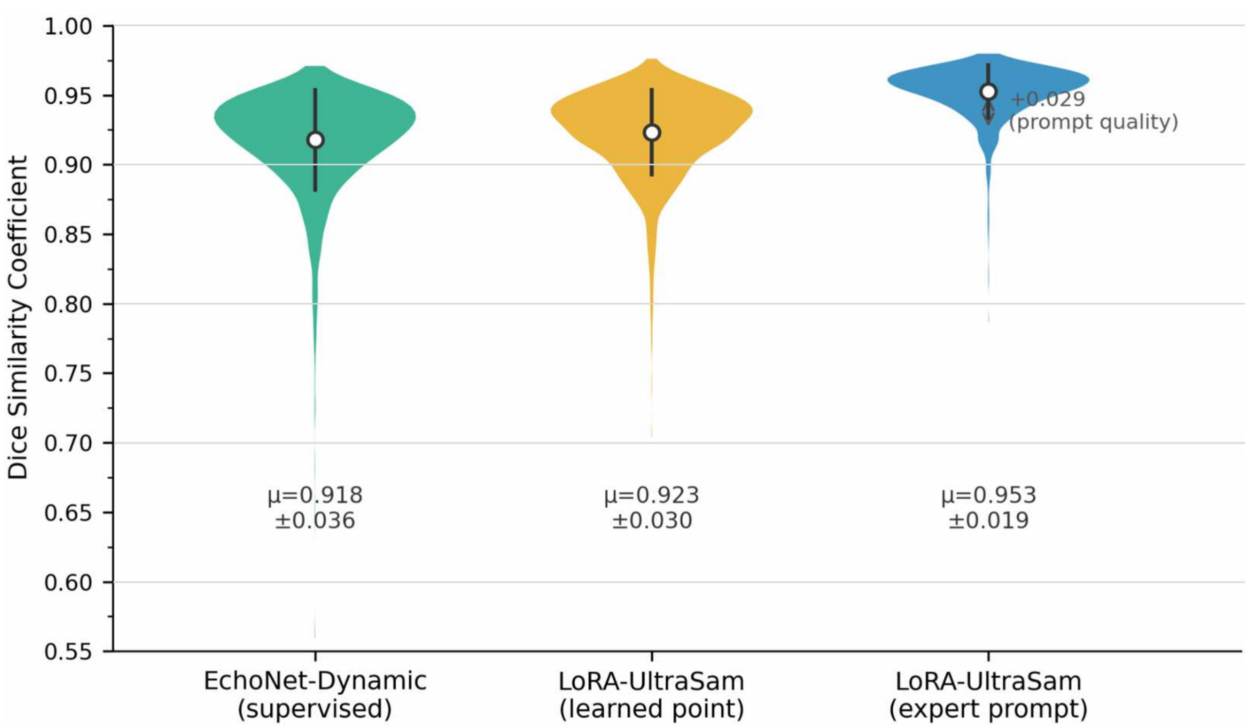

Fig. 2. Violin plots for the Left-Ventricular Segmentation Dice similarity scores tested on EchoNet-Dynamic. The plots represent ED/ES Dice scores averaged for the EchoNet-Dynamic's model (left), and the LoRA-UltraSAM model under two different types of prompt: the automated point prompting (middle), and the expert annotation (right). The .029 difference between the prompting choice indicates the prompt quality is the bottleneck.

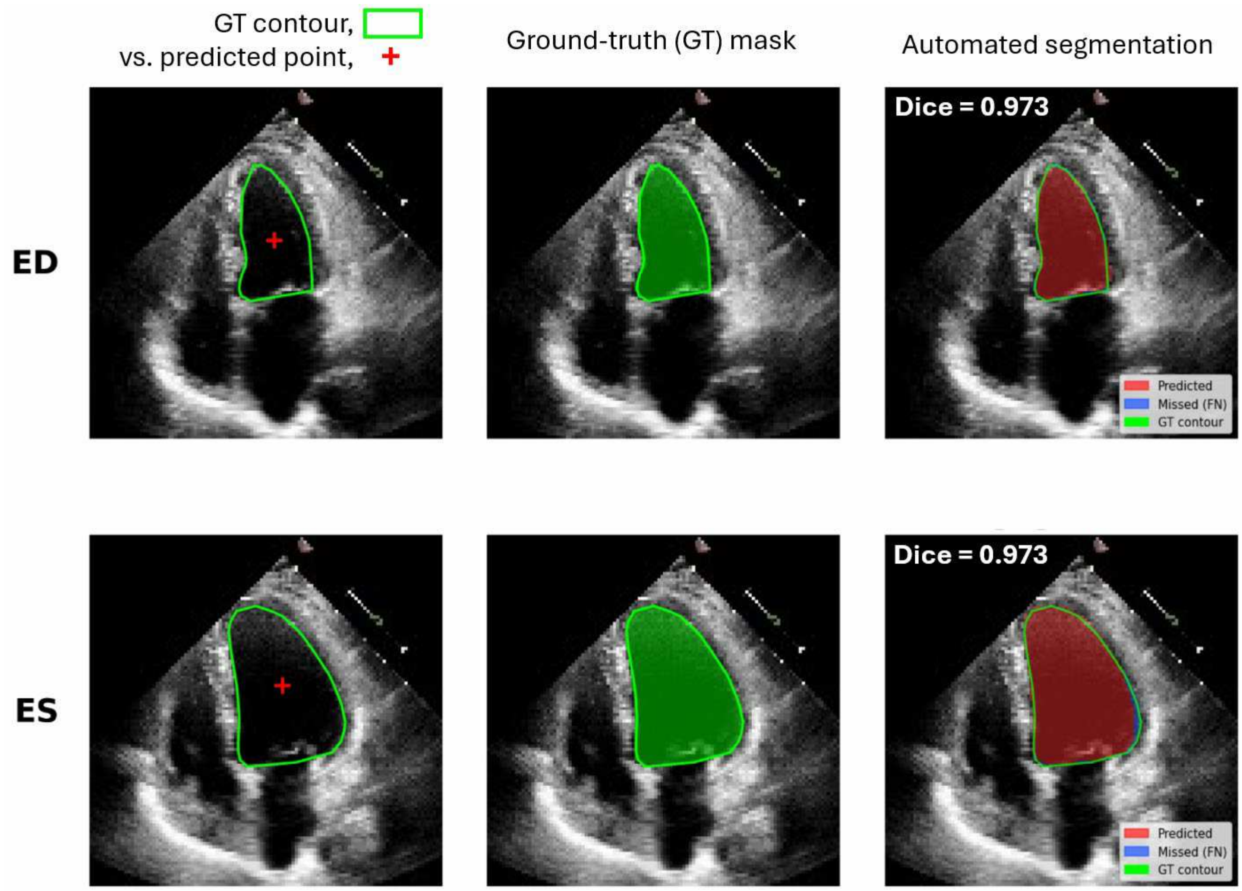

Fig. 3. Representative case of automatic LV segmentation using LoRA-UltraSam with learned automated prompting. The red and blue areas show the predicted area, the difference from the prediction and the ground truth, respectively. ED: End diastolic, ES: End systolic.

Keywords

Echocardiography; ultrasound segmentation; automated prompting; UltraSAM; LoRA fine tuning; ejection fraction