High-Frequency Guided Feature Refinement for Semi-Supervised Ultrasound Image Segmentation

Xiaming Wu[1], Jiezhang Cao[2], Xin Long[3], Bardia Rodd[4], Yonggang Liang[5], and Guoping Xu[1,6⋆]

[1] Hubei Key Laboratory of Intelligent Robot, School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan, China
[2] Shanghai Jiao Tong University, Shanghai, China
[3] Division of Hepato-Pancreato-Biliary Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[4] SUNY Upstate Medical University, Syracuse, NY, USA
[5] Department of Thoracic Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, China
[6] Intelligent Interconnected Systems Laboratory of Anhui Province, Hefei University of Technology, Hefei, China

Abstract. Semi-supervised segmentation faces an inherent trade-off between high-level semantic localization and precise boundary delineation due to limited annotated data, which is particularly severe in ultrasound imaging where strong speckle noise and low tissue contrast further degrade representation quality. While Vision Foundation Models (VFMs) such as DINOv3 provide robust global representations, their highly abstracted features tend to suppress fine-grained spatial details, resulting in suboptimal pseudo-label quality in semi-supervised settings. To address this limitation, we propose HG-SemiSeg, a high-frequency–guided feature refinement semi-supervised segmentation framework that improves pseudo-label reliability by explicitly enhancing boundary-sensitive information during upsampling. Specifically, we introduce a High-Frequency Component Enhancement (HFCE) module that randomly selects informative high-frequency components in the frequency domain to provide structural guidance for feature refinement. Built upon a frozen DINOv3 backbone and the feature-agnostic upsampling module AnyUp, HG-SemiSeg leverages the enhanced high-frequency cues produced by HFCE to facilitate spatial detail recovery. This design restores attenuated edge and texture information, thereby improving representational fidelity and pseudo-label quality. Extensive experiments demonstrate that HG-SemiSeg consistently outperforms state-of-the-art semi-supervised segmentation methods. In particular, using only 1% labeled data, HG-SemiSeg achieves average Dice Similarity Coefficient (DSC) improvements of +1.68%, +2.05%, and +1.50% over UniMatch on the CCAUI, BUS-BRA, and PSFHS ultrasound datasets, respectively, while significantly reducing boundary errors. These results validate the effectiveness of high-frequency feature enhancement during upsampling for

⋆ Corresponding author: xugp@wit.edu.cn

semi-supervised ultrasound image segmentation. Codes are available at `https://github.com/wxmadm/HG-SemiSeg`.

Keywords: Semi-supervised learning · Wavelet transform · High-frequency enhancement.

1 Introduction

Large-scale Vision Foundation Models (VFMs), such as DINOv3 [1], have significantly advanced visual representation learning by providing semantically rich features. Integrating VFMs into semi-supervised segmentation frameworks (e.g., UniMatch [2]) effectively reduces annotation costs, as these models can generate strong semantic representations that facilitate object recognition and coarse localization [3]. However, general-purpose VFMs are inherently biased toward high-level semantic abstraction at the expense of fine-grained spatial precision [4], which hampers accurate boundary delineation and degrades the quality of generated pseudo-labels. As illustrated in Fig. 1(b), features extracted from the DINOv3 backbone are spatially diffuse and struggle to capture subtle anatomical boundaries under the low contrast and strong speckle noise characteristic of ultrasound imaging [5]. Consequently, the model is limited in its ability to produce boundary-accurate pseudo-labels, thereby undermining the stability and effectiveness of semi-supervised training [6].

A critical bottleneck in pseudo-label quality arises during spatial feature recovery, where low-resolution semantic features must be reconstructed to the original image resolution by upsampling [7]. State-of-the-art upsampling methods, such as AnyUp [8] (Fig. 1(c)), typically rely on the raw input image as high-resolution guidance, which can introduce substantial irrelevant or misleading information during feature reconstruction [9]. This limitation is particularly pronounced in ultrasound imaging, where raw images often exhibit weak contrast at anatomical boundaries and severe noise contamination [10]. Consequently, directly adopting feature-agnostic upsampling strategies—such as AnyUp or bilinear interpolation—tends to produce over-smoothed predictions and imprecise pseudo-labels [11]. These observations highlight the necessity of generating more discriminative and boundary-aware feature representations during upsampling in semi-supervised settings, especially for ultrasound images with inherently low signal-to-noise ratios [12].

To address the limitations arising from low-resolution feature representations and suboptimal pseudo-label generation in ultrasound imaging, we propose a high-frequency–guided feature refinement framework for semi-supervised segmentation, termed HG-SemiSeg. It enhances pseudo-label quality by explicitly incorporating high-frequency information into the feature upsampling process to promote accurate boundary refinement. Specifically, we introduce a High-Frequency Component Enhancement (HFCE) module to strengthen boundary-sensitive features during the upsampling process, thereby producing more representative high-resolution feature maps. To this end, we leverage the Discrete

Wavelet Transform (DWT) to perform spectral decomposition and isolate high-frequency components associated with morphological edges [13]. These components are subsequently refined by the HFCE module, which amplifies structurally meaningful details while suppressing non-semantic noise for upsampled feature maps [14]. Built upon a pre-trained DINOv3 backbone and the feature-agnostic upsampling module AnyUp, HG-SemiSeg injects the enhanced high-frequency components into the upsampling pipeline to guide spatial detail recovery. This design effectively reintroduces edge and texture cues attenuated by the frozen backbone, thereby improving the representational fidelity of the reconstructed features within the HG-SemiSeg framework. As illustrated in Fig. 1(d), the refined high-frequency cues are integrated to guide the reconstruction of DINOv3 features during upsampling. By explicitly restoring boundary details suppressed by high-level abstraction, the proposed framework enables the teacher model to generate more anatomically consistent pseudo-labels, which in turn provide higher-quality supervision for the student model. Extensive experiments on three ultrasound datasets—CCAUI, BUS-BRA, and PSFHS—demonstrate that HG-SemiSeg consistently outperforms state-of-the-art semi-supervised segmentation methods. By improving the guidance source for spatial recovery, HG-SemiSeg achieves superior contour accuracy and more stable training under limited annotation settings. Our main contributions are summarized as follows: (1) We propose a High-Frequency Component Enhancement (HFCE) module that explicitly reinforces boundary-sensitive features to facilitate accurate structural recovery. (2) We develop a high-frequency–guided semi-supervised segmentation framework, HG-SemiSeg, which incorporates the proposed HFCE module to guide feature upsampling and enhance pseudo-label quality. (3) Extensive experiments on three ultrasound imaging datasets demonstrate that HG-SemiSeg consistently improves segmentation performance compared with state-of-the-art semi-supervised methods.

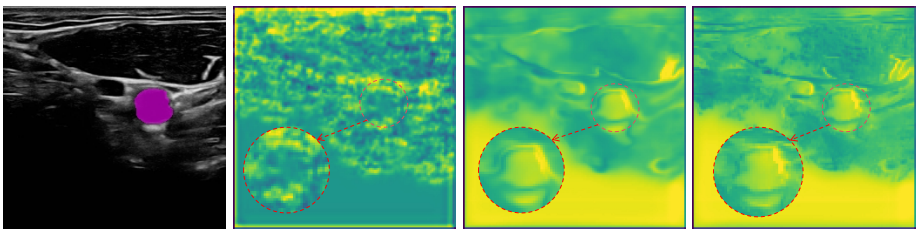

Fig. 1. Visualization of intermediate feature representations under different model configurations. (a) Input ultrasound image. (b) Feature maps extracted by the frozen DINOv3 encoder. (c) Feature maps from DINOv3 with the frozen AnyUp upsampling module. (d) Feature maps from our HG-SemiSeg model (DINOv3 + AnyUp + HFCE).

2 Methods

2.1 Overall Architecture

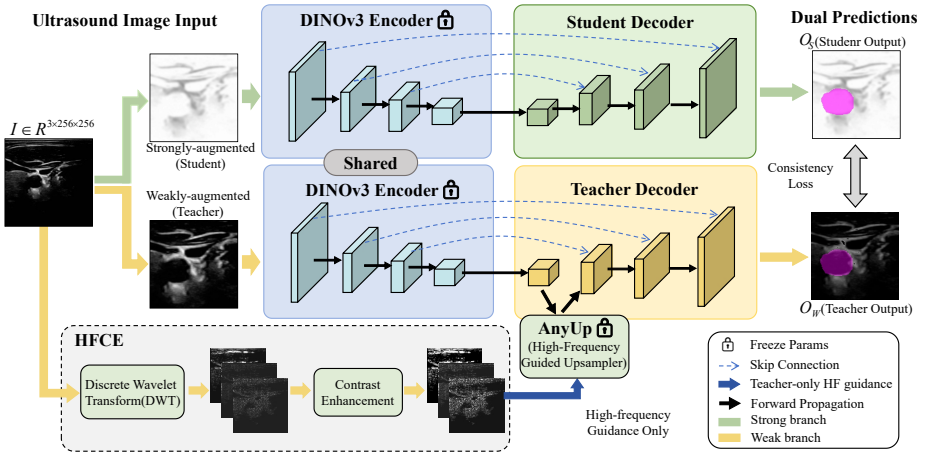

Fig. 2. Overview of the proposed HG-SemiSeg framework. A frozen DINOv3 encoder extracts multi-scale features from weakly and strongly augmented inputs. In the weakly augmented path, high-frequency components obtained via Discrete Wavelet Transform (DWT) are enhanced by the High-Frequency Component Enhancement (HFCE) module and used as structural guidance for the AnyUp upsampling module to sharpen reconstructed high-resolution features and improve boundary delineation. Although depicted as two branches for clarity, both inputs are processed by the same segmentation network with shared parameters under a weak-to-strong consistency learning paradigm. A consistency loss between the teacher prediction (O_w) and the student prediction (O_s) enables robust semi-supervised learning. The weak and strong pipelines are highlighted in yellow and green, respectively.

The overall architecture of the proposed HG-SemiSeg framework is illustrated in Fig. 2. HG-SemiSeg follows a consistency-based semi-supervised paradigm [25], in which each input is processed through weakly and strongly augmented branches. Predictions from the weakly augmented branch on unlabeled data serve as pseudo-labels to supervise the strongly augmented branch.

Features are extracted by a frozen DINOv3 encoder, which provides robust semantic representations but tends to lose fine-grained spatial details. To recover structural information from low-resolution feature maps, the High-Frequency Component Enhancement (HFCE) module decomposes features via the Discrete Wavelet Transform, using high-frequency sub-band features to guide high-resolution feature generation. Guided upsampling is applied through the AnyUp module at the first upsampling layer of the weakly augmented branch to emphasize boundary information during feature upsampling. The decoder then pro-

gressively restores full-resolution features, integrating richer boundary details compared with standard AnyUp or bilinear upsampling in a U-Net–like manner. Finally, a consistency loss is employed on unlabeled data to enforce agreement between predictions from the weakly and strongly augmented branches during semi-supervised training.

2.2 High-Frequency Component Enhancement (HFCE)

To compensate for boundary-sensitive cues lost in frozen foundation encoders, we introduce a High-Frequency Component Enhancement (HFCE) module within our framework (Fig. 2). HFCE operates in the frequency domain, selectively amplifying high-frequency information to provide structural guidance for boundary-aware decoding.

Given an input image $\mathbf{I} \in \mathbb{R}^{3\times256\times256}$, a single-level 2D discrete wavelet transform (2D-DWT) with Daubechies-3 yields

$$(\mathbf{Y}_l, \mathbf{Y}_H) = \mathrm{DWT}(\mathbf{I}; \mathrm{wave} = \mathrm{db3}, J = 1), \tag{1}$$

where $\mathbf{Y}_l$ is the low-frequency sub-band and $\mathbf{Y}_H = \{\mathbf{Y}_{\mathrm{LH}}, \mathbf{Y}_{\mathrm{HL}}, \mathbf{Y}_{\mathrm{HH}}\}$ are high-frequency sub-bands, each $\in \mathbb{R}^{1\times128\times128}$.

HFCE discards $\mathbf{Y}_l$ and applies an independent stochastic gain to each high-frequency sub-band:

$$\mathbf{Y}'_i = g_i \cdot \mathbf{Y}_i, \quad g_i \sim \mathcal{U}(g_{\min}, g_{\max}), \quad i \in \{\mathrm{LH}, \mathrm{HL}, \mathrm{HH}\}, \tag{2}$$

where $g_{\min} = 1.0$ and $g_{\max} = 5.0$ in all experiments, followed by concatenation:

$$\mathbf{Y}'_{\mathrm{H}} = \mathrm{Concat}(\mathbf{Y}'_{\mathrm{LH}}, \mathbf{Y}'_{\mathrm{HL}}, \mathbf{Y}'_{\mathrm{HH}}) \in \mathbb{R}^{3\times128\times128}. \tag{3}$$

The tensor **HF** is upsampled to the original resolution:

$$\mathbf{HF}' = \mathrm{Up}(\mathbf{Y}'_{\mathrm{H}}) \in \mathbb{R}^{3\times256\times256}. \tag{4}$$

and used as a purified, edge-focused guidance signal for the frozen AnyUp module during decoding. By replacing raw-image guidance with frequency-enhanced cues, HFCE constrains feature reconstruction to high-contrast boundaries, compensating for spatial detail loss in frozen encoders.

3 Experiments

3.1 Datasets and Implementation Details

Experiments were conducted on three publicly available ultrasound datasets:

– **CCAUI** [15]: 1,100 B-mode carotid ultrasound images with pixel-level vessel wall annotations, split into 770/110/220 images for training, validation, and testing, respectively, targeting carotid segmentation under speckle noise and low contrast.

Table 1. Quantitative comparison on the CCAUI dataset under different labeled ratios.

Method	1% labeled			3% labeled		
	DSC (%)↑	HD95↓	ASD↓	DSC (%)↑	HD95↓	ASD↓
Fullsup	95.04±0.47	3.16	1.10	95.04±0.47	3.16	1.10
ADVENT	89.15±1.22	13.90	4.97	91.06±1.04	14.16	4.23
DAN	86.10±2.13	27.63	9.40	91.12±1.60	9.30	2.97
ICT	85.32±3.71	19.23	6.07	91.52±1.45	13.16	4.03
MT	85.55±2.25	26.60	9.12	89.69±1.76	14.35	4.50
UA-MT	87.40±1.43	24.29	7.77	91.30±0.89	13.99	4.44
BCP	89.54±1.27	19.68	6.18	91.31±1.15	10.71	3.65
FixMatch	89.36±0.76	22.24	7.60	90.47±0.76	11.58	3.98
UniMatch	89.67±0.66	9.95	3.45	91.40±0.44	8.06	2.52
Ours	**91.65±0.20**	**3.66**	**1.55**	**92.78±0.07**	**4.92**	**1.93**

– **BUS-BRA** [16]: 1,875 breast ultrasound images with expert-annotated tumor boundaries, divided into 1,312/188/375 training, validation, and test samples, focusing on lesion delineation under heterogeneous echogenicity and shadowing artifacts.
– **PSFHS** [17]: 1,358 intrapartum perineal ultrasound images annotated for the fetal head and pubic symphysis, with 1,000/120/238 images for training, validation, and testing, enabling multi-structure segmentation for fetal head position assessment.

All experiments were performed on an NVIDIA RTX 4090 GPU with input images resized to 256×256. Our framework was trained for 300 epochs using SGD with an initial learning rate of 0.01, momentum of 0.9, and a batch size of 12 (6 labeled and 6 unlabeled samples).

3.2 Evaluation

To quantitatively evaluate segmentation performance, we adopt both overlap-based and boundary-based metrics, including the Dice Similarity Coefficient (DSC), 95% Hausdorff Distance (HD95), and Average Surface Distance (ASD). DSC measures global region agreement, whereas HD95 and ASD assess boundary alignment, which is particularly critical for clinically meaningful anatomical delineation. In the result tables, the fully supervised upper bound is highlighted in red, while the best-performing semi-supervised method is indicated in bold.

3.3 Quantitative Comparison

Extensive experiments on CCAUI, BUS-BRA, and PSFHS with 1% and 3% labeled data show that the proposed framework consistently outperforms representative semi-supervised baselines, including ADVENT [18], DAN [19], ICT [20],

Table 2. Quantitative comparison on the BUS-BRA dataset under different labeled ratios.

Method	1% labeled			3% labeled		
	DSC (%) ↑	HD95 ↓	ASD ↓	DSC (%) ↑	HD95 ↓	ASD ↓
Fullsup	87.70±1.69	15.18	4.86	87.70±1.69	15.18	4.86
ADVENT	54.07±6.38	58.38	22.30	65.90±6.50	54.24	20.55
DAN	55.33±5.46	70.06	29.00	68.49±6.31	40.20	15.03
ICT	60.29±6.25	60.74	23.98	71.78±5.51	38.15	14.29
MT	54.12±6.66	77.39	33.13	70.27±6.22	32.43	11.00
UA-MT	60.20±6.16	55.31	20.80	66.59±8.05	39.36	13.65
BCP	59.94±5.75	60.09	23.78	70.27±6.21	39.65	14.24
FixMatch	60.39±5.87	57.49	22.94	64.24±6.52	51.93	20.80
UniMatch	75.95±4.74	37.23	13.61	78.50±5.20	26.50	8.20
Ours	**78.45±4.36**	**32.15**	**10.46**	**80.00±4.94**	**21.57**	**6.34**

Table 3. For the PSFHS dataset, we report overall segmentation performance under different labeled ratios without category-wise breakdown, as aggregated metrics are sufficient to reflect the model's robustness and overall segmentation capability.

Method	1% labeled			3% labeled		
	DSC (%)↑	HD95↓	ASD↓	DSC (%)↑	HD95↓	ASD↓
Fullsup	93.21±1.46	6.10	2.43	93.21±1.46	6.10	2.43
ADVENT	80.32±1.68	35.86	12.38	87.91±1.44	12.33	4.28
DAN	81.08±1.95	34.89	10.86	87.62±1.54	17.45	6.01
ICT	83.17±1.69	29.12	9.91	87.96±1.48	18.39	6.20
MT	81.52±2.32	27.37	9.64	87.28±1.76	13.37	4.87
UA-MT	80.95±2.07	29.85	10.66	87.16±1.90	14.27	4.95
BCP	82.41±1.53	21.59	7.56	87.43±1.27	18.79	6.42
FixMatch	78.17±1.59	35.86	12.76	86.38±2.02	17.65	5.86
UniMatch	85.87±1.47	15.26	7.35	87.12±1.51	12.41	4.92
Ours	**87.69±0.93**	**11.54**	**4.34**	**88.30±1.03**	**10.73**	**3.96**

MT [21], UA-MT [22], BCP [23], FixMatch [24], and UniMatch, under identical splits, training, and evaluation protocols. On CCAUI (Table 1), it achieves the highest DSC and reduced boundary errors, preserving vascular structures and ensuring training stability. On BUS-BRA (Table 2), it excels under 1% labels and narrows the gap to fully supervised performance as annotations increase. On PSFHS (Table 3), it also delivers superior results, particularly on boundary-sensitive metrics. Overall, these results demonstrate the framework's effectiveness and robustness for low-annotation ultrasound segmentation.

3.4 Ablation Study

Effectiveness of Individual Modules Table 4 reports the ablation results on the CCAUI dataset, evaluating the contributions of DINOv3, AnyUp, and HFCE. Using DINOv3 alone yields a DSC of 87.18% with 1% labeled data. Introducing AnyUp improves spatial reconstruction, increasing DSC to 88.67%. Further incorporating HFCE leads to the largest gain, especially in boundary accuracy, reducing HD95 from 10.73 to 3.66 under the 1% setting. The full model consistently achieves the best performance across all metrics for both labeling ratios, highlighting the complementary roles of the three components.

Table 4. Ablation study of different model components under different labeled ratios.

Model Components			1% labeled			3% labeled		
DINOv3	AnyUp	HFCE	DSC(%)↑	HD95↓	ASD↓	DSC(%)↑	HD95↓	ASD↓
✓	✓	✓	**91.65**	**3.66**	**1.55**	**92.78**	**4.92**	**1.93**
✓	✓		88.67	9.81	3.85	89.06	7.99	2.80
✓			87.18	10.73	4.64	90.31	9.63	3.84

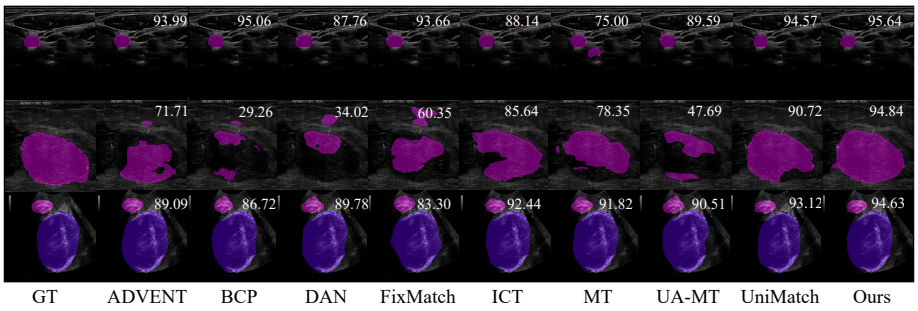

Fig. 3. Qualitative results on CCAUI, BUS-BRA, and PSFHS. Columns: GT, ADVENT, BCP, DAN, FixMatch, ICT, MT, UA-MT, UniMatch, Ours. Mean DSC (%) is shown below each prediction.

3.5 Visualization Results

To qualitatively evaluate the proposed framework, we present visual comparisons of segmentation results on three representative ultrasound datasets: CCAUI, BUS-BRA, and PSFHS. The first row shows results from CCAUI, the second row from BUS-BRA, and the third row from PSFHS. As illustrated in Fig. 3, the proposed method generates segmentation masks with improved topological

consistency and clearer boundary delineation, showing closer agreement with the ground truth (GT) than existing semi-supervised segmentation approaches.

4 Conclusion

In this work, we presented HG-SemiSeg, a high-frequency–guided feature refinement framework for semi-supervised ultrasound image segmentation. By integrating wavelet-based high-frequency structural guidance into the feature upsampling process, HG-SemiSeg enhances boundary fidelity and improves the reliability of pseudo-label generation under limited annotation settings. Extensive experimental results demonstrate consistent and robust performance gains over existing semi-supervised methods. Overall, this study underscores the importance of frequency-domain structural cues as an effective complement to frozen vision foundation models for semi-supervised medical image segmentation.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Siméoni, Oriane, Huy V. Vo, Maximilian Seitzer, Federico Baldassarre, Maxime Oquab, Cijo Jose, Vasil Khalidov et al. "Dinov3." arXiv preprint arXiv:2508.10104 (2025).
2. Yang, Lihe, Lei Qi, Litong Feng, Wayne Zhang, and Yinghuan Shi. "Revisiting weak-to-strong consistency in semi-supervised semantic segmentation." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 7236-7246. 2023.
3. Mirjalili, Vahid, Ramin Giahi, Sriram Kollipara, Akshay Kekuda, Kehui Yao, Kai Zhao, Jianpeng Xu et al. "Spatial Reasoning in Foundation Models: Benchmarking Object-Centric Spatial Understanding." arXiv preprint arXiv:2509.21922 (2025).
4. Dong, Caixia, Duwei Dai, Xinyi Han, Fan Liu, Xu Yang, Zongfang Li, and Songhua Xu. "Unleashing Vision Foundation Models for Coronary Artery Segmentation: Parallel ViT-CNN Encoding and Variational Fusion." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 647-657. Cham: Springer Nature Switzerland, 2025.
5. Yang, Sicheng, Hongqiu Wang, Zhaohu Xing, Sixiang Chen, and Lei Zhu. "Segdino: An efficient design for medical and natural image segmentation with dino-v3." arXiv preprint arXiv:2509.00833 (2025).
6. Gao, Yifan, Haoyue Li, Feng Yuan, Xiaosong Wang, and Xin Gao. "Dino u-net: Exploiting high-fidelity dense features from foundation models for medical image segmentation." arXiv preprint arXiv:2508.20909 (2025).
7. Huang, Haiwen, Anpei Chen, Volodymyr Havrylov, Andreas Geiger, and Dan Zhang. "LoftUp: Learning a Coordinate-Based Feature Upsampler for Vision Foundation Models." arXiv preprint arXiv:2504.14032 (2025).
8. Wimmer, Thomas, Prune Truong, Marie-Julie Rakotosaona, Michael Oechsle, Federico Tombari, Bernt Schiele, and Jan Eric Lenssen. "AnyUp: Universal Feature Upsampling." In Proceedings of the International Conference on Learning Representations (ICLR), 2026.

9. Wang, Zhihao, Jian Chen, and Steven CH Hoi. "Deep learning for image super-resolution: A survey." IEEE transactions on pattern analysis and machine intelligence 43, no. 10 (2020): 3365-3387.
10. Gago, Lucas, Miguel A. Fernández González, Justin Engelmann, Beatriz Remeseiro, and Laura Igual. "Bridging the quality gap: Robust colon wall segmentation in noisy transabdominal ultrasound." Computers in Biology and Medicine 197 (2025): 111077.
11. Chen, Fang, Lingyu Chen, Wentao Kong, Weijing Zhang, Pengfei Zheng, Liang Sun, Daoqiang Zhang, and Hongen Liao. "Deep semi-supervised ultrasound image segmentation by using a shadow aware network with boundary refinement." IEEE transactions on medical imaging 42, no. 12 (2023): 3779-3793.
12. He, Qi, Xianghao Cui, Qingjing Fei, Wen Xiong, Yongjie Pang, Wenying Liu, Zhi Chen, and Fang Hou. "Masked pretraining of U-Net for ultrasound image segmentation." Scientific Reports 15, no. 1 (2025): 31713.
13. Yue, Wenbo, Xiaming Wu, Qing Huang, Xinglong Wu, Chang Li, Yajun Yu, Jun Lyu, and Guoping Xu. "Wavelet-Based Frequency Replacement and Edge Enhancement for Semi-Supervised Fetal Ultrasound Image Segmentation." Journal of Ultrasound in Medicine (2026).
14. Leng, Xuesong, Xiaxia Wang, Wenbo Yue, Jianxiu Jin, and Guoping Xu. "Structural tensor and frequency guided semi-supervised segmentation for medical images." Medical Physics 51, no. 12 (2024): 8929-8942.
15. Momot, Agata (2022), "Common Carotid Artery Ultrasound Images", Mendeley Data, V1, doi: 10.17632/d4xt63mgjm.1.
16. Gómez-Flores, Wilfrido, Maria Julia Gregorio-Calas, and Wagner Coelho de Albuquerque Pereira. "BUS-BRA: a breast ultrasound dataset for assessing computer-aided diagnosis systems." Medical Physics 51, no. 4 (2024): 3110-3123.
17. Chen, Gaowen, Jieyun Bai, Zhanhong Ou, Yaosheng Lu, and Huijin Wang. "PSFHS: intrapartum ultrasound image dataset for AI-based segmentation of pubic symphysis and fetal head." Scientific Data 11, no. 1 (2024): 436.
18. Vu, Tuan-Hung, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Pérez. "Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2517-2526. 2019.
19. Zhang, Yizhe, Lin Yang, Jianxu Chen, Maridel Fredericksen, David P. Hughes, and Danny Z. Chen. "Deep adversarial networks for biomedical image segmentation utilizing unannotated images." In International conference on medical image computing and computer-assisted intervention, pp. 408-416. Cham: Springer International Publishing, 2017.
20. Verma, Vikas, Kenji Kawaguchi, Alex Lamb, Juho Kannala, Arno Solin, Yoshua Bengio, and David Lopez-Paz. "Interpolation consistency training for semi-supervised learning." Neural Networks 145 (2022): 90-106.
21. Tarvainen, Antti, and Harri Valpola. "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results." Advances in neural information processing systems 30 (2017).
22. Yu, Lequan, Shujun Wang, Xiaomeng Li, Chi-Wing Fu, and Pheng-Ann Heng. "Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation." In International conference on medical image computing and computer-assisted intervention, pp. 605-613. Cham: Springer International Publishing, 2019.
23. Bai, Yunhao, Duowen Chen, Qingli Li, Wei Shen, and Yan Wang. "Bidirectional copy-paste for semi-supervised medical image segmentation." In Proceedings of the

IEEE/CVF conference on computer vision and pattern recognition, pp. 11514-11524. 2023.

24. Sohn, Kihyuk, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A. Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li. "Fixmatch: Simplifying semi-supervised learning with consistency and confidence." Advances in neural information processing systems 33 (2020): 596-608.
25. Guoping Xu, Jayaram K. Udupa, Jax Luo, Songlin Zhao, Yajun Yu, Scott B. Raymond, Xiaoxue Qian, Nian Wang, Hao Peng, Steve Jiang, Weiguo Lu, Lipeng Ning, Yogesh Rathi, Wei Liu, You Zhang. "Is the medical image segmentation problem solved? A survey of current developments and future directions."*ArXiv Preprint ArXiv:2508.20139.* (2025)