Deep Conditional Adversarial Learning for Polyp Segmentation

Debapriya Banik, Debotosh Bhattacharjee
Department of Computer Science and Engineering, Jadavpur University, Kolkata-700032, India

ABSTRACT

This approach has addressed the *Medico automatic polyp segmentation challenge*, a part of Multimedia Evaluation 2020 (*MediaEval 2020*). We have proposed a deep conditional adversarial learning-based network for the automatic polyp segmentation task. The network is composed of two interdependent models, namely a generator and a discriminator. The generator network is a Fully Convolutional Network (FCN) employed for the prediction of the polyp mask while the discriminator enforces the segmentation to be as similar as the real segmented mask (ground truth). Our proposed model achieved a Jaccard index of 0.713, Dice Similarity Coefficient of 0.801, recall of 0.835, precision of 0.826, accuracy of 0.944, and F2 measure of 0.812 quantitatively evaluated by the organizers on the test dataset.

1 INTRODUCTION

Colorectal cancer (CRC) is a malignancy developed from the non-cancerous protrusions called polyps that develop in the colorectum's inner tissue lining [4]. Colonoscopy is considered a "gold standard" for the early diagnosis of CRC [14]. However, polyp miss rates are high during the exploration of the colorectum due to human factors such as lethargy, inadequate concentration, and workload. There is an estimated average miss-rate of 20% of polyps during the examination [3].

Polyp segmentation is a significant problem for the early diagnosis of CRC. So, a great deal of research has been dedicated to automatic polyp segmentation. In the last few years, deep learning techniques based on Generative Adversarial Networks (GANs)[5] have further revolutionized the state-of-the-art techniques in the computer vision domain [10]. Recently, a significant improvement has been seen in the field of medical image analysis [8]. Some GAN based approaches for polyp segmentation have been reported in the works of [11–13].

The task provided for the *Medico automatic polyp segmentation challenge* is to develop a computer-aided diagnosis system for automatic polyp segmentation with high efficiency and accuracy[6].

So, motivated by the insights of GANs in the domain of medical image analysis [8], we have proposed a deep conditional adversarial learning based network for automatic polyp segmentation.

2 APPROACH

This section describes our approach and proposed solution to the *Medico automatic polyp segmentation challenge* [6].

The proposed framework in Figure 1 is based on conditional generative adversarial learning, which involves two interdependent modules, a generator network (G) and the discriminator network (D). The generator network is responsible for the segmentation task, while the discriminator network enforces the segmented mask to be similar to the ground truth.

MediaEval'20, December 14-15 2020, Online

2.1 Preprocessing

The training challenge dataset [7] contains a limited number of samples, so to increase the size of the training dataset, we have applied 3 general transformation approaches: rotation (15° to 350°), flipping (vertical and horizontal), and translation (-5 to 30). Such transformations are invariant, and it is anticipated that clinicians can analyze medical images from different perspective without tampering the diagnostic results.

2.2 Proposed cGAN Architecture

The proposed cGAN network works in two stages. In the 1^{st} stage, the input grayscale image (P) is fed into the generator (G_{θ_G}), which is an FCN parameterized by θ_G where $\theta_G = \{W_G, b_G\}$ denotes the weights and biases of the encoder and decoder network. The encoder part consists of 12 convolutional layers, each with a kernel size of (3×3), while the decoder part contains 12 deconvolutional layers. Each convolutional layer in the encoder part is followed by batch normalization to avoid the risk of overfitting and is empowered by a Rectified Linear Unit (ReLu) activation function [1]. After a sequence of 3 convolutional layers, a MaxPooling layer[2] is used. The decoder part has a sequence of deconvolutional layers followed by batch normalization and ReLu as an activation function. An Upsampling strategy is followed in the decoder part to adjust the dimensions. The deconvolutional layers in the decoder part contain strides of different sizes respectively. We further implemented skip connections between the layers in the encoder and the decoder part to concatenate coarse, deep, and semantic information which will enhance the semantic representation of the output generated mask. Finally, the last layer in the generator network contains a convolutional layer with a kernel size of (1×1). The generator network is trained to produce the binary segmented generated mask(M).

In the 2^{nd} stage, the input image (P) and the binary mask containing the polyp region, which can either be the output binary generated mask (M) or the ground truth mask (Q) is concatenated and fed into the discriminator (D_{φ_D}) parameterized by (φ_D) where $\varphi_D = \{W_D, b_D\}$ denotes the weights and biases of the discriminator network. In the discriminator network, we have stacked a total number of 10 convolutional layers with batch normalization and ReLu activation function. Each of the layer has a kernel of size 3×3 and different strides. A skip connection is also applied so that the features learned in the initial layers can be sustained. Finally, the sigmoid function is applied to the output which takes real values varying between 0 and 1, where 0 signifies fake, and 1 signifies real.

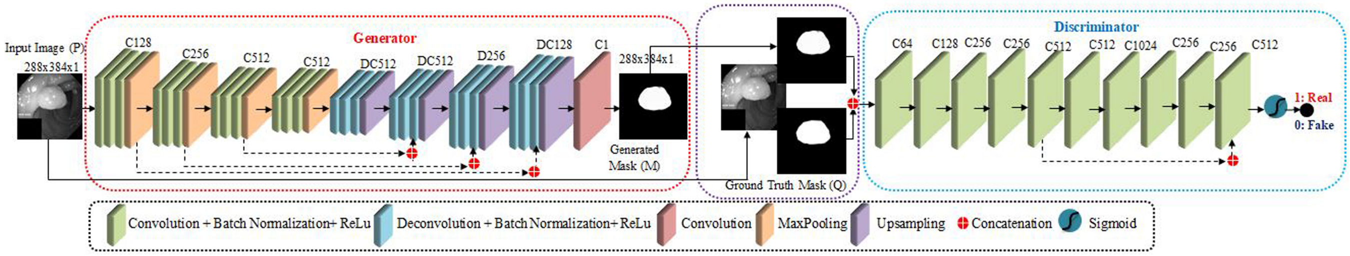

Figure 1: Proposed Framework for segmentation of polyp

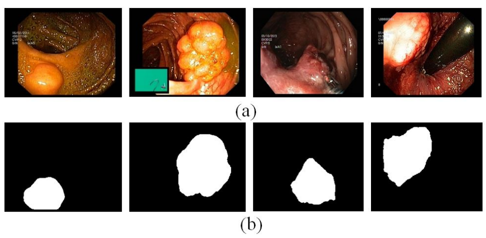

Figure 2: (a) Sample polyp images, and (b) Corresponding segment masks generated by the PA

2.3 Training Strategy

The training process of the proposed network alternates between 2 steps: (1) The generator is trained to produce a predicted synthetic mask by freezing the discriminator, and (2) the discriminator is trained while freezing the generator. The adversarial learning process of the network iteratively optimizes the parameters of the generator(θ_G) and the discriminator (φ_D) through the loss function. The error produced by the discriminator is backpropagated through the generator such that the segmented mask is consistent with the ground truth and can fool the discriminator. In this work, to optimize the generator network, we have used Binary Cross Entropy (BCE) and Mean Squared Error (MSE) loss functions weighted with an empirical weighing factor λ to minimize the training instability where λ is chosen to be 0.5. Similarly, we have used BCE as a loss function for the discriminator network. The total GAN loss is the sum of the loss for the generator network and the discriminator network. In this study, we have trained each of the generator and the discriminator networks for 35 epochs. We have used the Adam optimizer [9] with a learning rate of $1e-4$ and a batch size of 12.

3 RESULTS AND ANALYSIS

Our proposed segmentation technique is developed and evaluated on "Kvasir-Seg" dataset[7] provided by the organizers of the challenge.

In Figure 2, (a) we have included some polyp frames from the test dataset showing different variations of polyp in terms of shapes, sizes, and color, whereas (b) shows the corresponding predicted segmented mask by our Proposed Approach (PA). It can be well visualized that our proposed technique can efficiently segment the abnormal region (polyp) from the frame.

In Table 1, we have tabulated the official results evaluated by organizers of the challenge on the test dataset for our proposed approach (PA) towards the automatic polyp segmentation task. To justify the clinical relevance of the method, different measures have been considered, such as Jaccard Index (JI), Dice Similarity Coefficient (DSC), Precision (PRE), Recall(REC), Accuracy(ACC), and F2 score(F2). The JI and DSC determine the degree of overlap between the predicted mask and the ground-truth where 1 signifies perfect segmentation. Although, there were polyps with complex shapes and sizes, we achieved a considerable JI and DSC score by our PA. There is a significant trade-off between the PRE and REC, which justifies the consistency of the model, and as the REC is higher, the miss diagnosis rate can be considered as very low, and we have got a noticeable higher F2. We have also secured a competitive pixel ACC, which justifies that our PA can well classify the polyp pixels.

Table 1: Results of our approach on the first task of Medico automatic polyp segmentation challenge

Method	JI ↑	DSC ↑	PRE ↑	REC ↑	ACC ↑	F2 ↑
PA	0.713	0.801	0.828	0.835	0.944	0.812

4 CONCLUSION AND FUTURE WORK

In this study, we have proposed an approach for polyp segmentation by a deep adversarial learning technique. Our experimental results on the test dataset justify the effectiveness of our proposed technique. However, there were few polyp frames where polyps were imperceptible due to low lighting conditions, so in such cases, our algorithm fails to segment the polyp region. We believe that a much stronger GAN with an effective optimization technique can be specifically designed for the medical image segmentation task. Our work's future direction would be to implement an effective GAN and to evaluate our method on datasets of other medical image modalities to justify our method's robustness.

5 ACKNOWLEDGEMENTS

The first author is grateful to the Council of Scientific and Industrial Research (CSIR) for providing Senior Research Fellowship (SRF) under the SRF-Direct fellowship program (ACK No. 143416/2K17/1, File No. 09/096(0922)2K18 EMR-I). The authors are thankful for the Indo-Austrian joint project grant No. INT/AUSTRIA/BMWF/P-25/2018 funded by the DST, GOI, and the SPARC project (ID: 231) funded by MHRD, GOI.

REFERENCES

[1] Abien Fred Agarap. 2018. Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375* (2018).
[2] Nadeem Akhtar and U Ragavendran. 2020. Interpretation of intelligence in CNN-pooling processes: A methodological survey. *Neural Computing and Applications* (2020), 1–20.
[3] Sharib Ali, Felix Zhou, Barbara Braden, Adam Bailey, Suhui Yang, Guanju Cheng, Pengyi Zhang, Xiaoqiong Li, Maxime Kayser, Roger D Soberanis-Mukul, et al. 2020. An objective comparison of detection and segmentation algorithms for artefacts in clinical endoscopy. *Scientific reports* 10, 1 (2020), 1–15.
[4] Debapriya Banik, Kaushiki Roy, Debotosh Bhattacharjee, Mita Nasipuri, and Ondrej Krejcar. 2020. Polyp-Net: A Multi-model Fusion Network for Polyp Segmentation. *IEEE Transactions on Instrumentation and Measurement* (2020).
[5] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.
[6] Debesh Jha, Steven A. Hicks, Krister Emanuelsen, Håvard D. Johansen, Dag Johansen, Thomas de Lange, Michael A. Riegler, and Pål Halvorsen. [n.d.]. Medico Multimedia Task at MediaEval 2020:Automatic Polyp Segmentation.
[7] Debesh Jha, Pia H Smedsrud, Michael A Riegler, Pål Halvorsen, Thomas de Lange, Dag Johansen, and Håvard D Johansen. 2020. Kvasir-SEG: A Segmented Polyp Dataset. In *Proc. of International Conference on Multimedia Modeling (MMM)*. 451–462.
[8] Salome Kazeminia, Christoph Baur, Arjan Kuijper, Bram van Ginneken, Nassir Navab, Shadi Albarqouni, and Anirban Mukhopadhyay. 2020. GANs for medical image analysis. *Artificial Intelligence in Medicine* (2020), 101938.
[9] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
[10] Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. 2016. Semantic segmentation using adversarial networks. *arXiv preprint arXiv:1611.08408* (2016).
[11] Konstantin Pogorelov, Olga Ostroukhova, Mattis Jeppsson, Håvard Espeland, Carsten Griwodz, Thomas de Lange, Dag Johansen, Michael Riegler, and Pål Halvorsen. 2018. Deep learning and hand-crafted feature based approaches for polyp detection in medical videos. In *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 381–386.
[12] JM Poomeshwaran, Kumar S Santhosh, Keerthi Ram, Jayaraj Joseph, and Mohanasankar Sivaprakasam. 2019. Polyp Segmentation using Generative Adversarial Network. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 7201–7204.
[13] Younghak Shin, Hemin Ali Qadir, and Ilangko Balasingham. 2018. Abnormal colon polyp image synthesis using conditional adversarial networks for improved detection performance. *IEEE Access* 6 (2018), 56007–56017.
[14] David Vázquez, Jorge Bernal, F Javier Sánchez, Gloria Fernández-Esparrach, Antonio M López, Adriana Romero, Michal Drozdzal, and Aaron Courville. 2017. A benchmark for endoluminal scene segmentation of colonoscopy images. *Journal of healthcare engineering* 2017 (2017).