Segmentation of Polyp Instruments using UNet based deep learning model

Rishav Kumar Rajak[1], Ashar Beg Mirza[1]

1. *Chemical Engineering, Indian Institute of Technology, Delhi, India*

Abstract

In this paper, we present a UNet-based architecture used to segment polyps and instruments from the KVASIR Polyp Dataset [1] and KVASIR Intruments Dataset [2] provided at the MedAI Challenge 2021. For the polyp segmentation task, we developed a UNet-based algorithm for segmenting polyps in images taken from endoscopies. The main focus of this task is to achieve high segmentation metrics on the supplied test dataset. Similar to the polyp segmentation task, we also developed a UNet-based algorithm for segmenting instruments in the instrument segmentation task.

Keywords: Artificial Intelligence; Machine Learning; Deep Learning; Biomedical Image Segmentation

Introduction

The 2021 MedAI Challenge [3] focuses on three tasks based on input from experts in the field to address unique gastrointestinal picture segmentation issues. Two alternative segmentation situations are included, as well as a challenge on transparent machine learning systems, which underlines the importance of explainable and interpretable deep learning methods. Segmentation, as opposed to categorization and object identification, provides a more exact region of interest for a given class.

The task of categorising each pixel of an object of interest in medical images is known as medical image segmentation. Clinicians can use medical picture segmentation to focus on a specific area of a condition and extract comprehensive information for a more accurate diagnosis. The lack of a significant number of annotated, high-quality labelled images for training [2], low image quality, the lack of a consistent segmentation technique, and large variances in images among patients are the main issues associated with medical image segmentation. To estimate the performance of additional applications, segmentation accuracy and uncertainty must be quantified [4].This indicates the requirement for an automatic, generalizable, and efficient semantic image segmentation approach.

For automated medical image segmentation, Convolutional Neural Networks (CNNs) have demonstrated state-of-the-art performance [4]. A Fully Convolutional Network is an early Deep Learning (DL) architecture that was trained end-to-end for pixel-wise prediction for semantic segmentation tasks (FCN). Another prominent image segmentation architecture for pixel-wise prediction is U-Net, which is taught end-to-end. The U-Net architecture is divided into two sections: analysis and synthesis. Deep features are learned in the analysis path, and segmentation is conducted based on the learned features in the synthesis path.

For the Instrument Segmentation Development Task, we were given 590 images and their corresponding masks in JPEG and PNG format. For the Poly Segmentation Development Task, we were given 1000 images and their corresponding masks in JPEG format. These datasets were provided for training purpose.

Materials and methods

In this section, we discuss the approach and data pre-processing steps that carried out our model formulation and helped in getting the results.
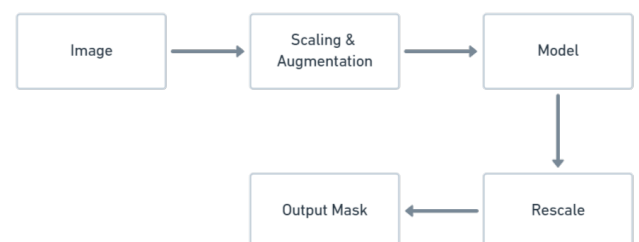

Figure 1: Data Pipeline

Data Pre-Processing

As the images in the KVASIR Polyp Dataset [1] and KVASIR Intruments Dataset [2] were of various shapes, we scaled them to 256×256 dimensions in order to train it. When trained on a larger dataset, neural networks perform better in the training and test sets. However, we have insufficient data for the instrumentation task. To address this, we used the albumentation library [5]. We used HorizontalFlip, VerticalFlip, ElasticTransform, GridDistortion and OpticalDistortion to increase out the dataset by 5 fold. Then, we divided our data into three parts: training, validation, and testing in an 80:10:10 ratio respectively randomly using scikit-learn library [6],
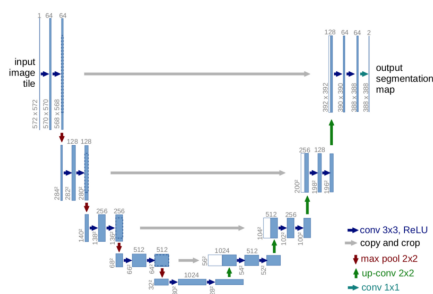
Model

Figure 2: UNet Architecture

UNet [7] is the state-of-the-art deep learning models based on neural networks that offers the best metrics score for biomedical picture segmentation, according to extensive literature surveys [8] [9] [10] in the field of biomedical image analysis. The U-Net architecture is a semantic segmentation system. The primary idea behind a "fully convolutional network" is to add layers to an existing contracting network, replacing pooling operators with upsampling operators. It can take two different paths: one that contracts and one that expands.

Model Tuning

As shown in the diagram below, the parameters to tune U-Net model for better performance falls into three categories. 1. Model Parameters: This includes 23 convolution layers, 64 filters at the first convolution layer and Batch Normalization 2. Image Parameters: In this we have used Pixel normalization as min-max, image cropping was set to be 256*256 and as discussed in the data preprocessing step, augmentation was used to be 5 times 3. Model Hyperparameters: This comprises of dice loss as loss function, optimizer was set to be adam and batch size was taken as 2. Finally the model was evaluated on Dice Coefficient, IoU, Accuracy, Precision and Recall.

Results

As shown in the figure 4 for instance, dice coefficient is 2 *times* the area of overlap divided by the total number of pixels in both ground truth and segmented mask for
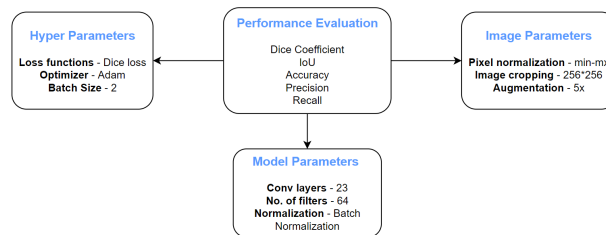

Figure 3: Parameters

Polyp task was found to be 0.7986 and 0.6298 for training and internal testing dataset respectively. Similarly IoU is the area of overlap between the predicted segmentation and the ground truth divided by the union area between the predicted mask and the ground truth came out to be 0.8955 and 0.1381 respectively.Here, the training column specifies the training data and test data corresponds internal test data which we kept during splitting.

Metrics	Polyp Task		Instrument Segmentation	
	Training	Test	Training	Test
Accuracy	0.9014	0.89586	0.98	0.81271
Dice Coefficient	0.7986	**0.62975**	0.9407	**0.276**
Precision	0.8652	0.77139	0.9608	0.28253
Recall	0.7175	0.63449	0.8988	0.3113
IoU	0.6864	0.52259	0.8955	0.18310

Figure 4: Performance Metrics

Discussion

The UNet design performs admirably in a variety of biomedical segmentation applications. We trained the model in our personal computers till 28 epochs due to lack of computation power. The test accuracy is of 10% dataset that we kept for evaluation during splitting. The UNet model has shown good accuracy on the polyp segmentation task, but we can observe that the score was not good in the instrument segmentation task. One of the main reason behind it is polyp masks are continuous circular kind of shape. But, in case of instrument dataset, the masks are abrupt, sparse and discontinuous. This can be resolved by increasing the complexity of the model like increasing neurons, use of transfer learning, etc. Our model gave similar results on testing dataset as on internal testing dataset in the case of intrumentation task which is around 0.3. Results got differ in Polyp case, here test score was low compare to our internal testing dataset.

We tracked our model on two main scores: Dice coefficient and IoU because these are the scores that give us an accurate idea about the quality of the mask predicted by the model.This metric ranges from 0–1, with 0 signifying no overlap and 1 signifying perfectly overlapping segmentation.

References

1. 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. 2020

2. Jha D, Ali S, Emanuelsen K, Hicks SA, Thambawita V, Garcia-Ceja E, Riegler MA, Lange T de, Schmidt PT, Johansen HD, Johansen D, and Halvorsen P. Kvasir-Instrument: Diagnostic and Therapeutic Tool Segmentation Dataset in Gastrointestinal Endoscopy. *MultiMedia Modeling.* 2021 :218–29

3. 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

4. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JA, van Ginneken B, and Sánchez CI. A survey on deep learning in medical image analysis. Medical Image Analysis 2017; 42:60–88. DOI: https://doi.org/10.1016/j.media.2017.07.005

5. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, and Kalinin AA. Albumentations: Fast and Flexible Image Augmentations. Information 2020; 11

6. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, and Duchesnay E. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 2011; 12:2825–30

7. Ronneberger O, Fischer P, and Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. LNCS 2015 Oct; 9351:234–41. DOI: 10.1007/978-3-319-24574-4_28

8. Shelhamer E, Long J, and Darrell T. Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 2017; 39:640–51. DOI: 10.1109/TPAMI.2016.2572683

9. Weng Y, Zhou T, Li Y, and Qiu X. NAS-Unet: Neural Architecture Search for Medical Image Segmentation. IEEE Access 2019; 7:44247–57. DOI: 10.1109/ACCESS.2019.2908991

10. Jha D, Riegler MA, Johansen D, Halvorsen P, and Johansen HD. DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS).* 2020 :558–64