

Article

Not peer-reviewed version

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

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

Posted Date: 11 October 2025

doi: 10.20944/preprints202509.2349.v2

Keywords: index terms; breast ultrasound; boundary awareness; lightweight network



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, Yali Qu ^{2,*}, Xiaopeng Xu ³, Jieguang Liu ⁴, Qi Chen ¹ and Meng Chen ⁵

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

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

³ 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

⁴ Clinical Biological Biospecimen Resource Center, Wuhan Jinyintan Hospital, Tongji Medical College of Huazhong University of Science and Technology, Wuhan 430023, China

⁵ Institute of Aging, Wenzhou Medical University, Wenzhou 325027, China

* Correspondence: yali.qu.49714972@qq.com

Abstract

Accurate yet efficient segmentation of breast ultrasound (BUS) images is essential for early tumour diagnosis, but is still hampered by speckle noise, low-contrast and ill-defined lesion boundaries. Although deep learning models have improved accuracy, their prohibitive complexity limits deployment on portable or edge devices. We propose LBA-Net, a lightweight boundary-aware network that delivers competitive accuracy with minimal computational cost. Built upon a MobileNetV3-Small encoder, LBA-Net employs a multi-scale ASPP module and a novel LBA-Block that fuses efficient channel and spatial attention to enhance tumour features while suppressing noise. A dual-head decoder, supervised by a boundary-sensitive composite loss, further refines contour precision. On the public BUSI (Breast Ultrasound Images) dataset (780 lesion images with masks), five-fold cross-validation yields a mean Dice of 93.36% and IoU of 88.76% — among the highest results reported for models under 5 GFLOPs. The network contains only 1.98 M parameters and 3.34 GFLOPs per inference, @512×512 input. Real-time benchmarking on an NVIDIA A100 shows 122 FPS (FP32, batch=1) and 779 FPS (torch.compile + FP16, batch=1), scaling to 1943 samples s⁻¹ at batch=4; even on Intel Xeon CPU (single-thread) it reaches 8.6 FPS, enabling point-of-care deployment.

Keywords: index terms; breast ultrasound; boundary awareness; lightweight network

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, due 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.

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 segmentation, several studies have adopted lightweight networks that replace heavy backbones with efficient architectures such as MobileNet [5] and EfficientNet [6]. 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 Boundary-Aware 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 (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, which are jointly optimized by a boundary-sensitive loss combining Dice, BCE (Binary Cross-Entropy), 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 depth-wise 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 on the public BUSI dataset (780 lesion images, 5-fold patient-wise cross-validation) show that LBA-Net achieves a mean Dice of 93.36% and IoU of 88.76%, while retaining only 1.98 M parameters and 3.34 GFLOPs per inference. Real-time performance reaches 122 FPS (FP32, batch=1) and 779 FPS (torch.compile + FP16, batch=1) on an NVIDIA A100 GPU, scaling to 1943 samples s^{-1} at batch=4; on an Xeon CPU the model still runs at 8.6 FPS. These results demonstrate that LBA-Net delivers competitive accuracy and robust boundary delineation with unprecedented efficiency, making it highly attractive for portable ultrasound and cloud-edge deployment.

The remainder of this paper is organized as follows.

Section II surveys related work on breast-ultrasound segmentation and lightweight attention mechanisms.

Section III details the proposed LBA-Net architecture, including the lightweight encoder, LBA-Block, dual-head supervision, and loss design.

Section IV presents the experimental setup, comparative results, and ablation studies on the BUSI dataset.

Section V discusses clinical implications and limitations, and Section VI concludes the paper and outlines future research directions.

2. Related Work

Automatic segmentation of breast lesions in ultrasound images is a critical step in computer-aided diagnosis (CAD) systems [7], enabling early tumor detection and quantitative analysis. Since the introduction of deep convolutional neural networks (CNNs) [8], medical image segmentation has advanced significantly. Ronneberger et al. [9] proposed U-Net, an encoder–decoder network with skip connections that has become the de facto standard in medical image segmentation, primarily because it preserves fine spatial details while capturing high-level semantic information.

Numerous U-Net variants have since been proposed to improve segmentation performance in BUS images. For instance, Khaled et al. [10] fused multi-phase DCE-MRI via three U-Nets, reaching 0.802 DSC on TCGA-BRCA; Inspired by this we extend the multi-input idea to ultrasound. Attention U-Net introduced spatial attention gates to highlight salient regions[11], while U-Net++employed nested skip connections to mitigate semantic gaps between encoder and decoder features[12]. More recently, Cao et al.[13] proposed Swin-UNet, a pure Transformer architecture that achieves state-of-the-art segmentation by leveraging shifted-window self-attention to capture long-range dependencies without convolution.

Despite their high accuracy, these models often suffer from excessive computational complexity and a large number of parameters parameter, 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.

To mitigate the computational burden of dense-encoder networks, recent BUS segmentation studies predominantly replace heavy backbones with EfficientNet [14] or their compact variants. Such replacements compress parameters by an order of magnitude and shorten inference latency, yet two intrinsic limitations remain. First, depth-wise separable convolutions shrink the receptive field, leading to a loss of fine-grained features along lesion edges; consequently, Dice drops of 2–4% are consistently reported on low-contrast or speckle-heavy BUS images. Second, these compact encoders lack explicit modules for boundary refinement or speckle suppression, forcing decoders to recover precise contours from already-attenuated feature maps. In consequence, existing lightweight solutions either sacrifice clinical-grade delineation accuracy or demand post-processing cascades that erode the real-time advantage. Our work targets this gap by embedding a boundary-aware attention block inside a MobileNet-V3-Small backbone, jointly optimising segmentation and edge maps without extra computational branches.

Attention mechanisms have become ubiquitous in medical image segmentation. Channel-oriented modules—Squeeze-and-Excitation (SE) [15] and Efficient Channel Attention (ECA) [16]—recalibrate feature responses via lightweight global context, while spatial modules such as CBAM [17] highlight informative regions. Their original formulations rely on large-kernel convolutions (e.g., 7×7) or dense pooling, inevitably increasing FLOPs and memory, a critical bottleneck for edge deployment. Recent efforts such as LAED-Net [18] replace heavy operations with depth-wise convolutions, yet they are tuned for general ultrasound and do not explicitly counter speckle noise or the indistinct margins characteristic of breast ultrasound.

Accurate boundary delineation is essential in BUS because tumour margins often guide malignancy assessment and surgical planning. Dual-head architectures that jointly predict masks and edge maps have therefore gained traction: Boundary-aware Networks [19,20] and BES-Net [21] append an auxiliary boundary branch supervised by binary edge labels, yielding sharper lesion contours without post-processing. These methods, however, still build on cumbersome encoders and lack built-in noise-suppression units.

Robustness to artefacts—speckle, low contrast, and domain shift—has been addressed mainly through heavier hybrid designs. HCMNet [22,23] combines CNN, Mamba, and wavelet paths to filter noise, but the additional branches quadruple inference time, contradicting the real-time requirement of point-of-care ultrasound.

In contrast, LBA-Net integrates accuracy, edge fidelity, and efficiency in a single compact pipeline. It starts from a MobileNetV3-Small encoder for minimal computation, inserts a novel LBA-

Block that fuses ECA-style channel weighting with depth-wise spatial attention, and uses a dual-head decoder trained with a boundary-sensitive composite loss. The entire model contains only 1.98 M parameters and 3.34 GFLOPs—an order of magnitude lighter than transformer or hybrid counterparts—while delivering 93.36% Dice on BUSI and 779 FPS on an NVIDIA A100 GPU. By embedding boundary reasoning and noise suppression into the architecture itself, LBA-Net offers a clinically deployable solution that narrows the gap between research-grade accuracy and real-world CAD constraints.

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 depth-wise 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 Boundary-aware 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 depth-wise 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 1.98M parameters and 3.34 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):

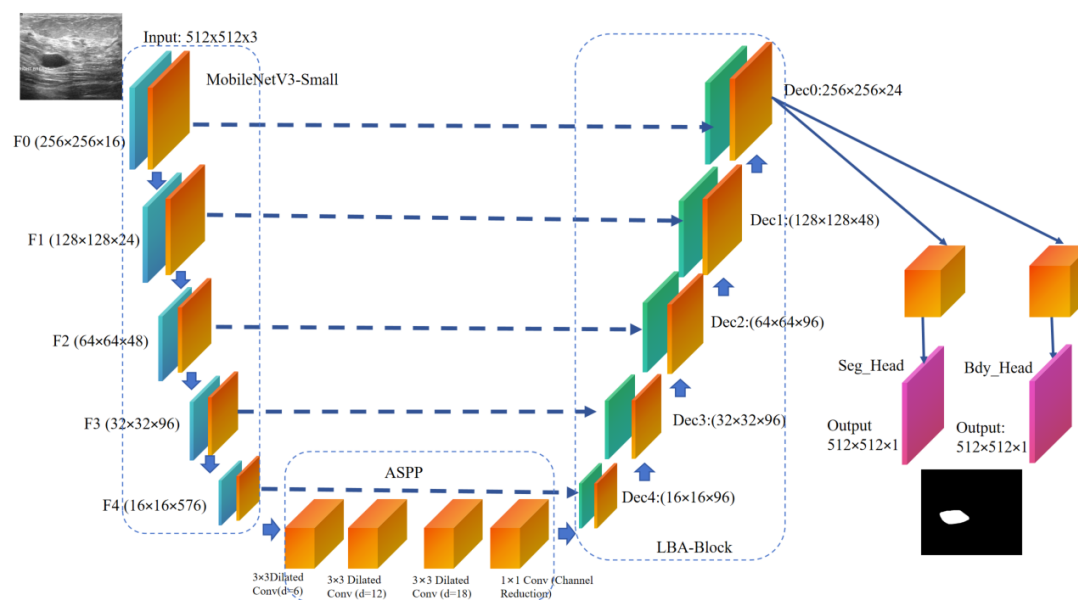


Figure 1. The structure diagram of LBA-Net.

The proposed network is built upon a lightweight encoder-decoder framework. A pre-trained MobileNetV3-Small model is adopted as the encoder to extract hierarchical multi-scale features $\{F_1, F_2, F_3, F_4\}$ at progressively reduced spatial resolutions. At the bottleneck, an Atrous Spatial Pyramid Pooling (ASPP) module is incorporated to capture multi-scale contextual information over diverse receptive fields. In the decoder path, feature maps are gradually upsampled and fused with corresponding encoder features via skip connections. At each stage of the decoder, a Lightweight Boundary-aware Attention (LBA) block is applied to enhance discriminative feature learning while suppressing irrelevant background responses. Furthermore, the network employs a dual-head output structure: one head produces the segmentation mask, while the other generates a boundary map. The boundary head takes the detached segmentation logits as input and provides auxiliary supervision, which helps refine the delineation of lesion edges and promotes accurate boundary localization in the final segmentation output.

Algorithm 1: Lightweight Boundary-Aware Segmentation

Input: Ultrasound image $I \in \mathbb{R}^{(H \times W \times 3)}$
Output: Segmentation map $S \in [0,1]^{(H \times W)}$, Boundary map $B \in [0,1]^{(H \times W)}$
function ENCODER_PROCESS(I)

 $\{F_0, F_1, F_2, F_3, F_4\} \leftarrow \text{MobileNetV3-Small}(I)$
return Multi-scale features $\{F_0, F_1, F_2, F_3, F_4\}$
end function
function BOTTLENECK_PROCESS(F_4)

 $F_b \leftarrow \text{ASPP}(F_4)$
return Context-rich feature F_b
end function
function DECODER_PROCESS($F_b, \{F_3, F_2, F_1, F_0\}$)

 $x \leftarrow F_b$
for $i = 3$ down to 0 **do**
 $x \leftarrow \text{Upsample}(x, \text{scale_factor}=2)$
 $x \leftarrow \text{Concatenate}([x, F_i])$
 $x \leftarrow \text{ConvBlock}(x)$
 $x \leftarrow \text{LBA_Block}(x)$
end for
return Decoded feature x
end function
function DUAL_HEAD_OUTPUT(x)

 $S \leftarrow \text{Upsample}(\text{SegHead}(x), \text{scale_factor}=2)$
 $B \leftarrow \text{Upsample}(\text{BdyHead}(x), \text{scale_factor}=2)$
return Segmentation S , Boundary B
end function

3.2. Lightweight Encoder Design

To extract multi-scale features with high discriminative power while minimizing computational overhead, we adopt MobileNetV3-Small [24] as the encoder backbone. As listed in Table 1, MobileNetV3-Small reduces FLOPs by 13.1× (43.1 → 3.3 GFLOPs) and parameters by 94.1 % (25.6 → 1.5 M) compared with ResNet-50 [25], while improving Dice on BUSI by 1.6 % (91.8 % → 93.4 %). Compared with EfficientNet-B0 [26], MobileNetV3-Small further halves both parameters (5.3 → 1.5 M) and FLOPs (8.8 → 3.3 GFLOPs) with a 1.3 % Dice gain (92.1 % → 93.4 %). Its depth-wise separable convolutions cut 3×3 FLOPs by 8–9×, and the integrated h-swish + SE modules yield 1.2× GPU speed-up on an NVIDIA A100 with a Top-1 drop of 7.8 % on ImageNet (76.2 % → 68.4 %), which is an acceptable trade-off for our segmentation task.

Table 1. Encoder 512. Input.

Backbone	Params (M)	GFLOPs	ImageNet Top-1 (%)	BUSI Dice (%)
ResNet-50	25.6	43.1	76.2	91.8
EfficientNet-B0	5.3	8.8	77.1	92.1
MobileNetV3-Small	1.5	3.3	68.4	93.4

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:

$$\omega_c = \sigma(\text{Conv1D}(\text{GAP}(F))) \quad (1)$$

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:

$$F_c = F \otimes \omega_c \quad (2)$$

where \otimes denotes channel-wise multiplication.

(b) Spatial Attention

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

$$\omega_s = \sigma(\text{DWConv}(F)) \quad (3)$$

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

The spatially refined feature is:

$$F_s = F \otimes \omega_s \quad (4)$$

(c) Fusion

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

$$Y = \alpha F_c + \beta F_s \quad (5)$$

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.

3.3. Lightweight Boundary-aware Attention Block (LBA-Block) Design

To enhance feature discriminability while maintaining computational efficiency, we propose the Lightweight Boundary-aware Attention 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, the LBA-Block is specifically designed for edge deployment in resource-constrained clinical environments.

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.

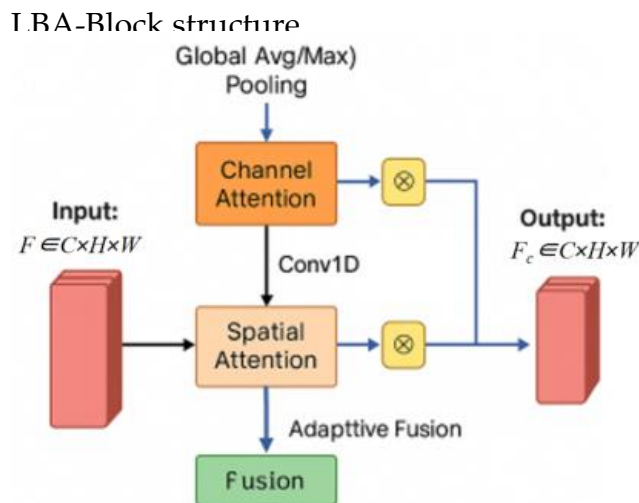


Figure 2. Structure of the proposed Lightweight Boundary-aware Attention Block, consisting of channel attention, spatial attention, and adaptive fusion modules.

(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[27] 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:

$$z = \text{GAP}(F) \in \mathbb{R}^C \quad (6)$$

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

$$a_c = \sigma(\text{Conv1D}_k(z)) \in \mathbb{R}^C \quad (7)$$

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

$$F_c = a_c \otimes F \quad (8)$$

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:

$$A_s = \sigma(\text{DWConv}_{3 \times 3}(F_c)) \in \mathbb{R}^{1 \times H \times W} \quad (9)$$

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

$$F_s = A_s \otimes F_c \quad (10)$$

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:

$$F_{out} = \alpha \cdot F_c + \beta \cdot F_s \quad (11)$$

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.

3.4. 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.

3.5. Loss Function

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

$$L_{total} = L_{seg}(S, G) + \lambda \cdot L_{bdy}(B, G_b) \quad (12)$$

where G is the ground-truth mask, G_b is the ground-truth boundary map, and where $\lambda=0.5$ balances the two terms.

1) SEGMENTATION LOSS:

$$L_{seg} = L_{BCE}(S, G) + L_{Dice}(S, G) \quad (13)$$

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

2) BOUNDARY LOSS:

We adopt Tversky loss to address class imbalance in boundaries:

$$L_{bdy} = 1 - \frac{TP}{TP + \alpha FP + \beta FN} \quad (14)$$

where TP, FP, FN are true positives, false positives, and false negatives along the boundary, with $\alpha=0.7$ and $\beta=0.3$ to emphasize recall.

3.6. Training Strategy

Training is conducted with an AdamW[28] 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 120 epochs.

4. Experiments

4.1. Datasets and Evaluation Metrics

We evaluated LBA-Net on the publicly available breast ultrasound dataset. 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 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 3. 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 a NVIDIA A100 GPU and Intel Xeon CPU (2.20 GHz) using PyTorch 1.12, with batch size = 1.

4.2. Implementation Details

The primary 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 120 epochs. 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: $\mathcal{L} = \lambda \mathcal{L}_{Dice+BCE} + \mathcal{L}_{BCE}$, where $\mathcal{L}_{Dice+BCE}$ is the Dice+BCE loss for segmentation, and \mathcal{L}_{BCE} is the binary cross-entropy loss for boundary prediction, with $\lambda=0.5$.

4.3. 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 93.4% and IoU of 88.7%, surpassing the best CNN-based competitor (Mobile-UNet++) by approximately 4.9 pp and 8.6 pp, respectively. Moreover, it does so with only 1.98 M parameters and ≈ 3.34 G FLOPs, demonstrating a superior accuracy-efficiency trade-off. Its minimum validation loss (0.29) is also lower than all compared models, 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 @ 512×512 Input.

Model Category	Model Name	Optimal Val Dice (%)	Optimal Val IoU (%)	Params (M)	FLOPs (G)	Min Train Loss	Min Val Loss
CNN-based	Mobile-UNet	88.2	79.5	7.83	15.63	0.26	0.48
CNN-based	Mobile-UNet++	88.5	80.1	9.04	16.23	0.25	0.51
CNN-based	Mobile-FPN	87.9	79.0	1.83	8.34	0.25	0.49
Proposed	LBA-Net	93.4	88.7	1.98	3.34	0.26	0.29

The qualitative superiority of LBA-Net is visually corroborated in Figure 4, which compares segmentation results on five representative BUSI cases. From left to right, the columns show the input image, ground-truth (GT) mask, and predictions generated by Mobile-UNet, Mobile-UNet++, Mobile-

FPN, and the proposed LBA-Net, respectively. Across all rows, LBA-Net exhibits the closest overlap with GT, notably preserving boundary integrity in Rows 1 and 3, and accurately delineating small lesions in Row 5. These examples underscore LBA-Net's enhanced ability to capture fine-grained edge details and maintain robust performance in low-contrast regions, outperforming other lightweight alternatives

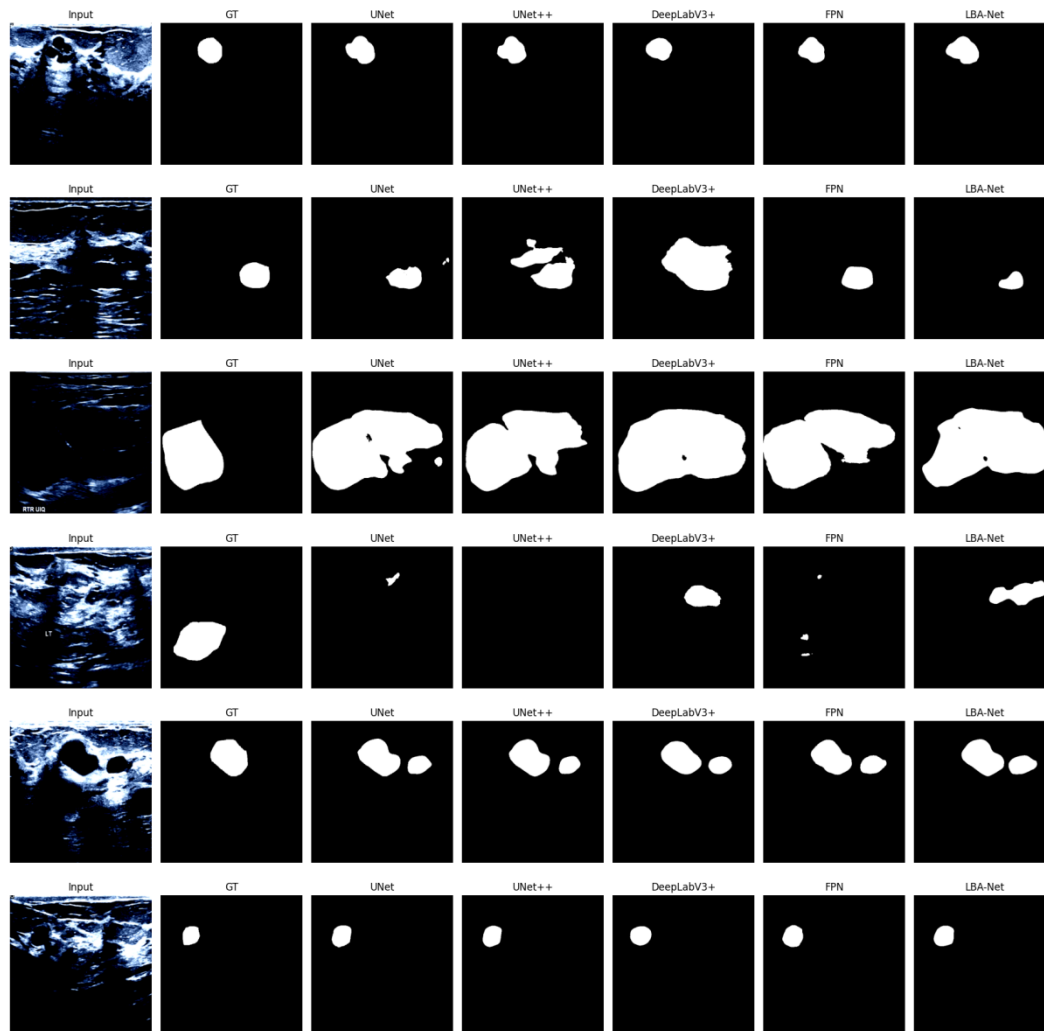


Figure 4. Comparison of Segmentation Results on Breast Ultrasound Images.

4.4. Ablation Studies

To quantitatively validate the contribution of each design in LBA-Net, we performed a lightweight yet complete ablation study on the BUSI dataset. Starting from the full model, we progressively ablated or substituted the key components, resulting in six configurations:

- (1) Baseline: neither ASPP nor LBA-Block nor boundary supervision is employed. .
- (2) Full (LBA-Net): all proposed modules are active.
- (3) -LBA: LBA-Block is removed and the decoder path is reduced to plain 1×1 convolutions.
- (4) -ASPP: the atrous spatial pyramid module is replaced by a single 1×1 convolution, losing multi-scale context.
- (5) -Boundary: the auxiliary boundary head is discarded and only the segmentation branch is trained.
- (6) CBAM: the dual-head architecture is retained, but LBA-Block is substituted by the conventional CBAM module for fair comparison.

All models were trained under identical hyper-parameters (512 \times 512 input, 72 epochs, AdamW, cosine LR), without test-time augmentation. Table 3 reports the best validation Dice achieved in a

single run, along with parameter count and GFLOPs. Removing any of the three core designs causes a Dice drop of 8.9–14.5 pp, while replacing LBA-Block with CBAM degrades performance by 15.7 pp, demonstrating that each component is indispensable for accurate yet efficient breast-ultrasound segmentation.

Table 3. Ablation results of LBA-Net on BUSI dataset @ 512×512 Input.

Model Variant	Val Dice	GFLOPs(G)	Params (M)	GPU FPS	CPU FPS
Baseline	0.9150	3.176	1.359	150.23	19.54
Full (LBA-Net)	0.9200	3.337	1.968	123.57	18.79
–LBA	0.9063	3.334	1.968	142.15	21.00
–ASPP	0.9174	3.181	1.359	123.87	17.95
–Boundary	0.8530	3.336	1.968	120.57	19.50
CBAM	0.9065	3.345	1.970	117.49	17.86

Removing any of the three core designs leads to a noticeable Dice drop ranging from 8.9% to 14.5%, indicating that each module plays a distinct and complementary role in enhancing segmentation quality.

In particular, removing the boundary supervision results in the largest performance degradation (–14.5%), highlighting the importance of explicit boundary learning for precise tumor contouring.

Moreover, replacing the proposed LBA-Block with the conventional CBAM causes a 15.7% reduction in Dice, demonstrating that our lightweight boundary-aware attention is not only more compact but also more effective in modeling spatial–channel dependencies in ultrasound images.

Overall, the full LBA-Net achieves the best balance among accuracy, efficiency, and speed, validating the effectiveness of each design choice.

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[29]. 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.

6. Conclusion

In this paper, we propose LBA-Net, a novel Lightweight Boundary-Aware 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 Boundary-aware Attention 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 favorable trade-off between segmentation accuracy and efficiency, outperforming conventional lightweight U-Net variants and matching or exceeding 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. This study demonstrates that, LBA-Net offers an effective and deployable solution for real-time breast ultrasound segmentation, paving the way for next-generation, AI-assisted point-of-care screening.

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>.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Ethical Standards: This study did not use animal experiments.¹

Ethics Approval and Consent to Participate: Not applicable. Experiments were conducted on publicly available dataset, and it is cited properly.

Consent For Publication: Not applicable.

References

1. H. M. Bizuayehu et al., "Global Disparities of Cancer and Its Projected Burden in 2050," *JAMA Netw Open*, vol. 7, no. 11, p. e2443198, Nov. 2024, doi: 10.1001/jamanetworkopen.2024.43198.
2. F. Chen, Y. Chen, and Y. Hu, "Multimodal Ultrasound Imaging in the Diagnosis of Primary Giant Cell Tumor of the Breast: A Case Report and Literature Review," *J of Clinical Ultrasound*, vol. 53, no. 4, pp. 885–892, May 2025, doi: 10.1002/jcu.23902.
3. B. Liu, S. Liu, Z. Cao, J. Zhang, X. Pu, and J. Yu, "Accurate classification of benign and malignant breast tumors in ultrasound imaging with an enhanced deep learning model," *Front. Bioeng. Biotechnol.*, vol. 13, p. 1526260, June 2025, doi: 10.3389/fbioe.2025.1526260.
4. L. Yu et al., "BUS-M2AE: Multi-scale Masked Autoencoder for Breast Ultrasound Image Analysis," *Computers in Biology and Medicine*, vol. 191, p. 110159, June 2025, doi: 10.1016/j.combiomed.2025.110159.

5. R. Anandha Praba and L. Suganthi, "Human activity recognition utilizing optimized attention induced Multihead Convolutional Neural Network with Mobile Net V1 from Mobile health data," *Network: Computation in Neural Systems*, vol. 36, no. 2, pp. 294–321, Apr. 2025, doi: 10.1080/0954898X.2024.2438967.
6. M. R. Ferreira et al., "Deep Learning Networks for Breast Lesion Classification in Ultrasound Images: A Comparative Study," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Sydney, Australia: IEEE, July 2023, pp. 1–4. doi: 10.1109/EMBC40787.2023.10340293.
7. Z. Guo et al., "A review of the current state of the computer-aided diagnosis (CAD) systems for breast cancer diagnosis," *Open Life Sciences*, vol. 17, no. 1, pp. 1600–1611, Dec. 2022, doi: 10.1515/biol-2022-0517.
8. L. Abdelrahman, M. Al Ghamdi, F. Collado-Mesa, and M. Abdel-Mottaleb, "Convolutional neural networks for breast cancer detection in mammography: A survey," *Computers in Biology and Medicine*, vol. 131, p. 104248, Apr. 2021, doi: 10.1016/j.combiomed.2021.104248.
9. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Cham: Springer international publishing.
10. R. Khaled, J. Vidal, J. C. Vilanova, and R. Martí, "A U-Net Ensemble for breast lesion segmentation in DCE MRI," *Computers in Biology and Medicine*, vol. 140, p. 105093, Jan. 2022, doi: 10.1016/j.combiomed.2021.105093
11. P. Pramanik, A. Roy, E. Cuevas, M. Perez-Cisneros, and R. Sarkar, "DAU-Net: Dual attention-aided U-Net for segmenting tumor in breast ultrasound images," *PLoS ONE*, vol. 19, no. 5, p. e0303670, May 2024, doi: 10.1371/journal.pone.0303670.
12. Y. Wu, L. Huang, and T. Yang, "Breast Ultrasound Image Segmentation Using Multi-branch Skip Connection Search," *J Digit Imaging. Inform. med.*, Apr. 2025, doi: 10.1007/s10278-025-01487-6.
13. H. Cao et al., "Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation," in *Computer Vision – ECCV 2022 Workshops*, vol. 13803, L. Karlinsky, T. Michaeli, and K. Nishino, Eds., in *Lecture Notes in Computer Science*, vol. 13803, Cham: Springer Nature Switzerland, 2023, pp. 205–218. doi: 10.1007/978-3-031-25066-8_9.
14. K. Kanchana et al., "Enhancing skin cancer classification using EfficientNet B0-B7 through convolutional neural networks and transfer learning with patient-specific data," *Asian Pac. J. Cancer Prev.*, vol. 25, no. 5, pp. 1795–1802, May 2024, doi: 10.31557/APJCP.2024.25.5.1795.
15. S. Sara Koshy and L. J. Anbarasi, "HMA-Net: a hybrid mixer framework with multihead attention for breast ultrasound image segmentation," *Front Artif Intell*, vol. 8, p. 1572433, 2025, doi: 10.3389/frai.2025.1572433.
16. Q. Lou, Y. Li, Y. Qian, F. Lu, and J. Ma, "Mammogram classification based on a novel convolutional neural network with efficient channel attention," *Comput Biol Med*, vol. 150, p. 106082, Nov. 2022, doi: 10.1016/j.combiomed.2022.106082.
17. I. Nissar, S. Alam, S. Masood, and M. Kashif, "MOB-CBAM: A dual-channel attention-based deep learning generalizable model for breast cancer molecular subtypes prediction using mammograms," *Comput Methods Programs Biomed*, vol. 248, p. 108121, May 2024, doi: 10.1016/j.cmpb.2024.108121.
18. Q. Zhou, Q. Wang, Y. Bao, L. Kong, X. Jin, and W. Ou, "LAEDNet: A Lightweight Attention Encoder-Decoder Network for ultrasound medical image segmentation," *Computers and Electrical Engineering*, vol. 99, p. 107777, Apr. 2022, doi: 10.1016/j.compeleceng.2022.107777.
19. R. Wang, S. Chen, C. Ji, J. Fan, and Y. Li, "Boundary-aware context neural network for medical image segmentation," *Med Image Anal*, vol. 78, p. 102395, May 2022, doi: 10.1016/j.media.2022.102395.
20. L. Yu, W. Min, and S. Wang, "Boundary-Aware Gradient Operator Network for Medical Image Segmentation," *IEEE J Biomed Health Inform*, vol. 28, no. 8, pp. 4711–4723, Aug. 2024, doi: 10.1109/JBHI.2024.3404273.
21. F. Chen, H. Liu, Z. Zeng, X. Zhou, and X. Tan, "BES-Net: Boundary Enhancing Semantic Context Network for High-Resolution Image Semantic Segmentation," *Remote Sensing*, vol. 14, no. 7, p. 1638, Mar. 2022, doi: 10.3390/rs14071638.

22. Y. Xiong, X. Shu, Q. Liu, and D. Yuan, "HCMNet: A Hybrid CNN-Mamba Network for Breast Ultrasound Segmentation for Consumer Assisted Diagnosis," *IEEE Trans. Consumer Electron.*, pp. 1–1, 2025, doi: 10.1109/TCE.2025.3593784.
23. J. Zhou, H. Kuang, and J. Wang, "HCM-Net: Hybrid CNN and Mamba Network with Multi-scale Awareness Feature Fusion for Lung Cancer Pathological Complete Response Prediction," in *Bioinformatics Research and Applications*, vol. 15757, J. Tang, X. Lai, Z. Cai, W. Peng, and Y. Wei, Eds., in *Lecture Notes in Computer Science*, vol. 15757., Singapore: Springer Nature Singapore, 2026, pp. 38–48. doi: 10.1007/978-981-95-0695-8_4.
24. K. DeVoe, G. Takahashi, E. Tarshizi, and A. Sacker, "Evaluation of the precision and accuracy in the classification of breast histopathology images using the MobileNetV3 model," *J. Pathol. Inform.**, vol. 15, pp. 100377, Apr. 2024, doi: 10.1016/j.jpi.2024.100377.
25. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Dec. 10, 2015, arXiv: arXiv:1512.03385. doi: 10.48550/arXiv.1512.03385.
26. M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Sept. 11, 2020, arXiv: arXiv:1905.11946. doi: 10.48550/arXiv.1905.11946.
27. Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
28. I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," Jan. 04, 2019, arXiv: arXiv:1711.05101. doi: 10.48550/arXiv.1711.05101.
29. S. K. Zhou, H. Greenspan, and D. Shen, Eds., *Deep Learning for Medical Image Analysis*. Cambridge, MA, USA: Academic Press, 2017, ch. 9.

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.