IEEE International Conference on Image Processing
27 - 30 October 2024, Abu Dhabi, UAE

Accurate Colon Segmentation Using 2D Convolutional Neural Networks With 3D Contextual Information

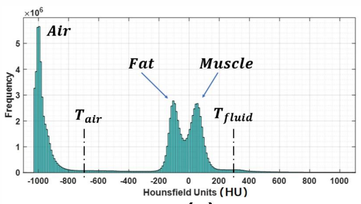
Samir Harb Ahmed Elsayed Mohamed Yousuf Islam Alkabbany Asem Ali Salwa Elshazley Aly Farag
Computer Vision and Image Processing Laboratory (CVIP), University of Louisville, Louisville, KY

ABSTRACT:

This study introduces an innovative 2D framework for accurate colon segmentation in abdominal CT scans, addressing unique challenges in this task. The proposed architecture integrates 2D and 3D contextual information by generating an attention map for each slice, considering neighboring slices. A per-pixel weighted dice loss is designed to enhance segmentation precision, giving more weight to pixels near the colon boundaries. This approach achieves effective segmentation using 2D CNNs, avoiding the complexity of 3D CNNs or LSTM networks. The framework is validated on a dataset of 98 CT scans from 49 patients, achieving a Dice coefficient of 98.67%.

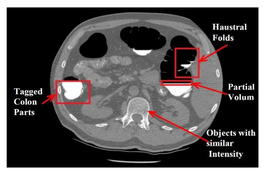
Background

- The colon segmentation step plays a crucial role in the CTC pipeline, as errors in this phase can affect subsequent steps.
- Colon segmentation is performed to extract the lumen in abdominal CT and is a challenging task due to several factors as shown in the Figure



(a)

(b)

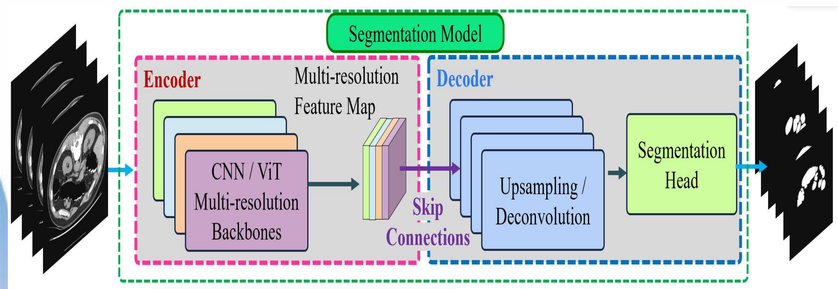


(c)

(d)

Related Work

- The traditional meta-architecture (illustrated in the Figure) for segmentation involves an encoder-decoder structure.



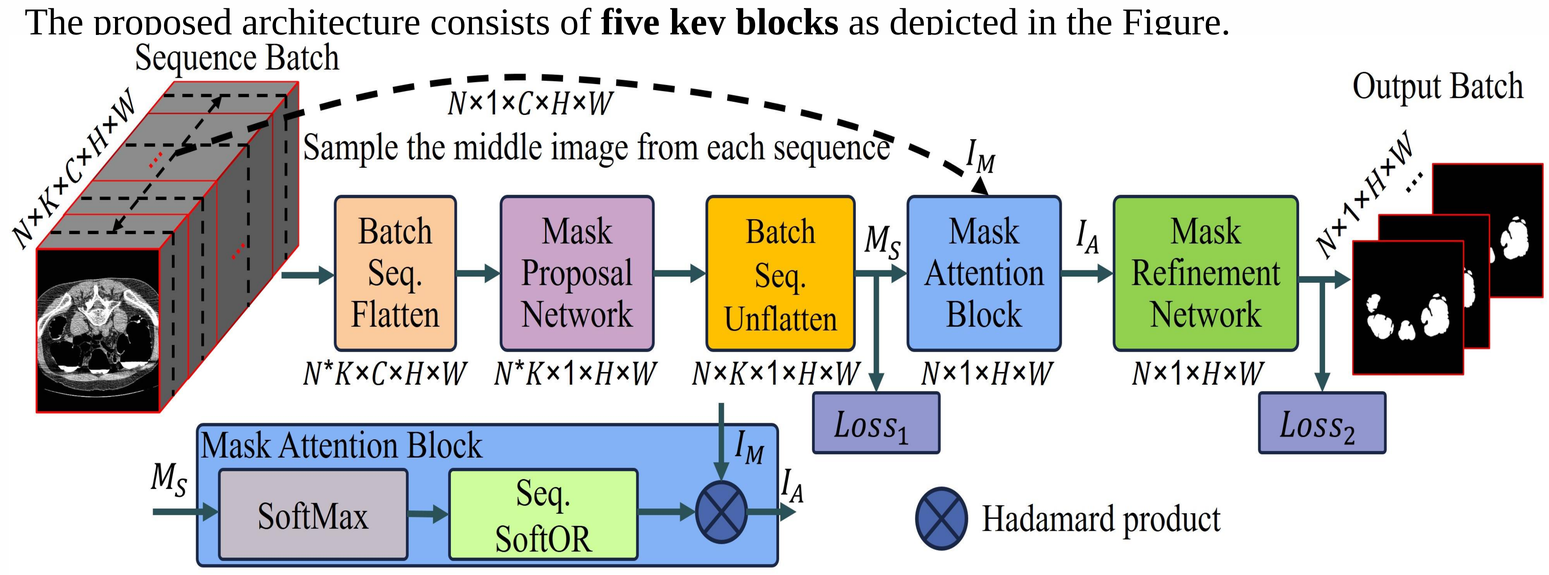
Segmentation Models	
• UNet++	• PAN
• Manet	• DeepLabV3+
• Linknet	

Backbones	
CNN	tu-resnest269e
ViT	maxvit_xlarge_tf_512 mobilevitv2_200

- Based on various experiments, Table 1 the best one found was the base model Unet++ paired with resnest269.
- A key limitation of this approach is its inability to incorporate temporal information, which is critical for processing image sequences like those in colon CT scans.
- So, we aimed to propose an approach to mitigate challenges encountered by the traditional encoder-decoder approaches, integrating both 2D and 3D contextual information using 2D networks.

Methodology

- The proposed architecture consists of **five kev blocks** as depicted in the Figure.



Where: N is batch size, K is Sequence length, C is number of channels, W is the image width, and H is the image height

- **BSF** block prepares sequences of CT images, and **MPN** generates initial segmentation masks. These masks are then reshaped back into sequences by the BSU block. **MA** block creates attention masks that focus on the most relevant areas of each CT slice, effectively incorporating 3D contextual information from neighboring slices. **MRN** refines these focused areas into the final segmentation masks.

Results

- We conducted our experiments on a private dataset comprised scans of 49 patients in both supine and prone positions. Experts annotated the colon segments in these 98 CT scans.
- we designed a per-pixel weighted dice loss that gives more weight to pixels near the boundaries of the colon (boundary weight)

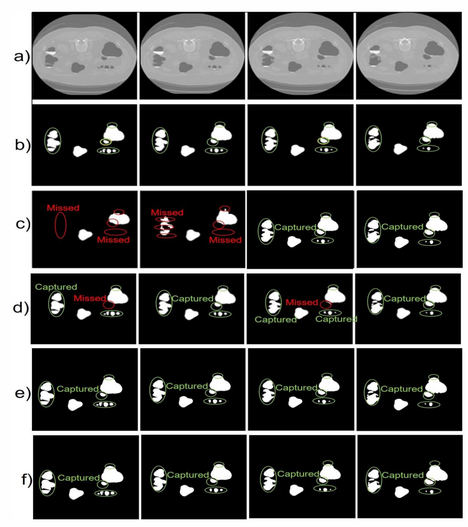
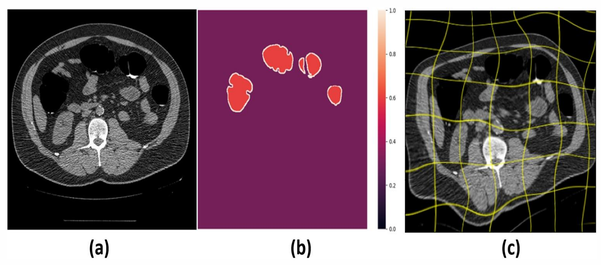


Table 1. Validation Dice and Jaccard scores for multiple architectures with different backbones to choose the baseline

Architecture/Backbone	Val. Dice Score	JACRD
PAN/resnest269e	97.1%	95.46%
DeepLabV3+/resnest269e	97.35%	95.74%
Linknet/resnest269e	97.36%	95.96%
UNet++/resnest269e	97.44%	95.75%



(a) (b) (c)

Table 2. Evaluation of volume segmentations of the proposed model compared to different approaches.

Model	DICE	JACRD	bAVD	KAPPA	ICCORR	MUTINF	SNSVTY	SPCFTY
nnUNet [20]*	98.79%	97.61%	0.0133	98.73%	98.79%	25.78%	98.85%	99.94%
Base Model	97.44%	95.75%	0.3136	97.39%	97.47%	22.72%	98.01%	99.89%
C-LSTM	89.16%	80.68%	0.39	88.62%	91.10%	21.30%	96.38%	99.08%
Ours	98.76%	97.60%	0.03	98.70%	98.76%	22.90%	98.60%	99.90%

* Due to computation burden, it was evaluated on half the resolution.

Conclusions

- We introduced a novel 2D deep learning approach for colon segmentation.
- These advancements provide a robust foundation for future applications in clinical settings, particularly in the 3D reconstruction of colon structures, which could significantly enhance the efficiency and accuracy of CRC screening.

References

- Isensee, Fabian et al., "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation“, Nature methods, 2021.
- Zang Hang, et al., “Resnest: Split-attention networks”, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022 **This work has been funded by the NSF grants 1602333 and 2124316**