



Improving Polyp Segmentation in Colonoscopy using Deep Learning

Saurab Rauniar¹, Vabesh Kumar Jha², Ritika Kumari Jha³, Debesh Jha⁴ and Ashish Rauniar⁵

1. Vyobotics, Singapore

2. Khwopa College of Engineering, Department of Electrical Engineering, Kathmandu, Nepal

3. Care Medical Center, Janakpur, Nepal

4. UiT The Arctic University of Norway, Tromsø, Norway

5. SINTEF Digital, Trondheim, Norway. Correspondence:ashish.rauniar@sintef.no

Abstract

Colorectal cancer is one of the major causes of cancer-related deaths globally. Although colonoscopy is considered the gold standard for examination of colon polyps, there is a significant miss rate of around 22-28%. Deep learning algorithms such as convolutional neural networks can aid in the detection and describe abnormalities in the colon that clinicians might miss during endoscopic examinations. The "MedAI: Transparency in Medical Image Segmentation" competition provides an opportunity to develop accurate and automated polyp segmentation algorithms on the dataset provided by the challenge organizer. We participate in the polyp segmentation task of the challenge and provide a solution based on the dual decoder attention network (DDANet). The DDANet is an encoder-decoder architecture based on a dual decoder attention network. Our experimental results on the organizers' dataset showed a dice coefficient of 0.7967, a Jaccard index of 0.7220, a recall of 0.8214, a precision of 0.8359, and an accuracy of 0.9557. Our results on test datasets suggest that deep learning and computer vision-based methods can effectively solve automated polyp segmentation tasks.

Keywords: Deep learning; medical image segmentation; endoscopy; colorectal polyps

Introduction

Polyps can be found anywhere in the gastrointestinal tract and can develop into colorectal cancer. The quality of a colonoscopy examination is largely determined by the experience and skill of the endoscopist [1]. Polyps are mostly divided into non-neoplastic and neoplastic. Non-neoplastic polyps are non-cancerous and do not require removal. However, neoplastic polyps such as adenomas and serrated polyps can develop into cancer. In addition, based on their size, colorectal polyps can be categorized into diminutive ($\leq 5\text{mm}$), small (6 to 9 mm), and large

($\geq 10\text{mm}$) [2]. Usually, larger polyps are visible and can be easily detected and resected by gastroenterologists. However, gastroenterologists have a problem detecting diminutive and smaller sized polyps. A framework that could detect such small or diminutive polyps abnormalities in real-time could potentially assist clinicians by providing extra pair of eyes to the physicians. This could also improve the clinical workflow and also reduce the clinical burden.

To tackle the problem of polyp high miss-detection, we propose a solution based on DDANet [3], specifically designed to achieve a decent performance and real-time speed. DDANet showed a dice coefficient of 0.8576 and a real-time speed of 69.59. In our work, we have carefully performed data augmentation and trained a smaller network that requires low GPU training time and only has a lower number of trainable parameters. Thus, our aim is to achieve decent evaluation metrics while achieving a decent frame per second (FPS) speed.

Related Works

There have been several attempts made to develop automated systems for colorectal cancer screening [4, 5, 6]. Fan et al. [7] proposed a parallel reverse attention network (PraNet) that applied area and boundary cues in a reverse attention module and achieved a good performance on the still frames. Jha et al. [8] performed an extensive study on the different publicly available polyp still and video datasets using ResUNet++ [9], Conditional Random Field (CRF) and Test-Time Augmentation (TTA) in order to improve its prediction performance on the polyp image and video datasets. The use of CRF and TTA techniques further improved the performance of ResUNet++. Another architecture called MSRF-Net [10] was built on a dual-scale dense fusion (DSDF) block

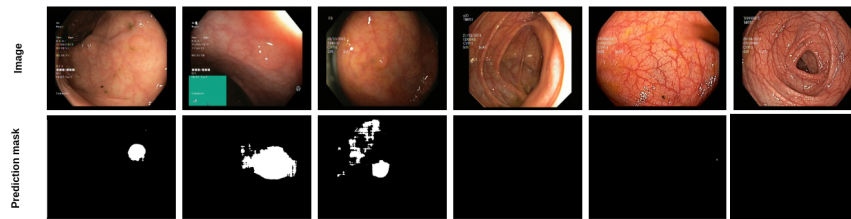


Figure 1: Qualitative results of the proposed solution. For first three images (left) algorithm have shown the predicted masks where as for other three (right), the algorithm does not produce any mask.

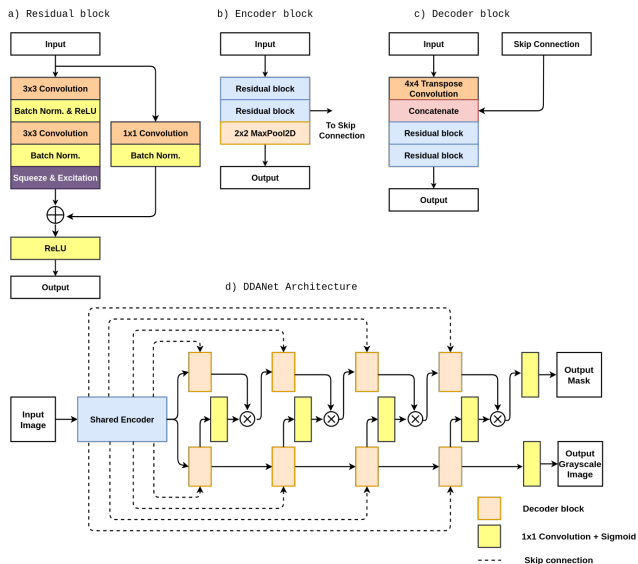


Figure 2: DDANet architecture and its components [3]

that consist of dense residual connections and is capable of transferring data across different scales. FANet [11] uses iterative mechanism to improve the predicted mask during the evaluation. All of these architectures have improved the baseline results. However, the architectural design considers only performance metrics such as dice coefficient or mean intersection over union.

Methodology

The block diagram of DDANet architecture and its components are shown in Figure 2. DDANet is an encoder-decoder architecture like ResUNet++ [9]. The DDANet combines the residual learning [12] and squeezing and excitation [13] networks' strengths. It is a fully convolutional network with a single encoder that is shared by two decoders. Detail explanation of the architecture and each component used in the DDANet can be found in [3]. Due to space limitations, we omit the detail explanations here, and the informed reader should refer [3] for more details on DDANet.

Results and Analysis

We present the results of our approach in Table 1. From Table 1, it can be observed that the DDANet model achieved a dice coefficient of 0.7967, Jaccard index of

Table 1: Results of our polyp segmentation method on the test set provided by MedAI 2021 challenge

Dataset	Dice	Jaccard	Recall	Prec.	Acc.
Development	0.8576	0.7800	0.8880	0.8643	-
Test	0.7967	0.7220	0.8214	0.8359	0.9557

0.7220 on the 2021 MedAI organizer's test dataset. Similarly, the model produced a DSC of 0.8576 and 0.7800 on the development dataset. Kvasir-SEG [14] was used as the development dataset in our study. The MedAI 2021 challenge organizers' provided 300 new, unseen images for testing. Although the result on the test dataset is less as compared to the development dataset. However, on the new test dataset, the performance is acceptable as the source dataset is not from the same distribution. It should also be noted that the scores presented here are calculated by the competition organizers. Figure 1 shows the qualitative results for the test samples. As we do not have the ground truth, we are only showing the images and the corresponding prediction masks. Here, we have selected the example polyps (flat polyps having less ≤ 10 mm) that are usually challenging to detect by a medical expert. The qualitative results show that our model is able to detect such types of polyps for two examples (please refer to Figure 1), whereas for the third example, it shows over-segmentation. Similarly, the test dataset has non-polyp images as well. Our trained segmentation model did not produce the segmentation masks for such types of image samples (please refer to the right side of Figure 1). We believe that these images are non-polyp frames. Thus, we can conclude from both qualitative and quantitative results that our model can act as the strong baseline for automated colorectal polyp segmentation.

Conclusions & Future Work

We have presented a DDANet-based architecture for automatic polyp segmentation. With the organizers' test dataset, we achieved a dice coefficient of 0.7967 and a Jaccard index of 0.7220. Thus, the proposed solution could be a strong baseline for the polyp segmentation tasks. In the future, DDANet could also be investigated on other bowel conditions and other biomedical datasets to demonstrate the efficiency of the proposed model.

References

1. Yamada M, Saito Y, Imaoka H, Saiko M, Yamada S, Kondo H, Takamaru H, Sakamoto T, Sese J, Kuchiba A, et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Scientific reports* 2019; 9:1–9
2. Lee J. Resection of diminutive and small colorectal polyps: What is the optimal technique? *Clinical endoscopy* 2016; 49:355
3. Tomar NK, Jha D, Ali S, Johansen HD, Johansen D, Riegler MA, and Halvorsen P. DDANet: Dual decoder attention network for automatic polyp segmentation. *International Conference on Pattern Recognition*. Springer. 2021 :307–14
4. Wang P, Xiao X, Brown JRG, Berzin TM, Tu M, Xiong F, Hu X, Liu P, Song Y, Zhang D, et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nature biomedical engineering* 2018; 2:741–8
5. Wei JW, Suriawinata AA, Vaickus LJ, Ren B, Liu X, Lisovsky M, Tomita N, Abdollahi B, Kim AS, Snover DC, et al. Evaluation of a deep neural network for automated classification of colorectal polyps on histopathologic slides. *JAMA network open* 2020; 3:e203398–e203398
6. Lieberman DA, Holub JL, Moravec MD, Eisen GM, Peters D, and Morris CD. Prevalence of colon polyps detected by colonoscopy screening in asymptomatic black and white patients. *Jama* 2008; 300:1417–22
7. Fan DP, Ji GP, Zhou T, Chen G, Fu H, Shen J, and Shao L. Pranet: Parallel reverse attention network for polyp segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2020 :263–73
8. Jha D, Smedsrud PH, Johansen D, Lange T de, Johansen HD, Halvorsen P, and Riegler MA. A comprehensive study on colorectal polyp segmentation with ResUNet++, conditional random field and test-time augmentation. *IEEE journal of biomedical and health informatics* 2021; 25:2029–40
9. Jha D, Smedsrud PH, Riegler MA, Johansen D, De Lange T, Halvorsen P, and Johansen HD. Resunet++: An advanced architecture for medical image segmentation. *2019 IEEE International Symposium on Multimedia (ISM)*. 2019 :225–2255
10. Srivastava A, Jha D, Chanda S, Pal U, Johansen HD, Johansen D, Riegler MA, Ali S, and Halvorsen P. MSRF-Net: A Multi-Scale Residual Fusion Network for Biomedical Image Segmentation. *arXiv preprint arXiv:2105.07451* 2021
11. Tomar NK, Jha D, Riegler MA, Johansen HD, Johansen D, Rittscher J, Halvorsen P, and Ali S. Fanet: A feedback attention network for improved biomedical image segmentation. *arXiv preprint arXiv:2103.17235* 2021
12. He K, Zhang X, Ren S, and Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016 :770–8
13. Hu J, Shen L, and Sun G. Squeeze-and-excitation networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018 :7132–41
14. Jha D, Smedsrud PH, Riegler MA, Halvorsen P, Lange T de, Johansen D, and Johansen HD. Kvasir-seg: A segmented polyp dataset. *International Conference on Multimedia Modeling*. 2020 :451–62
15. Hicks S, Jha D, Thambawita V, Riegler M, Halvorsen P, Singstad B, Gaur S, Pettersen K, Goodwin M, Parasa S, and Lange T de. MedAI: Transparency in Medical Image Segmentation. *Nordic Machine Intelligence* 2021