Master's Thesis Presentation for the 144th Dissertation Evaluation

# Dwarf-Net :Encoder-Decoder Neural Network for Small Polyp Segmentation by Shallow Decoder Capability

## Dwarf-Net: 단일 디코더 신경망 특성을 활용한 소형 용종 분할 연구


Donghee Kim

donghee6@cau.ac.kr


24th July, 2025

# Contents

# Introduction

## Polyp Image Segmentation

- Polyps are small tumors that form on the mucosal lining of internal organs and are classified as non-neoplastic or neoplastic [1]. Non-neoplastic polyps include hyperplastic and inflammatory types, while neoplastic polyps include adenomatous and serrated polyps [2].
- Most organ cancers start from adenomatous polyps and may become malignant without early diagnosis. Precise boundary segmentation is vital for early detection and timely treatment of colorectal cancer [3].

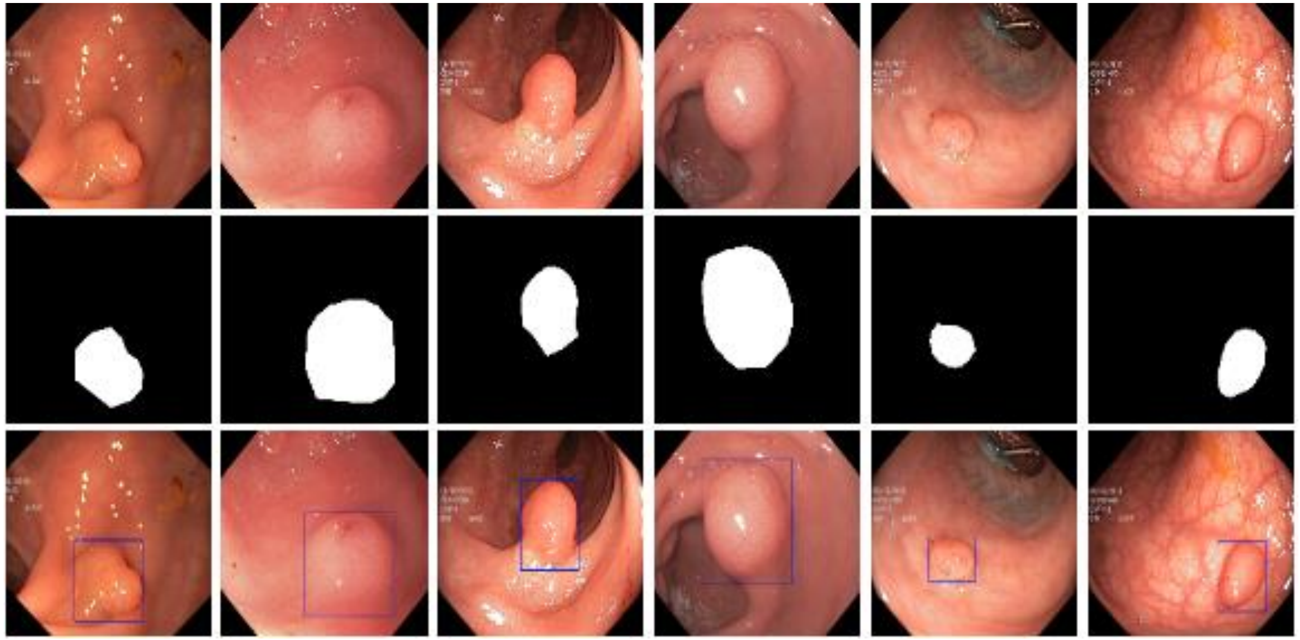


Fig. (a) Polyp image segmentation

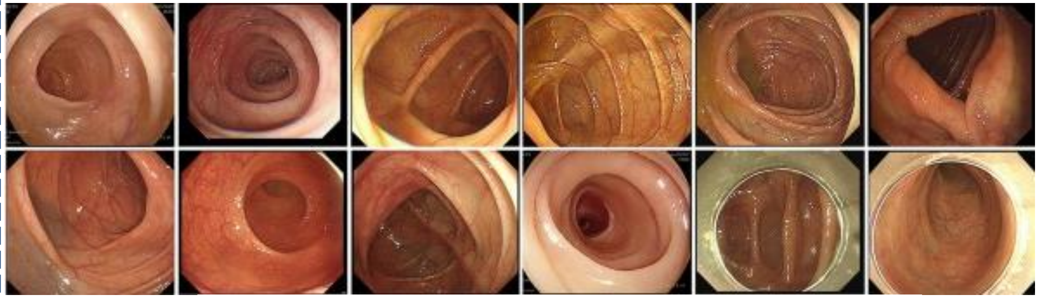


Fig. (b) Normal Polyp images


Fig. (c) Abnormal Polyp images

[1] Fan et al. "Pranet: Parallel reverse Attention Network for Polyp Segmentation." International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer International Publishing, 2020.
[2] M Gupta et al. "A Systematic Review of Deep Learning based Image Segmentation to detect Polyp." Artificial Intelligence Review 57.1 (2024): 7.
[3] Duc et al. "Colonformer: An Efficient Transformer based Method for Colon polyp Segmentation." *IEEE Access* 10 (2022): 80575-80586.

# Introduction

## Ideal Goal and Challenge in Polyp Segmentation

- Polyp segmentation is challenging because each polyp changes in size and shape at different growth rates over time. Additionally, the high intrinsic similarity between polyps and the surrounding mucosa makes accurate segmentation difficult [4].
- Polyps vary in shape and size [3], making it difficult for models to generate precise masks. Since polyps are generally smaller than surrounding tissues, they may disappear during repeated downsampling and are difficult to recover [5].

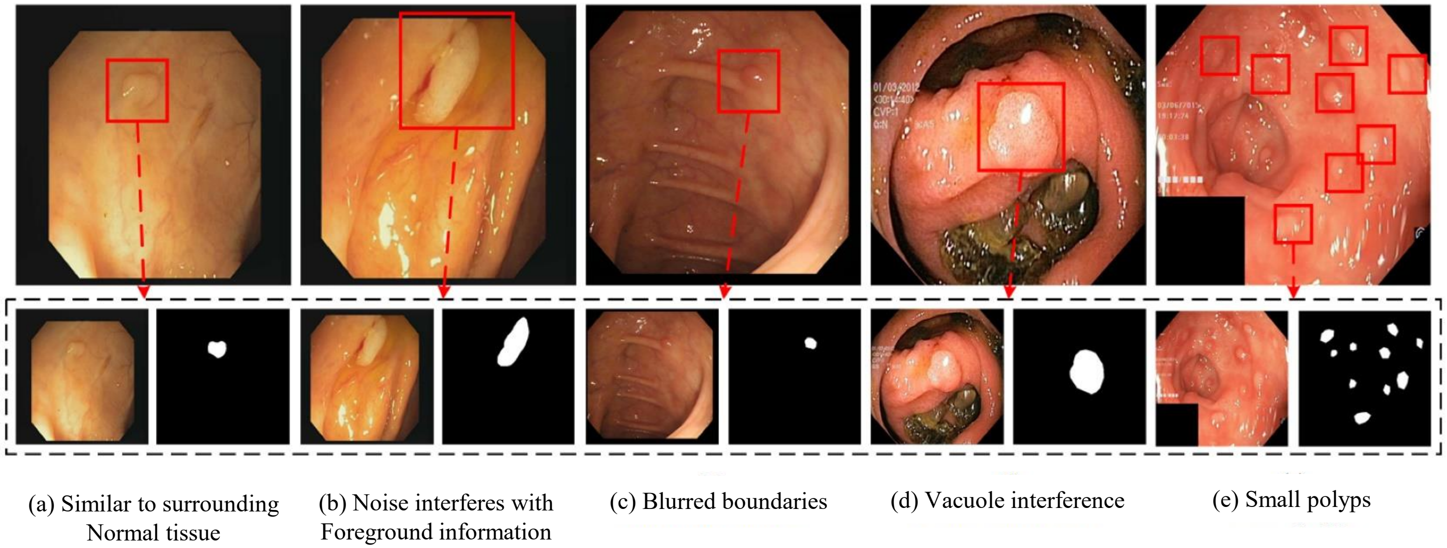



Fig. (d) Various Type of  Polyp Image

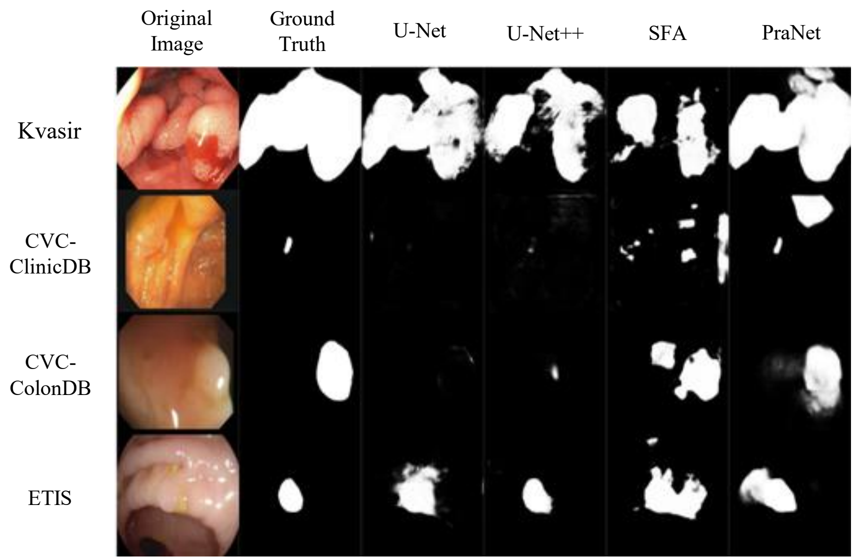



Fig. (e) Comparison of polyp segmentation

[3] Duc et al. "Colonformer: An Efficient Transformer based Method for Colon Polyp Segmentation." *IEEE Access* 10 (2022): 80575-80586.
[4] Zhou et al. "Cross-Level Feature aggregation Network for Polyp Segmentation." *Pattern Recognition* 140 (2023): 109555.
[5] Xiao et al. "CTNet: Contrastive Transformer Network for Polyp Segmentation." *IEEE Transactions on Cybernetics* (2024).

# Introduction

## Objective and Standard Procedure

- The aim of this study is to provide precise segmentation masks for small polyps in colonoscopy images, which play a critical role in the early diagnosis and treatment of colorectal cancer [6].
- Using colonoscopy images as input data, the segmentation neural network targets non-cancerous or cancerous protrusions within the colon wall tissue and generates masks that identify polyp regions in the target organ [6].

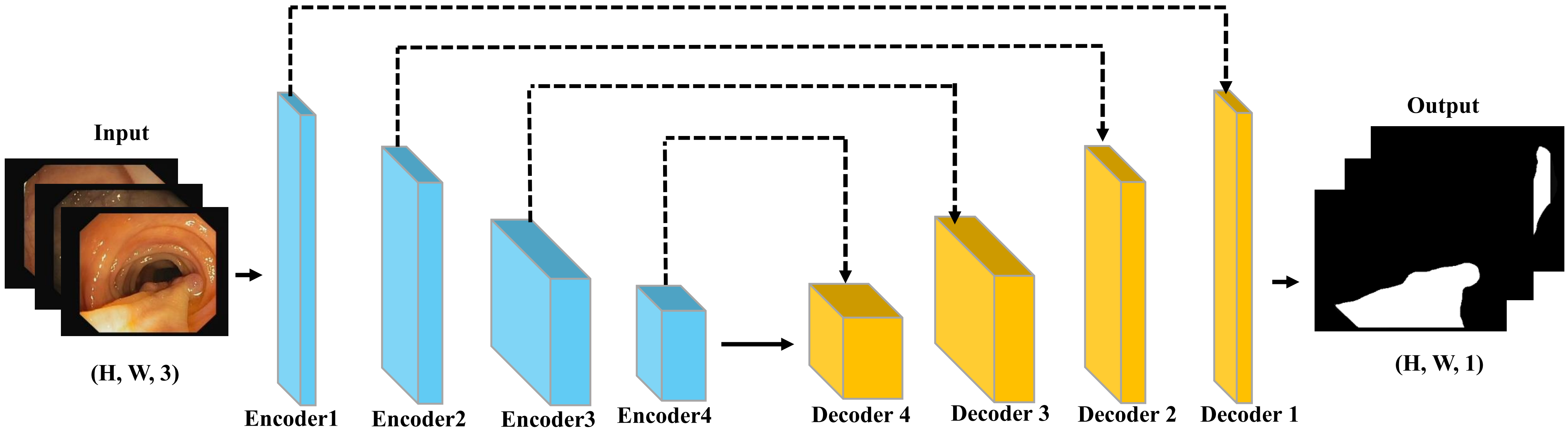


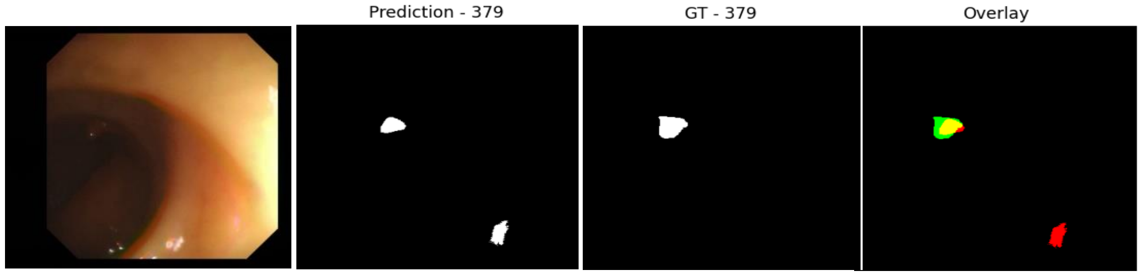

Fig. (f) The standard procedure for Polyp Segmentation

[6] Mei et al. "A Survey on Deep Learning for Polyp Segmentation: Techniques, Challenges and Future Trends." *Visual Intelligence* 3.1 (2025): 1.

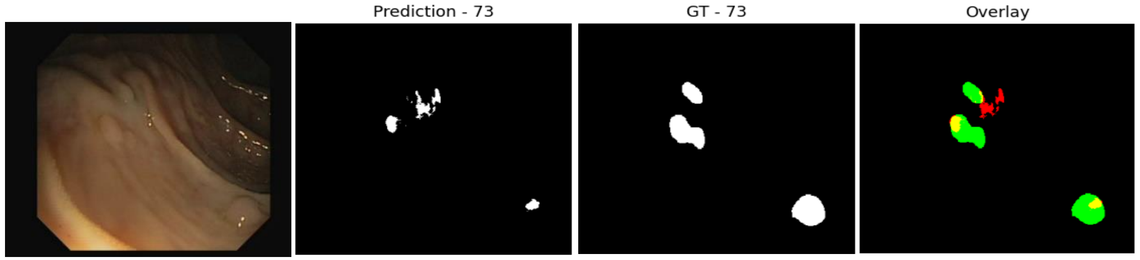
# Introduction

## Problem and Strategy

- Since small polyps occupy very few pixels in the entire image, they are often overwhelmed by the large number of background pixels. This information dominance makes it difficult for the model to learn the detailed boundaries of small polyps.
- To accurately segment small polyps without loss, the depth of the decoder should be reduced to preserve fine-grained low-level semantic information of the polyps.

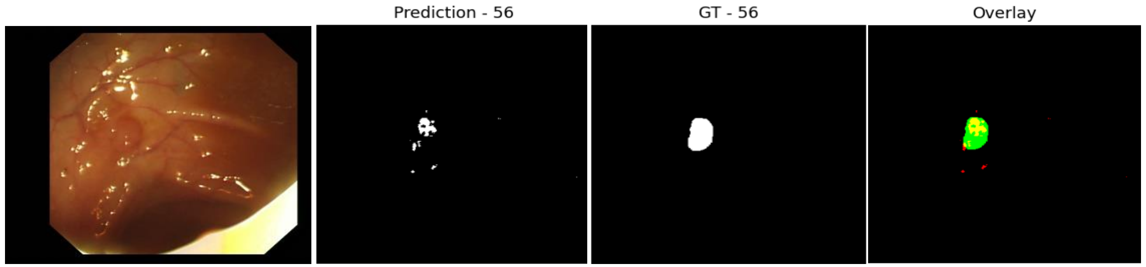



Patient 379 Original image & Comparison of prediction and ground truth (GT) through overlay




Patient 73 Original image & Comparison of prediction and ground truth (GT) through overlay


Patient 359 Original image & Comparison of prediction and ground truth (GT) through overlay




Patient 56 Original image & Comparison of prediction and ground truth (GT) through overlay

# CRC-Net (2023) [7]

- CRCNet aims to recognize polyps of various sizes by utilizing small kernels in the Local Context Module (LCM) to extract fine-grained local information while integrating the Global Context Module (GCM) to capture broader context.
- Due to the inability of neural networks to preserve the boundary information of small polyps, accurate segmentation becomes challenging. This results in poor detection and weak boundary definition, negatively impacting overall model performance.

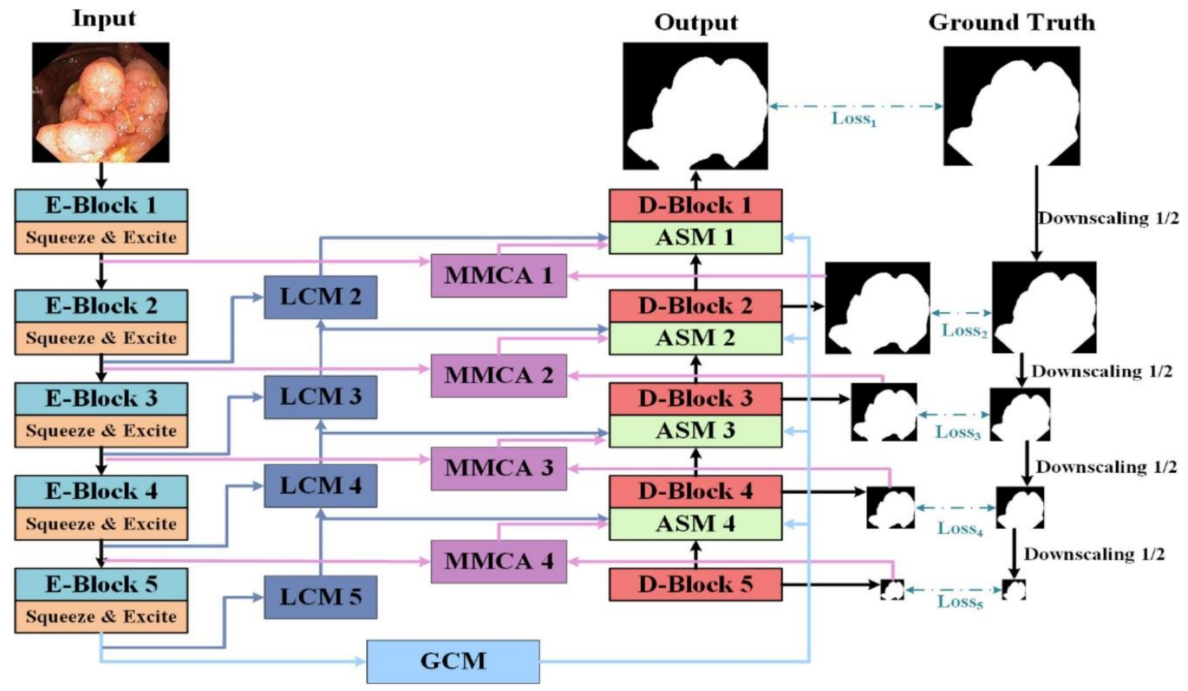



Fig. (k) CaraNet Architecture

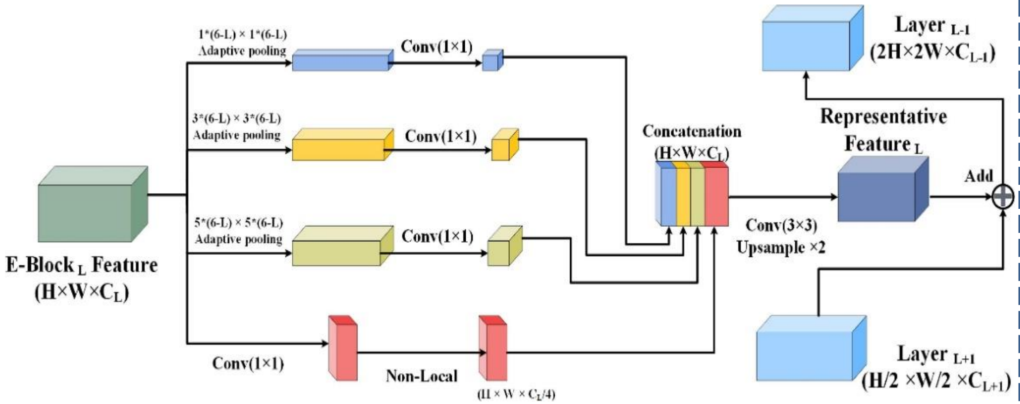



Fig. (l) Architecture of LCM

[7] Zhu et al. "CRCNet: Global-Local Context and Multi-Modality Cross Attention for Polyp Segmentation." *Biomedical Signal Processing and Control* 83 (2023): 104593.

# Related Work

## CaraNet (2023) [9]

- CaraNet merges a partial decoder, channel-based feature pyramid (CFP), and axial reverse attention (A-RA) for small-object segmentation. Reverse attention recovers missed lesions, while CFP ensures multi-scale context, improving accuracy.
- Due to the inability of neural networks to preserve the boundary information of small polyps, accurate segmentation becomes challenging. This results in poor detection and weak boundary definition, negatively impacting overall model performance.

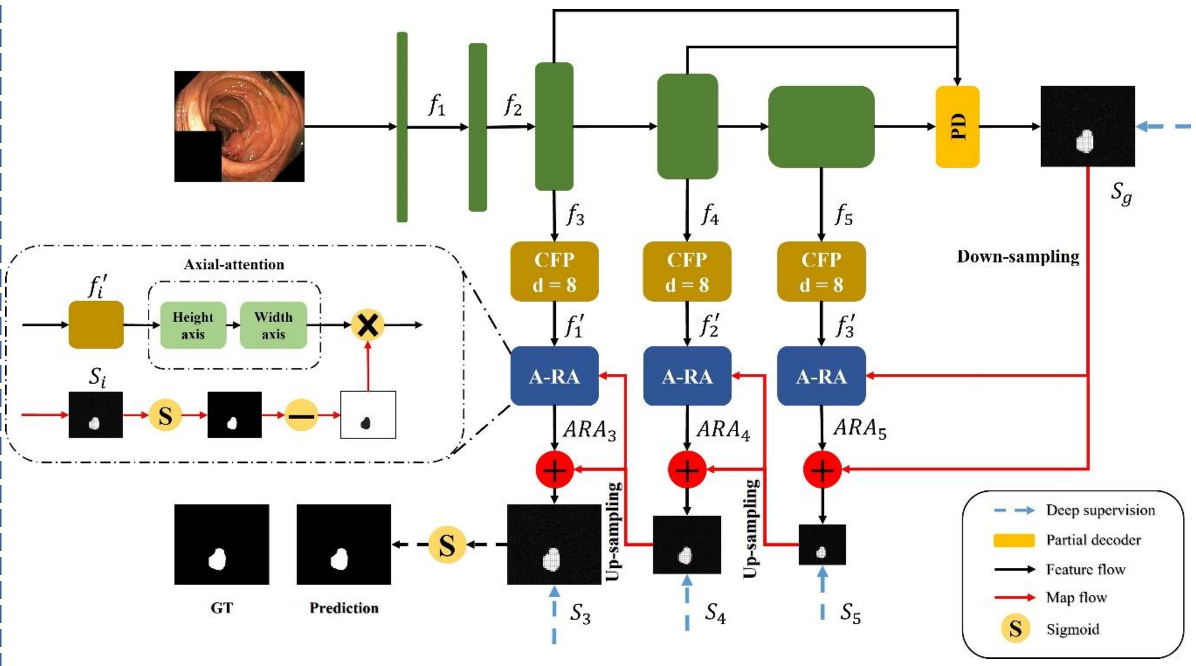



Fig. (k) CaraNet Architecture

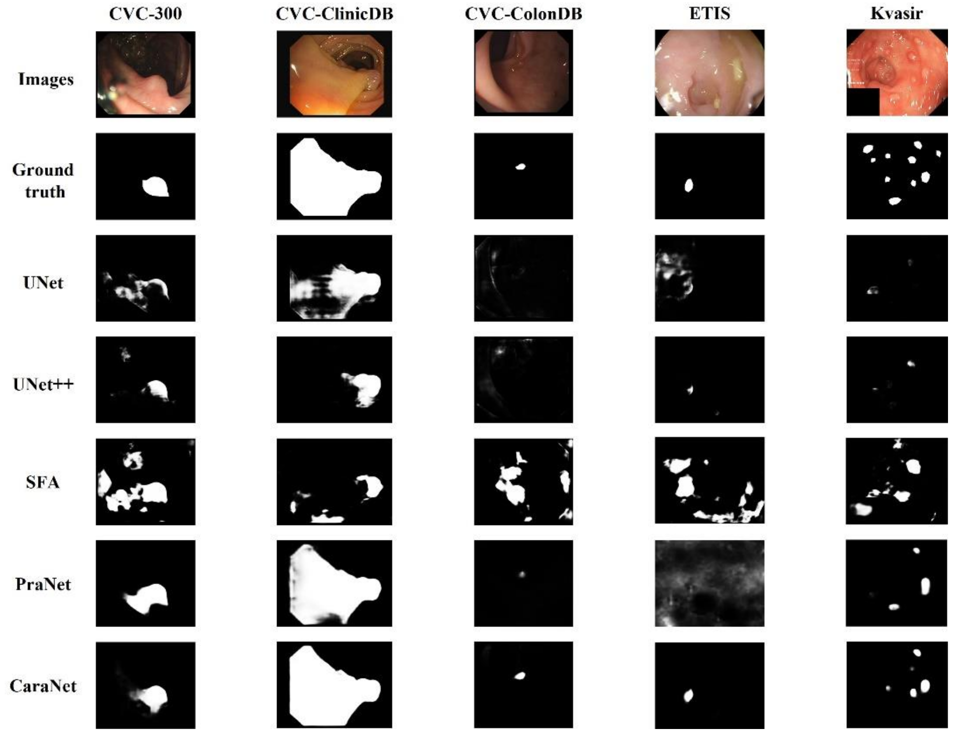



Fig. (l) Results of Cara-Net

[9] Lou et al. "CaraNet: Context Axial Reverse Attention Network for Segmentation of Small Medical Objects." *Journal of Medical Imaging* 10.1 (2023): 014005-014005.

# Related Work

## ConvSegNet (2023) [10]

- ConvSegNet extracts multi-scale contextual information from the input feature map through the Context Feature Refinement module, which utilizes convolution kernels of various sizes (1×1, 3×3, 7×7, 11×11) in parallel.
- Due to the inability of neural networks to preserve the boundary information of small polyps, accurate segmentation becomes challenging. This results in poor detection and weak boundary definition, negatively impacting overall model performance.

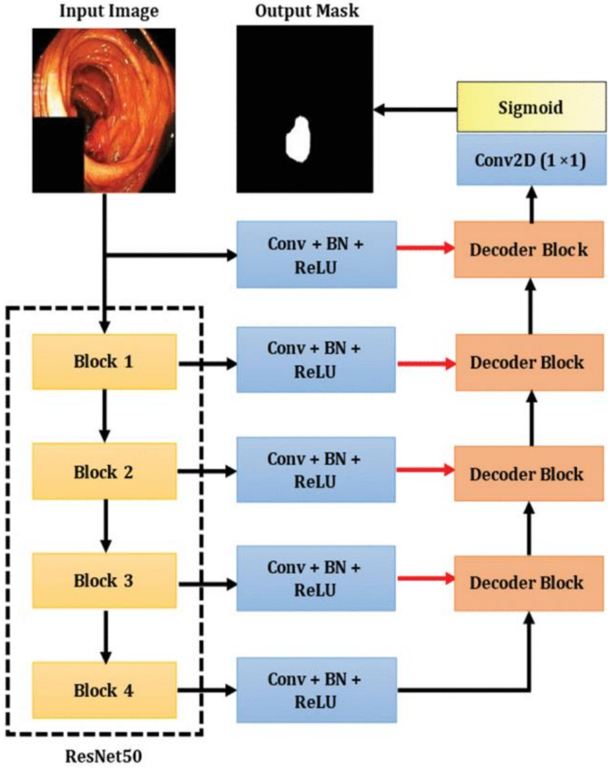


Fig. (k) ConvSegNet Architecture

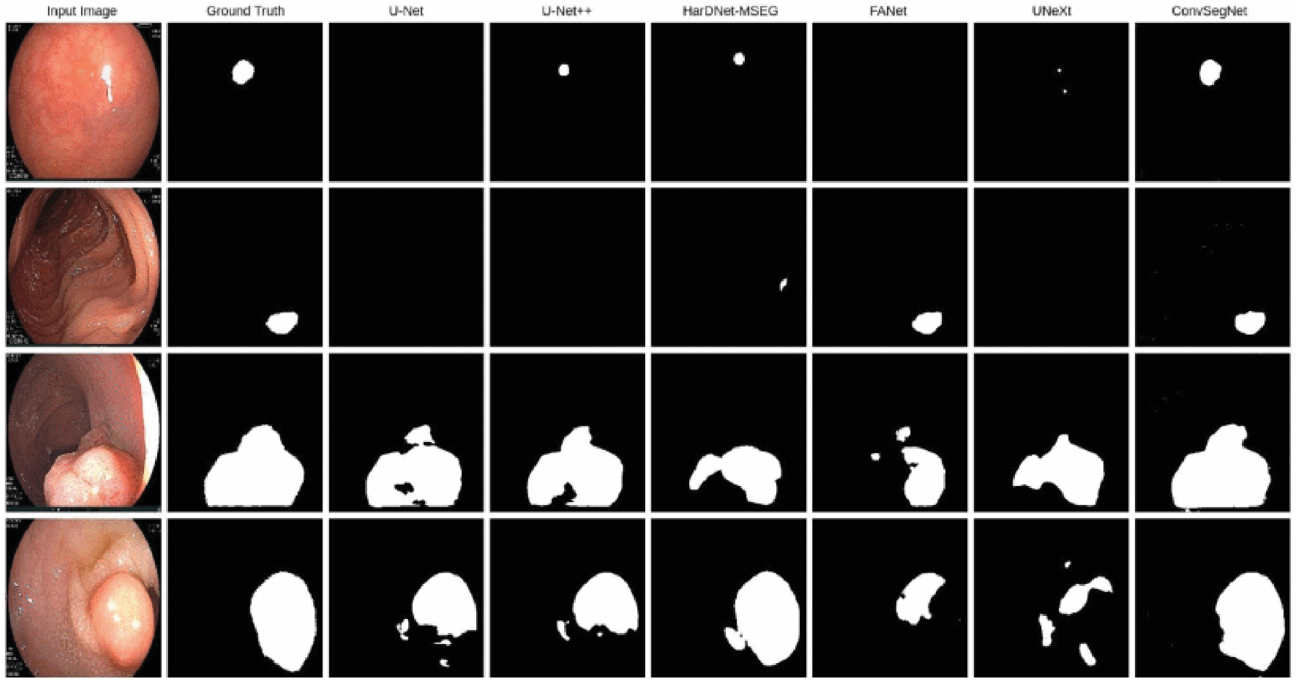


Fig. (l) Results of ConvSegNet

[10] Ige et al. "Convsegnet: Automated Polyp Segmentation from Colonoscopy using Context Feature refinement with Multiple Convolutional Kernel Sizes." *IEEE Access* 11 (2023): 16142-16155.

# Related Work

## CFHA-Net (2023) [11]

- CFHA-Net combines features from different layers in the encoder via the Cross-Scale Fusion (CCF) module and captures contextual information for both large and small polyps using multiple expansion kernels in the Dense-Receptive Feature Fusion (DFF) module.
- Due to the inability of neural networks to preserve the boundary information of small polyps, accurate segmentation becomes challenging. This results in poor detection and weak boundary definition, negatively impacting overall model performance.

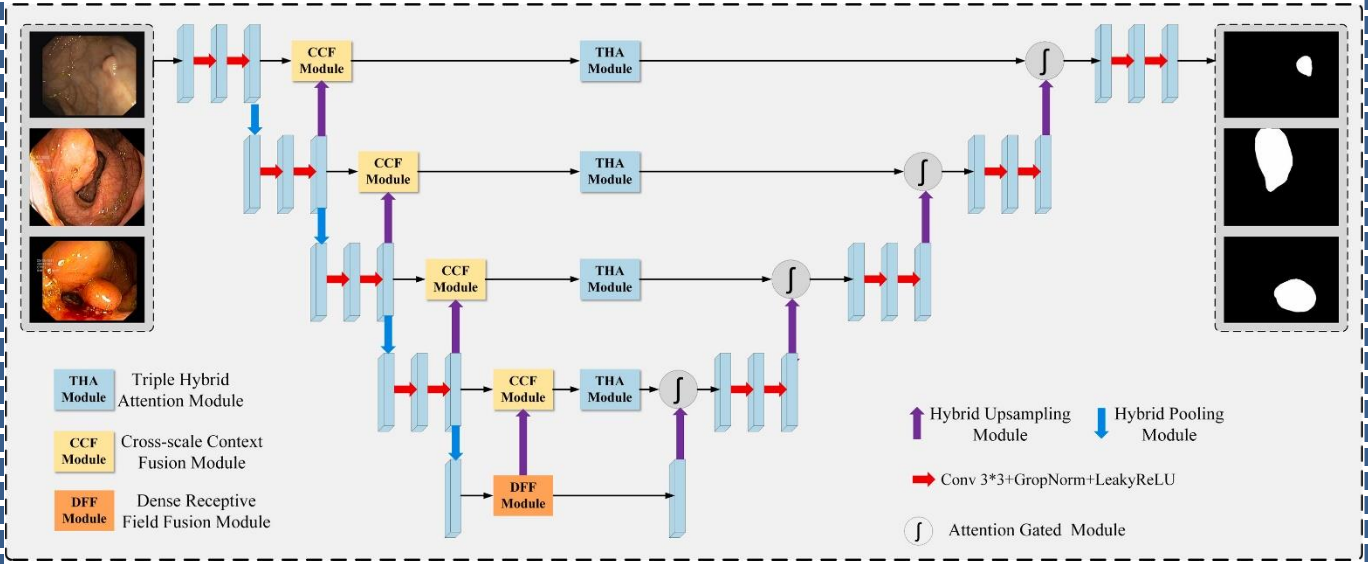



Fig. (k) CFHA-Net Architecture

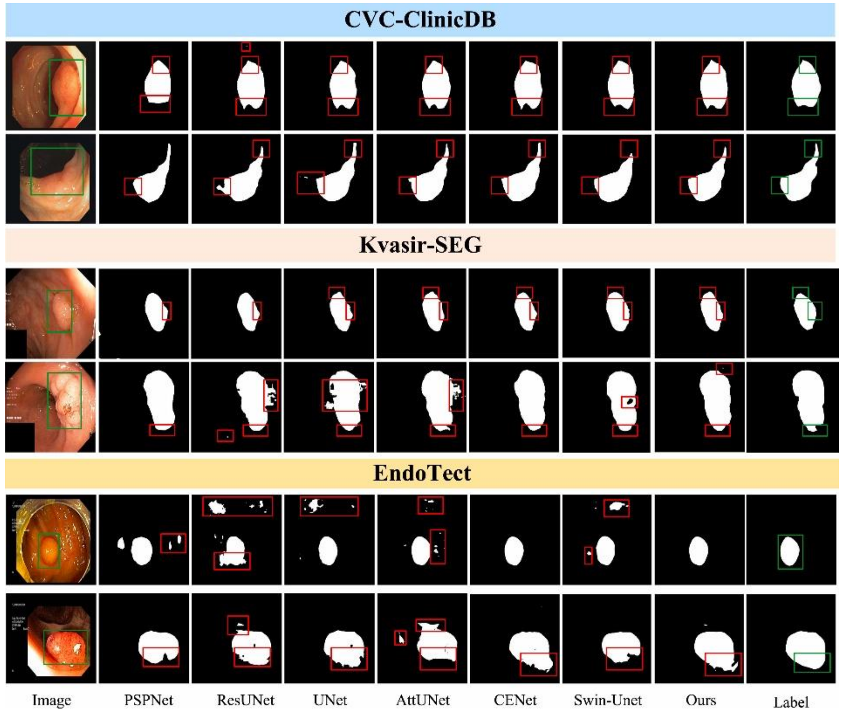



Fig. (l) Results of CFHA-Net

[11] Yang et al. "CFHA-Net: A Polyp Segmentation Method with Cross-Scale Fusion Strategy and Hybrid Attention." *Computers in Biology and Medicine* 164 (2023): 107301.

# Related Work

## Polyper (2024) [13]

- This model improves polyp boundary identification with the Boundary Sensitive Attention technique, which separates boundary and inner regions in the initial segmentation map using morphological operations and supplements boundary with inner region features.
- Due to the inability of neural networks to preserve the boundary information of small polyps, accurate segmentation becomes challenging. This results in poor detection and weak boundary definition, negatively impacting overall model performance.

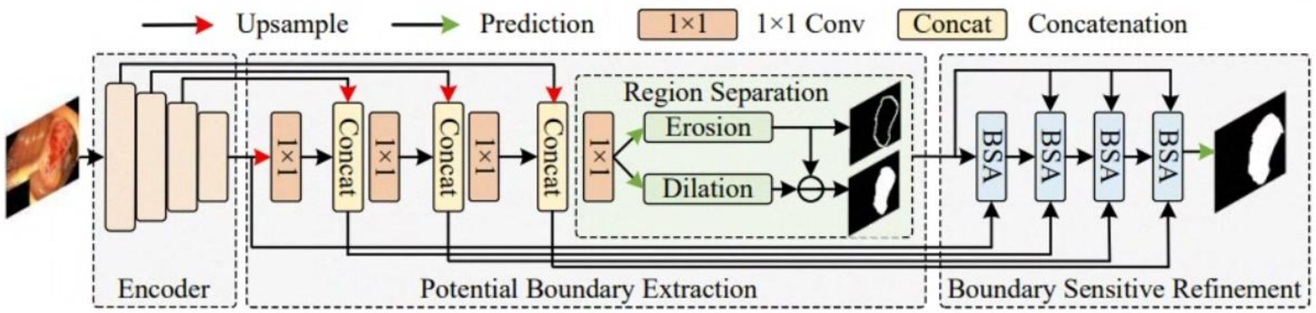



Fig. (k) Polyper Architecture

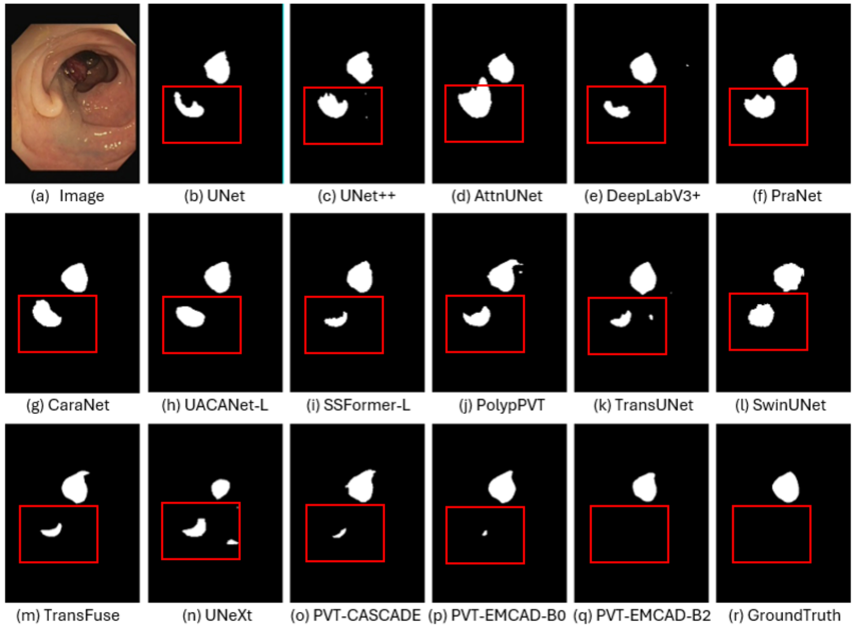



Fig. (l) Results of Polyper

[13] Shao et al. "Polyper: Boundary Sensitive Polyp Segmentation." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 38. No. 5. 2024.

# Related Work

## Common Drawback of Conventional Methods

- Conventional neural networks often fail to preserve the boundary information of small polyps, making accurate segmentation challenging. This leads to poor detection and weak boundary definition, negatively impacting overall model performance.
- To reduce the risk of losing fine details from deep network layers, we adopt a shallower architecture that better preserves low-level features critical for accurate small polyp segmentation.

| Model | Author | Publication | Summary | Drawback |
|---|---|---|---|---|
| CaraNet | Lou et al. (2023) | Journal of Medical Imaging | CaraNet merges a partial decoder, channel-based feature pyramid (CFP), and axial reverse attention (A-RA) for small-object segmentation. Reverse attention recovers missed lesions, while CFP ensures multi-scale context, improving accuracy. | Conventional neural networks often fail to preserve the boundary information of small polyps, making accurate segmentation challenging. This leads to poor detection and weak boundary definition, negatively impacting overall model performance. |
| ConvSegNet | Ige et al. (2023) | IEEE Access | ConvSegNet extracts multi-scale contextual information from the input feature map through the Context Feature Refinement (CFR) module, which utilizes convolution kernels of various sizes (1×1, 3×3, 7×7, 11×11) in parallel. | |
| CFHA-Net | Yang et al. (2023) | Computers in Biology a Medicine | CFHA-Net combines features from different layers in the encoder via the Cross-Scale Fusion (CCF) module and captures contextual information for both large and small polyps using multiple expansion kernels in the Dense-Receptive Feature Fusion (DFF) module. | |
| Polyper | Shao et al. (2024) | AAAI | This model improves polyp boundary identification with the Boundary Sensitive Attention technique, which separates boundary and inner regions in the initial segmentation map using morphological operations and supplements boundary with inner region features. | |

# Proposed Method

## Notation and Definition

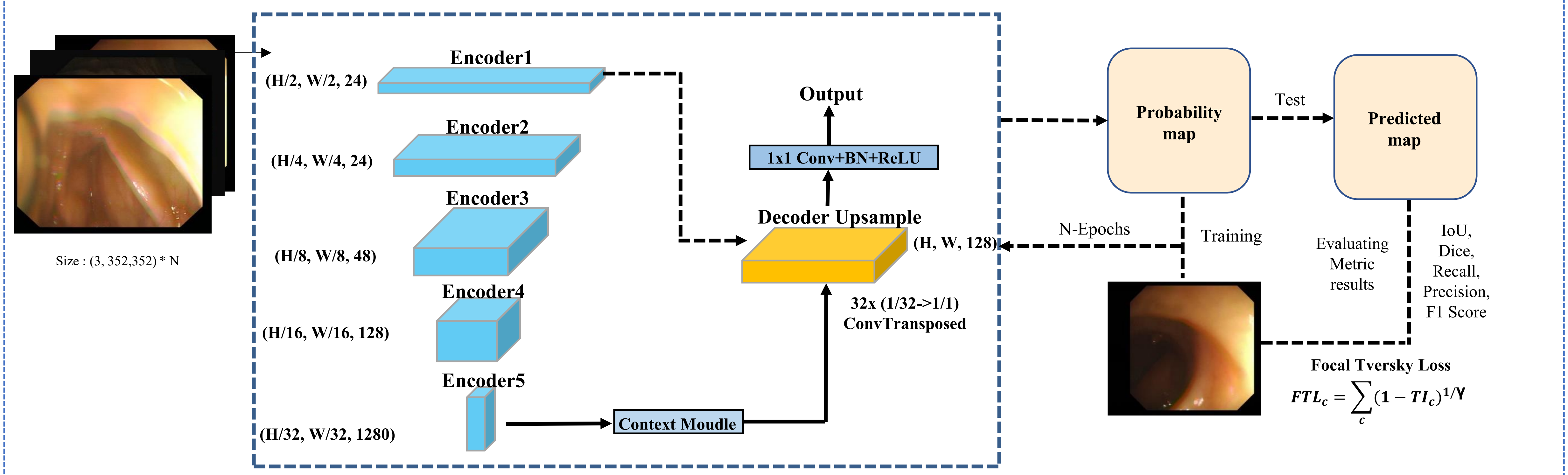
- Given input, $X = \{X_i\}_{i=1}^{m}$ and their corresponding Ground Truth (GT) $Y = \{Y_i\}_{i=1}^{m}$, consider a 2D image $X_i \in R^{H\times W}$ with its associated GT C-class segmentation masks, $Y_i \in \{1,2,\dots,C\}^{H\times W}$, where H, W are height and width.
- The goal of small polyp segmentation is to design a function $P$ so that $\widehat{Y_i} = P(X_i)$ closely approximates $Y_i$, with the learnable parameters are optimized to minimize to minimize the segmentation loss $L = \sum_c (1 - TI_c)^{1/\gamma}$ over the entire dataset.

| Notation | Name | Definition |
|---|---|---|
| $X = \{X_i\}_{i=1}^{m},\ X_i \in R^{H\times W}$ | Input images | Endoscopic images of colorectal polyps |
| $Y = \{Y_i\}_{i=1}^{m},\ Y_i \in \{1,2,\dots,C\}^{H\times W}$ | Ground Truth | Ground Truth of the Input images |
| $TI_c = \frac{\sum_{i=1}^{N} p_{ic} g_{ic} + \varepsilon}{\sum_{i=1}^{N} p_{ic} g_{ic} + \alpha \sum_{i=1}^{N} p_{i\bar{c}} g_{ic} + \beta \sum_{i=1}^{N} p_{ic} g_{i\bar{c}} + \varepsilon}$<br>$FTL_c = \sum_c (1 - TI_c)^{1/\gamma}$<br>where, $p_{ic}$ is the probability that pixel $i$ is of the lesion class $c$ and $p_{i\bar{c}}$ is the probability pixel $i$ is of the non-lesion class, $\bar{c}$. The same is true for $g_{ic}$ and $g_{i\bar{c}}$, respectively. $\alpha$ and $\beta$ are hyper-parameters. | Focal Tversky Loss | Focal loss function based on the Tversky index to address the issue of data imbalance |

# Proposed Method

## Procedural Overview

- In this polyp segmentation study, endoscopic images were resized to 352 × 352 for consistent input dimensions. This helps prevent small lesions from being overly reduced and allows the model to better learn polyp shapes and boundaries.
- The data is input into the proposed encoder-decoder-based neural network model for training and optimization using Focal -Tversky loss, with performance measured using IoU, Dice-Coefficient, F1 score, Precision and Recall metrics on the evaluation dataset.



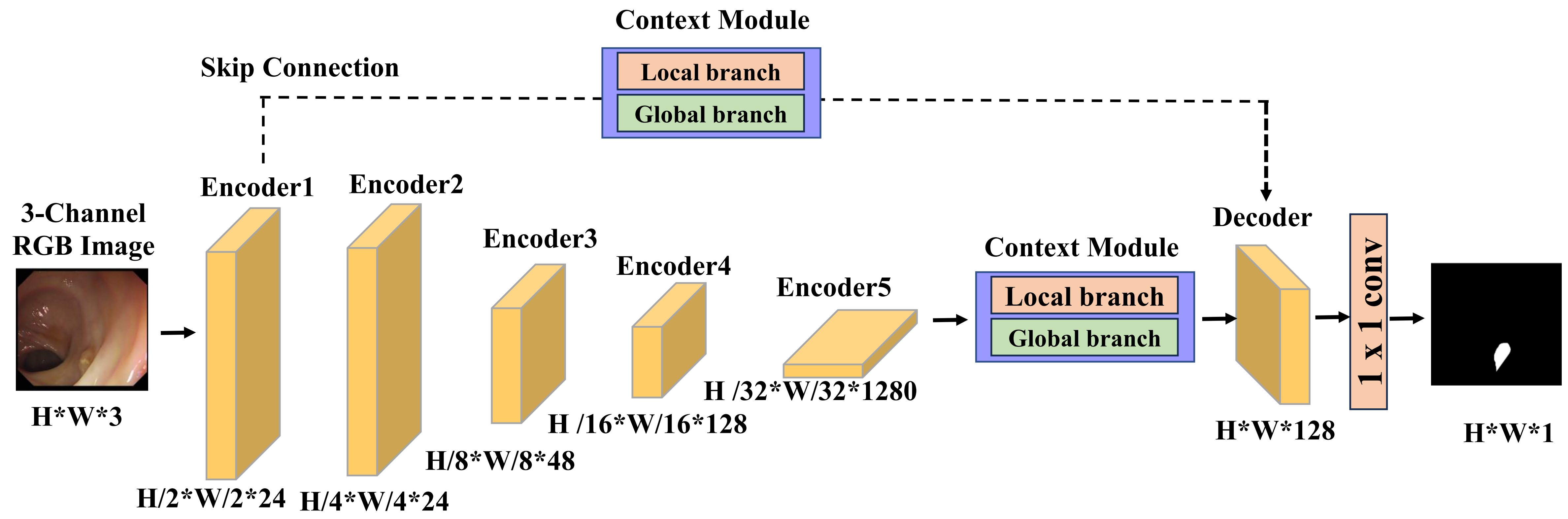

Overview of segmentation procedure of Proposed Method

# Proposed Method

## Details

- To better preserve the detailed structure of small polyps, we designed the decoder as a single-stage module. This simplified architecture helps reduce feature degradation typically caused by repeated upsampling operations in deep decoders.
- High-resolution skip connections from the early encoder layers were effectively incorporated during decoding. This strategy enhances the model's ability to retain fine spatial details and improves segmentation performance, particularly for small and ambiguous lesions.

Schematic overview of proposed encoder-decoder



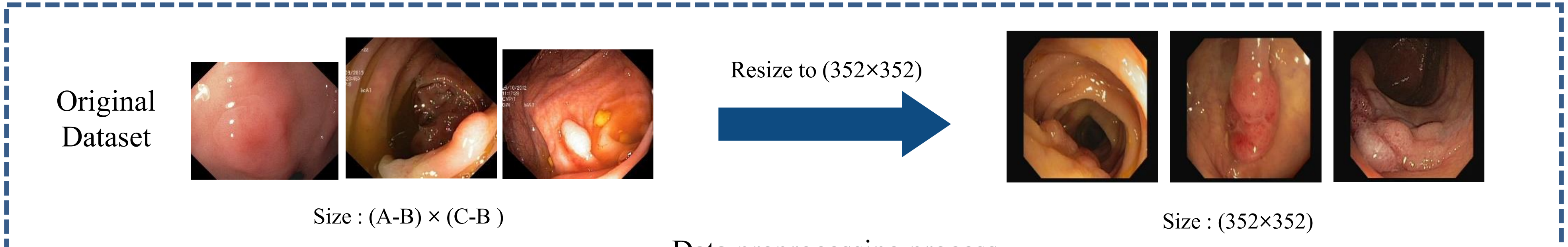
# Experimental Settings

## Data description

| Characteristics | CVC-Clinic DB | Kvasir-SEG | CVC-Colon DB | ETIS-LaribPolypDB |
|---|---|---|---|---|
| **Images** | 612 | 1000 | 380 | 196 |
| **Image Shape** | (384 × 288) | (401-672) × (415-547) | (574 × 500) | (1225 × 966) |

## Dataset

- The dataset consists of polyp images of various resolutions, which were resized to 352×352 to standardize the image size for research use.




Data preprocessing process

## Cross Validation

- Number of Iterations: 10
- Data Split Ratio (Train: Validation: Test) = 0.8 : 0.1: 0.1

# Experimental Settings

## Hyperparameter Settings

| Model | Batch size | Epochs | Loss Function | Opimizer | Learning Rate |
|---|---|---|---|---|---|
| Xu et al. (2023)<br>Zhang et al. (2023)<br>Rahman et al. (2023)<br>Aghdam et al. (2023)<br>Rahman et al. (2024) | 8 | 100 | Focal Tversky Loss | AdamW | 1e-4 |

## Evaluation Measures

- $Dice\ coefficient = \frac{2TP}{2TP+FP+FN}$, $IoU(Intersection\ of\ Union) = \frac{TP}{TP+FP+FN}$, $F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall}$, $Precision = \frac{TP}{TP+FP}$, $Recall = \frac{TP}{TP+FN}$ are the evaluation metrics, where TP is true positive, FN is False Negative, TN is True Negative, and FP is False Positive.

# Experimental Results

## Comparison to Conventional Methods

- The proposed single-decoder model, designed to effectively extract features of small polyps, achieves better performance in Dice Coefficient, IoU, Precision, and F1 Score compared to existing models.

| Evaluation measure | Dataset | Dwarf-Net | CRC | Cara | ConvSegnet | CFHA | Polyper |
|---|---|---|---|---|---|---|---|
| **Dice** | **CVC-CLINIC DB** | **0.9324 ± 0.0100** | 0.9103 ± 0.0523 | 0.8363 ± 0.0055 | 0.9023 ± 0.0204 | 0.8700 ± 0.0275 | 0.9137 ±0.0190 |
| | **KVASIR** | **0.8924 ± 0.0166** | 0.8794 ± 0.0153 | 0.7270 ± 0.0227 | 0.8839 ± 0.0334 | 0.8073 ± 0.0074 | 0.8669 ± 0.0145 |
| | **CVC-ColonDB** | **0.9052 ± 0.0287** | 0.8656 ± 0.0644 | 0.7607 ± 0.0559 | 0.8839 ± 0.0334 | 0.8367 ± 0.0331 | 0.8678 ± 0.0421 |
| | **ETIS** | **0.9104 ± 0.0225** | 0.4542 ± 0.0542 | 0.7507 ± 0.0389 | 0.8588 ± 0.0546 | 0.8441 ± 0.0527 | 0.8883 ± 0.0469 |
| **Iou** | **CVC-CLINIC DB** | **0.8602 ± 0.0133** | 0.8572 ± 0.0179 | 0.7023 ± 0.0144 | 0.8095 ± 0.0254 | 0.7742 ± 0.0320 | 0.8410 ± 0.0215 |
| | **KVASIR** | **0.8255 ± 0.0179** | 0.8154 ± 0.0176 | 0.6283 ± 0.0353 | 0.7552 ± 0.0444 | 0.7280 ± 0.0106 | 0.7960 ± 0.0207 |
| | **CVC-ColonDB** | **0.8235 ± 0.0333** | 0.7959 ± 0.0463 | 0.6510 ± 0.0377 | 0.7552 ± 0.0444 | 0.7163 ± 0.042 | 0.7957 ± 0.0351 |
| | **ETIS** | **0.7606 ± 0.0431** | 0.4910 ± 0.1099 | 0.4293 ± 0.0541 | 0.6665 ± 0.0768 | 0.6371 ± 0.0614 | 0.7603 ± 0.0466 |
| **Precision** | **CVC-CLINIC DB** | **0.9458 ± 0.0073** | 0.9179 ± 0.0904 | 0.8617 ± 0.0228 | 0.8922 ± 0.0252 | 0.9343 ± 0.0142 | 0.8490 ± 0.0088 |
| | **KVASIR** | **0.9079 ± 0.0179** | 0.9001 ± 0.0223 | 0.7833 ± 0.0286 | 0.8575 ± 0.0448 | 0.8640 ± 0.0218 | 0.8089 ± 0.0217 |
| | **CVC-ColonDB** | **0.9371 ± 0.0169** | 0.8541 ± 0.1078 | 0.8657 ± 0.0170 | 0.8575 ± 0.0448 | 0.8343 ± 0.0227 | 0.8286 ± 0.0238 |
| | **ETIS** | **0.9305 ± 0.0240** | 0.3211 ± 0.1693 | 0.8063 ± 0.0289 | 0.7810 ± 0.0825 | 0.7557 ± 0.0758 | 0.8385 ± 0.0191 |
| **Recall** | **CVC-CLINIC DB** | **0.9192 ± 0.0184** | 0.9076 ± 0.0232 | 0.8147 ± 0.0139 | 0.9078 ± 0.0097 | 0.8154 ± 0.0430 | 0.8816 ± 0.0342 |
| | **KVASIR** | **0.8777 ± 0.0280** | 0.8604 ± 0.0327 | 0.6783 ± 0.0192 | 0.8240 ± 0.0307 | 0.7580 ± 0.0082 | 0.8285 ± 0.0161 |
| | **CVC-ColonDB** | **0.8766 ± 0.0502** | 0.8758 ± 0.0487 | 0.6870 ± 0.0851 | 0.8240 ± 0.0307 | 0.8410 ± 0.0485 | 0.8166 ± 0.0659 |
| | **ETIS** | **0.8927 ± 0.0425** | 0.8341 ± 0.0397 | 0.7043 ± 0.0517 | 0.8597 ± 0.0287 | 0.8600 ± 0.0280 | 0.8471 ± 0.0839 |

# Experimental Results

## Results of Model Experiments with 5% of Small Size Polyps in Each Dataset

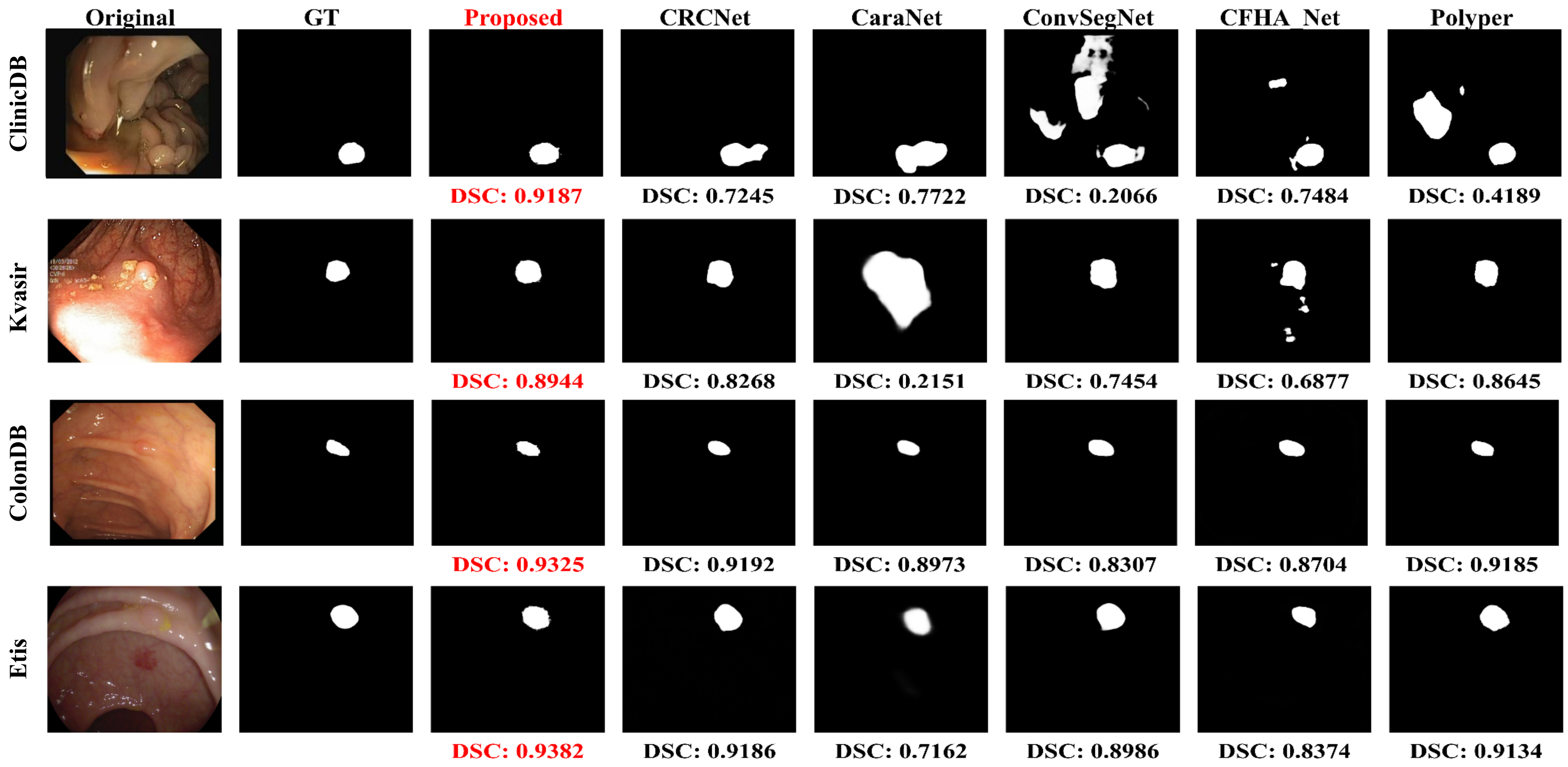
- The proposed single-decoder model, which focuses on effectively extracting features of small polyps occupying less than 5% of the image, achieves superior performance compared to existing models.

| Evaluation measure | Dataset | Dwarf-Net | CRC | Cara | ConvSegnet | CFHA | Polyper |
|---|---|---|---|---|---|---|---|
| Dice | CVC-CLINIC DB | **0.8948 ± 0.0279** | 0.8926 ± 0.0225 | 0.5010 ± 0.0350 | 0.8278 ± 0.0348 | 0.8217 ± 0.0531 | 0.8812 ± 0.0329 |
| | KVASIR | **0.8717 ± 0.0574** | 0.8637 ± 0.0517 | 0.5823 ± 0.0522 | 0.7645 ± 0.0771 | 0.7939 ± 0.0691 | 0.8155 ± 0.0598 |
| | CVC-ColonDB | **0.8776 ± 0.0323** | 0.8540 ± 0.0450 | 0.6938 ± 0.0527 | 0.7982 ± 0.0483 | 0.6975 ± 0.2155 | 0.8341 ± 0.0881 |
| | ETIS | **0.8243 ± 0.0483** | 0.4908 ± 0.1218 | 0.4597 ± 0.1162 | 0.6553 ± 0.1447 | 0.3509 ± 0.1617 | 0.8163 ± 0.0591 |
| Iou | CVC-CLINIC DB | **0.8370 ± 0.0240** | 0.8291 ± 0.0210 | 0.3777 ± 0.0305 | 0.7440 ± 0.0376 | 0.7480 ± 0.0508 | 0.8151 ± 0.0379 |
| | KVASIR | **0.7933 ± 0.0677** | 0.7895 ± 0.0584 | 0.4932 ± 0.0502 | 0.6607 ± 0.0842 | 0.7005 ± 0.0691 | 0.7395 ± 0.0624 |
| | CVC-ColonDB | **0.8066 ± 0.0352** | 0.7829 ± 0.0446 | 0.6039 ± 0.0512 | 0.7021 ± 0.0583 | 0.6295 ± 0.2103 | 0.7553 ± 0.1002 |
| | ETIS | **0.7411 ± 0.0454** | 0.4142 ± 0.1191 | 0.3780 ± 0.1069 | 0.5392 ± 0.1451 | 0.2938 ± 0.1401 | 0.7452 ± 0.0567 |
| Precision | CVC-CLINIC DB | **0.8925 ± 0.0274** | 0.8864 ± 0.0254 | 0.5753 ± 0.0398 | 0.7661 ± 0.0396 | 0.8315 ± 0.0433 | 0.8819 ± 0.0436 |
| | KVASIR | **0.8705 ± 0.0823** | 0.8582 ± 0.0709 | 0.5463 ± 0.0565 | 0.6802 ± 0.0833 | 0.7681 ± 0.0797 | 0.8065 ± 0.0720 |
| | CVC-ColonDB | **0.8784 ± 0.0274** | 0.8723 ± 0.0534 | 0.7294 ± 0.0607 | 0.7175 ± 0.0655 | 0.6991 ± 0.2368 | 0.8488 ± 0.1057 |
| | ETIS | **0.8605 ± 0.0545** | 0.4404 ± 0.1316 | 0.5569 ± 0.1216 | 0.5568 ± 0.1487 | 0.3452 ± 0.1627 | 0.8575 ± 0.0630 |
| Recall | CVC-CLINIC DB | **0.9070 ± 0.0323** | 0.8959 ± 0.0240 | 0.4752 ± 0.0337 | 0.8738 ± 0.0308 | 0.8386 ± 0.0481 | 0.8976 ± 0.0359 |
| | KVASIR | **0.8913 ± 0.0445** | 0.8815 ± 0.0310 | 0.7378 ± 0.0477 | 0.8855 ± 0.0402 | 0.8708 ± 0.0715 | 0.8600 ± 0.0470 |
| | CVC-ColonDB | **0.8953 ± 0.0378** | 0.8652 ± 0.0354 | 0.6988 ± 0.0557 | 0.8932 ± 0.0294 | 0.7832 ± 0.0980 | 0.8449 ± 0.0611 |
| | ETIS | **0.9107 ± 0.0539** | 0.9098 ± 0.0522 | 0.4404 ± 0.1259 | 0.9086 ± 0.0554 | 0.7646 ± 0.1578 | 0.7973 ± 0.0614 |

# Experimental Results

## In-depth Analysis

- The proposed model achieves segmentation accuracy close to the ground truth, surpassing three comparison models. In contrast, comparison models often segment non-target areas or fail to produce results.



# Experimental Results

## In-depth Analysis

- Small polyps occupy only a few pixels in the image, making their features prone to loss during encoding. To address this, skip connections use high-resolution features extracted solely from the initial encoder.

| Evaluation measure | Dataset | One-Skip connection | Two-Skip connection | Three-Skip connection | Four-Skip connection |
|---|---|---|---|---|---|
| Dice | CVC-CLINIC DB | **0.9324 ± 0.0100** | 0.9239 ± 0.0211 | 0.9296 ± 0.0108 | 0.9231 ± 0.0150 |
| | KVASIR | **0.8924 ± 0.0166** | 0.8890 ± 0.0138 | 0.8904 ± 0.0149 | 0.8884 ± 0.0142 |
| | CVC-ColonDB | **0.9052 ± 0.0287** | 0.8962 ± 0.0449 | 0.8743 ± 0.0354 | 0.9049 ± 0.0228 |
| | ETIS | **0.9104 ± 0.0225** | 0.7803 ± 0.1996 | 0.6516 ± 0.2927 | 0.6979 ± 0.2837 |
| Iou | CVC-CLINIC DB | **0.8602 ± 0.0133** | 0.8519 ± 0.0178 | 0.8588 ± 0.0111 | 0.8256 ± 0.0150 |
| | KVASIR | **0.8255 ± 0.0179** | 0.8229 ± 0.0148 | 0.8230 ± 0.0161 | 0.8215 ± 0.0139 |
| | CVC-ColonDB | **0.8235 ± 0.0333** | 0.8111 ± 0.0408 | 0.8064 ± 0.0309 | 0.8104 ± 0.0297 |
| | ETIS | **0.7606 ± 0.0431** | 0.6187 ± 0.1498 | 0.6125 ± 0.1453 | 0.6275 ± 0.1747 |
| Precision | CVC-CLINIC DB | 0.9458 ± 0.0073 | 0.9415 ± 0.0078 | 0.9477 ± 0.0083 | **0.9490 ± 0.0071** |
| | KVASIR | 0.9079 ± 0.0179 | 0.9058 ± 0.0215 | 0.9158 ± 0.0145 | **0.9229 ± 0.0143** |
| | CVC-ColonDB | 0.9371 ± 0.0169 | **0.9441 ± 0.0132** | 0.9237 ± 0.0309 | 0.9389 ± 0.0193 |
| | ETIS | **0.9305 ± 0.0240** | 0.7709 ± 0.2527 | 0.6323 ± 0.3687 | 0.7065 ± 0.3598 |
| Recall | CVC-CLINIC DB | **0.9192 ± 0.0184** | 0.9076 ± 0.0359 | 0.9126 ± 0.0189 | 0.8990 ± 0.0275 |
| | KVASIR | **0.8777 ± 0.0280** | 0.8735 ± 0.0257 | 0.8669 ± 0.0189 | 0.8568 ± 0.0309 |
| | CVC-ColonDB | **0.8766 ± 0.0502** | 0.8570 ± 0.0806 | 0.8248 ± 0.0584 | 0.8738 ± 0.0379 |
| | ETIS | **0.8927 ± 0.0425** | 0.8265 ± 0.0948 | 0.8110 ± 0.1406 | 0.8555 ± 0.0985 |

# Conclusion and Contribution

- Improved segmentation performance for polyps with unclear boundaries and small size.
- In this study, we proposed a novel neural network with a single decoder structure that can accurately segment sma polyps without losing information.
- The proposed model demonstrated the best performancecompared to other models, significantly outperforming existing models when evaluated using four datasets: CVC-ClinicDB, CVC-ColonDB, Kvasir and Etis.

## Contribution

- Conferences
  - Analyzing Small Polyps through Endoscopic Image Segmentation, ***IEIE***, 2025
  - A Crux on Audio-Visual Emotion Recognition in the Wild with Fusion Methods, ***ICCE***, 2025

# Thank you

# Reference


[1] Fan et al. "Pranet: Parallel reverse Attention Network for Polyp Segmentation." International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer International Publishing, 2020.

[2] M Gupta et al. "A Systematic Review of Deep Learning based Image Segmentation to detect Polyp." Artificial Intelligence Review 57.1 (2024): 7.

[3] Duc et al. "Colonformer: An Efficient Transformer based Method for Colon polyp Segmentation." IEEE Access 10 (2022): 80575-80586.

[4] Zhou et al. "Cross-Level Feature aggregation Network for Polyp Segmentation." Pattern Recognition 140 (2023): 109555.

[5] Xiao et al. "CTNet: Contrastive Transformer Network for Polyp Segmentation." IEEE Transactions on Cybernetics (2024).

[6] Mei et al. "A Survey on Deep Learning for Polyp Segmentation: Techniques, Challenges and Future Trends." Visual Intelligence 3.1 (2025): 1.

[7] Zhu et al. "CRCNet: Global-Local Context and Multi-Modality Cross Attention for Polyp Segmentation." Biomedical Signal Processing and Control 83 (2023): 104593.

[8] Yu et al.. "HarDNet-CPS: Colorectal Polyp Segmentation based on Harmonic Densely United Network." Biomedical Signal Processing and Control 85 (2023): 104953.

[9] Lou et al. "CaraNet: Context Axial Reverse Attention Network for Segmentation of Small Medical Objects." Journal of Medical Imaging 10.1 (2023): 014005-014005.

[10] Ige et al. "Convsegnet: Automated Polyp Segmentation from Colonoscopy using Context Feature refinement with Multiple Convolutional Kernel Sizes." IEEE Access 11 (2023): 16142-16155.

[11] Yang et al. "CFHA-Net: A Polyp Segmentation Method with Cross-Scale Fusion Strategy and Hybrid Attention." Computers in Biology and Medicine 164 (2023): 107301.

[12] Iqbal et al. "Rethinking Encoder-Decoder Architecture using Vision Transformer for Colorectal Polyp and Surgical Instruments Segmentation." Engineering Applications of Artificial Intelligence 136 (2024): 108962.