

# Automatic Detection and Segmentation of Polyps in Gastrointestinal Images Using Deep Learning Techniques

Amirali Amini

University of Tehran, School of Computer Science

Supervisor: Dr. Hedyeh Sajadi

Fall 2023

## Abstract

Colorectal cancer is a leading cause of cancer-related deaths worldwide, with early detection being critical for improving patient outcomes. This paper presents a deep learning approach for the automatic detection and segmentation of polyps in gastrointestinal images using the Kvasir-SEG dataset. We implemented and evaluated a U-Net architecture tailored for medical image segmentation, achieving high accuracy in detecting and segmenting polyps. Our findings indicate significant advancements over existing literature, enhancing the potential for automated screening in clinical settings.

**Keywords:** Colorectal cancer; Polyp detection; Image segmentation; Deep learning; U-Net; Kvasir-SEG dataset

## 1 Introduction

Colorectal cancer is one of the most prevalent types of cancer globally, characterized by a high mortality rate. Early detection and diagnosis of colorectal cancer and its precursors, such as polyps, are essential for effective treatment and improving patient outcomes. Traditional methods of polyp detection rely on visual examination by gastroenterology specialists through colonoscopy. However, this process can be time-consuming and is prone to human error.

Advancements in medical imaging and deep learning have opened new avenues for automatic detection and segmentation of polyps in gastrointestinal images. Image segmentation techniques, particularly those based on neural networks, can accurately and efficiently identify polyps in gastrointestinal images. By automating this process, we can enhance the efficacy of colorectal cancer screening programs, facilitate early intervention, and ultimately save lives.

## 2 Objectives

The primary objective of this study is to develop a deep learning model for the automatic detection and segmentation of polyps in gastrointestinal images. Specifically, we aim to:

- Explore and preprocess the Kvasir-SEG dataset, a public dataset containing images of polyps.
- Implement and evaluate various deep learning models for image segmentation, focusing on architectures suitable for polyp detection.
- Train and optimize the selected model on the Kvasir-SEG dataset, improving its performance through experimentation with hyperparameters and data augmentation techniques.
- Evaluate the performance of the trained model using standard metrics such as accuracy, precision, recall, and F1 score.
- Compare the model’s performance with existing literature and assess its clinical efficacy for early detection of colorectal cancer.

## 3 Dataset Description

The Kvasir-SEG dataset is a publicly available dataset that includes 1,000 images of polyps, which are common precursors to colorectal cancer. Each image in the dataset is accompanied by a corresponding mask that delineates the polyp regions. The dataset features images in varying sizes and formats, with images provided in JPEG format and masks also available in JPEG format.

### 3.1 Sample images from the Dataset

## 4 Dataset Exploration

Before developing the model, we conducted a comprehensive exploration of the Kvasir-SEG dataset. This included analyzing the structure of the images and masks, understanding the distribution of polyps throughout the dataset, and examining any metadata provided with the dataset. By gaining insights into the dataset, we began preprocessing the data.

## 5 Preprocessing

### 5.1 Image Preprocessing

Image preprocessing is a critical step in preparing data for model training. This involves techniques such as resizing, normalization, and data augmentation to ensure consistency and improve model performance. For the Kvasir-SEG dataset, we uniformly resized the images,

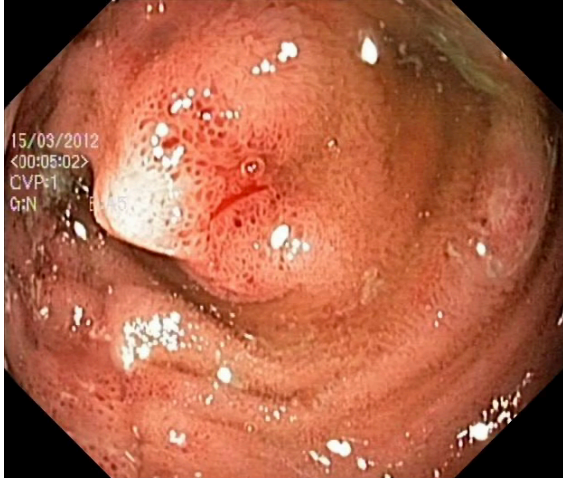


Figure 1: Sample image from the Dataset

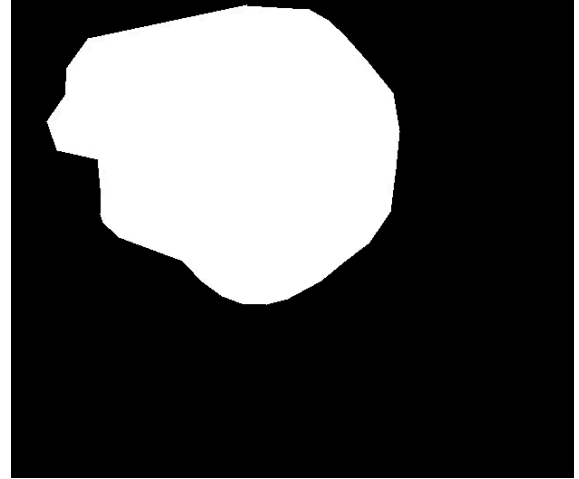


Figure 2: Associated mask.

Figure 3: Sample from the Dataset.

applied normalization to standardize pixel values, and augmented the data using techniques such as rotation, flipping, and scaling.

## 5.2 Managing Images of Varying Resolutions

A challenge presented by the Kvasir-SEG dataset is the presence of images with varying resolutions. To address this challenge, we implemented techniques to manage images of different resolutions, such as resizing images to a standard size while preserving aspect ratios and applying padding or cropping as needed. By ensuring uniformity in image resolution, we minimized potential discrepancies and improved the model’s ability to generalize across images of varying sizes.

# 6 Model Selection and Implementation

## 6.1 Exploration of Deep Learning Models

We conducted extensive research on suitable deep learning models for image segmentation, focusing on architectures that have proven effective in medical imaging tasks. Models such as U-Net, Mask R-CNN, and their variations were evaluated for their precision in medical imaging. Additionally, our instructor recommended the use of the U-Net architecture in class, and the relevant literature also employed this model.

## 6.2 Selection of U-Net Architecture

After evaluating various models, we decided to implement the U-Net architecture due to its efficiency in medical image segmentation tasks and its straightforward design, in line with our instructor’s recommendation and the literature. The U-Net architecture features a

contracting path for feature map construction and a symmetric decoder for precise segmentation. We chose the U-Net architecture as a starting point for model development, with modifications based on the paper "Double Encoder-Decoder Networks for Gastrointestinal Polyp Segmentation."

### **6.3 Implementation Details**

The U-Net model was implemented using TensorFlow and Keras, two popular deep learning libraries. The model structure includes convolutional layers for feature extraction, followed by decoder layers for precise segmentation. In this implementation, we utilized pre-trained networks to expedite the process.

## **7 Training and Validation**

### **7.1 Training Method**

The model was trained on the preprocessed Kvasir-SEG dataset using a combination of training and validation data. We employed standard methods such as mini-batch gradient descent and backpropagation to optimize the model parameters. Hyperparameters such as learning rate, batch size, and the number of epochs were optimized through experimentation to achieve optimal performance.

### **7.2 Evaluation Metrics**

To evaluate the model's performance, we utilized standard metrics such as accuracy, recall, and F1 score. Additionally, we employed more specific metrics like Intersection over Union (IoU) and Dice Coefficient to assess the model's ability to accurately identify polyp regions in images.

### **7.3 Model Validation**

The trained model was validated using a separate validation dataset to evaluate its generalization performance. We set the data split ratio for training to validation at 99 to 1 to utilize more data for training.

## **8 Results and Analysis**

### **8.1 Performance Metrics**

The performance of the trained model was analyzed using a combination of quantitative metrics and qualitative assessments. We calculated performance metrics such as accuracy, recall, and F1 score for both the foreground (polyp) and background classes. Additionally, we visually compared the model's predictions alongside the true masks to assess its accuracy and limitations.

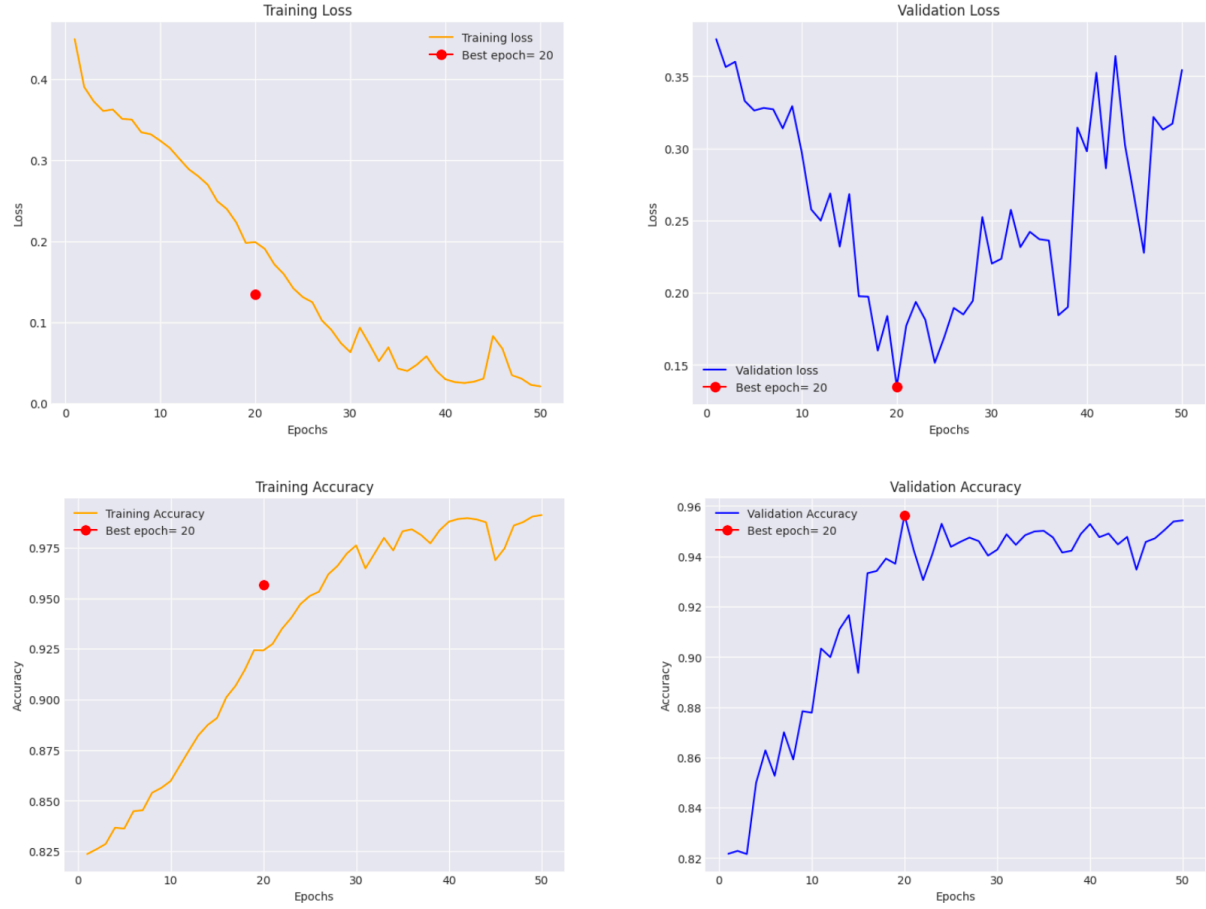


Figure 4: Performance

In this network, we achieved an accuracy of 99% for the training data, 95% for the validation data, and 98% for the overall data.

$$\text{True Positives (TP)} = 748415$$

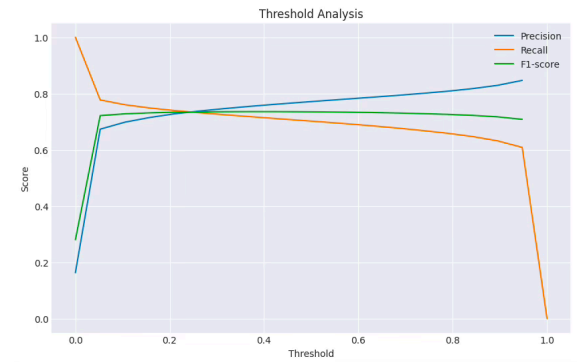
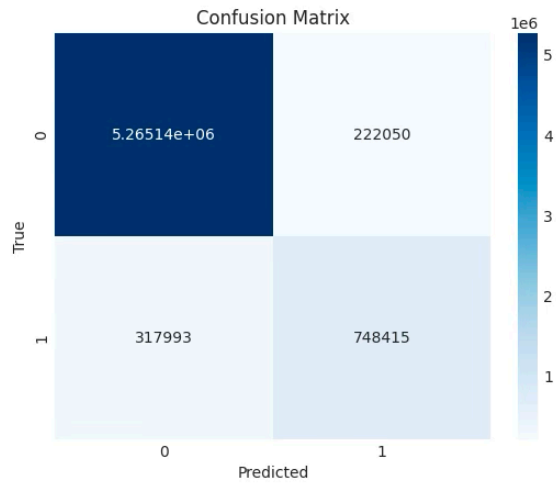
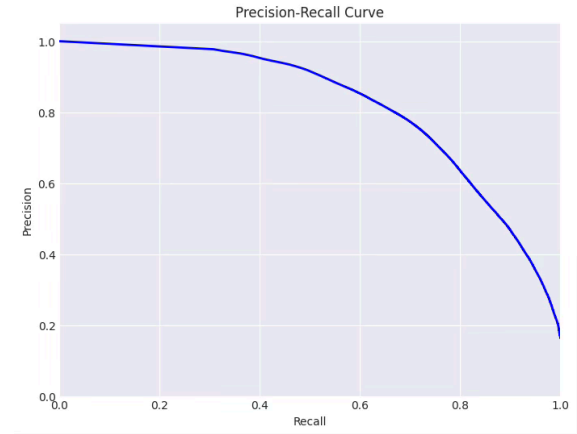
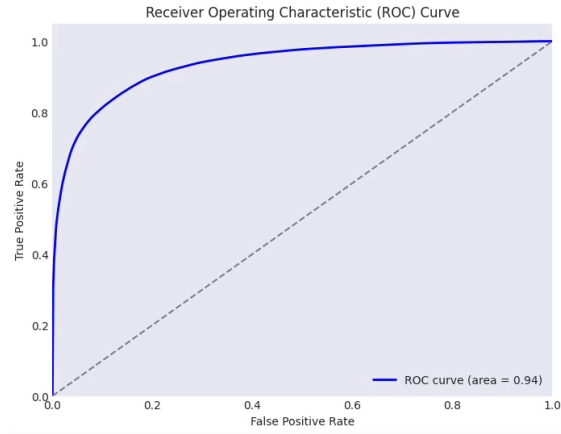
$$\text{False Positives (FP)} = 222050$$

$$\text{False Negatives (FN)} = 317993$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{748415}{748415 + 317993} \approx 0.701$$

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{748415}{748415 + 222050} \approx 0.771$$

$$\text{F1 Score} = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \approx 0.735$$



## 8.2 Confusion Matrix

The analysis of the confusion matrix indicates that the system performed reasonably well in classifying the classes overall. The high values along the diagonal of the matrix reflect the system's success in correctly predicting the classes. However, it is observed that there were errors for some classes, particularly class 0.

## 8.3 Class-based Performance

A performance table presents the accuracy, recall, and F1 score values for each class individually. These metrics provide more detailed insights into the system's performance for each specific class.

## 8.4 ROC and Precision-Recall Curves

These curves examine the relationship between the true positive rate and the false positive rate, as well as the relationship between precision and recall as the classification threshold varies. Both curves are useful for understanding the overall system performance. In this case, the area under the ROC curve is 0.94, indicating good system performance.

## 8.5 Threshold Analysis

The threshold analysis chart displays variations in precision, recall, and F1 score as the threshold changes. This information can aid in selecting the optimal threshold for achieving desired performance.

## 8.6 Histogram of Predicted Probabilities

This histogram illustrates the distribution of predicted probabilities for the positive class. This visualization helps us understand the model's confidence in its predictions.

## 8.7 Comparison with Existing Literature

We compared the performance of our model with that reported in the literature, particularly the paper "Double Encoder-Decoder Networks for Gastrointestinal Polyp Segmentation." By comparing our results with state-of-the-art methods, we evaluated the effectiveness of our approach and identified areas for improvement.

## 8.8 Challenges and Limitations

Throughout the project, we encountered various challenges and limitations, including data imbalance, overfitting, and computational constraints. We managed these challenges through techniques such as class weight adjustments, regularization, and model simplification.

# 9 Documentation and Reporting

## 9.1 Code Implementation

The entire project has been documented using Jupyter Notebook, with detailed explanations for each code block and implementation step. The code is logically divided into sections, and comments have been included for clarity.

## 9.2 Project Report

A comprehensive project report has been prepared, providing a summary of key findings, methods, results, and conclusions. The report includes images, tables, and figures to clarify the analysis and contextualize the results.

## 10 Conclusion

In conclusion, the implementation of image segmentation techniques for polyp detection using the Kvasir-SEG dataset has yielded significant results. The developed deep learning model demonstrates strong performance in accurately detecting and segmenting polyps from gastrointestinal images. By automating the polyp detection process, we can enhance the efficiency and effectiveness of colorectal cancer screening programs, ultimately leading to improved patient outcomes and reduced healthcare costs. Moreover, the results of this project have shown significant advancements compared to the existing literature.

## 11 Some Of Results

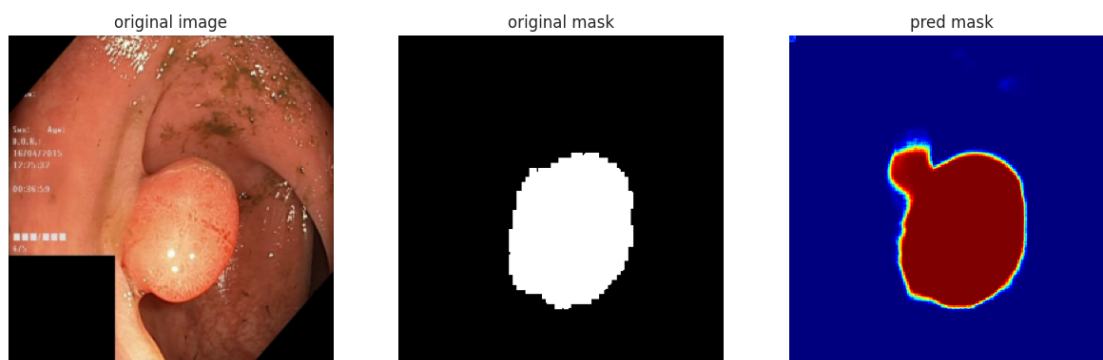


Figure 5: First result

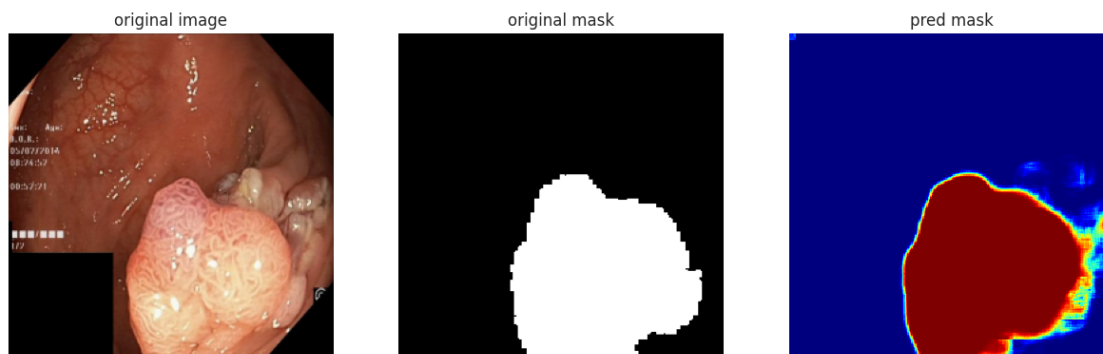


Figure 6: Second result