Dual-Domain Meta Conditioning Network for Fetal Head and Pubic Symphysis Segmentation in Ultrasound Images Analysis

Xinglong Hu[1,†], Lan Dong[2,†], Lu Zhang[3,†], Zhaoyi Ye[1], Du Wang[1], Liye Mei[5,6,7,*], and Cheng Lei[1,4,8,*]

[1] School of Integrated Circuits, Wuhan University, Wuhan, China
liyemei@whu.edu.cn, leicheng@whu.edu.cn
[2] Department of Gynecology, Renmin Hospital of Wuhan University, Wuhan, China
[3] Department of Obstetrics, Renmin Hospital of Wuhan University, Wuhan, China
[4] Department of Hematology, Zhongnan Hospital of Wuhan University, Wuhan, China
[5] The Institute of Technological Sciences, Wuhan University, Wuhan, China
[6] School of Computer Science and Artificial Intelligence, Hubei University of Technology, Wuhan, China
[7] Xinguang Zhiying (Wuhan) Technology Co., Ltd., Wuhan, China
[8] Shenzhen Institute of Wuhan University, Shenzhen, China

Abstract. Accurate segmentation of the fetal head (FH) and pubic symphysis (PS) in intrapartum ultrasound is crucial for Angle of Progression (AoP) estimation. However, speckle noise, weak boundary contrast, and anatomical variability hinder segmentation. We propose DMC-Net, a dual-domain meta-conditioning framework that integrates structural and spectral priors in a stage-specific manner. The encoder uses a Structural Domain Meta Conditioning Module to inject anatomical cues, reducing noise and enhancing boundary coherence. A Frequency-aware Cross-scale Attention Module in deeper layers models spatial-frequency interactions, suppressing noise while maintaining structural continuity. In the decoder, a second DMCM stabilizes reconstruction with refined frequency information. Finally, a Progressive Spectral Reconstruction Decoder restores Fourier domain components for precise boundary recovery. Experiments on the PSFHS2023 dataset show DMC-Net outperforms state-of-the-art methods in segmentation accuracy and AoP estimation, demonstrating its effectiveness in challenging ultrasound scenarios. Our code and data are available at https://github.com/XingLongH/DMC_Net.

Keywords: Intrapartum ultrasound · Deep Learning · Fetal biometry · Meta information

1 Introduction

Intrapartum ultrasound (US) has become a key tool in obstetrics due to its non-invasive, reproducible, and real-time imaging capabilities [5]. With the global cesarean section (CS) rate exceeding 20% and continuing to rise, there is an increasing need for objective tools to aid labor management and reduce unnecessary operative deliveries [2]. Conventional assessment of fetal head (FH) station and position by vaginal examination is subjective and associated with substantial inter-observer variability, with misclassification rates reported above 30% [25]. Accordingly, the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG) advocates intrapartum US for standardized and quantitative evaluation during labor [9].

Among ultrasound-derived parameters, the angle of progression (AoP) is a robust predictor of delivery outcome [19]. Larger AoP values are strongly associated with successful vaginal birth, and thresholds around 120° are frequently considered clinically meaningful [8,1]. However, manual AoP measurement is time-consuming and operator-dependent, with notable intra- and inter-observer variability [23]. Accurate AoP estimation requires precise delineation of the pubic symphysis (PS) axis and the lowest point of the FH contour, making reliable segmentation of PS and FH structures essential. This task is challenging in intrapartum US due to speckle noise, acoustic shadowing, low soft-tissue contrast, anatomical variability, dynamic fetal head rotation, and partial occlusion [7,20].

Deep learning has substantially advanced medical image segmentation. Convolutional neural network (CNN)-based models, particularly U-Net and its variants [15,14], serve as foundational frameworks in biomedical image analysis. Extensions incorporating multi-scale fusion and attention mechanisms have improved performance in obstetric US tasks [21]. Nevertheless, CNNs emphasize local receptive fields, limiting their ability to capture long-range anatomical dependencies [16,27]. Transformer-based models such as Swin Transformer [13] and hybrid architectures like Swin UNETR [6] enable global context modeling and show promising results in medical segmentation. More recently, structured state space models [22] and foundation frameworks such as the Segment Anything Model (SAM) [10] have demonstrated strong long-range modeling capacity. However, their performance in intrapartum US remains constrained by domain-specific noise and limited annotated data [17].

To address speckle noise and ambiguous boundaries in fetal ultrasound, we propose the Dual-Domain Meta Conditioning Network (DMC-Net). Unlike spatial-only models, DMC-Net integrates structural and frequency priors via two independent Domain Meta Conditioning Modules (DMCM). In the encoder, the first DMCM injects low-level structural cues as meta-information into deeper layers to mitigate noise-induced bias and ensure anatomical consistency. Subsequently, a Frequency-aware Cross-scale Attention module (FCAM) models spatial-frequency interactions, suppressing artifacts while generating refined frequency meta-information. In the decoding phase, this frequency prior is injected into a second, independent DMCM to stabilize segmentation profiles. Finally, the Progressive Spectral Reconstruction Decoder (PSRD) restores fine-grained details through spectral magnitude and phase refinement. This dual-stage conditioning enables robust boundary recovery in challenging ultrasound scenarios. The main contributions of this work are summarized as follows:

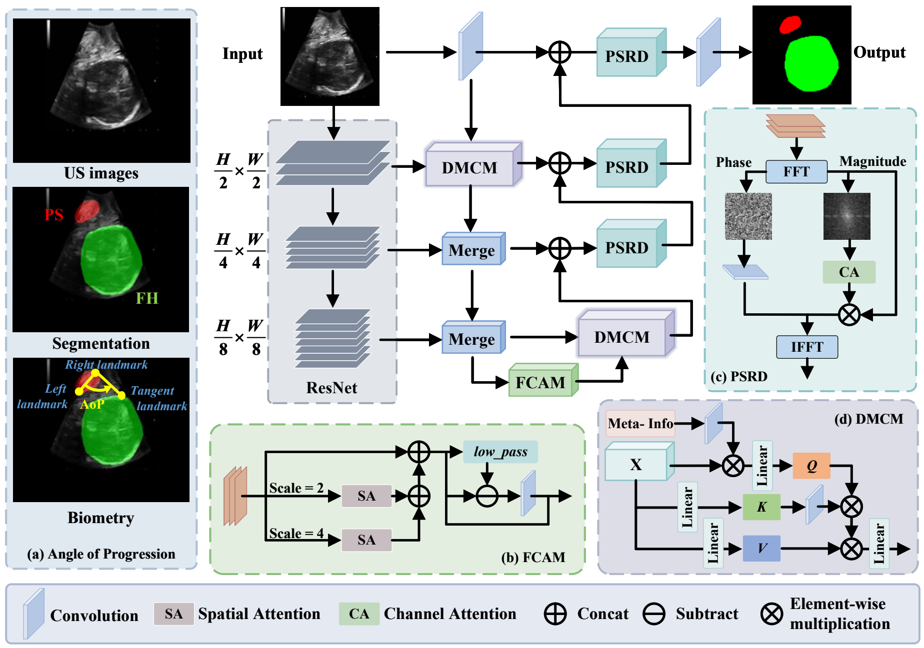

Fig. 1. Framework of DMC-Net. (a) Schematic diagram of Angle of Progression; (b) Frequency-aware Cross-scale Attention Module; (c) Progressive Spectral Reconstruction Decoder; (d) Domain Meta Conditioning Module

- We propose DMC-Net, a dual-domain meta-conditioning framework that explicitly models structural and frequency priors for robust fetal ultrasound segmentation.
- We design a stage-specific Domain Meta Conditioning Module (DMCM) that enhances feature representations by injecting structural and spectral meta-information through query-guided modulation.
- We introduce a Frequency-aware Cross-scale Attention Module and a Progressive Spectral Reconstruction Decoder to suppress ultrasound artifacts and recover fine-grained boundary details in the Fourier domain.

2 Proposed Method

2.1 Overview

Given a fetal ultrasound image $I \in \mathbb{R}^{3\times H\times W}$, the goal of DMC-Net is to segment both the FH and PS regions for accurate Angle of Progression (AoP) estimation. The architecture is based on a multi-scale encoder-decoder framework, where the encoder generates hierarchical feature maps:

$$\{F_1, F_2, F_3\} = Res(I), \tag{1}$$

where $Res(\cdot)$ represents the ResNet and $F_l \in \mathbb{R}^{C_i \times H_l \times W_l}$ denotes the feature map with C_i channels at scale l. For robustness under noise and structural ambiguity, dual-domain meta conditioning is applied. At the shallow level, DMCM maintains anatomical boundaries by structural conditioning:

$$\tilde{F}_1 = \mathcal{M}(F_1, \phi_0(I)), \tag{2}$$

where $\phi_0(\cdot)$ is the initial convolution block, $\mathcal{M}(\cdot,\cdot)$ is DMCM, and $\tilde{F}_1$ is the structurally conditioned feature. At the deepest level, frequency-aware refinement is performed:

$$P_3 = \mathcal{F}_{freq}(F_3), \quad D_3 = \mathcal{M}(F_3, P_3), \tag{3}$$

where $\mathcal{F}_{freq}(\cdot)$ denotes FCAM, P_3 is the frequency-aware prior, and D_3 is the initial decoder feature. During decoding, PSRD restores frequency details for high-resolution reconstruction:

$$D_{l-1} = U\left(\mathcal{R}(D_l) + F_{l-1}\right), \quad l = 3, 2, \tag{4}$$

where D_l denotes the decoder feature at scale l, $U(\cdot)$ denotes upsampling, and $\mathcal{R}(\cdot)$ denotes PSRD.

2.2 Domain Meta Conditioning Module

In DMCM, meta-information denotes task-specific prior features extracted from complementary domains rather than external metadata. Shallow structural features provide encoder-side structural priors, while FCAM generates decoder-side frequency-aware priors for boundary refinement. These priors modulate only the query branch, enabling structure- and frequency-aware attention without disturbing the value representation, thereby reducing noise amplification.

Given an input feature map $X \in \mathbb{R}^{C \times H \times W}$ and meta-information $M \in \mathbb{R}^{C_m \times H \times W}$, where C_m is the number of meta channels, we first project M to the same channel dimension:

$$\tilde{M} = W_m * M, \tag{5}$$

where W_m is a 1×1 convolution and $*$ denotes convolution. This projected prior modulates the input feature map spatially:

$$X_m = X \odot \tilde{M}, \tag{6}$$

where $\odot$ indicates element-wise multiplication. The modulated feature generates the query Q, while the original feature generates the key K and value V:

$$Q = \phi_q(X_m), \quad K = \phi_k(X), \quad V = \phi_v(X), \tag{7}$$

where ϕ_q, ϕ_k, and ϕ_v are 1×1 projection layers. The attention output is computed as:

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^\top}{\sqrt{d}}\right)V, \tag{8}$$

where d is the projected channel dimension of Q and K. By conditioning only the query on meta-information, attention weights become prior-aware while stable value representations are preserved.

2.3 Frequency-aware Cross-scale Attention Module

Ultrasound images are often affected by high-frequency noise, which can obscure fine anatomical details. To mitigate this, we propose a Frequency-aware Cross-scale Attention Module (FCAM) that integrates multi-scale contextual modeling with frequency decomposition for adaptive noise suppression.

Given an input feature map $x \in \mathbb{R}^{C\times H\times W}$, multi-scale representations are constructed using average pooling with scale factors $s \in \{4, 2\}$:

$$x_s = \mathrm{Pool}_s(x), \tag{9}$$

where $\mathrm{Pool}_s(\cdot)$ denotes average pooling with downsampling factor s. Refinement via spatial attention is performed:

$$\hat{x}_s = \mathcal{A}(\mathrm{Conv}_{1\times1}(x_s)), \tag{10}$$

where $\mathcal{A}(\cdot)$ denotes the spatial attention operator and $\hat{x}_s$ is the refined feature at scale s. The refined features are aggregated and upsampled to the original resolution:

$$x_{\mathrm{agg}} = x + \sum_s \mathrm{Up}(\hat{x}_s), \tag{11}$$

where $\mathrm{Up}(\cdot)$ denotes upsampling to $H\times W$, and the summation is over $s \in \{4, 2\}$. To suppress high-frequency noise, a learnable low-pass filter extracts x_{low}, and the residual high-frequency component x_{high} is computed as:

$$x_{\mathrm{low}} = W * x_{\mathrm{agg}}, \quad x_{\mathrm{high}} = x_{\mathrm{agg}} - x_{\mathrm{low}},$$

where W is the learnable low-pass kernel. The high-frequency component is refined and added back to obtain the FCAM output:

$$x_{\mathrm{out}} = x_{\mathrm{agg}} + \mathrm{Conv}_{3\times3}(x_{\mathrm{high}}). \tag{12}$$

This decomposition reflects that speckle noise mainly appears as high-frequency fluctuations, whereas anatomical contours retain more stable low-frequency structures. FCAM thus refines residual details while avoiding over-smoothing.

2.4 Progressive Spectral Reconstruction Decoder

To recover fine-grained frequency details and achieve high-resolution reconstruction, we further design a Progressive Spectral Reconstruction Decoder (PSRD), which explicitly models spectral magnitude and phase information in the Fourier domain. Given an intermediate feature map $x \in \mathbb{R}^{C\times H\times W}$, we first transform it into the frequency domain via 2D real-valued Fast Fourier Transform (FFT):

$$\mathcal{F}(x) = \mathrm{FFT}(x) = M \cdot e^{j\Phi}, \tag{13}$$

where j is the imaginary unit, and $M = |\mathcal{F}(x)|$ and $\Phi = \angle\mathcal{F}(x)$ represent the magnitude and phase, respectively. The magnitude component is refined using convolutional spectral enhancement and channel attention:

$$\tilde{M} = \mathcal{C}_m(M) \odot \sigma(\mathcal{A}_c(\mathcal{C}_m(M))), \tag{14}$$

where $\mathcal{C}_m(\cdot)$ is a convolutional mapping, $\mathcal{A}_c(\cdot)$ denotes channel attention, and $\sigma(\cdot)$ is the sigmoid function. The phase component is refined via:

$$\tilde{\Phi} = \mathcal{C}_p(\Phi), \tag{15}$$

where $\mathcal{C}_p(\cdot)$ is a learnable convolutional mapping. The enhanced spectral representation is transformed back to the spatial domain using inverse FFT:

$$x_{\text{rec}} = \text{IFFT}(\tilde{M} \cdot e^{j\tilde{\Phi}}), \tag{16}$$

where x_{rec} denotes the reconstructed spatial feature. By progressively refining both magnitude and phase components, the PSRD restores high-frequency details while maintaining global structural consistency, ensuring accurate segmentation of FH and PS regions.

3 Experiments

3.1 Datasets and Comparative Methods

We used the PSFHS2023 dataset from the MICCAI 2023 Grand Challenge [3], which includes ultrasound images from the second stage of labor, with segmentation masks for the pubic symphysis (PS) and fetal head (FH). The dataset consists of 4,000 training images (305 subjects), 400 validation images (325 subjects), and 700 test images (545 subjects), with no overlap between sets to prevent data leakage.

To evaluate DMC-Net, we compared its performance against state-of-the-art methods, including fully convolutional approaches: CMU [24], Rolling [12]; hybrid convolution-transformer models: DCSAU [26], Mamba-Sea [4], U-KAN [11]. Segmentation performance was assessed using Dice Similarity Coefficient (DSC), Hausdorff Distance (HD), Average Surface Distance (ASD), ΔAoP, and Score [18]. For ΔAoP calculation (Fig. 1(a)), we segment the PS and FH regions, define the PS long axis as the primary reference line, and construct a second reference line tangent to the FH contour. The angle difference between these two lines and the ground-truth angle defines ΔAoP. Experiments were conducted on a PyTorch framework with an NVIDIA 3090 GPU (24 GB memory), batch size 16, 8 worker threads, and a random seed of 42. The Adam optimizer was used with a weight decay of 1e-4 and an initial learning rate of 1e-4. All experiments were conducted under a fixed random seed, and performance trends were consistently observed across validation runs.

3.2 Comparative Analysis

We compare DMC-Net with several state-of-the-art segmentation methods on the PSFH dataset. Fig. 2 shows qualitative results, where red contours indicate segmented PS regions, green areas denote FH, and yellow/blue contours represent ground-truth boundaries. Five representative cases with different angles

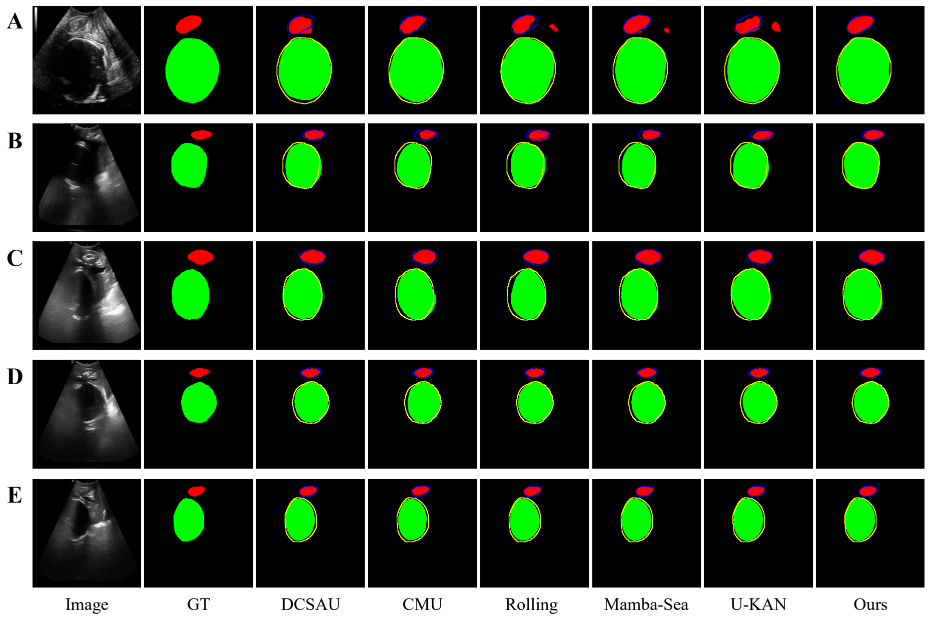

Fig. 2. Visualization of comparison results on the PSFH2023 dataset.

are presented for clarity. In Group A, DCSAU, Rolling, Mamba-Sea, and U-Kan show notable segmentation errors, while CMU produces less complete regions compared to DMC-Net. As reported in Table 1, DMC-Net achieves the best overall Dice and Score performance and competitive improvements in boundary metrics compared with recent state-of-the-art approaches.

These improvements arise from the synergy of our proposed modules. By injecting structural meta-information from shallow features, the DMCM enhances anatomical consistency and reduces noise-induced deviations. Complementing this, the FCAM integrates frequency responses with multi-scale spatial features, suppressing irrelevant high-frequency noise while preserving crucial edge details. During upsampling, the PSRD gradually restores fine-grained structures, mitigating semantic loss and maintaining structural integrity. This coordinated design enables DMC-Net to produce fewer false positives and more accurate boundary delineation, particularly in challenging regions.

3.3 Ablation Analysis

We conducted ablation experiments on the PSFH dataset to assess the contributions of the DMCM, FCAM, and PSRD modules. As shown in Table 2, incrementally adding these modules improved performance across metrics. Specifically, FCAM and PSRD enhanced DSC and boundary precision, demonstrating that noise suppression and spectral restoration improve segmentation continuity. The

Table 1. Quantitative experiments on the PSFHS2023 dataset. The best results are shown in **bold**.

Method	DSC (%)↑	HD↓	ASD↓	ΔAoP↓	Score(%)
$DCSAU_{23}$	90.30	20.72	4.68	11.55	90.57
CMU_{23}	91.32	20.39	4.13	10.24	91.57
$Rolling_{24}$	89.54	23.56	5.09	11.40	90.15
Mamba-Sea_{25}	90.14	19.58	4.73	12.58	90.31
U-KAN_{25}	90.96	20.69	4.41	9.17	91.64
Ours	**92.35**	**18.91**	**3.72**	**9.15**	**92.45**

Table 2. Effectiveness of DMCM, FCAM and PSRD in DMC-Net.

Components			Evaluation Metrics				
DMCM	FCAM	PSRD	DSC (%)↑	HD↓	ASD↓	ΔAoP↓	Score(%)
			89.19	20.19	5.20	11.37	89.29
	✓		90.99	18.43	4.31	10.68	90.95
		✓	90.13	20.69	4.79	10.63	90.66
	✓	✓	91.64	18.32	4.05	9.50	92.06
✓	✓		91.99	18.66	3.92	9.73	91.87
✓		✓	90.12	22.30	4.75	10.12	91.17
✓	✓	✓	**92.35**	**18.91**	**3.72**	**9.15**	**92.45**

full DMC-Net achieved the best results, highlighting the synergy between structural meta-conditioning and dual-domain feature refinement in fetal ultrasound analysis. All results were averaged over three independent runs with consistent performance trends.

3.4 Discussion

Although DMC-Net introduces additional attention and spectral refinement modules, the performance gains suggest that explicit dual-domain conditioning effectively addresses the intrinsic noise and structural ambiguity of intrapartum ultrasound images. The integration of structural and spectral priors enables more stable boundary delineation, particularly in regions affected by acoustic shadowing and speckle interference. Failure cases mainly occur when fetal head boundaries are severely blurred or strongly shadowed, with reliable anatomical cues partially missing. This further highlights the importance of structural and frequency priors for robust intrapartum ultrasound segmentation.

4 Conclusion

In this work, we proposed DMC-Net, a dual-domain meta-conditioning framework for fetal head and pubic symphysis segmentation in intrapartum ultrasound images. By integrating structural meta-information in the encoder and spectral refinement in the decoder, the proposed model enhances anatomical consistency and boundary precision under severe noise conditions. Experimental results on the PSFHS2023 benchmark demonstrate competitive improvements over recent segmentation approaches and improved AoP estimation accuracy. The proposed dual-domain conditioning strategy provides a principled direction for robust ultrasound representation learning. Future work will focus on lightweight optimization and spatiotemporal modeling for real-time clinical deployment.

Acknowledgments

The research work of this paper was supported by National Natural Science Foundation of China (62475198); Hubei Provincial Natural Science Foundation of China (2025AFB055); Shenzhen Science and Technology Program (JCYJ20230 807090207014); Science and Technology Research Project of Education Department of Hubei Province (Q20241403); Graduate Education Reform Research Project of Hubei University of Technology (2025025).

Disclosure of Interests

The authors declare that they have no conflicts of interest.

References

1. Adeli, K., Higgins, V., Trajcevski, K., White-Al Habeeb, N.: The canadian laboratory initiative on pediatric reference intervals: a caliper white paper. Critical reviews in clinical laboratory sciences **54**(6), 358–413 (2017)
2. Angolile, C.M., Max, B.L., Mushemba, J., Mashauri, H.L.: Global increased cesarean section rates and public health implications: A call to action. Health science reports **6**(5), e1274 (2023)
3. Chen, G., Bai, J., Ou, Z., Lu, Y., Wang, H.: Psfhs: intrapartum ultrasound image dataset for ai-based segmentation of pubic symphysis and fetal head. Scientific Data **11**(1), 436 (2024)
4. Cheng, Z., Guo, J., Zhang, J., Qi, L., Zhou, L., Shi, Y., Gao, Y.: Mamba-sea: A mamba-based framework with global-to-local sequence augmentation for generalizable medical image segmentation. IEEE Transactions on Medical Imaging (2025)
5. Ferreira, I., Simões, J., Pereira, B., Correia, J., Areia, A.L.: Ensemble learning for fetal ultrasound and maternal–fetal data to predict mode of delivery after labor induction. Scientific reports **14**(1), 15275 (2024)
6. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: International MICCAI brainlesion workshop. pp. 272–284. Springer (2021)
7. Huang, J., Mao, Y., Deng, J., Ye, Z., Zhang, Y., Zhang, J., Dong, L., Shen, H., Hou, J., Xu, Y., et al.: Emganet: Edge-aware multi-scale group-mix attention network for breast cancer ultrasound image segmentation. IEEE Journal of Biomedical and Health Informatics (2025)

8. Kalache, K.D., Dückelmann, A., Michaelis, S., Lange, J., Cichon, G., Dudenhausen, J.: Transperineal ultrasound imaging in prolonged second stage of labor with occipitoanterior presenting fetuses: how well does the 'angle of progression'predict the mode of delivery? Ultrasound in Obstetrics and Gynecology **33**(3), 326–330 (2009)
9. Khalil, A., Sotiriadis, A., D'Antonio, F., Da Silva Costa, F., Odibo, A., Prefumo, F., Papageorghiou, A., Salomon, L., et al.: Isuog practice guidelines: performance of third-trimester obstetric ultrasound scan. Ultrasound in Obstetrics & Gynecology **63**(1), 131–147 (2024)
10. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 4015–4026 (2023)
11. Li, C., Liu, X., Li, W., Wang, C., Liu, H., Liu, Y., Chen, Z., Yuan, Y.: U-kan makes strong backbone for medical image segmentation and generation. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 4652–4660 (2025)
12. Liu, Y., Zhu, H., Liu, M., Yu, H., Chen, Z., Gao, J.: Rolling-unet: Revitalizing mlp's ability to efficiently extract long-distance dependencies for medical image segmentation. In: Proceedings of the AAAI conference on artificial intelligence. vol. 38, pp. 3819–3827 (2024)
13. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B.: Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 10012–10022 (2021)
14. Mei, L., Hu, X., Xu, C., Huang, T., Ye, Z., Wang, Y., Yang, W.: Learning to optimize unsupervised image fusion with learnable loss and fusion strategy. Pattern Recognition p. 113279 (2026)
15. Mei, L., Hu, X., Ye, Z., Tang, L., Wang, Y., Li, D., Liu, Y., Hao, X., Lei, C., Xu, C., et al.: Gtmfuse: Group-attention transformer-driven multiscale dense feature-enhanced network for infrared and visible image fusion. Knowledge-Based Systems **293**, 111658 (2024)
16. Mei, L., Hu, X., Ye, Z., Ye, Z., Xu, C., Liu, S., Lei, C.: Visual fidelity and full-scale interaction driven network for infrared and visible image fusion. Pattern Recognition **165**, 111612 (2025)
17. Mei, L., Lian, C., Han, S., Jin, S., He, J., Dong, L., Wang, H., Shen, H., Lei, C., Xiong, B.: High-accuracy and lightweight image classification network for optimizing lymphoblastic leukemia diagnosis. Microscopy Research and Technique **88**(2), 489–500 (2025)
18. Mei, M., Li, Y., An, X., Ye, Z., Mei, L.: Gmanet: Gate mamba attention for fetal head and pubic symphysis segmentation in ultrasound images analysis. Digital Signal Processing p. 105533 (2025)
19. Nassr, A., Hessami, K., Berghella, V., Bibbo, C., Shamshirsaz, A., Shirdel Abdolmaleki, A., Marsoosi, V., Clark, S., Belfort, M., Shamshirsaz, A.: Angle of progression measured using transperineal ultrasound for prediction of uncomplicated operative vaginal delivery: systematic review and meta-analysis. Ultrasound in Obstetrics & Gynecology **60**(3), 338–345 (2022)
20. Noble, J.A., Boukerroui, D.: Ultrasound image segmentation: a survey. IEEE Transactions on medical imaging **25**(8), 987–1010 (2006)
21. Qiu, R., Zhou, M., Bai, J., Lu, Y., Wang, H.: Psfhsp-net: an efficient lightweight network for identifying pubic symphysis-fetal head standard plane from intrapartum ultrasound images. Medical & Biological Engineering & Computing **62**(10), 2975–2986 (2024)

22. Ruan, J., Li, J., Xiang, S.: Vm-unet: Vision mamba unet for medical image segmentation. arXiv preprint arXiv:2402.02491 (2024)
23. Sarris, I., Ioannou, C., Chamberlain, P., Ohuma, E., Roseman, F., Hoch, L., Altman, D., Papageorghiou, A., Fetal, I., for the 21st Century (INTERGROWTH-21st), N.G.C.: Intra-and interobserver variability in fetal ultrasound measurements. Ultrasound in obstetrics & gynecology **39**(3), 266–273 (2012)
24. Tang, F., Wang, L., Ning, C., Xian, M., Ding, J.: Cmu-net: a strong convmixer-based medical ultrasound image segmentation network. In: 2023 IEEE 20th international symposium on biomedical imaging (ISBI). pp. 1–5. IEEE (2023)
25. Tutschek, B., Torkildsen, E., Eggebø, T.: Comparison between ultrasound parameters and clinical examination to assess fetal head station in labor. Ultrasound in Obstetrics & Gynecology **41**(4), 425–429 (2013)
26. Xu, Q., Ma, Z., Duan, W., et al.: Dcsau-net: A deeper and more compact split-attention u-net for medical image segmentation. Computers in Biology and Medicine **154**, 106626 (2023)
27. Ye, Z., Zhang, Y., Huang, J., Wang, D., Liu, S., Mei, L., Lei, C.: Rethinking multi-center semi-supervised breast cancer ultrasound image segmentation: An intermediate-domain perspective. IEEE Journal of Biomedical and Health Informatics **30**(6), 5156–5166 (2026)