

Photorealistic Colonoscopy Image Generation

Ángel Villanueva Agudo • Clara Tomasini • O. León Barbed • Ana C. Murillo

RoPeRT Group • Instituto de Investigación en Ingeniería de Aragón (I3A) • Universidad de Zaragoza

Motivation & Background

Colorectal cancer is one of the leading causes of **cancer-related mortality** worldwide. Automatic **polyp segmentation** in colonoscopy enables **early detection** and can save lives.

! Key challenge

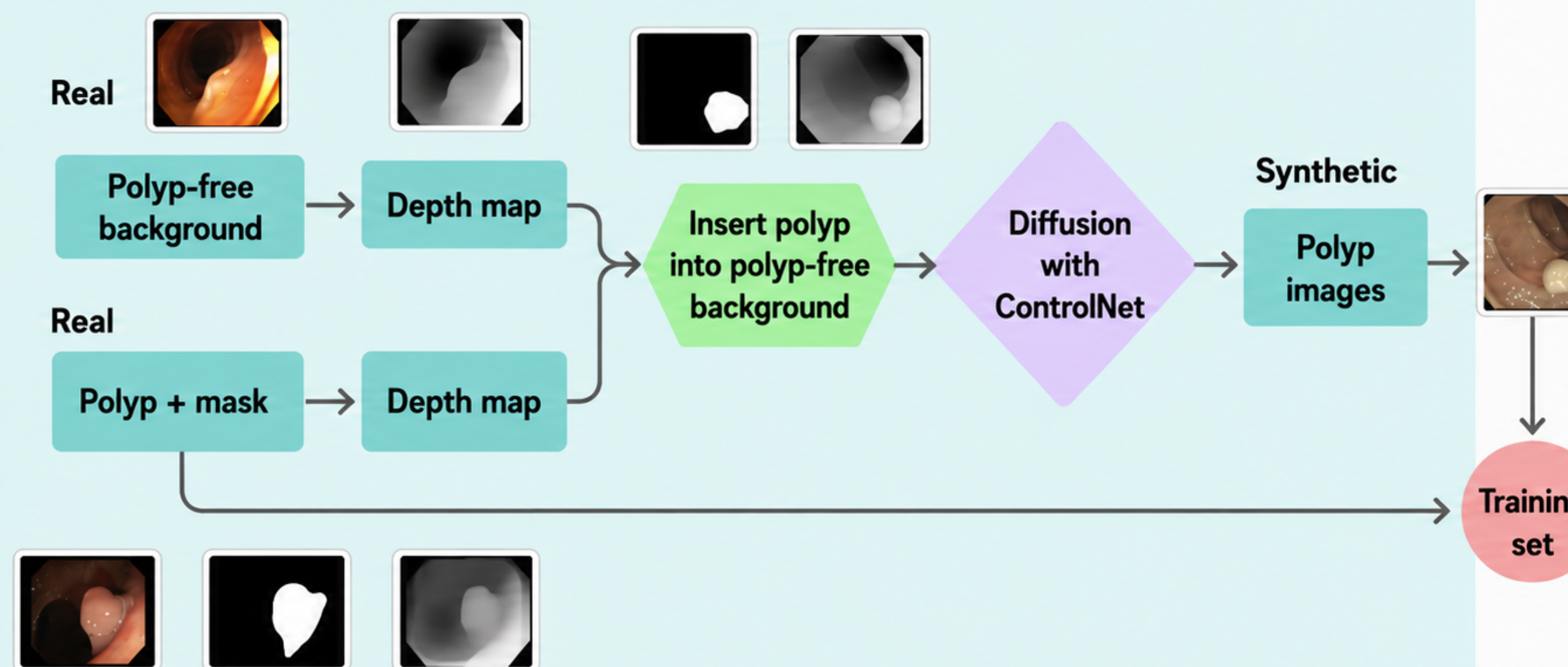
- Scarce clinical data and privacy issues
- Costly expert masks

+ Contribution

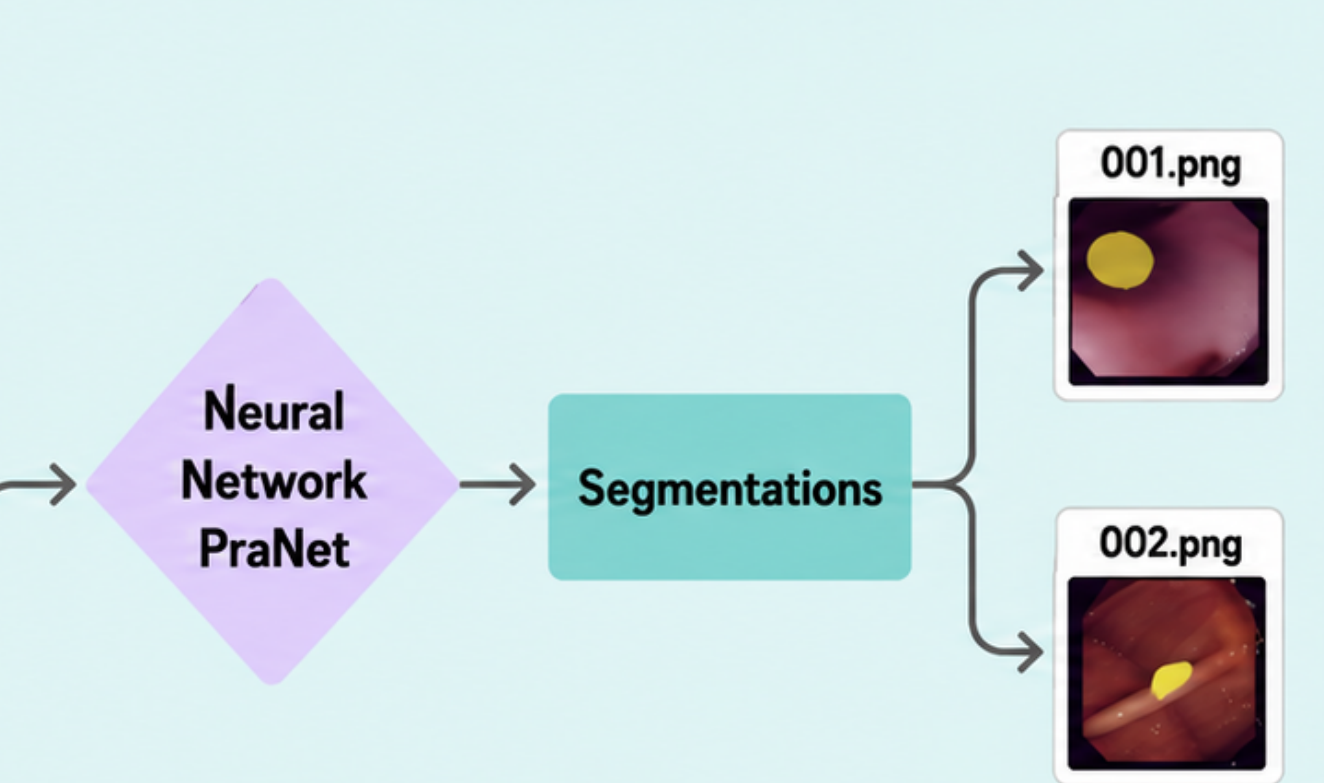
- Synthetic image-masks pairs
- SSIM-based placement
- Small-polyp curation

Methodology

Synthetic Image Generation



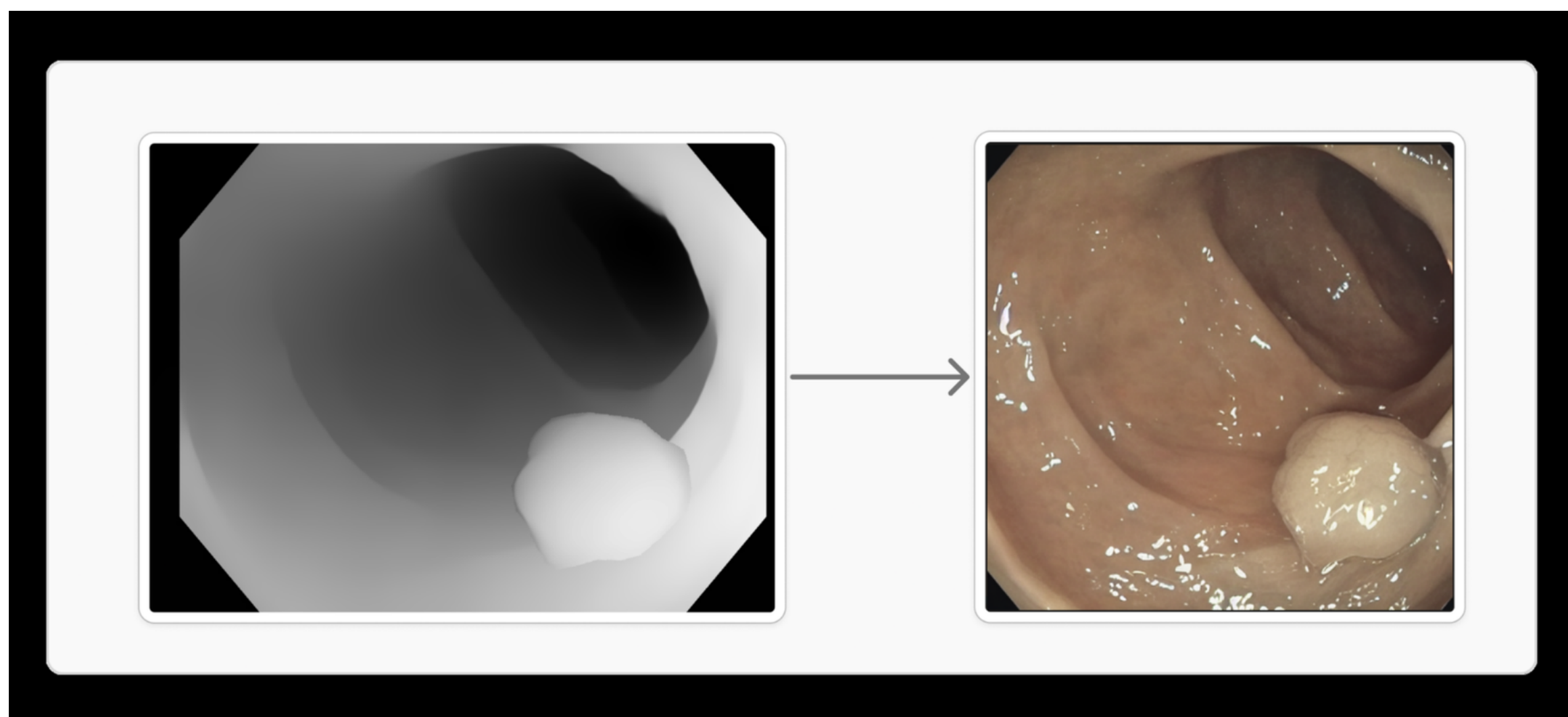
Polyp Segmentation Model Training and Evaluation



Overview of the proposed synthetic generation pipeline for data synthesis and segmentation model training, showing the involved datasets, algorithms, models, and example outputs.

Results - image generation

Diffusion example



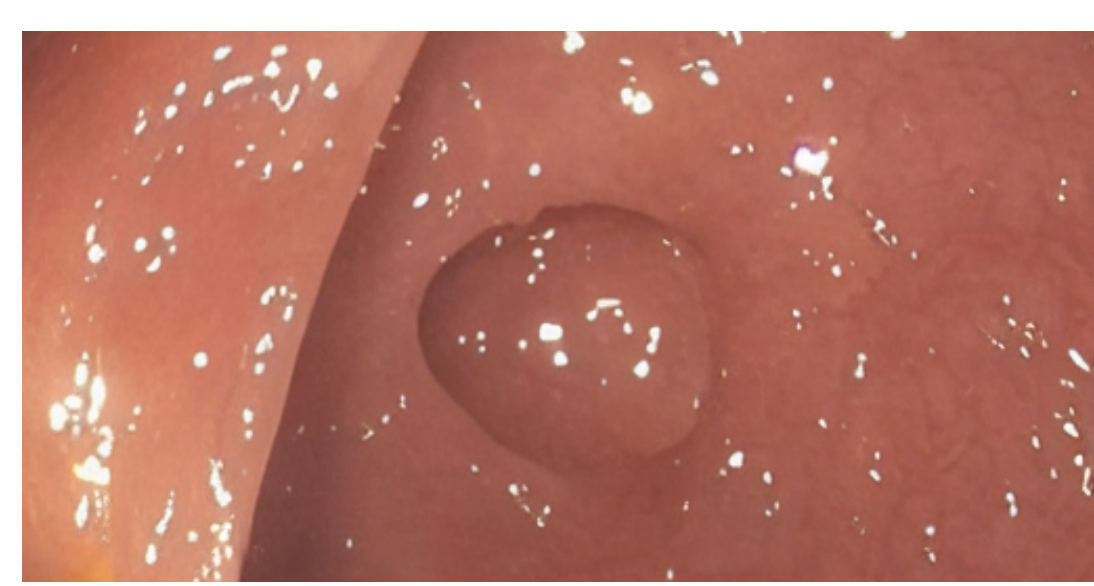
Depth[3] control signal + inserted polyp - synthetic colonoscopy image

Examples by set

PolypMax



PolypCore



PolypMini



RealMiniMax



Representative samples showing the key feature of each training set.

Results - segmentation

Training sets

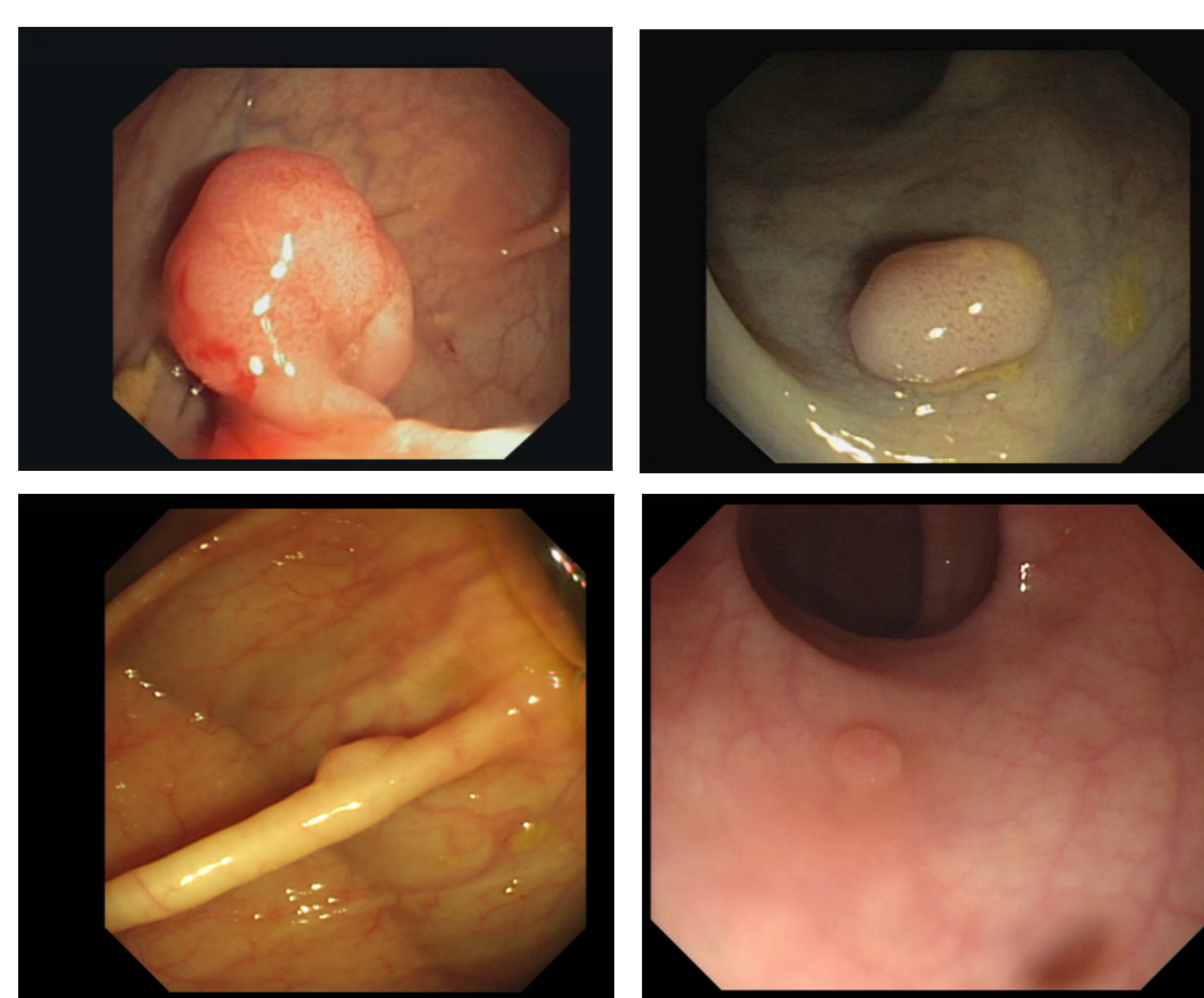
PraNet[2] trained on 2900 images, evaluated on CVC-300[5] and CVC-ClinicDB[4]:

Dataset	Type	Small	Rare	SSIM	Local Scale	CVC-300 Dice/IoU	ClinicDB Dice/IoU
PolypMax	S	✓	✓	✗	✗	0.51/0.43	0.50/0.41
PolypCore	S	✗	✗	✓	✓	0.80/0.72	0.45/0.39
PolypMini ★	S	✓	✗	✓	✓	0.84/0.75	0.52/0.45
RealMiniMax	M	—	—	—	—	0.89/0.82	0.91/0.85
RealExp (ref.)	R	—	—	—	—	0.88/0.80	0.90/0.85

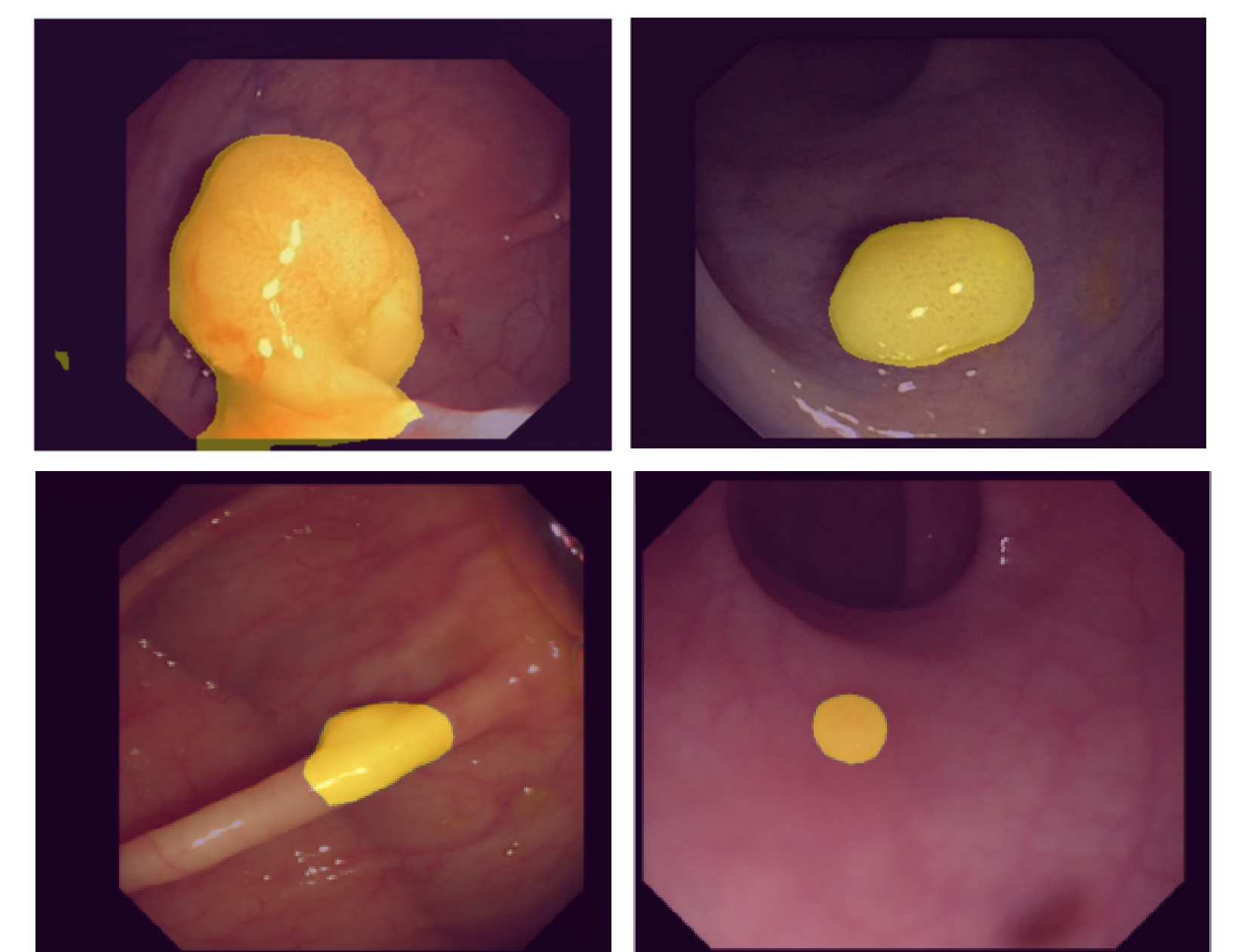
S: fully synthetic • M: 50% real + 50% synthetic • R: real only • ★: best synthetic config

Prediction examples

Input colonoscopy



PraNet prediction



Examples of real colonoscopy inputs and PraNet predicted polyp masks.

Conclusions

- **Depth-guided ControlNet**[1] generates useful synthetic data for **polyp segmentation**.
- The pipeline reduces reliance on **real clinical data** while keeping **competitive performance**.
- **PolypMini** is the best fully synthetic set.
- **RealMiniMax** matches **real-data training** performance.
- Synthetic data supports **privacy-preserving** medical image reuse.

References

- [1] Zhang & Agrawala. ControlNet. ICCV 2023.
- [2] Fan et al. PraNet. MICCAI 2020.
- [3] Yang et al. Depth Anything. CVPR 2024.
- [4] Bernal et al. CVC-ClinicDB. CVC 2015.
- [5] CVC & Hospital Clínic Barcelona. CVC-300. 2012.