Detection Of Elusive Polyp Using U-Net for Polyp Segmentation

N. Sravan Kumar*
Department of Computational Intelligence
SRM Institute of Science and Technology, Faculty of Engineering and Technology, KTR campus,
Chennai, India
*nn0800@srmist.edu.in

C. Gokul Krishna Reddy *
Department of Computational Intelligence
SRM Institute of Science and Technology, Faculty of Engineering and Technology, KTR campus,
Chennai, India
*cc1733@srmist.edu.in

A. Bhuvan Sai *
Department of Computational Intelligence
SRM Institute of Science and Technology, Faculty of Engineering and Technology, KTR campus,
Chennai, India
*as9964@srmist.edu.in

B. Phanindra Babu *
Department of Computational Intelligence
SRM Institute of Science and Technology, Faculty of Engineering and Technology, KTR campus,
Chennai, India
*bb7944@srmist.edu.in

Abstract – **Colon cancer is a prevalent and potentially deadly form of cancer that originates as polyps on the surface of the colon. Early detection of these polyps and malignancies is crucial for effective screening. Deep Convolutional Neural Network (DCNN) techniques, specifically using U-Net architecture, are employed to improve colorectal cancer screening. Data augmentation enhances the dataset, aiding in accurate detection. Challenges such as inconsistent polyp presence, non-uniform colour patterns, and reflection effects in images from colonoscopies and endoscopies can lead to over or under-segmentation. To address this, a segmentation method considers these factors, recognizing areas through edge and valley detection. Post-processing techniques refine segmentation boundaries and group neighbouring objects to improve detection accuracy. This automated approach facilitates easier, faster, and more accurate polyp identification by clinicians. Such methods are integral to the development of Computer-Aided Diagnosis (CAD) systems for automated polyp detection. Our proposed U-Net strategy offers an effective solution for DCNN design in this context, enhancing automated polyp detection methods.**

***Keywords* – Colorectal cancer, polyp, DCNN, U-Net, post-processing, segmentation, augmentation.**

I. INTRODUCTION

Colorectal cancer presents a formidable global health challenge, accounting for a substantial portion of cancer-related fatalities annually. As the second leading cause of cancer mortality, its impact reverberates globally. Nonetheless, there exists a glimmer of hope: timely detection and intervention. By identifying and managing colorectal cancer in its initial stages, patients can experience significantly improved outcomes and a reduced mortality rate.

At the forefront of this endeavor lies colonoscopy, a cornerstone procedure for both screening and diagnosis. By meticulously examining the colon's interior using a flexible tube equipped with a camera, clinicians can detect abnormalities, including polyps—the precursors to colorectal cancer. These polyps, often small and inconspicuous, mark a critical juncture in the disease's progression. Detecting and precisely localizing these lesions during colonoscopy are pivotal in early detection and curbing the advancement of colorectal cancer.

Polyp Segmentation:

Recent years have witnessed remarkable strides in medical imaging, particularly in automated polyp segmentation. Advanced learning models, exemplified by U-Net, have emerged as frontrunners in this domain, offering unparalleled precision and efficiency in segmenting polyps within colonoscopy frames.

The adoption of advanced learning methodologies holds immense promise in reshaping colorectal disease diagnosis and management. By leveraging extensive datasets and sophisticated algorithms, these models can discern intricate details in polyp morphology and appearance, facilitating accurate segmentation even in challenging circumstances. Comparative assessments have underscored U-Net's efficacy in high-precision semantic segmentation tasks, positioning it as a pivotal tool in enhancing colorectal cancer detection and management.

As medical imaging advances, the incorporation of deep learning-based techniques signifies a pivotal shift in precision medicine. By enhancing diagnostic accuracy and empowering clinicians, these innovations are poised to redefine the management of colorectal diseases, offering renewed optimism globally. Colorectal cancer, a significant global health concern and the second leading cause of cancer mortality, necessitates early detection and intervention. Procedures like colonoscopy enable meticulous examination of the colon for polyps, crucial precursors to colorectal cancer. Automated polyp segmentation, exemplified by models like U-Net, showcases the transformative potential of deep learning in accurately identifying these lesions.

II. LITERATURE SURVEY

Mahmud et al. [1] proposed PolypSegNet uses multiple sequential depth dilation initiation (DDI) blocks, a deep fusion skip module (DFSM), and a deep reconstruction module (DRM) to extract polyps from colonoscopy images. Generate automatically. SegNet architecture modified for segmentation. Fan et al. [2] In order to improve an FCN-like model for medical picture segmentation, we suggest PraNet, which uses parallel partial decoding and reverse attention modules. UNet is suggested as an alternative to the classic FCN (Fully Convolutional Network) architecture's single encoder, which significantly boosts FCN's speed and has become a popular option for medical image segmentation. The features from the encoding and decoding layers are concatenated using skip connections in the encoder-decoder-based structure known as UNet.

A number of variations that were suggested for polyp segmentation and produced encouraging results were motivated by UNet's performance. Jha et al. [3] provided Double UNet, a UNet combination. A pre-trained VGG-19 serves as the foundation of the first UNet. To effectively collect more semantic data, the second UNet is added below the first UNet. Additionally, they use Atrous Spatial Pyramid Pooling (ASPP) to gather contextual data from the network. Kang and Gwak [4] employed Mask-RCNN as the main framework for automatic polyp detection and segmentation, that is based on ResNet50 and ResNet101.

Using a fully convolutional neural network (FCN), pixel-level segmentation was achieved. Zhou et al. [5] introduced UNet++, a highly supervised encoderdecoder network that joins UNet together using several layered, dense skip routes. Zhang et al. [6] employed the FCN-8S to divide potential polyp area candidates into discrete regions, and a random forest classifier used the texton characteristics computed from each region to make the final determination.

Cai et al. (2020) applied deep learning to gastrointestinal imaging and interpreted major findings in the National Polyp Study (NPS) dataset [7]. Li et al. (2021) conducted a systematic review and meta-analysis of deep learning in video-based colonoscopy for polyp detection, highlighting its potential in improving diagnostic accuracy [8]. Additionally, Mahmood et al. (2019) utilized deep convolutional neural networks for polyp detection in colonoscopy images, underscoring the relevance of deep learning in this domain [9].

III. PROPOSED ARCHITECTURE

Our proposed model's primary objective is to enhance the accuracy of polyp detection. To achieve this goal and improve the effectiveness of colorectal cancer screening, we employ deep convolutional neural network (DCNN) techniques based on the U-Net architecture. Data augmentation is utilized to enrich our dataset, providing U-Net with access to additional details for more precise detection. U-Net, specifically designed for biological image segmentation, enables more accurate segmentation with fewer training images due to its extended and upgraded network architecture. By integrating complex DCNN construction methods with U-Net methodology, we present an automated polyp detection method in this thesis, exploring advanced techniques and algorithms essential for creating a highly effective U-Net model.

A) Existing Systems:

Numerous methodologies have been proposed to achieve accurate polyp segmentation, broadly categorized into deep learning and conventional machine learning approaches. These methods typically analyze the polyp's edge, color, and texture for segmentation purposes, showcasing a growing sophistication in polyp segmentation research. However, a notable challenge persists in the field—the lack of research utilizing cross-dataset testing to assess model generalizability. Many recent studies rely on small, unbalanced, and hand-selected datasets, hindering the development of reliable and widely applicable polyp segmentation methods. Moreover, the medical profession faces obstacles due to limited training datasets and dataset imbalances, complicating the creation of robust and scalable segmentation techniques. Addressing these challenges, our study aims to develop an algorithm capable of achieving high performance in polyp segmentation.

B) Fundamental Concepts:

Deep Learning: Deep learning, a subset of machine learning, forms the basis of artificial intelligence (AI) systems. It revolves around artificial neural networks (ANN), intricate structures designed to analyze vast datasets by passing them through multiple layers of interconnected neurons.

UNET: U-Net, a convolutional neural network (CNN), specializes in segmenting biological images. Its architecture has been refined to allow for precise segmentation with minimal training images. Modern GPUs can segment a 512x512 image in less than a second.

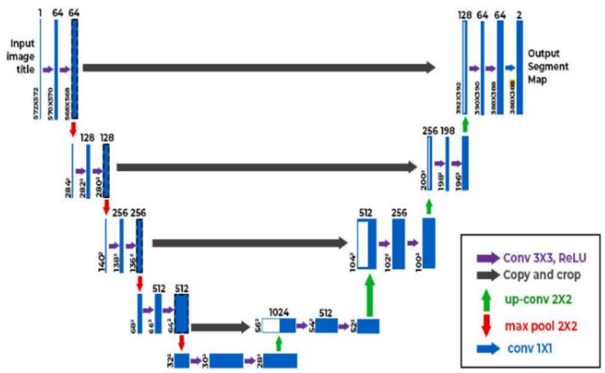

Fig. 1. U-Net Architecture

The architecture involves operations conducted across multiple convolutional layers. To activate neurons within the convolutional network, the architecture utilizes the ReLU activation function. Additionally, the Adam optimizer is employed to minimize the loss/cost function. The difference between predicted and actual output is determined by the loss function, allowing the model to reduce loss or error through various convolutional techniques. Furthermore, pooling techniques are implemented to reduce network computations and parameters that require learning.

DCNN: Deep convolutional neural networks (DCNNs) derive their strength from layered processing. Processing red, green, and blue components simultaneously using a three-dimensional neural network significantly reduces the number of artificial neurons required. Images are processed through convolutions instead of matrix multiplication.

Data Augmentation: Data augmentation prevents overfitting by enriching training datasets when they are insufficient. Techniques such as rotation, mirroring, shifting, and zooming create diverse images from a limited set, enhancing model accuracy.

Convolution Operation: Convolution, a mathematical process, combines two functions to produce a third describing how one function's shape changes due to the other. This operation resembles the analysis of metrics to optimize results, resulting in weighted averages of similar metrics.

Pooling: Pooling operations generate images based on prior layer outputs. Max pooling selects the highest value within a rectangular area, useful for identifying brighter pixels in dark backgrounds. Pooling layers reduce feature representation resolution and computation, aiding in overfitting prevention.

Block Diagram:

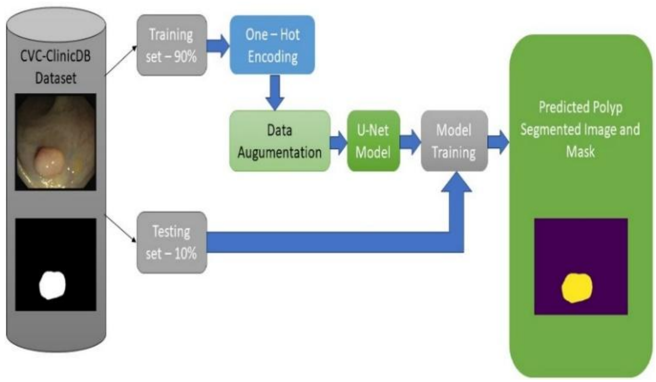

Fig. 2. Block Diagram of Polyp Segmentation

Mathematical Representation of U-Net Segmentation Process

The segmentation process can be expressed as:

$$M_{output} = PostProcess(f_{U-Net}(X)),$$

where:

- X: Input colonoscopy image.
- $f_{U-Net}(X)$: U-Net segmentation output.
- M_{output}: Final mask after applying post-processing techniques ($PostProcess$) to refine boundaries and group objects.

IV. DATASET

The CVC-ClinicDB dataset, used for training and testing, contains 612 colonoscopy images with corresponding ground truth masks. The images include:

- **Size**: 612 images in total, split into 80% for training and 20% for testing.
- **Diversity**: Frames capture polyps under varying lighting conditions, angles, and sizes, ensuring robust model generalization.
- **Quality**: Images are resolution with 384x288 pixels, and annotations are provided by clinical experts to ensure precision.

The dataset is widely used in segmentation challenges, such as the MICCAI 2015 Sub-Challenge on Automatic Polyp Detection in Colonoscopy Videos, validating its relevance.

CVC-ClinicDB database consists of two different types of images:

- Original images: original/frame_number.tiff
- Polyp mask: ground truth/frame_number.tiff

Contains the Frames' & GTs' sequence / file path information:

vpn_keyframe_idsort:
Frame ID (corresponds directly to image & mask names)
vpn_keysequence_idsort:
Video Sequence to which the frame belongs to
text_formattif_image_pathsort:
TIFF file path of Frame
text_formattif_mask_pathsort:
TIFF file path of GT Mask
text_formatpng_image_pathsort:
PNG file path of Frame
text_formatpng_mask_pathsort:
PNG file path of GT Mask

We have 612 paths for each column and 612 images of original polyps and ground truths.

- Images from folder 'Original' are property of Hospital Clinic, Barcelona, Spain.
- Images from folder 'Ground Truth' are propery of Computer Vision Center, Barcelona, Spain.

V. RESULTS AND DISCUSSIONS

A. *Evaluation Criteria:*

The performance of the U-Net model was assessed using two key evaluation metrics: Mean Intersection over Union (IoU) and Mean Dice Loss. These metrics offer a thorough evaluation of the models' capability to accurately segment polyps in colonoscopy images.

B. *U-Net Performance:*

The U-Net model undergoes training on a portion of the CVC-ClinicDB dataset. A stratified random division of the dataset is implemented to ensure an equitable distribution of images between the training and testing sets.

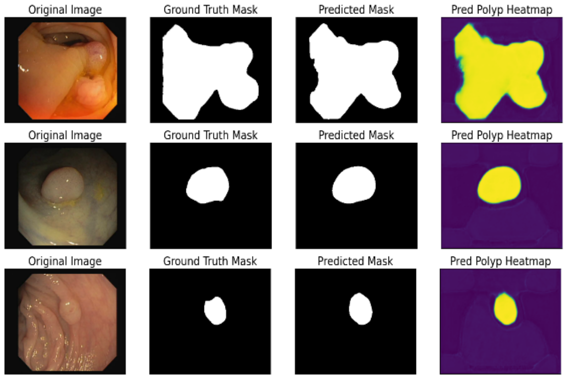

Fig. 3. Predicted Polyp Mask and Heatmap

The model attained an impressive Mean Intersection over Union (IoU) Score of 0.9730, signifying a substantial overlap between the predicted and ground truth polyp masks. Additionally, the Mean Dice Loss for U-Net was determined to be 0.4186, reinforcing its precision.

Impact of Post-Processing Techniques:
To evaluate the role of post-processing in refining segmentation boundaries, we compared model performance with and without these techniques. Post-processing involves edge and valley detection to group neighbouring objects, improving boundary accuracy.

Table 1. Evaluation on Test Data:

Metric	Without Post-Processing	With Post-Processing
Mean IoU Score	0.9730	0.9860
Mean Dice Loss	0.4186	0.2716

The results indicate a significant improvement in both Mean IoU and Dice Loss with post-processing, underscoring its importance in segmentation accuracy.

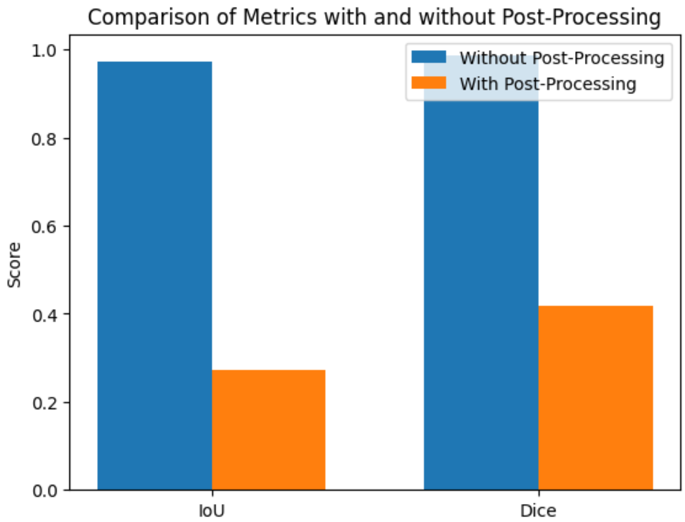

Fig. 4. Comparision of Metrics with and without Post-Processing

The comparison of measures such as IoU and Dice, illustrated in bar charts, demonstrates the effect of post-processing on segmentation precision. Elevated bars for "With Post-Processing" signify enhanced outcomes, quantifying its efficacy in refining initial forecasts. This objective assessment ascertains the value of post-processing and its impact on the model's overall efficacy.

Analyzing the score discrepancies provides insights into the model's strengths and limitations. Significant improvements with post-processing demonstrate it effectively tackles difficulties like fragmented forecasts, whereas small or negative changes can reflect the technique's ineffectiveness or even deleterious impacts. This comparison informs model design and post-processing decisions for optimal segmentation outcomes.

To understand the relation between the Mean IoU Score and Mean Dice Loss of training and validation data, see the graphs below:

The IoU scores for U-Net displayed uniform high performance across the Training and Validation samples, ranging from 0.928061 to 0.981176. The high IoU scores indicate the model's accuracy in capturing polyp boundaries.

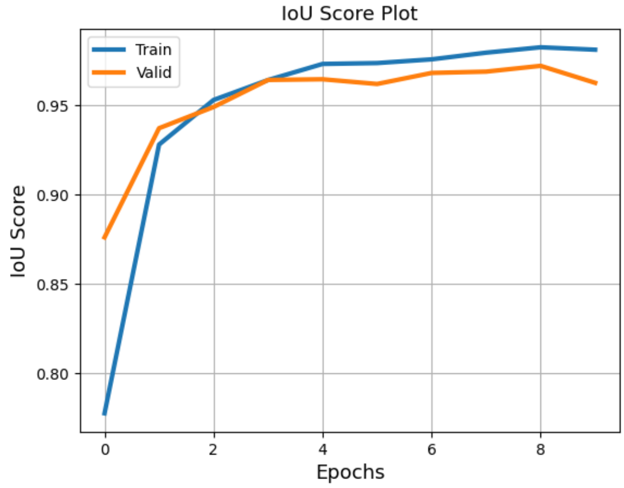

Fig. 5. IoU score Plot on Training Data and Validation Data.

The Dice loss for U-Net demonstrated consistency throughout the test dataset, with values ranging from 0.065844 to 0.344598. This consistent performance underscores the model's resilience in polyp segmentation.

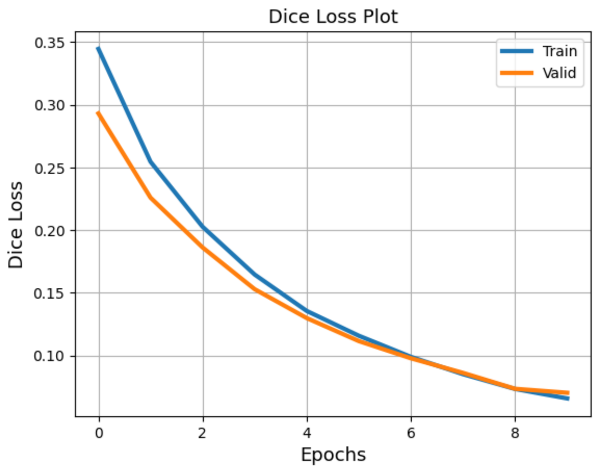

Fig. 6. Dice Loss Plot on Training Data and Validation Data.

Table 2. IoU score and Dice Loss in each Epoch.

dice_loss	iou_score
0.344598	0.777588
0.254641	0.928061
0.202756	0.953035
0.164642	0.964206
0.135607	0.973205
0.11584	0.973697
0.099017	0.975775
0.085182	0.979512
0.07338	0.982528
0.065844	0.981176

VI. CONCLUSION

The primary goal of our proposed model is to enhance the accuracy of polyp detection. By incorporating DCNN techniques based on the U-Net architecture, our model aims to improve the robustness and effectiveness of colorectal cancer screening, along with segmentation precision. The model's performance is evaluated in terms of IoU score and Dice loss. The IoU score measures the overlap between predicted and ground truth segmentations, while Dice loss addresses the imbalance between foreground and background. However, it overlooks the imbalance between easy and difficult examples, which can also hinder the model's training process.

In the domain of medical image segmentation, accurately detecting polyps in colonoscopy videos is crucial for early diagnosis and treatment of colorectal diseases. This study presents an evaluation of the state-of-the-art deep learning model, U-Net, for automatically segmenting polyps in colonoscopic frames using the CVC-ClinicDB dataset. Leveraging its diverse collection of colonoscopy images, the CVC-ClinicDB dataset provides a strong foundation for this comparative analysis. The primary evaluation metrics, Mean Intersection over Union (IoU) and Mean Dice Loss, offer a comprehensive assessment of model performance.

Evaluation results demonstrate that U-Net achieves strong performance in polyp segmentation, with a Mean IoU of 0.9730 and a Mean Dice Loss of 0.4186. Post-processing further enhances these results, leading to a marginal improvement in IoU to 0.9860 and a substantial decrease in Dice Loss to 0.2716, indicating refined segmentation accuracy and agreement. These findings underscore the effectiveness of U-Net, especially when combined with post-processing techniques, for accurate and robust polyp segmentation in clinical applications.

VII. REFERENCES

[1] T.Mahmud, B. Paul, and S. A. Fattah, "PolypSegNet: A modified encoder-decoder T architecture for automated polyp segmentation from colonoscopy images," Comput. Biol. Med., vol. 128, Jan. 2021, Art. no. 104119.

[2] D.-P. Fan, G.-P. Ji, T. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "PraNet:Parallel reverse attention network for polyp segmentation," in Proc. Int.Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2020, pp. 263–273.

[3] Jha, M. A. Riegler, D. Johansen, P. Halvorsen, and H. D. Johansen, "DoubleU-Net: A deep convolutional neural network for medical image segmentation," in Proc. IEEE 33rd Int. Symp. Comput.-Based Med. Syst. (CBMS), Jul. 2020, pp. 558–564.

[4] J. Kang and J. Gwak, "Ensemble of instance segmentation models for polyp segmentation in colonoscopy images," IEEE Access, vol. 7, pp. 26440–26447, 2019.

[5] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham, Switzerland: Springer, 2018, pp. 3–11

[6] L. Zhang, S. Dolwani, and X. Ye, "Automated polyp segmentation in colonoscopy frames using fully convolutional neural network and textons," in Proc. Annu. Conf. Med. Image Understand. Anal. Cham, Switzerland: Springer, 2017, pp. 707–717.

[7] Cai, S. L., Yao, Y. Q., Cao, S. K., Liu, L. R., Zhang, Y. F., Sun, J. J., ... & Chen, G. (2020). Application of deep learning in gastrointestinal imaging: Interpretation of the major findings in the National Polyp Study (NPS) dataset. European Radiology, 30(1), 61-68.

[8] Li, L., Zeng, X., Jin, X., Lu, L., Zhang, H., & Niu, L. (2021). Deep learning in video-based colonoscopy for polyp detection: A systematic review and meta-analysis. Endoscopy International Open, 9(06), E746-E755.

[9] Mahmood, T., Niazi, T., & Jahanzaib, M. (2019). Polyp detection in colonoscopy images using deep convolutional neural networks. Biomedical Signal Processing and Control, 47, 114-126.