

# **Identification and Segmentation of Barrett's Esophagus Using Deep Learning Methods**

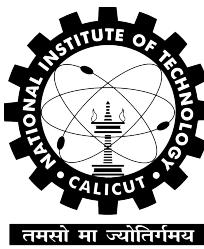
**CS4099D Project  
End Semester Report**

*Submitted by*

Akhil Sarwar T H (B190486CS)  
Tom Saju (B191290CS)  
Nishant Kumar Bhardwaj (B190427CS)

*Under the Guidance of*

**Dr. Jayaraj P B  
Assistant Professor**

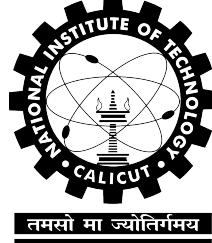


**Department of Computer Science and Engineering  
National Institute of Technology Calicut  
Calicut, Kerala, India - 673 601**

**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 SEGMENTATION OF BARRETT'S  
ESOPHAGUS USING DEEP LEARNING METHODS**

*done by*

**Akhil Sarwar T H**

**Tom Saju**

**Nishant Kumar Bhardwaj**

*of Eighth Semester B. Tech, during the Winter Semester 2022-'23, 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 Segmentation of Barrett's Esophagus 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 which 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 acknowledgment and reference has been made in the text.

Place : NIT Calicut

Date: 12-05-2023

Name: Akhil Sarwar T H

Roll. No.: B190486CS

Name: Tom Saju

Roll. No.: B191290CS

Name: Nishant Kumar Bhardwaj

Roll. No.: B190427CS

## **Abstract**

Barrett's Esophagus is a condition in which the flat pink lining of the esophagus becomes damaged by acid reflux, which causes the lining to thicken and become red. Barrett's Esophagus is associated with an increased risk of developing esophageal cancer. Upper endoscopy is currently the standard test for diagnosing Barrett's Esophagus due to its ability to remove (biopsy) tissue samples from any abnormal areas using special instruments through the scope. Despite advances in endoscopic imaging modalities, there are still significant miss rates of dysplasia and cancer in Barrett's Esophagus, moreover, its performance is significantly operator dependent. Our project uses deep learning methods to detect dysplasia, in the Oesophageal region, particularly the Lower Esophageal Sphincter. Studies have shown AI systems have a sensitivity of more than 90% and specificity of more than 80% in detecting Barrett's related dysplasia and cancer. Our aim is to develop a deep-learning computer-aided system for the detection of Barrett's Esophagus using endoscopic images.

## **ACKNOWLEDGEMENT**

We would like to express our sincere and heartfelt gratitude to our guide and mentor Dr.Jayaraj P B, who guided us throughout the final year project. We also thank Dr Sunil Kumar (Gastroenterologist, Kozhikode Medical College) for his guidance. We also thank Mr.Gopakumar G and Dr.Manjusha K for evaluating us and for their valuable feedback and supervision which helped us to enhance our approach to completing the project. Without their active guidance, help, cooperation, and encouragement, we would not have made headway in the project. We would like to thank our 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

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Problem Statement</b>	<b>4</b>
<b>3</b>	<b>Literature Survey</b>	<b>5</b>
3.0.1	Use of AI with white light and virtual chromoendoscopy	5
3.0.2	Use of AI in Volumetric Laser Endomicroscopy (VLE)	6
3.0.3	Published systematic reviews and meta-analyses . . . . .	7
3.0.4	Histopathology interpretation and AI . . . . .	8
3.0.5	Quality control and AI . . . . .	8
<b>4</b>	<b>Proposed Work</b>	<b>9</b>
4.1	Design . . . . .	9
4.1.1	UNet . . . . .	9
<b>5</b>	<b>Experimental Results</b>	<b>13</b>
5.0.1	UNet based models . . . . .	13

<i>CONTENTS</i>	iii
<b>6 Conclusion</b>	<b>16</b>
<b>References</b>	<b>17</b>

# List of Figures

4.1	UNet Architecture . . . . .	10
4.2	Sample output from the model (original image, predicted mask) . . . . .	12
5.1	Mean IoU vs Epoch . . . . .	14
5.2	Comparison of segmentation by different models(1) . . . . .	15
5.3	Comparison of segmentation by different models(2) . . . . .	15
5.4	Comparison of segmentation by different models(3) . . . . .	15
5.5	Comparison of segmentation by different models(4) . . . . .	15

# List of Tables

5.1 Evaluation metrics of different segmentation models . . . . .	14
---	----

# Chapter 1

## Introduction

Barrett's Esophagus (BE) is the only known premalignant precursor for Esophageal Adenocarcinoma (EAC) and is thought to follow a linear progression from nondysplastic BE to low-grade dysplasia to high-grade dysplasia (HGD) and finally to cancer.[1] Early detection of dysplastic lesions and cancer confined to the mucosa allows for minimally invasive curative endoscopic treatment, which provides a less invasive method of treatment than surgical resection and/or neo adjuvant therapy for advanced lesions.[2]

However, neoplasia within BE may be subtle and difficult to recognize, with a recent meta-analysis showing high miss rates of around 25% for HGD and cancer within 1 year of a negative index examination.[3] The reasons for this are likely multi-factorial, including the lack of recognition of subtle lesions, lack of detailed inspection of the esophageal mucosa, non optimum cleaning techniques, and less experienced endoscopists.[4] In addition, the currently recommended Seattle biopsy protocol for tissue sampling has been shown to sample 5% or less of the entire BE mucosa thereby potentially missing focal areas of neoplasia.[5] Faced with these challenges, the need for

better detection of neoplastic areas is, thus, paramount. One possible way to improve detection of these lesions is through enhanced imaging technologies such as volumetric laser endomicroscopy (VLE), confocal laser endomicroscopy (CLE), virtual chromoendoscopy, and dye-based chromoendoscopy. The American Society of Gastroenterology (ASGE) has released criteria for the preservation and incorporation of valuable endoscopic innovations (PIVI) to evaluate the use of advanced imaging techniques. It requires that an imaging modality have a per patient sensitivity of 90%, specificity of 80%, and NPV of 98% in detecting HGD and EAC in order to replace four quadrant biopsies.[6] Of the currently available technologies, dye-based and virtual chromoendoscopy are the only two technologies that meet these criteria.[6] However, the use of these image enhancing technologies is operator dependent and subjective, especially by non experts.[7,8] Artificial intelligence (AI) has emerged in recent years as a promising tool in improving clinical performance in gastrointestinal (GI) endoscopy. The hope is that computer-aided diagnosis will play an adjunct role to endoscopists in the early detection and characterization of neoplastic lesions in BE patients.

# **Chapter 2**

## **Problem Statement**

Develop a deep-learning-based system for detecting Barrett's Esophagus using endoscopic images.

# Chapter 3

## Literature Survey

### 3.0.1 Use of AI with white light and virtual chromoendoscopy

One of the first studies that demonstrated the use of AI with white light and chromoendoscopy in the detection of Lesion was done by using 100 images from 44 Barrett's Esophagus(BE) patients[9]. The algorithm was able to detect neoplastic lesions with specificity of 87% and sensitivity of 86%. A Real Time use of AI in the evaluation of BE was done to validate a CNN system[10] to detect esophageal adenocarcinoma (EAC) in real time with the endoscopic examination of 14 patients using 62 images, and showed a sensitivity of 83.7% and specificity of 100%. Hashimoto[11] developed a CNN trained by 916 images of BE and validated with 458 images with a reported accuracy of 95.4% for detection of early neoplasia.

Going beyond still images, de Groof et al. [19] performed one of the initial studies of CAD on live endoscopic procedures on 20 patients; 10 with BE dysplasia and 10 without dysplasia. The sensitivity of the system per

level analysis was 91% and specificity was 89% for detection of BE neoplasia.

**Lesion Characterization**—A more recent study used a computer aided diagnostic (CAD) system (Hybrid ResNet-UNet) that classified images as nondysplastic or dysplastic with sensitivity of 90% and specificity of 88% and achieved higher accuracy than nonexpert endoscopists[9]

**Determining depth of invasion**—In a study by Ebigbo et al., a deep learning system was trained and tested to differentiate between T1a and T1b Barrett's cancer using 230 white light endoscopic images (108 T1a and 122 T1b) from three tertiary care centers and compared to experts' classification. The sensitivity, specificity, and accuracy of the AI system 0.77, 0.64, and 0.71, respectively, in differentiating T1a from T1b, with no significant difference to that of experts, indicating that accurate prediction of submucosal invasion remains challenging for both experts and AI. [12]

### **3.0.2 Use of AI in Volumetric Laser Endomicroscopy (VLE)**

Interpretation of VLE images from BE patients can be quite difficult and requires a steep learning curve. An AI software called intelligent real-time image segmentation has been developed to identify VLE features by different color schemes. A pink color scheme indicates a hyper-reflective surface which implies increased cellular crowding, increased maturation, and a greater nuclear to cytoplasmic ratio. A blue color scheme indicates a hypo-reflective surface which implies abnormal BE epithelial gland morphology. An orange color scheme indicates lack of layered architecture which differentiates squamous epithelium from BE[13]. Another study created an algorithm to identify early BE neoplasia on ex vivo VLE images showing a sensitivity of 90% and specificity of 93% in detection with better performance than the clinical

VLE prediction score.[14] A CAD system reported by Struveynberg et al. analyzed multiple neighboring VLE frames and showed improved neoplasia detection in BE with an area under the curve of 0.91 [15]

### 3.0.3 Published systematic reviews and meta-analyses

The main drawback of majority of the published studies is the lack of external validity, generalizability, as well as the limited sample sizes to adequately power for diagnostic accuracy as there is a need to annotate large test datasets. A recent systematic review and meta-analysis by Bang et al.[16] included 19 studies (10 image-based and 9 patient-based) on the use of AI in detecting esophageal cancer. The majority of the included studies used white light examination (except 7 with narrow-band imaging) and a convolutional neural network CAD algorithm (except 5 studies that used support vector machines). For the image-based studies, the sensitivity was 94% (95% CI, 89–96%) and specificity was 88% (95% CI, 76–94%) for detection of neoplasia. For the patient-based studies, the sensitivity was 93% (95% CI, 86–96%) and the specificity was 85% (95% CI, 78–89%) [14] for detection of neoplasia.

Another systematic review and meta-analysis by Lui et al. included 561 endoscopic images of patients with Barrett's esophagus (6 studies) showing the pooled sensitivity of detection of neoplastic lesions to be 88% (95% CI, 82–92.1%), specificity to be 90.4% (95% CI, 85.6–94.5%), and area under the curve of 0.96 (95% CI, 0.93–0.99). From the included studies, 3 used CNN and 3 studies non-CNN models with no significant difference in the performance between the two models.[17]

### 3.0.4 Histopathology interpretation and AI

Histological diagnosis of BE associated neoplasia is challenging and is another area where AI may prove to be useful. [18] This may be especially true with low-grade dysplasia which has been shown to have a very low interobserver agreement even among expert histopathologists. [18] A study by Tomita et al. [19] using a CNN-based algorithm to classify histology images into nondysplastic BE, dysplastic BE, and EAC showed a classification accuracy of 0.85, 0.89, and 0.88, respectively.

### 3.0.5 Quality control and AI

An AI system named ENDOANGEL has been studied by Chen et al. [20] to prompt identification of blind spots during upper endoscopy, provide a grading score of percentage of mucosa that is adequately visualized, and measure inspection time to precisely determine the quality of the examination. Such use of AI will be beneficial moving forward to monitor and record the quality of exams as physician reimbursement rates likely will be increasingly based on outcome measures performance with the goal of providing value-based care. [21] In addition, as these AI quality systems become more sophisticated, it is possible that their use will expand to other applications, such as providing real-time feedback and objective metrics that can be used to train young endoscopists.

# **Chapter 4**

## **Proposed Work**

### **4.1 Design**

For barrett's segmentation we have used Unet based models like Unet based on vanilla CNN, VGG16, Resnet50 and Segnet

#### **4.1.1 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**

Unet is the architecture used for semantic segmentation. It consists of a contracting network(downsampling) and an expanding network(upsampling).

The downsampling network consists of repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2. The feature channels are doubled. Every step in the expansive path consists of an upsampling of the feature map followed by a 2x2 convolution (“up-convolution”) that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers[24].

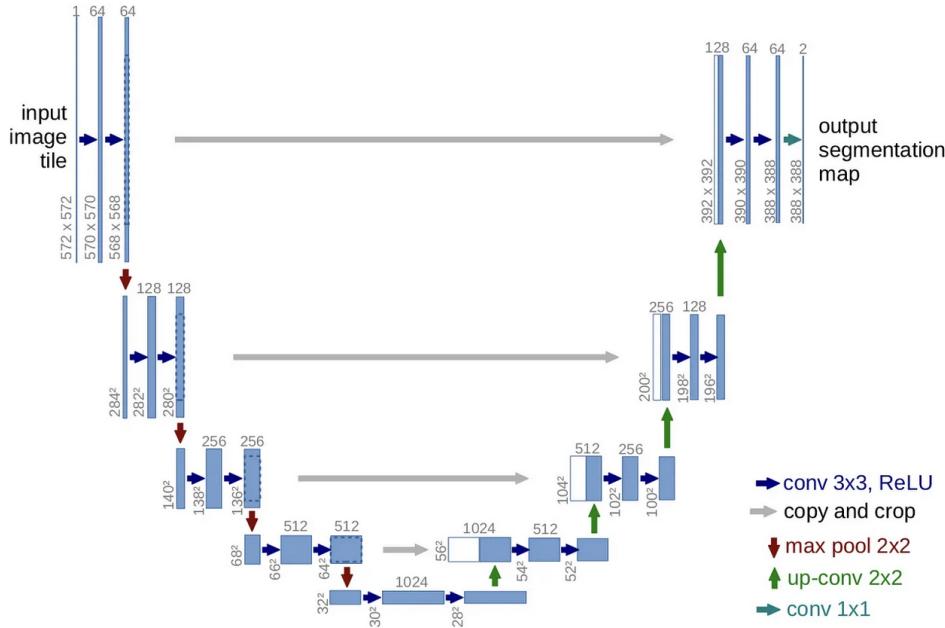


Figure 4.1: UNet Architecture

## Implementation

- 1. Collection of Dataset** We had initially collected 30-40 endoscopic images of Lower Esophageal Sphincter (LES) from Medical College Kozhikode. Later about 92 samples having barretts were identified from Hyper-Kvasir[25] dataset which is publically available and it was used for segmentation model.
- 2. Dataset Annotation** BE mucosa were annotated using an online tool ([https://www.robots.ox.ac.uk/~vgg/software/via/via\\_demo.html](https://www.robots.ox.ac.uk/~vgg/software/via/via_demo.html)). The SCJ (Squamous Columnar Junction) was marked using this tool and resulting annotations were downloaded as a json file. We then used OpenCv to convert the resulting annotations to black and white masks which will then be used as the ground truth for segmentation.
- 3. Preprocessing** As the amount of data available for training is limited we have applied augmentation techniques to increase the amount of data. We augmented the collected 92 sample images to 296 images for training the model. Augmentations include random brightness alteration, horizontal flip, and random rotation. Alpha value (contrast) was chosen to be between (1, 1.5). Beta value (brightness) was chosen to be between (-3, 5). Rotation is random between (0, 180) degree. Although rotation transformation is not safe on images showing 6 and 9 in digit recognition applications, it is generally safe on medical images[26]. Image format is converted from jpg to png as there is a chance of loss of information when using JPG[26]. Image flipping is only done horizontally. Vertical flipping is rarely preferred since the bottom and top regions of an image may not always be interchangeable [26].
- 4. Training** Pre-trained models of keras-segmentation libarary was used for

training and prediction. Different variations of Unet (vanilla CNN, vgg16, resnet50) and segnet were used. All the models were trained on the dataset for 10 epochs. Images were resized to 608 \* 416. Model was configured to predict 2 classes - Either mucosa or not.

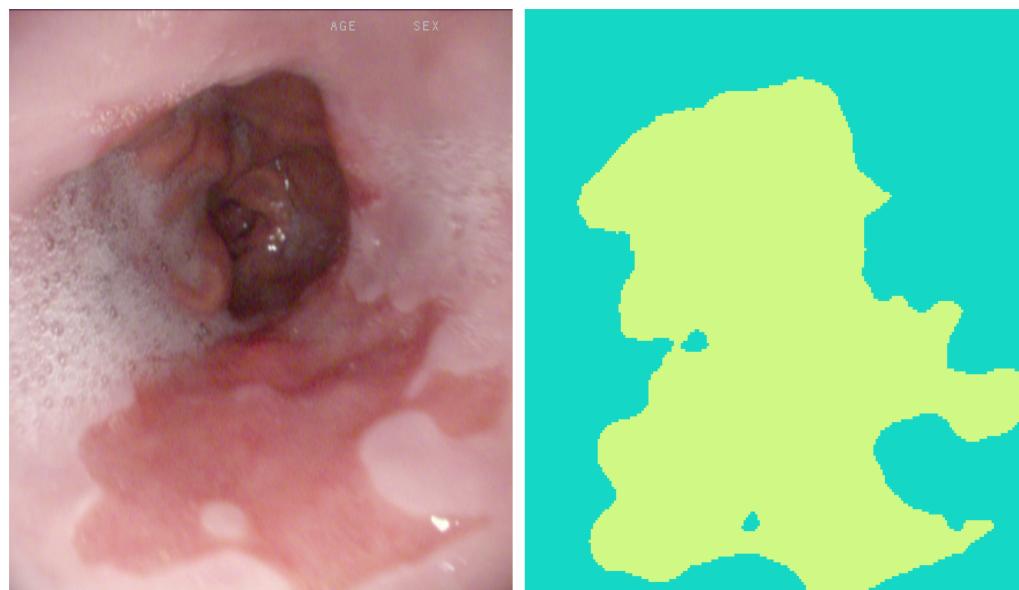


Figure 4.2: Sample output from the model (original image, predicted mask)

**5. Post Processing** Post processing step involves the conversion of color values of different classes to 0 and 255 (black and white). Comparisons are made by generating collaged images.

# Chapter 5

## Experimental Results

### 5.0.1 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 (Training)	Mean IoU (Training)
SEGNET	83.04%	0.706	0.657
UNET	83.54%	0.775	0.713
RESNET50_UNET	88.57%	0.728	0.679
VGG_UNET	84.76%	0.748	0.671

Table 5.1: Evaluation metrics of different segmentation models

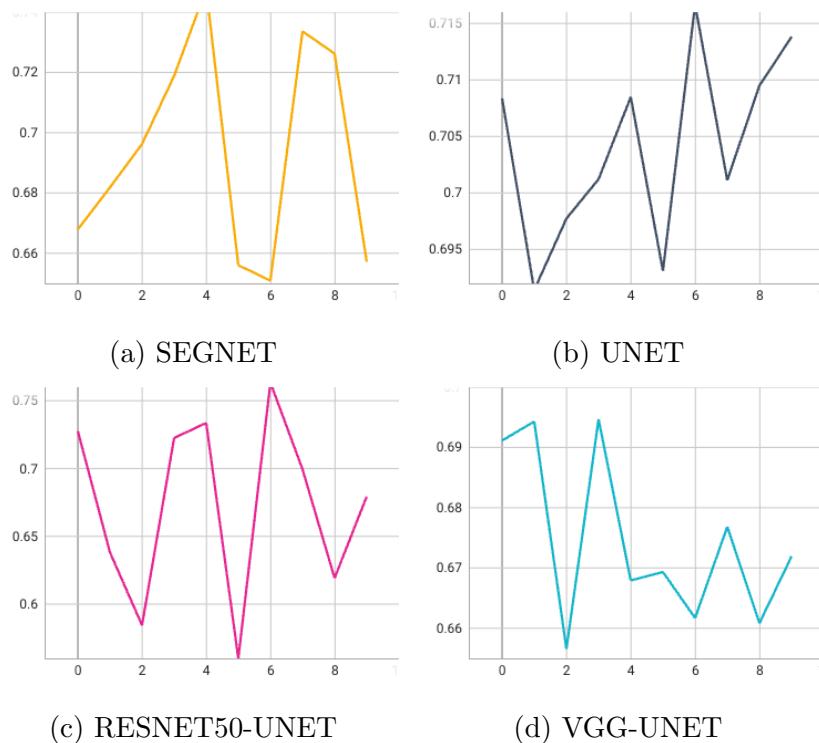
**Mean IoU vs Epoch:**

Figure 5.1: Mean IoU vs Epoch

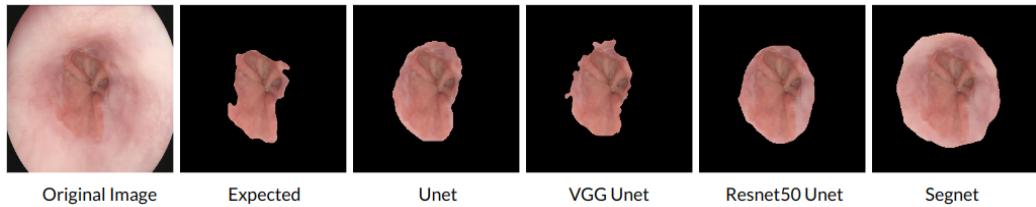
**Comparing the output of different models:**

Figure 5.2: Comparison of segmentation by different models(1)

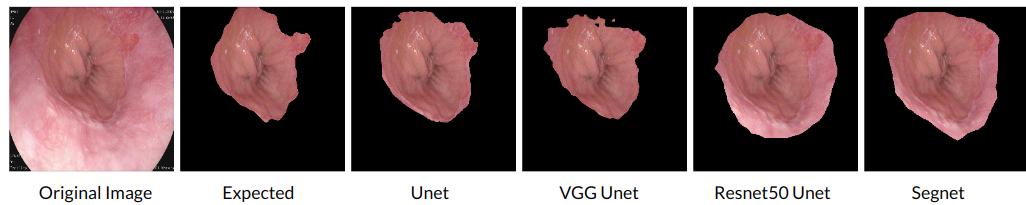


Figure 5.3: Comparison of segmentation by different models(2)

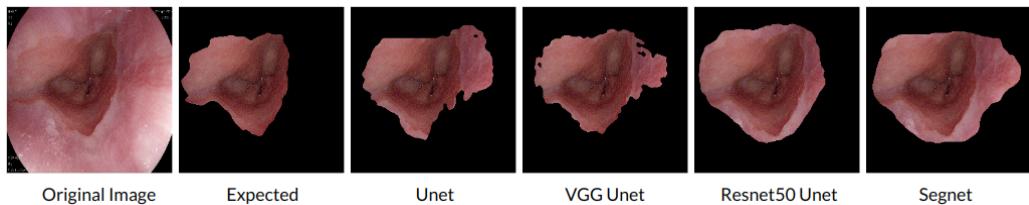


Figure 5.4: Comparison of segmentation by different models(3)

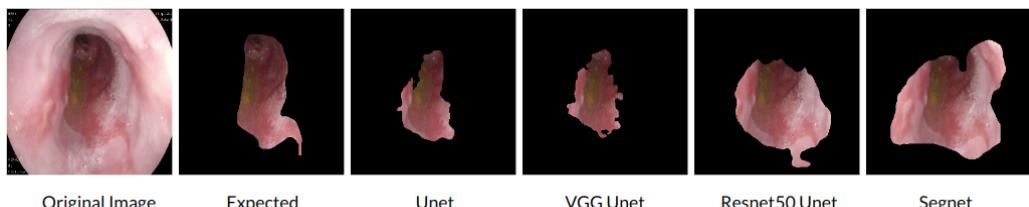


Figure 5.5: Comparison of segmentation by different models(4)

# **Chapter 6**

## **Conclusion**

On testing with different Unet based models and Segnet, we have found that Unet based models performs better than that of Segnet. Resnet50 and VGG architectures gave better prediction than the vanilla CNN with a training accuracy of 88.54 and 84.76 respectively. However training accuracy more or less remained constant after 8th epoch for all models.

Esophageal Adenocarcinoma is the main histological type of esophageal cancer in the west and BE is the only known histological precursor of EAC. Endoscopic biopsy is the most commonly used method for diagnosis and monitoring of BE[24]. This process relies on the experience of individuals with inevitable misjudgments, variations, and time consumption. Hence we require efficient AI models to segment and detect BE for effective treatment.

## References

- [1] Bhat S, Coleman HG, Yousef F, et al. Risk of malignant progression in Barrett's esophagus patients: results from a large population-based study. *J Natl Cancer Inst* 2011; 103: 1049– 1057.
- [2] Wu J, Pan YM, Wang TT, et al. Endotherapy versus surgery for early neoplasia in Barrett's esophagus: a meta-analysis. *Gastrointest Endosc* 2014; 79: 233–241.e2.
- [3] Visrodia K, Singh S, Krishnamoorthi R, et al. Magnitude of missed esophageal adenocarcinoma after Barrett's esophagus diagnosis: a systematic review and meta-analysis. *Gastroenterology* 2016; 150: 599–607.e7; quiz e14–e15.
- [4] Rodríguez de Santiago E, Hernanz N, Marcos Prieto HM, et al. Rate of missed oesophageal cancer at routine endoscopy and survival outcomes: a multicentric cohort study. *United European Gastroenterol J* 2019; 7: 189–198.
- [5] Jankowski M and Wani S. Diagnostic and management implications of basic science advances in Barrett's esophagus. *Curr Treat Options Gastroenterol* 2015; 13: 16–29.
- [6] Sharma P, Savides TJ, Canto MI, et al. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valu-

- able Endoscopic Innovations) on imaging in Barrett's esophagus. Gastrointest Endosc 2012; 76: 252–254.
- [7] Everson MA, Lovat LB, Graham DG, et al. Virtual chromoenendoscopy by using optical enhancement improves the detection of Barrett's esophagus-associated neoplasia. Gastrointest Endosc 2019; 89: 247–256.e4.
  - [8] Subramaniam S, Kandiah K, Schoon E, et al. Development and validation of the international Blue Light Imaging for Barrett's Neoplasia Classification. Gastrointest Endosc 2020; 91: 310–320.
  - [9] de Groof AJ, Struyvenberg MR, van der Putten J, et al. Deep-learning system detects neoplasia in patients with Barrett's esophagus with higher accuracy than endoscopists in a multistep training and validation study with benchmarking. Gastroenterology 2020; 158: 915–929.e4.
  - [10] Ebigo A, Mendel R, Probst A, et al. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. Gut 2019; 69: 615–616.
  - [11] Hashimoto R, Requa J, Dao T, et al. Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video). Gastrointest Endosc 2020; 91: 1264–1271.e1.
  - [12] de Groof AJ, Struyvenberg MR, Fockens KN, et al. Deep learning algorithm detection of Barrett's neoplasia with high accuracy during live endoscopic procedures: a pilot study (with video). Gastrointest Endosc 2020; 91: 1242–1250.
  - [13] Ebigo A, Mendel R, Rückert T, et al. Endoscopic prediction of submucosal invasion in Barrett's cancer with the use of artificial intelligence: a pilot study. Endoscopy 2021; 53: 878–883.

- [14] Trindade AJ, McKinley MJ, Fan C, et al. Endoscopic surveillance of Barrett's esophagus using volumetric laser endomicroscopy with artificial intelligence image enhancement. *Gastroenterology* 2019; 157: 303–305.
- [15] Swager AF, van der Sommen F, Klomp SR, et al. Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy. *Gastrointest Endosc* 2017; 86: 839–846.
- [16] Struyvenberg MR, van der Sommen F, Swager AF, et al. Improved Barrett's neoplasia detection using computer-assisted multiframe analysis of volumetric laser endomicroscopy. *Dis Esophagus* 2020; 33: doz065.
- [17] Bang CS, Lee JJ and Baik GH. Computer-aided diagnosis of esophageal cancer and neoplasms in endoscopic images: a systematic review and meta-analysis of diagnostic test accuracy. *Gastrointest Endosc* 2021; 93: 1006–1015.e13. e9.
- [18] Lui TKL, Tsui VWM and Leung WK. Accuracy of artificial intelligence-assisted detection of upper GI lesions: a systematic review and meta-analysis. *Gastrointest Endosc* 2020; 92: 821–830. e9.
- [19] Vennalaganti P, Kanakadandi V, Goldblum JR, et al. Discordance among pathologists in the United States and Europe in diagnosis of low-grade dysplasia for patients with Barrett's esophagus. *Gastroenterology* 2017; 152: 564–570. e4.
- [20] Tomita N, Abdollahi B, Wei J, et al. Attention-based deep neural networks for detection of cancerous and precancerous esophagus tissue on histopathological slides. *JAMA Netw Open* 2019; 2: e1914645.
- [21] Chen D, Wu L, Li Y, et al. Comparing blind spots of unsedated ultra-fine, sedated, and unsedated conventional gastroscopy with and without artificial intelligence: a prospective, single-blind, 3-parallel-group, randomized, single-center trial. *Gastrointest Endosc* 2020; 91: 332–339.e3.

- [22] Sinonquel P, Eelbode T, Bossuyt P, et al. Artificial Intelligence and its impact on quality improvement in upper and lower gastrointestinal endoscopy. *Dig Endosc* 2021; 33: 242–253.
- [23] Pan, W., Li, X., Wang, W. et al. Identification of Barrett's esophagus in endoscopic images using deep learning. *BMC Gastroenterol* 21, 479 (2021). <https://doi.org/10.1186/s12876-021-02055-2>
- [24] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv.org. Available at: <https://arxiv.org/abs/1505.04597>
- [25] Borgli, H., Thambawita, V., Smedsrud, P.H. et al. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Sci Data* 7, 283 (2020). <https://doi.org/10.1038/s41597-020-00622-y>
- [26] Goceri, E. Medical image data augmentation: techniques, comparisons and interpretations. *Artif Intell Rev* (2023). <https://doi.org/10.1007/s10462-023-10453-z>.