



BAMPolyp: Bi-Axial Mamba Bottleneck for Gastrointestinal Polyp Segmentation

Md. Farhadul Islam, Tashik Ahmed, Partho Chanda, Joyanta Jyoti Mondal, Meem Arafat Manab, Sarah Zabeen, Jannatun Noor



Introduction:

Colorectal cancer remains one of the leading causes of cancer-related deaths in the US [1], and early detection of polyps during colonoscopy is vital for reducing mortality. However, small, flat, or camouflaged polyps are often missed due to their low contrast and irregular boundaries. Automated segmentation can assist clinicians, but existing models struggle to balance fine boundary precision with the need for global contextual understanding. Convolutional networks tend to capture local details but lack long-range dependencies, while Transformer-based models improve global reasoning but are computationally expensive and may blur boundaries. This creates a need for an efficient solution that preserves boundary accuracy while incorporating global context.

Objective:

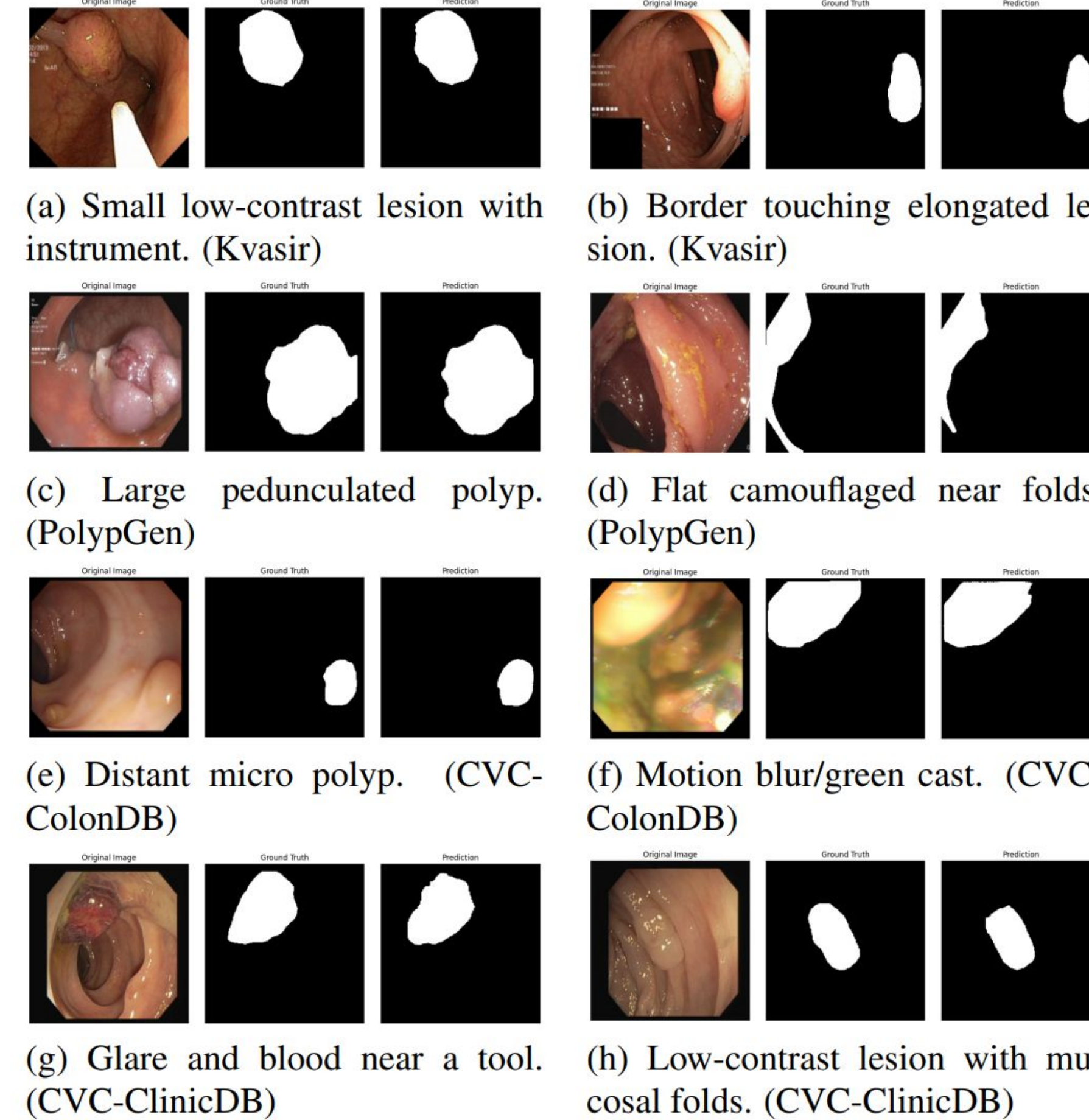
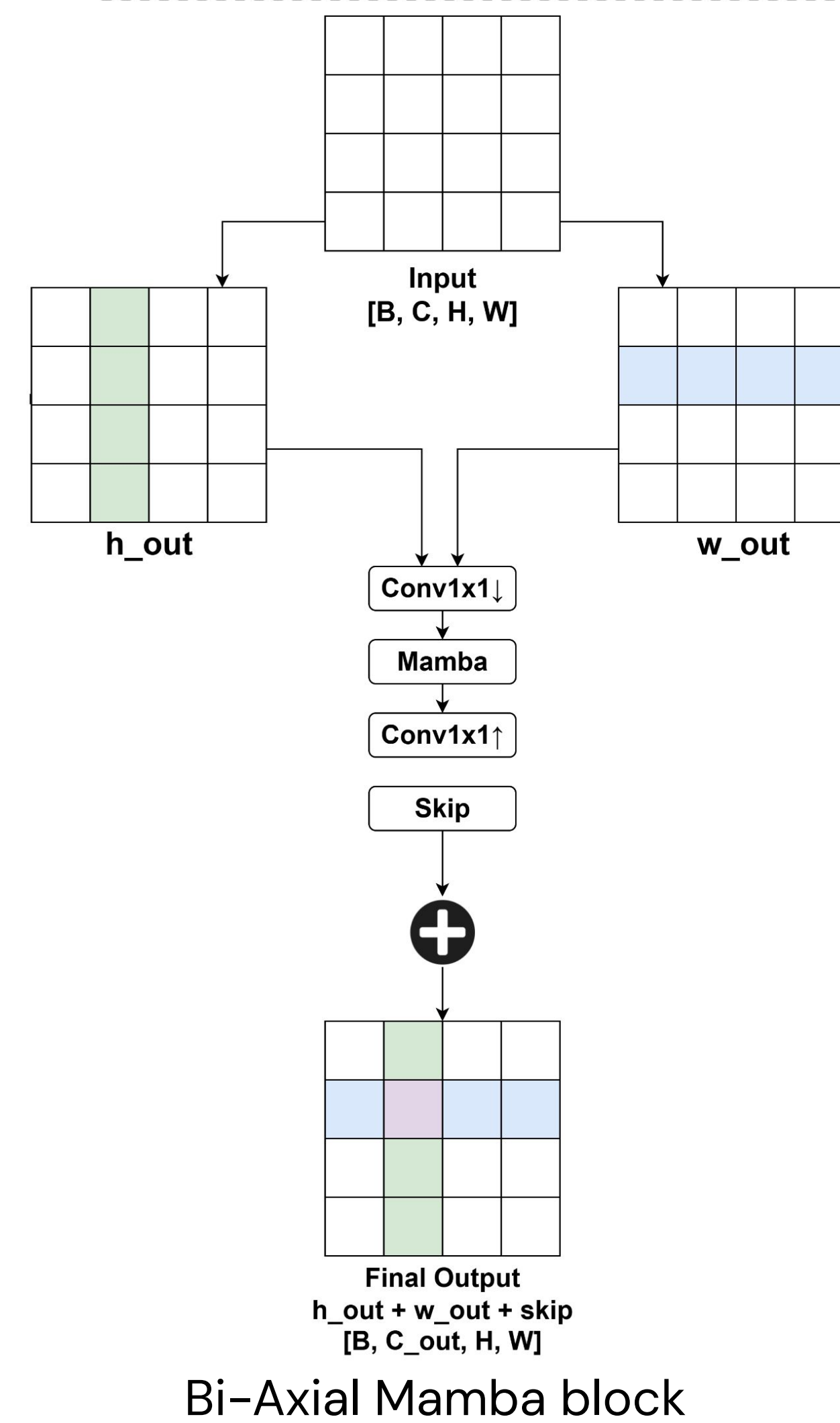
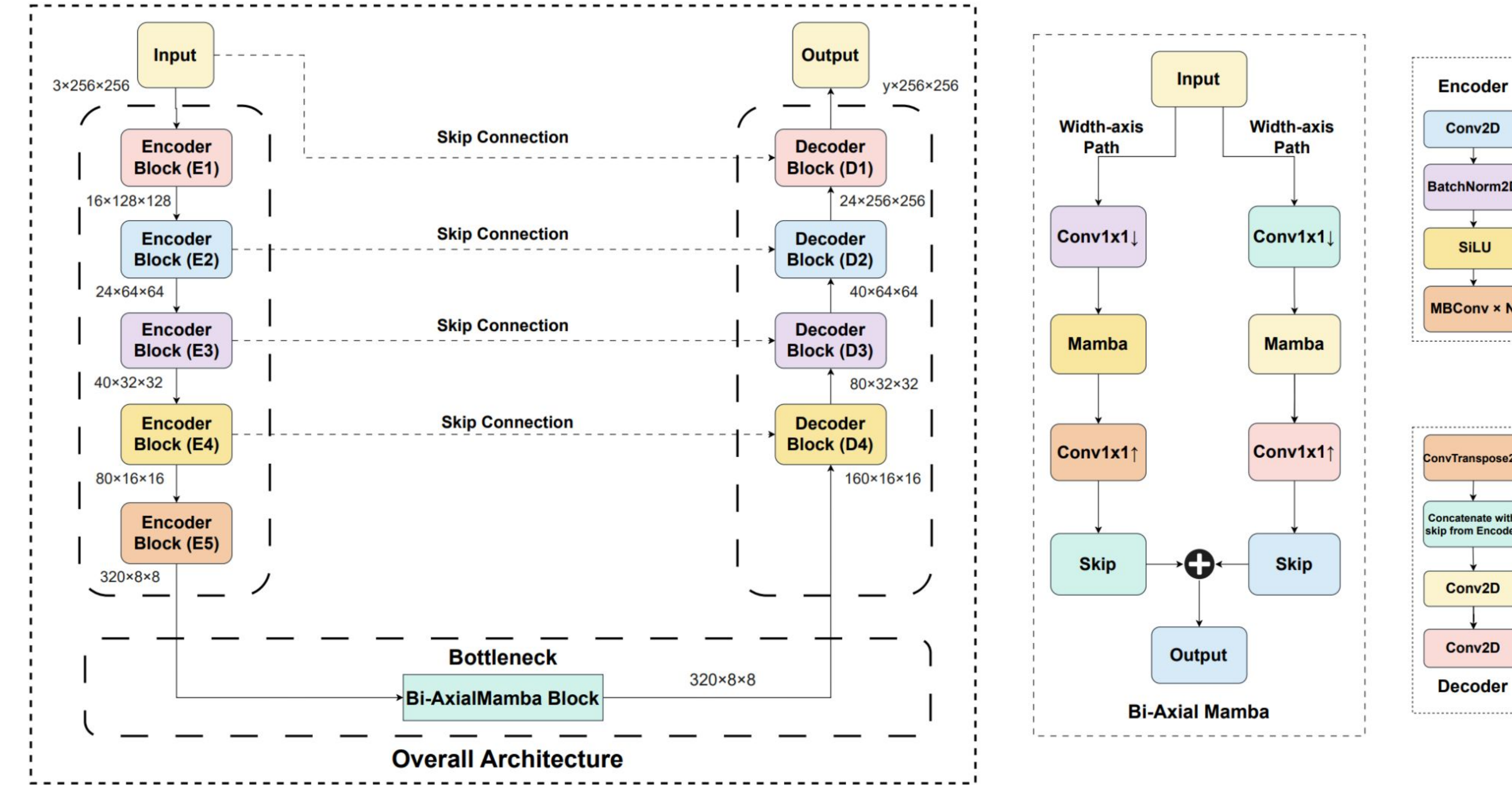
- Preserve spatial structure via directional expansion
- Enhancing Local-Global Contextual feature extraction
- Bottleneck Enhancement with axial state-space modeling
- Balanced Performance and Efficiency

Dataset:

Kvasir-SEG, CVC-ClinicDB, CVC-ColonDB, and PolypGen

Methodology:

- U-Net style framework with a Bi-Axial Mamba bottleneck.
- Leverages VMamba-style state-space mixer [2].
- Axis-wise state-space mixing:
 - Height-axis Mamba captures dependencies along columns.
 - Width-axis Mamba captures dependencies along rows.
- Preserves local continuity while integrating global context.
- EfficientNet-BO encoder for feature extraction.



The original colonoscopy frame (left), ground-truth mask (middle), and our model prediction (right)

Results:

- Dice / IoU: Kvasir-SEG 0.938 / 0.888, CVC-ClinicDB 0.944 / 0.894, CVC-ColonDB 0.926 / 0.866, PolypGen 0.868 / 0.821
- Outperforms PVT-Cascade (0.926/0.878) and TransFuse-L (0.920/0.870) on Kvasir, and rivals PVT-EMCAD-B2 (0.952 Dice) on ClinicDB while staying far lighter.
- Robust generalization across polyps of varied size, texture and contrast including flat or camouflaged region.
- 6.51 M parameters, 3.13 G FLOPs.

Ablation Studies:

- Bi-Axial Mamba > Plain Mamba, Axial Attention, Convolution, V2M.
- Gains of up to +5.9% Dice and +8.9% IoU on challenging datasets.

Bottleneck	Kvasir		ClinicDB		ColonDB		PolypGen	
	Dice	IoU	Dice	IoU	Dice	IoU	Dice	IoU
PlainMamba	0.9307	0.8801	0.9339	0.8810	0.9114	0.8530	0.8222	0.7980
VMamba	0.9361	0.8852	0.9389	0.8843	0.9198	0.8611	0.8444	0.8157
Axial Attention	0.9335	0.8810	0.9366	0.8852	0.9229	0.8632	0.8544	0.8134
Plain Conv	0.9273	0.8767	0.9301	0.8772	0.9074	0.8333	0.8088	0.7322
V2M	0.9319	0.8782	0.9330	0.8789	0.9180	0.8572	0.8474	0.8101
Bi-Axial Mamba (Ours)	0.9380	0.8881	0.9437	0.8939	0.9255	0.8659	0.8683	0.8211

Conclusion

- BAMPolyp achieves a strong balance of **accuracy, efficiency, and generalization.**
- Effective for real-time and resource-constrained clinical applications.
- Future work: Exploring optimal architectures for other layers and other medical segmentations, leveraging BAMPolyp's advantage.

References

1. <https://www.cancer.org/cancer/types/colon-rectal-cancer/about/key-statistics.html>
2. Liu et al., Vmamba: Visual state space model, NeurIPS, 2024.

Contact

