ICIT'24 Presentation

Focusing the View: Enhancing U-Net with Convolutional Block Attention for Superior Medical Image Segmentation

Nhu-Tai Do[1], Dat Nguyen Khanh[2], Tram-Tran Nguyen-Quynh[2], Quoc-Huy Nguyen[1]

[1]Saigon University, Vietnam
[2]Ho Chi Minh City University of Foreign Language-Information Technology, Vietnam

September 2024

Agenda

1. INTRODUCTION

Our Problem:

- **Input:** High-resolution colonoscopic images
- **Output:** Binary masks that delineate the exact boundaries of polyps

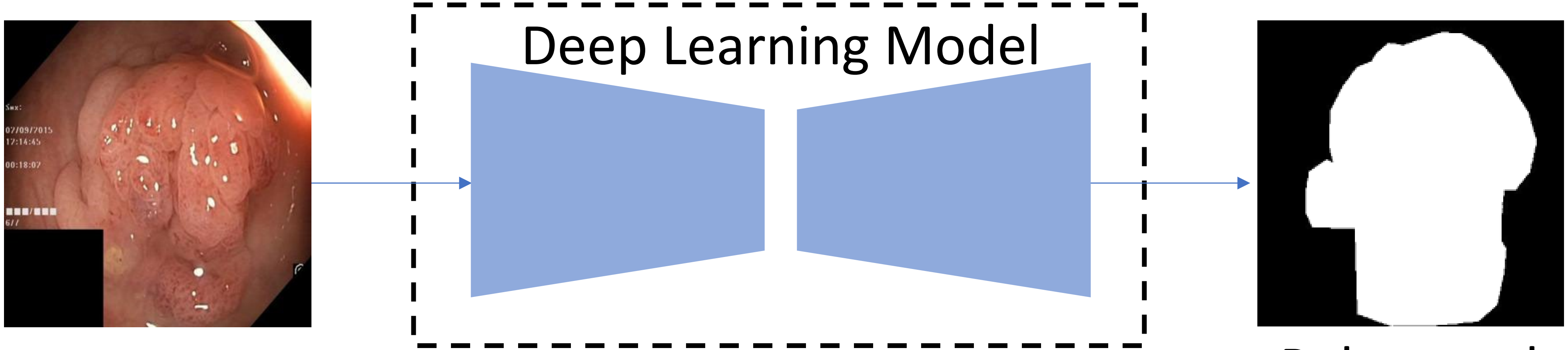


Fig 1. Polyp segmentation problem

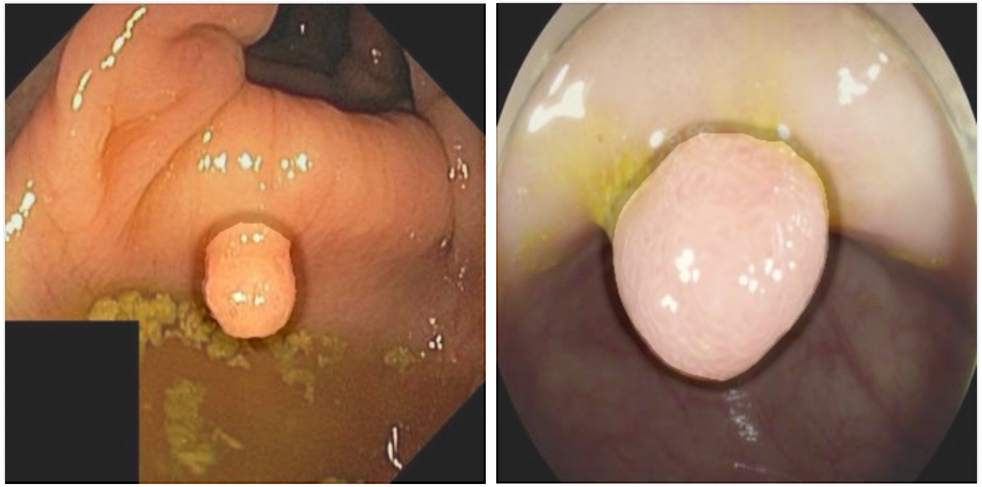
Application:

- Early detection to improving patient outcomes
- Endoscopy Support
- Drug Research and Development
- Reducing Endoscopy Time and Cost
- Medical Education and Training

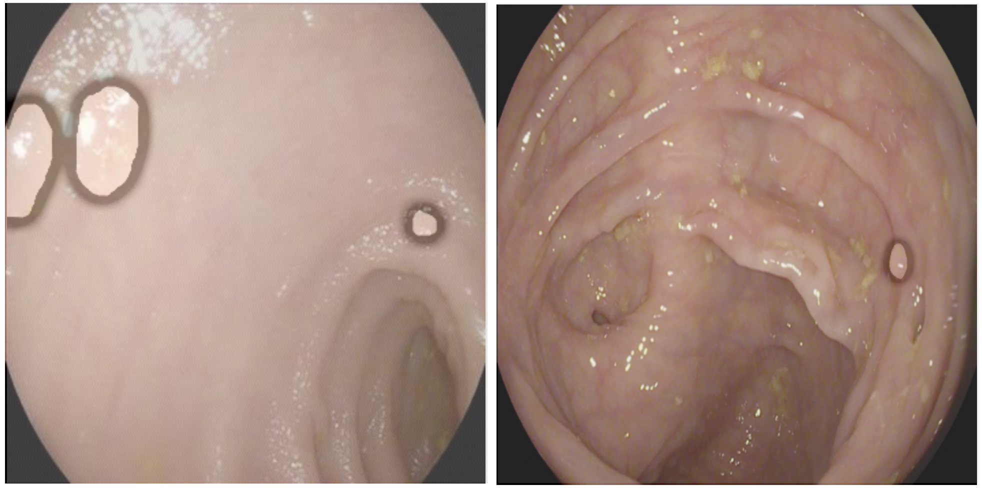
1. INTRODUCTION

Challenges:

- Polyps can vary greatly in shape, size, and color.
- Similarity between polyp and surrounding tissue.



(a)



(b)

Fig 2. (a) Varies greatly in shape, size and color (b) Similarity between polyps and surrounding tissue.

2. Related works

Common Attention Blocks:

- **WOO, Sanghyun, et al CBAM [1]**

Áp dụng các lớp tích chập trên sự kết hợp của các tính năng đỉnh và trung bình theo kênh

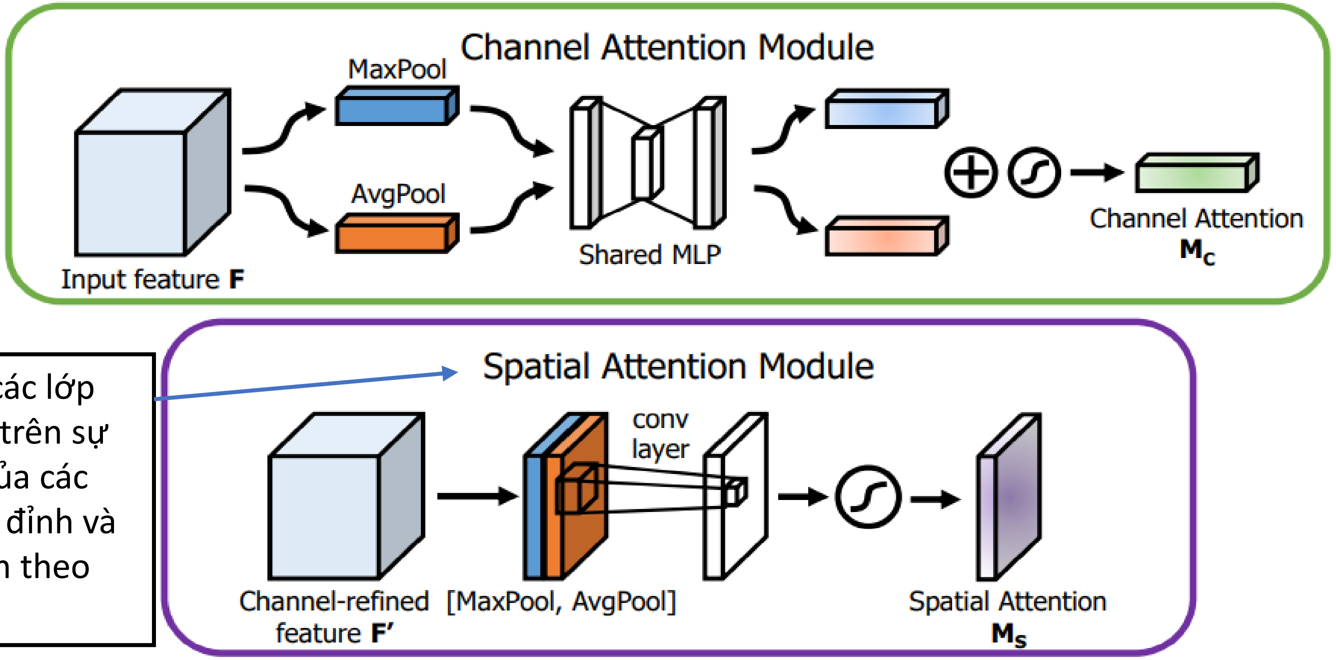


Fig 2. CBAM Architecture

- Optimize performance with Chanel Attention and Spatial Attention
- Focus on important features
- Use aggregated information from the entire feature space to determine importance
- Through global aggregation (both average and maximum) over spatial features.

- **HE, Kaiming, et al Residual Block [2]**

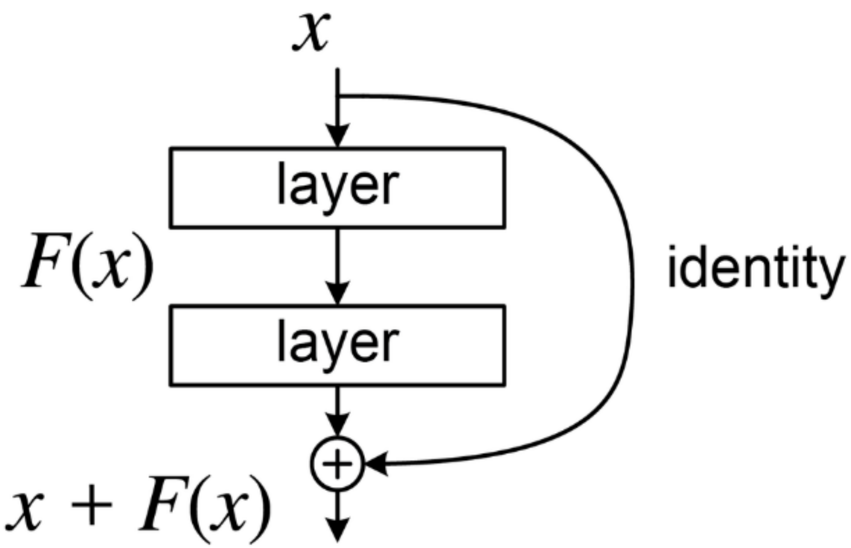


Fig 3. Residual Block

- Reduce vanishing gradient problem
- Speed up learning time
- Shortcut connections allow gradients to propagate directly through the network during backpropagation.

[1] WOO, Sanghyun, et al. **Cbam: Convolutional block attention module**. In: Proceedings of the European conference on computer vision (ECCV). 2018. p. 3-19.
[2] HE, Kaiming, et al. **Deep residual learning for image recognition**. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 770-778.

2. Related Works

Common Attention Block:

- **CORDONNIER, Self-Attention (SA) [1]**

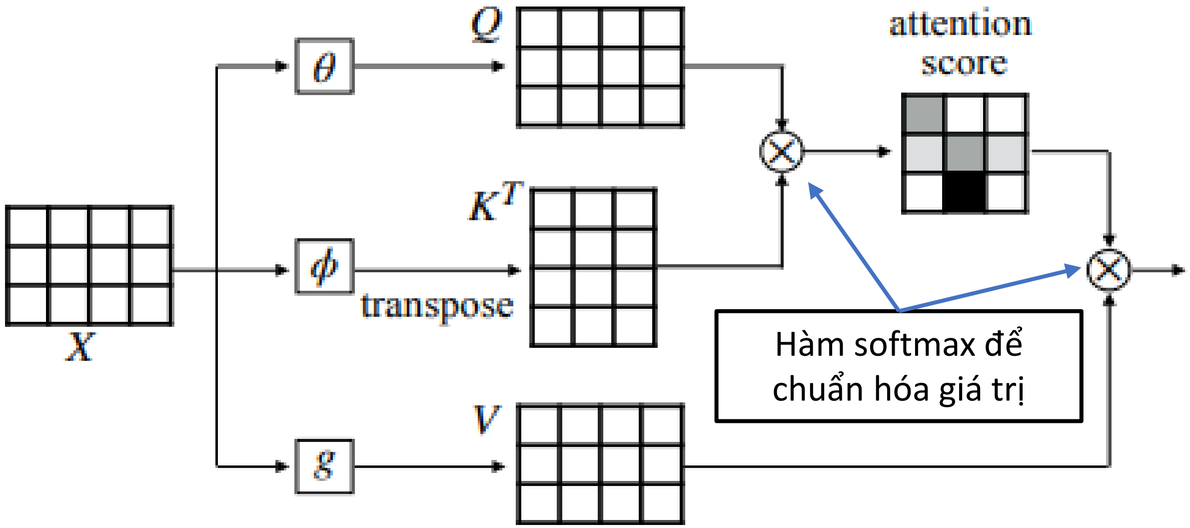


Fig 4. Self-Attention

- Understanding Context
- Flexibility
- Query (Q), Key (K) and Value (V) are computed by multiplying the input by three matrices.

[1] CORDONNIER, Jean-Baptiste; LOUKAS, Andreas; JAGGI, Martin. ***Multi-head attention: Collaborate instead of concatenate***. arXiv preprint arXiv:2006.16362, 2020.

3. Proposed Method

Unet Focus

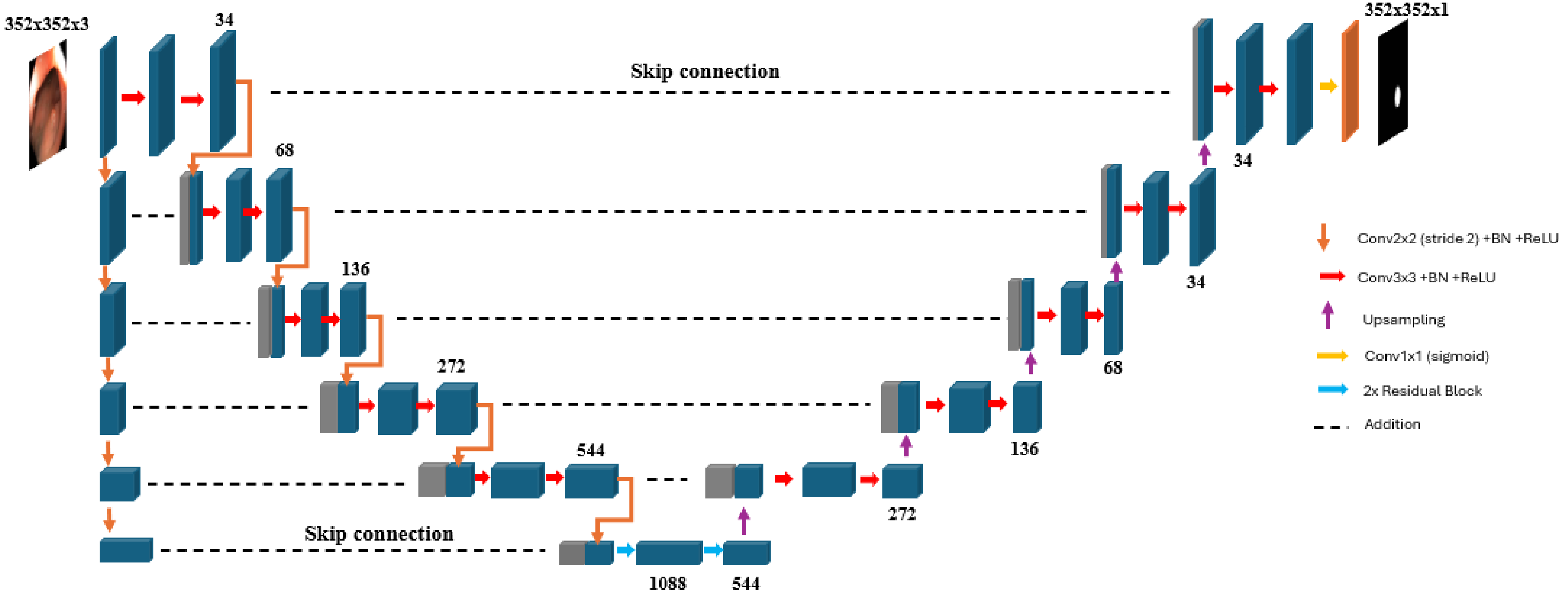


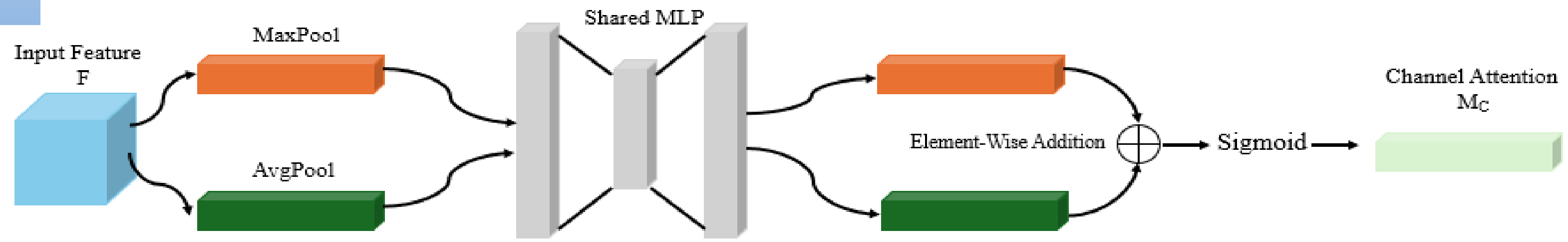
Fig 5. Our method - Unet Focus

- **Feature Enhancement:** Integrates CBAM in convolutional layers for superior feature refinement.
- **Residual Blocks:** Maintain effective training and prevent model degradation.
- **Advanced Decoder Path:** Utilizes up-sampling and skip connections to recover details, improving segmentation accuracy.
- **Unique CBAM Placement:** Positioned in the decoder and after each down-sampling to focus on critical features.

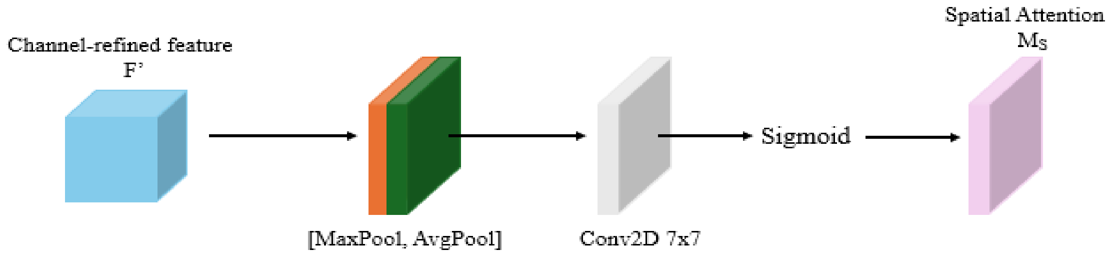
3. Proposed Method

Unet Focus + CBAM

Channel Attention



Spatial Attention



CBAM Block

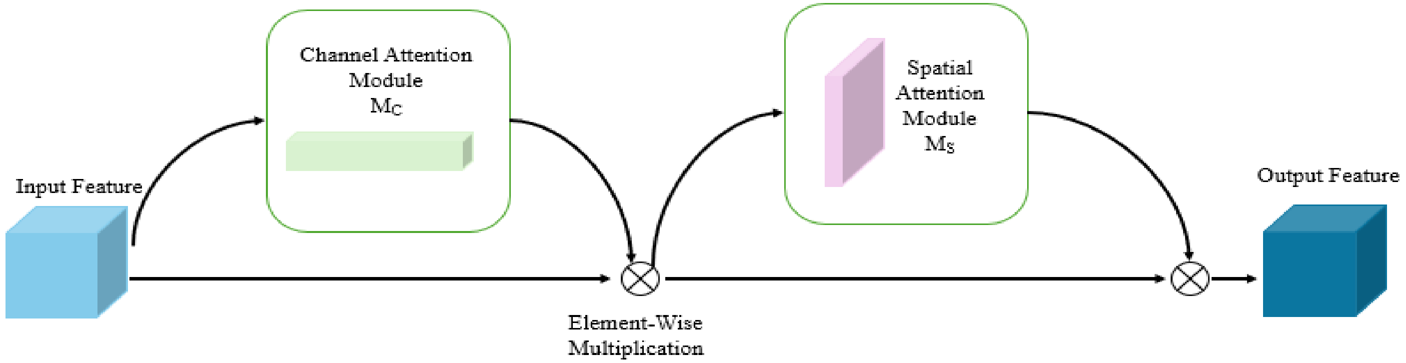


Fig 5. CBAM Architecture

3. Proposed Method

Unet Focus + CBAM

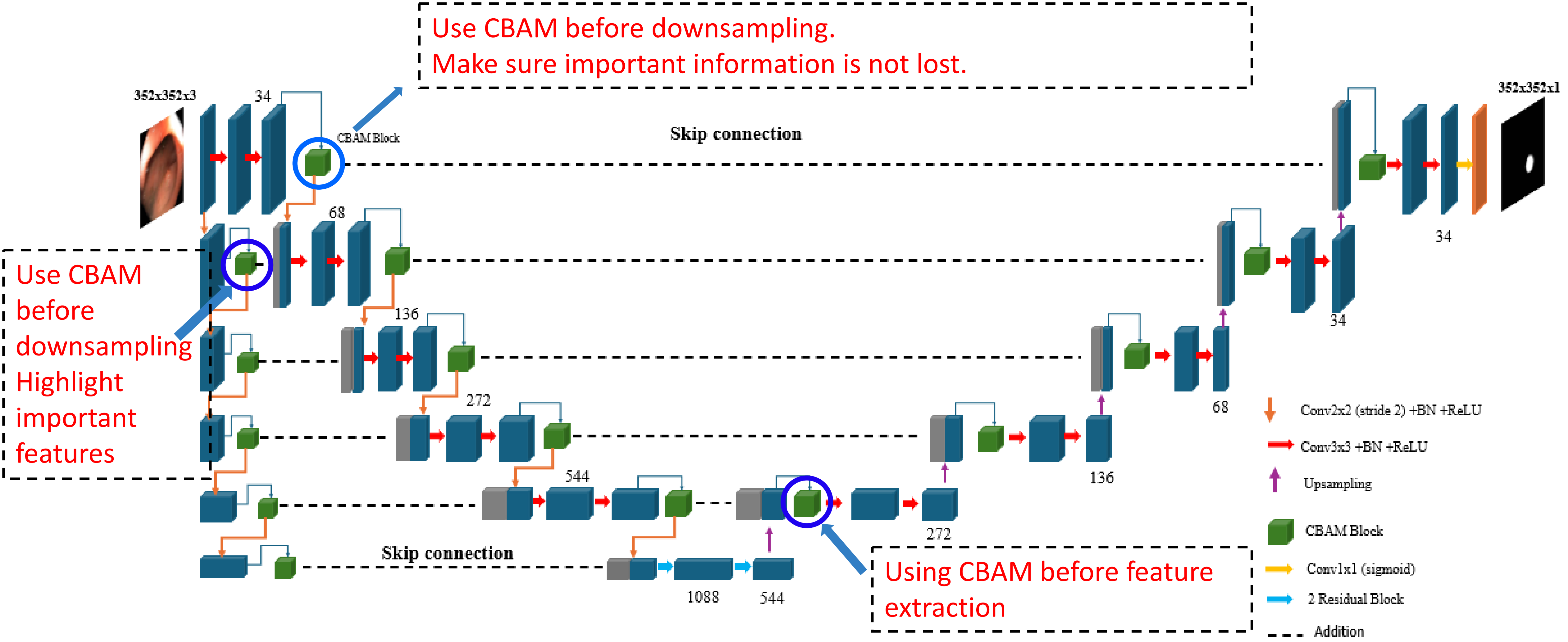
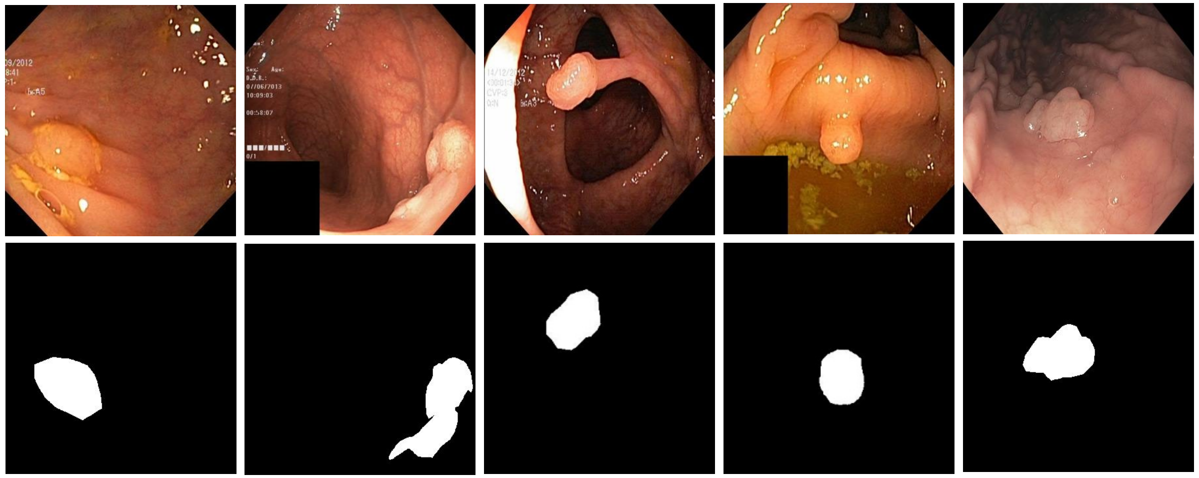


Fig 6. More Details about our method

4. Experiments and Discussion

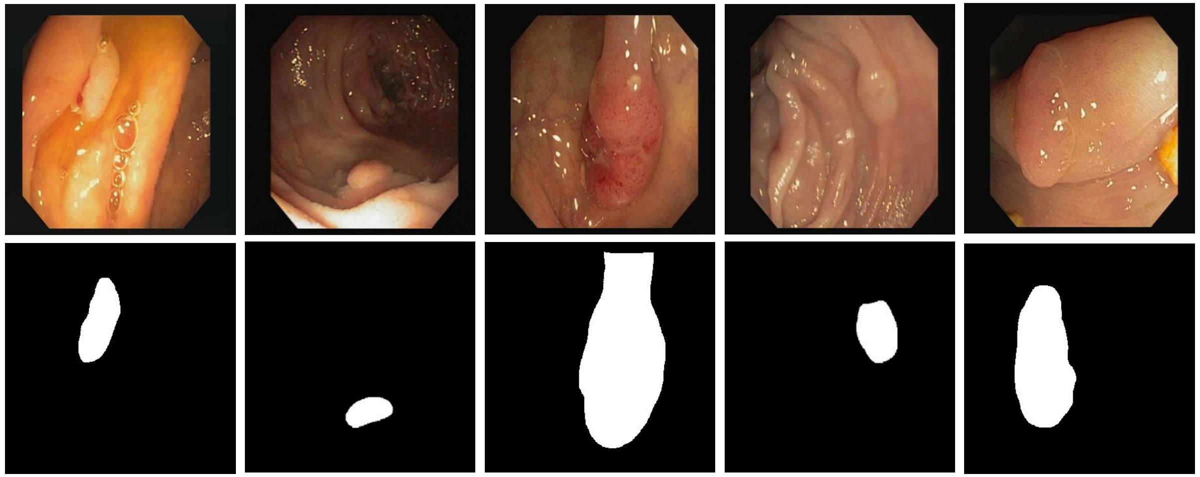
Datasets: Endoscopic imaging collection

Kvasir-SEG [1] – **1000** images



Train: 800
Val: 100
Test: 100

CVC-ClinicDB [2] – **612** images



Train: 488
Val: 62
Test: 62

Fig 7. Example Images of Datasets

[1] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, "CVC-ClinicDB," 2015. [Online]. Available: https://polyp.grand-challenge.org/CVCClinicDB/
[2] D. Jha, P. H. Smedsrud, M. A. Riegler, P. Halvorsen, T. de Lange, D. Johansen, and H. D. Johansen, "Kvasir-Seg: A Segmented Polyp Dataset," in *International Conference on Multimedia Modeling*, 2020, pp. 451-462.

4. Experiments and Discussion

Experiment setup:

- Image Size: **352x352x3**
- Learning rate: **1e-4 (0.0001)**
- Optimize: **AdamW** [1]
- Loss: **Dice Loss** [2]
- **500 epochs**

Metrics: Dice Score and mIOU

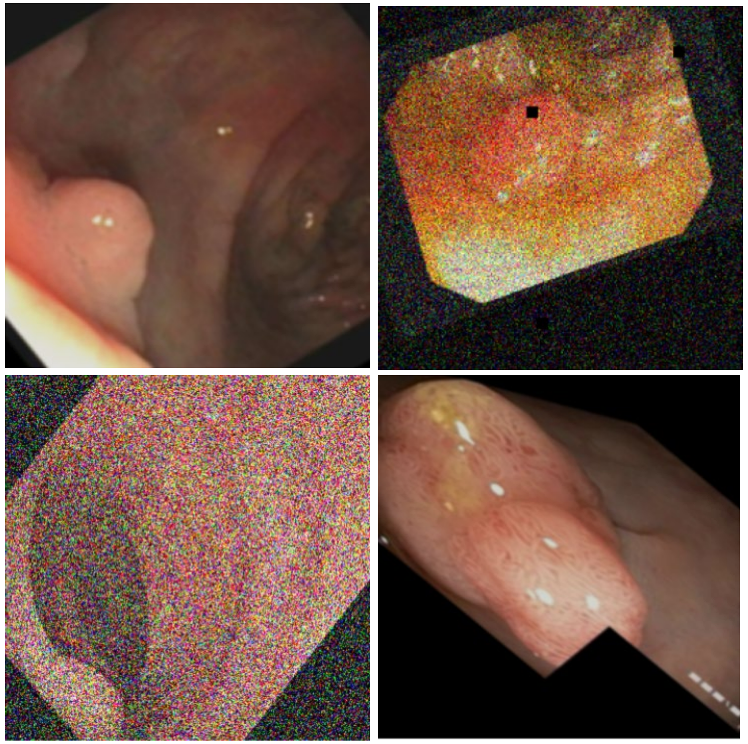


Fig 8. Data agumentation (Use many add-ons like Flip Horizontal, Flip Vertical, Color Jitter,…)

[1] LLUGSI, Ricardo, et al. Comparison between Adam, AdaMax and Adam W optimizers to implement a Weather Forecast based on Neural Networks for the Andean city of Quito. In: *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*. IEEE, 2021. p. 1-6.

[2] ZHAO, Rongjian, et al. Rethinking dice loss for medical image segmentation. In: *2020 IEEE International Conference on Data Mining (ICDM)*. IEEE, 2020. p. 851-860.

[3] SINGH, Owen; SENGAR, Sandeep Singh. BetterNet: An Efficient CNN Architecture with Residual Learning and Attention for Precision Polyp Segmentation. arXiv preprint arXiv:2405.04288, 2024.

4. Experiments and Discussion

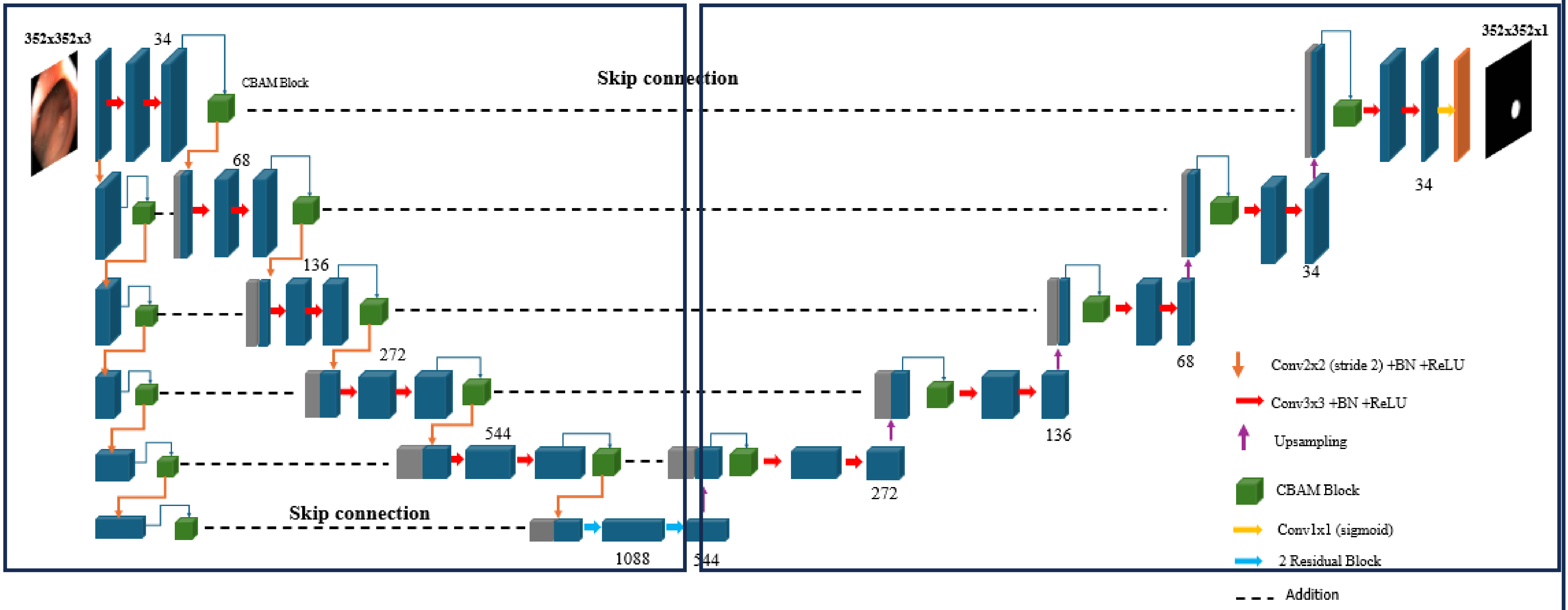


Fig 9. Ablation Studies on our method

4. Experiments and Discussion

Unet Focus + CBAM

Table 1: Ablation Studies on CBAM

Models	CVC-ClinicDB		Kvasir-SEG	
	DSC	mIOU	DSC	mIOU
U-Net Focus + CBAM Encoder	0.911	0.836	0.862	0.762
U-Net Focus + CBAM Decoder	0.919	0.852	0.883	0.801
U-Net Focus + CBAM Encoder/Decoder	**0.935**	**0.877**	**0.891**	**0.803**

Using CBAM on both Encode and Decode branches achieves the highest rate.
Improves Dice score by 2% on CVC-ClinicDB and 1% on Kvasir-SEG

4. Experiments and Discussion

Unet Focus + CBAM

Table 2: SOTA Comparision

Model	Year	CVC-ClinicDB		Kvasir-SEG	
		DSC	mIOU	DSC	mIOU
U-Net	2015	0.710	0.627	0.818	0.746
ResUnet++	2019	0.763	0.701	0.813	0.793
HRNetV2	2019	0.778	0.636	0.853	0.744
DCRNet	2022	0.856	0.788	0.886	**0.825**
MSRF-Net	2022	0.906	0.828	0.851	0.740
Ours	2024	**0.935**	**0.877**	**0.891**	0.803

The segmentation results show that my proposed method achieves the best results with a Dice score of 0.935 on CVC-ClinicDB and 0.891 on Kvasir-SEG.

4. Experiments and Discussion

Unet Focus + CBAM

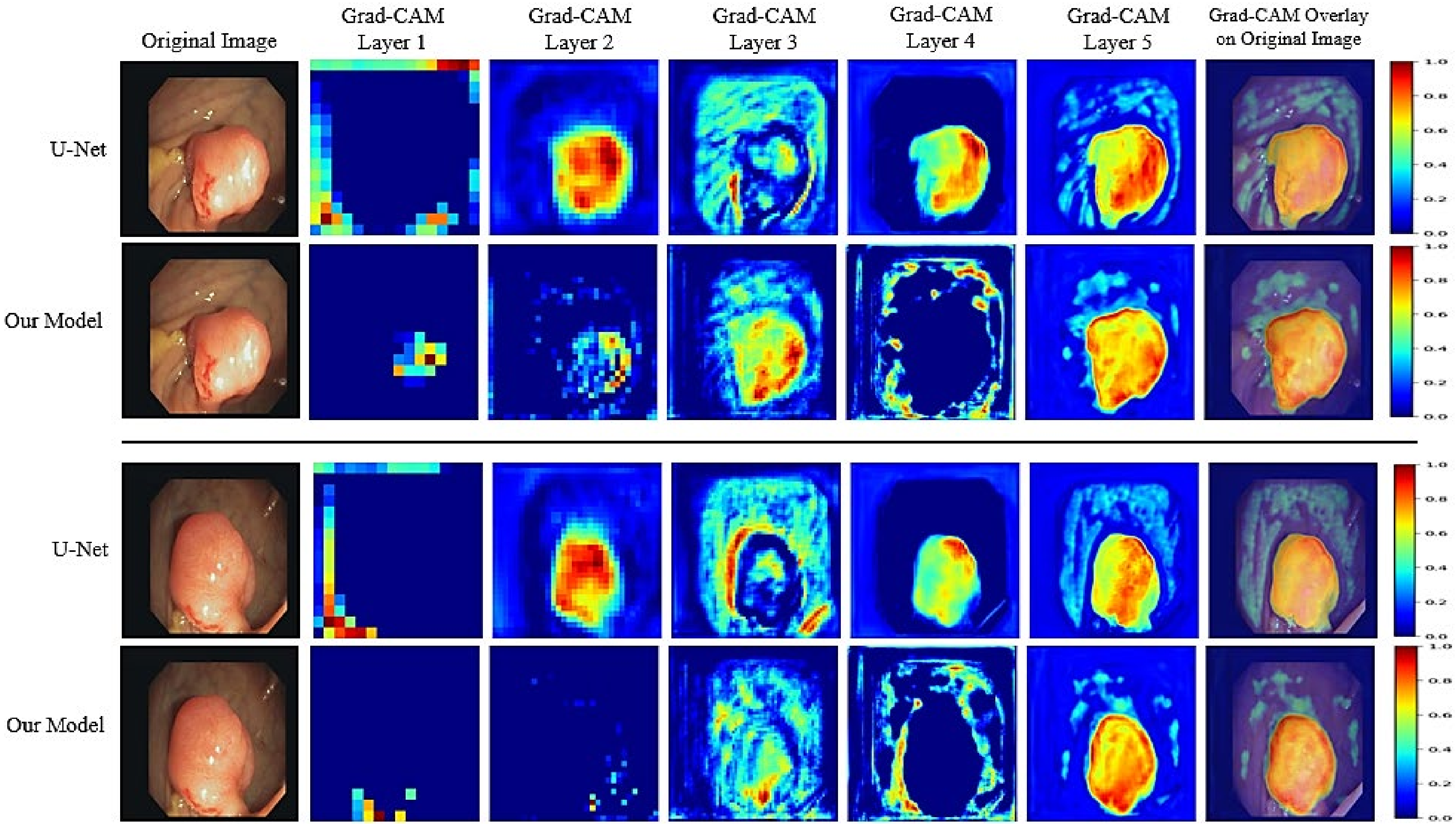


Fig. 10. Grad-CAM comparison highlights our model-based enhancement focus

```python
    st.image("https://production-media.paperswithcode.com/datasets/Screenshot_from_2021-05-05_23-44-10.png", use_column_width=True)
    st.header("How to Use This Tool")
    st.write("""
        This tool allows healthcare professionals to upload colonoscopy images and receive automated polyp segmentation.
        - **Original Image**: Shows the uploaded colonoscopy image.
        - **Predicted Mask**: Shows the segmentation of the polyp.
        Supported file formats: jpg, jpeg.
    """)

    contrast = st.sidebar.slider("Adjust Contrast", 1.0, 3.0, 1.5)
    brightness = st.sidebar.slider("Adjust Brightness", 1.0, 2.0, 1.2)

    uploaded_files = st.file_uploader("Choose images", accept_multiple_files=True, type=['jpg', 'jpeg'])
    temp_dir = './temp_app_data/app_images/test'
    if uploaded_files:
        if st.button('Process Images'):
            with st.spinner('Processing images...'):
                models = load_models()
                process_and_display_all_models(models, uploaded_files, temp_dir, contrast, brightness)
                st.success('Processing complete!')

if __name__ == "__main__":
    main()

```

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 401ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 1s 549ms/step
1/1 ━━━━━━━━━━━━━━━━━━━━ 1s 718ms/step

 *  History restored

PS C:\Users\talas\OneDrive\Desktop\Demo_polyp>
 *  History restored

PS C:\Users\talas\OneDrive\Desktop\Demo_polyp>
 *  History restored

PS C:\Users\talas\OneDrive\Desktop\Demo_polyp> strea
```

5. Conclusion

- **Integration of CBAM:** The addition of the Convolutional Block Attention Module (CBAM) to the U-Net Focus architecture has led to significant enhancements in medical image segmentation.
- **Enhanced Accuracy:** CBAM enables more precise focus on critical image areas, greatly improving the accuracy and sensitivity of object detection and segmentation.
- **Impressive Results:** Testing on CVC-ClinicDB and Kvasir-SEG datasets showed that U-Net Focus with CBAM significantly outperforms traditional U-Net and competes well with other advanced models.
- **Future Potential:** Demonstrates the value of CBAM, suggesting potential for further advancements in attention mechanisms within practical applications.

Thank you