Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

DMPB-Net: A Mamba-Driven Semi-Supervised Network for Precise Breast Ultrasound Lesion Segmentation

Haiyang Sun

sunhy@niit.edu.cn

Nanjing University of Industry Technology

Xuanbo Xu

Shantou Power Supply Bureau of Guangdong Power Grid

Li Wang

Nanjing University of Industry Technology

Article

Keywords: Breast Ultrasound, Lesion Segmentation, Mamba, Semi-supervised Learning

Posted Date: July 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10095241/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

DMPB-Net: A Mamba-Driven Semi-Supervised Network for Precise Breast Ultrasound Lesion Segmentation

Haiyang Sun[1]*, Xuanbo Xu[2], Li Wang[1]

1 School of Computer and Software, Nanjing University of Industry Technology, Nanjing 210023, China; sunhy@niit.edu.cn; li1019wang@niit.edu.cn;

2 Shantou Power Supply Bureau of Guangdong Power Grid, Shantou, 515000, China 15651731350@163.com

* Correspondence: sunhy@niit.edu.cn

Abstract: Accurate segmentation of breast lesions in ultrasound images is critical for early cancer diagnosis, yet it remains challenging due to speckle noise, acoustic artifacts, and blurred boundaries. In this paper, we propose DMPB-Net, a novel semi-supervised framework designed for precise breast lesion segmentation under limited supervision. The framework utilizes a Mamba-driven encoder to capture long-range global dependencies, effectively overcoming the receptive field limitations of conventional CNNs. To mitigate ultrasound-specific interference, we introduce a Semantic Decoupling (SD) module that separates lesion-related features from task-irrelevant background noise. Furthermore, a Confidence-Ranked Prototype Attention (CR-PA) module is developed to facilitate robust semantic alignment by dynamically filtering high-confidence prototypes. To address boundary ambiguity, a Boundary Iterative Refinement (BIR) module is employed to progressively polish masks through a generative denoising process. Finally, a composite loss function comprising Consistency, Fuzzy Rough Set Boundary (FRSB), and Structural Similarity Contrastive (SSC) losses is introduced to supervise both local details and global topology. Extensive experiments on BUSI and UDIAT datasets demonstrate that DMPB-Net significantly outperforms state-of-the-art methods, achieving a Dice score of 83.20% and 91.51% respectively, showcasing its high clinical potential for computer-aided diagnosis. Code is available at https://github.com/HengfanLi/DMPB-Net2.

Keywords: Breast Ultrasound; Lesion Segmentation; Mamba; Semi-supervised Learning

1.Introduction

Breast cancer represents the most prevalent malignancy among females worldwide, posing a formidable challenge to global public health due to its substantial incidence rate. Clinical evidence underscores that early intervention and diagnosis serve as pivotal pathways for reducing patient mortality [1,2]. Consequently, the development of diagnostic instruments capable of facilitating timely detection and precise assessment is of paramount strategic significance for arresting disease progression, optimizing clinical therapeutic regimens, and improving patient prognoses [3].

Early and precise diagnosis of breast tumors is critical for optimizing clinical outcomes. While B-mode ultrasound (B-US) is the preferred screening modality due to its real-time capability and cost-effectiveness, automated lesion segmentation remains challenging. This difficulty primarily arises from the intricate interplay between ultrasound's physical artifacts and the complex biological morphology of breast tumors [4,5].

These challenges manifest in two primary dimensions: morphologically, invasive growth leads to blurred interfaces and spiculate boundaries lacking clear contours; acoustically, varying pathological types—such as fibroadenomas versus invasive ductal carcinomas—exhibit distinct impedance characteristics. These variations result in diverse echo intensities, posterior acoustic artifacts, and heterogeneous textures. Such uncertainties elevate the risk of misinterpretation in computer-aided diagnosis (CAD), potentially leading to overdiagnosis or missed detections [6].

Clinical practice remains heavily dependent on subjective radiological interpretation, leading to significant inter-observer variability. Factors such as cognitive fatigue, varying expertise levels, and the complexity of multi-modal data integration often undermine diagnostic consistency [7]. Consequently, developing robust, interpretable automated segmentation algorithms is essential to provide objective, reproducible evidence for malignancy classification and surgical planning [8].

To address these challenges, deep learning-driven CAD systems have emerged as a vital pathway for enhancing diagnostic standardization [9]. Recently, Convolutional Neural Networks (CNNs) have demonstrated superior feature representation in medical image analysis. Their performance in lesion detection, classification, and segmentation now rivals or surpasses traditional methods, supported by extensive clinical validation [10,11].

This progress stems from continuous architectural innovations: FCN [12] pioneered end-to-end pixel-level prediction, while U-Net [13] optimized segmentation on limited data via its encoder-decoder structure and skip connections. SegNet [14] and DeconvNet [15] introduced diverse upsampling strategies to restore spatial resolution, and the DeepLab series employed atrous convolution and multi-scale fusion to refine target perception. Collectively, these advancements provide a robust algorithmic foundation for precise breast lesion segmentation in ultrasound imaging [16].

Yet, the efficacy of these frameworks remains tied to the fully supervised learning paradigm, which demands large-scale, high-quality pixel-level annotations. In breast ultrasound, acquiring such precise manual segmentations is labor-intensive and requires specialized clinical expertise, creating a major bottleneck for model generalizability. To mitigate this, semi-supervised learning (SSL) has gained attention for its ability to leverage limited labeled data alongside vast unlabeled pools. However, current SSL frameworks often struggle with standard convolutional kernels' inability to capture long-range spatial dependencies. Furthermore, consistency-based methods remain highly susceptible to pseudo-label noise, particularly near fuzzy tumor boundaries. To address morphological heterogeneity and label scarcity in breast ultrasound (BUS) images, we propose DMPB-Net, a semi-supervised framework featuring a Mamba-inspired encoder-decoder.

By leveraging State-Space Models (SSM), the network achieves a global receptive field with linear complexity, while a semantic decoupling mechanism isolates lesion-specific features from noisy sonographic backgrounds. To maximize the utility of limited supervision, the Confidence-Ranked Prototype Attention (CR-PA) module is incorporated. It dynamically ranks pixel-level features based on prediction confidence to generate representative prototypes. This mechanism aligns labeled and unlabeled data, ensuring robust semantic consistency across the semi-supervised latent space. To resolve boundary ambiguity, the Boundary Iterative Refinement (BIR) stage treats mask optimization as a generative denoising process. Inspired by diffusion models, it views initial fuzzy masks as noisy states and progressively "sculpts" sharp, anatomically accurate contours through successive residual corrections, effectively filtering out pseudo-label noise. Finally, the framework is optimized via a composite loss strategy featuring Fuzzy Rough Set Boundary (FRSB) and Structural Similarity Contrastive (SSC) losses. These components explicitly model edge uncertainty and enforce global structural integrity, providing a superior solution for accurate breast cancer segmentation under limited supervision. Our contributions are summarized as follows:

(1) We propose DMPB-Net, a novel semi-supervised BUS segmentation framework that combines a Mamba-inspired encoder for efficient long-range dependency modeling with a semantic decoupling mechanism to isolate lesion features from complex sonographic backgrounds.

(2) We introduce the CR-PA module, which ensures semantic consistency by ranking pixel-level features by confidence to generate representative class prototypes. This mechanism facilitates robust knowledge transfer and feature alignment between labeled and unlabeled data under limited supervision.

(3) We design a Boundary Iterative Refinement (BIR) stage that treats mask optimization as a generative denoising process to sharpen ambiguous contours. This is supervised by a composite loss (incorporating FRSB and SSC), which explicitly models edge uncertainty and ensures global structural integrity.

(4) Our method achieves state-of-the-art performance on two benchmark datasets (BUSI and UDIAT), consistently outperforming competitive methods across various partition protocols and clinical scenarios.

2. Related work

2.1 Semi-Supervised BUS Image Segmentation

The synergistic advancement of Breast Ultrasound (BUS) segmentation and semi-supervised learning frameworks has catalyzed a growing research interest in their integration. This integrated approach aims to overcome the intrinsic challenges of BUS imaging — namely multiscale morphological variations, ill-defined boundaries, low contrast, and speckle artifacts — thereby enhancing clinical diagnostic precision. Recently, Jiang et al. [17] propose a novel semi-supervised framework (CoBiNet) that enhances the segmentation of ambiguous lesion regions by integrating confidence-ranked features with bi-level prototypes to facilitate both directional contrastive learning

and cross-guided consistency learning. To address the inherent limitations of pseudo-labels, Liu et al. [18] introduce a semi-supervised framework (CGDM) that leverages a contrast-guided diffusion model to restore lesion areas to healthy-like tissues, thereby alleviating the model's over-reliance on noisy supervision through structural similarity contrast (SSC) learning. In a different approach to semantic enhancement, Zhang et al. [19] develop SDC-Net, which addresses low-contrast challenges by decoupling semantic information into distinct components to refine lesion boundary delineation. Building on the synergy of different architectures, Yang et al. [20] introduce a multi-paradigm synergy and cross-cyclic distillation (MSCD) method that integrates CNN and Transformer structures to maximize data utilization in semi-supervised BUS segmentation. Parallel to these developments, Wen et al. [21] develop a class-aware multi-level attention learning model specifically designed to mitigate class imbalance and enhance feature extraction under limited label distribution. Furthermore, Ding et al. [22] present CSC-PA, a single-network framework that leverages foreground and edge prototype attentions to capture cross-image semantic correlations for more accurate tumor segmentation. Beyond pure segmentation, several studies have focused on integrated detection and classification frameworks. Richa et al. [23] propose a deep learning framework (DLF) that combines CNNs with feature selection and fusion methods to improve early detection accuracy. Most recently, Zhang et al. [24] propose RA-BUSSeg, a relation-aware semi-supervised framework that enhances the robustness of BUS image segmentation by leveraging intrinsic pixel relations through adjacent propagation and cross-layer alignment to refine pseudo-label quality. Despite these advancements, existing methods still struggle to balance long-range modeling with semantic decoupling, and they lack a generative mechanism to iteratively refine precise boundaries from noisy, fuzzy masks.

2.2 Mamba-based Architectures in Medical Image Segmentation

Recently, the advent of State-Space Models (SSMs), particularly the Mamba architecture, has revolutionized medical image segmentation by offering a compelling alternative to Transformers. Unlike traditional CNNs with limited receptive fields or Transformers with quadratic complexity, Mamba-based models leverage selective scan mechanisms to achieve global context modeling with linear complexity. In the realm of volumetric analysis, Xing et al. [25] propose SegMamba, a novel 3D medical image segmentation model that effectively captures long-range dependencies within whole-volume features at multiple scales while maintaining superior memory efficiency. Expanding upon foundation models, Liang et al. [26] introduce MambaSAM, a framework that integrates a lightweight Visual Mamba adapter with the pre-trained SAM ViT encoder to enhance multi-scale contextual correlation and mitigate ambiguities. Focusing on architectural refinements, Wu et al. [27] develop SK-VM++, which incorporates Mamba blocks into skip-connections to facilitate the adaptive fusion of high- and low-level features, thereby optimizing the information flow within U-shaped architectures. To further address morphological variability, Li et al. [28] present a selective and multi-scale fusion Mamba network that dynamically aggregates local and global features through a specialized multi-scale state-space modeling approach. In a similar vein of hybrid

modeling, Zhong et al. [29] introduce VMAXL-UNet, a network that synergizes Visual State Space (VSS) modules with lightweight LSTMs (xLSTM) to bridge local boundary details and global semantic context. This integration of local and global strengths is also evident in the work of Xu et al. [30], who propose Mamba VesselNet++, a hybrid CNN-Mamba architecture specifically designed to capture intricate vascular structures by combining convolutional feature extraction with efficient long-range modeling. Moving toward model robustness, Cheng et al. [31] develop Mamba-Sea, a generalizable framework that incorporates global-to-local sequence augmentation to improve performance against distribution shifts in medical imaging. Finally, emphasizing computational efficiency, Kui et al. [32] present Gl-MambaNet, a dual-path network that fuses CNN and Mamba architectures through a global-local feature fusion module to capture comprehensive spatial information with minimal overhead. Despite these advancements, most Mamba-based methods focus on fully supervised architectural changes and struggle to decouple lesion features from noisy backgrounds, while lacking an iterative generative mechanism to refine precise boundaries from fuzzy, low-quality masks under limited supervision. Despite these advancements, most Mamba-based methods focus on fully supervised architectures, struggling to decouple features from noisy backgrounds or handle the pseudo-label noise inherent in semi-supervised scenarios.

2.3 Prototype Consistency Learning

Prototype learning has emerged as a powerful paradigm in semi-supervised medical image segmentation (SSMIS), addressing the limitations of traditional pixel-wise consistency by capturing representative class-level semantic embeddings. Unlike standard methods that rely on voxel-level predictions, prototype-based approaches represent each category as a latent cluster center (prototype) in the feature space, facilitating more robust feature alignment between labeled and unlabeled data.

Early integration of prototypes into the Mean Teacher framework, such as PAMT [33], demonstrated that non-learnable prototypes could significantly enhance feature-space supervision, moving beyond simple label-space consistency. To further stabilize the training process, Zhang et al. [34] introduced a stability-aware dual-head network that utilizes prototype-based consistency to handle the inherent noise in medical images. By comparing pixel-to-prototype distances, these methods convert the segmentation task into a feature matching problem, which is inherently more resilient to class imbalance and intra-class variation.

Recent advancements have focused on more sophisticated alignment strategies, particularly through contrastive learning. PCCS [35] and Yue et al. [36] proposed dual-prototype contrastive learning, which enforces intra-class compactness and inter-class separability by pulling similar semantic prototypes together while pushing disparate ones apart. This strategy is further extended to cross-image scenarios; for instance, DuCiSC [37] explicitly aligns prototypes across labeled, unlabeled, and even fused images to establish a unified semantic space, thereby mitigating the feature discrepancy issue caused by limited supervision. Similarly, CPL [38] leverages clustering-guided queries in a Mask Transformer to learn more discriminative prototypes, effectively

addressing the "pseudo-label bias" in long-tail medical datasets.

Furthermore, ensuring the reliability of prototypes has become a critical research direction. EPCL-JUDA [39] integrates joint uncertainty quantification with data augmentation to enhance the semantic expression of prototypes, ensuring that only high-quality feature embeddings contribute to the global class representation. In the context of evolving clinical requirements, PRIME [40] demonstrates the versatility of prototype learning by applying it to class-incremental segmentation, using prototypes to preserve knowledge of old categories without requiring their original labels. Collectively, these studies suggest that prototype consistency learning provides a more comprehensive semantic regularizer than pixel-level methods, enabling models to achieve superior generalization in challenging domains like breast ultrasound (BUS) segmentation where boundary ambiguity and data scarcity are prevalent.

2.4 Iterative Boundary Refinement

Precise boundary delineation remains a formidable challenge in medical image segmentation, particularly for breast ultrasound (BUS) images characterized by severe boundary ambiguity and low contrast. Traditional one-step inference methods often struggle to capture intricate anatomical contours. Consequently, Iterative Boundary Refinement has emerged as a robust paradigm to progressively rectify initial coarse predictions.

Early research focuses on explicit boundary modeling and feature alignment. BFA [41] introduces a boundary feature alignment method to resolve the imbalance between global consistency and local edge precision in semi-supervised learning. To force the network to focus on sparse edge pixels, BOCLIS [42] proposes a boundary-guided patch sampling strategy, enabling the framework to learn more discriminative representations in transition zones. These methods establish the foundation for boundary-aware learning by emphasizing the importance of local geometric features.

To mitigate the accumulation of pseudo-label noise during training, multi-scale and dynamic refinement mechanisms are explored. MBPLR [43] proposes a multi-scale boundary enhanced refinement module that dynamically adjusts pixel-level weights to amplify correct boundary signals while suppressing incorrect ones across different scales. Similarly, CBDPL [44] utilizes confidence-based dynamic pseudo-labeling to ensure that only reliable boundary information is propagated during the iterative optimization process. These strategies transition from static feature enhancement to dynamic label correction.

Recent advancements reframe boundary refinement as a progressive evolution or generative process. DPGNet [45] introduces a three-stage progressive refinement strategy, leveraging an Edge Difference Attention (EDA) module to implicitly quantify boundary uncertainties without explicit supervision. Taking a step further, MARL-MambaContour [46] models boundary refinement as a multi-step reinforcement learning task, employing multi-agent collaboration for active contour optimization. This approach effectively “sculpts” anatomically accurate contours through successive residual corrections, mimicking the cognitive process of clinical experts. By integrating

such iterative mechanisms, contemporary frameworks can effectively filter pseudo-label noise and reconstruct sharp boundaries even in the presence of intense speckle noise and artifacts.

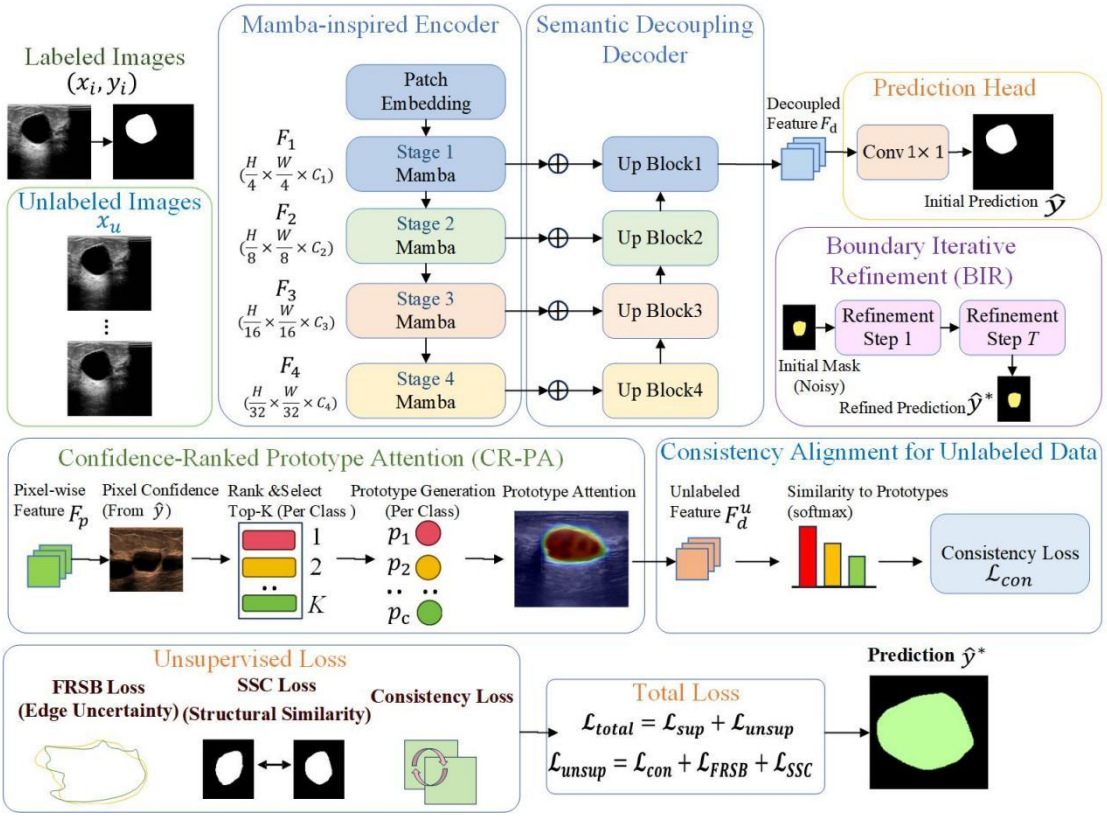


Figure. 1. Architecture of the proposed DMPB-Net.

3. Method

3.1 Overview

The inherent characteristics of breast ultrasound (BUS) images, such as low contrast, shadows, and blurry boundaries, coupled with the expensive cost of expert annotations, hinder the development of robust segmentation models. To address these limitations, we propose a novel semi-supervised framework named DMPB-Net, which integrates long-range dependency modeling with generative boundary refinement.

As illustrated in Figure. 1, unlike conventional CNN-based architectures, DMPB-Net utilizes a Mamba-inspired encoder-decoder backbone with a semantic decoupling mechanism to separate lesion-specific representations from complex background noise. To leverage cross-image information in a semi-supervised manner, we incorporate a Confidence-Ranked Prototype Attention (CR-PA) module. This module dynamically maintains a global memory bank of lesion and background prototypes, facilitating the transfer of high-confidence semantic knowledge to guide the learning of unlabeled data.

Furthermore, to resolve the issue of boundary ambiguity, we introduce a Boundary Diffusion Refinement (BDR) stage. By modeling the segmentation task as a conditional denoising process, the BDR effectively suppresses pseudo-label noise and sharpens the lesion contours through iterative refinement. To optimize the framework, we design a Fuzzy Rough Set Boundary (FRSB)

loss and a Structural Similarity Contrastive (SSC) loss. These functions explicitly model the uncertainty of transition regions and enforce structural consistency between initial and refined masks, thereby significantly enhancing the model's discriminative power and generalization across multi-center clinical datasets.

3.2. Mamba-driven Encoder with Semantic Decoupling

The proposed encoder is designed to capture long-range contextual dependencies while maintaining computational efficiency, which is critical for identifying lesions with diverse scales in breast ultrasound (BUS) images. Unlike traditional CNN-based encoders that suffer from a limited receptive field or Transformer-based models with quadratic complexity, our encoder leverages the State Space Model (SSM) backbone, specifically the Mamba architecture, to achieve linear-time sequence modeling.

3.2.1. Mamba-based Feature Extraction

The core of our encoder consists of a series of Mamba blocks. Given an input feature map $\mathbf{F} \in \mathbb{R}^{H\times W\times C}$, we first flatten it into a sequence $\mathbf{x} \in \mathbb{R}^{L\times C}$ where $L = H \times W$. Each Mamba block processes the sequence through a selective scan mechanism. The continuous system of SSM is discretized as follows:

$$\mathbf{h}_t = \bar{\mathbf{A}}\mathbf{h}_{t-1} + \bar{\mathbf{B}}\mathbf{x}_t \tag{1}$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{h}_{t-1} + \mathbf{D}\mathbf{x}_t \tag{2}$$

where $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ are discretized parameters derived from the continuous state matrices **A** and **B** via a timescale parameter Δ. By employing the bidirectional selective scan (Bi-SSM), the encoder effectively aggregates spatial information from both horizontal and vertical directions, ensuring a comprehensive global receptive field that is essential for perceiving the overall structure of breast tumors.

3.2.2. Semantic Decoupling Mechanism

Medical ultrasound images are often corrupted by speckle noise and shadow artifacts, which are frequently entangled with lesion features in the latent space. To address this, we introduce a Semantic Decoupling (SD) strategy within the encoder stages to explicitly separate task-relevant tumor information from task-irrelevant background interference.

Specifically, at each hierarchy of the encoder, the latent feature $\mathbf{f}_l$ is decomposed into two orthogonal components: a tumor-specific feature $\mathbf{f}_{ts}$ and a background-dominant feature $\mathbf{f}_{bg}$. This decoupling is implemented through a learnable projection mechanism:

$$\mathbf{f}_{ts} = \mathbf{\Phi}_{ts}(\mathbf{f}_l),\ \mathbf{f}_{bg} = \mathbf{\Phi}_{bg}(\mathbf{f}_l) \tag{3}$$

where $\mathbf{\Phi}_{ts}$ and $\mathbf{\Phi}_{bg}$ represent distinct transformation branches. To ensure effective separation, we apply a decoupling constraint that minimizes the mutual information between $\mathbf{f}_{ts}$ and $\mathbf{f}_{bg}$. By suppressing the background-dominant features and emphasizing the lesion-related semantics during the encoding phase, the model enhances its sensitivity to subtle mass boundaries and internal echo patterns, even in the presence of low-contrast artifacts. This decoupled representation provides a

clean and highly discriminative embedding for the subsequent prototype generation and boundary refinement stages.

3.3. Confidence-Ranked Prototype Attention

To mitigate the reliance on extensive pixel-level annotations and bridge the distribution gap between labeled and unlabeled data, we propose the Confidence-Ranked Prototype Attention (CR-PA) module. This module dynamically aggregates semantic information into a set of representative class prototypes, which serve as global anchors to guide feature alignment in the semi-supervised latent space.

As illustrated in Figure. 2, the Confidence-Ranked Prototype Attention (CR-PA) module is designed to bridge the distribution gap between labeled and unlabeled data while mitigating the reliance on extensive pixel-level annotations. The detailed workflow of the module is as follows: First, the module extracts Initial Prototypes from labeled feature via class-specific mask pooling and generates Candidate Prototypes from unlabeled features through grid-based aggregation. Subsequently, a Confidence-Ranked Filtering mechanism is introduced to evaluate the reliability of these candidates by calculating similarity-based confidence scores. This process effectively identifies and discards low-confidence features corrupted by noise or acoustic artifacts, thereby constructing a high-quality Confident Prototype Bank. Finally, this bank serves as a set of global anchors that undergo pixel-wise dot-product operations with the Query Feature Map to generate a precise Attention Weight Map. Through this dynamic filtering and alignment strategy, the CR-PA module not only achieves robust feature alignment in the latent space but also significantly enhances the model's discriminative power for key semantic features under limited supervision.

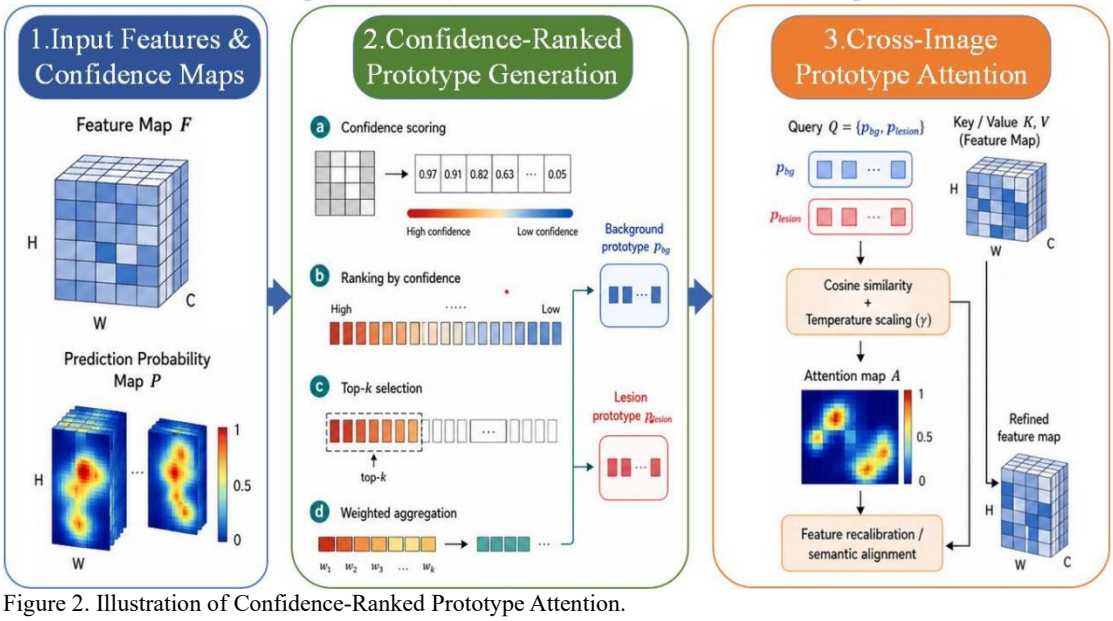


Figure 2. Illustration of Confidence-Ranked Prototype Attention.

3.3.1. Confidence-Ranked Prototype Generation

Conventional prototype generation methods often suffer from noisy pseudo-labels, where low-confidence predictions introduce irrelevant features into the class representation. To ensure the purity of the prototypes, we introduce a Confidence-Ranking mechanism. For each class $c \in \{0,1\}$

(background and lesion), we first obtain the prediction probability map $\mathbf{P}_c$ from the student network. For unlabeled images, the pixel-level features f_i are filtered and ranked based on their corresponding confidence scores $p_{i,c} \in \mathbf{P}_c$.

We define a dynamic threshold τ to select the top-k most confident pixels for prototype estimation. The prototype $\mathbf{v}_c$ for class c is computed as a weighted average of these high-confidence feature vectors:

$$\mathbf{v}_c = \frac{\sum_{i \in \Omega_c} 1(p_{i,c} > \tau) \cdot w_i \cdot \mathbf{f}_i}{\sum_{i \in \Omega_c} 1(p_{i,c} > \tau) \cdot w_i} \tag{4}$$

where Ω_c denotes the spatial domain of class c, and w_i is a normalized weight proportional to the confidence $p_{i,c}$. This ranking ensures that the resulting prototypes are robust to the "pseudo-label bias" inherent in the early stages of semi-supervised training.

3.3.2. Cross-Image Prototype Attention

Once the stable prototypes $\{\mathbf{v}_0, \mathbf{v}_1\}$ are generated from both labeled and unlabeled batches, we employ a Cross-Image Prototype Attention mechanism to enforce semantic consistency. We treat the prototypes as the Query (Q) and the local feature maps as Key (K) and Value (V) to perform a global-to-local alignment.

Specifically, the attention map $\mathbf{A}_c$ is calculated by measuring the cosine similarity between the prototype $\mathbf{v}_c$ and all pixel features $\mathbf{f}_i$ in a target image:

$$\mathbf{A}_c(i) = \frac{\exp(\mathrm{sim}(\mathbf{v}_c, \mathbf{f}_i)/\sigma)}{\sum_j \exp(\mathrm{sim}(\mathbf{v}_c, \mathbf{f}_i)/\sigma)} \tag{5}$$

where sim denotes cosine similarity and σ is a temperature scaling factor. This attention mechanism re-calibrates the feature response of unlabeled samples based on the reliable semantic guidance from the prototypes. By pulling the features of unlabeled pixels closer to their corresponding high-confidence prototypes and pushing them away from disparate ones, CR-PA achieves robust semantic alignment and effectively transfers knowledge across the labeled-unlabeled boundary.

3.4. Generative Boundary Iterative Refinement

To overcome the limitations of static single-pass segmentation, which often fails to capture precise contours in low-contrast BUS images, we introduce the generative Boundary Iterative Refinement (BIR) stage. Inspired by the denoising diffusion paradigm, BIR reframes mask optimization as a multi-step generative process, where an initial coarse segmentation is treated as a "noisy" estimate that is progressively refined into a sharp, anatomically accurate boundary.

3.4.1. Iterative Denoising Formulation

Let $\mathbf{M}_0$ be the initial coarse mask generated by the decoder. We define a sequence of refinement steps $t = 1, \ldots, T$, where each step produces a more refined mask $\mathbf{M}_t$ by estimating the structural residual between the current estimate and the latent true boundary. The refinement process can be formulated as a conditional transition:

$$\mathbf{M}_t = \mathbf{M}_{t-1} + R(\mathbf{f}_l, \mathbf{M}_{t-1}; \theta_{BIR}) \tag{6}$$

Where R is a lightweight residual refinement network parameterized by θ_{BIR}, and $\mathbf{f}_l$ represents the multi-scale latent features from the Mamba-driven encoder. By incorporating $\mathbf{f}_l$ as a spatial condition, R learns to identify boundary pixels that are misclassified due to acoustic shadows or speckle noise.

3.4.2. Progressive Structural Recovery

Unlike standard iterative methods that merely smooth the mask, our BIR module explicitly treats the refinement as a generative structural recovery. At each iteration *t*, the module predicts a boundary-aware correction map $\mathbf{E}_t$. To handle the high uncertainty at tumor edges, we introduce a stochastic sampling strategy during the early iterations to explore potential boundary candidates, which gradually transitions to a deterministic sharpening phase.

The iterative update is supervised by comparing the final refined mask $\mathbf{M}_T$ with the ground truth (for labeled data) or the high-confidence pseudo-labels (for unlabeled data). This progressive mechanism allows the model to "sculpt" the lesion contours through successive residual corrections, effectively filtering out the artifacts that often plague single-stage ultrasound segmentation. The final output $\mathbf{M}_T$ exhibits significantly higher structural integrity and sharper edges, providing a more reliable basis for clinical diagnosis.

3.5. Composite Loss Function

The overall optimization objective of DMPB-Net is formulated as a combination of supervised and unsupervised components to effectively leverage both pixel-level annotations and unlabeled structural information. The total loss function $\mathcal{L}_{total}$ is defined as:

$$\mathcal{L}_{total} = \mathcal{L}_{sup} + \lambda \mathcal{L}_{unsup} \tag{7}$$

Where $\mathcal{L}_{sup}$ denotes the supervised loss calculated on labeled data using a combination of Dice and Cross-Entropy losses. The term $\mathcal{L}_{unsup}$ represents the comprehensive unsupervised loss applied to unlabeled samples, which is further decomposed into three specialized components:

$$\mathcal{L}_{unsup} = \mathcal{L}_{con} + \alpha \mathcal{L}_{FRSB} + \beta \mathcal{L}_{SSC} \tag{8}$$

Here, λ, α, β are weighting coefficients to balance the contribution of each term.

3.5.1. Consistency Loss

The term $\mathcal{L}_{con}$ serves as the foundational consistency constraint in our semi-supervised framework. It enforces prediction invariance between the student network and the teacher network under different stochastic perturbations. By minimizing the discrepancy between the student's output and the teacher's generated pseudo-labels, $\mathcal{L}_{con}$ ensures that the model learns stable and noise-robust representations from the large-scale unlabeled BUS pool.

3.5.2. Fuzzy Rough Set Boundary Loss

The inherent blurriness of breast tumor boundaries in ultrasound images often leads to ambiguous pixel classifications. To explicitly model this uncertainty, we introduce the Fuzzy Rough Set Boundary (FRSB) loss. Drawing on fuzzy rough set theory, we define the boundary region as the difference between the fuzzy upper approximation *R*(c) and the fuzzy lower approximation *R*(c)

of a class c:

$$\mathcal{L}_{FRSB} = \sum_{i\in\Omega}\left|\left(\bar{R}(i,c) - R(i,c)\right)\cdot\psi\left(p_{i,c}, y_{i,c}\right)\right| \quad (9)$$

Where $p_{i,c}$ is the predicted probability and $y_{i,c}$ is the target (ground truth or pseudo-label). The FRSB loss assigns higher penalties to pixels located in the "vague" boundary zone, encouraging the model to reduce the entropy of the fuzzy approximation and produce sharper, more confident transitions between the lesion and the stroma.

3.5.3. Structural Similarity Contrastive Loss

While FRSB focuses on local edges, the Structural Similarity Contrastive (SSC) loss is designed to preserve the global topological integrity of the lesion. Standard contrastive losses often operate on individual pixels, ignoring the spatial correlation between them. Our SSC loss constructs sample pairs based on the structural similarity (SSIM) of mask patches.

For a predicted patch $\mathbf{m}_i$ and its corresponding target patch $\mathbf{g}_i$, we define the contrastive objective as:

$$\mathcal{L}_{SSC} = -\log\frac{\exp(\mathrm{SSIM}(\mathbf{m}_i,\mathbf{g}_i)/\tau)}{\exp(\mathrm{SSIM}(\mathbf{m}_i,\mathbf{g}_i)/\tau)+\sum_{j\in N^-}\exp(\mathrm{SSIM}(\mathbf{m}_i,\mathbf{g}_i)/\tau)} \quad (10)$$

where N^- represents negative patches (*e.g.*, healthy tissue or disparate image regions). By maximizing the structural similarity between the prediction and the true morphology while minimizing resemblance to non-target regions, the SSC loss ensures that the BIR stage maintains global shape constraints and avoids fragmented segmentations common in noisy BUS images.

4. Experiments

4.1. Datasets

BUSI dataset The Breast Ultrasound Images (BUSI) dataset is collected in 2018 at Al-Diwaniyah Teaching Hospital, Iraq, using a LOGIQ E9 ultrasound system. It contains 780 grayscale ultrasound images from 600 female patients, divided into three categories: 133 normal, 437 benign, and 210 malignant cases. The images are stored in PNG format with varying spatial resolutions, typically around 500 × 500 pixels (most commonly 500 × 500, though the exact dimensions may slightly vary). For benign and malignant cases, the dataset provides corresponding pixel-level ground truth segmentation masks, where lesion regions are manually annotated by medical experts. The images are grayscale (single-channel) with 8-bit depth, and lesion sizes, shapes, and boundary clarity vary significantly, reflecting real clinical variability [47].

UDIAT The UDIAT Breast Ultrasound Dataset is collected at the UDIAT Diagnostic Centre in Sabadell, Spain, and is designed for research on breast lesion analysis in ultrasound imaging. It contains 163 grayscale ultrasound images acquired from breast tumor cases, including both benign and malignant lesions. The images are stored in PNG format with a typical resolution of approximately 760 × 570 pixels (the resolution may slightly vary across samples). Each image is accompanied by expert-annotated pixel-level ground truth masks outlining the lesion boundaries, enabling segmentation, detection, and classification research [48].

4.2 Implementation Details

The proposed DMPB-Net is implemented using PyTorch 1.7.0 and Python 3.8.13, and all experiments are conducted on an NVIDIA GeForce RTX 3090 GPU. The training process follows a two-stage strategy—backbone warm-up and joint optimization with the boundary diffusion module—for a total of 100 epochs. We utilize the AdamW optimizer (weight decay of 1e-2) with an initial learning rate of 1×10^{-4}, which is adjusted via a CosineAnnealingLR scheduler. In the semi-supervised setting, the labeled-to-unlabeled ratio is maintained at 1:3, utilizing a momentum coefficient of $\gamma = 0.999$ for global prototype updates and T=1000 timesteps for the diffusion process. To enhance robustness against speckle noise and varying morphologies in the BUSI, BUS-UCLM, and UDIAT datasets, extensive data augmentations are applied, including random flipping, rotation (up to 30°), and elastic transformations, ensuring superior generalization across multi-center clinical data [17].

4.3 Comparison with state-of-the-arts

To evaluate the effectiveness of the proposed framework, we conduct comparative experiments on the BUSI dataset across three different annotation ratios: 1/2, 1/4, and 1/8. As shown in Table 1, DMPB-Net consistently achieves state-of-the-art performance and outperforms existing semi-supervised methods such as Hermes [49], SDC-Net [19], and CoBiNet [17] across all evaluation metrics. Specifically, with 50% labeled data (1/2), our method achieves a Dice score of 83.20% and an IoU of 71.74%, representing a significant improvement over the runner-up model, Hermes. Notably, under the most challenging scenario with only 1/8 labeled data, DMPB-Net maintains superior robustness with a Dice score of 77.34%, whereas other models exhibit more pronounced performance degradation. Furthermore, our method demonstrates lower standard deviations, indicating higher stability in handling the inherent noise and boundary ambiguity of breast ultrasound images. These results validate that the integration of Mamba-driven global modeling and generative boundary refinement effectively enhances the model's discriminative power and generalization under limited supervision.

The performance evaluation on the UDIAT dataset further demonstrates the superiority of the proposed framework across all labeling scenarios. As presented in Table 2, our method consistently achieves the highest scores in both Dice and IoU metrics compared to other advanced semi-supervised models, including Hermes and SDC-Net. Specifically, under the 1/2 partition protocol, our approach reaches a Dice of 91.51% and an IoU of 83.34%, setting a new performance benchmark for this dataset. Even when the labeled data is reduced to 1/8, our model maintains a high Dice score of 89.54%, which still outperforms most competing methods operating at much higher supervision levels. Additionally, the significantly lower standard deviations (*e.g.*, ±0.23 at 1/2 labeling) highlight the remarkable stability and reliability of our framework in accurately segmenting breast lesions within the UDIAT clinical environment.

Figure 3. illustrates the qualitative segmentation results of various models on the UDIAT and BUSI datasets, clearly demonstrating the superior performance of the proposed DMPB-Net in handling complex lesion boundaries. For highly infiltrative tumors with irregular morphologies in

the UDIAT dataset (rows 1-2) and low-contrast lesions with blurry interfaces in the BUSI dataset (rows 3-4), comparative methods such as CGDM [18] and CoBiNet often suffer from noticeable over-segmentation or under-segmentation, producing coarse and anatomically inaccurate contours. In contrast, by leveraging the Mamba-inspired encoder's global contextual modeling and the Boundary Iterative Refinement (BIR) mechanism, our method (Ours) effectively filters out speckle noise and acoustic artifacts. The resulting segmentation masks not only maintain high topological consistency with the ground truth but also exhibit sharper, more precise edges that better capture the subtle spiculate boundaries of breast tumors. This visual evidence corroborates the quantitative findings, validating that DMPB-Net provides a more robust and clinically reliable solution for breast lesion segmentation under limited supervision.

Table 1. Quantitative comparisons using different SOTA methods on the BUSI test sets.

Method	1/2			1/4			1/8		
	Dice (%)	IoU (%)	P-value	Dice (%)	IoU (%)	P-value	Dice (%)	IoU (%)	P-value
CGDM	77.84±0.84	63.97±0.67	0.018	75.86±0.68	61.84±0.64	0.019	73.78±0.74	58.49±0.74	0.013
CSC-PA	78.22±0.71	64.27±0.45	0.024	76.55±0.41	62.02±0.42	0.028	74.06±0.56	58.87±0.58	0.021
CoBiNet	78.44±0.36	64.54±0.27	0.031	76.80±0.44	62.35±0.39	0.032	74.57±0.31	59.47±0.22	0.038
SDC-Net	81.18±0.37	68.32±0.52	0.026	79.69±0.45	66.24±0.62	0.025	74.61±0.67	59.50±0.85	0.021
Hermes	81.65±0.28	69.14±0.34	0.051	79.95±0.38	66.87±0.35	0.049	75.02±0.48	59.84±0.41	0.037
Ours	83.20±0.21	71.74±0.25	-	81.15±0.24	68.54±0.31	-	77.34±0.28	60.24±0.32	-

Table 2. Quantitative comparisons using different SOTA methods on the UDIAT test sets.

Method	1/2			1/4			1/8		
	Dice (%)	IoU (%)	P-value	Dice (%)	IoU (%)	P-value	Dice (%)	IoU (%)	P-value
CGDM	87.84±0.91	78.74±0.84	0.028	86.10±0.98	76.28±0.98	0.026	85.10±0.91	74.10±0.76	0.017
CSC-PA	88.11±0.52	79.03±0.36	0.039	86.35±0.75	76.52±0.87	0.044	85.38±0.64	74.49±0.48	0.026
CoBiNet	88.78±0.51	79.86±0.40	0.042	87.39±0.29	77.66±0.21	0.038	85.95±0.44	75.39±0.35	0.048
SDC-Net	89.18±0.37	80.32±0.52	0.036	88.69±0.45	77.24±0.62	0.033	86.61±0.67	76.50±0.85	0.031
Hermes	89.65±0.28	81.14±0.34	0.048	88.95±0.38	77.87±0.35	0.053	87.02±0.48	76.84±0.41	0.057
Ours	91.51±0.23	83.34±0.29	-	90.77±0.31	79.61±0.38	-	89.54±0.34	77.35±0.28	-

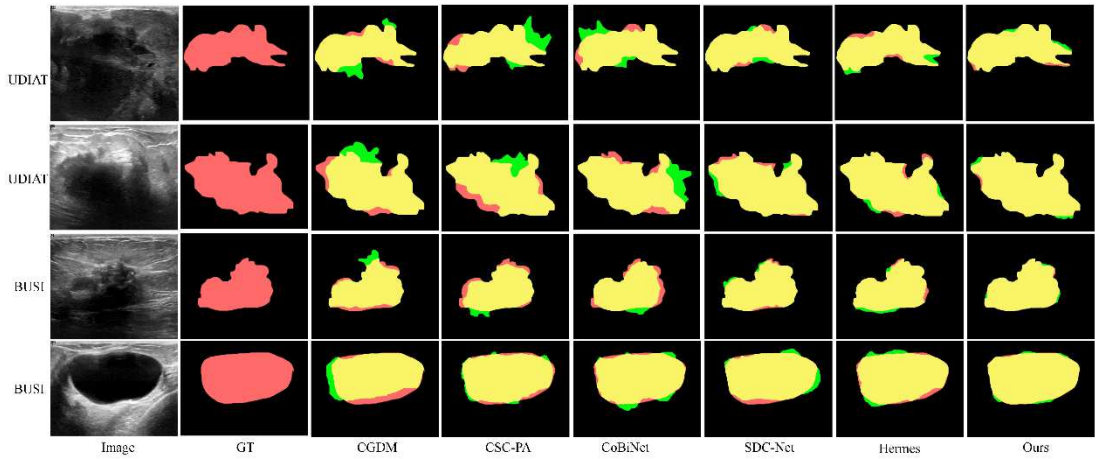


Figure 3. Visual comparison with different SOTA methods on two benchmark BUS test sets on 1/2 partitioning protocol. Red, green and yellow represent the ground truth, prediction and their overlapping regions, respectively.

4.4 Ablation Studies

To evaluate the individual contribution of each core component within DMPB-Net, we conduct comprehensive ablation studies on both the BUSI and UDIAT datasets. As summarized in Table 3, the "Suponly" baseline, which relies exclusively on labeled data, yields the lowest performance across all metrics. The integration of the Semantic Decoupling (SD) mechanism significantly boosts the Dice score on the BUSI dataset from 78.46% to 81.94%, validating its effectiveness in isolating lesion-specific features from complex sonographic backgrounds. Furthermore, the addition of the Confidence-Ranked Prototype Attention (CRPA) module facilitates robust semantic alignment between labeled and unlabeled data, leading to additional performance gains. The Boundary Iterative Refinement (BIR) stage also proves critical; when independently added or combined with other modules, it effectively refines fuzzy masks into sharp, anatomically accurate contours. Ultimately, the complete DMPB-Net achieves the highest Dice scores of 83.20% and 89.65% on BUSI and UDIAT, respectively, while exhibiting the lowest standard deviations. These results demonstrate that the synergy of SD, CRPA, and BIR provides a powerful and stable solution for breast ultrasound segmentation under limited supervision.

Table 3. Ablation studies on DMPB-Net. SD: Semantic Decoupling, CRPA: Confidence-Ranked Prototype Attention, BIR: Boundary Iterative Refinement.

Method	SD	CRPA	BIR	BUSI		UDIAT	
				Dice±std(%)	IoU±std(%)	Dice±std(%)	IoU±std(%)
Suponly				78.46±1.45	65.18±1.67	82.74±1.80	76.66±1.38
1	✓			81.94±0.92	69.94±1.14	88.15±1.05	80.12±0.97
2	✓	✓		82.75±0.61	70.47±0.52	88.86±0.61	80.61±0.68
3			✓	82.14±0.82	70.10±0.87	88.58±0.74	80.24±0.84
4		✓	✓	82.87±0.45	70.68±0.31	89.23±0.48	80.87±0.55
5	✓	✓	✓	83.20±0.21	71.74±0.25	89.65±0.28	81.14±0.34

Impact of Semantic Decoupling The Semantic Decoupling (SD) module plays a pivotal role in enhancing the model's discriminative capability by explicitly separating lesion-related features from complex sonographic backgrounds. In breast ultrasound imaging, pervasive artifacts such as speckle noise and acoustic shadows often overlap with the semantic representation of tumors, leading to false positives or ambiguous boundaries. By employing an orthogonal projection mechanism, the SD module decomposes the latent feature space into task-relevant and task-irrelevant components. As demonstrated in the ablation study (Table 3), the integration of SD leads to a substantial performance boost (*e.g.*, improving Dice from 78.46% to 81.94% on BUSI). This suggests that filtering out non-essential background interference allows the encoder to focus on the intrinsic morphology of the lesion, thereby significantly improving the sensitivity and robustness of the segmentation framework in low-contrast clinical scenarios.

Effectiveness of CR-PA Module The Confidence-Ranked Prototype Attention (CR-PA)

module is designed to bridge the semantic gap between labeled and unlabeled data in a semi-supervised setting. Traditional prototype-based methods often suffer from "prototype drift" caused by noisy pseudo-labels in unlabeled samples. To address this, our CR-PA module introduces a dynamic ranking mechanism that selects only high-confidence pixels to compute representative class prototypes. By performing cross-attention between these refined prototypes and the query features, the module enforces a strong semantic consistency across the entire dataset. As shown in the experimental results, the inclusion of CR-PA (Method 1 vs. Method 2 in Table 3) leads to a consistent improvement in segmentation accuracy. This demonstrates that the module effectively mitigates the bias introduced by limited supervision, ensuring that the model learns a robust and discriminative feature space even when high-quality manual annotations are scarce.

Figure 4. illustrates the qualitative comparison of prototype-based attention heatmaps. While Figure 4. (a) displays the original ultrasound image and the ground truth, Figure 4. (b) presents the attention maps generated by standard prototypes, which suffer from scattered activation and high sensitivity to background artifacts. In contrast, Figure 4. (c) demonstrates the heatmaps optimized by our proposed Confidence-Ranked Prototype Attention (CR-PA) module. By dynamically filtering high-confidence pixels to construct more representative prototypes, the model achieves a more concentrated and accurate response on the lesion area while effectively suppressing non-target noise. This visual evidence validates that the CR-PA module substantially enhances semantic consistency and feature discriminative power in semi-supervised scenarios, leading to more precise lesion localization aligned with the ground truth.

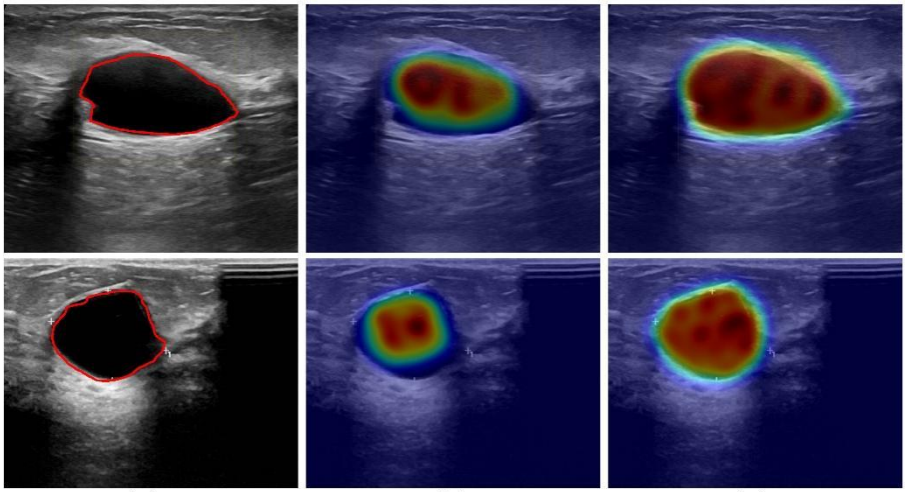


Figure 4. Visualization of CR-PA module. (a) Image and ground truth. (b) w/ standard prototypes. (c) w/ CR-PA.

Impact of Boundary Iterative Refinement To further validate the effectiveness of the Boundary Iterative Refinement (BIR) module, we analyze the progressive performance enhancement throughout the iterative process, as summarized in Table 4. Starting from the initial coarse mask (Step 0), which often lacks precise boundary details, the BIR module iteratively applies residual corrections to the uncertain transition zones. We observe a consistent upward trend in both

Dice and IoU metrics as the iteration count increases. Notably, the most substantial gain occurs during the first two steps, where the model effectively recovers the spiculate margins of the breast lesions. By the final step, the cumulative improvement in Dice and IoU reaches 1.85% and 3.50%, respectively. This evolution demonstrates that the BIR module functions as a powerful generative polisher, capable of transforming initial fuzzy predictions into high-quality, anatomically accurate segmentation results.

Table 4. Progressive Performance Evolution during the Boundary Iterative Refinement (BIR) Process.

Refinement Stage	Iteration Step	Dice (%)	Gain (Dice)
Initial Prediction	Step 0	81.35	-
BIR Progress	Step 1	82.54	1.19%
BIR Progress	Step 2	83.08	0.54%
Final Mask (Ours)	Step 3	83.20	0.12%

Ablation of Composite Loss Functions To rigorously evaluate the contribution of our multi-objective optimization strategy, we conduct an ablation study focusing on the individual and combined effects of the loss components. As presented in Table 5, the baseline model supervised by the standard cross-entropy loss and basic Consistency Loss (L_{con}) serves as the starting point. Integrating the Fuzzy Rough Set Boundary (FRSB) loss (L_{FRSB}) yields a significant performance boost, demonstrating its superior capability in modeling boundary uncertainty and suppressing noise in transition zones. Furthermore, the addition of the Structural Similarity Contrastive (SSC) loss (L_{SSC}) ensures global topological consistency, leading to the optimal performance of 83.20% Dice and 71.74% IoU. These quantitative results confirm that the synergy of L_{con}, L_{FRSB}, and L_{SSC}effectively compensates for the limitations of simple pixel-wise supervision, providing more robust guidance for segmenting breast lesions with complex morphologies.

Table 5. Performance Ablation of Composite Loss Functions on the BUSI Dataset.

Method	$\mathcal{L}_{con}$	$\mathcal{L}_{FRSB}$	$\mathcal{L}_{SSC}$	Dice (%)	IoU (%)	Gain (DIce)
Baseline	✓			79.45	66.21	-
Case 1	✓	✓		81.86	70.08	+2.41%
Case 2	✓		✓	80.92	68.54	+1.47%
Ours (Full)	✓	✓	✓	83.20	71.74	+3.75%

5 Conclusion

In this paper, we propose DMPB-Net, a novel semi-supervised framework specifically designed for breast ultrasound lesion segmentation under limited supervision. By integrating a Mamba-inspired encoder for global contextual modeling and a Semantic Decoupling (SD) module for artifact suppression, our method effectively overcomes the challenges of low contrast and complex background interference. Furthermore, the Confidence-Ranked Prototype Attention (CR-PA) and Boundary Iterative Refinement (BIR) modules facilitate robust semantic alignment and precise boundary delineation, respectively. Extensive experiments on the BUSI and UDIAT datasets demonstrate that DMPB-Net significantly outperforms state-of-the-art methods, achieving superior

segmentation accuracy and topological consistency. The results confirm that our framework provides a reliable and efficient solution for computer-aided breast cancer diagnosis, offering significant potential for reducing the workload of radiologists and improving clinical diagnostic precision.

Funding: This work was supported by Natural Science Foundation of the Jiangsu Higher Education Institutions of China（No.24KJD510005）, Start-up Fund for New Talented Researchers of Nanjing Vocational University of Industry Technology （No. YK24-05-02).

Declarations

Competing interests The authors declare no competing interests.

Data Availability Statement The datasets analyzed during the current study are publicly available repository datasets. The BUSI dataset can be accessed via https://scholar.cu.edu.eg/?q=afahmy/pages/dataset, and the UDIAT dataset can be obtained from http://www2.docm.mmu.ac.uk/STAFF/M.Yap/dataset.php.

Ethics Approval This study is conducted in strict accordance with the ethical standards of the Declaration of Helsinki. Formal ethics approval is not required for this study as it exclusively utilizes publicly available, fully anonymized, and de-identified breast ultrasound datasets (the BUSI dataset and the UDIAT dataset). The secondary use of these open-access benchmarks does not constitute human subjects research.

References

[1] Aumente-Maestro, Carlos, Jorge Díez, and Beatriz Remeseiro. "A multi-task framework for breast cancer segmentation and classification in ultrasound imaging." Computer methods and programs in biomedicine 260 (2025): 108540.

[2] Xu, Qiaoyi, et al. "A Review of Advanced Deep Learning Methods of Multi-target segmentation for breast cancer WSIs." IEEE Access (2025).

[3] Adepoju, Adekola George, and Daniel Adeyemi Adepoju. "Biomarker discovery in clinical biology enhances early disease detection, prognosis, and personalized treatment strategies." Int. J. Adv. Res. Publ. Rev 2 (2025): 229-252.

[4] Zhou, Jianmin, Yanchun Zhang, and Shaohua Shi. "Ultrasound elastography: advances and challenges in early detection of breast cancer." Frontiers in Oncology 15 (2025): 1589142.

[5] Szumiejko, Adrianna, Mariusz Ptak, and Bartosz Dołęga-Kozierowski. "Computational Methods in Breast Cancer Diagnostics and Surgery Planning: A Review." Archives of Computational Methods in Engineering (2025): 1-33.

[6]Chen, Xue. Influence of Predictions of an AI-based Computer-aided Diagnosis System on Radiologists in Reading Mammograms. Diss. Universitäts-und Stadtbibliothek Köln, 2025.

[7]Avwerosuoghene, Emmanuella. "Deep Learning-Based Computer Vision Models for Early Cancer Detection Using Multimodal Medical Imaging and Radiogenomic Integration Frameworks."

[8]Khosravi, Pegah, Thomas J. Fuchs, and David Joon Ho. "Artificial intelligence–driven cancer

diagnostics: enhancing radiology and pathology through reproducibility, explainability, and multimodality." Cancer Research 85.13 (2025): 2356-2367.
[9]Chen, Haohao, Ying Zeng, and De Cai. "Deep learning for cardiovascular management: optimizing pathways and cost control under diagnosis-related group models." Frontiers in Artificial Intelligence 8 (2025): 1580445.
[10]Das, Surajit, and Rajat Subhra Goswami. "Advancements in brain tumor analysis: a comprehensive review of machine learning, hybrid deep learning, and transfer learning approaches for MRI-based classification and segmentation." Multimedia Tools and Applications 84.23 (2025): 26645-26682.
[11]Birjais, Roshan. "Challenges and future directions for segmentation of medical images using deep learning models." Deep learning applications in medical image segmentation: overview, approaches, and challenges (2025): 243-264.
[12]Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
[13]Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Cham: Springer international publishing, 2015.
[14]Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation." IEEE transactions on pattern analysis and machine intelligence 39.12 (2017): 2481-2495.
[15]Noh, Hyeonwoo, Seunghoon Hong, and Bohyung Han. "Learning deconvolution network for semantic segmentation." Proceedings of the IEEE international conference on computer vision. 2015.
[16]Aniq, Elmehdi, Mohammed Jetti, and Mohamed Chakraoui. "Innovative breast cancer classification through semantic segmentation of ultrasound images: Advancing diagnostic precision and clinical utility." Biomedical Signal Processing and Control 122 (2026): 110371.
[17]Jiang, Siyao, et al. "Semi-Supervised Breast Lesion Segmentation Using Confidence-Ranked Features and Bi-Level Prototypes." IEEE Transactions on Neural Networks and Learning Systems (2025).
[18]Liu, Chunyuan, et al. "Semi-Supervised Medical Lesion Image Segmentation Based on a Contrast-Guided Diffusion Model." IEEE Transactions on Radiation and Plasma Medical Sciences (2025).
[19]Ma, Chaoqun, et al. "SDC-Net: A novel selective dilated convolution network for medical images segmentation." Computer Vision and Image Understanding (2025): 104526.
[20]Yang, Haonan, et al. "Semi-supervised breast ultrasound image segmentation via multi-paradigm synergy and cross-cyclic distillation." Expert Systems with Applications (2025): 130462.
[21]Wen, Renjun, et al. "Class-aware multi-level attention learning for semi-supervised breast

cancer diagnosis under imbalanced label distribution." Medical & Biological Engineering & Computing 63.7 (2025): 1997-2009.
[22]Ding, Zhenhui, et al. "CSC-PA: Cross-image Semantic Correlation via Prototype Attentions for Single-network Semi-supervised Breast Tumor Segmentation." Proceedings of the Computer Vision and Pattern Recognition Conference. 2025.
[23]Patro, Bachu Dushmanta Kumar. "Improved early detection accuracy for breast cancer using a deep learning framework in medical imaging." Computers in Biology and Medicine 187 (2025): 109751.
[24]Zhang, Wanting, et al. "RA-BUSSeg: Relation-aware Semi-supervised Breast Ultrasound Image Segmentation via Adjacent Propagation and Cross-layer Alignment." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2025.
[25]Dong, Zhiqi, et al. "A Mamba-optimized decoder for improved medical image segmentation." Machine Vision and Applications 37.1 (2026): 4.
[26]Liang, Pengchen, et al. "Mambasam: A visual mamba-adapted sam framework for medical image segmentation." IEEE journal of biomedical and health informatics (2025).
[27]Wu, Renkai, et al. "SK-VM++: Mamba assists skip-connections for medical image segmentation." Biomedical Signal Processing and Control 105 (2025): 107646.
[28]Li, Guangju, et al. "Selective and multi-scale fusion Mamba for medical image segmentation." Expert Systems with Applications 261 (2025): 125518.
[29]Zhong, Xin, Gehao Lu, and Hao Li. "Vision Mamba and xLSTM-UNet for medical image segmentation." Scientific reports 15.1 (2025): 8163.
[30]Liu, Tianyong, et al. "Mambavesselnet: a novel approach to blood vessel segmentation based on state-space models." IEEE Journal of Biomedical and Health Informatics 29.3 (2024): 2034-2047.
[31]Cheng, Zihan, et al. "Mamba-sea: A mamba-based framework with global-to-local sequence augmentation for generalizable medical image segmentation." IEEE Transactions on Medical Imaging (2025).
[32]Kui, Xiaoyan, et al. "Gl-MambaNet: A global-local hybrid Mamba network for medical image segmentation." Neurocomputing 626 (2025): 129580.
[33]Zhu, Shengqian, et al. "PRIME: Prototype-Driven Class Incremental Learning for Medical Image Segmentation." Proceedings of the 33rd ACM International Conference on Multimedia. 2025.
[34]Wu, Han, Chong Wang, and Zhiming Cui. "Dual cross-image semantic consistency with self-aware pseudo labeling for semi-supervised medical image segmentation." IEEE Transactions on Medical Imaging (2025).
[35]Zhang, Huaikun, et al. "Prototype-augmented mean teacher for robust semi-supervised medical image segmentation." Pattern Recognition 166 (2025): 111722.
[36]Zhang, Leyi, et al. "A Stability-Aware Dual-Head Network with Prototype-Based Consistency for Semi-Supervised Medical Image Segmentation." International Journal of Neural Systems 35.10

(2025): 2550054.
[37]Yue, Tianai, et al. "Dual prototypes contrastive learning based semi-supervised segmentation method for intelligent medical applications." Engineering Applications of Artificial Intelligence 154 (2025): 110905.
[38]Li, Lijian, Yuanpeng He, and Chi-Man Pun. "Efficient prototype consistency learning in semi-supervised medical image segmentation via joint uncertainty and data augmentation." 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2024.
[39]He, Shihuan, et al. "Prototype contrastive consistency learning for semi-supervised medical image segmentation." arXiv preprint arXiv:2502.06650 (2025).
[40]Lv, Zihe, Zhengda Wu, and Jinghua Zhu. "Clustering-guided contrastive prototype learning: Towards semi-supervised medical image segmentation." Pattern Recognition (2025): 112321.
[41]Wang, Huafeng, et al. "DPGNet: A Boundary-Aware Medical Image Segmentation Framework Via Uncertainty Perception." IEEE Journal of Biomedical and Health Informatics (2025).
[42]Zhang, Ruicheng, et al. "Marl-mambacontour: Unleashing multi-agent deep reinforcement learning for active contour optimization in medical image segmentation." Proceedings of the 33rd ACM International Conference on Multimedia. 2025.
[43]Yang, Yang, et al. "Boundary-guided contrastive learning for semi-supervised medical image segmentation." IEEE Transactions on Medical Imaging (2025).
[44]Zhang, Xingkuo, et al. "Multi-Scale Boundary Enhanced Pseudo Label Refinement for Semi-Supervised Medical Image Segmentation." 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). IEEE, 2025.
[45]Huang, Yigeng, et al. "Boundary feature alignment for semi-supervised medical image segmentation." Pattern Recognition 170 (2026): 111946.
[46]Hemalatha, K., and Shradhanjali Das. "Federated Learning model using CBDPL for Medical Image Segmentation." IEEE Access (2025).
[47]Yang, Haonan, et al. "Semi-supervised breast ultrasound image segmentation via multi-paradigm synergy and cross-cyclic distillation." Expert Systems with Applications (2025): 130462.
[48]Liu, Jiayi, et al. "BMFNet: A boundary-aware multi-scale fusion network for accurate breast tumor segmentation." Biomedical Signal Processing and Control 113 (2026): 108855.
[49]Zhang, Peng, Zhihui Lai, and Heng Kong. "Semi-Supervised Dual-Threshold Contrastive Learning for Ultrasound Image Classification and Segmentation." arXiv preprint arXiv:2508.02265 (2025).