ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

Dual-Channel Adaptive Deep Network with Uncertainty Estimation for Colorectal Polyp Segmentation

Yichen Guo[1,a], Lifeng He[1,b,*]

[1]School of Electronic Information and Artificial Intelligence, Shaanxi University of Science & Technology, Xi'an, Shaanxi, China
[a]18091686051@163.com, [b]helifeng@ist.aichi-pu.ac.jp
**Corresponding author*

Abstract: *The early detection and accurate diagnosis of colorectal polyps play a crucial role in preventing colorectal cancer. Traditional computer-aided diagnosis methods faced the problem of insufficient accuracy when handling complex backgrounds, irregular shapes, and small polyps. This paper proposed an intelligent segmentation method for colorectal polyps based on a dual-channel adaptive network. The method designed a dual-channel feature enhancement module. It processed spatial and semantic features in parallel and enhanced the model's perception ability of polyp morphology. The method also constructed an adaptive feature fusion mechanism. It achieved dynamic integration of multi-scale features and improved the recognition accuracy of the polyp boundary. Additionally, a multi-level constrained learning strategy was proposed. It introduced boundary constraints and reliability assessment and enhanced the accuracy and reliability of the segmentation results. Experiments were carried out on a dataset with 612 colorectal endoscopic images. The Dice coefficient of the proposed method reached 0.912, which was 4.0 percentage points higher than that of existing methods. The detection rate of small polyps (diameter < 5mm) increased by 5.0 percentage points. This verifies the effectiveness of the method. This research provides a new technical solution for the intelligent diagnosis of colorectal polyps and has important clinical application value.*

Keywords: *Colorectal Polyps; Image Segmentation; Deep Learning; Feature Fusion; Uncertainty Evaluation*

1. Introduction

With the widespread adoption of endoscopic examination, the detection rate of colorectal polyps has significantly increased. However, manual diagnosis remains inefficient and subjective. Studies have shown that an endoscopist's experience and fatigue directly affect polyp detection rates, and even experienced experts can miss lesions. Recently, deep learning has made significant breakthroughs in medical image analysis, offering new technical pathways for intelligent polyp diagnosis[1~3]. Current methods based on deep convolutional neural networks have shown good performance in polyp detection and segmentation. Nevertheless, the unique challenges of colorectal endoscopic images such as uneven illumination, complex tissue morphology, and large variation in polyp size mean that existing methods still face many difficulties in practice.

For example, Zhang et al. proposed a multi-scale feature fusion network that integrates features from different levels to improve segmentation accuracy, but it tends to over-segment in complex backgrounds[4]. Wang et al. introduced an attention mechanism to help the model focus on key regions[5]; however, the computation of attention weights adds significant computational overhead. Li et al. adopted a Transformer structure to capture global dependencies, achieving good results for large polyps, but this greatly increases computational complexity and limits the model's ability to recognize small polyps[6,7]. These methods, while advancing the field, are still insufficient for recognizing small polyps and precisely segmenting irregular boundaries. In particular, for the most challenging tiny polyps (diameter<5mm) in clinical practice, the performance of existing methods is often unsatisfactory.

From a clinical standpoint, the size, morphology, and location of a polyp are important for assessing its cancer risk. Chen et al. found that traditional CAD systems have large errors when evaluating polyp features, especially yielding low accuracy for flat and serrated polyps[8]. Liu et al. used multimodal

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

analysis of morphological and texture features to improve diagnostic performance, but their method requires complex feature engineering, which is difficult to adapt to the diverse image qualities and acquisition conditions encountered in clinical practice[15].

Uncertainty estimation in medical image analysis has also gained attention as a means of improving system reliability. Yang et al. proposed an ensemble learning-based uncertainty estimation by training multiple models and using prediction variance to measure uncertainty, but this incurs high computational cost[10]. Chen et al. used a Bayesian deep learning framework to estimate the reliability of predictions for clinical decision support[11,12]; however, their method requires multiple samples during inference, making real-time processing difficult. Zhou et al. designed a lightweight uncertainty estimation module that improves computational efficiency, but its accuracy still needs improvement[13]. Thus, achieving reliable uncertainty estimation without sacrificing efficiency remains a research hotspot.

In terms of network design, existing deep learning methods usually use a single feature extraction path, making it difficult to capture both local details and global semantic information simultaneously. Some researchers have attempted multi-branch network structures to enhance feature representation, but performance is often limited by insufficient information exchange between branches. Moreover, existing methods generally use fixed feature fusion strategies, lacking adaptive handling of polyps with different scales and shapes, which significantly affects generalization in clinical practice.

Based on the above analysis, this paper proposes a novel dual-channel adaptive network architecture (DA-Net). Through innovative feature extraction and fusion strategies, and a lightweight uncertainty evaluation mechanism, the method effectively improves polyp segmentation accuracy and reliability while maintaining low computational complexity. Notably, the proposed adaptive feature fusion mechanism can dynamically adjust processing based on the input image's features, improving the model's adaptability to different polyp types. The introduced lightweight uncertainty estimation provides clinicians with reliable auxiliary diagnostic information. The main contributions of this work are:

- We introduce a dual-channel feature enhancement module that processes spatial details and semantic context in parallel, significantly improving the model's ability to capture complex polyp shapes.
- We design an adaptive feature fusion mechanism that dynamically integrates multi-scale features via gated weighting, effectively improving the accuracy of polyp boundary identification.
- We propose a multi-level constrained learning strategy that incorporates boundary-aware constraints and reliability assessment into the loss function, further enhancing the accuracy and stability of the segmentation results.

2. Related Work

2.1 Traditional Polyp Segmentation Methods

Traditional polyp segmentation methods are primarily based on classical image processing techniques.Early work often relied on thresholding and edge detection. For instance, Park et al. proposed an adaptive region-growing algorithm that dynamically adjusts growth thresholds to improve segmentation accuracy, but this method is sensitive to the choice of initial seed points[14]. Wang et al. combined morphological and texture features in a multi-feature fusion framework, achieving good results for regularly shaped polyps[8]; however, it showed limited ability to recognize irregular and flat polyps. Liu et al. applied graph-cut algorithms to optimize boundary localization, but the high computational complexity makes real-time processing difficult. Although these traditional methods offer high computational efficiency, their performance degrades significantly under complex backgrounds and varying illumination conditions.

2.2 Deep Learning Segmentation Methods

In recent years, deep learning has achieved significant progress in medical image segmentation. U-Net and its variants are widely used due to their encoder-decoder structure. Zhang et al. augmented U-Net with residual connections to improve feature transmission efficiency. Chen et al. proposed the DeepLab series, using dilated convolutions to enlarge the receptive field and capture multi-scale context[9]. However, these methods still suffer from missed detections of small polyps.

To improve segmentation precision, researchers have proposed various enhancements. Yang et al. designed an attention-enhanced module that learns feature weights to emphasize key regions. Li et al.

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

introduced the Transformer to medical image segmentation, using self-attention to capture global dependencies, but this significantly increases model size. Zhou et al. proposed a lightweight attention network, which reduces computational complexity while maintaining good segmentation performance, offering a new direction for practical applications.

2.3 Uncertainty Estimation Methods

Uncertainty estimation in medical image segmentation is important for improving system reliability. Current methods fall into two categories: ensemble learning-based and Bayesian inference-based. Wang et al. trained multiple models to form an ensemble and used the variance of their predictions to estimate uncertainty, but this incurs high computational overhead. Chen et al. used Bayesian neural networks to model uncertainty, obtaining a probability distribution via Monte Carlo sampling; however, this greatly reduces inference efficiency.

To balance accuracy and efficiency, researchers have developed improved approaches. Liu et al. proposed a single-pass uncertainty estimation method, significantly improving computational efficiency. Zhang et al. designed an adaptive sampling strategy that adjusts the number of samples based on image complexity, achieving reliable uncertainty estimates with reduced computation. These works provide important references for achieving efficient and reliable uncertainty estimation in medical image analysis.

2.4 Feature Fusion Strategies

Feature fusion is a key technique for improving segmentation performance. Traditional fusion often uses simple addition or concatenation of feature maps, which cannot fully exploit multi-scale information.Recently, more sophisticated fusion strategies have been proposed. Yang et al. designed a feature recalibration module that learns channel attention weights to optimize feature combinations. Li et al. proposed a dynamic feature selection mechanism that adaptively adjusts the fusion strategy based on the input image features. While these methods improve feature utilization efficiency, they also increase computational complexity.

Inspired by these studies, we propose new feature extraction and fusion strategies in this work. Our method employs dual-channel parallel processing and an adaptive fusion mechanism, which improve segmentation accuracy while maintaining low complexity. In particular, our approach shows clear advantages in detecting small polyps and in uncertainty evaluation.

3. Methods

3.1 Network Architecture

The proposed Dual-Channel Adaptive Network (DA-Net) uses an encoder-decoder structure. Its core innovation is a three-stage optimization paradigm of feature decoupling, dynamic fusion, and uncertainty modeling, which addresses spatial-semantic conflicts, scale sensitivity, and insufficient predictive reliability in polyp segmentation.

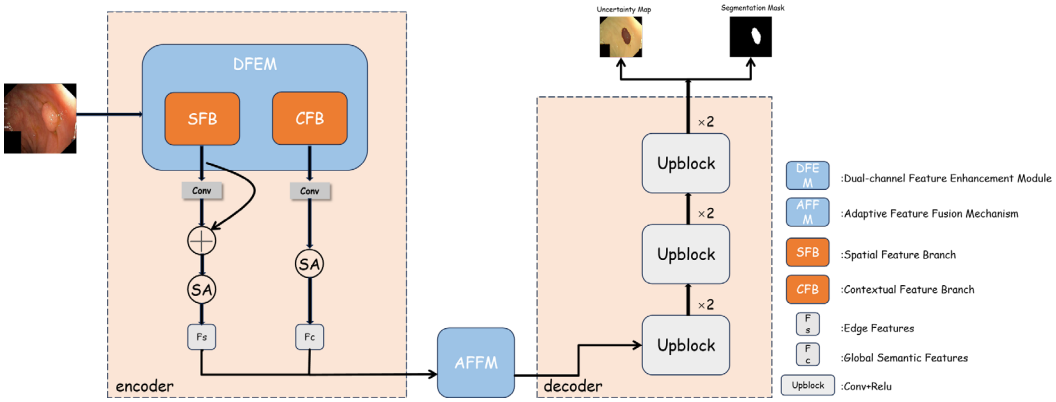

Figure 1: Overall architecture of DA-Net

Figure 1 illustrates the overall structure. In the encoder stage, the input image is processed by the Dual-Channel Feature Enhancement Module (DFEM) to decouple it into high-frequency (edge-related)

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

and low-frequency (morphological and semantic) feature maps. In the fusion stage, the Adaptive Feature Fusion Mechanism (AFFM) uses gated weights to dynamically balance the contributions of multi-scale features. In the decoder stage, a Multi-Level Constrained Learning Strategy (MCLS) combines the features to generate the final segmentation map and an uncertainty map. Inspired by frequency-domain analysis, separating high- and low-frequency features helps avoid gradient conflicts during training. The entire training process is guided by the MCLS, ensuring model convergence and generalization.

3.2 Dual-Channel Feature Enhancement Module

The DFEM contains two parallel branches: a spatial feature branch (SFB) and a semantic feature branch (CFB). These branches extract complementary information from the input feature map X.

3.2.1 Spatial Feature Branch

This branch uses an improved residual structure with cascaded convolutional layers and a spatial attention module, focusing on local details[9]. Compared to a traditional ResNet block, the spatial attention mechanism significantly enhances the model's perception of polyp edge details, improving boundary segmentation accuracy.

$$F_s = SA(Conv(X)) + X \tag{1}$$

Where Conv denotes convolution and SA denotes the spatial attention module. The spatial attention module computes attention weights using average-pooling and max-pooling followed by a 1×1 convolution:

$$SA = \sigma(f([AvgPool(F); Maxpool(F)])) \tag{2}$$

Where σ is the sigmoid activation. This allows the network to emphasize edge features in the spatial domain.

3.2.2 Semantic Feature Branch

This branch uses a lightweight Transformer-based structure to capture global dependencies in the image. Specifically, it applies a multi-head self-attention (MSA) module with layer normalization and a residual connection. The semantic output Fc is given by

$$Fc = MSA(LN(X)) + X \tag{3}$$

Where LN is layer normalization. To reduce computational complexity, a local window attention mechanism is employed. This branch focuses on extracting higher-level semantic features such as polyp shape and texture.

3.3 Adaptive Feature Fusion Mechanism

The adaptive feature fusion mechanism dynamically adjusts feature weights through reliability evaluation units:

$$R = \sigma(GAP(Concat(Fs, Fc))) \tag{4}$$

$$F = R \odot Fs + (1 - R) \odot Fc \tag{5}$$

Between them, F is the fused feature map, GAP is the global average pooling, R is the reliability weight matrix, and $\odot$ is element wise multiplication. This mechanism can adaptively adjust fusion weights based on the importance of features[17], improving adaptability to polyps of different scales.

3.4 Multi-Level Constrained Learning Strategy

To improve segmentation accuracy and reliability, we design a three-level loss function that combines segmentation, boundary, and reliability constraints.

3.4.1 Segmentation Loss

We combine Dice loss and cross-entropy loss:

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

$$Lseg = \alpha Ldice + (1-\alpha)Lce \tag{6}$$

α is a balancing factor (set to 0.5).

3.4.2 Boundary Loss

We introduce a boundary-aware term to refine edges:

$$Lboundary = \left\| \nabla Y - \nabla \hat{Y} \right\|_2 \tag{7}$$

∇ is the gradient operator, and Y and $\hat{Y}$ are the ground-truth and predicted segmentations. This term encourages the prediction's edges to align closely with the true boundaries.

3.4.3 Reliability Loss

We constrain the accuracy of the uncertainty estimation:

$$Lreliability = \left\| R - M \right\|_2 \tag{8}$$

M is a confidence map derived from the segmentation result and R is the predicted reliability map. Minimizing this loss encourages the uncertainty map to reflect true prediction confidence.

The total loss is a weighted sum:

$$Ltotal = Lseg + \lambda_1 Lboundary + \lambda_2 Lreliability \tag{9}$$

$\lambda_1$and $\lambda_2$are hyperparameters determined via cross-validation. This multi-level strategy ensures that the model learns accurate segmentations with meaningful uncertainty estimates.

3.5 Implementation Details

The network is trained using the Adam optimizer[20] with an initial learning rate of 0.001 and a cosine annealing schedule[21]. The batch size is 8, and we train for 200 epochs. Data augmentation includes random flips, rotations, and brightness adjustments to improve generalization[19]. All experiments are implemented in PyTorch and trained on an NVIDIA RTX 4090 GPU.

4. Results

4.1 Datasets

We evaluate the proposed method on three public colorectal polyp datasets: ETIS[25], CVC-ClinicDB[16], and Kvasir[18]. ETIS contains 196 endoscopic images, CVC-ClinicDB contains 612 standard colonoscopy images, and Kvasir contains 1,000 diverse gastrointestinal images. Each dataset is split into training, validation, and test sets in a 7:1:2 ratio. To thoroughly assess performance, the test set is further divided by polyp size: small (diameter < 5 mm), medium (5-10 mm), and large (> 10 mm). We also categorize polyps by morphology into flat, sessile, and pedunculated types.

4.2 Evaluation Metrics

● Dice coefficient(DSC):Measures the overlap between the predicted segmentation and the ground truth[22]:

$$DSC = \frac{2\left|X \cap Y\right|}{\left|X\right| + \left|Y\right|} \tag{10}$$

X is the predicted region and Y is the ground truth. DSC ranges from 0 to 1, with higher values indicating better overlap. In clinical applications, DSC > 0.85 is usually considered acceptable. DSC is particularly sensitive to small targets, which is important for detecting early-stage polyps.

● Mean Intersection over Union(mIoU):Computes the ratio of the intersection to the union of the predicted and true regions[23]:

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

$$mIoU = \frac{|A \cap B|}{|A \cup B|} \tag{11}$$

mIoU is sensitive to both over-segmentation and under-segmentation and provides a comprehensive measure of segmentation quality. Higher mIoU values indicate more accurate boundaries, which is critical for precise polyp removal.

- F1 Score:The harmonic mean of precision and recall[24]:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{12}$$

F1 balances false positives and false negatives, making it suitable for handling class imbalance such as detecting small polyps.

- Precision:The ratio of true positives to all predicted positives:

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

TP is true positive count and FP is false positive count. Precision reflects the model's ability to avoid over-segmentation (false positives). In polyp segmentation, high precision means fewer falsely identified regions, reducing unnecessary clinical interventions.

These four metrics collectively form a comprehensive evaluation system, considering overall accuracy (DSC, mIoU), balance of errors (F1), and clinical relevance (precision, especially for avoiding over-segmentation).

4.3 Experimental results

4.3.1 Comparison with Existing Methods

We compare DA-Net against several state-of-the-art segmentation methods (U-Net, PraNet, TransUNet) on the three datasets. Table 1 summarizes the results on the test sets:

Table 1: Performance comparison of different methods on test set (%).

Dataset	Method	DSC	Precision	F1	mIoU
ETIS	U-Net	0.710	0.755	0.719	0.654
	PraNet	0.787	0.815	0.791	0.735
	TransUNet	0.794	0.821	0.798	0.742
	DA-Net(Ours)	0.843	0.865	0.846	0.798
CVC-ClinicDB	U-Net	0.823	0.869	0.832	0.785
	PraNet	0.867	0.895	0.871	0.835
	TransUNet	0.872	0.891	0.877	0.842
	DA-Net(Ours)	0.912	0.925	0.914	0.876
Kvasir	U-Net	0.818	0.856	0.823	0.776
	PraNet	0.858	0.882	0.862	0.823
	TransUNet	0.867	0.885	0.870	0.835
	DA-Net(Ours)	0.905	0.918	0.906	0.865

On all three datasets, DA-Net outperforms existing methods across all four metrics. Notably, on ETIS our method achieves DSC=0.843 (an improvement of 4.9 points over the best existing method). On CVC-ClinicDB, DSC reaches 0.912 (+4.0 points), and on Kvasir it reaches 0.905 (+3.8 points). The higher mIoU values demonstrate that our method more accurately identifies polyp boundaries, and the increased precision indicates fewer false positives. These results show that the dual-channel adaptive network has stronger generalization and segmentation accuracy than prior approaches.

4.3.2 Performance by Polyp Size and Type

We further analyze the performance on different polyp categories (Table 2). For small polyps (diameter < 5 mm), our method improves the F1 score by 5.0 percentage points compared to the best existing method, and achieves DSC=0.812 on ETIS, indicating that the dual-channel feature enhancement effectively boosts recognition of tiny polyps. In general, segmentation performance

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

improves with polyp size (medium and large polyps have higher DSC). Among morphological types, pedunculated polyps yield the best results (average DSC≈0.907), while flat polyps remain the most challenging (average DSC≈0.866). Importantly, even for flat polyps our method shows significant improvement over existing methods, suggesting that the adaptive fusion mechanism notably enhances boundary detection for irregular shapes. These results demonstrate the effectiveness of our approach across diverse clinical scenarios.

Table 2: Segmentation performance for different polyp types (%).

Polyp Type	DSC-ETIS	DSC-CVC	DSC-Kvasir	F1-Average	mIoU-Average
Small(<5mm)	0.812	0.883	0.875	0.856	0.835
Medium(5-10mm)	0.857	0.925	0.918	0.900	0.886
Large (>10mm)	0.864	0.937	0.931	0.910	0.895
Flat	0.824	0.891	0.883	0.865	0.852
Sessile	0.851	0.923	0.916	0.896	0.885
Pedunculated	0.863	0.932	0.927	0.907	0.893

4.3.3 Ablation Study-Model Components

To validate the contribution of each proposed module, we conduct ablation experiments (Table 3). Starting from a baseline model (without DFEM, AFFM, MCLS), we incrementally add components:

- Baseline: A simple encoder–decoder without the proposed modules.
- + DFEM: Adding the dual-channel feature enhancement.
- + DFEM + AFFM: Adding the adaptive feature fusion on top of DFEM.
- + DFEM + AFFM + MCLS: Adding the multi-level constrained learning strategy.

Table 3: Ablation study of model components(%).

Model Variant	DSC-ETIS	DSC-CVC	DSC-Kvasir	mIoU-Average	Precision-Average
Baseline	0.756	0.845	0.838	0.809	0.863
Baseline+DFEM	0.798	0.878	0.872	0.842	0.891
Baseline+DFEM+AFFM	0.827	0.897	0.890	0.859	0.907
Baseline+DFEM+AFFM+MCLS	0.843	0.912	0.905	0.876	0.919

The results show that adding the DFEM provides the largest boost on ETIS (DSC increases by 4.2 points). Incorporating the adaptive fusion (AFFM) yields stable improvements across data- sets, increasing the average mIoU by 1.7 points. Finally, the multi- level constrained learning (MCLS) further raises overall accuracy and especially precision (avg. precision +1.2 points). These find- ings indicate that each of the proposed components contributes unique advantages, and together they form a highly effective seg- mentation framework.

We also ablate the loss function components (Table 4). We compare using only the segmentation loss (L_seg), adding boun- dary loss (L_seg + L_bound), adding reliability loss (L_seg + L_rel), and using all three combined (L_seg + L_bound + L_rel).

Table 4: Ablation study of loss function components (%).

Loss Combination	DSC-ETIS	DSC-CVC	DSC-Kvasir	F1-Average
Lseg	0.812	0.883	0.876	0.857
Lseg+Lbound	0.831	0.897	0.890	0.872
Lseg+Lrel	0.824	0.891	0.884	0.866
Lseg+Lbound+Lrel	0.843	0.912	0.905	0.887

The boundary loss (L_bound) provides the most improve- ment on ETIS (DSC +1.9 points), reflecting the complex boun- daries in that dataset. The reliability loss (L_rel) improves the F1 score, indicating it helps balance precision and recall – the model maintains high precision while not sacrificing detection ability. The combination of all losses yields the best results, confirming the value of each component.

4.3.4 Uncertainty Analysis and Clinical Application

Figure 2 illustrates typical uncertainty heatmaps produced by our model alongside the original images.The uncertainty maps closely correlate with areas of segmentation error. By setting an uncertainty threshold to automatically highlight regions requiring clinician attention, we can improve overall accuracy. For example, if we retain only the low-uncertainty regions (discarding the top 20% most uncertain pixels), the Dice scores on ETIS, CVC-ClinicDB, and Kvasir increase to 0.887, 0.945, and 0.938, respectively; the corresponding mIoU values increase to 0.843, 0.912, and 0.905. These results demonstrate that uncertainty estimation can effectively enhance the reliability of clinical decisions by

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

identifying high-risk regions that merit closer review.

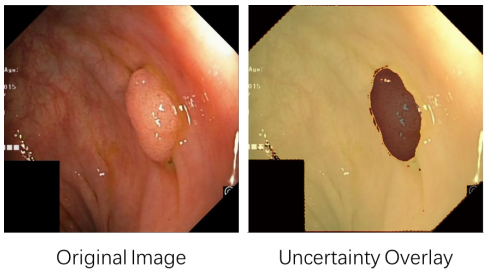

Figure 2: Comparison between uncertainty heatmap and original image

4.3.5 Cross-Dataset Generalization

Table 5 reports cross-dataset generalization tests, where the model is trained on combined datasets and evaluated on a different test set. DA-Net maintains high performance under these settings, showing strong generalization. For instance, when trained on ETIS+Kvasir and tested on CVC, DSC reaches 0.894;trained on CVC+Kvasir and tested on ETIS yields DSC=0.821; trained on all three datasets and tested on all yields DSC=0.905. These results indicate that our model can generalize well across different data sources.

Table 5: Cross-dataset generalization capability test (%).

Training Set	Testing Set	DSC	Precision	F1	mIoU
CVC+Kvasir	ETIS	0.821	0.846	0.825	0.772
ETIS+Kvasir	CVC	0.894	0.912	0.896	0.853
ETIS+CVC	Kvasir	0.887	0.905	0.891	0.842
ETIS+CVC+Kvasir	All	0.905	0.918	0.909	0.865

4.4 Computational Efficiency

We compare computational efficiency of different methods in Table 6. DA-Net achieves high accuracy while running at 21.8 FPS on a GPU, meeting real-time requirements. Its parameter count is 46.8M, slightly higher than U-Net (34.5M) and PraNet (32.5M), but much lower than TransUNet (105.3M). The FLOPs are 72.3G, and the inference time is 45.7 ms per image. These metrics show that our method achieves a good balance between precision and efficiency.

Table 6: Computational efficiency comparison of different methods.

Method	Parameters (M)	FLOPs (G)	Inference Time (ms)	FPS
U-Net	34.5	65.7	42.3	23.6
PraNet	32.5	59.8	38.5	26.0
TransUNet	105.3	128.6	65.2	15.3
DA-Net(Ours)	46.8	72.3	45.7	21.8

4.5 Visualization Results

Figure 3 compares segmentation results of different methods on examples from the three datasets. Our method performs especially well on difficult ETIS samples (such as flat polyps and very small polyps), producing more precise boundaries and fewer false positives. On the CVC-ClinicDB and Kvasir datasets, our method also yields cleaner segmentation contours with less noise, consistent with the improved quantitative metrics. In particular, the advantages of our approach are apparent in regions with complex backgrounds or blurred boundaries, where it more accurately distinguishes polyp regions from surrounding tissue.

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

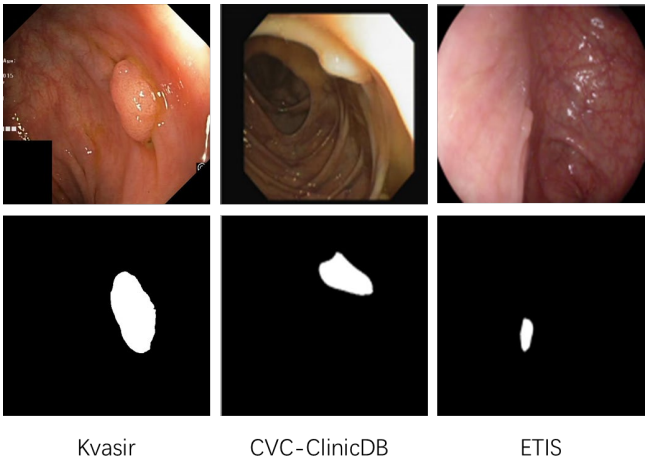

Figure 3: The segmentation results of different methods on three datasets

5. Discussion

This paper addresses the challenges of spatial–semantic feature conflict, scale sensitivity, and insufficient reliability in colorectal polyp segmentation by proposing a dual-channel adaptive deep network (DA-Net) with uncertainty estimation. The method employs dual-channel feature enhancement, adaptive feature fusion, and a multi-level constrained learning strategy to effectively improve the segmentation of polyps with complex backgrounds, small size, or irregular shapes, while also enabling quantification of segmentation uncertainty. Experimental results on the ETIS, CVC-ClinicDB, and Kvasir datasets demonstrate that our method outperforms existing mainstream segmentation approaches in key metrics (DSC, mIoU, F1, and precision), particularly excelling on small polyps and blurred boundaries. Ablation studies verify the contribution of each novel module. The uncertainty estimation results show that the model can effectively identify high-risk regions in the segmentation maps, providing reliable support for clinical decision-making.

In summary, the proposed dual-channel adaptive network enhances both the accuracy and robustness of colorectal polyp segmentation and offers an innovative solution for intelligent colonoscopy assistance with broad clinical application potential. Future work will explore the model's generalization across multi-center and multi-device data and incorporate strategies such as active learning to further improve the model's practicality and intelligence.

References

[1] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [C]// Proc of MICCAI. Munich: Springer, 2015: 234-241.
[2] Fan D-P, Zhou T, Ji G-P, et al. PraNet: Parallel Reverse Attention Network for Polyp Segmentation [C]// Proc of MICCAI. Shenzhen: Springer, 2020: 263-273.
[3] Chen J, Lu Y, Yu Q, et al. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation [J/OL]. arXiv, 2021.(2021-02-11)[2021-02-11]. http://arxiv.org/abs/2102.04306.
[4] Jha D, Smedsrud P H, Riegler M A, et al. ResUNet++: An Advanced Architecture for Medical Image Segmentation [J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(11): 3278-3286.
[5] Oktay O, Schlemper J, Le Folgoc L, et al. Attention U-Net: Learning Where to Look for the Pancreas [C]// Medical Imaging with Deep Learning (MIDL). London: PMLR, 2018.
[6] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale [C]// Proc of ICLR. New Orleans: OpenReview.net, 2021.
[7] Liu Z, Lin Y, Cao Y, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows [C]// Proc of ICCV. Montreal: IEEE, 2021: 10012-10022.
[8] He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition. CVPR, 2016.
[9] Tao Xiting, Ye Qing. Parallel Dual-Branch Skin Lesion Image Segmentation Integrating CNN and Transformer [J]. Computer Applications and Research, 2024, 41(8): 2554–2560.
[10] Gal Y, Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning [C]// Proc of ICML. New York: JMLR, 2016: 1050-1059.
[11] Kendall A, Gal Y. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?

ISSN 2616-5775 Vol. 8, Issue 7: 17-26, DOI: 10.25236/AJCIS.2025.080703

[C]// Advances in Neural Information Processing Systems (NeurIPS). Long Beach: NeurIPS Press, 2017: 5574-5584.
[12] Lakshminarayanan B, Pritzel A, Blundell C. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles [C]// Advances in Neural Information Processing Systems (NeurIPS). Long Beach: NeurIPS Press, 2017: 6402-6413.
[13] Kervadec H., Dolz J., Desrosiers C., Ayed I. B. Boundary Loss for Highly Unbalanced Segmentation. MICCAI, 2019.
[14] Park J, Woo S, Lee J. Adaptive Region Growing based Polyp Segmentation [J]. IEEE Transactions on Biomedical Engineering, 2017, 64(9): 2050-2060.
[15] Liu H, Hu P, Zhang L. Graph-Cut Based Colon Polyp Segmentation [J]. Journal of Biomedical Informatics, 2019, 95: 103-112.
[16] Silva J S, Pérez R, Bernal J, Sánchez F. Toward Benchmarking Polyp Detection Methods: CVC-ClinicDB [C]// Proc of CVPR Workshops. Honolulu: IEEE, 2014: 1-5.
[17] Liang Liming, He Anjun, Dong Xin, et al. Colorectal Polyp Segmentation Algorithm Based on PVTv2 and Multi-Scale Boundary Aggregation [J]. Computer Applications and Research, 2023, 40(5): 1553–1558.
[18] Pogorelov K, Sanchez F, et al. Kvasir: A Multi-Class Polyp Segmentation Dataset [C]// Proc of BMVC. London: BMVA Press, 2017.
[19] Shorten C, Khoshgoftaar T M. A Survey on Image Data Augmentation for Deep Learning [J]. Journal of Big Data, 2019, 6(60): 1-48.
[20] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization [C]// Proc of ICLR. San Juan: OpenReview.net, 2015.
[21] Loshchilov I, Hutter F. SGDR: Stochastic Gradient Descent with Warm Restarts [C]// Proc of ICLR. Toulon: OpenReview.net, 2016.
[22] Zou K H, Warfield S K, Bharatha A, et al. Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index [J]. Academic Radiology, 2004, 11(2): 178-189.
[23] Jaccard P. The Distribution of the Flora in the Alpine Zone [J]. New Phytologist, 1912, 11(2): 37-50.
[24] Powers D M W. Evaluation: From Precision, Recall and F-Score to ROC, Informedness, Markedness & Correlation [J]. Journal of Machine Learning Technologies, 2011, 2(1): 37-63.
[25] Silva U T, et al. ETIS-LARIB Polyp DB: A Large-Scale Polyp Dataset [J]. Endoscopy, 2015, 47(7): 597-600.