BiGAMR-Net: Bidirectional Gated Attention and Multi-scale Residual Network for Polyp Segmentation

Liuyi Yang, Shao-Chi Pao, Xiaoxing Yang, Rui Li, Lixin Liang(✉), and Bingding Huang(✉)

College of Big Data and Internet, Shenzhen Technology University, Shenzhen, China
{Lianglixin,huangbingding}@sztu.edu.cn

Abstract. Colorectal cancer (CRC) screening critically relies on precise polyp segmentation. However, existing methodologies often struggle with lesion heterogeneity, indistinct boundaries, and elevated rates of missed diagnoses. In response to these challenges, this study proposes BiGAMR-Net, a novel deep learning framework that synergistically combines Bidirectional Gated Attention with Multi-scale Residual Learning to enhance segmentation performance. The proposed architecture comprises three core components: (1) a channel-wise Squeeze-and-Excitation attention mechanism designed to suppress irrelevant noise and emphasize discriminative features; (2) Multi-scale Residual (MR) blocks, which utilize parallel atrous convolutions to capture diverse receptive fields while maintaining stable gradient propagation; and (3) A Bidirectional Fusion Module (BFM) facilitates dynamic, cross-scale feature integration to align spatial and semantic information. Comprehensive evaluations across five benchmark datasets (NBI, Kvasir-SEG, CVC-ClinicDB, CVC-ColonDB, ETIS) demonstrate that BiGAMR-Net achieves state-of-the-art performance, with mean Dice scores of 0.956 on Kvasir-SEG and 0.958 on ETIS exceeding the performance of competitive models such as nnUNet and DuckNet by up to 4.4%. The framework performs robustly in challenging scenarios, including narrow-band imaging and complex mucosal patterns. Ablation studies further validate the effectiveness of the BFM in refining lesion boundaries (3.8% Dice improvement) and the MR module in detecting small lesions (2.9% IoU gain). These findings underscore the clinical promise of BiGAMR-Net in reducing diagnostic oversights and improving the reliability of CRC screening. Future research will focus on enabling real-time inference, enhancing cross-modality generalization, and conducting large-scale clinical trials to support broader clinical translation.

Keywords: Polyp segmentation · Bidirectional attention · Multi-scale residual learning · Colonoscopy · Convolutional neural network (CNN)

L. Yang and S.-C. Pao–contributed equally to this work.

D.-S. Huang et al. (Eds.): ICIC 2025, LNBI 15866, pp. 165–176, 2025.
https://doi.org/10.1007/978-981-95-0027-7_15

1 Introduction

Colorectal cancer (CRC) remains a significant global health burden and is one of the leading causes of cancer-related mortality, accounting for a substantial proportion of cancer deaths annually [1]. Most CRC cases originate from adenomatous polyps, underscoring the critical importance of early detection and removal in effective prevention and treatment [2]. Colonoscopy is widely regarded as the gold standard for detecting and excising polyps before malignant transformation [3]. However, the accurate identification and segmentation of polyps during colonoscopy pose substantial challenges due to the considerable polyp morphology, size, and texture variability. Furthermore, polyp detection is highly dependent on the subjective assessment of endoscopists, leading to a considerable rate of missed or misdiagnosed polyps, which can adversely affect patient outcomes.

Image segmentation, a fundamental task in computer vision, aims to assign each pixel in an image to a specific category, thereby facilitating precise object delineation [4]. In recent years, deep learning-based image segmentation techniques have demonstrated remarkable success across various domains, including medical imaging [5]. Salient object detection, which focuses on identifying and segmenting the most visually distinctive objects within an image, has inspired similar methodologies in medical image analysis [6]. Specifically, automated polyp segmentation has emerged as a promising approach to improve detection efficiency and accuracy while reducing the cognitive burden on clinicians [7].

Despite the notable progress in deep learning-driven polyp segmentation, several challenges persist. First, the substantial heterogeneity in the polyp morphology complicates the generalization of models across diverse datasets [8]. Second, polyps often exhibit blurred or indistinct boundaries, making their differentiation from surrounding normal tissue difficult [9]. Third, small polyps are particularly prone to be overlooked, especially in complex endoscopic environments where they might be obscured by mucosal folds [10].

To overcome these challenges, this study proposes BiGAMR-Net, a Bidirectional Gated Attention and Multi-scale Residual Network, as an advanced deep learning framework designed to facilitate accurate and automated polyp segmentation. The proposed architecture synergistically integrates three core innovations: (1) Bidirectional Gated (BiG) mechanisms that enable dynamic fusion of hierarchical features through our BFM module, effectively bridging low-level spatial details with high-level semantic information; (2) Attention (A) mechanisms implemented via squeeze-and-excitation (SE) blocks that perform channel-wise recalibration, where adaptive feature refinement amplifies discriminative channel responses while suppressing noisy activations; (3) Multi-scale Residual (MR) structures combining improved atrous convolutions with stabilized residual connections, achieving comprehensive receptive field coverage while maintaining gradient flow integrity. Extensive experimental evaluations on multiple benchmark datasets demonstrate the proposed approach's superiority over existing state-of-the-art methods, highlighting its potential for improving clinical outcomes in colorectal cancer screening and diagnosis.

2 Methodology

In this section, we introduce our BiGAMR-Net's architecture and the details of its constituent modules. We first provide an overview of the network structure, followed by a detailed description of its fundamental components, including SEMR and BFM.

2.1 SEMR Block

The proposed SEMR block is designed to improve multiscale feature extraction and enhance segmentation accuracy, particularly for medical image segmentation tasks. To address the limitations of conventional residual and atrous convolutional modules, the SEMR block incorporates three key enhancements: an improved atrous convolution block, an optimized residual connection, and the integration of Squeeze-and-Excitation (SE) attention mechanisms. Together, these modifications enable more effective feature representation and segmentation performance.

As illustrated in 错误未找到引用源。, the SEMR block consists of a parallel atrous convolution module, a residual connection block, and an SE attention module. The enhanced atrous block replaces the sequential dilated convolutions of traditional architectures with parallel dilated convolutions (dilation rates: 1, 2, 3), which are subsequently fused through concatenation and a 1 × 1 convolution. This parallel structure facilitates extracting fine-grained details and global contextual information, thus improving the network's ability to capture critical features at multiple scales. The residual connection block stabilizes gradient propagation while integrating 1 × 1 convolutions to align channel dimensions, ensuring compatibility across varying feature map resolutions (Fig. 1).

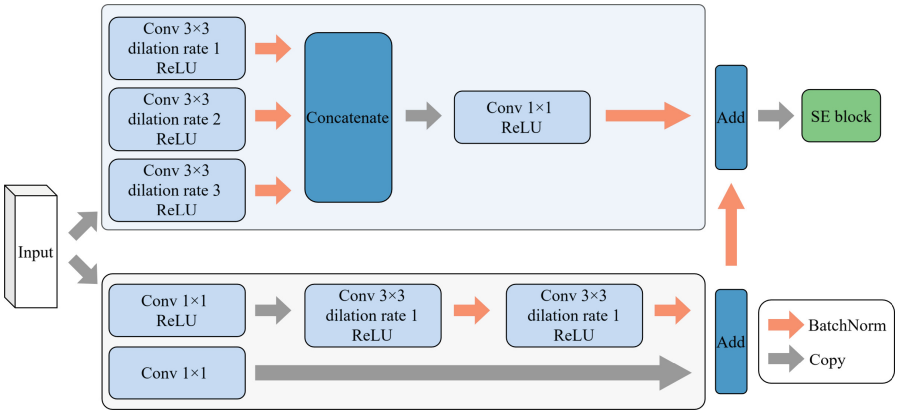

Fig. 1. The architecture of the proposed SEMR block.

To further refine the extracted features, the SEMR block employs an SE attention mechanism that adaptively recalibrates channel-wise feature responses. This mechanism significantly enhances the contrast between foreground and background regions, particularly benefiting polyp segmentation by emphasizing informative features and suppressing less relevant ones.

SEMR block effectively balances local and global feature extraction while maintaining computational efficiency. Combining parallel atrous convolutions, enhanced residual connections, and adaptive attention mechanisms improves segmentation accuracy and robustness, particularly in complex medical imaging scenarios.

2.2 BFM

The Bidirectional Fusion Module (BFM) is designed to enhance multi-scale feature aggregation by effectively integrating low-level and high-level features. It improves segmentation accuracy by addressing the limitations of conventional skip connections and enhancing cross-scale information flow.

As shown in Fig. 2, BFM consists of a bidirectional gating mechanism that adaptively regulates the exchange of features between low-level and high-level inputs. Unlike the SBA module [22], which relies solely on element-wise weighting, BFM introduces two gating functions (A_L and A_H) to refine feature fusion selectively. These gating mechanisms act as feature selectors, allowing the model to enhance relevant information while suppressing redundancy.

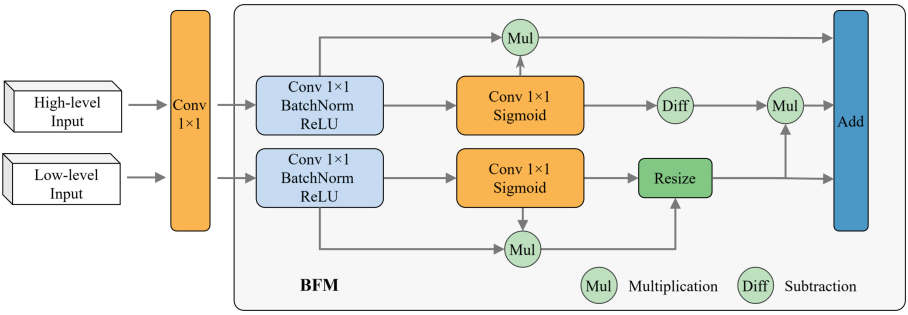

Fig. 2. The architecture of the proposed BFM.

BFM modulates the information using A_L and integrates high-level guidance via an upsampled high-level representation for low-level features. This effectively preserves fine-grained details, such as polyp boundaries and textures. In contrast, BFM incorporates spatially aligned low-level information via A_H for high-level features, refining global semantic understanding while maintaining structural consistency.

The bidirectional design facilitates smoother cross-scale feature interaction and retains essential details throughout the segmentation process. By leveraging multiplicative gating functions, BFM provides a more stable gradient flow, improving training convergence and reducing optimization difficulties.

Finally, the fused features from both branches are concatenated and processed through a 3×3 convolutional layer, refining the final representation before segmentation prediction. This adaptive fusion strategy significantly improves segmentation performance, particularly in challenging medical imaging tasks, where precise boundary delineation is crucial.

3 Experiments

3.1 Datasets

To comprehensively evaluate the proposed BiGAMR Model, extensive experiments were conducted on five benchmark datasets: NBI (Narrow Band Imaging) dataset, Kvasir-SEG[12], CVC-ClinicDB dataset [13], CVC-ColonDB dataset [14], and ETIS dataset [15]. 错误未找到引用源。 presents the distribution of the dataset across the training, validation, and test sets, along with the pixel dimensions of the images.

The NBI dataset consists of 1,000 images, with 700 allocated for training, 100 for validation, and 200 for testing, all at a resolution of 1280 × 1080. This dataset is derived from the ISICDM2024 Challenge Task 1 on NBI Polyp Segmentation. Similarly, the Kvasir-SEG dataset comprises 1,000 images, divided into 800 for training, 100 for validation, and 100 for testing, with image resolutions ranging from 332 × 487 to 1920 × 1072. The partitioning of both datasets follows the official splits provided by their respective sources.

For the remaining datasets, an 8:1:1 ratio was applied for random partitioning. The CVC-ClinicDB dataset contains 612 images, with 489 assigned to training, 62 to validation, and 61 to testing, at a resolution 384 × 288. The CVC-ColonDB dataset consists of 380 images, divided into 304 for training, 38 for validation, and 38 for testing, with a resolution of 574 × 500. Lastly, the ETIS-LaribPolypDB dataset includes 196 images, with 156 designated for training, 21 for validation, and 19 for testing, at a resolution 1255 × 966.

3.2 Implementation Details

All experiments used the same dataset to ensure a fair and consistent evaluation. The model was implemented using the TensorFlow framework and deployed on multiple NVIDIA A800 GPUs, each with 80GB of memory. Multi-GPU training was managed through TensorFlow's MirroredStrategy, with the number of available GPUs verified at runtime. Data augmentation used Albumentations to enhance model generalization, improve robustness, mitigate overfitting, and optimize performance on previously unseen data.

All input images were resized to 352 × 352 pixels, and the model was trained for 300 epochs with a mini-batch size of 8. A polynomial decay learning rate schedule was employed, starting at 1×10^{-4} and progressively decreasing to 1×10^{-6} over 1,000 steps, with a decay power of 0.2. The AdamW optimizer from TensorFlow Addons was utilized alongside this learning rate strategy. The Dice loss function was adopted to maximize the overlap between predicted segmentation and ground truth, effectively mitigating class imbalance, particularly given the relatively small proportion of polyps within the images. This process is mathematically formulated in (1).

$$DiceLoss = 1 - Dice \tag{1}$$

3.3 Evaluation Metrics

Four widely adopted metrics comprehensively evaluate the proposed polyp segmentation model's performance: mean Dice Coefficient (mDice) [11], mean Intersection over Union (mIoU), Precision, and Recall. These metrics quantitatively assess segmentation accuracy by measuring the degree of overlap between predicted masks and ground truth annotations. The Dice Coefficient and IoU are key indicators of overall segmentation similarity, capturing spatial agreement and segmentation completeness. Meanwhile, Precision and Recall provide insights into the model's capability to detect polyp regions while accurately minimizing false positives and negatives. Collectively, these metrics facilitate a rigorous and holistic assessment of segmentation effectiveness.

4 Result

4.1 Qualitative Results

In this study, we conducted a series of experiments utilizing five datasets, including NBI, Kvasir-SEG, CVC-ClinicDB, CVC-ColonDB, and ETIS-LaribPolypDB, to evaluate the learning efficacy of BiGAMR-Net. We systematically compared the segmentation performance of BiGAMR-Net with that of previously reported state-of-the-art (SOTA) models using a range of evaluation metrics.

For this comparative analysis, we incorporated a total of six baseline models from the field of polyp segmentation, namely U-Net [16], UNet++ [17], ResUNet [19], ResUNet++ [18], nnUNet [20], and DuckNet [21]. To ensure a rigorous evaluation, all models were trained and assessed on the training, validation, and test partitions derived from the same dataset. The results of this performance comparison are presented in Table 2 and Table 3. It is noteworthy that, due to the random partitioning of the dataset and the potential for data leakage, obtaining performance metrics that accurately align with those reported in the original literature presents considerable challenges despite the availability of experimental code provided by the authors.

Based on the results presented in Table 2 and Table 3, the experimental findings indicate that the proposed model achieves state-of-the-art performance in polyp segmentation across four benchmark datasets, consistently surpassing existing methodologies in all evaluation metrics. Compared to widely adopted architectures, the proposed approach demonstrates superior generalization capabilities, achieving the highest mDice and mIoU scores across relatively straightforward datasets (CVC-ClinicDB, Kvasir) and more complex ones (CVC-ColonDB, ETIS). Notably, while nnUNet and DuckNet exhibit competitive performance, their segmentation accuracy remains marginally lower than that of the proposed model. Moreover, the performance of conventional models such as ResUNet and ResUNet++ deteriorates significantly on complex datasets, underscoring the necessity of advanced architectural designs. The pronounced performance decline observed across all models on the ETIS dataset highlights its challenging nature; however, the proposed model maintains the highest accuracy, demonstrating its robustness (Table 1).

Concerning the mDice metric, BiGAMR-Net surpasses the second-best-performing DuckNet model by 4.1% and 2.4% on the NBI and Kvasir datasets, respectively, and

Table 1. Dataset distribution and image resolutions.

Dataset	Train	Validation	Test	Resolution
NBI	700	100	200	1280 × 1080
Kvasir-SEG	800	100	100	332 × 487 to 1920 × 1072
CVC-ClinicDB	489	62	61	384 × 288
CVC-ColonDB	304	38	38	574 × 500
ETIS	156	21	19	1255 × 966

Table 2. Polyp segmentation performance comparison on CVC-ClinicDB and CVC-ColonDB.

Model	CVC-ClinicDB				CVC-ColonDB			
	mDice	mIoU	mPrec	mRec	mDice	mIoU	mPrec	mRec
U-Net	0.868	0.795	0.910	0.850	0.793	0.721	0.839	0.765
U-Net++	0.926	0.862	0.939	0.915	0.933	0.874	0.951	0.915
ResUNet	0.726	0.613	0.844	0.717	0.637	0.504	0.623	0.695
nnUNet	0.948	0.905	0.958	0.944	0.931	0.881	0.934	0.936
ResUNet++	0.839	0.766	0.873	0.873	0.780	0.693	0.874	0.791
DuckNet	0.940	0.887	0.916	0.965	0.915	0.843	0.923	0.906
Ours	**0.968**	**0.938**	**0.970**	**0.966**	**0.958**	**0.920**	**0.964**	**0.953**

Table 3. Polyp segmentation performance comparison on Kvasir and ETIS datasets.

Model	Kvasir Dataset				ETIS Dataset			
	mDice	mIoU	mPrec	mRec	mDice	mIoU	mPrec	mRec
U-Net	0.838	0.744	0.871	0.838	0.503	0.426	0.665	0.469
U-Net++	0.887	0.801	0.907	0.875	0.522	0.387	0.911	0.407
ResUNet	0.648	0.510	0.672	0.692	0.369	0.261	0.580	0.365
nnUNet	0.929	0.883	0.952	0.924	0.812	0.750	0.924	0.769
ResUNet++	0.343	0.273	0.705	0.329	0.524	0.450	0.604	0.518
DuckNet	0.933	0.874	0.933	0.932	0.737	0.584	0.670	0.818
Ours	**0.956**	**0.917**	**0.959**	**0.954**	**0.958**	**0.920**	**0.957**	**0.960**

outperforms both DuckNet and nnUNet by 4.4% and 2.7% on the CVC-ColonDB dataset. These results validate the model's stability and generalization capability.

Figure 3 presents line graphs illustrating the mDice and mIoU performance of the above models on the entire dataset. These graphs demonstrate the performance variations among the different models across various datasets, with the proposed method consistently achieving high scores.

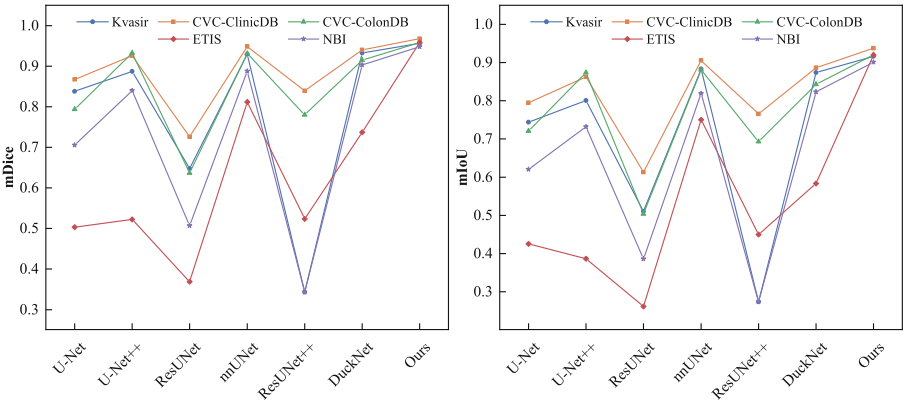

Fig. 3. Compare the Dice and IoU metrics across all models evaluated on the entire dataset.

The segmentation results for a challenging NBI task, comparing the proposed BiGAMR-Net with conventional deep-learning models, are presented in Fig. 4. The distinct imaging characteristics of NBI enhance mucosal microvascular structures and reduce the color contrast between polyps and the surrounding tissue, thereby limiting the effectiveness of color-based segmentation approaches. Additionally, lighting artifacts, including reflections and mucus-induced highlights, introduce interference that can degrade the performance of segmentation methods reliant on color and texture cues. Furthermore, the complex intestinal background, characterized by pronounced folds and uneven textures, increases the risk of false segmentation.

Our proposed model exhibits robust segmentation performance across a diverse range of cases. Nevertheless, some images with high overall segmentation accuracy, particularly those shown in the second, third, and sixth rows, still pose additional challenges. In these cases, the ground truth masks of the polyps display inconsistent contrast levels and blurred boundary definitions, impeding the model's ability to delineate polyp contours accurately. Such inconsistencies in annotation introduce ambiguity into the learning process, adversely affecting segmentation performance. Pronounced local contrast variations disrupt the model's capacity to extract stable boundary features, resulting in incomplete edge segmentation, partial omission of polyp regions, and mis-segmentation or missed detections.

Furthermore, uneven illumination and noise interference exacerbate these challenges, leading to incorrect boundary predictions in certain areas. Therefore, enhancing the model's boundary detection capability in regions with contrast inhomogeneity is crucial for optimizing segmentation performance. The local contrast variance values corresponding to the six samples in Fig. 4 are presented in Fig. 5, and the corresponding

formulas are shown in (2) and (3).

$$\sigma_c^2 = \frac{1}{M}\sum_{j=1}^{M}\left(L_j - \overline{L}\right)^2 \tag{2}$$

$$\overline{L} = \frac{1}{M}\sum_{j=1}^{M} L_j \tag{3}$$

The contrast variance, denoted as σ_c^2, is computed using the Laplacian operator, which quantifies intensity variations within the masked region of a grayscale image. Specifically, L_j represents the Laplacian-transformed grayscale intensity at pixel j within the mask region, while $\overline{L}$ denotes the mean Laplacian value across the masked area. Additionally, M corresponds to the total number of pixels within the mask.

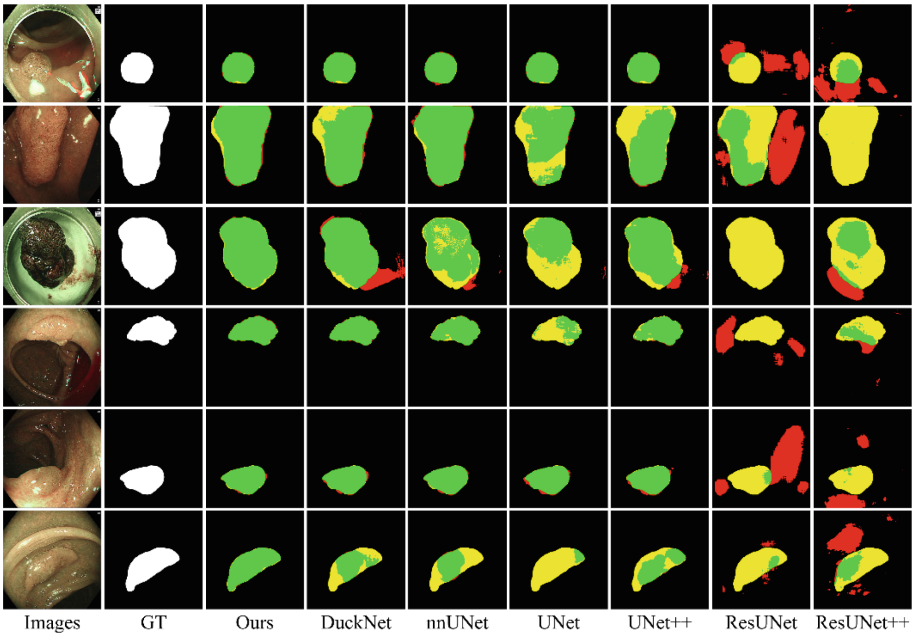

Fig. 4. Comparison of the proposed model with various segmentation methods on the NBI polyp dataset. Green, red, and yellow regions indicate true positives, false positives, and false negatives, respectively.

4.2 Ablation Study

To demonstrate the effectiveness of the SE-AR and BFM modules, we conduct an ablation study to evaluate the performance of BIGAMR-Net. For comparative analysis, we substitute the primary module with a standard convolutional block comprising a 3×3 convolution, batch normalization, and ReLU activation. The ablation study is performed

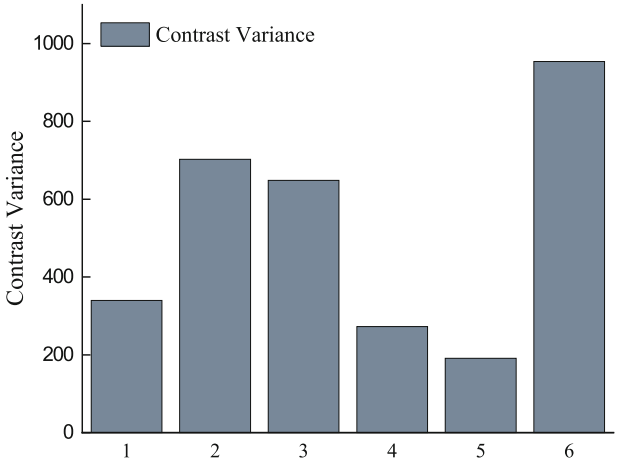

Fig. 5. Local contrast variance values for the six samples.

on the NBI and CVC-ClinicDB datasets. As presented in Table 4, the experimental results indicate that the AMR module significantly enhances the model's sensitivity to small lesions through multi-scale feature fusion, yielding an improvement of 1.9% in mean Recall (mRec) on the NBI dataset. Furthermore, the bidirectional gating mechanism of BiGAMR-Net further refines segmentation accuracy, particularly in delineating complex boundaries, resulting in a 1.0% increase in mIoU on the ClinicDB dataset. The cross-dataset evaluation confirms the model's robustness and generalization capability across narrow-band imaging (NBI) and conventional endoscopy (ClinicDB).

Table 4. Quantitative results for ablation studies on NBI and CVC-ClinicDB datasets.

Method	NBI				CVC-ClinicDB			
	mDice	mIoU	mPrec	mRec	mDice	mIoU	mPrec	mRec
CNN	0.930	0.869	0.942	0.918	0.955	0.914	0.960	0.950
AMR-Net	0.940	0.887	0.943	0.937	0.963	0.928	0.961	0.964
BiGAMR-Net	**0.948**	**0.901**	**0.951**	**0.944**	**0.968**	**0.938**	**0.970**	**0.966**

5 Conclusion

In this paper, we propose BiGAMR-Net, a polyp segmentation framework that combines bidirectional gated attention and multi-scale residual modules to enhance feature representation. Extensive experiments on five benchmarks demonstrate state-of-the-art performance, especially in cases with blurred boundaries. Ablation studies confirm the effectiveness of BFM and MR in improving accuracy. BiGAMR-Net offers a reliable

tool to reduce missed detections and assist endoscopic screening. Future work will focus on real-time applications, cross-modality adaptation, and broader clinical validation.

Acknowledgments. We sincerely thank the ISICDM2024 organizing committee for providing the NBI dataset from Challenge Project 1, which has greatly supported this research.

Funding. This work was supported by the Shenzhen Science and Technology Program (No. KJZD20240903095605007), the Higher Education Stability Support Program General Project (No. 20220715114836001), and the Natural Science Foundation of Top Talent of SZTU (No. GDRC202213).

References

1. Sung, H., et al.: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. **71**(3), 209–249 (2021)
2. Favoriti, P., Carbone, G., Greco, M., Pirozzi, F., Pirozzi, R.E.M., Corcione, F.: Worldwide burden of colorectal cancer: a review. Updates Surg. **68**, 7–11 (2016)
3. Siegel, R.L., Miller, K.D., Jemal, A.: Cancer statistics, 2019. CA Cancer J. Clin. **69**(1), 7–34 (2019)
4. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431--3440 (2015)
5. Litjens, G., Kooi, T., Bejnordi, B.E., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
6. Borji, A., Cheng, M.M., Hou, Q., Jiang, H., Li, J.: Salient object detection: a survey. Comput. Vis. Media **5**, 117–150 (2019)
7. Jha, D., Ali, S., Tomar, N.K., et al.: Real-time polyp detection, localization and segmentation in colonoscopy using deep learning. IEEE Access **9**, 40496–40510 (2021)
8. Akbari, M., Mohrekesh, M., Nasr-Esfahani, E., et al.: Polyp segmentation in colonoscopy images using the fully convolutional network. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 69–72 (2018)
9. Fan, D.P., Ji, G.P., Zhou, T., et al.: PraNet: parallel reverse attention network for polyp segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 263–273 (2020)
10. Zhang, Y., Liu, H., Hu, Q.: TransFuse: fusing transformers and CNNs for medical image segmentation. In: MICCAI 2021: 24th International Conference, pp. 14–24 (2021)
11. Milletari, F., Navab, N., Ahmadi, S.A.: V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision (3DV), pp. 565–571 (2016)
12. Jha, D., Smedsrud, P.H., Riegler, M.A., et al.: Kvasir-SEG: a segmented polyp dataset. In: MultiMedia Modeling: 26th International Conference, MMM 2020, pp. 451–462 (2020)
13. Bernal, J., S\'anchez, F.J., Fern\'andez-Esparrach, G., et al.: WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians. Comput. Med. Imaging Graph. **43**, 99–111 (2015)
14. Tajbakhsh, N., Gurudu, S.R., Liang, J.: Automated polyp detection in colonoscopy videos using shape and context information. IEEE Trans. Med. Imaging **35**(2), 630–644 (2015)

15. Silva, J., Histace, A., Romain, O., Dray, X., Granado, B.: Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. Int. J. Comput. Assist. Radiol. Surg. **9**, 283–293 (2014)
16. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: MICCAI 2015: 18th International Conference, pp. 234–241 (2015)
17. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested u-net architecture for medical image segmentation. In: DLMIA and ML-CDS 2018, MICCAI, pp. 3–11 (2018)
18. Jha, D., Smedsrud, P.H., Riegler, M.A., et al.: Resunet++: an advanced architecture for medical image segmentation. In: 2019 IEEE International Symposium on Multimedia (ISM), pp. 225–2255 (2019)
19. Zhang, Z., Liu, Q., Wang, Y.: Road extraction by deep residual U-net. IEEE Geosci. Remote Sens. Lett. **15**(5), 749–753 (2018)
20. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: NnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods **18**(2), 203–211 (2021)
21. Dumitru, R.G., Peteleaza, D., Craciun, C.: Using DUCK-net for polyp image segmentation. Sci. Rep. **13**(1), 9803 (2023)
22. Tang, F., et al.: DuAT: dual-aggregation transformer network for medical image segmentation. In: Chinese Conference on Pattern Recognition and Computer Vision (PRCV), pp. 343–356. Springer, Cham (2023)