Jurnal Kecerdasan Buatan dan Teknologi Informasi
Vol. 5, No. 2, May 2026, pp. 235-242
e-ISSN: 2964-2922, p-ISSN: 2963-6191
DOI : https://doi.org/10.69916/jkbti.v5i2.476

COLONOSCOPIC POLYP SEGMENTATION USING SEGFORMER-B0 WITH A DICE-BCE HYBRID LOSS

Ahmad Yani*[1], San Sudirman[2], M. Zulpahmi[3], Emi Suryadi[4], Bahtiar Imran[5]

[1]Information Technology, Faculty of Information and Communication Technology, University of Technology Mataram, Mataram, Indonesia
[2]Information Systems, Faculty of Information and Communication Technology, University of Technology Mataram, Mataram, Indonesia
[3]Informatics Engineering, Faculty of Information and Communication Technology, University of Technology Mataram, Mataram, Indonesia
[4,5]Computer Systems Engineering, Faculty of Information and Communication Technology, University of Technology Mataram, Mataram, Indonesia
Email: [1]m4dy45@gmail.com, [2]sansudirman43@gmail.com, [3]pahmijorge04@gmail.com, [4]emisuryadi@gmail.com, [5]bahtiarimranlombok@gmail.com

(Received: April 30, 2026; Revised: May 9, 2026; Published: May 10, 2026)

Abstract

Colorectal cancer is one of the leading causes of cancer-related deaths worldwide, with most cases originating from early lesions such as colon polyps. Early detection through colonoscopy is essential to reduce mortality rates; however, accurate polyp identification remains challenging due to variations in shape, size, texture, and illumination conditions. This study aims to implement and evaluate the SegFormer-B0 architecture combined with a Dice-BCE hybrid loss function for polyp segmentation in colonoscopy images. The study utilized the public Kvasir-SEG dataset consisting of 1,000 colonoscopy images with pixel-level annotations. The dataset was divided into 80% training data and 20% validation data. Image preprocessing included resizing to 256×256 pixels and normalization using ImageNet statistics. The model was trained for 25 epochs using the AdamW optimizer with a learning rate of 1×10^{-4}. Performance evaluation was conducted using Dice Coefficient, Intersection over Union (IoU), Sensitivity, and Specificity metrics. The experimental results demonstrated that the proposed model achieved a Dice Coefficient of 89.92%, Mean IoU of 81.90%, Sensitivity of 89.12%, and Specificity of 98.51%. The training process also showed stable convergence, supported by a training loss of 7.53% and validation loss of 23.30%. The findings indicate that the integration of SegFormer-B0 with the Dice-BCE hybrid loss effectively improves segmentation accuracy and stability while addressing class imbalance issues in colonoscopy images. Therefore, the proposed approach has strong potential to support computer-aided diagnosis systems for colorectal cancer screening.

Keywords: computer-aided diagnosis; dice-bce loss; kvasir-seg; medical image segmentation; segformer-b0.

1. INTRODUCTION

Colorectal cancer is one of the leading causes of cancer death globally, and most cases arise from early lesions such as polyps in the colon. Early detection through colonoscopy plays a crucial role in reducing mortality rates, but identifying polyps remains challenging. Variations in shape, size, texture, and lighting conditions in colonoscopy images make it difficult to distinguish polyps from normal tissue. In addition, the observation process, which relies on the experience of medical personnel, increases the risk of missed polyps, which can result in delayed diagnosis [1], [2]. With the development of deep learning, various approaches have been developed to improve the accuracy of medical image segmentation. Convolutional Neural Network (CNN)-based architectures, such as U-Net and its variants, have become the dominant method used due to their ability to extract spatial features effectively [3], [4]. Various developments, such as the use of attention mechanisms, residual connections, and dilated convolutions, have been carried out to improve segmentation performance [5], [6]. However, CNNs have limitations in capturing the global context of images, making them less than optimal in handling objects with unclear boundaries and complex structures such as polyps [7], [8]. To address these limitations, transformer-based approaches have become increasingly used in image segmentation tasks. These models utilize self-attention mechanisms to capture global relationships between pixels, resulting in more comprehensive feature representations [9], [10].

Several architectures, such as the Vision Transformer, Swin Transformer, and SegFormer, have demonstrated significant performance improvements in various medical image segmentation tasks [11], [12]. One architecture that has demonstrated strong performance is SegFormer, which combines computational efficiency with multi-scale feature extraction capabilities [13]. A lightweight variant of this model, SegFormer-B0, boasts resource efficiency,

making it suitable for implementation in cloud-based computing environments like Google Colab. However, the main problem in polyp segmentation lies not only in the model architecture but also in class imbalance, where the polyp area tends to be much smaller than the background [14], [15]. This condition causes the model to be biased toward the majority class and degrades segmentation performance. Therefore, selecting the right loss function is crucial for improving the quality of segmentation results.

The Binary Cross Entropy (BCE) function is widely used due to its stability in pixel-based optimization processes, but it is less sensitive to the fit of the segmentation area. In contrast, Dice Loss is more effective in measuring the degree of overlap between predictions and ground truth, but can experience instability in the early stages of training [16]. Combining the two in the form of a hybrid loss function (Dice-BCE) is a potential approach to overcome the weaknesses of each method and improve overall segmentation performance [17], [18], [19], [20].

Several recent studies have examined various approaches to improve polyp segmentation performance. CNN-based approaches such as U-Net and its variants remain the main baseline, but they have limitations in capturing global dependencies [6], [7]. Meanwhile, transformer-based approaches have shown significant improvements in understanding the global context of medical images [19]. Furthermore, recent research has begun to explore combinations of loss functions to improve segmentation performance [20]. However, most studies have focused on high-complexity models or have not specifically evaluated the effectiveness of lightweight transformer architectures such as SegFormer-B0 in combination with hybrid loss functions in computationally limited environments.

Based on the review, there is still a need to evaluate the effectiveness of a lightweight transformer architecture combined with a hybrid loss function in the context of polyp segmentation in colonoscopy images. The novelty in this study lies in the integrated evaluation between the SegFormer-B0 architecture and the Dice-BCE hybrid loss function for polyp segmentation in colonoscopy images, which is implemented in a cloud-based computing environment with a focus on efficiency and reproducibility. Therefore, this study aims to implement and evaluate the SegFormer-B0 model with the Dice-BCE loss function on the Kvasir-SEG dataset. Model performance evaluation is carried out using the Dice Coefficient, Intersection over Union (IoU), sensitivity, and specificity metrics.

2. RESEARCH METHODS

2.1. Research Methods

The research methodology flow includes the stages of dataset loading, pre-processing, model training, and performance evaluation. The research methodology flow can be seen in Figure 1.

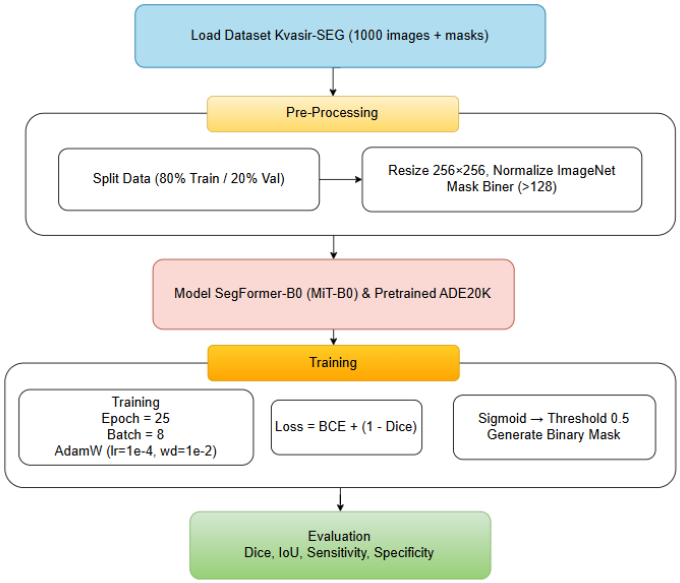

Figure 1. Research methodology flow

2.2. Dataset and Pre-processing

The study used the public Kvasir-SEG dataset, consisting of 1,000 high-resolution endoscopic images with pixel-level annotations validated by gastroenterologists. The dataset was fully anonymized and randomly split with a ratio of 80% for training and 20% for validation. All images and masks were resized to 256×256 pixels using bilinear and nearest-neighbor interpolation. Intensity normalization was performed using ImageNet mean and standard deviation statistics.

2.3. Model Architecture

The model utilizes SegFormer-B0 with a MiT-B0 backbone, generating multi-scale feature maps at four resolution levels. The decoder aggregates features through cascaded upsampling and 1×1 convolution, producing a binary output aligned to the input dimensionality. The initial weights are initialized from the ADE20K pretrained model and then adapted for the binary segmentation task by replacing the final classification head with two channels. The model architecture can be seen in Figure 2.

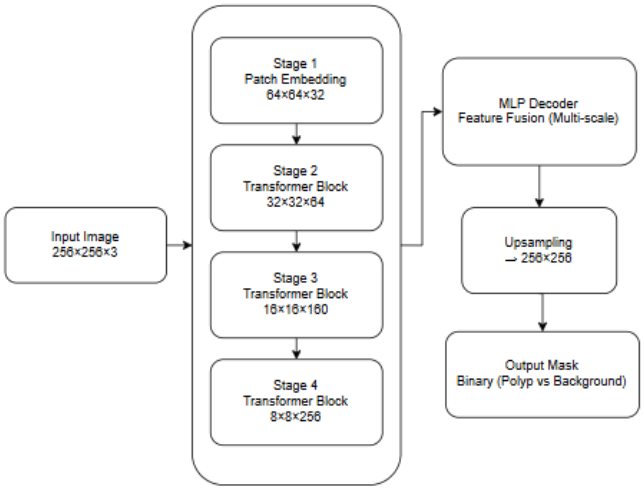

Figure 2. SegFormer model architecture

The model architecture used in this study is based on the SegFormer framework, specifically designed for efficient and accurate image segmentation. The process begins with receiving an Input Image with a resolution of 256 256 pixels with 3 color channels (RGB). The image is then processed through a Hierarchical Transformer Encoder consisting of four main stages. In Stage 1, patch embedding is performed, which produces a feature map with a resolution of 64 64 32, which is then progressively reduced in spatial resolution but deepened in dimension through Stage 2 (32 32 64), Stage 3 (16 16 160), until it reaches Stage 4 with the smallest resolution of 8 8 256. This hierarchical structure allows the model to capture detailed spatial information in the initial layers as well as broader and abstract semantic information in deeper layers. Furthermore, the features from the four stages are integrated in the MLP Decoder through the Feature Fusion (Multi-scale) mechanism. At this stage, feature representations from various scales are combined using a Multi-Layer Perceptron layer to unify local and global information without the need for traditional positional encoding. After the feature fusion process is complete, an Upsampling stage is performed to return the feature map resolution to its original size of 256 256 pixels. This architectural flow ends with the formation of an Output Mask in the form of a binary image, which precisely separates the Polyp area as the target object from the Background or normal colon tissue as the background, thus providing accurate segmentation visualization results for medical needs.

2.4. Hybrid Loss Function

The loss function is defined as the weighted sum of BCE and Dice Loss:

$$\mathcal{L} = \mathcal{L}_{\text{BCE}} + (1 - \mathcal{L}_{\text{Dice}}) \qquad (1)$$

With:

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \qquad (2)$$

$$\mathcal{L}_{\text{Dice}} = \frac{2\sum_{i=1}^{N} y_i \hat{y}_i + \epsilon}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} \hat{y}_i + \epsilon} \qquad (3)$$

Where y_i is the ground truth value, $\hat{y}_i$ is the probability of prediction, N is the number of pixels, and $\epsilon = 10^{-5}$ smoothing constant.

2.5. Training and Evaluation Protocol

The training was carried out for 25 epochs using the AdamW optimizer with a learning rate of 1×10^{-4} and weight decay 1×10^{-2} All experiments were run in PyTorch with a fixed seed value of 42 to ensure reproducibility. To objectively measure model performance, we evaluated four standard metrics based on True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) values. The following are the details of the metrics used:

a. Dice Coefficient
This metric measures the degree of similarity or overlap between the model's predicted mask and the ground truth mask (the doctor's original annotation).

$$DC = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (4)$$

b. Mean IoU
Often referred to as the Jaccard Index, this metric calculates the ratio between the intersection area (meeting) divided by the combined area between the prediction and target.

$$IoU = \frac{TP}{TP + FP + FN} \quad (5)$$

c. Sensitivity
Measures the ability of the model to detect all pixels that are actually part of a polyp.

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

d. Specificity
Measures the ability of the model to correctly identify non-polyp (healthy tissue) pixels.

$$Spesificity = \frac{TN}{TN + FP} \quad (7)$$

3. RESULTS AND DISCUSSION

3.1. Quantitative Results

The performance evaluation of the SegFormer-B0 model with the Dice-BCE hybrid loss function was performed on the validation data using the Dice Coefficient, Intersection over Union (IoU), Sensitivity, and Specificity metrics. A summary of the final evaluation results is presented in Table 1.

Table 1. Segmentation Evaluation Results

Metric	Value
Dice Coefficient	89,92%
Mean IoU	81,90%
Sensitivity	89,12%
Specificity	98,51%
Training Loss	7,53%
Validation Loss	23,30%

Based on Table 1, the model obtained a Dice Coefficient of 89.92% and a Mean IoU of 81.90%, indicating a high level of spatial agreement between the segmentation results and ground truth annotations. This indicates that the SegFormer-B0 architecture is able to effectively capture the global context of the image while preserving local details, resulting in precise polyp segmentation. A Sensitivity value of 89.12% indicates that the model has a good ability to detect polyp areas. However, there are still challenges in detecting small polyps or polyps with low contrast to the surrounding tissue. Meanwhile, a Specificity of 98.51% indicates that the model is very good at identifying non-polyp areas, thus minimizing detection errors (false positives). This capability is an important aspect in clinical applications to improve the reliability of diagnostic aid systems. In terms of the learning process, the training loss value of 7.53% and validation loss of 23.30% indicate that the model is able to learn well and has a fairly good generalization ability. The difference in loss values that is still within reasonable limits indicates that the model does not experience significant overfitting.

3.2. Training Convergence Analysis

To analyze the dynamics of the model training process, the performance at each epoch during 25 iterations is presented in Table 2.

Table 2. Training Process

epoch	train_loss	val_loss	val_dice	val_iou	val_sens	val_spec
1	77.09%	49.56%	75.69%	61.90%	91.54%	91.26%
2	38.62%	40.14%	80.52%	68.54%	72.25%	99.15%
3	27.87%	30.78%	86.47%	76.57%	82.48%	98.63%
4	22.11%	29.32%	86.11%	76.35%	82.97%	98.51%
5	19.22%	28.24%	86.57%	76.71%	91.60%	96.61%
6	17.14%	27.02%	88.03%	79.22%	86.79%	98.32%
7	16.71%	24.80%	88.67%	79.83%	89.08%	98.01%
8	14.84%	25.77%	87.58%	78.53%	90.55%	97.35%
9	12.84%	25.26%	88.84%	80.19%	87.64%	98.44%
10	12.38%	24.94%	88.76%	80.08%	87.63%	98.46%
11	11.76%	22.58%	89.60%	81.47%	90.26%	98.18%
12	11.77%	31.36%	86.09%	76.19%	94.83%	95.59%
13	11.64%	27.88%	88.81%	80.11%	85.41%	98.90%
14	11.04%	23.89%	89.62%	81.50%	90.22%	98.17%
15	9.51%	25.18%	89.48%	81.25%	89.05%	98.35%
16	8.99%	24.49%	89.37%	81.11%	86.86%	98.76%
17	11.50%	28.07%	87.91%	78.93%	90.85%	97.35%
18	10.91%	26.21%	88.93%	80.37%	87.49%	98.57%
19	10.53%	24.59%	89.33%	81.01%	89.53%	98.25%
20	9.51%	25.09%	89.29%	80.95%	89.31%	98.28%
21	9.58%	31.53%	87.84%	79.03%	84.75%	98.81%
22	9.00%	24.81%	89.50%	81.24%	87.32%	98.77%
23	7.62%	25.19%	89.71%	81.54%	87.57%	98.78%
24	7.79%	23.81%	90.10%	82.18%	88.16%	98.73%
25	7.53%	23.30%	89.92%	81.90%	89.12%	98.51%

Based on Table 2, the training process for the SegFormer-B0 model with the Dice-BCE loss shows a stable convergence pattern with significant performance improvements in the early stages. In the first epoch, the Dice Coefficient remained at 75.69%, but increased rapidly to 88.67% in the 7th epoch. This improvement indicates that the model is able to quickly learn basic feature representations from colonoscopy images. Entering the middle phase (epochs 8–15), model performance began to stabilize, with Dice values ranging from 87.58% to 89.62%. However, there were fluctuations in the validation loss, particularly in the 12th epoch, where it increased to 31.36%, accompanied by a decrease in Dice to 86.09%. This indicates temporary instability in the optimization process, likely influenced by variations in the complexity of polyp shape and texture. In the final phase (epochs 16–25), the model entered the convergence stage with relatively stable and high performance. The Dice Coefficient peaked at 90.10% in the 24th epoch, with an IoU of 82.18%. After that, in the 25th epoch, performance dropped slightly to 89.92%, but remained at the optimal level. Performance fluctuations were also observed in the 21st epoch, where validation loss increased to 31.53%, and Dice decreased to 87.84%. However, in subsequent epochs, performance improved again, indicating that this phenomenon was a local variation in the training process and not an indication of significant overfitting.

Overall, the training convergence pattern can be categorized into three main phases:

a. Epochs 1–7: rapid learning phase
b. Epochs 8–15: stabilization phase
c. Epochs 16–25: convergence phase

Furthermore, the consistent downward trend in training loss from 77.09% to 7.53% indicates that the model can minimize prediction errors during training. Meanwhile, the validation loss tends to fluctuate but remains within a stable range, with a final value of 23.30%, indicating the model's good generalization ability. The use of the Dice-BCE hybrid loss function significantly contributed to the convergence stability. Dice Loss helps optimize the fit of the segmentation areas (overlap), while Binary Cross Entropy (BCE) maintains gradient stability at the pixel level. The combination of the two allows the model to achieve a balance between spatial accuracy and learning stability, particularly in addressing class imbalance between polyp areas and the background.

3.3. Segmentation Results Analysis (Qualitative)

To visually evaluate the model performance, several examples of segmentation results are shown in Figure 3, which consist of input images, ground truth, and the predicted results of the SegFormer-B0 model with Dice-BCE loss.

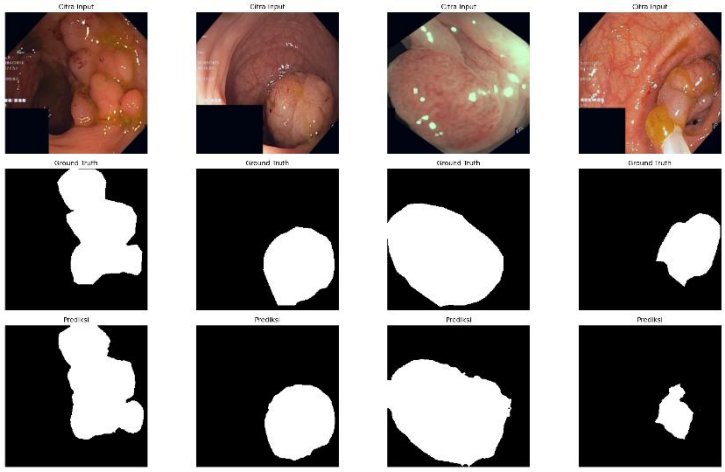

Figure 3. Results Of Polyp Segmentation On Colonoscopy Images

Based on Figure 3, the model is generally able to produce segmentations that are close to the ground truth annotations. In most samples, the polyp area is successfully identified with a consistent shape and size, demonstrating the model's ability to accurately capture the spatial structure of the object. The first and second columns show almost complete overlap with the ground truth, especially for polyps with clear boundaries and relatively high contrast compared to the surrounding tissue. This indicates that the model is able to effectively recognize the main morphological features of polyps. In the third column, the model still manages to follow the shape of relatively large and homogeneous polyps, although there are slight differences at the edges. This difference is likely due to the smooth and ambiguous nature of polyp boundaries, which is a common challenge in medical image segmentation. Meanwhile, the fourth column shows slight discrepancies in the prediction results, particularly at the bottom edge of the object, where the model produces segmentations that do not fully match the ground truth. This indicates that the model still has limitations in handling areas with low contrast or lighting noise. Overall, these visual results reinforce previous quantitative findings, where the model demonstrated high performance in polyp segmentation with a good level of accuracy. The combination of the SegFormer-B0 architecture, capable of capturing global context, and the Dice-BCE loss function that optimizes the suitability of the segmentation area has proven effective in producing precise and stable predictions.

3.4. Discussion

The findings of this study demonstrate that the implementation of the SegFormer-B0 architecture combined with the Dice-BCE hybrid loss function is effective for polyp segmentation in colonoscopy images. The quantitative results show high performance, achieving a Dice Coefficient of 89.92% and Mean IoU of 81.90%, indicating a strong spatial agreement between predicted segmentation masks and ground truth annotations. These results confirm that transformer-based architectures can effectively capture both local and global contextual information in medical images. Unlike conventional CNN-based approaches, which mainly focus on local feature extraction, SegFormer-B0 utilizes hierarchical transformer encoding to model long-range dependencies between image regions, enabling more accurate boundary delineation and contextual understanding of polyp structures. This capability is particularly important in colonoscopy images, where polyps often exhibit irregular shapes, blurred edges, and heterogeneous textures.

The high sensitivity score of 89.12% indicates that the proposed model can effectively detect most polyp regions, which is essential for minimizing missed diagnoses in clinical applications. Missed polyp detection may delay treatment and increase the risk of colorectal cancer progression. However, several segmentation limitations remain, especially for small polyps and low-contrast regions that blend with surrounding tissue. These challenges were also reflected in the qualitative evaluation, where slight discrepancies appeared at object boundaries or under difficult lighting conditions. Despite these limitations, the specificity value of 98.51% demonstrates that the model has excellent capability in distinguishing non-polyp tissue, reducing false-positive segmentation that could potentially interfere with clinical interpretation. High specificity is particularly important in supporting reliable computer-aided diagnosis systems to reduce unnecessary medical interventions.

The convergence analysis further confirms the effectiveness of the proposed approach. The training process demonstrated a stable learning pattern, categorized into rapid learning, stabilization, and convergence phases. The significant decrease in training loss from 77.09% to 7.53% indicates successful optimization during training. Although fluctuations in validation loss occurred at several epochs, such as epochs 12 and 21, the model quickly recovered and maintained high performance, suggesting good generalization capability rather than overfitting. This stability can be attributed to the Dice-BCE hybrid loss function, which combines the strengths of Binary Cross Entropy and Dice Loss. BCE contributes to pixel-level optimization stability, while Dice Loss improves overlap accuracy between predictions and annotations. This combination is particularly effective for handling class imbalance problems, where the polyp area occupies a significantly smaller portion of the image than the background.

Compared to previous studies that mainly emphasized CNN-based architectures or computationally expensive transformer models, this study highlights the effectiveness of a lightweight transformer architecture in a resource-constrained environment. SegFormer-B0 demonstrated strong segmentation performance while maintaining computational efficiency, making it suitable for implementation in cloud-based environments such as Google Colab. Therefore, the proposed approach not only contributes to improving segmentation accuracy but also offers practical advantages in terms of efficiency, reproducibility, and accessibility for future medical imaging research and computer-aided diagnostic systems.

4. CONCLUSION

This study successfully implemented and evaluated the SegFormer-B0 model combined with the Dice-BCE hybrid loss function for polyp segmentation in colonoscopy images using the Kvasir-SEG dataset. The experimental results demonstrated that the proposed approach achieved high segmentation performance, with a Dice Coefficient of 89.92%, Mean IoU of 81.90%, Sensitivity of 89.12%, and Specificity of 98.51%. These results indicate that the model is capable of accurately identifying polyp regions while effectively minimizing false-positive predictions. The SegFormer-B0 architecture proved effective in capturing both local and global contextual information through hierarchical transformer encoding, enabling more precise segmentation of complex polyp structures. The training convergence analysis also showed stable optimization performance, where the model successfully minimized training loss while maintaining good generalization ability. The use of the Dice-BCE hybrid loss function significantly contributed to overcoming class imbalance issues and improving segmentation overlap accuracy. Although minor limitations were observed in detecting small or low-contrast polyps, the overall results indicate that the proposed model has strong potential to support computer-aided diagnosis systems for colorectal cancer screening.

In future research, improvements can be made by exploring more advanced transformer variants, applying data augmentation strategies, and integrating attention refinement modules to improve segmentation accuracy for difficult cases. Further studies may also compare different hybrid loss functions, evaluate model robustness on multiple medical datasets, and optimize deployment in real-time clinical environments to support more efficient and accurate colonoscopy diagnosis.

REFERENCES

[1] X. Liu and others, “Colorectal Polyp Segmentation Based on Deep Learning Methods: A Systematic Review,” J. Imaging, vol. 11, no. 9, p. 293, 2025, doi: 10.3390/jimaging11090293.

[2] E. Goceri, “Polyp Segmentation Using a Hybrid Vision Transformer and a Hybrid Loss Function,” Journal of Imaging Informatics in Medicine, 2024, doi: 10.1007/s10278-024-00915-7.

[3] K. He and others, “Deep Residual Learning for Image Recognition,” CVPR, 2016, doi: 10.1109/CVPR.2016.90.

[4] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” May 2015, [Online]. Available: http://arxiv.org/abs/1505.04597

[5] O. Oktay and others, “Attention U-Net: Learning Where to Look for the Pancreas,” arXiv preprint arXiv:1804.03999, 2018.

[6] S. Kumar and others, “Dilated U-Net-Seg for Polyp Segmentation,” Biomed. Signal Process. Control, vol. 92, p. 105975, 2024, doi: 10.1016/j.bspc.2024.105975.

[7] J. Mei and others, “A Survey on Deep Learning for Polyp Segmentation,” Visual Intelligence, 2025, doi: 10.1007/s44267-024-00071-w.

[8] Z. Zhou and others, “UNet++: A Nested U-Net Architecture for Medical Image Segmentation,” Deep Learning in Medical Image Analysis, 2018, doi: 10.1007/978-3-030-00889-5_1.

[9] A. Dosovitskiy and others, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” ICLR, 2021.

[10] Z. Liu and others, “Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows,” ICCV, 2021, doi: 10.1109/ICCV48922.2021.00986.

[11] H. Cao and others, “Transformer-Based Medical Image Segmentation: A Review,” Med. Image Anal., vol. 85, p. 102762, 2023, doi: 10.1016/j.media.2023.102762.

[12] D. Liu and others, "NA-SegFormer: A Multi-Level Transformer Model for Medical Image Segmentation," Sci. Rep., vol. 14, p. 21567, 2024, doi: 10.1038/s41598-024-71523-4.

[13] E. Xie and others, "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers," NeurIPS, 2021.

[14] H. Cao et al., "Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation," May 2021, [Online]. Available: http://arxiv.org/abs/2105.05537

[15] S. Taghanaki and others, "Combo Loss: Handling Input and Output Imbalance in Multi-Organ Segmentation," Computerized Medical Imaging and Graphics, vol. 85, p. 101744, 2021, doi: 10.1016/j.compmedimag.2020.101744