Boundary-Aware Attention U-Net for Retinal Blood Vessel Segmentation

Subhasish Bandyopadhyay

June 2025

Abstract

Accurate segmentation of retinal blood vessels is crucial for the diagnosis and monitoring of various ophthalmological and cardiovascular diseases. While U-Net architectures have shown promising results in medical image segmentation, they often struggle with preserving fine vessel boundaries and topology. We propose a boundary-aware attention U-Net that incorporates edge detection mechanisms within attention gates to enhance vessel boundary preservation. Our approach applies Sobel edge detection to high-resolution skip connections, allowing the model to focus on boundary-rich regions during feature fusion. Experimental evaluation on retinal vessel datasets demonstrates modest but consistent improvements over baseline U-Net, with a 2.9% increase in centerline Dice (clDice) and 2.7% improvement in vessel IoU, while maintaining computational efficiency.

Keywords: Retinal vessel segmentation, Attention mechanisms, Boundary detection, Medical image analysis, U-Net

1 Introduction

Retinal vessel segmentation is a fundamental task in automated ophthalmological diagnosis, enabling the detection and monitoring of diabetic retinopathy, hypertensive retinopathy, and cardiovascular diseases. The retinal vasculature provides a non-invasive window into systemic health, making accurate vessel delineation critical for clinical decision-making.

Traditional segmentation approaches often fail to preserve fine vessel structures and accurate boundaries, which are essential for reliable clinical analysis. U-Net architectures, while effective for medical image segmentation, face challenges in maintaining vessel topology and precise boundary delineation due to information loss during the encoding-decoding process.

This work addresses these limitations by introducing boundary-aware attention gates that incorporate edge detection mechanisms to enhance boundary preservation during feature fusion. Our contribution lies in the strategic application of edge detection to high-resolution skip connections, allowing the model to selectively emphasize boundary-rich features.

2 Related Work

2.1 U-Net Architectures for Medical Segmentation

U-Net has become the de facto standard for medical image segmentation due to its encoder-decoder architecture with skip connections that preserve spatial information. However, standard U-Net implementations often struggle with fine-grained boundary detection in vessel segmentation tasks.

2.2 Attention Mechanisms in Medical Imaging

Attention gates have been successfully applied to medical image segmentation to focus on relevant regions while suppressing irrelevant background features. These mechanisms typically use gating signals from the decoder path to modulate skip connections from the encoder.

2.3 Boundary-Aware Segmentation

Recent work in boundary-aware segmentation has shown that explicit boundary information can improve segmentation quality, particularly for structures with complex topology like blood vessels. However, most approaches require separate boundary detection networks or post-processing steps.

3 Methodology

3.1 Theoretical Foundation

3.1.1 Boundary Preservation in Vessel Segmentation

Retinal blood vessels exhibit several characteristics that make boundary preservation critical:

1. **Multi-scale Structure:** Vessels range from large arteries and veins to capillaries with widths of only a few pixels

2. **Complex Topology:** Bifurcations, crossings, and tortuous paths require precise boundary delineation

3. **Low Contrast:** Fine vessels often have minimal contrast against the background, making boundary detection challenging

The fundamental challenge in vessel segmentation lies in preserving spatial details while achieving semantic understanding. Standard U-Net skip connections provide spatial information but do not explicitly emphasize boundary regions.

3.1.2 Edge Detection for Boundary Enhancement

Edge detection identifies regions of rapid intensity change, which correspond to vessel boundaries. In the context of attention mechanisms, edge information can serve as a spatial prior to guide feature selection. We hypothesize that incorporating edge detection into attention gates will:

1. **Enhance Boundary Sensitivity:** Focus attention on regions with high gradient magnitude

2. **Preserve Fine Structures:** Maintain spatial details of thin vessels during feature fusion

3. **Improve Topology:** Better preserve vessel connectivity and branching patterns

3.1.3 Mathematical Formulation

Standard Attention Gate: Traditional attention gates compute attention coefficients $\alpha \in [0, 1]$ through a gating mechanism:

$$g_1 = W_g * g + b_g \tag{1}$$
$$x_1 = W_x * x + b_x \tag{2}$$
$$\psi_{\text{raw}} = \text{ReLU}(g_1 + x_1) \tag{3}$$
$$\psi = W_\psi * \psi_{\text{raw}} + b_\psi \tag{4}$$
$$\alpha = \sigma(\psi) \tag{5}$$
$$x_{\text{att}} = \alpha \odot x \tag{6}$$

where $g \in \mathbb{R}^{C_g \times H_g \times W_g}$ is the gating signal, $x \in \mathbb{R}^{C_x \times H_x \times W_x}$ is the skip connection, W_g, W_x, W_ψ are learnable convolution weights, and $\odot$ denotes element-wise multiplication.

Boundary-Aware Extension: Our boundary-aware attention gate incorporates edge information $E(x)$:

$$\alpha_{\text{ba}} = \sigma(\psi + \lambda \cdot E(x)) \tag{7}$$
$$x_{\text{att}} = \alpha_{\text{ba}} \odot x \tag{8}$$

where $\lambda \in \mathbb{R}^+$ is the edge weighting factor ($\lambda = 0.1$ in our implementation) and $E(x)$ is the normalized edge map.

Edge Detection Operator: The edge map $E(x)$ is computed using Sobel operators $S_x, S_y \in \mathbb{R}^{3\times3}$:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Given the grayscale conversion of feature maps:

$$x_{\text{gray}} = \frac{1}{C_x} \sum_{i=1}^{C_x} x_i \tag{9}$$
$$G_x = S_x * x_{\text{gray}} \tag{10}$$
$$G_y = S_y * x_{\text{gray}} \tag{11}$$
$$E_{\text{raw}}(x) = \sqrt{G_x^2 + G_y^2 + \epsilon} \tag{12}$$

where $*$ denotes convolution and $\epsilon = 10^{-6}$ for numerical stability.

Normalization: To ensure stable training, edge maps are normalized per batch:

$$E(x) = \frac{E_{\text{raw}}(x) - \min(E_{\text{raw}}(x))}{\max(E_{\text{raw}}(x)) - \min(E_{\text{raw}}(x)) + \epsilon} \tag{13}$$

3.2 Architecture Design

3.2.1 Boundary-Aware Attention Gate

Our attention gate implementation computes edge maps using Sobel operators:

1. **Grayscale Conversion:** Convert multi-channel features to single-channel by averaging
2. **Sobel Edge Detection:** Apply horizontal and vertical Sobel kernels
3. **Gradient Magnitude:** Compute edge strength as $\sqrt{G_x^2 + G_y^2}$
4. **Normalization:** Scale edge values to [0,1] range for stable training

3.2.2 Skip Connection Enhancement

We apply edge detection specifically to skip connections (high-resolution encoder features) rather than gating signals (low-resolution decoder features) because:

1. **Spatial Resolution:** Skip connections maintain original spatial resolution with fine boundary details
2. **Information Content:** Encoder features contain raw spatial information before semantic processing
3. **Computational Efficiency:** Edge detection on single pathway reduces computational overhead

3.2.3 Network Architecture

Our model follows the standard U-Net encoder-decoder structure with boundary-aware attention gates replacing standard skip connections:

- **Encoder:** Four encoding blocks with progressive downsampling (64, 128, 256, 512 channels)
- **Bottleneck:** 1024-channel processing layer
- **Decoder:** Four decoding blocks with boundary-aware attention gates
- **Output:** Single-channel vessel probability map

3.3 Training Configuration

3.3.1 Loss Function

We employ Focal Tversky Loss to address class imbalance:

$$\mathrm{FTL} = (1 - \mathrm{TI})^{\gamma}$$

$$\mathrm{TI} = \frac{\mathrm{TP} + \epsilon}{\mathrm{TP} + \alpha \cdot \mathrm{FP} + \beta \cdot \mathrm{FN} + \epsilon}$$

with parameters $\alpha = 0.7$, $\beta = 0.3$, $\gamma = 0.75$, designed to emphasize recall while maintaining precision balance.

3.3.2 Implementation Details

- **Input Resolution:** 512×512 pixels
- **Batch Size:** 4
- **Optimizer:** Adam with learning rate 10^{-4}
- **Normalization:** Instance normalization for attention gates, batch normalization for baseline
- **Training Epochs:** 50 with early stopping based on validation loss

4 Experimental Setup

4.1 Dataset

We evaluate our approach on a retinal vessel segmentation dataset containing training and test splits with corresponding ground truth annotations. Images are resized to 512×512 pixels and normalized to [0,1] range.

4.2 Evaluation Metrics

We employ comprehensive metrics covering different aspects of segmentation quality:

4.2.1 Primary Metrics

- **Centerline Dice (clDice):** Measures topology preservation using skeletonized vessels
- **Dice Coefficient:** Standard overlap metric for binary segmentation

4.2.2 Vessel-Specific Metrics

- **Vessel IoU:** Intersection over Union for vessel class
- **Precision/Recall:** Class-wise performance metrics

- **F1-Score:** Harmonic mean of precision and recall

4.2.3 Boundary Metrics

- **Boundary F1:** Evaluates boundary detection accuracy with spatial tolerance

4.2.4 Topology Metrics

- **Skeleton Recall:** Measures preservation of vessel centerlines
- **Spurious Branches:** Quantifies false positive vessel branches

4.2.5 Threshold-Independent Metrics

- **AUPRC:** Area under precision-recall curve for robust evaluation

4.3 Baseline Comparison

We compare against a standard U-Net implementation with identical training configuration, differing only in the attention mechanism. This ensures fair comparison and isolates the contribution of boundary-aware attention.

5 Results

5.1 Quantitative Results

Our boundary-aware attention U-Net achieves consistent improvements over the baseline across most metrics:

Table 1: Performance comparison of baseline U-Net and boundary-aware attention U-Net

Metric	Baseline U-Net	Boundary-Aware	Improvement
clDice	0.3852 ± 0.0646	0.3962 ± 0.0646	+2.9%
Dice Coefficient	0.7586 ± 0.0636	0.7702 ± 0.0717	+1.5%
Vessel IoU	0.6152 ± 0.0794	0.6318 ± 0.0958	+2.7%
Precision	0.8063 ± 0.0821	0.8241 ± 0.0975	+2.2%
Recall	0.7225 ± 0.0764	0.7275 ± 0.0709	+0.7%
Boundary F1	0.8949 ± 0.0439	0.9095 ± 0.0338	+1.6%
Spurious Branches	0.1218 ± 0.0784	0.1082 ± 0.0802	-11.2%

5.2 Topology Preservation

The boundary-aware model demonstrates improved topology preservation:

- **Skeleton Recall:** 0.8126 vs 0.8057 (+0.9%)
- **Component Count Error:** 17.8 vs 19.1 components (-6.8%)

- **Breaks per Image:** Similar performance (6.58 vs 6.55)

5.3 Trade-offs

While showing improvements in most metrics, the boundary-aware model exhibits slightly lower AUPRC (0.7848 vs 0.8058, -2.6%), suggesting potential sensitivity to threshold selection.

6 Discussion

6.1 Performance Analysis

Our results demonstrate that boundary-aware attention provides modest but consistent improvements in vessel segmentation. The 2.9% improvement in clDice indicates better topology preservation, while the 2.7% increase in vessel IoU suggests improved boundary delineation.

6.2 Mechanism Effectiveness

The reduction in spurious branches (-11.2%) validates our hypothesis that edge-enhanced attention helps distinguish true vessel boundaries from noise. The improved precision (+2.2%) further supports this, indicating fewer false positive detections.

6.3 Limitations

6.3.1 Magnitude of Improvements

The improvements, while consistent, are modest in magnitude (1-3% range). This suggests that boundary-aware attention provides incremental rather than transformative benefits.

6.3.2 Threshold Sensitivity

The decrease in AUPRC indicates potential threshold sensitivity, suggesting the model may optimize for specific operating points rather than providing robust threshold-independent performance.

6.3.3 Computational Overhead

While minimal, edge detection adds computational overhead to the attention mechanism. The benefits must be weighed against increased inference time.

6.4 Clinical Relevance

Despite modest improvements, the enhanced boundary preservation and reduced spurious branches have clinical relevance. Accurate vessel topology is crucial for:

- Vessel diameter measurements for cardiovascular risk assessment
- Bifurcation analysis for retinopathy grading
- Vessel tracking for longitudinal monitoring

6.5 Future Directions

Several directions could enhance the approach:

1. Multi-scale edge detection using different kernel sizes
2. Learnable edge operators replacing fixed Sobel kernels
3. Adaptive edge weighting based on local image characteristics
4. Integration with vessel-specific loss functions for topology preservation

7 Conclusion

We presented a boundary-aware attention U-Net that incorporates edge detection mechanisms to enhance retinal vessel segmentation. By applying Sobel edge detection to high-resolution skip connections, our approach achieves modest but consistent improvements over baseline U-Net across multiple evaluation metrics.

The key contributions are:

1. Theoretical justification for edge-enhanced attention in vessel segmentation
2. Practical implementation of boundary-aware attention gates
3. Comprehensive evaluation demonstrating consistent improvements in topology preservation and boundary delineation

While the improvements are incremental rather than transformative, they represent meaningful progress in preserving vessel topology and reducing false positive detections. The approach maintains computational efficiency while providing clinically relevant enhancements in boundary accuracy.

Future work should explore learnable edge detection mechanisms and multi-scale boundary enhancement to further improve vessel segmentation performance.