Mr. S. Sinimoxon Lee, 2024, 12:5
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

# Polyp Detection Using U-Net Deep Learning Method

**Mr. S. Sinimoxon Lee, Professor Arpita Das**
School of Computer Science Engineering
Vaishnavi Institutes of Technology and Science Bhopal, India

**Abstract- Polyp detection in colonoscopy images plays a crucial role in the early diagnosis and prevention of colorectal cancer. This project implements a U-Net model to segment polyp regions from colonoscopy frames using the CVC-ClinicDB dataset, which includes annotated ground truth masks. The U-Net architecture, with its encoder-decoder structure and skip connections, is particularly well-suited for medical image segmentation, enabling precise identification of polyp regions. The model's performance is evaluated using Intersection over Union (IoU) to measure segmentation accuracy, alongside precision-recall curves to assess detection reliability. Training progress is monitored through the visualization of training and validation loss curves, as well as accuracy curves, ensuring the model's effectiveness and generalization. The results demonstrate that the U-Net model can significantly improve the accuracy of polyp detection, contributing to more reliable colorectal cancer screening.**

**Keywords- Polyp Detection, U-Net Model, Medical Image Segmentation, Colonoscopy**

## I. INTRODUCTION

Colorectal cancer stands as a formidable global health challenge, ranking as the third most prevalent cancer and the second leading cause of cancer-related mortality worldwide [1]. The significance of early detection cannot be overstated, as it has been conclusively demonstrated that the timely identification and removal of precancerous polyps during colonoscopy screenings can reduce the risk of colorectal cancer development by up to 80% [2].

However, despite the effectiveness of colonoscopy as a screening tool, manual detection and segmentation of polyps present significant challenges. These challenges include the variability in polyp appearance, size, and location within the colon, as well as the subjective nature of detection, which is influenced by factors such as endoscopists' experience and visual fatigue. The emergence of automated polyp detection and segmentation systems holds promise in overcoming these challenges. Leveraging advanced computer vision and deep learning techniques, these systems offer real-time feedback to endoscopists, enhancing the accuracy and efficiency of polyp detection.

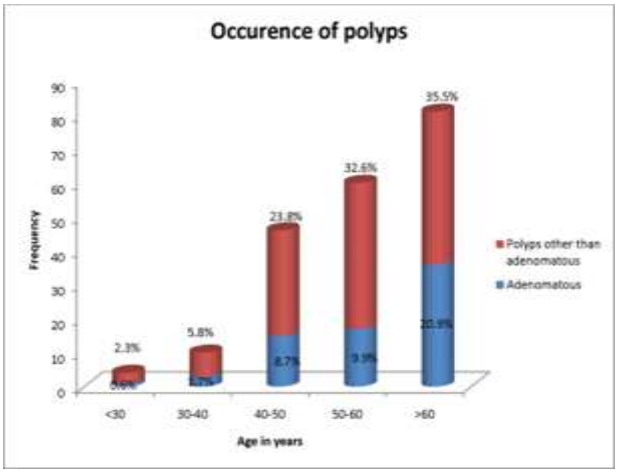


Fig. 1: Bar chart showing the occurrence of polyps categorized by age in years and types.

Among the various methodologies, the U-Net architecture has shown particular promise in medical image segmentation tasks. This project

aims to address the limitations of manual polyp detection by designing and implementing a deep learning model based on the U-Net architecture, tailored specifically for polyp segmentation in colonoscopy images. Through rigorous preprocessing, model training, and evaluation, this project endeavors to provide insights into the potential of automated systems in improving colorectal cancer screening outcomes. Furthermore, the project aims to identify avenues for future research, including the exploration of additional data augmentation techniques and advanced model architectures, to further enhance automated polyp detection and segmentation. The main objectives of this project are:

- Pre-process and prepare the CVC-ClinicDB dataset, a collection of frames extracted from colonoscopy videos, along with ground truth masks for polyp regions.
- Design and implement a deep learning model based on the U-Net architecture, tailored specifically for polyp segmentation in colonoscopy images.
- Train and evaluate the model using appropriate loss functions and evaluation metrics, such as binary cross-entropy, accuracy, recall, precision, and Intersection over Union (IoU).
- Analyse and visualize the model's performance, including qualitative results showcasing input images, ground truth masks, and predicted masks, as well as quantitative evaluations of the model's segmentation accuracy.
- Discuss the limitations and potential improvements for future work in automated polyp detection and segmentation, including incorporating additional data augmentation techniques, exploring advanced model architectures, and integrating post-processing methods to refine the segmentation results further.

Our paper was organised in the following way: In section 2 we have given the literature survey. In section 3 we have discussed our proposed methodology. In section 4 the results are discussed. Finally, we concluded in section 5.

## II. LITERATURE SURVEY

Several studies have explored the application of deep learning models, particularly convolutional neural networks (CNNs), for polyp detection and segmentation tasks. This section highlights some notable case studies in this field.

In [3] the U-Net architecture was employed for the segmentation of polyps in colonoscopy images. The authors proposed a modified U-Net model with additional convolutional layers and dropout regularization to improve performance. The model was trained and evaluated on a dataset of 379 colonoscopy images, with manually annotated ground truth masks for polyp regions. The proposed approach achieved promising results, with an average Dice similarity coefficient of 0.76 and a mean intersection over union (IoU) of 0.68,

In [4] a convolutional multilayer perceptron (MLP) network is introduced for polyp segmentation to improve accuracy in colonoscopy images. Our network utilizes a convolutional MLP encoder and enhances low-level features through a parallel self-attention module. Additionally, rather than directly incorporating encoder features into the decoder, we employ a cascaded context aggregation module to combine high-level semantic and low-level local features. Finally, a channel-guided group reverse attention mechanism is applied to refine structural and textural details by exploring the relationship between regions and boundary cues.

In [5], the authors proposed a computer-aided system for colorectal polyp classification, aiming to assist in timely resection without the need for histological analysis. They leveraged deep learning techniques, addressing the gap in existing methods by incorporating hyperparameter optimization, which is critical for improving model performance. Unlike traditional binary classification approaches, the authors focused on multiclass classification, specifically tackling the challenge of classifying serrated adenomas, which are difficult due to their hybrid nature. To enhance accuracy, they introduced a weighted ensemble model that combines weak learners, optimizing both high-level

and low-level features for better classification. The proposed method was evaluated on the UCI and PICCOLO datasets, achieving superior performance with accuracy rates of 96.3% and 81.2%, along with strong precision, recall, and F1-scores. Additionally, Cohen's Kappa Coefficient indicated high model reliability. The authors demonstrated that integrating a weighted ensemble of optimized networks, alongside data augmentation, significantly improves the classification accuracy over existing models, as seen in comparisons with SVM and other deep learning methods.

In [6], the authors addressed the challenge of colon cancer, which is the third most commonly diagnosed cancer after breast and lung cancer. Most cases of colon cancer arise from adenocarcinomas, which develop from adenomatous polyps on the colon's inner lining. Colonoscopy is the standard method for polyp detection, but its success largely depends on the experience of the colonoscopist and environmental factors. Despite this, between 8-37% of polyps are missed during colonoscopy due to human error, potentially leading to colorectal cancer.

To mitigate this issue, various deep learning-based computer-aided detection (CAD) systems have been proposed in recent years for automated polyp detection, localization, and segmentation. However, the need for more accurate polyp detection and segmentation remains to reduce missed polyps. In response, the authors proposed a Super-Resolution Generative Adversarial Network (SRGAN)-assisted Encoder-Decoder model for fully automated colon polyp segmentation from colonoscopic images.

The SRGAN was integrated into the up-sampling process to improve segmentation accuracy. The model was evaluated on benchmark datasets, CVC-ColonDB and Warwick-QU, achieving a dice score of 0.948 on CVC-ColonDB, outperforming recent state-of-the-art methods. On the Warwick-QU dataset, the model achieved dice scores of 0.936 on Part A and 0.895 on Part B, demonstrating superior accuracy, particularly for sessile and smaller-sized polyps.

## III. PROPOSED METHODOLOGY

Before delving into the intricacies of employing the U-Net architecture for automated polyp detection and segmentation in colonoscopy images, it's essential to grasp the fundamental context of this endeavor. Colorectal cancer stands as a significant global health concern, ranking as the third most common cancer and the second leading cause of cancer-related deaths worldwide. Early detection plays a pivotal role in mitigating its impact, with studies demonstrating that timely identification and removal of precancerous polyps during colonoscopy screenings can reduce the risk of colorectal cancer development by up to 80% [2]. However, manual detection of polyps during these screenings poses challenges due to the variability in polyp appearance, size, and location within the colon, as well as the subjective nature of detection influenced by endoscopists' experience and visual fatigue. This chapter elucidates the methodology employed to address these challenges through the utilization of the U-Net architecture, encompassing model design, implementation, training procedures, and evaluation metrics.

### 1. CVC-ClinicDB Dataset

The CVC-ClinicDB dataset [7] is a publicly available database of frames extracted from colonoscopy videos, specifically designed for the task of automatic polyp detection and segmentation. This dataset was created by the Computer Vision Center (CVC) and Hospital Clínic in Barcelona, Spain, and has been widely used in research and challenges related to polyp detection and segmentation.

Original Images: These are the raw frames extracted from colonoscopy videos, stored in TIFF format. The dataset contains a total of 612 original images, each representing a different frame from various colonoscopy procedures.

Ground Truth Masks: For each original image, a corresponding ground truth mask is provided, also in TIFF format. These masks indicate the precise regions within the image that contain polyps, serving as the target output for the segmentation task.

**2. U-Net Architecture**

The U-Net architecture, introduced in [3], has become a popular choice for various medical image segmentation tasks due to its exceptional performance and ability to capture contextual information effectively. The U-Net model is well-suited for polyp segmentation in colonoscopy images for several reasons:

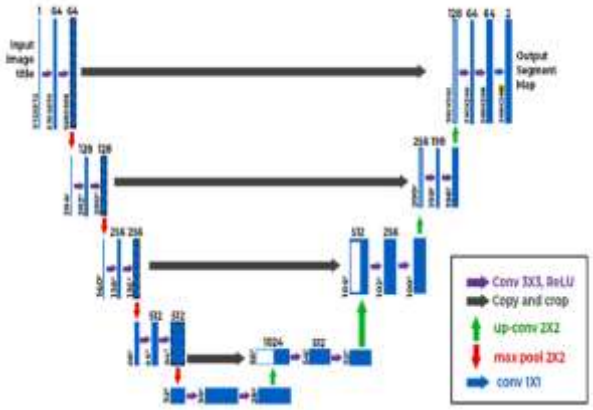


Fig.2. U-Net Architecture

**Encoder Path**

The encoder path of the U-Net architecture is responsible for capturing and downsampling the input image's features. It consists of a series of convolutional blocks, each composed of two consecutive convolutional layers, followed by batch normalization and ReLU activation. After each convolutional block, a max-pooling operation is applied to downsample the feature maps, effectively reducing their spatial dimensions while increasing the receptive field and capturing higher-level semantic information. The number of filters in the convolutional layers gradually increases as the depth of the encoder path increases, allowing the model to capture more complex and higher-level features.

**Decoder Path**

The decoder path of the U-Net architecture is designed to be symmetric to the encoder path, upsampling is performed using transpose convolutional layers (also known as deconvolution), which are followed by concatenation with the corresponding feature maps from the encoder path via skip connections.

**Skip Connections**

One of the key features of the U-Net architecture is the use of skip connections, which allow the fusion of low-level and high-level features from the encoder and decoder paths, respectively. These skip connections concatenate the feature maps from the encoder path with the corresponding up sampled feature maps in the decoder path, enabling the model to combine both coarse, high- level semantic information and fine-grained, low-level details. The skip connections play a crucial role in the precise segmentation of polyps with varying sizes and appearances. By incorporating low-level features from the encoder path, the model can capture intricate details and fine boundaries, while the high-level features from the decoder path provide the necessary contextual information for accurate localization and delineation of polyp regions.

## IV. RESULT & DISCUSSION

As we delve into the detailed analysis of the automated polyp detection and segmentation system based on the U-Net architecture, it becomes imperative to scrutinize both its qualitative and quantitative performance metrics. This chapter meticulously evaluates the model's effectiveness in delineating polyp regions through sample predictions and comprehensive visualization techniques. Moreover, it quantifies the model's performance using key evaluation metrics such as accuracy, recall, precision, and Intersection over Union (IoU) on the test set, providing insights into its strengths and areas for improvement. Additionally, a thorough performance analysis elucidates the system's capabilities and limitations, paving the way for potential enhancements through advanced data augmentation, transfer learning, ensemble methods, post-processing techniques, and multi- task learning. By dissecting the model's performance and proposing avenues for refinement, this chapter aims to foster a deeper understanding of automated polyp detection and segmentation in colonoscopy images, ultimately striving for improved diagnostic outcomes in colorectal cancer screening.

To quantitatively evaluate the model's performance, the evaluation metrics defined in the methodology section were calculated on the test set. The following table summarizes the obtained results.

Table.1 Quantitative Evaluation

| Model | Accuracy | Precision | Recall | Loss | F1 Score | IOU |
|---|---|---|---|---|---|---|
| UNET | 0.9835 | 0.9595 | 0.8870 | 0.1127 | 0.9219 | 0.4446 |

The accuracy metric measures the overall performance of the model in correctly classifying pixels as either belonging to a polyp region or not. A high accuracy value indicates that the model is capable of accurately distinguishing between polyp and non-polyp regions across the entire dataset.

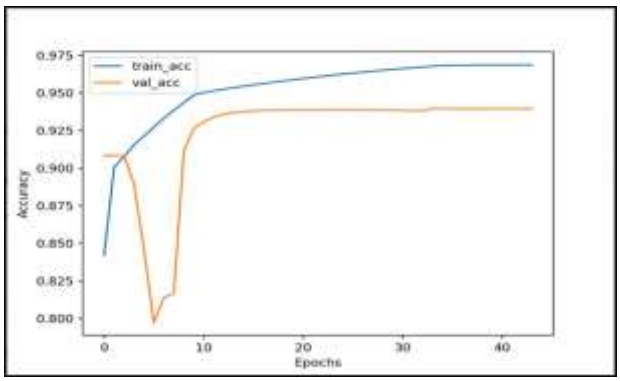


Fig.3. Accuracy vs Epochs Graph

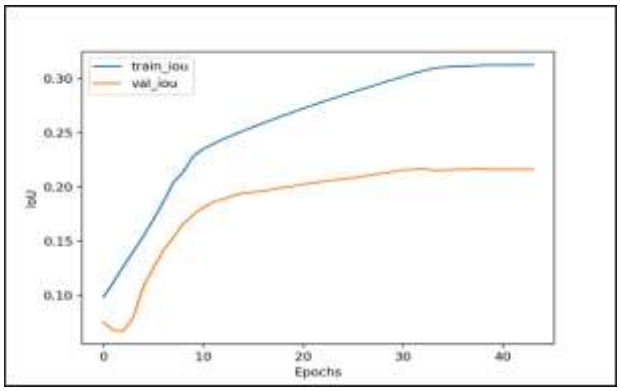


Fig.4. IoU vs Epochs Graph

The Intersection over Union (IoU) metric is widely used in segmentation tasks to measure the overlap between the predicted segmentation masks and the ground truth masks. It quantifies the model's ability to accurately delineate the boundaries and extent of polyp regions. A higher IoU value suggests that the predicted segmentation masks closely match the ground truth masks, indicating precise localization and delineation of polyp regions.

## V. CONCLUSION

This work focused on developing an automated polyp segmentation system for colonoscopy images using the U-Net architecture, a deep learning model that has shown remarkable success in various medical image segmentation tasks. The motivation behind this work is to contribute to the advancement of computer-assisted diagnosis systems for colorectal cancer screening, which can potentially improve patient outcomes and reduce the burden on healthcare professionals. The methodology involved preprocessing the CVC-ClinicDB dataset, designing and implementing the U-Net model with appropriate convolutional blocks and skip connections, and training the model using binary cross-entropy loss and evaluation metrics such as accuracy, recall, precision, and Intersection over Union (IoU). Extensive experiments were conducted to evaluate the model's performance on the test set, and the results were presented through qualitative visualizations and quantitative evaluations. The qualitative results demonstrated the model's ability to accurately segment polyps of varying sizes, shapes, and appearances, with the predicted segmentation masks closely matching the ground truth masks. The quantitative evaluation further validated the model's performance, achieving promising values for accuracy, recall, precision, and IoU metrics.

## REFERENCES

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, 68(6), 394-424.
2. Mandel, J. S., Bond, J. H., Church, T. R., Snover, D. C., Bradley, G. M., Schuman, L. M., & Ederer, F. (1993). Reducing mortality from colorectal cancer by screening for fecal occult blood. New

England Journal of Medicine, 328(19), 1365-1371.
3. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241).
4. Jin, Y., Hu, Y., Jiang, Z., & Zheng, Q. (2023). Polyp segmentation with convolutional MLP. The Visual Computer, 39(10), 4819-4837.
5. Younas, F., Usman, M., & Yan, W. Q. (2023). A deep ensemble learning method for colorectal polyp classification with optimized network parameters. Applied Intelligence, 53(2), 2410-2433.
6. Mahanty, M., Bhattacharyya, D., & Midhunchakkaravarthy, D. (2021). SRGAN Assisted Encoder-Decoder Deep Neural Network for Colorectal Polyp Semantic Segmentation. Rev. d'Intelligence Artif., 35(5), 395-401.
7. Bernal, J., Sánchez, F. J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., & Vilariño, F. (2015). WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. Computerized Medical Imaging and Graphics, 43, 99-111.