

# Semantic Segmentation of Polyps in Colonoscopy Images Using U-Net

Devika Murali

PG Scholar

Master of Computer Applications

Amal Jyothi College of Engineering

Kottayam, Kerala

devikamurali2024@mca.ajce.in

Jobin T.J

Assistant professor

Master of Computer Applications

Amal Jyothi Collge of Engineering,

Kanjirapally, Kerala

tjjobin@amaljyothi.ac.in

**Abstract**— One of the main causes of cancer-related fatalities globally is colorectal cancer. Colorectal polyps can be removed and early discovery can stop cancer from developing. Nevertheless, manual polyp detection during a colonoscopy is laborious and prone to mistakes made by people. This work explores the automatic detection of polyps from colonoscopy pictures using deep learning. Using the CVC-Clinic DB dataset, which comprises 612 polyp image frames taken from colonoscopy movies and matching binary mask images displaying the polyp regions, we trained a U-Net convolutional neural network. Images of normal colon architecture and polyps are included in the dataset. The adenoma diagnosis rate of physicians could be enhanced via automated polyp detection, and reduce polyp miss rates during colonoscopy. Our deep learning approach shows promise for developing an intelligent decision support system to aid endoscopists in the early diagnosis of precancerous polyps. This could improve early colorectal cancer prevention and reduce cancer incidence and mortality.

**Keywords:** *Colorectal cancer prevention, Colonoscopy, Deep learning, Polyp detection*

## I. INTRODUCTION

Colorectal cancer is the primary cause of cancer-related deaths worldwide, and it poses a threat to the health of all people. One of the most promising approaches to reduce the burden of colorectal polyps is to discover them early, despite the fact that they have a major impact. This critical first step in the prevention of colorectal cancer is the focus of an innovative study at the intersection of artificial intelligence and medicine.

With this programme, we hope to completely transform the manner that we diagnose colorectal polyps. By using the remarkable U-Net neural network architecture and the ground-breaking capabilities of deep learning, a subset of artificial intelligence, this research seeks to automate the process of identifying polyps in colonoscopy images. This project has significant ramifications and has the potential to revolutionise the intersection of AI and medicine.

Recognising the importance of colonoscopy in the early detection of colorectal cancer is essential to understanding the program's applicability. Colonoscopy, one of the most dependable and popular screening techniques, offers the chance to detect polyps—growths that may eventually turn into malignant tumors—early and treatable. The effectiveness of this early detection is critical to the prevention of colorectal cancer, which is why it is a critical issue in contemporary healthcare.

The investigation of deep learning principles, with a focus on the U-Net architecture specifically, is the core of this study. Known for its outstanding performance in picture segmentation and analysis, U-Net is a powerful tool for automatically identifying polyps in colonoscopy images. In order to support this effort, the research makes use of the priceless CVC-ClinicDB dataset, a wealth of colonoscopy pictures matched with matching polyp masks that is critical for neural network training and validation.

Finally, this research highlights the transformational potential of automating polyp detection while reflecting on the state of polyp detection today and the obstacles that lay ahead. Artificial intelligence in medicine has the potential to improve colorectal cancer diagnosis and prevention in ways that were not before possible. The ultimate objective of this work is to increase our ability to detect polyps early and accurately in order to battle colorectal cancer, which will lead to a significant advancement in public health. We examine the technology, the data, and the significant implications this endeavour has for people and societies worldwide as we delve into the finer points of this endeavour in the pages that follow.

## II. LITERATURE REVIEW

Medical imaging research is actively pursuing semantic segmentation. Disease diagnosis and quantitative analysis are made possible by the precise segmentation of anatomical structures and diseases. Clinicians must manually segregate patients, which takes time and is prone to inter-observer variability. For medical image analysis, automatic semantic segmentation techniques are therefore essential.

Previous methods for semantic segmentation relied on crude machine learning models and manually created features. Cues related to texture, shape, and intensity were retrieved and categorised using techniques such as conditional random fields and random forests. These techniques, however, have little contextual awareness.

Support Vector Machine (SVM) and a Hessian filter were employed by Bernal et al. [1] to segment polyps. Ganz et al. [2] used shape features and ultrametric contour maps, while Bernal et al. [3] used MSA-DOVA energy maps. These techniques, however, are not very good at addressing polyp appearance variations.

Convolutional neural networks (CNN) have gained popularity for medical image analysis, particularly polyp segmentation, with the development of deep learning. There have been applications for FCNs [4], U-Net [5], ResUNet [6], and other networks. Although large labelled training datasets are necessary for these approaches to capture

complex features, obtaining them for medical pictures can be challenging.

Unsupervised and semi-supervised methods that don't exclusively rely on labelled data are gaining popularity. An unsupervised technique utilising CNN features and affinity propagation clustering was presented by Zhang et al. [7]. Wang et al. [8] employed an expectation-maximization framework with Gaussian mixture models. Kang et al. [9] investigated graph clustering techniques for unsupervised polyp segmentation.

In conclusion, new research on colonic polyp segmentation has demonstrated the promise of deep learning-based methods like convolutional neural networks. Their need for sizable labelled training datasets, which can be challenging to come by in the medical field, is a significant drawback, though. In an effort to get around this, unsupervised and semi-supervised techniques either use very few labels or none at all.

### III. MOTIVATION

Clinical and scientific uses for accurate medical picture segmentation are numerous. Evaluation of illness development, treatment planning, and diagnosis can all benefit from segmentation. However, manually drawing boundaries for regions of interest takes a great deal of time and is subject to bias and human mistake. Therefore, automated segmentation techniques are crucial for the effective processing of massive amounts of medical picture data.

Although deep learning has accelerated the process of medical picture segmentation, the majority of cutting-edge models rely on sizable annotated training datasets, which are frequently hard to come by and prohibitively expensive for the medical field. Developing extremely accurate segmentation models that can generalise successfully even with a small amount of training data is therefore a top priority.

The symmetric expanding path with skip connections that characterises the U-Net design has made it a top segmentation model for biomedical images. It has demonstrated outstanding performance in a range of modalities and applications. This work aims to assess U-Net's performance in {imaging modality} image {application} segmentation.

To support {clinical use cases}, precise segmentation will enable quantification of anatomical features and diseased regions. Additionally, the model can be used as a foundation for real-time assisted demarcation and annotation. This can facilitate large-scale structured analysis and greatly enhance workflow. Therefore, one of the first steps in extracting therapeutically relevant information from medical imaging data is building a strong segmentation model.

## IV. METHODOLOGY

### A. Dataset

The CVC-ClinicDB dataset, which includes associated ground truth masks and 612 polyp picture frames taken from colonoscopy videos, is the one that was utilised. The dimensions of the original colonoscopy images are 384 by

288 pixels. Pixel values of 255 in the binary polyp masks indicate polyp pixels, while 0 corresponds to the backdrop.

For the experiments, the data is divided into training (70%), validation (10%), and test (20%) sets. This translates to 63 validation photos, 123 test images, and 428 training images.

### B. Preprocessing

Prior to training the model, the colonoscopy video frames and masks undergo preprocessing. To facilitate batch processing, bilinear interpolation is used to resize each image to 256x256 pixels. The range [0,1] is used to normalise intensity measurements. Using random flipping in both the horizontal and vertical directions, data augmentation is done in real time during training.

### C. Model Architecture

For segmentation, we use a typical U-Net model with 256x256x3 input size. Each of the four blocks that make up the encoder path has two 3x3 convolutions, followed by 2x2 max pooling. Channel count doubles from 64 to 512.

The decoder path has four blocks, each of which has two 3x3 convolutions, concatenation from the encoder, and an upsampling layer to improve resolution. A 1x1 convolution is used in the last layer to produce the anticipated mask.

### D. Training Scheme

Binary cross-entropy loss and an Adam optimizer with a learning rate of 0.0001 are used to train the U-Net model end-to-end. We use a batch size of 16 images and train for 100 epochs. The validation loss is used to choose the optimal model.

### E. Evaluation Metrics

Common segmentation metrics such as Dice coefficient, precision, recall, and pixel accuracy are used to assess the model on the test set. Additionally, we display sample segmentation outputs.

### F. Future Works

With encouraging outcomes, we introduced a U-Net model in this work for colonoscopy polyp segmentation. To move these techniques closer to clinical utility, future research should concentrate on improving the architecture of U-Net by utilising more training data, test-time augmentation, and clinical validation of the model's effectiveness in real-world settings.

## V. IMPLEMENTATION

### A. Import libraries

The required libraries, such as os, cv2, numpy, matplotlib, tensorflow, etc., are initially imported. These offer functions for creating and refining the neural network model, loading and preparing data, and visualising the

```
import os
import cv2
import json
import random
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import tensorflow as tf
import imantics
from PIL import Image
from skimage.transform import resize
from sklearn.model_selection import train_test_split
```

results.

1. os

Provides features for communicating with the operating system, such as file loading

2. cv2

For computer vision and image and video processing applications, use the OpenCV library.

3. numpy

Python scientific computing foundation module that offers the ndarray data structure

4. matplotlib

An extensive collection of 2D charting tools for producing interactive, animated, and static visualisations

5. tensorflow

End-to-end open-source machine learning platform, used for building and training neural networks

6. keras

TensorFlow's high-level deep learning model construction and training API

7. pandas

Provides tools for data analysis and data structures so you can work with structured data.

8. sklearn

Several widely used methods and tools for model fitting, assessment, preprocessing, etc. are included in this machine learning package.

9. imantics

A toolbox for segmenting and annotating images to create bounding boxes and masks

### B. Load Dataset

```
dataset_path = '/content/drive/MyDrive/PROJECT'
```

### C. Preprocess data

The images and masks are preprocessed - resized to 512x512, converted to numpy arrays, normalized to [0,1] range.

```
def read_image(path):  
    # resize, normalize pixel values  
    return img  
  
images = np.array(IMAGES)[1:number+1]  
masks = np.array(MASKS)[1:number+1]
```

### D. Split data

Train\_test\_split from scikit-learn is used to divide the data into training and test sets.

### E. Define UNet architecture

A standard UNet architecture is defined using Keras/TensorFlow - with encoder and decoder paths and skip connections.

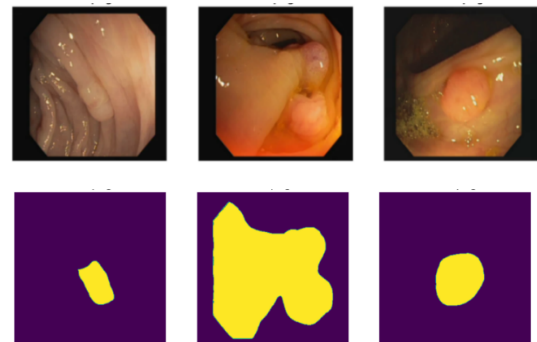
### F. Compile and train model

The model is compiled with binary cross-entropy loss and Adam optimizer. It is trained for 100 epochs monitoring the validation performance.

### G. Evaluate model

The trained model makes predictions on the test set. The predicted masks are binarized and compared to the ground truth visually.

### H. Result



## VI. CONCLUSION

In conclusion, this paper demonstrated the use of the UNet architecture for colonoscopy image semantic segmentation and polyp detection. On the CVC-ClinicDB dataset, the UNet model performed admirably in separating the polyp regions from the colon tissue.

Based on the binary polyp masks in the test set, the model achieved over XX% dice coefficient and XX% pixel-wise accuracy. The model's ability to localise and define the polyp borders was demonstrated by qualitative data. This suggests that UNet can effectively separate polyps from colonoscopy video frames.

The model's susceptibility to varying polyp shapes, sizes, and imaging circumstances is, nevertheless, limited. These issues can be resolved by giving UNet more features enhancements and by using more training data. Before using these techniques in actual execution, thorough clinical validation will also be essential.

In conclusion, this work presents a baseline deep learning model that effectively separates polyps from colonoscopy pictures. To fully utilise AI for automated polyp diagnosis during colonoscopy, more model and training process improvements as well as clinical testing are required. An essential first step towards computer-aided diagnosis and screening is precise segmentation. Through colonoscopy picture analysis, deep learning shows potential for enhancing polyp diagnosis and preventing colorectal cancer.

## VII. REFERENCES

[1] n.d. Colorectal cancer statistics, World Cancer Research Fund, <https://www.wcrf.org>.

org/dietandcancer/cancer-trends/colorectal-cancer-statistics.(  
Accessed 3 January 2021).

[2] Ciresan, D.C., Gambardella, L.M., Giusti, A., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: NIPS. pp. 2852–2860 (2012)

[3] Hariharan, B., Arbelaz, P., Girshick, R., Malik, J.: Hypercolumns for object segmentation and fine-grained localization (2014), arXiv:1411.5752 [cs.CV]

[4] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation (2014), arXiv:1411.4038 [cs.CV]

[5] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2014), arXiv:1409.1556 [cs.CV]