

Identification and Characterization of Colorectal Polyps Using Deep Learning Methods

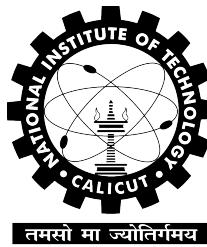
**CS4099D Project
End Semester Report**

Submitted by

**Abhinav P (B190460CS)
Athul Joseph (B190010CS)
Vishnu Ajay (B190368CS)**

Under the Guidance of

**Dr. Jayaraj P B
Assistant Professor**

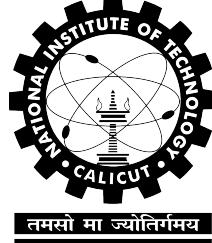


**Department of Computer Science and Engineering
National Institute of Technology Calicut
Calicut, Kerala, India - 673 601**

May 2023

**NATIONAL INSTITUTE OF TECHNOLOGY CALICUT
KERALA, INDIA - 673 601**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



CERTIFICATE

Certified that this is a bonafide report of the project work titled

**IDENTIFICATION AND CHARACTERIZATION OF
COLORECTAL POLYPS USING DEEP LEARNING
METHODS**

done by

Abhinav P

Athul Joseph

Vishnu Ajay

*of Eighth Semester B. Tech, during the Winter Semester 2020-'21, in
partial fulfillment of the requirements for the award of the degree of
Bachelor of Technology in Computer Science and Engineering of the
National Institute of Technology, Calicut.*

12-05-2023

Date

(Dr. Jayaraj P B)
(Assistant Professor)
Project Guide

DECLARATION

I hereby declare that the project titled, **Identification and Characterization of Colorectal Polyps Using Deep Learning Methods**, is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material that has been accepted for the award of any other degree or diploma of the university or any other institute of higher learning, except where due acknowledgement and reference has been made in the text.

Place : NIT Calicut

Date : 12-05-2023

Name : Abhinav P

Roll. No. : B190460CS

Name : Athul Joseph

Roll. No. : B190010CS

Name : Vishnu Ajay

Roll. No. : B190368CS

Abstract

Colonoscopy is currently the standard exam for colorectal cancer screening due to its ability to find and detect pre-cancerous lesions in the colon. However, its performance is significantly operator dependent. Our project uses deep learning methods to detect small mucoid growth, known as polyps, in the rectal region of the colon. The deep learning model would be partially trained for human anatomy detection in the initial stages of the project and would be trained further for successful polyp detection in subsequent phases. The input to our final model would be an endoscopic image, and the output would be either a diagram locating the positions of malignant colorectal polyps or a report.

ACKNOWLEDGEMENT

We would like to express our sincere and heartfelt gratitude to our guide and mentor Dr. Jayaraj P B and Dr Sunil Kumar, who have guided us throughout the course of the final year project. Without their active guidance, help, cooperation, and encouragement, we would not have made headway in the project. We would like to thank our parents and the faculty members for motivating us and being supportive throughout our work. We also take this opportunity to thank our friends who have cooperated with us throughout the course of the project.

Contents

1	Introduction	2
2	Problem Statement	3
3	Literature Survey	4
3.0.1	Filter bank-based techniques	5
3.0.2	Fast R-CNN	5
3.0.3	Improved Single Shot multi-box Detector(SSD)	6
3.0.4	U-Net: Convolutional Networks for Biomedical Image Segmentation	6
4	Proposed Work	8
4.1	Design	8
4.1.1	SSD - single shot multi-box detector	8
4.1.2	UNet	11
5	Experimental Results	15
5.0.1	SSD - single shot multibox detector	15

<i>CONTENTS</i>	iii
5.0.2 UNet based models	16
6 Conclusion	22
References	23

List of Figures

4.1	Command line utility for prepossessing	10
4.2	UNet Architecture	12
4.3	Sample output from the model (original image, predicted mask)	13
4.4	Post-processing to reduce noise.	14
5.1	Results of SSD (1) polyp detection	15
5.2	Results of SSD (2) multiple polyp detections	16
5.3	Mean IoU vs Epoch	17
5.4	Results of segmentation using UNet model (1)	18
5.5	Results of segmentation using UNet model (2)	18
5.6	Results of segmentation using UNet model (3)	19
5.7	Results of segmentation using UNet model (4)	19
5.8	Results of segmentation using UNet model (5)	19

5.9 Comparison of segmentation by different models(1)	20
5.10 Comparison of segmentation by different models(2)	20
5.11 Comparison of segmentation by different models(3)	20
5.12 Comparison of segmentation by different models(4)	21
5.13 Comparison of segmentation by different models(5)	21
5.14 Comparison of segmentation by different models(6)	21

List of Tables

5.1	Evaluation metrics of different segmentation models	17
5.2	Peak Mean IoU of different segmentation models	18

Chapter 1

Introduction

Colorectal Cancer (CRC) is one of the most dangerous cancers in the world and is the third most cause of death in India and the fourth cause of cancer deaths in the world [1,6]. It has been found that 85% of colorectal cases grow from adenomas because of genetic or epigenetic reasons. It can be reduced by endoscopic resection of colorectal polyps[7]. According to pathology, the polyps are categorized into four major categories: adenoma, sessile serrated adenoma/polyp (SSAP), and hyperplastic, including inflammation and juvenile polyps. Each of these categories has a different risk of developing cancer. The adenoma and SSAP polyps have a very high possibility of converting into cancer. On the other hand, hyperplastic polyps are less likely to develop into cancer[8]. According to the Preservation and Incorporation of Valuable Endoscopic Innovation (PIVI) strategy, the polyps smaller than 5 mm after resection can be omitted. Also, the hyperplastic polyps in the colon and rectum do not require sampling or endoscopic resection as they are non-malignant. Hence, the accurate classification of polyps can save a lot of risks, resources, and efforts of the patients and medical authorities[2].

Chapter 2

Problem Statement

Detect and classify colorectal polyps from colonoscopy images using a deep convolutional neural network.

Chapter 3

Literature Survey

The initial CADe systems were reported in the early 2000s. These systems were designed with a handcrafted algorithm based on certain polyp features and provided an accuracy of more than 90%. Several other groups designed and evaluated different handcrafted CADe solutions using small numbers of static images [9]. While these systems typically showed high accuracy on carefully chosen data sets, they were limited in real-world applications due to low sensitivity, high false-positive rates, and long processing time. More recently, deep-learning algorithms such as convolutional neural networks (CNNs) have been utilized to develop CADe systems, enabling the continuous recognition of abnormal lesions without needing external input. Using 50 polyp and 85 non-polyp videos, Misawa and colleagues developed a three-dimensional CNN-based CADe with a sensitivity and specificity of 90% and 63%, respectively[10].

3.0.1 Filter bank-based techniques

In [12], a filter bank-based technique was applied for the classification of CRC polyps. Filter-bank used filter masks for classifying different polyps [13]. In [14], a local binary pattern variant was used for the automatic classification of endoscopic images. The similarity among neighboring pixels was utilized to create a color vector field. A kNN classifier was used for classification. [15] used a wavelet transform for feature extraction and proposed three different approaches that can automatically classify colonic polyps. In [16], local features for detection and support vector machine (SVM) for classification were utilized. In [17], monogenic local binary pattern combined with Gabor filter was used to generate a new feature. The new feature was able to extract shape and edge details at multiresolution and prevented the color information. The use of linear discriminant analysis was made for reducing features. In [18], the features were extracted using segmentation techniques and further utilized for classification. CNN has reduced the use of handcrafted features for the extraction of features and classification.

3.0.2 Fast R-CNN

In [19], two-stage detection idea was used, in the positioning stage, loss function was improved based on intersection over Union (IoU) for bounding box regression, and use bilinear interpolation to improve the regions of interest (RoI) pooling operation to solve the problem of positioning deviation, in the recognition stage, the multi-scale convolution feature fusion was used to make the feature map contain more information, and used the improved non-maximum suppression (NMS) algorithm to avoid loss of overlapping objects. The results show that the algorithm has good performance on traffic signs whose resolution is in the range of (0, 32], the algorithm's recall rate reaches

90%, and the accuracy rate reaches 87%.

3.0.3 Improved Single Shot multi-box Detector(SSD)

The fastest algorithm which uses a single layer of convolutional network to detect the objects from the image is single shot multi-box detector (SSD) algorithm. In [20], classification accuracy of detecting objects was increased by improving the SSD algorithm while keeping the speed constant. These improvements have been done in their convolutional layers, by using depth-wise separable convolution along with spatial separable convolutions generally called multilayer convolutional neural networks. This method uses these multilayer convolutional neural networks to develop a system model which consists of multilayers to classify the given objects into any of the defined classes. The schemes then use multiple images and detect the objects from these images, labeling them with their respective class label. To speed up the computational performance, the algorithm is applied along with the multilayer convolutional neural network which uses a larger number of default boxes and results in more accurate detection. The accuracy in detecting the objects is checked by different parameters such as loss function, frames per second (FPS), mean average precision (mAP), and aspect ratio. Experimental results confirm that our proposed improved SSD algorithm has high accuracy.

3.0.4 U-Net: Convolutional Networks for Biomedical Image Segmentation

The U-Net paper[21] proposes a convolutional neural network architecture for biomedical image segmentation tasks. The network consists of a con-

tracting path to capture context and a symmetric expanding path to enable precise localization. The architecture uses skip connections to transfer information from the contracting path to the expanding path, which helps to preserve spatial information and improve segmentation accuracy. The authors demonstrate the effectiveness of the U-Net architecture on several biomedical image segmentation tasks, including cell segmentation, nuclei segmentation, and medical image segmentation. The U-Net architecture has become a widely used and influential method in the field of biomedical image analysis, with numerous applications and extensions.

Chapter 4

Proposed Work

4.1 Design

We have used 2 different approaches to segment the polyp - Single Shot multi-box Detector(SSD), Transfer learning on UNet-based models.

4.1.1 SSD - single shot multi-box detector

This model consists of multiple convolutional neural network layers, with each layer processing the input image and generating feature maps. These feature maps are then passed through subsequent layers for further processing, which helps extract increasingly higher-level features from the input image.

Model overview

The model starts with two convolutional layers (conv1_1 and conv1_2), followed by a max pooling layer (pool1). Then, the feature maps generated from the previous layer are passed through two more convolutional layers (conv2_1 and conv2_2), followed by another max pooling layer (pool2). This pattern is repeated for the following two sets of layers (conv3_x and pool3, and conv4_x and pool4) to generate feature maps of increasing complexity.

The final three layers of the model (conv5_x, pool5, and conv6) further process the feature maps generated from the previous layers. Finally, the output of the last convolutional layer (conv7) is fed into two branches, where the output from conv4_3 is used for object detection at lower resolution and conv7 is used for detection at higher resolution.

Algorithm

We used the SSD model mentioned above to train the dataset containing 1000 images. We also tried fine-tuning the VGG16 model from the Pytorch models database to fit the purpose of detecting colorectal polyps. The VGG16 model is an object detection model; this model was introduced for the SSD model to grasp the concept of bounding boxes quickly.

1. Dataset We have collected 90 images of the Colon(with and without polyp) from Govt. Medical College, Kozhikode. We have also used the publicly available dataset - CVC-ClinicDB and Kvasir-SEG. The dataset contains two folders - images and masks. The CVC-ClinicDB dataset had 612 PNG images, and the Kvasir-SEG dataset had 1000 JPG images. So in total, we got 1702 images.

2. Pre-processing The images in the Kvasir-SEG dataset were initially in

jpg format. we converted the images to PNG format. The datasets contained an image and the corresponding mask of the image depicting the colon area. A script was written to find the bounding box from the mask and convert it into VGG annotations format, an XML file.

A command line utility was created to easily convert images to png and annotate them in the VGG annotations format.

```
PS B:\projects\polyp\git upload\Colerectal-polyp-detection\src> python.exe .\pascalVOCAnnotator.py -h
usage: pascalVOCAnnotator.py [-h] [--convert] -original ORIGINAL -mask MASK [-csv CSV]

options:
  -h, --help            show this help message and exit
  --convert             convert jpg to png
  -original ORIGINAL   relative path of original image directory
  -mask MASK           relative path of mask image directory
  -csv CSV             relative path of bounding box csv
PS B:\projects\polyp\git upload\Colerectal-polyp-detection\src>
```

Figure 4.1: Command line utility for prepossessing

The images were resized to 300 x 300 pixels, and the mask values were converted from BGR to grayscale.

3. Training The model was trained for 50 epochs with two classes: class_1, the background class, a default class for the background, and class_2, the polyp class, which denotes the cancer region. The model was trained with a batch size of 4 and a learning rate of 0.001, which was found to be optimal using the trial and error method.

The model was trained to a 90:10 train-to-test ratio. During each epoch, the model is trained on batches of data, calculating the loss and performing backpropagation to update the model parameters. After each epoch, a checkpoint of that epoch is saved.

4. Post-processing The model takes a 300 x 300 normalized image as input and outputs the predicted locations, and scores are used to detect polyps in

the image. Many predicted (tensor coordinated) locations would have overlapping regions depicting the same polyp. Non-maximum suppression (NMS) removes all overlapping boxes other than the highest confidence box. The removal of unwanted boxes is done using the IOU strategy. This helps reduce the number of false positives and improves the model’s accuracy. Finally, the bounding box locations are scaled to the original image dimensions, and the detected boxes are drawn on the image.

4.1.2 UNet

U-Net is a convolutional neural network that was developed for biomedical image segmentation at the Computer Science Department of the University of Freiburg.

UNet Architecture

The network consists of a contracting path and an expansive path, which gives it the u-shaped architecture. The contracting path is a typical convolutional network consisting of repeated convolutions, followed by a rectified linear unit (ReLU) and a max pooling operation. During the contraction, the spatial information is reduced while feature information is increased. The expansive pathway combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the contracting path.

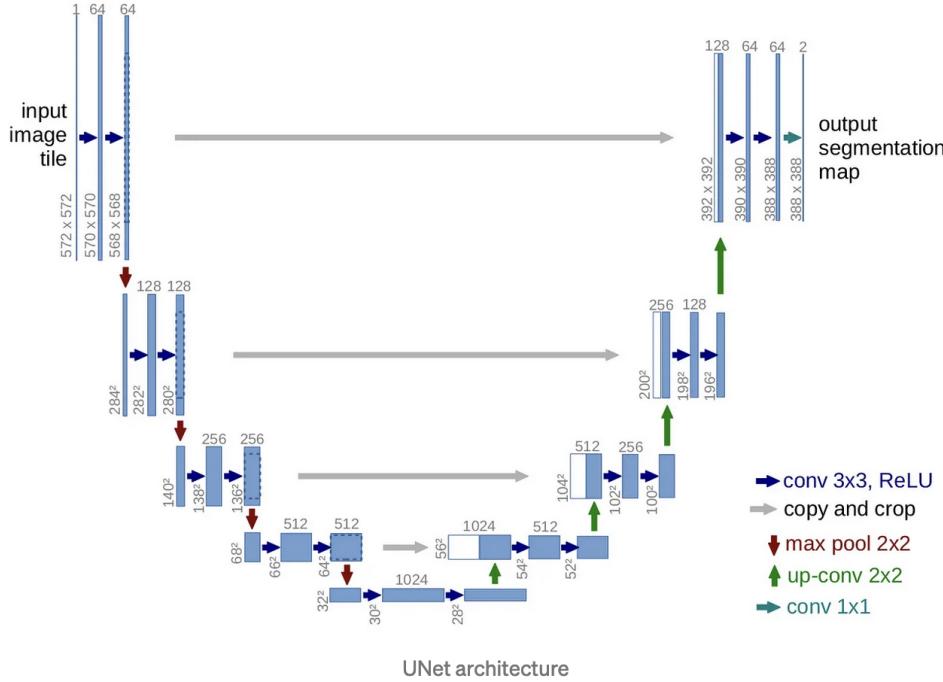


Figure 4.2: UNet Architecture

Algorithm

We have used pre-trained U-Net models such as - vgg_unet, resnet50_unet, and U-Net for segmentation. We have also used SegNet to compare the result against various U-Net models.

- 1. Dataset** We have collected 90 images of the Colon(with and without polyp) from Govt. Medical College, Kozhikode. We have also used the publicly available dataset - CVC-ClinicDB and Kvasir-SEG. The dataset contains two folders - images and masks. CVC-ClinicDB dataset had 612 PNG images, and the Kvasir-SEG dataset had 1000 JPG images. So in

total, we got 1702 images.

2. Pre-processing The images in the Kvasir-SEG dataset were converted into png format. The dataset was split using a 90-10 split. For training, the images were augmented to increase the dataset count. We have augmented the images by adding random translation, rotation, and flips. We have also resized the images to 256 x 256. The mask values are mapped to {0, 1}. After augmentation, the dataset count was 53,687.

3. Training We have used pre-trained models from the keras_segmentation library. We have tried different models like - unet, vgg-unet, resnet50-unet, segnet. All segmentation models take in similar parameters - n_classes and image size. Here n_classes represent the no of classes the image needs to be segmented into. We have used a value of 2 for n_classes - a region with polyp (1) and without polyp (0). The image size is 256 x 256. The model was trained for 50 epochs.

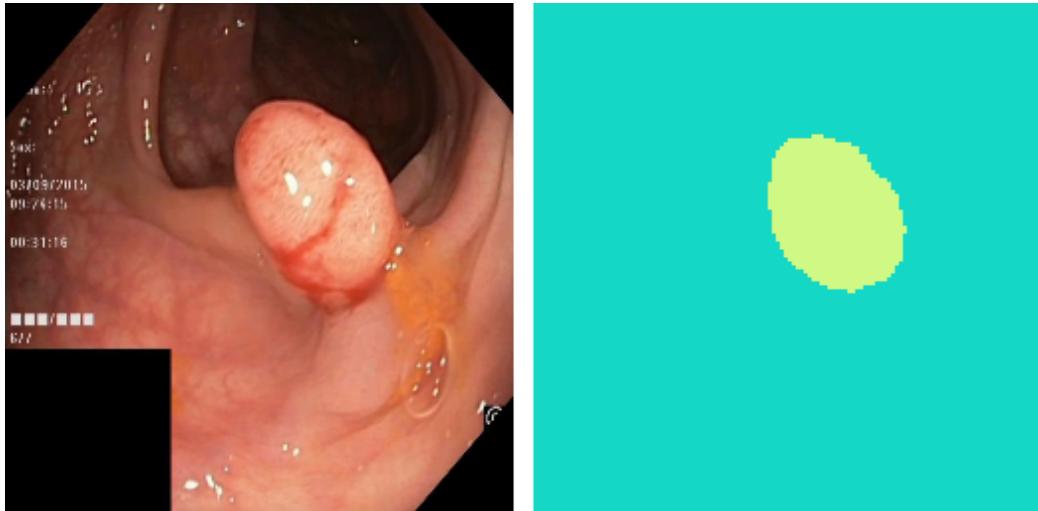


Figure 4.3: Sample output from the model (original image, predicted mask)

4. Post-processing The output from the model is noisy, and the saved mask is in BGR form. We read the saved mask and applied Gaussian blur with kernel size (7, 7). This has been shown to smoothen out the noise in the prediction. Then the values were mapped to the mask to {0, 1}. The resultant mask was multiplied by the actual image to get the segmented image.

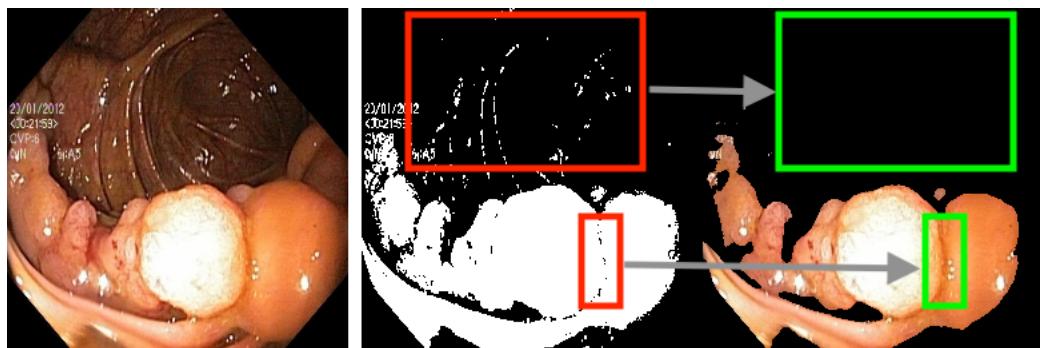


Figure 4.4: Post-processing to reduce noise.

Chapter 5

Experimental Results

5.0.1 SSD - single shot multibox detector

SSD gave a loss of 2.45 on the last epoch (50) during the training. The model was evaluated using IOU during the testing phase and gave an accuracy of 64.3%. There was no improvement in using VGG16 as the base model. The SSD trained only on the dataset we had provided good enough results.



Figure 5.1: Results of SSD (1) polyp detection

The SSD model gave a confidence score for each prediction it makes. The highest confidence score obtained is 0.7. The highest confidence score box is selected and drawn on the image. For better predictions, we show the first 4 top confidence score boxes to detect multiple polyps in the image.



Figure 5.2: Results of SSD (2) multiple polyp detections

5.0.2 UNet based models

The evaluation metrics used to evaluate the performance of the models were:

Intersection over Union(IoU): The IoU score for a given class is calculated as the intersection of the predicted and ground truth masks for that class divided by the union of the two masks.

Frequency Weighted IoU: This measure ensures that the classes that occur more frequently in the dataset have a greater impact on the overall score by weighing regions in the order of the frequency of their class.

Mean IoU: It is calculated as the average of the IoU scores across all classes.

Following table summarizes the performance of different models. The models were trained for 50 epochs.

Models	Training Accuracy	FW IoU (Testing)	Mean IoU (Testing)
SEGNET	94.10%	0.925	0.462
UNET	91.13%	0.842	0.656
RESNET50_UNET	93.12%	0.860	0.698
VGG_UNET	90.02%	0.838	0.641

Table 5.1: Evaluation metrics of different segmentation models

Mean IoU vs Epoch:

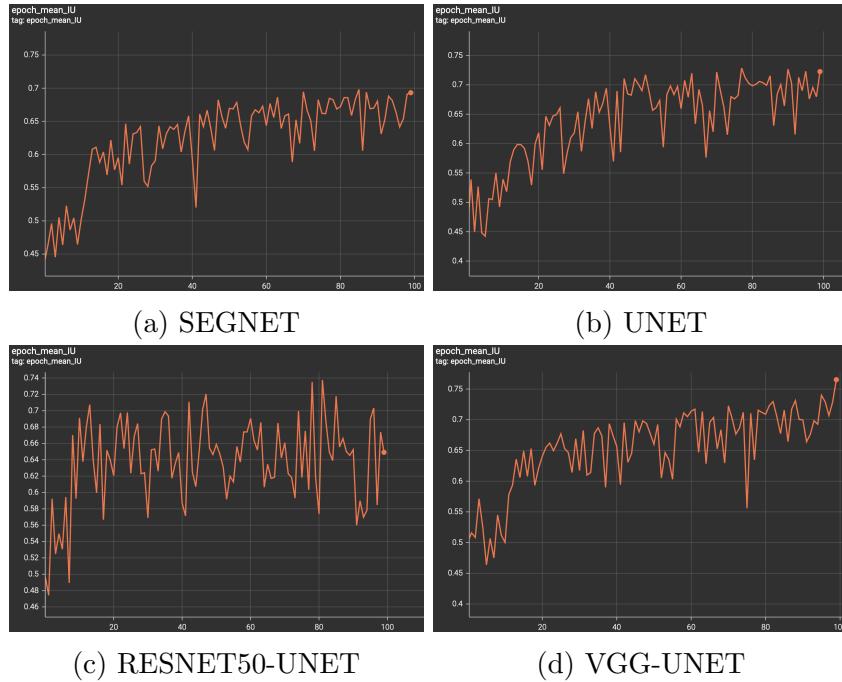


Figure 5.3: Mean IoU vs Epoch

Following table summarizes the performance of different models. The models were trained for 100 epochs.

Models	Peak Mean IoU	Epoch	Training Accuracy at that Epoch
SEGNET	0.6982	85	91.25%
UNET	0.7231	95	92.91%
RESNET50_UNET	0.7376	81	94.50%
VGG_UNET	0.7379	99	91.55%

Table 5.2: Peak Mean IoU of different segmentation models

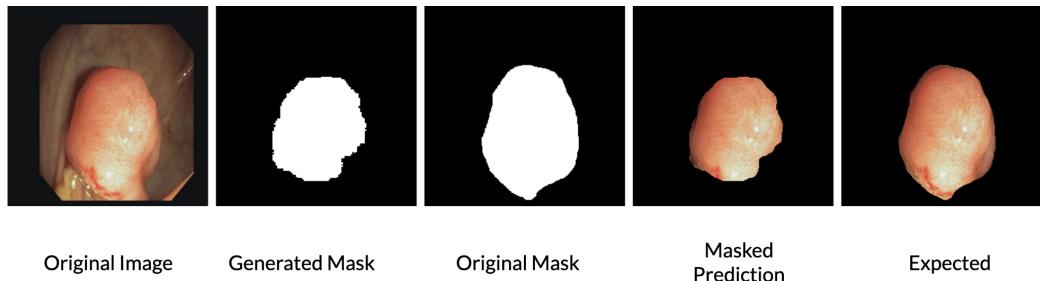
Results from UNet model:

Figure 5.4: Results of segmentation using UNet model (1)

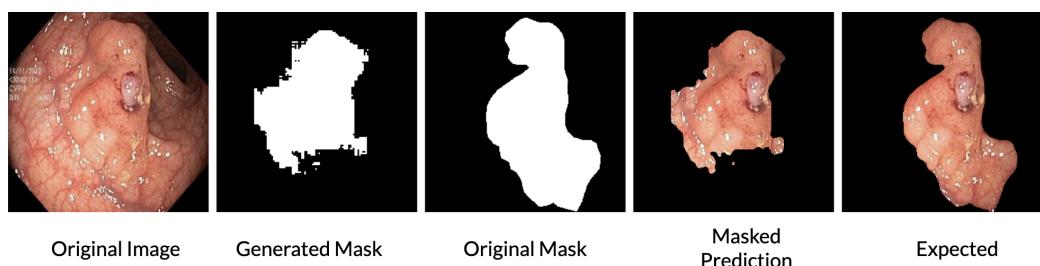


Figure 5.5: Results of segmentation using UNet model (2)

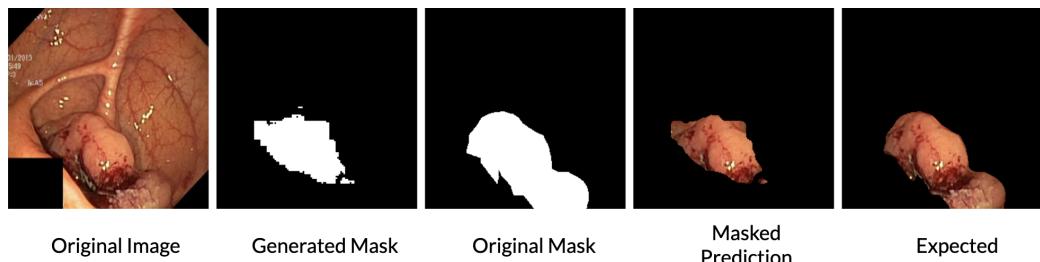


Figure 5.6: Results of segmentation using UNet model (3)

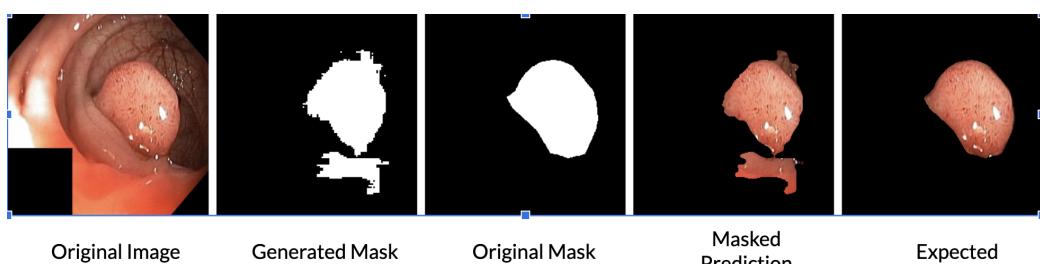


Figure 5.7: Results of segmentation using UNet model (4)

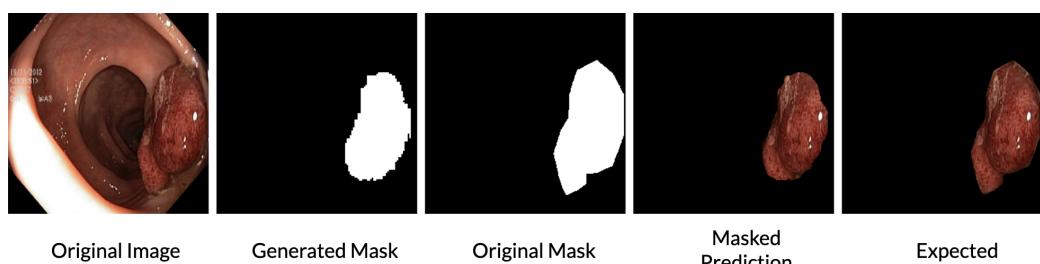


Figure 5.8: Results of segmentation using UNet model (5)

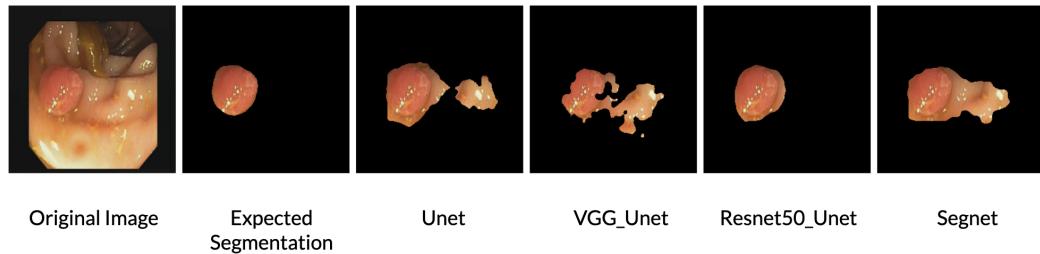
Comparing the output of different models:

Figure 5.9: Comparison of segmentation by different models(1)

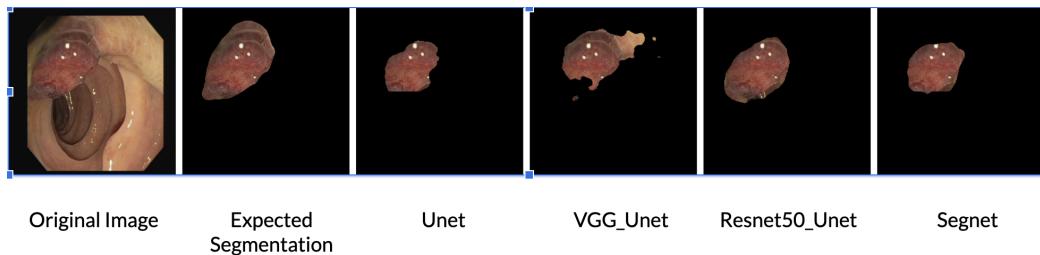


Figure 5.10: Comparison of segmentation by different models(2)

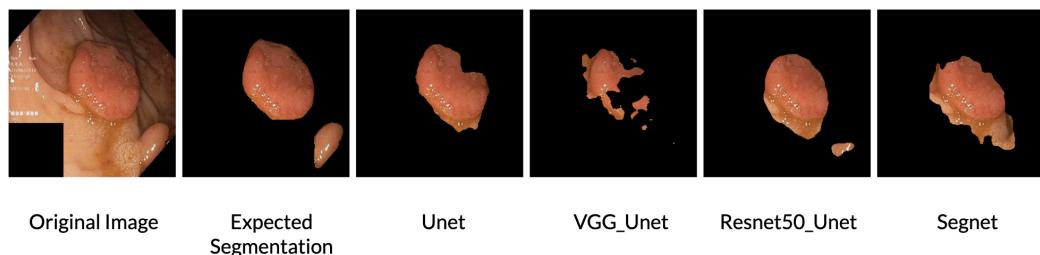


Figure 5.11: Comparison of segmentation by different models(3)

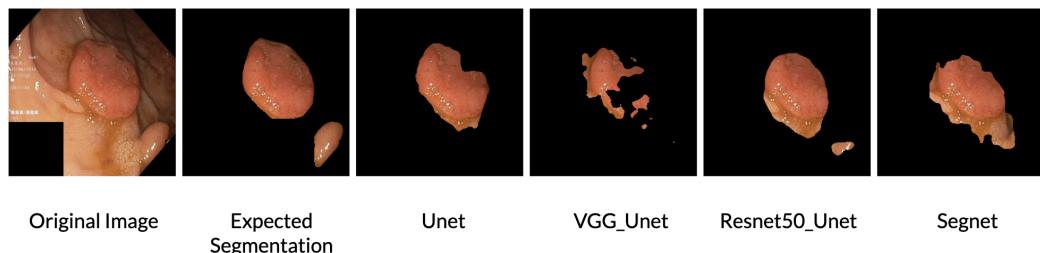


Figure 5.12: Comparison of segmentation by different models(4)

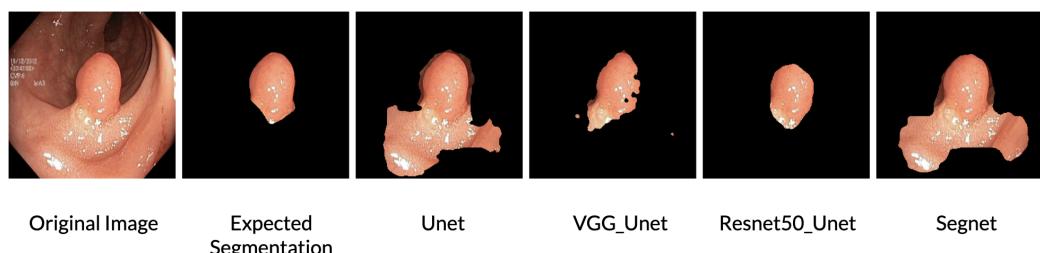


Figure 5.13: Comparison of segmentation by different models(5)

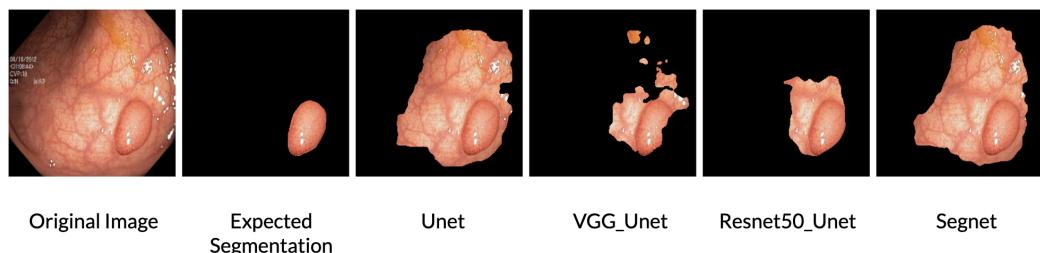


Figure 5.14: Comparison of segmentation by different models(6)

Chapter 6

Conclusion

We have tested different methods for the segmentation of colonic polyp - SSD and UNet-based segmentation models. On testing, we found that segmentation using UNet-based models was better than SSD. The UNet-based models also give a more precise boundary of the polyp.

Among UNet models, we found that RESNET50_UNET and VGG_UNET architecture performed the best. We trained these models for 100 epochs and compared their results. RESNET50_UNET model's performance degraded when training beyond 30 epochs, while the VGG_UNET model's performance improved as the epoch count increased. VGG_UNET model gave the best performance at the 99th epoch.

References

- [1] Parsa N, Byrne MF. Artificial intelligence for identification and characterization of colonic polyps. Ther Adv Gastrointest Endosc. 2021 Jun 29;14:26317745211014698. doi: 10.1177/26317745211014698. PMID: 34263163; PMCID: PMC8252334.
- [2] Tanwar S, Vijayalakshmi S, Sabharwal M, Kaur M, AlZubi AA, Lee HN. Detection and Classification of Colorectal Polyp Using Deep Learning. Biomed Res Int. 2022 Apr 15;2022:2805607. doi: 10.1155/2022/2805607. PMID: 35463989; PMCID: PMC9033358.
- [3] **Machine Learning — Andrew Ng, Stanford University [FULL COURSE]** YouTube playlist.
- [4] Kaggle free deep learning resources by Google.
- [5] **Artificial intelligence in Barrett's Esophagus** Nour Hamade and Prateek Sharma
- [6] **Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** CA Cancer J Clin 2018; 68: 394–424.

- [7] Brenner H and Stock CHM. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis.
- [8] Robertson DJ, Lieberman DA, Winawer SJ, et al. Colorectal cancers soon after colonoscopy: a pooled multicohort analysis. Gut 2014; 63: 949–956.
- [9] Parsa N, Byrne MF. Artificial intelligence for identification and characterization of colonic polyps. Ther Adv Gastrointest Endosc. 2021 Jun 29;14:26317745211014698. doi: 10.1177/26317745211014698. PMID: 34263163; PMCID: PMC8252334.
- [10] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394–424. [PubMed] [Google Scholar]
- [11] Brenner H, Stock CHM. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. BMJ 2014; 348: g2467.
- [12] Häfner M., Tamaki T., Tanaka S., Uhl A., Wimmer G., Yoshida S. Local fractal dimension based approaches for colonic polyp classification. Medical Image Analysis. 2015;26(1):92–107. doi: 10.1016/j.media.2015.08.007.
- [13] Wimmer G., Uhl A., Häfner M. A novel filterbank especially designed for the classification of colonic polyps. 2016 23rd International Conference on Pattern Recognition; Dec. 2016; Cancun, Mexico.
- [14] Häfner M., Liedlgruber M., Uhl A., Vécsei A., Wrba F. Color treatment in endoscopic image classification using multi-scale local

- color vector patterns. *Medical Image Analysis.* 2012;16(1):75–86. doi: 10.1016/j.media.2011.05.006.
- [15] Wimmer G., Tamaki T., Tischendorf J. J. W., et al. Directional wavelet-based features for colonic polyp classification. *Medical Image Analysis.* 2016;31:16–36. doi: 10.1016/j.media.2016.02.001.
- [16] Tamaki T., Yoshimuta J., Kawakami M., et al. Computer-aided colorectal tumor classification in NBI endoscopy using local features. *Medical Image Analysis.* 2013;17(1):78–100. doi: 10.1016/j.media.2012.08.003.
- [17] Yuan Y., Meng M. Q.-H. A novel feature for polyp detection in wireless capsule endoscopy images. 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems; Sept. 2014; Chicago, IL, USA0. pp. 5010–5015.
- [18] Stehle T., Auer R., Gross S., et al. Classification of colon polyps in NBI endoscopy using vascularization features. *Medical Imaging 2009: Computer-Aided Diagnosis.* 2009;7260:774–785.
- [19] C. Cao et al., "An Improved Faster R-CNN for Small Object Detection," in *IEEE Access*, vol. 7, pp. 106838-106846, 2019, doi: 10.1109/ACCESS.2019.2932731.
- [20] Kumar, A., Zhang, Z.J. Lyu, H. Object detection in real-time based on improved single shot multi-box detector algorithm. *J Wireless Com Network* 2020, 204 (2020).
- [21] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pp. 234-241.