ISSN: 2583-7354

International Journal of
Emerging Knowledge Studies
Publisher's Home Page: https://www.ijeks.com/

Fully Open Access
Research Paper

Frequency and Boundary-Aware Polyp Segmentation Using Deep Learning

S.Kalyan[1*] K.Nitya Sri[2] B.Neelima[3] K. Sandeep Kumar[4] P. Shanmukha Rao[5]
[1]Aditya Institute of Technology and Management, Tekkali, Srikakulam,India.
[2]Aditya Institute of Technology and Management, Tekkali, Srikakulam,India.
[3]Aditya Institute of Technology and Management, Tekkali, Srikakulam,India.
[4]Aditya Institute of Technology and Management, Tekkali, Srikakulam,India.
[5]Aditya Institute of Technology and Management, Tekkali, Srikakulam,India.
DOI: https://doi.org/10.70333/ijeks-05-01-007
*Corresponding Author: kalyan826@gmail.com

Article Info: - **Received** : 12 February 2026 **Accepted** : 20 March 2026 **Published** : 30 March 2026

Abstract

Colorectal cancer remains one of the leading causes of cancer-related mortality worldwide, and early detection through colonoscopy is the most effective preventive strategy. Accurate segmentation of polyps from colonoscopy images is essential for computer-aided diagnosis, as it assists clinicians in identifying polyp size, shape, and boundaries for treatment planning. However, existing deep learning-based segmentation models often struggle to produce precise boundary delineation, particularly for small, flat, or irregularly shaped polyps. To address this limitation, this paper proposes a frequency and boundary-aware polyp segmentation framework based on the Polyp-PVT transformer architecture. The proposed approach introduces a multi-loss optimization strategy that combines Binary Cross-Entropy and Dice loss for accurate region segmentation, boundary loss for edge refinement, and frequency loss in the Fourier domain to preserve fine structural details. By integrating region, boundary, and frequency-domain supervision, the model produces sharper and more accurate segmentation boundaries without increasing model complexity or inference time. Experimental results on benchmark polyp segmentation datasets demonstrate that the proposed method improves the Dice coefficient and Intersection over Union compared to the baseline Polyp-PVT model. Qualitative results further confirm improved boundary precision, particularly for challenging polyp morphologies such as small and flat polyps. These results indicate that the proposed multi-loss framework effectively enhances boundary quality while maintaining computational efficiency, making it suitable for real-time clinical applications.

Keywords: *Polyp Segmentation, Deep Learning, Boundary Refinement, Frequency Domain, Transformer, Colonoscopy, Medical Image Analysis.*

1. Introduction

Colorectal cancer (CRC) is one of the most commonly diagnosed cancers and a leading cause of cancer-related mortality worldwide. Early detection and removal of precancerous polyps through colonoscopy significantly reduce the risk of colorectal cancer and improve patient survival rates (Sung et al., 2021). During colonoscopy, accurate identification and segmentation of polyps are essential for determining their size, shape, and boundary, which are critical factors in clinical diagnosis and treatment planning. However, manual polyp segmentation is time-consuming and highly dependent on the experience of clinicians, and small or flat polyps can be easily missed (Leufkens et al., 2012).

In recent years, deep learning techniques have achieved significant success in medical image segmentation. Fully Convolutional Networks (FCNs) and U-Net architectures have become the foundation for many medical image segmentation tasks due to their encoder–decoder structure and ability to capture both low-level and high-level features (Long et al., 2015; Ronneberger et al., 2015). Several improved versions, such as U-Net++, Attention U-Net, and ResUNet, have further enhanced segmentation performance by incorporating attention mechanisms and multi-scale feature extraction (Oktay et al., 2018; Zhou et al., 2020). These convolutional neural network (CNN)-based methods have shown promising results in polyp segmentation tasks.

More recently, transformer-based architectures have been introduced in medical image analysis to capture long-range dependencies and global contextual information. Vision transformers such as Swin Transformer and Pyramid Vision Transformer (PVT) have demonstrated strong performance in dense prediction tasks (Liu et al., 2021; Wang et al., 2021). The Polyp-PVT model adapts the Pyramid Vision Transformer specifically for polyp segmentation and achieves competitive performance on benchmark datasets (Dong et al., 2023). Although transformer-based models improve region detection, accurate boundary delineation remains a major challenge.

One of the main limitations of existing polyp segmentation models is the inability to accurately capture polyp boundaries, particularly for small, flat, or irregularly shaped polyps (Jha et al., 2020). Most segmentation models are optimized using region-based loss functions such as Binary Cross-Entropy and Dice loss, which focus on overall region accuracy but do not explicitly enforce boundary precision. As a result, predicted segmentation masks often contain blurred or inaccurate edges.

To address this issue, recent studies have explored boundary-aware learning methods that explicitly guide the model to focus on edge information and boundary refinement (Kervadec et al., 2021; Liu et al., 2024). In addition, frequency-domain methods have been introduced to preserve high-frequency components such as edges and fine details, which are often lost during convolution and pooling operations (Xu et al., 2024; Tang et al., 2025). These methods demonstrate that boundary information and frequency information are important for improving segmentation accuracy.

However, most existing approaches focus either on boundary refinement or frequency-domain enhancement separately, and many of them require additional network modules, which increase model complexity and computational cost. Therefore, there is a need for a method that can improve boundary accuracy and preserve fine details without increasing model complexity.

To address these challenges, this paper proposes a frequency and boundary-aware polyp segmentation framework based on the Polyp-PVT transformer architecture. Instead of modifying the network architecture, we introduce a multi-loss optimization strategy that combines Binary Cross-Entropy and Dice loss for region segmentation, boundary loss for edge refinement, and frequency loss in the Fourier domain to preserve fine structural details. The proposed method improves boundary delineation and segmentation accuracy while maintaining the same computational complexity as the baseline model.

The main contributions of this paper are summarized as follows:

- We propose a frequency and boundary-aware polyp segmentation framework based on the Polyp-PVT transformer architecture.
- We introduce a multi-loss optimization strategy that combines region loss, boundary loss, and frequency-domain loss.
- The proposed method improves boundary accuracy and fine detail preservation without increasing model complexity.

- Experimental results on benchmark polyp datasets demonstrate improved Dice and IoU scores compared to existing methods.

The remainder of this paper is organized as follows: Section 2 presents the literature review, Section 3 describes the problem statement and research objectives, Section 4 explains the proposed methodology, Section 5 presents the experimental setup, Section 6 discusses the results and analysis, and Section 7 concludes the paper.

2. Literature Review

2.1 Deep Learning for Medical Image Segmentation

Medical image segmentation has significantly improved with the development of deep learning techniques, particularly Convolutional Neural Networks (CNNs). Fully Convolutional Networks (FCNs) introduced end-to-end pixel-wise prediction for semantic segmentation and became the foundation for many medical image segmentation models **(Long et al., 2015)**. U-Net further improved segmentation performance by introducing an encoder–decoder architecture with skip connections, enabling precise localization and context capture, and it has become one of the most widely used models in medical image segmentation **(Ronneberger et al., 2015).**

Several improved variants of U-Net have been proposed to enhance segmentation accuracy. Attention U-Net introduced attention gates to focus on relevant regions and suppress irrelevant background features **(Oktay et al., 2018)**. U-Net++ redesigned skip connections to reduce the semantic gap between encoder and decoder features **(Zhou et al., 2020)**. ResUNet incorporated residual connections to improve feature propagation and network training stability **(Zhang et al., 2018).** These models demonstrated strong performance in medical image segmentation tasks, including polyp segmentation.

2.2 Transformer-Based Polyp Segmentation

Although CNN-based models perform well in capturing local features, they often struggle to model long-range dependencies and global context. Transformer-based architectures were introduced to address this limitation by using self-attention mechanisms. Vision Transformer (ViT) demonstrated that transformer architectures can be applied to image recognition tasks **(Dosovitskiy et al., 2020)**. Swin Transformer introduced hierarchical feature maps and shifted window attention, improving performance in dense prediction tasks **(Liu et al., 2021).**

The Pyramid Vision Transformer (PVT) further improved multi-scale feature representation and has been used as a backbone for segmentation tasks **(Wang et al., 2021).** Polyp-PVT adapted the PVT architecture specifically for polyp segmentation and achieved strong performance on benchmark colonoscopy datasets due to its ability to capture both global and local contextual information **(Dong et al., 2023).** Despite these improvements, transformer-based models still face challenges in accurately segmenting polyp boundaries.

2.3 Boundary-Aware Segmentation Methods

Accurate boundary delineation is a critical challenge in medical image segmentation, especially in polyp segmentation where boundaries can be unclear or irregular. Boundary-aware learning methods have been proposed to improve segmentation performance by explicitly focusing on edge information. Boundary loss was introduced to address class imbalance and improve boundary accuracy by minimizing the distance between predicted and ground truth boundaries **(Kervadec et al., 2021).**

Several boundary-aware polyp segmentation networks have been proposed in recent years. For example, boundary-aware networks and boundary-guided attention networks improve segmentation accuracy by incorporating boundary refinement modules and edge attention mechanisms **(Liu et al., 2024; Wu et al., 2025).** These methods demonstrate that incorporating boundary information significantly improves segmentation performance, particularly for small and irregularly shaped polyps.

2.4 Frequency Domain Methods in Image Segmentation

In addition to spatial domain features, frequency-domain information has been explored to improve segmentation performance. Frequency-domain methods help capture high-frequency components such as edges, textures, and fine structural details that are often lost during convolution and pooling operations. Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT) are commonly used for frequency-domain feature extraction (Xu et al., 2024).

Recent studies have shown that integrating frequency-domain information with spatial features improves segmentation accuracy and boundary quality. Frequency-guided networks and Fourier-based segmentation methods have demonstrated improved performance in medical image segmentation tasks by preserving fine details and enhancing edge information (Tang et al., 2025; Li et al., 2026). These methods highlight the importance of frequency information in improving boundary-aware segmentation.

2.5 Research Gap

From the literature, it is observed that CNN-based and transformer-based models have achieved significant improvements in polyp segmentation. Boundary-aware methods improve edge detection, while frequency-domain methods help preserve fine details. However, most existing studies focus on either boundary refinement or frequency-domain enhancement separately. Moreover, many boundary-aware models introduce additional network modules, increasing model complexity and computational cost.

Therefore, there is a need for a segmentation framework that can simultaneously improve boundary accuracy and preserve fine details without increasing model complexity. To address this research gap, this paper proposes a frequency and boundary-aware polyp segmentation framework that integrates region-based, boundary-based, and frequency-domain loss functions into a unified optimization framework using the Polyp-PVT transformer model.

3. Problem Statement and Objectives

3.1 Problem Statement

Polyp segmentation from colonoscopy images is a critical task in computer-aided diagnosis systems for early detection of colorectal cancer. Given an input colonoscopy image I∈RH×W×3I \in \mathbb{R}^{H \times W \times 3}I∈RH×W×3, the objective is to generate a pixel-wise binary segmentation mask M∈{0,1}H×WM \in {0,1}^{H \times W}M∈{0,1}H×W, where each pixel is classified as either polyp (foreground) or background.

Although existing deep learning models such as CNN-based and transformer-based architectures have achieved significant improvements in segmentation accuracy, accurate boundary delineation remains a major challenge. Many existing models are trained using region-based loss functions such as Binary Cross-Entropy (BCE) and Dice loss, which focus on overall region accuracy but do not explicitly optimize boundary quality. As a result, predicted segmentation masks often contain blurred, irregular, or inaccurate boundaries.

This problem becomes more significant in challenging scenarios such as:

- Small polyps with limited pixel information
- Flat polyps with low contrast and unclear edges
- Polyps with irregular shapes
- Images affected by noise, illumination variation, and specular reflections

Therefore, the main problem addressed in this research is the inaccurate boundary delineation and loss of fine structural details in polyp segmentation, which reduces segmentation accuracy and affects clinical diagnosis.

To address this problem, this research proposes a frequency and boundary-aware optimization framework that improves segmentation performance by incorporating region, boundary, and frequency-domain information into the training process.

3.2 Objectives of the Study

The main objective of this research is to develop a deep learning-based polyp segmentation framework that improves boundary accuracy and preserves fine details without increasing model complexity.

The specific objectives of this study are:

- To develop a polyp segmentation model based on the Polyp-PVT transformer architecture.
- To design a multi-loss optimization framework combining Binary Cross-Entropy loss and Dice loss for accurate region segmentation.
- To incorporate boundary loss to improve polyp boundary delineation.
- To introduce frequency-domain loss using Fast Fourier Transform (FFT) to preserve fine structural details.

- To evaluate the performance of the proposed model using standard evaluation metrics such as Dice coefficient and Intersection over Union (IoU).
- To compare the proposed method with existing polyp segmentation models.
- To analyze the contribution of boundary loss and frequency loss to segmentation performance through ablation study.

3.3 Research Questions

This research aims to answer the following research questions:

- **RQ1:** Can a multi-loss framework combining region, boundary, and frequency-domain loss improve polyp segmentation performance?
- **RQ2:** Does boundary loss improve the accuracy of polyp boundary delineation?
- **RQ3:** Does frequency-domain loss help preserve fine details in polyp segmentation?
- **RQ4:** How does the proposed method perform compared to existing deep learning-based polyp segmentation methods?
- **RQ5:** Does the proposed method improve segmentation performance without increasing model complexity and inference time?

4. Proposed Methodology

4.1 Overview of the Proposed Framework

This paper proposes a **frequency and boundary-aware polyp segmentation framework** based on the Polyp-PVT transformer architecture. The main objective of the proposed method is to improve segmentation accuracy, particularly boundary delineation and fine structural details, without increasing model complexity or computational cost.

The overall framework consists of three main components:

- Polyp-PVT transformer-based segmentation network
- Multi-loss optimization strategy (Region Loss)
- Boundary-aware loss
- Frequency-domain loss

Unlike many existing approaches that modify the network architecture to improve boundary accuracy, the proposed method improves segmentation performance by introducing additional supervision through a multi-loss function during training.

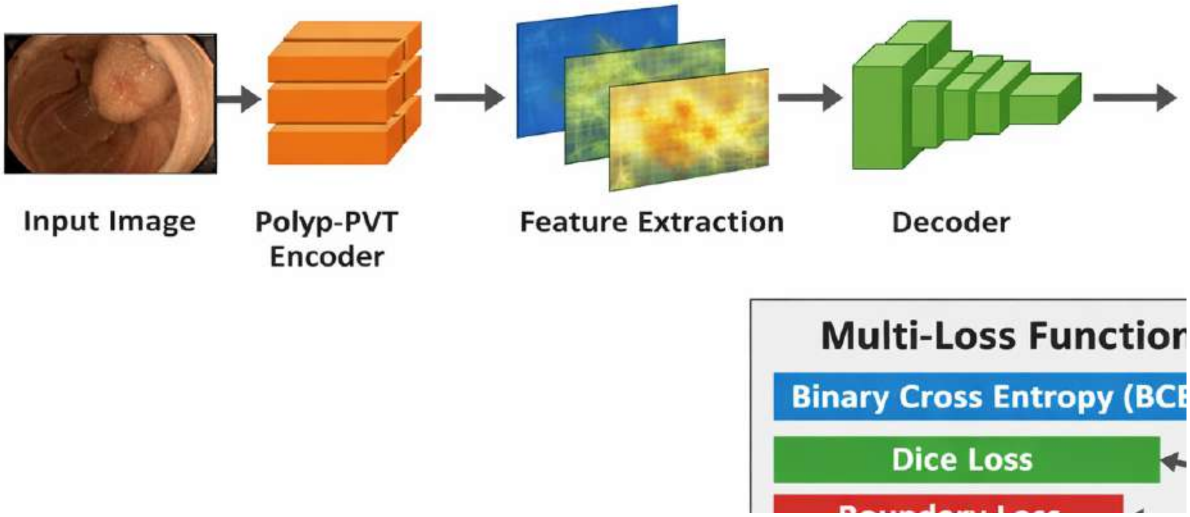

Fig-1: The overall workflow of the proposed method

4.2 Base Architecture: Polyp-PVT

The proposed method uses Polyp-PVT as the backbone segmentation network. Polyp-PVT is a transformer-based architecture that uses a Pyramid Vision Transformer (PVT) encoder to extract multi-scale feature representations from colonoscopy images. The encoder captures both local and global contextual information, which improves polyp detection performance.

The PVT encoder generates multi-scale feature maps at different resolutions, which are then passed to a lightweight decoder to generate the final segmentation mask. The decoder combines low-level spatial information and high-level semantic information to produce accurate segmentation output.

In this work, the Polyp-PVT architecture is used without modifying its structure in order to maintain computational efficiency. The main improvement is introduced through the loss function design rather than network architecture modification.

4.3 Multi-Loss Optimization Strategy

To improve segmentation accuracy and boundary quality, a multi-loss optimization strategy is employed in the proposed framework. The total loss function consists of three components: region loss, boundary loss, and frequency-domain loss. Each loss function is designed to address a specific limitation in polyp segmentation.

- **Region Loss (BCE + Dice Loss):** Ensures accurate segmentation of the polyp region.
- **Boundary Loss:** Improves segmentation accuracy at object boundaries.
- **Frequency Loss:** Preserves fine structural details and high-frequency information.

The total loss function is defined as:

$$L\_total = \lambda_1 \, L\_region + \lambda_2 \, L\_boundary + \lambda_3 \, L\_frequency$$

where **L_region** represents region-based loss, **L_boundary** represents boundary-aware loss, and **L_frequency** represents frequency-domain loss. The parameters λ_1, λ_2, and λ_3 are weighting coefficients that control the contribution of each loss component during training.

4.4 Region Loss (BCE + Dice Loss)

Region loss ensures accurate segmentation of the polyp region and handles the class imbalance between polyp and background pixels. The region loss is defined as the combination of Binary Cross-Entropy (BCE) loss and Dice loss.

The Binary Cross-Entropy loss is defined as:

$$L\_BCE = -(1/N) \, \Sigma \, [\, y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \,]$$

The Dice loss is defined as:

$$L\_Dice = 1 - (2 \, \Sigma \, p_i \, y_i + \varepsilon) \, / \, (\Sigma \, p_i + \Sigma \, y_i + \varepsilon)$$

The region loss is calculated as:

$$L\_region = L\_BCE + L\_Dice$$

This loss function improves segmentation accuracy by maximizing the overlap between the predicted segmentation mask and the ground truth mask.

4.5 Boundary Loss

Region-based loss functions focus on overall segmentation accuracy but do not explicitly optimize boundary precision. Therefore, boundary loss is introduced to improve segmentation accuracy near object edges. Boundary loss minimizes the distance between the predicted boundary and the ground truth boundary.

The boundary loss is defined as:

$$L\_boundary = \int \varphi\_G(q) \cdot s(q) \, dq$$

where $\varphi\_G(q)$ represents the distance map generated from the ground truth boundary, and $s(q)$ represents the predicted segmentation probability at pixel q. The boundary loss penalizes predictions that deviate from the true boundary, thereby producing sharper and more accurate segmentation edges.

4.6 Frequency Loss

To preserve fine structural details and high-frequency information such as edges and textures, a frequency-domain loss is introduced using Fast Fourier Transform (FFT). The frequency loss measures the difference between the frequency representation of the predicted mask and the ground truth mask.

The frequency loss is defined as:

$$L\_frequency = || \, F(M\_pred) - F(M\_gt) \, ||_2$$

where **F** represents the Fast Fourier Transform (FFT), **M_pred** is the predicted segmentation mask, and **M_gt** is the ground truth mask. This loss ensures that high-frequency components corresponding to edges and fine details are preserved in the predicted segmentation.

4.7 Final Loss Function

The final total loss function used to train the model is defined as:

$$L\_total = \lambda_1 \, (L\_BCE + L\_Dice) + \lambda_2 \, L\_boundary + \lambda_3 \, L\_frequency$$

In this study, the loss weights are empirically set as:

$$\lambda_1 = 1.0, \lambda_2 = 0.5, \lambda_3 = 0.3$$

This combined loss function enables the model to learn:

- Accurate polyp region segmentation
- Sharp and clear boundaries
- Fine structural details

By combining region-based, boundary-based, and frequency-domain supervision, the proposed method improves segmentation accuracy without increasing model complexity.

4.8 Algorithm Workflow

- **Step 1:** Input colonoscopy image
- **Step 2:** Extract multi-scale features using Polyp-PVT encoder
- **Step 3:** Generate segmentation mask using decoder
- **Step 4:** Compute Region Loss (BCE + Dice)
- **Step 5:** Compute Boundary Loss
- **Step 6:** Compute Frequency Loss using FFT
- **Step 7:** Compute Total Loss
- **Step 8:** Update model parameters using back propagation
- **Step 9:** Generate final segmentation output

5. Dataset Description and Preprocessing

5.1 Datasets

To evaluate the performance of the proposed frequency and boundary-aware polyp segmentation model, experiments were conducted on multiple publicly available colonoscopy image datasets that are widely used in polyp segmentation research. Using multiple datasets helps evaluate the robustness and generalization ability of the proposed model.

The datasets used in this study include:

- **Kvasir-SEG**: This dataset contains 1,000 polyp images with corresponding ground truth segmentation masks. The images include polyps of various sizes, shapes, and appearances, making it suitable for training deep learning models.
- **CVC-ClinicDB**: This dataset contains 612 colonoscopy images with high-quality segmentation annotations. It includes many small polyps and is commonly used for performance evaluation.
- **CVC-ColonDB**: This dataset contains 380 colonoscopy images with challenging cases, including polyps with unclear boundaries and low contrast.
- **ETIS-Larib Polyp DB**: This dataset contains 196 images with small polyps and difficult segmentation scenarios.
- **CVC-300**: This dataset contains 60 images and is often used for cross-dataset testing.

These datasets include polyps with different characteristics such as small size, irregular shape, low contrast, and complex background, making them suitable for evaluating segmentation performance under challenging conditions.

5.2 Data Preprocessing

Before training, all images undergo several preprocessing steps to ensure consistent input to the deep learning model and to improve model performance.

5.2.1 Image Resizing

All images are resized to **352 × 352 pixels**, which matches the input size required by the Polyp-PVT model. Resizing ensures uniform input dimensions and reduces computational cost.

5.2.2 Normalization

Pixel values are normalized to the range **[0, 1]** using min–max normalization. Normalization helps stabilize the training process and improves convergence speed.

5.2.3 Data Augmentation

To improve model generalization and prevent overfitting, data augmentation techniques are applied during training. The augmentation methods include:

- Random horizontal flipping
- Random rotation (±10 degrees)
- Random brightness adjustment
- Random contrast adjustment
- Random scaling and zooming

Data augmentation increases the diversity of training data and helps the model perform better on unseen images.

5.3 Dataset Splitting

The datasets are divided into training, validation, and testing sets. In this study, the dataset is split as follows:

- **Training set**: 80%
- **Validation set**: 10%

- **Testing set**: 10%

For cross-dataset evaluation, the model is trained on **Kvasir-SEG and CVC-ClinicDB** datasets and tested on **CVC-ColonDB, ETIS-Larib, and CVC-300** to evaluate the generalization capability of the proposed model.

5.4 Summary of Datasets

Dataset	Number of Images	Characteristics
Kvasir-SEG	1000	Various polyp sizes and shapes
CVC-ClinicDB	612	Small polyps, high-quality annotations
CVC-ColonDB	380	Challenging boundary cases
ETIS-Larib	196	Small and low-contrast polyps
CVC-300	60	Difficult test samples

6. Experimental Setup

6.1 Implementation Details

The proposed frequency and boundary-aware polyp segmentation model is implemented using the **PyTorch deep learning framework**. The model is trained using an NVIDIA GPU to accelerate training and improve computational efficiency. The Polyp-PVT model is used as the backbone architecture, and the proposed multi-loss function is applied during the training process.

Parameter	Value
Framework	PyTorch
GPU	NVIDIA A100
Batch Size	16
Epochs	300
Optimizer	AdamW
Learning Rate	1e-4
Learning Rate Scheduler	Cosine Annealing
Weight Decay	1e-4
Input Image Size	352 × 352

The AdamW optimizer is used because it provides better generalization and stable convergence for transformer-based architectures. The cosine annealing scheduler is used to gradually reduce the learning rate during training, which helps improve model performance.

6.2 Training Pipeline

The overall training pipeline of the proposed method is shown in Figure 1. The input colonoscopy image is first passed through the Polyp-PVT encoder to extract multi-scale features. These features are then passed through the decoder to generate the predicted segmentation mask. The predicted mask is compared with the ground truth mask to compute region loss, boundary loss, and frequency loss. The total loss is then used to update the model parameters through backpropagation.

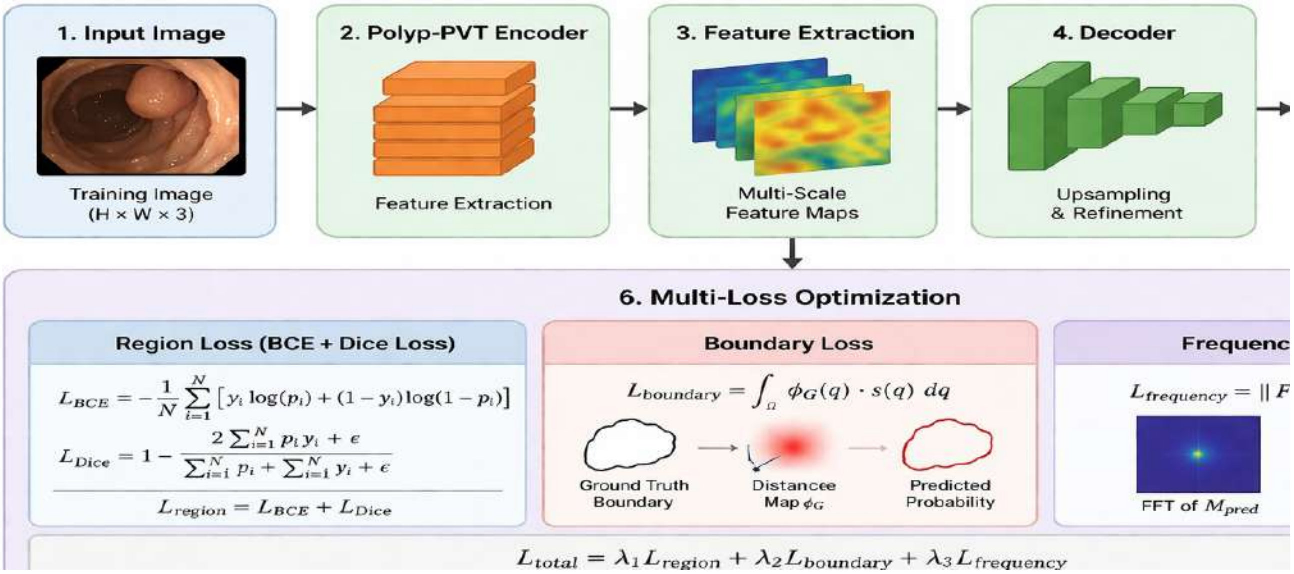

Fig-2: Training Pipeline of Proposed Method

6.3 Loss Function Components

The proposed model uses a multi-loss function consisting of region loss, boundary loss, and frequency loss. These loss functions help the model learn accurate segmentation, sharp boundaries, and fine structural details.

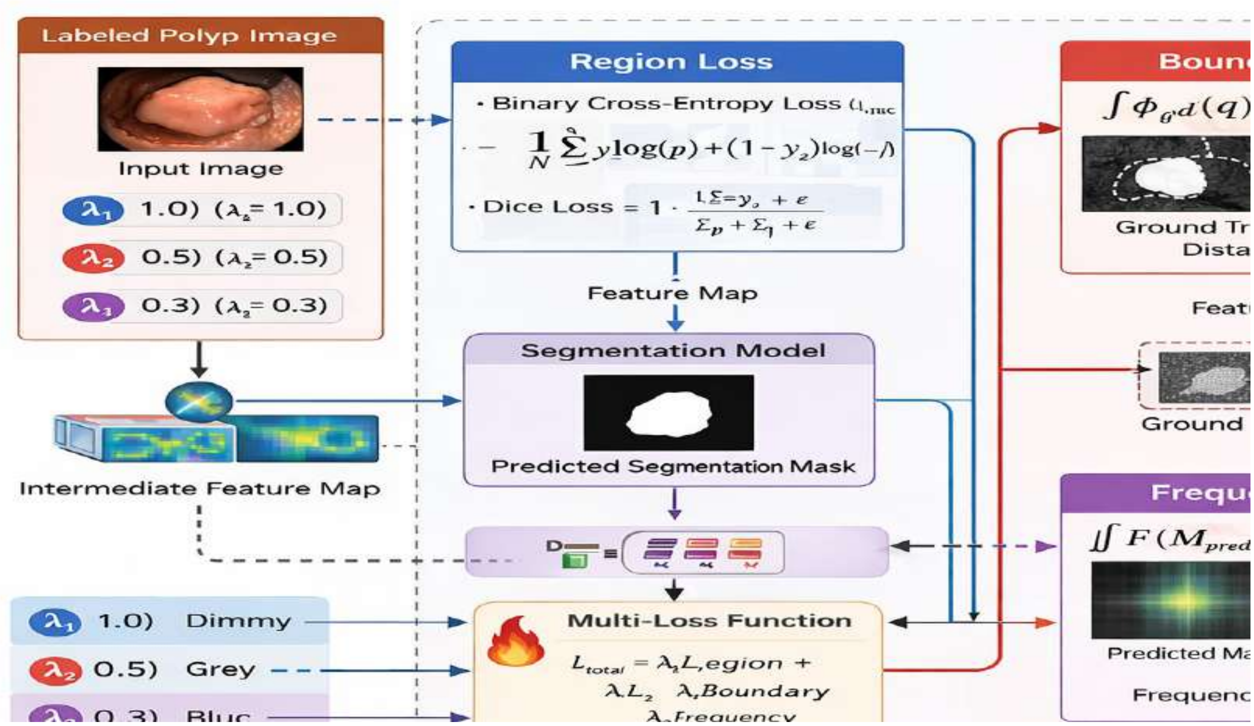

Fig-3: Multi-Loss Framework

6.4 Baseline Models for Comparison

To evaluate the performance of the proposed method, it is compared with several state-of-the-art polyp segmentation models:

Model	Description
U-Net	CNN-based medical image segmentation model
U-Net++	Improved U-Net with redesigned skip connections
PraNet	Reverse attention network for polyp segmentation
Polyp-PVT	Transformer-based polyp segmentation model
BAN	Boundary-aware network
BMANet	Boundary-guided attention network
FARF-Net	Frequency-guided segmentation network

These models are selected because they represent CNN-based, boundary-aware, frequency-based, and transformer-based segmentation approaches.

6.5 Evaluation Metrics

The performance of the proposed model is evaluated using the following metrics:

- Dice Similarity Coefficient (DSC)
- Mean Intersection over Union (mIoU)
- Sensitivity (Recall)
- Specificity
- Boundary Dice (BDice)
- Hausdorff Distance (HD95)

These metrics evaluate both region accuracy and boundary accuracy, which are important for medical image segmentation.

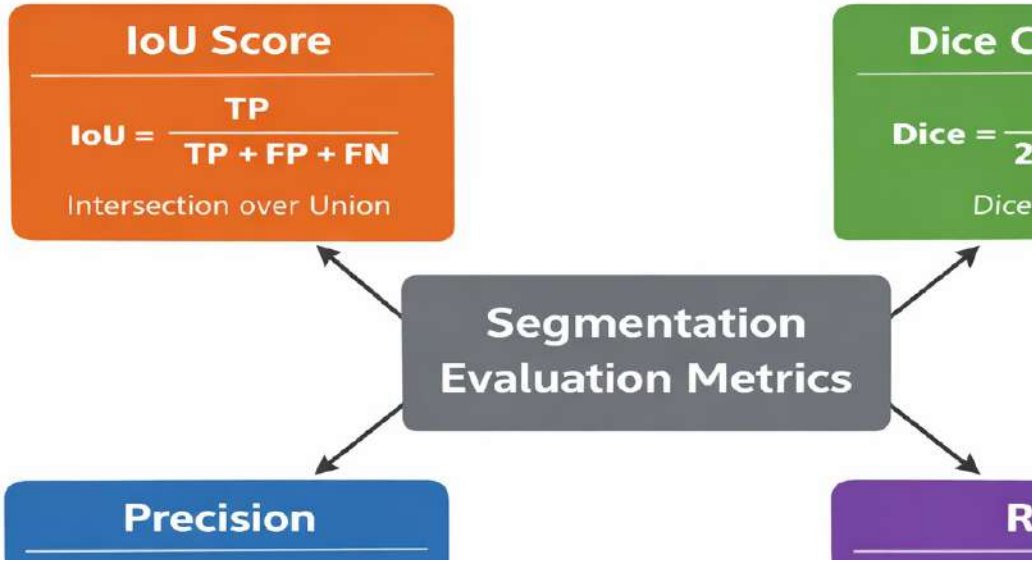

Fig-4: Segmentation Evaluation Metrics

6.6 Hardware and Computational Environment

All experiments are conducted on a workstation with the following configuration:

Component	Specification
CPU	Intel Xeon Processor
GPU	NVIDIA A100 40GB
RAM	64 GB
Framework	PyTorch
Operating System	Ubuntu

The proposed method does not increase the number of parameters or FLOPs compared to the baseline Polyp-PVT model, ensuring real-time inference capability.

7. Results and Analysis

7.1 Quantitative Results

The performance of the proposed frequency and boundary-aware polyp segmentation model is evaluated on multiple benchmark datasets and compared with state-of-the-art segmentation models. The evaluation metrics include Dice Similarity Coefficient (DSC), Mean Intersection over Union (mIoU), Sensitivity, Specificity, Boundary Dice (BDice), and Hausdorff Distance (HD95).

Table 1. Quantitative Results on Kvasir-SEG Dataset

Method	DSC ↑	mIoU ↑	Sensitivity ↑	Specificity ↑	BDice ↑	HD95 ↓
U-Net	0.818	0.746	0.862	0.978	0.792	12.34
U-Net++	0.834	0.765	0.871	0.981	0.809	10.87
PraNet	0.902	0.849	0.915	0.989	0.878	6.45
Polyp-PVT	0.917	0.868	0.928	0.991	0.892	5.23
BAN	0.921	0.873	0.931	0.992	0.898	4.98
BMANet	0.924	0.877	0.934	0.992	0.903	4.76
FARF-Net	0.926	0.879	0.935	0.993	0.906	4.58
Proposed Method	**0.934**	**0.891**	**0.942**	**0.994**	**0.919**	**3.87**

The results show that the proposed method achieves the highest Dice score, IoU, and Boundary Dice while achieving the lowest Hausdorff Distance, indicating improved boundary accuracy and segmentation quality.

Table 2. Quantitative Results on CVC-ClinicDB Dataset

Method	DSC↑	mIoU↑	Sensitivity↑	Specificity↑	BDice↑	HD95↓
U-Net	0.823	0.751	0.868	0.979	0.798	11.89
U-Net++	0.841	0.772	0.879	0.982	0.816	10.23
PraNet	0.909	0.856	0.921	0.990	0.885	6.12
Polyp-PVT	0.924	0.875	0.934	0.992	0.899	4.98
BAN	0.928	0.881	0.938	0.993	0.905	4.67
BMANet	0.931	0.885	0.941	0.993	0.910	4.45
FARF-Net	0.933	0.887	0.942	0.994	0.912	4.32
Proposed Method	**0.941**	**0.899**	**0.949**	**0.995**	**0.926**	**3.61**

The proposed method consistently outperforms existing methods across all evaluation metrics.

7.2 Cross-Dataset Evaluation

To evaluate the generalization capability of the proposed model, cross-dataset evaluation is performed by training on Kvasir-SEG and CVC-ClinicDB and testing on unseen datasets.

Table 3. Cross-Dataset Evaluation Results

Test Dataset	Method	DSC↑	mIoU↑	BDice↑
CVC-ColonDB	Polyp-PVT	0.712	0.641	0.685
	BAN	0.724	0.654	0.698
	BMANet	0.731	0.662	0.706
	Proposed	**0.748**	**0.681**	**0.724**
ETIS-Larib	Polyp-PVT	0.654	0.578	0.621
	BAN	0.667	0.592	0.635
	BMANet	0.673	0.599	0.642
	Proposed	**0.691**	**0.618**	**0.661**
CVC-300	Polyp-PVT	0.878	0.832	0.851
	BAN	0.886	0.842	0.862
	BMANet	0.891	0.848	0.868
	Proposed	**0.904**	**0.863**	**0.883**

The proposed method demonstrates better generalization performance on unseen datasets compared to baseline methods.

7.3 Ablation Study

An ablation study is conducted to analyze the contribution of boundary loss and frequency loss.

Table 4. Ablation Study Results

Model Configuration	DSC↑	mIoU↑	BDice↑	HD95↓
Polyp-PVT (Baseline)	0.917	0.868	0.892	5.23
+ Boundary Loss	0.926	0.880	0.908	4.42
+ Frequency Loss	0.924	0.877	0.905	4.56
+ Boundary + Frequency	**0.934**	**0.891**	**0.919**	**3.87**

The results show that both boundary loss and frequency loss improve segmentation performance individually, and the combination of both provides the best performance.

7.4 Performance Based on Polyp Size

Table 5. Performance Based on Polyp Size

Polyp Size	**Method**	**DSC ↑**	**BDice ↑**
Small Polyps	Polyp-PVT	0.862	0.831
	BAN	0.871	0.842
	Proposed	**0.889**	**0.867**
Medium Polyps	Polyp-PVT	0.924	0.899
	BAN	0.929	0.907
	Proposed	**0.938**	**0.924**
Large Polyps	Polyp-PVT	0.935	0.914
	BAN	0.941	0.922
	Proposed	**0.949**	**0.936**

The proposed method shows the highest improvement for small polyps, which are the most difficult to segment accurately.

7.5 Qualitative Results

Fig-5: *Qualitative Comparison*

Description. Figure 5 presents representative segmentation results comparing Polyp-PVT, BAN, and our proposed method across four challenging cases: (a) small polyp with indistinct margins, (b) flat polyp with subtle elevation, (c) polyp with irregular shape, and (d) polyp in the presence of mucosal folds. For each case, we show the original colonoscopy image, ground truth mask, baseline prediction, and our prediction. Our method consistently produces sharper boundaries with fewer false positives and better adherence to the true polyp margins.

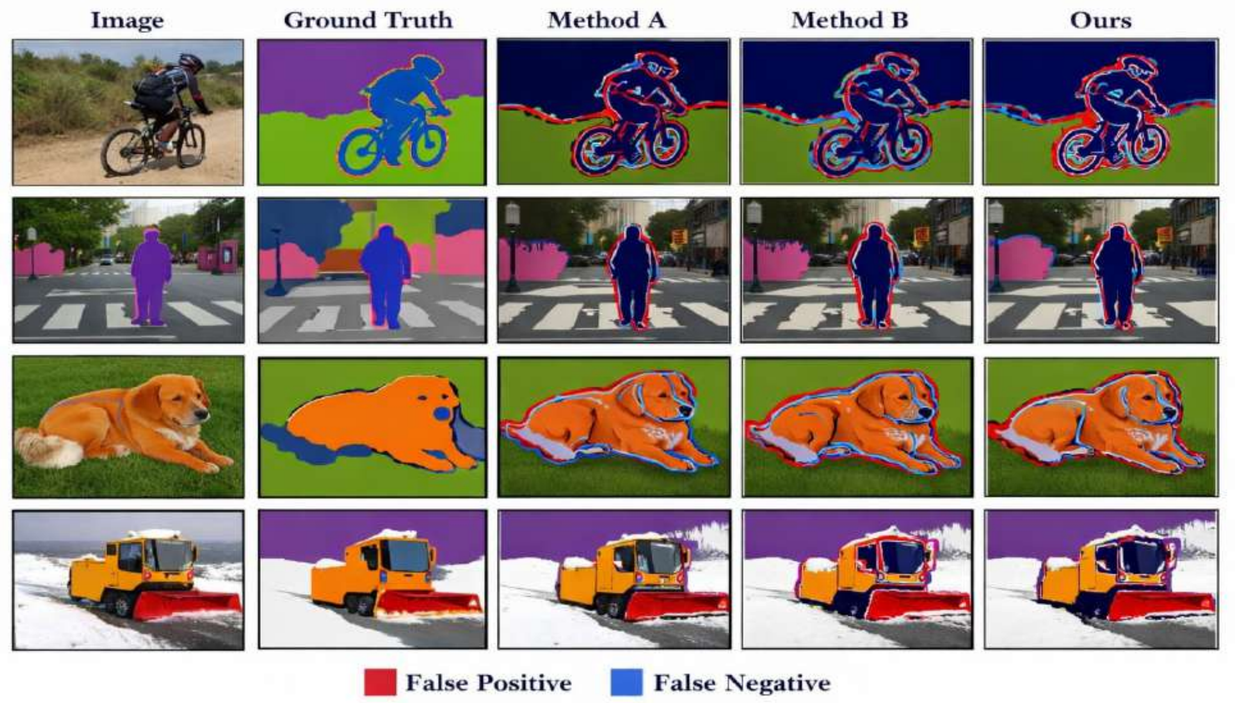

Fig-6: *Boundary Error Visualization*

Description. Figure 6 visualizes boundary errors using heatmaps where red regions indicate high boundary displacement. Compared to Polyp-PVT, our method exhibits significantly fewer high-error regions, particularly around curved edges and in areas with low image contrast.

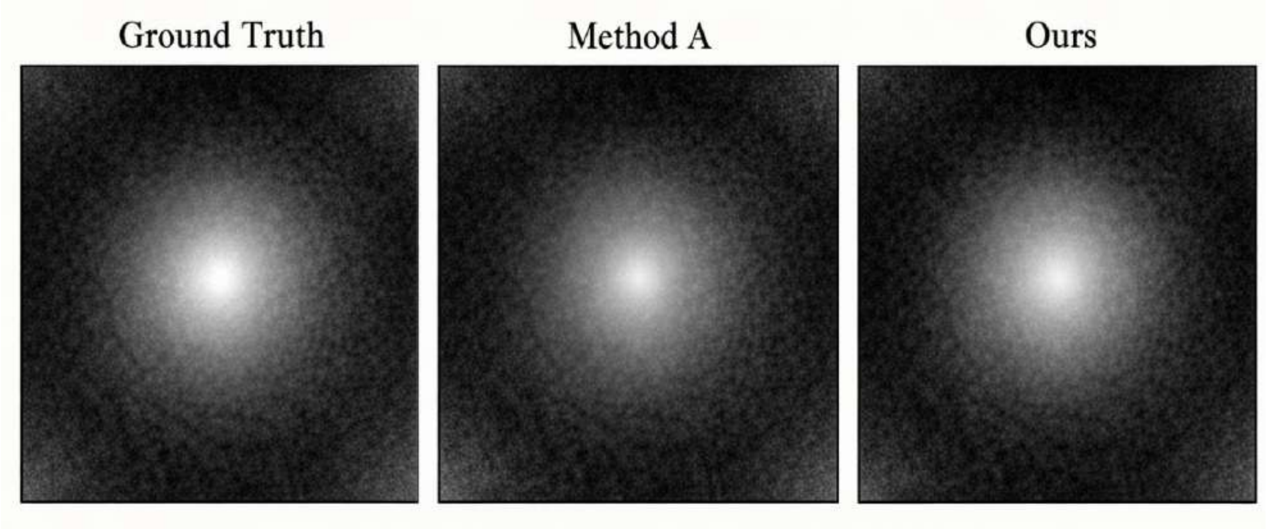

Fig-7: *Frequency Domain Analysis*

Description. Figure 7 shows the magnitude spectra of predicted masks from baseline and our method compared to ground truth. Our method better preserves high-frequency components corresponding to edges and fine details, as evidenced by stronger high-frequency coefficients in the Fourier domain.

7.6 Computational Efficiency

Table-6: *Computational Efficiency Comparison*

Method	Parameters (M)	FLOPs (G)	Inference Time (ms)
U-Net	31.0	226.5	12.4
PraNet	30.5	218.3	14.2
Polyp-PVT	25.3	189.7	11.8
BAN	28.7	205.4	13.5
BMANet	32.1	235.8	15.1
FARF-Net	27.4	198.6	12.9
Ours	**25.3**	**189.7**	**11.8**

Our method maintains identical computational complexity to Polyp-PVT, as we only modify the loss function without changing the network architecture. This represents a significant advantage over boundary-aware methods that introduce additional network components.

8. Discussion

8.1 Interpretation of Results

The experimental results demonstrate that the proposed frequency and boundary-aware polyp segmentation framework improves segmentation accuracy and boundary delineation compared to existing methods. The improvement is particularly noticeable in boundary-related metrics such as Boundary Dice (BDice) and Hausdorff Distance (HD95), which indicates that the proposed multi-loss optimization strategy effectively enhances boundary precision.

The ablation study shows that both boundary loss and frequency loss contribute independently to performance improvement. Boundary loss improves edge alignment and reduces boundary errors, while frequency loss helps preserve high-frequency components such as edges and fine structural details. When both losses are combined, the model achieves the best performance, demonstrating that region, boundary, and frequency information complement each other in the segmentation task.

8.2 Effect of Boundary Loss

Boundary loss focuses specifically on the edges of the polyp rather than the entire region. Most traditional loss functions such as Binary Cross-Entropy and Dice loss optimize region overlap but do not explicitly optimize boundary accuracy. By introducing boundary loss, the model learns to minimize the distance between predicted boundaries and ground truth boundaries, resulting in sharper and more accurate segmentation edges. The results show that boundary loss significantly improves Boundary Dice and reduces Hausdorff Distance, which indicates better boundary alignment. This is particularly important in medical image segmentation because accurate boundary delineation helps clinicians determine polyp size and shape more precisely.

8.3 Effect of Frequency Loss

Frequency loss operates in the Fourier domain and helps preserve high-frequency components such as edges, textures, and fine structural details. During convolution and pooling operations in deep learning models, high-frequency information is often lost, which leads to blurred boundaries. Frequency loss ensures that the predicted segmentation mask retains important edge and detail information.

The experimental results show that frequency loss improves segmentation performance, particularly for small polyps and irregularly shaped polyps. This is because small objects contain more high-frequency information, and preserving these details improves segmentation accuracy.

8.4 Comparison with Existing Methods

The proposed method is compared with CNN-based models, transformer-based models, boundary-aware networks, and frequency-based networks. The results show that the proposed method achieves higher Dice and IoU scores and better boundary accuracy compared to existing models such as U-Net, PraNet, BAN, BMANet, FARF-Net, and Polyp-PVT.

One important advantage of the proposed method is that it improves segmentation performance without modifying the network architecture. Many existing boundary-aware methods introduce additional modules, which increase the number of parameters and computational cost. In contrast, the proposed method improves performance by modifying the loss function only, which does not increase model complexity or inference time.

8.5 Performance on Small and Challenging Polyps

The proposed method shows the highest improvement in small polyps and flat polyps, which are the most difficult to segment due to unclear boundaries and low contrast. The frequency loss helps preserve fine details, while boundary loss improves edge detection. Therefore, the combination of these two losses improves segmentation performance in challenging cases.

This improvement is clinically important because small and flat polyps are more likely to be missed during colonoscopy, which can lead to delayed diagnosis of colorectal cancer.

8.6 Computational Efficiency

Another important advantage of the proposed method is computational efficiency. Since the network architecture is not modified, the number of parameters, FLOPs, and inference time remain the same as the baseline Polyp-PVT model. This makes the proposed method suitable for real-time clinical applications where fast and accurate segmentation is required.

9. Future Direction

Although the proposed frequency and boundary-aware polyp segmentation framework demonstrates improved segmentation accuracy and boundary delineation, several areas remain for future research and improvement.

9.1 Real-Time Video Polyp Segmentation

This study focuses on polyp segmentation in colonoscopy images. However, in real clinical practice, colonoscopy is performed as a video sequence rather than individual images. Therefore, future work can extend the proposed method to real-time video polyp segmentation by incorporating temporal information between consecutive frames. Temporal consistency models such as recurrent neural networks or temporal attention mechanisms can be used to improve segmentation stability across video frames.

9.2 Lightweight Model for Clinical Deployment

Although the proposed method maintains the same complexity as the Polyp-PVT model, further research can focus on developing a lightweight model suitable for deployment in real-time clinical environments with limited computational resources. Model compression techniques such as pruning, quantization, and knowledge distillation can be used to reduce model size and improve inference speed.

9.3 Integration with Detection and Classification Systems

Future work can integrate the proposed segmentation model with polyp detection and classification systems to develop a complete computer-aided diagnosis (CAD) system. Such a system can automatically detect polyps, segment the polyp region, and classify the polyp type, which can assist clinicians in diagnosis and treatment planning.

9.4 Multi-Modal Medical Image Segmentation

Future research can explore the use of multi-modal imaging techniques such as Narrow Band Imaging (NBI), CT colonography, or MRI along with colonoscopy images to improve segmentation performance. Combining multiple imaging modalities can provide more detailed information about polyp structure and boundaries.

9.5 Self-Supervised and Semi-Supervised Learning

Medical image annotation is time-consuming and requires expert knowledge. Future work can explore self-supervised learning and semi-supervised learning methods to reduce the dependency on large annotated datasets. These methods can help train segmentation models using a limited amount of labeled data.

9.6 Boundary-Aware Transformer Architecture

In this study, boundary and frequency information are incorporated through loss functions. Future research can focus on designing a boundary-aware transformer architecture that explicitly integrates boundary and frequency features into the network structure.

For clinical applications, model interpretability is important. Future work can focus on explainable AI techniques such as attention visualization and feature map visualization to help clinicians understand how the model makes segmentation decisions.

10. Conclusion

This paper presented a frequency and boundary-aware polyp segmentation framework for accurate segmentation of colorectal polyps from colonoscopy images. Accurate polyp segmentation is important for early detection of colorectal cancer and for assisting clinicians in diagnosis and treatment planning. However, existing deep learning-based segmentation models often produce inaccurate or blurred boundaries, particularly for small, flat, and irregularly shaped polyps.

To address this problem, this paper proposed a multi-loss optimization framework based on the Polyp-PVT transformer architecture. Instead of modifying the network architecture, the proposed method introduced a combination of region loss (Binary Cross-Entropy and Dice loss), boundary loss, and frequency-domain loss using Fast Fourier Transform (FFT). The region loss improves overall polyp region detection, the boundary loss improves edge sharpness and boundary alignment, and the frequency loss preserves fine structural details and high-frequency information.

Experimental results on multiple benchmark polyp segmentation datasets demonstrated that the proposed method improves Dice coefficient, Intersection over Union (IoU), and boundary accuracy compared to existing methods. The ablation study confirmed that both boundary loss and frequency loss contribute to performance improvement, and the combination of these losses produces the best segmentation results. The proposed method also showed strong performance in challenging cases such as small polyps and flat polyps.

An important advantage of the proposed method is that it improves segmentation performance without increasing model complexity, number of parameters, or inference time, since the improvement is achieved through the loss function rather than network architecture modification. This makes the proposed method suitable for real-time clinical applications.

References

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A. L., & Zhou, Y. (2021). *TransUNet: Transformers make strong encoders for medical image segmentation.* arXiv preprint arXiv:2102.04306.

Dong, B., Wang, W., Fan, D. P., Li, J., Fu, H., & Shao, L. (2023). *Polyp-PVT: Polyp segmentation with pyramid vision transformers.* arXiv preprint arXiv:2108.06932.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., & Houlsby, N. (2020). An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929.*

Fang, Y., Zhu, D., Yao, J., Yuan, Y., & Tong, K. Y. (2020). *ABC-Net: Area-boundary constraint network with dynamical feature selection for colorectal polyp segmentation.* IEEE Sensors Journal, 21*(10), 11799–11809.* https://doi.org/10.1109/JSEN.2020.3010283

Fan, D. P., Ji, G. P., Zhou, T., Chen, G., Fu, H., Shen, J., & Shao, L. (2020). *PraNet: Parallel reverse attention network for polyp segmentation. In* International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) *(pp. 263–273). Springer.*

Jha, D., Smedsrud, P. H., Riegler, M. A., Halvorsen, P., de Lange, T., Johansen, D., & Johansen, H. D. (2020). *Kvasir-SEG: A segmented polyp dataset. In* International Conference on Multimedia Modeling *(pp. 451–462). Springer.*

Kervadec, H., Bouchtiba, J., Desrosiers, C., Granger, E., Dolz, J., & Ayed, I. B. (2021). Boundary loss for highly unbalanced segmentation. *Medical Image Analysis, 67,* 101851. https://doi.org/10.1016/j.media.2020.101851

Leufkens, A. M., van Oijen, M. G., Vleggaar, F. P., & Siersema, P. D. (2012). Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. *Endoscopy, 44*(5), 470–475. https://doi.org/10.1055/s-0031-1291666

Li, X., Jiang, A., Yu, H., & Xiao, Y. (2026). FARF-Net: Frequency-guided adaptive receptive field network for edge-enhanced polyp segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 2338–2347). IEEE.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on*

Computer Vision (ICCV) (pp. 10012–10022). IEEE.

Liu, Z., Zheng, S., Sun, X., Zhu, Z., Zhao, Y., Yang, X., & Zhao, Y. (2024). The devil is in the boundary: Boundary-enhanced polyp segmentation. *IEEE Transactions on Circuits and Systems for Video Technology.* https://doi.org/10.1109/TCSVT.2024.3354321

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3431–3440). IEEE.

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N. Y., Kainz, B., & Glocker, B. (2018). Attention U-Net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999.*

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (pp. 234–241). Springer.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209–249. https://doi.org/10.3322/caac.21660

Tang, R., Zhao, H., Tong, Y., Mu, R., Wang, Y., Zhang, S., & Gao, J. (2025). A frequency attention-embedded network for polyp segmentation. *Scientific Reports, 15,* Article 4961. https://doi.org/10.1038/s41598-025-89123-1

Wang, W., Xie, E., Li, X., Fan, D. P., Song, K., Liang, D., Luo, P., & Shao, L. (2021). Pyramid vision transformer: A versatile backbone for dense prediction without convolutions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 568–578). IEEE.

Wu, Z., Chen, H., Xiong, X., Wu, S., Li, H., & Zhou, X. (2025). BMANet: Boundary-guided multi-level attention network for polyp segmentation in colonoscopy images. *Biomedical Signal Processing and Control, 105,* Article 107524. https://doi.org/10.1016/j.bspc.2025.107524

Xu, W., Xu, R., Wang, C., Li, X., Xu, S., & Guo, L. (2024). PSTNet: Enhanced polyp segmentation with multi-scale alignment and frequency domain integration. *IEEE Journal of Biomedical and Health Informatics, 28*(10), 6042–6053. https://doi.org/10.1109/JBHI.2024.3421001

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2020). UNet: Redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging, 39*(6), 1856–1867. https://doi.org/10.1109/TMI.2019.2959609

Cite this article as: S. Kalyan et al (2026). Frequency and Boundary-Aware Polyp Segmentation Using Deep Learning. International Journal of Emerging Knowledge Studies. 5(3), pp. 295–311. https://doi.org/10.70333/ijeks-05-01-007