Iterative deep learning for improved segmentation of endoscopic images

Nikhil K Tomar¹, and Sharib Ali^{2,3}

1. *NepAI Applied Mathematics and Informatics Institute for Research (NAAMII), Kathmandu, Nepal*

2. *Department of Engineering Science, Big Data Institute, University of Oxford, Oxford, UK*

3. *E-mail any correspondence to: ali.sharib2002@gmail.com*

Abstract

Iterative segmentation is a unique way to prune the segmentation maps initialized by faster inference techniques or even unsupervised traditional thresholding methods. We used our previous feedback attention-based method for this work and demonstrate that with an optimal iterative procedure, our method can reach competitive accuracies in endoscopic imaging.

Keywords: deep learning; segmentation; polyps; instruments; endoscopy

Introduction

Endoscopic computer vision has been growing in the recent past years, and this has been possible with both the advent of deep learning methods and the availability of more publicly available datasets. In the clinical context, this is vital for improved patient care, especially when devising for challenging imaging technique such as endoscopy where the operator dependence is fairly high leading to nearly 6% to 28% miss-rates in polyps that can account for cancer deaths [1]. Procedures such as colonoscopy are considered the gold standard technique to find well known cancer precursors such as polyps. While the polyp localisation for large polyps $> 10mm$ might be trivial, diminutive, flat and serrated polyps are challenging [2] and are associated with a higher miss-rate. Even for large polyps that are usually removed by polypectomy, devising automated methods is of prime importance, mostly because of the probable polyp residues left is one of the major causes of interval cancer [3]. Similarly, associating instrument with resection and biopsy in endoscopic images can serve as information regarding possible locations of lesions in addition guide computer-aided tools to analyse such procedures.

In this work, we provide both validation and test results for the MedAI challenge [4]. We also demonstrate that our iterative deep learning-based pruning network can improve results of fast inference methods drastically. We experimented with our previously published DDANet [5] which has inference of nearly 70 frames-per-second (FPS) for polyp segmentation. We also provided results for our full iterative feedback attention network (FANet) [6]. In the 2021 MedAI challenge, we participated in both the polyp segmentation and instrument segmentation challenge tasks.

Materials and methods

We used the training and validation set splits as provided by the challenge organizers for both the Kvasir-SEG [7] and Kvasir-Instrument [8] dataset. The Kvasir-SEG dataset contains 1000 image and mask pairs, for which 880 images and masks pair were used for the training, and the remaining 120 were used for validation. For the Kvasir-Instrument dataset, we used 472 image and mask pairs for the training, and the remaining 118 for the validation purpose. Test dataset for both the polyp and instrument segmentation tasks included 300 images.

We used an iterative FANet architecture [6] and DDANet [5]. FANet is based on a feedback attention network that allows to rectify the predictions in an iterative fashion. The network consists of four encoder and four decoder layers. Similarly, DDANet is based on a dual decoder attention network with one shared encoder at each layer. While the iterative mechanism in full FANet (with 10 iterations) can adhere to larger computational time, DDANet has real-time performance (70 FPS), but sub-optimal output. We propose to use the segmentation maps from the DDANet output as input for the FANet iterative network for pruning. Even though our submissions for the challenge is based on a full FANet

Table 1: Results of our polyp segmentation methods for the validation set of Kvasir-SEG

Method	Iterations	F1	mIoU	Recall	Prec.	Acc.	F2
FANet	10	0.8599	0.7879	0.8928	0.8759	0.9600	0.8700
DDANet	NA	0.8238	0.7272	0.8607	0.8356	0.9502	0.8379
DDANet + FANet	3	0.8682	0.7971	0.8976	0.8844	0.9645	0.8775

Table 2: Results of our polyp segmentation methods on test set provided by MedAI challenge

Method	Iter.	F1	mIoU	Recall	Prec.	Acc.
FANet	10	0.7966	0.7334	0.8158	0.8374	0.9385
DDANet	NA	0.7935	0.7190	0.8182	0.8327	0.9555
DDANet + FANet	3	0.8100	0.7488	0.8275	0.8472	0.9506

iter. = iterations

Table 3: Results of our segmentation methods on test set provided by MedAI challenge

Class	Iter.	F1	mIoU	Recall	Prec.	Acc.
Polyp	10	0.7966	0.7334	0.8158	0.8374	0.9385
Tool	10	0.9073	0.8692	0.9236	0.9096	0.9854

iter., iterations

(single network), in this work, we further explore the pruning strength of FANet when initialized with masks from a faster network, in this case DDANet.

Results

It can be observed from our validation results (Table 1) that DDANet with pruned segmentation using iterative FANet produces a boost of nearly 4% on F1 score, and similar trends can be observed for other performance metrics. This is inline with the metric performance on the unseen test data presented in Table 2, where refined DDANet provided an improvement of nearly 2% on F1-score. Test results for both polyp and instrument data using iterative feedback and pruning method full FANet is shown in Table 3. Qualitative results (Figure 1) demonstrate the refinement ability of our iterative network using only 3 iterations over original DDANet segmentation maps. For example, most samples from DDANet (as shown Fig 1) had only partial segmentation maps which upon passing through FANet, it achieved results very close to ground truth masks. Similar, improvement over the UNet [9] approach on selected samples is presented in Figure 2.

Discussion

Combining faster but sub-optimal segmentation methods such as DDANet when fed to a FANet pruning network can provide improved segmentation maps at a near real-time speed. The iterative mechanism showed improved performance for both polyp and endoscopic tools when the output mask of DDANet was fed as an input to the FANet. Our experiments on both the polyp and instrument datasets demonstrate that the pruning

Figure 1: Qualitative results demonstrating the improvement of iterative pruning on DDANet, i.e., DDANet+FANet (only 3 iterations) for polyp segmentation predictions

Figure 2: Qualitative results demonstrating the improvement of iterative pruning on UNet, i.e., UNet+FANet (only 3 iterations) for instrument (tool) segmentation

mechanism are not dependent on the nature of the object of interest, for example, polyps which are most times blob like (Figure 1) and instruments which have mostly an elongated shape (Figure 2). Our proposed method showed improvement over individual algorithms (DDANet and FANet). Thus iterative mechanism improves the performance for the colorectal polyp segmentation and the instrument segmentation task.

References

1. Ahn SB, Han DS, Bae JH, Byun TJ, Kim JP, and Eun CS. The miss rate for colorectal adenoma determined by quality-adjusted, back-to-back colonoscopies. *Gut and liver* 2012; 6:64
2. Leufkens A, Van Oijen M, Vleggaar F, and Siersema P. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. *Endoscopy* 2012; 44:470–5
3. Anderloni A, Jovani M, Hassan C, and Repici A. Advances, problems, and complications of polypectomy. *eng. Clinical and experimental gastroenterology* 2014 Aug; 7:285–96. DOI: 10.2147/CEG.S43084. Available from: <https://doi.org/10.2147/CEG.S43084>
4. 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
5. 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
6. 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
7. 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*. Springer. 2020 :451–62
8. Jha D, Ali S, Emanuelsen K, Hicks SA, Thambawita V, Garcia-Ceja E, Riegler MA, Lange T de, Schmidt PT, Johansen HD, et al. Kvasir-instrument: Diagnostic and therapeutic tool segmentation dataset in gastrointestinal endoscopy. *International Conference on Multimedia Modeling*. Springer. 2021 :218–29
9. Ronneberger O, Fischer P, and Brox T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015 :234–41