FETAL-BCP: ADDRESSING EMPIRICAL DISTRIBUTION GAP IN SEMI-SUPERVISED FETAL ULTRASOUND SEGMENTATION

Ha-Hieu Pham[1,†] *Le Tran Quoc Khanh*[2,†] *Hoang-Thien Nguyen*[3,†] *Nguyen Lan Vi Vu*[4,†]
Quang-Vinh Dinh[†] *Thanh-Huy Nguyen*[5,6,†] *Xingjian Li*[6] *Min Xu*[6⋆]

[†] AI VIETNAM Lab, Vietnam
[1] University of Science, VNU-HCM, Vietnam
[2] University of Information Technology, VNU-HCM, Vietnam
[3] Posts and Telecommunications Institute of Technology, Ho Chi Minh City, Vietnam
[4] Ho Chi Minh University of Technology, Vietnam
[5] Université Bourgogne Europe, France
[6] Carnegie Mellon University, United States

A. METHODOLOGY

A.1. Preprocessing & Augmentation Details

Following [1], during data processing, we resize all images to one fixed size to ensure identical input dimensions. We also adjust intensity values to a set range to boost contrast and learning stability.

Based on [2, 3], to add variety and avoid overfitting, we use extra image enrichment steps. We flip images randomly on horizontal and vertical axes. We also rotate them up to ± 30° to mimic different scanning angles. To account for tissue deformations, elastic transformation is applied with parameters $\alpha = 10$ and $\sigma = 2$. Additionally, to replicate common ultrasound artifacts, a combination of Gaussian blur (kernel size: 3–5) and median blur (limit: 3) is introduced, along with Gaussian noise (variance range: 5.0–20.0) to simulate variations in signal quality.

We enhance contrast by using either random brightness and contrast tweaks (limit: 0.1) or CLAHE with a clip limit of 2.0 and an 8×8 tile grid [4]. These modifications apply randomly to each sample using the Albumentations library. They offer varied yet realistic changes that improve model robustness while preserving anatomical accuracy [2].

These preprocessing and augmentation methods play a crucial role in building deep learning models that automatically understand fetal ultrasound images, especially given the diversity in imaging conditions and the precision demanded in clinical diagnostics [5].

⋆Correspondence: mxu1@cs.cmu.edu (Min Xu)

A.2. Mean Teacher and Training Strategy

The Fetal-BCP framework employs a Mean Teacher model, where the Teacher network F_t and the Student network F_s are parameterized by Θ_t and Θ_s. The training process follows a three-step strategy. Initially, the model undergoes pretraining using only labeled data to establish a solid representation. Following this, the pretrained model functions as the Teacher network, generating pseudo-labels for the unlabeled data. During each iteration, the Student network is optimized via AdamW optimizer, while the Teacher network is updated using the exponential moving average (EMA) of the Student network's :

$$\Theta_t^{k+1} = \lambda \Theta_t^k + (1 - \lambda)\Theta_s^k \tag{1}$$

where λ represents the smoothing coefficient that governs the influence of previous Teacher parameters.

B. EXPERIMENTAL SETTINGS

B.1. Dataset

We use the dataset from the Fetal Ultrasound Grand Challenge: Semi-Supervised Cervical Segmentation [6], consisting of 500 transvaginal ultrasound images of the uterine cervix. Among these, 50 images are labeled and 450 images are unlabeled. The labeled set is split into 40 images for training and 10 images for validation, while the test set consists of 90 images from the challenge's Validation Phase [6].

All images were acquired using a curved transducer with a 2–10 MHz vaginal probe, commonly used for second-trimester cervical ultrasound screening. Scans were performed with patients in a low Fowler's position after bladder emptying. Sonographers were instructed to minimize post-processing and avoid artifacts such as smoothing, noise reduction, pointers, or calipers whenever possible. Image acquisition parameters, including gain, frequency, and time gain compensation, were adjusted at the operator's discretion.

Each image captures cervical anatomical structures, including the anterior and posterior lips of the cervix, providing a clinically relevant benchmark for semi-supervised cervical segmentation.

B.2. Implementation Details

Our experiments are conducted using PyTorch on an NVIDIA RTX A5000 GPU with 24GB VRAM. The model backbone is DeepLabV3+ with a ResNet101 encoder, which is well-suited for semantic segmentation tasks [7].

For optimization, we use the AdamW optimizer, with an initial learning rate of 1×10^{-4}, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1 \times 10^{-8}$, and a weight decay of 0.05 [8]. The learning rate is scheduled using a CosineAnnealingLR scheduler, ensuring smooth convergence throughout training.

To improve generalization and prevent overfitting, we apply data augmentation techniques following Section **??**. Additionally, we apply consistency training to the unlabeled set, where the consistency weight α is gradually increased following a ramp-up schedule, reaching a maximum value of 0.5 across all experiments.

C. REFERENCES

[1] R. Yousefpour Shahrivar, F. Karami, and E. Karami, "Enhancing fetal anomaly detection in ultrasonography images: A review of machine learning-based approaches," *Biomimetics (Basel, Switzerland)*, vol. 8, no. 7, pp. 519, 2023.

[2] E. J. Snider, S. I. Hernandez-Torres, and R. Hennessey, "Using ultrasound image augmentation and ensemble predictions to prevent machine-learning model overfitting," *Diagnostics (Basel, Switzerland)*, vol. 13, no. 3, pp. 417, 2023.

[3] C. Athalye and R. Arnaout, "Domain-guided data augmentation for deep learning on medical imaging," *PLOS ONE*, vol. 18, no. 3, pp. e0282532, 2023.

[4] H. R. Torres, P. Morais, B. Oliveira, C. Birdir, M. Rüdiger, J. C. Fonseca, and J. L. Vilaça, "A review of image processing methods for fetal head and brain analysis in ultrasound images," *Computer Methods and Programs in Biomedicine*, vol. 215, pp. 106629, 2022.

[5] C. Sendra-Balcells, V. M. Campello, J. Torrents-Barrena, Y. A. Ahmed, M. Elattar, B. Ohene-Botwe, P. Nyangulu, W. Stones, M. Ammar, L. N. Benamer, H. N. Kisembo, S. G. Sereke, S. Z. Wanyonyi, M. Temmerman, E. Gratacós, E. Bonet, E. Eixarch, K. Mikolaj, M. G. Tolsgaard, and K. Lekadir, "Generalisability of fetal ultrasound deep learning models to low-resource imaging settings in five african countries," *Scientific Reports*, vol. 13, no. 1, pp. 2728, 2023.

[6] Jieyun Bai, Ziduo Yang, Kamrul Hasan, Jie Gan, Zhuonan Liang, Weidong Cai, Tao Tan, Jing Ye, Mohammad Yaqub, Dong Ni, Saad Slimani, Benard Ohene-Botwe, Campello Manuel, and Karim Lekadir, "Fetal ultrasound grand challenge: Semi-supervised cervical segmentation (fugc25)," 12 2024.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[8] Ilya Loshchilov and Frank Hutter, "Decoupled weight decay regularization," in *International Conference on Learning Representations*, 2019.