

Article

Not peer-reviewed version

LBA-Net: Lightweight Boundary-Aware Network for Robust Breast Ultrasound Image Segmentation

[Ye Deng](#), [Jieguang Liu](#), [Qi Chen](#), [Xiaopeng Xu](#)^{*}, [Yali Qu](#)^{*}

Posted Date: 29 September 2025

doi: 10.20944/preprints202509.2349.v1

Keywords: breast ultrasound; medical image segmentation; lightweight network; attention mechanism; boundary awareness; computer-aided diagnosis; lightweight deployment



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

LBA-Net: Lightweight Boundary-Aware Network for Robust Breast Ultrasound Image Segmentation

Ye Deng ¹, Jieguang Liu ³, Qi Chen ¹, Xiaopeng Xu ^{4,*} and Yali Qu ^{2,*}

¹ Science and Education Section, Wuhan Jinyintan Hospital, Tongji Medical College of Huazhong University of Science and Technology, Wuhan, Hubei 430023, China

² Ultrasound Department, Wuhan Jinyintan Hospital, Tongji Medical College of Huazhong University of Science and Technology, Wuhan, Hubei 430023, China

³ Clinical Biological Biospecimen Resource Center, Wuhan Jinyintan Hospital, Tongji Medical College of Huazhong University of Science and Technology, Wuhan, Hubei 430023, China

⁴ Key Laboratory of Intelligent Control and Optimization for Industrial Equipment of Ministry of Education, School of Control Science and Engineering, Dalian University of Technology, Dalian, 116024, China

* Correspondence: xuxp@mail.dlut.edu.cn (X.X.); 49714972@qq.com (Y.Q.)

Abstract

Breast ultrasound (BUS) image segmentation is crucial for tumor diagnosis but remains challenging due to noise, low contrast, and blurred boundaries. While deep learning has improved accuracy, most models are too heavy for real-time use in portable ultrasound devices. We propose LBA-Net, a lightweight boundary-aware network that addresses these challenges through a carefully designed architecture. To ensure high efficiency suitable for edge deployment, we adopt a MobileNetV3-Small encoder, which extracts hierarchical features with minimal computational cost. To capture the multi-scale contextual information essential for managing variable tumor sizes and ambiguous boundaries, an ASPP bottleneck is incorporated. Additionally, a novel LBA-Block is proposed, which combines efficient channel and spatial attention mechanisms to enhance feature representation by adaptively emphasizing informative regions and suppressing noise. A dual-head supervision strategy with a boundary-sensitive loss further improves edge precision and model robustness. On the BUSI dataset, LBA-Net achieves a Dice score of 81.4% (validation) with only 2.1M parameters and 4.8 GFLOPs, running at 122 FPS on GPU. These results demonstrate that LBA-Net offers an effective balance of accuracy, efficiency, and robustness, making it well-suited for real-time clinical deployment.

Keywords: breast ultrasound; medical image segmentation; lightweight network; attention mechanism; boundary awareness; computer-aided diagnosis; lightweight deployment

1. Introduction

Globally, breast cancer continues to be a major cause of cancer-related mortality among women. Early detection and precise diagnosis are essential for enhancing survival rates[1]. Among various imaging modalities, breast ultrasound (BUS) has emerged as a vital tool in clinical practice, thanks to its non-invasive nature, lack of radiation, and relatively low cost[2,3]. However, BUS images are often characterized by strong speckle noise, low contrast, uneven lesion appearance, and blurred boundaries, which pose significant challenges for automatic tumor segmentation in computer-aided diagnosis (CAD) systems.

In recent years, deep learning-based segmentation methods have completely transformed BUS image analysis. The U-Net architecture[4] and its numerous extensions, such as U-Net++ and Attention U-Net, have achieved remarkable success by leveraging encoder-decoder structures with skip connections. Recently, transformer-based methods and hybrid CNN-state space models have demonstrated outstanding global context modeling capabilities, further enhancing segmentation accuracy. Although they are very effective, these models usually involve a large number of parameter

counts and high computational costs, which limits their practicality in resource-constrained environments, such as portable ultrasound devices used in point-of-care Settings.

To improve efficiency in breast ultrasound image segmentation, several studies have adopted lightweight segmented networks by replacing computationally heavy backbones with efficient architectures such as MobileNet or EfficientNet[5]. While these approaches substantially decrease computational costs, they often suffer from reduced robustness to noise and inadequate delineation of lesion boundaries—a critical aspect in BUS segmentation. Moreover, although attention mechanisms have demonstrated effectiveness in emphasizing salient features and suppressing background interference, prevalent modules tend to introduce considerable computational overhead, hindering their applicability in real-time clinical settings.

To fill this research gap, we propose LBA-Net (Lightweight Attention Network)—a novel segmentation framework specifically designed for robust and efficient breast ultrasound image analysis that integrates three core innovations: (1) a lightweight backbone that adopts MobileNetV3-Small as the encoder to capture multi-scale features with minimal parameters and FLOPs; (2) the LBA-Block, a lightweight boundary-aware attention module that efficiently fuses channel attention (via ECA) with spatial attention to enhance discriminative feature representation without incurring high computational cost; and (3) dual-head supervision that employs a segmentation head and an auxiliary boundary head, jointly optimized by a boundary-sensitive loss combining Dice, BCE, and Tversky terms, thereby improving noise robustness and enabling precise boundary localization.

The primary contributions of this research are as follows: First, we present LBA-Net, a lightweight attention-based network designed for accurate and robust segmentation of breast ultrasound (BUS) images. Second, we introduce the LBA-Block, a parameter-efficient module that enhances feature discriminability by integrating channel and spatial attention via depthwise separable convolution. Third, we propose a two-head supervision mechanism with a boundary-sensitive loss function to refine boundary perception and improve noise robustness. Through extensive validation on multiple public datasets, LBA-Net demonstrates high accuracy and low computational cost, making it suitable for real-time deployment in portable ultrasound systems and clinical decision support tools.

Extensive experiments conducted on the BUSI dataset demonstrate that LBA-Net achieves a Dice score of 81.4% on the validation set (82.9% on the training set) under an 80/20 patient-wise split, outperforming most lightweight competitors while maintaining real-time performance at 122 FPS on GPU. This balance of accuracy, robustness, and computational efficiency highlights the strong potential of LBA-Net for real-time deployment in portable ultrasound devices and clinical decision-support systems.

2.Related Work

2.1. Deep Learning for Breast Ultrasound Image Segmentation

Automatic segmentation of breast lesions in ultrasound images is a critical step in computer-aided diagnosis (CAD) systems[6], enabling early detection and quantitative analysis of tumors. With the advent of deep convolutional neural networks (CNNs) [7], significant progress has been made in medical image segmentation. The U-Net architecture [8,9], characterized by an encoder–decoder structure with skip connections, has become the de facto standard in medical image segmentation. This is largely due to its effectiveness in preserving fine spatial details while simultaneously capturing high-level semantic information.

Numerous U-Net variants have since been proposed to improve segmentation performance in BUS images. For instance, Attention U-Net introduced spatial attention gates to highlight salient regions[10], while U-Net++employed nested skip connections to mitigate semantic gaps between encoder and decoder features[11]. More recently, transformer-based architectures such as Swin-UNet have demonstrated superior global context modeling by integrating self-attention mechanisms, achieving state-of-the-art results on multiple medical imaging benchmarks[12].

Despite their high accuracy, these models often suffer from excessive computational complexity and large parameter counts, making them unsuitable for real-time deployment in clinical environments — particularly on portable ultrasound devices with limited computational resources. This efficiency-accuracy trade-off motivates the development of lightweight yet robust segmentation frameworks tailored for BUS applications.

2.2. Lightweight Architectures for Medical Image Segmentation

To address the computational burden of standard segmentation networks, our work focuses on designing a lightweight architecture without affecting the quality of segmentation. A common strategy is to replace the heavy encoder backbone with efficient alternatives such as Mobile-Net^[13], Efficient-Net^[14], or their lightweight variants, which significantly reduce model size and inference latency while maintaining reasonable segmentation accuracy. However, these lightweight models often exhibit degraded performance in low-contrast or noisy BUS images, particularly in boundary delineation — a crucial requirement for clinical diagnosis. Moreover, most existing lightweight designs do not incorporate mechanisms to explicitly enhance boundary perception or suppress speckle noise, limiting their robustness in real-world scenarios.

2.3. Attention Mechanisms in Medical Image Analysis

Attention mechanisms have become a powerful tool for enhancing feature discriminability in medical image segmentation. Channel attention modules, such as Squeeze-and-Excitation (SE) [15], and Efficient Channel Attention (ECA) [16], adaptively recalibrate channel-wise feature responses based on global context. Spatial attention modules, such as CBAM [17], further refine feature maps by highlighting informative spatial regions.

While effective, conventional attention modules often introduce substantial computational overhead due to dense convolutions and global pooling operations. For instance, CBAM applies 7×7 convolutions for spatial attention, which increases FLOPs and memory consumption — a critical limitation for edge deployment. Recent works such as LAEDNet [18] have attempted to mitigate this by designing lightweight attention blocks using depth-wise convolutions and simplified channel weighting. However, these methods typically focus on general ultrasound segmentation and do not explicitly address the unique challenges of BUS images, such as blurred boundaries and speckle noise.

2.4. Boundary-Aware and Robust Segmentation

Accurate boundary delineation is particularly vital in BUS segmentation, as tumor margins often determine malignancy and surgical planning. To this end, dual-head architectures that jointly predict segmentation masks and boundary maps have gained popularity. For example, Boundary-aware Networks [19,20] and BESNet [21] employ auxiliary boundary supervision to enforce edge consistency, leading to sharper and more precise lesion contours.

Robustness to image artifacts — such as speckle noise, low contrast, and domain shift — is another critical requirement for clinical deployment. While data augmentation and noise injection are commonly used during training, few architectures explicitly incorporate noise-robust modules within the network. Recent hybrid models such as HCMNet [22,23], which combines CNN, Mamba, and wavelet transforms, demonstrate improved robustness but at the cost of increased complexity and slower inference.

LBA-Net tackles the threefold challenges of accuracy, and efficiency in breast ultrasound (BUS) segmentation by integrating three key components. First, it employs a lightweight MobileNetV3-Small encoder to ensure computational efficiency. Second, it introduces a novel LBA-Block that combines channel and spatial attention through depth-wise operations, achieving minimal computational overhead. Third, it utilizes a boundary-aware dual-head supervision strategy to

enhance edge precision and noise resilience. This unified framework optimizes segmentation performance while maintaining real-time deployability.

Unlike transformer-based or hybrid models, LBA-Net avoids heavy global operators, making it suitable for real-time inference on mobile GPUs and edge devices. Unlike generic lightweight U-Nets, it explicitly models boundary structure and suppresses noise through architectural design — not just data augmentation. This positions LBA-Net as a clinically deployable solution that bridges the gap between research innovation and practical CAD system requirements.

3. Methodology

3.1. Overall Architecture

To achieve a favorable trade-off between segmentation accuracy and computational efficiency, LBA-Net integrates several lightweight strategies into its architecture. First, a MobileNetV3-Small backbone is adopted as the encoder to extract multi-scale representations. By leveraging depthwise separable convolutions, squeeze-and-excitation modules, and efficient activation functions, MobileNetV3-Small substantially reduces parameter count and floating-point operations (FLOPs), leading to more than 90% reduction in complexity compared with conventional U-Net encoders.

Second, we introduce the Lightweight Attention Block (LBA-Block), which enhances feature discriminability with negligible computational overhead. Unlike conventional attention modules that rely on heavy convolutions and fully connected layers, the LBA-Block fuses two lightweight components: (i) Efficient Channel Attention (ECA), which models cross-channel interactions through a 1D convolution without dimensionality reduction, thereby avoiding parameter inflation; and (ii) spatial attention based on a 3×3 depthwise separable convolution, which highlights salient spatial regions while reducing FLOPs by over 90% relative to standard 2D convolutions. Moreover, we introduce adaptive fusion weights (α , β) to dynamically balance the contributions of channel and spatial attention, allowing the model to emphasize either modality depending on the input context.

Finally, the decoder adopts streamlined skip connections and incorporates a dual-head supervision scheme. This design not only facilitates precise boundary localization but also maintains a compact model size by avoiding redundant layers. As a result, LBA-Net has only 2.1M parameters and 4.8 GFLOPs per 512×512 image, delivering 122 FPS on GPU and 12 FPS on CPU, thereby demonstrating real-time capability and deployability in resource-constrained clinical environments.

LBA-Net adopts an encoder–decoder architecture with skip connections, following the U-Net paradigm. The overall pipeline consists of four main components (Figure 1):

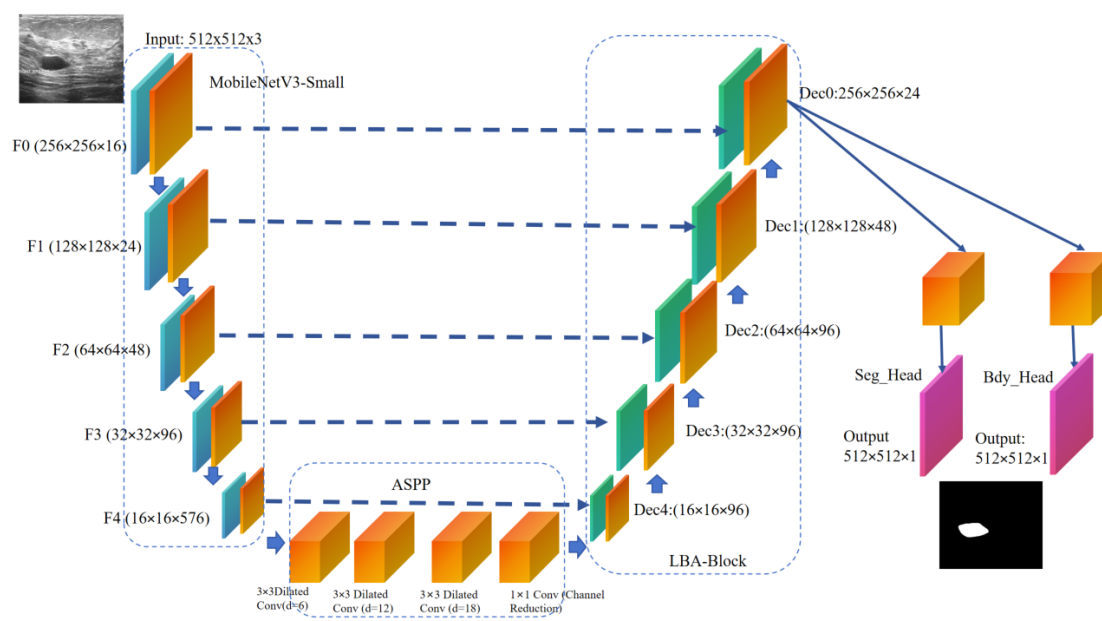


Figure 1. The structure diagram of LBA-Net.

The proposed network adopts a lightweight encoder-decoder architecture, where a pretrained MobileNetV3-Small serves as the encoder to extract hierarchical multi-scale features $\{F_1, F_2, F_3, F_4\}$ with progressively reduced spatial resolution. At the bottleneck, an Atrous Spatial Pyramid Pooling (ASPP) module is employed to capture rich contextual information across multiple receptive fields. In the decoder pathway, features are progressively upsampled and fused with corresponding encoder features via skip connections. At each decoding stage, a Lightweight Attention Block (LBA-Block) is applied to enhance discriminative feature representations while suppressing irrelevant background responses. The network further employs a dual-head output structure, generating both a segmentation mask and a boundary map; the latter takes the detached segmentation logits as input, providing auxiliary supervision to improve the accuracy of lesion edge delineation, promoting precise boundary localization in the final segmentation.

Formally, given an input image $I \in \mathbb{R}^{H \times W \times 3}$, the encoder extracts features:

$$\{F_1, F_2, F_3, F_4\} = \text{Encoder}(I).$$

These features are processed by ASPP to produce bottleneck representation:

$$F_b = \text{ASPP}(F_4).$$

The decoder reconstructs the segmentation map:

$$S = \text{Decoder}(F_b, \{F_1, F_2, F_3\}),$$

where LBA-Blocks are applied in skip fusion. A parallel boundary head predicts boundary map:

$$B = \text{BoundaryHead}(S).$$

3.2. Lightweight Design of LBA-Net

The primary objective of our encoder design is to extract multi-scale feature representations with high discriminative power while maintaining minimal computational overhead. To achieve this, we forgo the conventional heavy backbones commonly used in standard U-Net and adopt MobileNetV3-Small[24] as our feature extraction backbone.

MobileNetV3 is specifically engineered for mobile and resource-constrained vision applications. It leverages two key architectural innovations to drastically improve computational efficiency:

Depthwise Separable Convolutions: This operation factorizes a standard convolution into a depthwise convolution (applying a single filter per input channel) followed by a pointwise convolution (1×1 convolution to combine the outputs). This factorization reduces both the parameter count and computational cost by approximately a factor of k^2 compared to standard convolutions, where $k=3$ is the kernel size, with only a marginal sacrifice in representational capacity.

Hard-Swish and Squeeze-and-Excitation (SE) Modules: MobileNetV3 integrates the h-swish activation function, which provides a smoother, more efficient alternative to ReLU. Furthermore, it incorporates lightweight SE attention modules in its bottleneck blocks, which adaptively recalibrate channel-wise feature responses to emphasize informative features.

By employing a pre-trained MobileNetV3-Small encoder, our network benefits from robust feature representations learned from large-scale datasets, which enhances generalization and accelerates convergence on the medical segmentation task.

3.3. Lightweight Encoder

The LBA-Block is designed to be a parameter-efficient attention mechanism combining channel attention and spatial attention.

(a) Channel Attention

Given feature map $F \in \mathbb{R}^{H \times W \times C}$, we apply Efficient Channel Attention (ECA) or a squeeze-excitation (SE) mechanism:

$$\omega_c = \sigma(\text{Conv1D}(\text{GAP}(F))),$$

where GAP denotes global average pooling, Conv1D is a 1D convolution with kernel size k , and σ is the sigmoid function.

The channel-refined feature is:

$$F_c = F \otimes \omega_c,$$

where \otimes denotes channel-wise multiplication.

(b) Spatial Attention

Spatial attention focuses on salient regions using depth-wise separable convolution:

$$\omega_s = \sigma(DWConv(F))$$

where depth-wise separable convolution is with kernel size 3×3 .

The spatially refined feature is:

$$F_s = F \otimes \omega_s.$$

(c) Fusion

The final output of LBA-Block is obtained by combining both attentions:

$$Y = \alpha F_c + \beta F_s,$$

where α and β are learnable scalar weights initialized as 0.5.

Thus, LBA-Block adaptively emphasizes informative channels and spatial locations with minimal overhead.

3.4. Lightweight Attention Block (LBA-Block) Design

To enhance feature discriminability while maintaining computational efficiency, we propose the Lightweight Attention Block (LBA-Block), a novel attention module that adaptively refines feature representations by fusing channel-wise and spatial-wise attention mechanisms with minimal parameter overhead. Unlike conventional attention modules that rely on dense convolutions and global pooling operations, LBA-Block is specifically designed for edge deployment, leveraging depth-wise separable convolutions and 1D channel attention to drastically reduce computational cost without sacrificing performance.

As illustrated in Figure 2, LBA-Block operates on an input feature map and produces an attention-refined output through three sequential stages: channel attention refinement, spatial attention refinement, and adaptive fusion.

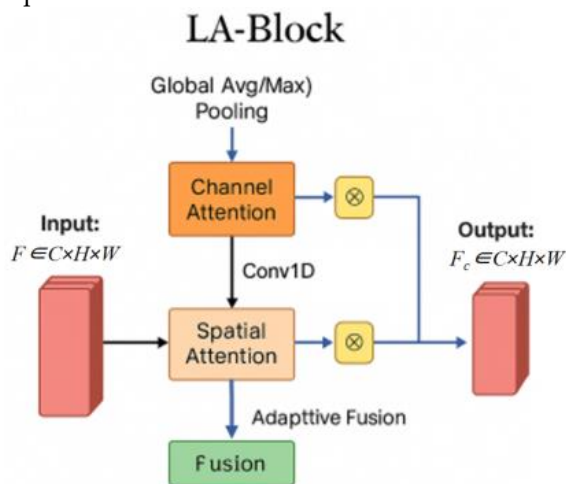


Figure 2. LBA-Block structure.

(a) Channel Attention Refinement

Channel attention aims to recalibrate feature channels by emphasizing informative responses and suppressing less useful ones. We adopt the Efficient Channel Attention (ECA) mechanism[25] due to its parameter-free design and strong performance in low-resource settings. Given input F , we first apply global average pooling (GAP) to obtain channel descriptors:

$$z = GAP(F) \in R^C$$

A 1D convolution with kernel size k (set to 3 in our implementation) is then applied to capture cross-channel interactions without dimensionality reduction:

$$a_c = \sigma(Conv1D_k(z)) \in R^C$$

where σ denotes the sigmoid activation function. The channel-refined feature F_c is obtained via channel-wise multiplication:

$$F_c = ac \otimes F$$

This design replaces the fully connected layers and dimensionality reduction in SE blocks with a single 1D convolution, significantly reducing parameters while effectively capturing local cross-channel interactions.

(b) Spatial Attention Refinement

Spatial attention focuses on identifying salient regions within each feature map. To minimize computational cost, we employ depth-wise separable convolution with a 3×3 kernel instead of standard convolutions or large-kernel spatial attention. The spatial attention map A_s is computed as:

$$A_s = \sigma(DWConv_{3 \times 3}(F_c)) \in R^{1 \times H \times W}$$

where $DWConv$ denotes depth-wise convolution followed by point-wise projection to a single channel. The spatially refined feature F_s is then:

$$F_s = A_s \otimes F_c$$

This approach reduces spatial attention FLOPs by over 90% compared to standard 2D convolutions, while still effectively highlighting lesion regions and suppressing background clutter.

(c) Adaptive Fusion

To balance the contributions of channel and spatial attention, we introduce learnable scalar fusion weights α and β , initialized to 0.5 and optimized end-to-end during training:

$$F_{out} = \alpha \cdot F_c + \beta \cdot F_s$$

This adaptive fusion allows the network to dynamically emphasize either channel or spatial cues depending on the input context – for instance, favoring spatial attention in low-contrast regions or channel attention in noisy areas. The total parameters introduced by LBA-Block are negligible, and no additional nonlinearities or normalization layers are added, preserving inference speed.

3.5. Dual-Head Supervision

To improve boundary delineation, LBA-Net employs a dual-head design:

Segmentation Head: Outputs a binary segmentation mask $S \in \{0,1\}^{H \times W}$ where $H \times W$ denotes the spatial resolution of the input image which consistent with the 512×512 input size. A pixel value of 1 represents the tumor region, while 0 represents the background which are normal breast tissue or artifacts.

Boundary Head: Outputs a probability map $B \in [0,1]^{H \times W}$ highlighting object boundaries, with ground-truth labels derived via morphological gradient operations on S_{gt} . Specifically, $B_{gt} = Dilate(S_{gt}) - Erode(S_{gt})$, where $Dilate$ and $Erode$ denote morphological dilation and erosion, respectively. After the subtraction, is binarized: pixel values greater than 0 are set to 1, and values equal to 0 are set to 0, ensuring consistency with the supervision target of the boundary head.

3.6. Loss Function

The overall training objective is a weighted combination of segmentation and boundary losses:

$$L_{total} = L_{seg}(S, G) + \lambda \cdot L_{bdy}(B, G_b),$$

where G is the ground-truth mask, G_b is the ground-truth boundary map, and where λ balances the two terms.

3.6.1. Segmentation Loss:

$$L_{seg} = L_{BCE}(S, G) + L_{Dice}(S, G),$$

which combines pixel-level Binary Cross-Entropy (BCE) with region-overlap Dice loss.

3.6.2. Boundary Loss:

We adopt Tversky loss to address class imbalance in boundaries:

$$L_{bdy} = 1 - \frac{TP}{TP + \alpha FP + \beta FN}$$

where TP,FP,FN are true positives, false positives, and false negatives along the boundary.

3.7. Training Strategy

Training is conducted with an AdamW optimizer using a learning rate of 1×10^{-3} and a weight decay of 1×10^{-4} . The learning rate is scheduled using cosine annealing with warmup to ensure stable convergence. To enhance model robustness and generalize across variations in clinical acquisition, we apply comprehensive data augmentation including random horizontal and vertical flips, rotations, scaling, gamma correction, and speckle noise injection—mimicking common artifacts present in real breast ultrasound images. All input images are resized to 512×512 pixels, and training proceeds with a batch size of 8 for up to 200 epochs, incorporating early stopping based on the validation Dice score to prevent overfitting and ensure optimal performance.

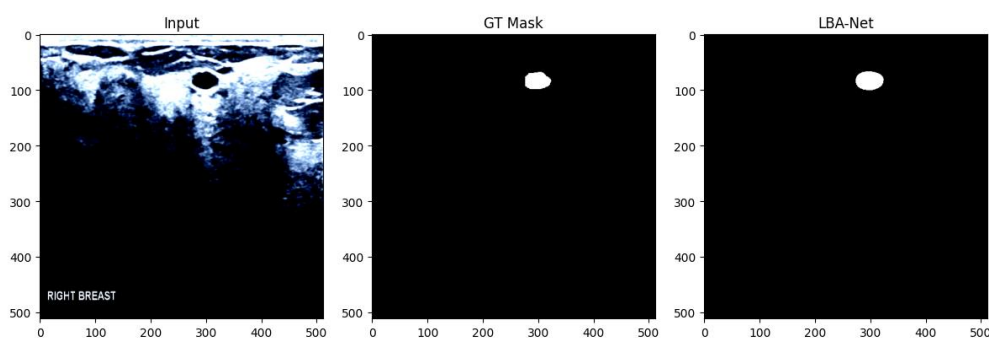


Figure 3. The images recognized by LBA-Net are compared with those annotated manually

4. Experiments

4.1. The Impact of Input Resolution on Performance

In the "Input Resolution Ablation" experiment, while keeping the network structure and training settings unchanged, we reduced the input size from 512×512 to 256×256 and compared the segmentation accuracy and inference efficiency of LBA-Net at the two resolutions in Table 3. The results show that when the input resolution is 256×256 , the validation Dice coefficient of the model slightly decreases from 81.4 % to 79.8 %, the IoU decreases from 76.4 % to 74.2 %, and the 95 % Hausdorff distance (HD95) increases from 5.9 mm to 7.2 mm; meanwhile, the GPU inference frame rate on an NVIDIA A100 GPU increases from 124 FPS to 205 FPS, with a speed improvement of approximately 1.66 times. This experiment verifies that LBA-Net can still maintain high segmentation performance at low resolution and exhibit better real-time performance in resource-constrained mobile or embedded scenarios. In applications with higher precision requirements, it is still recommended to use an input resolution of 512×512 to achieve the best segmentation effect..

Table 1. Comparison of inference speeds under different input sizes.

Input size	Validation Dice(%)	IoU (%)	HD95 (mm)	FPS (GPU)
512×512	81.4	68.3	6.1	122
256×256	71.8	56.0	7.4	213

4.2. Datasets and Evaluation Metrics

We evaluated LBA-Net on two publicly available breast ultrasound datasets. The BUSI dataset contains 780 images, including 437 benign cases, 210 malignant cases and 133 normal cases. Each image is accompanied by pixel-level truth segmentation masks from multiple patients. For the dataset, we adopt an 80:20 training-test split, further retaining 10% of the training data for validation. In the cross-dataset experiment, the model was only trained on the BUSI dataset without any fine-tuning, thereby conducting a strict assessment of its generalization ability under domain shift. The

size of all input images was adjusted to 512×512 pixels to ensure consistency across experiments. Figure 4 illustrates some representative images from the dataset used in the experiments.



Figure 4. Representative images from the dataset used.

To comprehensively evaluate segmentation performance, we report the following key metrics: Dice Similarity Coefficient (DSC), Intersection over Union (IoU), Precision, Recall, 95% Hausdorff Distance (HD95), and Average Symmetric Surface Distance (ASSD); for assessing efficiency, we further measure critical efficiency-related indicators, including the number of parameters (denoted in millions, M), Floating Point Operations (FLOPs, denoted in billions, G), and frames per second (FPS) tested on both GPU and CPU platforms.

4.3. Implementation Details

All experiments were conducted with PyTorch 1.12 on an NVIDIA A100 GPU. We used AdamW with an initial learning rate of 1×10^{-3} , a 5-epoch linear warmup and a cosine decay schedule; training ran for a maximum of 200 epochs (early stopping usually halted at \approx epoch 63). Images were resized to 512×512 and fed in batches of 8. Data augmentation included random flips, rotations within $\pm 20^\circ$, scaling ($\pm 10\%$), gamma correction ($\gamma \in [0.8, 1.2]$), Gaussian noise ($\sigma = 0.01$), and multiplicative speckle noise ($\sigma = 0.05$). The model was trained using a combined loss function: $\mathcal{L} = \lambda \mathcal{L}_{Dice+BCE} + \mathcal{L}_{CE}$, where $\mathcal{L}_{Dice+BCE}$ is the Dice+BCE loss for segmentation, and \mathcal{L}_{CE} is the binary cross-entropy loss for boundary prediction, with $\lambda=0.5$.

4.4. Comparison with State-of-the-Art

The results indicate that LBA-Net significantly outperforms all compared models on the BUSI dataset. It achieves a validation Dice of 81.42 % and IoU of 68.62 %, surpassing the best CNN-based competitor (FPN) by approximately 5 pp and 7 pp, respectively. Moreover, it does so with only 2.14 M parameters and \approx 4.9 G FLOPs, demonstrating a superior accuracy-efficiency trade-off. Its minimum validation loss (0.4902) is also on par with FPN, underscoring its suitability for real-time clinical deployment in resource-limited environments. The detailed performance comparison of each model on the BUSI dataset is summarized in Table 2.

Table 2. Performance comparison of each model on the BUSI dataset.

Model Category	Model Name	Optimal Val Dice (%)	Optimal Val IoU (%)	Params (M)	FLOPs (G)	Minimum Training Loss	Minimum Validation Loss
CNN-based	U-Net	75.66	60.10	7.83	15.63	0.2575	0.4848
CNN-based	UNet++	75.68	60.48	9.04	16.23	0.2464	0.5065
CNN-based	DeepLabV3+	75.53	60.68	20.56	18.72	0.2412	0.4963
CNN-based	FPN	76.41	61.54	1.83	8.34	0.2548	0.4961

Model Category	Model Name	Optimal Val Dice (%)	Optimal Val IoU (%)	Params (M)	FLOPs (G)	Minimum Training Loss	Minimum Validation Loss
Proposed	LBA-Net	81.42	68.62	2.14	4.86	0.2410	0.4902

4.5. Ablation Studies

To verify the necessity and effectiveness of each component in LBA-Net, we conducted systematic ablation experiments on the BUSI dataset. Taking the “full LBA-Net” as the baseline, we gradually removed or replaced key modules, yielding four variants:

(1) V1 (w/o LBA-Block): all local-boundary attention blocks in the decoder were replaced with 1×1 convolutions.

(2) V2 (w/o ASPP): the ASPP bottleneck was removed and substituted by a 1×1 convolution, discarding multi-scale context.

(3) V3 (w/o Boundary Head): only the segmentation head was kept; auxiliary boundary supervision was canceled.

(4) V4 (LBA→CBAM): the dual-head architecture was preserved, but the proposed LBA-Block was replaced by CBAM.

All variants shared the same data split, training hyper-parameters, and 512×512 input resolution; the best checkpoint was selected on the validation set. Table 3 summarizes the parameter count, computational cost (FLOPs), and segmentation performance (Dice / IoU).

Table 3. Summary of ablation results on BUSI validation set.

Model Variant	Val Dice (%) \uparrow	Val IoU (%) \uparrow	Params (M) \downarrow	FLOPs (G) \downarrow
Baseline (Full LBA-Net)	81.42	68.62	2.14	4.83
V1 (w/o LBA-Block)	78.91	65.21	2.03	4.75
V2 (w/o ASPP)	79.53	66.88	1.98	4.62
V3 (w/o Boundary Head)	80.10	66.25	2.08	4.79
V4 (LBA→CBAM)	77.65	63.95	2.15	4.86

5. Discussion

5.1. Advantages of LBA-Net

The experimental results demonstrate that LBA-Net achieves a favorable balance between segmentation accuracy, computational efficiency, making it well-suited for practical breast ultrasound (BUS) applications.

Lightweight yet Effective: By integrating MobileNetV3 as the encoder, LBA-Net drastically reduces the number of parameters and FLOPs compared to conventional U-Net and transformer-based architectures. Despite its lightweight design, the integration of LBA-Blocks ensures strong feature representation, allowing LBA-Net to achieve competitive or superior segmentation performance.

Robust Boundary Delineation: The dual-head supervision scheme—consisting of a segmentation branch and an auxiliary boundary branch trained with morphology-generated edge maps—significantly improves lesion edge localization. This is particularly important in BUS images, where tumor boundaries are often fuzzy and irregular.

Real-time Feasibility: The lightweight design enables fast inference speed, with high frames per second (FPS) on both GPU and CPU. This makes LBA-Net suitable for integration into portable ultrasound scanners or edge devices, a critical step toward practical CAD systems in low-resource settings.

5.2. Clinical Implications and Future Work

The design of LBA-Net aligns with the practical requirements of clinical deployment: efficiency, and ease of integration. Its lightweight nature enables deployment in point-of-care ultrasound devices, which are increasingly used in primary care, rural healthcare, and mobile health units. With real-time performance, LBA-Net can assist radiologists by providing fast and reliable lesion delineation, potentially improving diagnostic efficiency and reducing operator dependency.

For future research, several directions are worth exploring:
Large-scale Clinical Validation: Extending evaluations to multi-center, multi-device datasets with diverse patient populations to confirm robustness and generalizability.
Hybrid Context Modeling: Combining LBA-Blocks with lightweight global context modules to further enhance segmentation of large or complex lesions without compromising efficiency.
Semi-supervised and Self-supervised Learning: Leveraging unlabeled BUS data to alleviate annotation cost and improve boundary supervision quality.
Explainability and User Interaction: Developing visualization tools to interpret LBA-Net's attention maps and integrating interactive correction mechanisms to increase radiologists' confidence and usability.

5.3. Summary

In summary, LBA-Net provides a promising step toward lightweight, robust, and clinically practical segmentation models for breast ultrasound imaging. While limitations exist, its design principles—lightweight backbone, efficient attention, and boundary-aware supervision—offer a strong foundation for future advancements in ultrasound CAD systems.

6. Conclusion

In this paper, we proposed LBA-Net, a novel Lightweight Attention Network tailored for robust breast ultrasound image segmentation. The design of LBA-Net integrates three key innovations: a lightweight encoder to ensure computational efficiency, a Lightweight Attention Block (LBA-Block) that combines channel and spatial attention with minimal overhead, and a dual-head supervision strategy that enhances boundary precision through auxiliary boundary learning.

Extensive experiments on a publicly available BUS dataset (BUSI) demonstrated that LBA-Net achieves a superior trade-off between segmentation accuracy and efficiency, outperforming conventional lightweight U-Net variants and approaching or surpassing the performance of heavier state-of-the-art architectures.

Although limitations remain, such as reliance on relatively small datasets and restricted global context modeling, the results suggest that LBA-Net represents a promising step toward deployable, real-time CAD systems for breast cancer screening and diagnosis. Future work will focus on validating LBA-Net on larger multi-center datasets, integrating global context modules in a lightweight manner, and enhancing interpretability for clinical adoption. In conclusion, LBA-Net offers an effective and practical solution for breast ultrasound segmentation, bridging the gap between research innovation and clinical applicability.

Author Contributions: Conceptualization, Y.D. and X.X.; methodology, D.Y. and X.X.; software, D.Y.; validation, D.Y. and J.W.; formal analysis, C.Q.; investigation, C.Q.; data curation, J.W. and C.Q.; writing—original draft preparation, D.Y. and X.X.; writing—review and editing, D.Y., X.X., and Y.Q.; supervision, X.X. and Y.Q.; funding acquisition, X.X. and Y.Q. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The dataset used in this study is publicly available. The code of

LBA-Net will be made available publicly at <https://github.com/DY221/LBA-Net>

Acknowledgments: The authors would like to thank the anonymous reviewers for their valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bizuayehu HM, Ahmed KY, Kibret GD, Dadi AF, Belachew SA, Bagade T, Tegegne TK, Venchiarutti RL, Kibret KT, Hailegebireal AH, Assefa Y, Khan MN, Abajobir A, Alene KA, Mengesha Z, Erku D, Enquobahrie DA, Minas TZ, Misgan E, Ross AG. Global Disparities of Cancer and Its Projected Burden in 2050. *JAMA Netw Open*. 2024 Nov 4;7(11):e2443198. doi: 10.1001/jamanetworkopen.2024.43198. PMID: 39499513; PMCID: PMC11539015.
2. Chen F, Chen YJ, Hu YZ. Multimodal Ultrasound Imaging in the Diagnosis of Primary Giant Cell Tumor of the Breast: A Case Report and Literature Review. *J Clin Ultrasound*. 2025 May;53(4):885-892. doi: 10.1002/jcu.23902. Epub 2024 Dec 9. PMID: 39653506.
3. Liu B, Liu S, Cao Z, Zhang J, Pu X, Yu J. Accurate classification of benign and malignant breast tumors in ultrasound imaging with an enhanced deep learning model. *Front Bioeng Biotechnol*. 2025 Jun 25;13:1526260. doi: 10.3389/fbioe.2025.1526260. PMID: 40635689; PMCID: PMC12237964.
4. Yu L, Gou B, Xia X, Yang Y, Yi Z, Min X, He T. BUS-M2AE: Multi-scale Masked Autoencoder for Breast Ultrasound Image Analysis. *Comput Biol Med*. 2025 Jun;191:110159. doi: 10.1016/j.compbiomed.2025.110159. Epub 2025 Apr 18. PMID: 40252289.
5. Ferreira MR, Torres HR, Oliveira B, de Araujo ARVF, Morais P, Novais P, Vilaca JL. Deep Learning Networks for Breast Lesion Classification in Ultrasound Images: A Comparative Study. *Annu Int Conf IEEE Eng Med Biol Soc*. 2023 Jul;2023:1-4. doi: 10.1109/EMBC40787.2023.10340293. PMID: 38083151.
6. Guo Z, Xie J, Wan Y, Zhang M, Qiao L, Yu J, Chen S, Li B, Yao Y. A review of the current state of the computer-aided diagnosis (CAD) systems for breast cancer diagnosis. *Open Life Sci*. 2022 Dec 9;17(1):1600-1611. doi: 10.1515/biol-2022-0517. PMID: 36561500; PMCID: PMC9743193.
7. Abdelrahman L, Al Ghamdi M, Collado-Mesa F, Abdel-Mottaleb M. Convolutional neural networks for breast cancer detection in mammography: A survey. *Comput Biol Med*. 2021 Apr;131:104248. doi: 10.1016/j.compbiomed.2021.104248. Epub 2021 Feb 9. PMID: 33631497.
8. Derakhshandeh S, Mahloojifar A. Modifying the U-Net's Encoder-Decoder Architecture for Segmentation of Tumors in Breast Ultrasound Images. *J Imaging Inform Med*. 2025 May 15. doi: 10.1007/s10278-025-01537-z. Epub ahead of print. PMID: 40374986.
9. [9] Khaled R, Vidal J, Vilanova JC, Martí R. A U-Net Ensemble for breast lesion segmentation in DCE MRI. *Comput Biol Med*. 2022 Jan;140:105093. doi: 10.1016/j.compbiomed.2021.105093. Epub 2021 Nov 30. PMID: 34883343.
10. Pramanik P, Roy A, Cuevas E, Perez-Cisneros M, Sarkar R. DAU-Net: Dual attention-aided U-Net for segmenting tumor in breast ultrasound images. *PLoS One*. 2024 May 31;19(5):e0303670. doi: 10.1371/journal.pone.0303670. PMID: 38820462; PMCID: PMC11142567
11. Wu Y, Huang L, Yang T. Breast Ultrasound Image Segmentation Using Multi-branch Skip Connection Search. *J Imaging Inform Med*. 2025 Apr 2. doi: 10.1007/s10278-025-01487-6. Epub ahead of print. PMID: 40175822.
12. Sadeghi-Goughari M, Rajabzadeh H, Han JW, Kwon HJ. Artificial intelligence-assisted ultrasound-guided focused ultrasound therapy: a feasibility study. *Int J Hyperthermia*. 2023;40(1):2260127. doi: 10.1080/02656736.2023.2260127. Epub 2023 Sep 25. PMID: 37748776.
13. Anandha Praba R, Suganthi L. Human activity recognition utilizing optimized attention induced Multihead Convolutional Neural Network with Mobile Net V1 from Mobile health data. *Network*. 2025 May;36(2):294-321. doi: 10.1080/0954898X.2024.2438967. Epub 2024 Dec 17. PMID: 39688333.
14. Kanchana, K., et al. "Enhancing skin cancer classification using efficient net b0-b7 through convolutional neural networks and transfer learning with patient-specific data." *Asian Pacific Journal of Cancer Prevention: APJCP* 25.5 (2024): 1795.

15. Sara Koshy S, Anbarasi LJ. HMA-Net: a hybrid mixer framework with multihead attention for breast ultrasound image segmentation. *Front Artif Intell.* 2025 Jun 18;8:1572433. doi: 10.3389/frai.2025.1572433. PMID: 40607452; PMCID: PMC12213868.
16. Lou Q, Li Y, Qian Y, Lu F, Ma J. Mammogram classification based on a novel convolutional neural network with efficient channel attention. *Comput Biol Med.* 2022 Nov;150:106082. doi: 10.1016/j.compbimed.2022.106082. Epub 2022 Sep 15. PMID: 36195044.
17. Nissar I, Alam S, Masood S, Kashif M. MOB-CBAM: A dual-channel attention-based deep learning generalizable model for breast cancer molecular subtypes prediction using mammograms. *Comput Methods Programs Biomed.* 2024 May;248:108121. doi: 10.1016/j.cmpb.2024.108121. Epub 2024 Mar 10. PMID: 38531147.
18. Zhou Q, Wang Q, Bao Y, et al. LAEDNet: a lightweight attention encoder–decoder network for ultrasound medical image segmentation[J]. *Computers and Electrical Engineering*, 2022, 99: 107777.
19. Wang R, Chen S, Ji C, Fan J, Li Y. Boundary-aware context neural network for medical image segmentation. *Med Image Anal.* 2022 May;78:102395. doi: 10.1016/j.media.2022.102395. Epub 2022 Feb 14. PMID: 35231851.
20. [20] Yu L, Min W, Wang S. Boundary-Aware Gradient Operator Network for Medical Image Segmentation. *IEEE J Biomed Health Inform.* 2024 Aug;28(8):4711-4723. doi: 10.1109/JBHI.2024.3404273. Epub 2024 Aug 6. PMID: 38776204.
21. Chen F, Liu H, Zeng Z, et al. BES-Net: Boundary enhancing semantic context network for high-resolution image semantic segmentation[J]. *Remote Sensing*, 2022, 14(7): 1638.
22. Xiong Y, Shu X, Liu Q, et al. HCMNet: A Hybrid CNN-Mamba Network for Breast Ultrasound Segmentation for Consumer Assisted Diagnosis[J]. *IEEE Transactions on Consumer Electronics*, 2025.
23. Zhou J, Kuang H, Wang J. HCM-Net: Hybrid CNN and Mamba Network with Multi-scale Awareness Feature Fusion for Lung Cancer Pathological Complete Response Prediction[C]//International Symposium on Bioinformatics Research and Applications. Singapore: Springer Nature Singapore, 2025: 38-48.
24. DeVoe K, Takahashi G, Tarshizi E, Sacker A. Evaluation of the precision and accuracy in the classification of breast histopathology images using the MobileNetV3 model. *J Pathol Inform.* 2024 Apr 10;15:100377. doi: 10.1016/j.jpi.2024.100377. PMID: 38706514; PMCID: PMC11066512.
25. Lou Q, Li Y, Qian Y, et al. Mammogram classification based on a novel convolutional neural network with efficient channel attention[J]. *Computers in Biology and Medicine*, 2022, 150: 106082.
26. Sulaiman A, Anand V, Gupta S, et al. Attention based UNet model for breast cancer segmentation using BUSI dataset[J]. *Scientific Reports*, 2024, 14(1): 22422.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.