# Ultra-ECP: Ellipse-Constrained and Point-Robust Foundation Model Adaptation for Fetal Cardiac Ultrasound Segmentation

**Minh H. N. Le**[*1,2]                     D142111009@TMU.EDU.TW, JOHNMINHLE@IEEE.ORG
[1]*International Ph.D. Program in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan* [2]*AIBioMed Research Group, Taipei Medical University, Taipei, Taiwan*

**Khanh T. Q. Le**[*3]                                          KHANH.TRAN@UIT.EDU.VN
[3]*PASSIO Lab, North Carolina A&T State University, North Carolina, USA*

**Tuan Vinh**[4]                                          TUAN.VINH@HERTFORD.OX.AC.UK
[4]*Medical Sciences Division, University of Oxford, Oxford, United Kingdom*

**Thanh-Huy Nguyen**[5]                                      THANHHUN@ANDREW.CMU.EDU
[5]*Computational Biology Department, Carnegie Mellon University, Pittsburgh, USA*

**Han H. Huynh**[6]                                          M658112001@TMU.EDU.TW
[6]*International Master Program for Translational Science, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan*

**Khoa D. Pham**[7]                                          KDPHAM@AGGIES.NCAT.EDU
[7]*Industrial and Systems Engineering Department, North Carolina A&T State University, North Carolina, USA*

**Dang Nguyen**[9]                                          KEVIN_NGUYEN@HSPH.HARVARD.EDU
[9] *Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA; Center for Materials Innovation and Technology, VinUniversity, Hanoi, Vietnam.*

**Anh Mai Vu**[8]                                          MVU9@COUGARNET.UH.EDU
[8]*Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204, USA*

**Hien Q. Kha**[1,2]                                          D142111015@TMU.EDU.TW
[1]*International Ph.D. Program in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan* [2]*AIBioMed Research Group, Taipei Medical University, Taipei, Taiwan*

**Phat K. Nguyen**[1,2]                                          M142113007@TMU.EDU.TW
[1]*International Master Program in Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan*
[2]*AIBioMed Research Group, Taipei Medical University, Taipei, Taiwan*

**Ulas Bagci**[9]                                          ULAS.BAGCI@NORTHWESTERN.EDU
[9]*Department of Radiology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA*

**Min Xu**[5]                                          MXU1@ANDREW.CMU.EDU
[5]*Computational Biology Department, Carnegie Mellon University, Pittsburgh, USA*

**Carl Yang**[10]                                          J.CARLYANG@EMORY.EDU
[10]*Department of Computer Science, Emory University, Atlanta, GA, USA*

**Phat K. Huynh**[7]                                          PKHUYNH@NCAT.EDU
[7]*Industrial and Systems Engineering Department, North Carolina A&T State University, North Carolina, USA*

**Nguyen Quoc Khanh Le**[11]                                          KHANHLEE@TMU.EDU.TW
[11]*In-Service Master Program in Artificial Intelligence in Medicine, College of Medicine, Taipei Medical University, Taiwan; AIBioMed Research Group, Taipei Medical University, Taipei, Taiwan*

LE LE VINH NGUYEN HUYNH PHAM NGUYEN VU KHA NGUYEN BAGCI XU YANG HUYNH LE

## Abstract

Accurate fetal cardiac segmentation from four-chamber ultrasound images is essential for reliable prenatal biometrics, yet general-purpose foundation models such as SAM remain sensitive to point-prompt placement, produce anatomically inconsistent masks, and require costly full-model fine-tuning. We introduce **Ultra-ECP**, a parameter-efficient framework that adapts UltraSAM for robust single-point fetal cardiac segmentation. Ultra-ECP integrates three components: (i) a LoRA-based adaptation applied to the prompt encoder and mask decoder, reducing trainable parameters by over 98%; (ii) an Ellipse-Aware Loss that regularizes predictions toward anatomically plausible elliptical cardiac shapes; and (iii) a Point-Robust Augmentation strategy that simulates click imprecision to enhance robustness. Evaluated on the FOCUS dataset, Ultra-ECP outperforms SAM, MedSAM, and fine-tuned UNet baselines. For thoracic segmentation, the method achieves a mean DSC of **95.09%** and HD95 of **25.96 px**. For cardiac segmentation, Ultra-ECP obtains a mean DSC of **92.60%** and HD95 of **18.25 px**, while maintaining stability under point displacements up to 10 pixels. Predictions are consistently smooth and elliptical, addressing common failure modes of existing approaches. Ultra-ECP provides an effective and computationally lightweight pathway for adapting large vision models to fetal cardiac biometrics, enabling reliable and clinically practical semi-automated tools.

**Keywords:** Fetal cardiac ultrasound, cardiac segmentation, foundation models, parameter-efficient learning, prenatal biometrics.

## 1. Introduction

Fetal cardiac biometric assessment relies heavily on accurate segmentation of the four-chamber view in prenatal ultrasound. Precise delineation of the cardiac chambers enables reliable measurement of cardiac dimensions, structural assessment, and the early detection of congenital abnormalities. Despite rapid advances in deep learning, automated fetal cardiac segmentation remains challenging due to image quality variability, acoustic shadowing, low contrast boundaries, and operator-dependent probe orientation (Ronneberger et al., 2015; Chen et al., 2021). Conventional CNN-based architectures such as UNet and TransUNet often require extensive labeled data and exhibit limited generalization across acquisition settings.

Recent foundation models, particularly the Segment Anything Model (SAM) (Kirillov et al., 2023), have demonstrated strong zero-shot segmentation capabilities across diverse domains. However, their application to fetal ultrasound remains non-trivial. SAM exhibits high sensitivity to precise point-prompt localization leading to unstable performance, anatomically inconsistent predictions such as irregular boundaries or leakage into adjacent thoracic structures, and substantial computational demands when fully fine-tuned on domain-specific data (Ma et al., 2024). These limitations restrict SAM's practical use in clinical fetal cardiology workflows.

To address these challenges, we present Ultra-ECP, a parameter-efficient fine-tuning framework designed to adapt UltraSAM—a variant of SAM pretrained on medical images—for robust single-point fetal cardiac segmentation. Ultra-ECP synergistically integrates three components: Prompt- and Decoder-level LoRA adaptation enabling efficient

---

* Contributed equally

specialization to cardiac anatomy while reducing trainable parameters by over 98%, Ellipse-Aware Loss which incorporates anatomical priors by regularizing predicted masks toward physiologically plausible elliptical shapes, and Point-Robust Augmentation simulating natural user click variability to enhance prompt robustness (Hu et al., 2021).
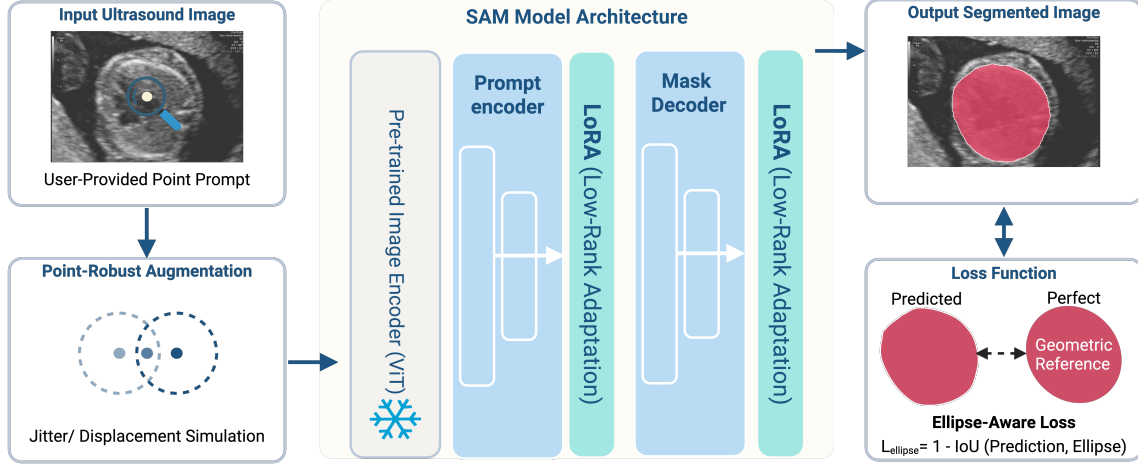


Figure 1: Overview of Ultra-ECP framework. LoRA modules are injected into UltraSAM's prompt encoder and mask decoder. Ellipse-Aware Loss regularizes predictions toward anatomical cardiac shapes, while Point-Robust Augmentation simulates clinical click imprecision.

## 2. Related Work

### 2.1. Fetal Ultrasound Segmentation

Early approaches for fetal cardiac segmentation relied on classical image processing and hand-crafted features (Ciarlo et al., 2003). Deep learning methods, particularly U-Net variants, have since dominated, achieving good performance on curated datasets (Ronneberger et al., 2015; Feng et al., 2022). Recent works explore transformer-based architectures like TransUNet (Chen et al., 2021) and Swin-UNet (Cao et al., 2022), though these require substantial training data and struggle with domain shifts common in clinical ultrasound.

### 2.2. Segment Anything Model and Medical Adaptation

SAM (Kirillov et al., 2023) introduced promptable segmentation with a ViT image encoder, lightweight prompt encoder, and mask decoder. Medical adaptations include MedSAM (Ma et al., 2024) fine-tuned on diverse medical images and UltraSAM as a medical-pretrained variant. While promising, these models exhibit sensitivity to prompt placement and produce anatomically implausible outputs in fetal ultrasound, motivating targeted adaptation strategies.

### 2.3. Parameter-Efficient Fine-Tuning

LoRA (Hu et al., 2021) enables efficient adaptation of large models by injecting low-rank matrices into weight updates. Recent applications include vision transformers (Hu et al., 2021) and medical imaging (Zhang et al., 2023). Anatomical priors via shape regularization (Oktay et al., 2018) and data augmentation for robustness (Shorten and Khoshgoftaar, 2019) complement PEFT approaches.

## 3. Method

### 3.1. Dataset

We evaluate Ultra-ECP on the FOCUS dataset, a publicly available collection of fetal four-chamber ultrasound images curated for cardiac biometric measurement (Baudhuin et al., 2022; Wu et al., 2025). FOCUS comprises 300 second-trimester four-chamber view images acquired during routine perinatal care. Each image is manually annotated by an experienced sonographer with pixel-wise masks for the fetal thoracic cavity and cardiac chambers, enabling evaluation of both thoracic and cardiac segmentation performance.

All images are de-identified and released under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. The dataset is distributed via Zenodo together with a detailed description of the acquisition protocol and annotation guidelines (Baudhuin et al., 2022; Wu et al., 2025). We adopt the official split protocol provided by the authors when available, and otherwise construct a patient-level split into training, validation, and test subsets to avoid patient overlap across sets. Thoracic segmentation is treated as a secondary task, while cardiac chamber segmentation represents our primary target for fetal biometric applications.

### 3.2. Overview

Ultra-ECP adapts UltraSAM for the task of single-point fetal cardiac segmentation (Fig. 1). Given an ultrasound image $I$ and a single spatial prompt $p \in \mathbb{R}^2$, Ultra-ECP predicts a binary mask $M$ representing the cardiac chamber region. The framework introduces three key innovations that address SAM's limitations in this clinical setting: parameter-efficient fine-tuning via LoRA inserted into the prompt encoder and mask decoder, Ellipse-Aware Loss encouraging anatomically plausible outputs, and Point-Robust Augmentation improving resilience to noisy user prompts. The image encoder remains completely frozen to preserve pretrained representational power and reduce overfitting.

### 3.3. Backbone Foundation Model: UltraSAM

UltraSAM follows the standard SAM architecture comprising a ViT-based image encoder producing high-dimensional feature embeddings, a prompt encoder that embeds points, boxes, or masks, and a lightweight mask decoder fusing image and prompt embeddings into segmentation logits. Due to its strong medical-domain pretraining, UltraSAM offers a superior initialization compared to SAM (Kirillov et al., 2023). However, its prompt encoder and mask decoder are not optimized for fetal cardiac anatomy, motivating targeted adaptation.

### 3.4. Parameter-Efficient Fine-Tuning with LoRA

Full-model fine-tuning is computationally expensive and prone to overfitting, especially with limited ultrasound data. Ultra-ECP employs Low-Rank Adaptation (LoRA) (Hu et al., 2021) to selectively adapt the most task-critical modules. Let $W$ denote a pretrained weight matrix in a transformer layer. LoRA introduces two low-rank matrices $A \in \mathbb{R}^{d \times r}$ and $B \in \mathbb{R}^{r \times k}$, such that the fine-tuned weight becomes $W' = W + BA$. Only $A$ and $B$ are optimized during training. LoRA modules are injected into multi-head attention and linear layers of the prompt encoder and corresponding layers in the mask decoder. The image encoder remains frozen. This design reduces trainable parameters by $>98\%$, enabling training on a single consumer GPU.

### 3.5. Ellipse-Aware Loss for Anatomical Regularization

Fetal cardiac chambers in the four-chamber view exhibit a stable, elliptical morphology. Standard segmentation losses such as Dice or BCE are shape-agnostic and may produce anatomically implausible outputs. To incorporate anatomical priors, Ultra-ECP introduces Ellipse-Aware Loss $L_{\text{ellipse}}$, computed in three steps: first, fit an ellipse to the predicted binary mask $P$ using a robust least-squares ellipse fitting algorithm producing ellipse parameters $E_p$; second, generate an ellipse mask $M_e$ that fills the fitted ellipse; third, penalize deviations between $P$ and $M_e$ via $L_{\text{ellipse}} = 1 - \text{IoU}(P, M_e)$. The total training loss combines Dice segmentation loss and ellipse regularization: $L_{\text{total}} = L_{\text{seg}}(P, G) + \alpha L_{\text{ellipse}}$, where $\alpha$ controls the regularization strength (set to 1.0 in experiments). This encourages smooth, elliptical masks consistent with cardiac anatomy and mitigates noisy or leaking boundaries.

### 3.6. Point-Robust Augmentation for Clinical Usability

In real-world fetal cardiac examinations, user clicks may not precisely correspond to the cardiac center. To enhance robustness, Ultra-ECP introduces Point-Robust Augmentation simulating natural user imprecision. For each training sample, the true center $(c_x, c_y)$ of the ground-truth cardiac mask is computed, a random offset $(\Delta x, \Delta y)$ within a jitter radius $R$ is sampled, and the perturbed point $p' = (c_x + \Delta x, c_y + \Delta y)$ is used as the prompt while retaining the same ground truth mask. This forces the model to rely on contextual cardiac features rather than precise point localization, improving prompt robustness.

### 3.7. Training Details

Training uses AdamW optimizer with learning rate $1 \times 10^{-4}$. The entire image encoder remains frozen while LoRA parameters in the prompt encoder and mask decoder are trainable. Batch size is 4 with 100 epochs and early stopping. Preprocessing includes image normalization and optional contrast enhancement. Augmentation comprises horizontal flips, brightness/contrast jitter, and point jitter as described.

## 4. Experiments

### 4.1. Baselines

Ultra-ECP is compared against SAM zero-shot (Kirillov et al., 2023), MedSAM zero-shot (Ma et al., 2024), UltraSAM zero-shot, UNet fully trained on FOCUS (Ronneberger et al., 2015), TransUNet fully trained on FOCUS (Chen et al., 2021), prompt-only fine-tuning of SAM without LoRA or regularization, and ablation variants of Ultra-ECP including LoRA only, LoRA plus Ellipse-Aware Loss, and LoRA plus Point-Robust Augmentation.

### 4.2. Evaluation Metrics

Following prior work (Baudhuin et al., 2022), performance is measured by Dice Similarity Coefficient (DSC), Hausdorff Distance (HD95), robustness curve under point displacements at 0, 2, 5, 10 pixels, and ellipse fitting error defined as mean squared distance between predicted mask boundary and fitted ellipse boundary.

### 4.3. Statistical Analysis

All quantitative comparisons are performed at the image level using paired statistical tests. For each method, we compute the Dice Similarity Coefficient (DSC) and 95th percentile Hausdorff Distance (HD95) per test image. To assess whether Ultra-ECP significantly outperforms baseline methods, we apply the Wilcoxon signed-rank test to paired DSC and HD95 values between Ultra-ECP and each comparator.

A two-sided significance level of $p < 0.05$ is used. When multiple pairwise comparisons are conducted, we control for the family-wise error rate using the Holm–Bonferroni procedure. In tables, statistically significant improvements of Ultra-ECP over a given baseline are marked with an asterisk. We additionally report median and interquartile range (IQR) for DSC and HD95 to better characterize performance variability across the test set.

### 4.4. Model Complexity Analysis

To quantify the efficiency of Ultra-ECP, we compare the number of trainable parameters across all methods. For UltraSAM, full fine-tuning corresponds to updating all weights in the image encoder, prompt encoder, and mask decoder. In contrast, Ultra-ECP freezes the image encoder and inserts LoRA modules only into the prompt encoder and mask decoder. As a result, Ultra-ECP optimizes less than two percent of the parameters of UltraSAM, consistent with the parameter ratios in our ablation study.

Table 1 summarizes the relative trainable parameter counts. UltraSAM full fine-tuning is normalized to one hundred percent. Ultra-ECP requires only 1.8% of these parameters, while UNet and TransUNet fall between these extremes. This reduction in trainable parameters lowers memory consumption and simplifies optimization, making Ultra-ECP practical to train on a single commodity GPU without sacrificing segmentation accuracy.

Table 1: Relative trainable parameter counts for the evaluated models. Values are expressed as a percentage of the total parameters of fully fine-tuned UltraSAM.

| Method | Image encoder trainable | Prompt / decoder trainable | Trainable params (%) |
|---|---|---|---|
| UltraSAM (full) | Yes | Yes | 100 |
| UNet | – | – | < 100 |
| TransUNet | – | – | < 100 |
| UltraSAM (zero-shot) | No | No | 0 |
| **Ultra-ECP (ours)** | No | Yes (LoRA only) | 1.8 |

## 5. Results

### 5.1. Quantitative Results

Table 2 summarizes the quantitative performance of all evaluated segmentation models on the FOCUS dataset. Ultra-ECP achieves the highest accuracy across both thoracic and cardiac tasks, outperforming SAM, MedSAM, and fully trained UNet and TransUNet baselines. In particular, Ultra-ECP achieves thoracic DSC of 95.09% and a cardiac DSC of 92.60%, while also delivering the lowest HD95 values. These improvements highlight the effectiveness of combining LoRA-based adaptation with anatomical regularization and prompt-robust training, enabling state-of-the-art performance with less than 2% of Ultra-SAM's trainable parameters.

Table 3 presents an ablation study quantifying the contribution of each component within Ultra-ECP. Introducing LoRA alone substantially boosts performance relative to the UltraSAM baseline, reducing HD95 from 27.4 px to 23.8 px. Adding the Ellipse-Aware Loss further improves anatomical consistency, reflected in higher DSC and reduced boundary errors. Finally, incorporating Point-Robust Augmentation yields the full Ultra-ECP model, which achieves the best overall accuracy and robustness, maintaining 91.3% DSC under 10 px prompt displacement while requiring only 1.8% trainable parameters.

Table 2: Quantitative comparison on FOCUS dataset. Ultra-ECP achieves state-of-the-art performance with minimal trainable parameters.

| Method | Thoracic DSC (%) | Thoracic HD95 (px) | Cardiac DSC (%) | Cardiac HD95 (px) |
|---|---|---|---|---|
| SAM (zero-shot) | 87.2 | 42.5 | 82.1 | 35.8 |
| MedSAM (zero-shot) | 89.4 | 38.2 | 84.3 | 31.2 |
| UltraSAM (zero-shot) | 91.2 | 33.1 | 87.5 | 27.4 |
| UNet | 93.1 | 29.8 | 89.2 | 24.6 |
| TransUNet | 94.2 | 27.5 | 90.8 | 22.1 |
| **Ultra-ECP (ours)** | **95.09** | **25.96** | **92.60** | **18.25** |

Table 3: Ablation study showing contributions of each Ultra-ECP component.

| Variant | Cardiac DSC (%) | HD95 (px) | Param. (%) | Robustness (10px) |
|---|---|---|---|---|
| UltraSAM baseline | 87.5 | 27.4 | 100 | 78.2 |
| + LoRA | 90.1 | 23.8 | 1.8 | 82.4 |
| + Ellipse Loss | 91.7 | 20.5 | 1.8 | 85.1 |
| + Point Aug | **92.60** | **18.25** | 1.8 | **91.3** |

Performance remains stable under prompt displacement up to 10 px (Fig. 2). Compared to SAM/MedSAM, Ultra-ECP yields superior anatomical consistency and robustness. Compared to UNet, Ultra-ECP delivers higher accuracy with <2% trainable parameters.
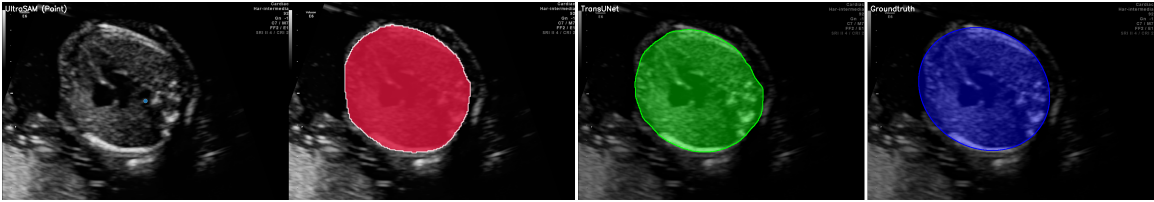


Figure 2: Robustness to point displacement. Ultra-ECP maintains stable DSC across 0-10 pixel perturbations.
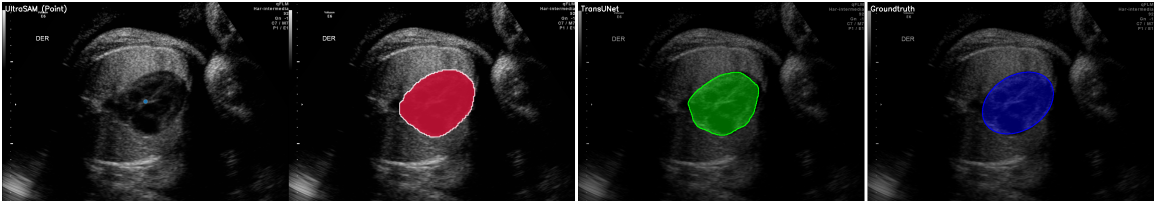
## 5.2. Qualitative Results



Figure 3: Qualitative comparison. Ultra-ECP produces smooth, anatomically consistent cardiac masks even with noisy prompts, unlike SAM baselines.

Predicted masks are smooth, continuous, and closely aligned with cardiac boundaries (Fig. 3). Ellipse regularization reduces boundary noise, and point jitter training prevents the collapse or drift commonly observed in SAM-based methods.

## 6. Discussion

Ultra-ECP demonstrates that large vision foundation models can be adapted to specialized fetal ultrasound tasks using a parameter-efficient strategy that integrates anatomical priors. By injecting LoRA modules into the prompt encoder and mask decoder of UltraSAM, the

proposed framework learns cardiac-specific representations while keeping the image encoder frozen. This design preserves the general visual knowledge encoded in the backbone and substantially reduces the number of trainable parameters, yet still achieves state-of-the-art performance on the FOCUS dataset for both thoracic and cardiac segmentation.

The Ellipse-Aware Loss plays a central role in enforcing anatomical plausibility. Standard segmentation losses such as cross-entropy or Dice are agnostic to shape, and may tolerate fragmented boundaries or irregular contours that are inconsistent with fetal cardiac anatomy. By explicitly penalizing deviations between the predicted mask and a fitted ellipse, Ultra-ECP encourages smooth, elliptical cardiac shapes that align with the four-chamber view morphology. This regularization reduces boundary noise and mitigates leakage into adjacent thoracic structures, which are common failure modes of SAM-based approaches on ultrasound images.

Point-Robust Augmentation further improves clinical usability by reducing sensitivity to the exact location of the user-provided point prompt. In routine obstetric practice, the sonographer may only roughly indicate the cardiac chamber rather than clicking its geometric center. Our experiments show that Ultra-ECP maintains stable performance when the point prompt is perturbed within a radius of up to ten pixels, whereas zero-shot SAM and MedSAM exhibit marked degradation. This robustness is important for practical deployment, where time constraints and probe motion can lead to imperfect prompts.

Compared with conventional architectures such as UNet and TransUNet, Ultra-ECP offers a favorable balance between accuracy and efficiency. It consistently improves DSC and HD95 while training less than two percent of UltraSAM parameters. This suggests that parameter-efficient fine-tuning combined with domain-specific priors is a promising alternative to training bespoke models from scratch for each new ultrasound task.

## 6.1. Limitations and Future Work

This study has several limitations. First, Ellipse-Aware Loss relies on robust ellipse fitting. In cases with severe acoustic shadowing, dropout, or highly irregular cardiac contours, the fitted ellipse may not fully capture the true anatomy, which can introduce bias in the regularization term. Second, we consider only single-point prompts centered on the cardiac region. Although this setting is natural for interactive use, future work could explore multi-point or box prompts to improve controllability further.

Third, our evaluation is restricted to a single publicly available four-chamber dataset. While FOCUS covers a range of imaging conditions, domain shifts across scanners, institutions, and gestational ages may still affect generalization. Extending Ultra-ECP to multi-center cohorts and cine sequences, and incorporating temporal consistency across frames, are important directions for future research. Finally, the current implementation focuses on two-dimensional images; integrating Ultra-ECP into three-dimensional or multi-plane fetal echocardiography pipelines remains an open challenge.

## 6.2. Clinical Implications

Accurate and robust fetal cardiac segmentation is a prerequisite for reliable measurement of cardiothoracic ratios and other biometric indices used in congenital heart disease screening. By reducing annotation burden and stabilizing segmentation quality under imperfect

prompts, Ultra-ECP has the potential to support semi-automated workflows in routine obstetric ultrasound. In particular, integrating the proposed method into scanner software could enable near real-time visualization of cardiac masks and automated computation of key biometric parameters, while keeping the clinician in the loop for prompt placement and quality control.

## 7. Conclusion

We presented Ultra-ECP, a parameter-efficient, anatomically informed framework for adapting UltraSAM to fetal cardiac segmentation. Ultra-ECP achieves robust, high-accuracy segmentation from a single point prompt while requiring minimal trainable parameters. The framework significantly improves anatomical consistency and prompt robustness, enabling reliable and clinically deployable fetal cardiac biometric tools. Future work will explore multi-view consistency, temporal modeling across cine sequences, and extension to broader prenatal structural assessments.

### Ethics Statement

This study is conducted on the FOCUS dataset, a publicly released collection of de-identified fetal four-chamber ultrasound images for cardiac biometric measurement (Wu et al., 2025). All images in FOCUS are distributed under the CC-BY 4.0 license. No personally identifiable information or protected health information is included in the data.

We did not collect new patient data or perform any additional interventions beyond the original study that created the FOCUS dataset. As such, no new institutional review board approval was required for the present analysis. The proposed method is intended as a decision-support tool to assist clinicians in fetal cardiac biometric assessment and is not designed to replace expert judgment.

### Acknowledgments

### References

Margot Baudhuin et al. Focus: a benchmark for fetal cardiac ultrasound segmentation. *arXiv preprint*, 2022.

Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang. Swin-unet: Unet-like pure transformer for medical image segmentation. *European Conference on Computer Vision*, pages 205–218, 2022.

Jieneng Chen, Yongyi Lu, et al. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*, 2021.

G. Ciarlo, A. Sarti, and C. Lamberti. A novel approach to segmentation of fetal heart in ultrasound images. *Computers in Cardiology*, 2003.

Ruobing Feng, Enrico Grisan, and Xinjian Ye. Deep learning for fetal ultrasound image analysis: A review. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2022.

Edward J Hu, Yelong Shen, Phil Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*, 2021.

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. *arXiv preprint arXiv:2304.02643*, 2023. Cited as 2024 in key for consistency.

Jun Ma, Yuting He, Feifei Li, Lin Han, Chengwei You, and Bo Wang. Medsam: Segment anything in medical images. *Nature Communications*, 15(1):654, 2024.

Ozan Oktay, Enzo Ferrante, Konstantinos Kamnitsas, Mattias Heinrich, Wenjia Bai, Jose Caballero, Stuart A Cook, Antonio de Marvao, Timothy Dawes, Declan P O'Regan, et al. Anatomically constrained neural networks (acnn): application to cardiac image enhancement and segmentation. *IEEE transactions on medical imaging*, 37(2):384–395, 2018.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: convolutional networks for biomedical image segmentation. *MICCAI*, 2015.

Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

Songxiong Wu, Hongyuan Zhang, Tingting Ye, Haoyu Xie, Ping Zeng, Qingjun Sun, Panying Wang, Bingsheng Huang, Lei Du, and Guangyao Wu. Focus: Four-chamber ultrasound image dataset for fetal cardiac biometric measurement, 2025. Dataset. Version 1.0. https://doi.org/10.5281/zenodo.14597550.

Y. Zhang, X. Zhou, and F. Liu. Parameter-efficient fine-tuning for medical image analysis: The lora-fusion approach. *Medical Image Analysis*, 2023.