Tec Empresarial
P-ISSN: 1659-2395; E-ISSN: 1659-3359

POLY DETECTION USING IMAGE SEGMENTATION

Kalyani Tiwari
Department of Computer Science and Engineering, IPS Academy IES Indore, India
ktiwari.official@gmail.com

Vandana Dubey*
Department of Electronics and Communication Engineering, IPS Academy IES Indore, India
vandanadubey110@gmail.com

Ms. Angeeta Hirwe
Department of Electronics and Communication Engineering, IPS Academy IES Indore, India
angita.hirwe@gmail.com

Ravi Sindal
DAVV,Institute of Engineering & Technology, DAVV Indore, India
ravi.sindal@gmail.com

Tanisha Dhakad
Department of Computer Science and Engineering, IPS Academy IES Indore, India
tanishadhakad17@gmail.com

Aniket Singh
Department of Computer Science and Engineering, IPS Academy IES Indore, India
aniketsinghn10@gmail.com
*Corresponding Author: **Vandana Dubey**

Abstract- in humans, one of the most commonly diagnosed cancer is Colorectal cancer (CRC) along with a major cause of cancer-related mortality. Colonoscopy is a widely used diagnostic procedure for detecting polyps. Early detection of polyps significantly improves treatment outcomes. However, the conventional method of polyp detection through colonoscopy is time-taking and errors are prone to take place.

In the proposed research, the paper presents an elementary approach for the purpose of polyp detection utilizing image segmentation based on the U-Net architecture. We employ the publicly available CVC-ClinicDB dataset, sourced from Kaggle, to develop our model. The proposed method generates a masked image, highlighting the detected polyps, thereby enhancing the accuracy and efficiency of polyp identification.

To improve detection performance, our approach incorporates preprocessing techniques such as contrast enhancement and noise reduction, ensuring optimal input quality for the segmentation model. The U-Net architecture is actually well-suited for medical image segmentation as a result its tendency to capture fine details through skip connections, which bridge high-resolution features from the encoder to the decoder. Additionally, we implement data augmentation strategies, including rotation and flipping, to increase the diversity of training samples and prevent overfitting.

Observational results demonstrate that our model achieves competitive performance in polyp segmentation, with high precision and recall rates. Comparative analysis against traditional methods highlights the superiority of deep learning-based approaches in reducing false positives and improving detection consistency. Future work will explore the integration of real-time processing capabilities and multi-modal imaging to further enhance diagnostic accuracy in clinical settings.

Keywords: Image Segmentation, Semantic Segmentation, Instance Segmentation, Deep Learning.

Introduction- Cancer is a deadly disease , there are many types of cancer but this paper focusses on the colorectal cancer , Cancer in the intestine is referred to as colorectal cancer. In human the large intestine is divided into several parts which are colon , rectum and anus . Among all types of cancer studies show that it is the 3rd major common cancer in men and 2nd in women [1].

In the initial stage of cancer small clumps of cells (also called polyps) are seen in the colon . Mostly these cells are just not harmful and a general lump of cells (benign) and are not cancer and cannot spread to rest of the intestine or body but If the cells are indeed cancer cells (malignant) the polyps can expand in size and conquer the deeper walls of the intestine , and if not treated and more over detected in time cancer cells could enter into the bloodstream and spread to nearing lymphatic nodes , and with time these polyps can take form of a tumour.

Generally colorectal cancer is inherited if your family has a history of colorectal cancer , or if your drink too much or aren't physically active or have a disease that leads to inflammation of colon like ulcerative colitis or Crohn's disease.[2]

Symptoms of Colorectal Cancer

- Change in bowl habits.
- Blood in the stool.
- Feeling unable to empty bowel.
- Belly pain, bloating, or cramps.
- Unexplained weight loss.
- Feeling very weak/tired.
- Vomiting

Role of Colonoscopy in Polyp Detection

Colonoscopy is the most accurate procedure for early stagesidentification and prevention of colorectal cancer (CRC). The procedure starts with an inspection of the anal canal and rectum, including a retroversion examination to check the mucosal areas above the Z-line and internal hemorrhoidal nodes.

The endoscope is then advanced through the colon—sigmoid, descending, transverse, and ascending—until it reaches the cecum. Any remaining contents are cleared during this process. A key step is capturing an image of the appendix opening. If possible, a retroversion examination is also done in the cecum.[3]

The right side of the colon is examined twice using narrow-band imaging (NBI), first from the cecum to the transverse colon and then back to the cecum. To detect flat polyps, methylene blue is applied to the mucosa. The endoscope is then slowly withdrawn, carefully inspecting the mucosa behind the folds[4].

Challenges in Manual Detection

Manual polyp detection in colonoscopy images is a sensitive and tough task to perform. Polyps can be of different shape, size and texture. Appearence of polyps can be low contrast against the surrounding sometimes which make it difficult to distinguish them. Also all the doctors don't have the same level of expertise which can lead to inconsistencies in the detection. Even the expert endoscopist can miss polyps during the diagnosis due to distraction or because of other reasons. Polyps have high risk of being missed in manual detection. Missed polyps can lead to the colorectal cancer which can causse death.[4]

Need for AI-Based Detection

Procedure of Colonoscopy needs high attention for a long period of a time , which can sometimes reduce an endoscopist's ability to focus leads to a mistake in a detection. AI based polyp detection is not only efficient but also essential due to the limitations in manual detection. AI models can identify even the smallest polyps that may be difficult for the human eye to detect[5].

Machine learning techniques, such as Convolutional Neural Networks (CNN), enhance polyp detection through image segmentation using the U-Net architecture.

UNET is convolution neural network architecture that was developed for biomedical image segmentation by Olaf Ronaberger in the year 2015 at the university of freiburg Germany [6]. It is one of the most popularly used architectures for any segmentation task. It is a fully convolution neural network that is designed to learn from a fewer training sample. One of the main feature of the UNET is the use of this encoder decoder structure . Here we have an image that displays an overall architecture of the UNET.[6]

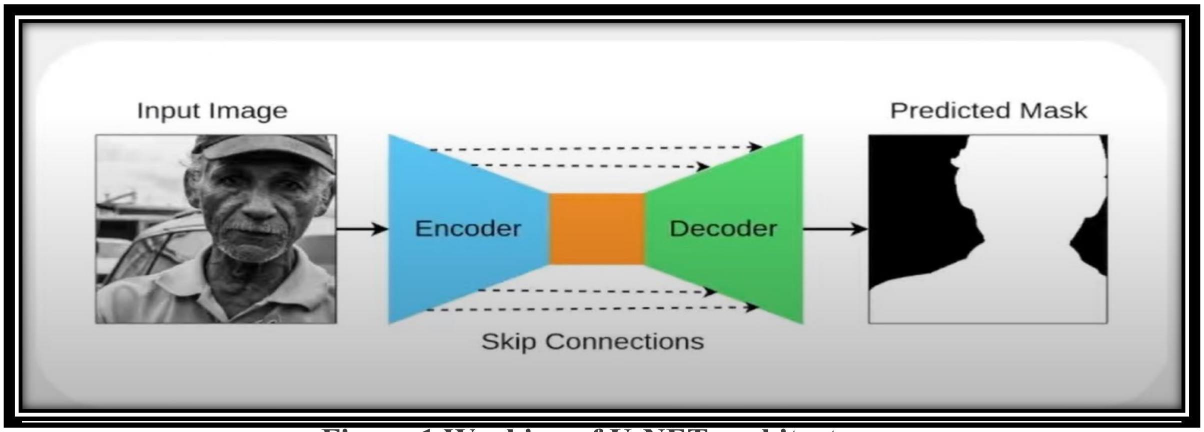

Figure 1 Working of U-NET architecture

Using multiple convolution layers, the image's useful features are fetched important features from the image. Then, the decoder then samples the features by utilising the transpose convolution and integrates them with a process referred as skip connection. Hence, in the result we obtain a segmentation mask from the system.[7]

Another important feature of UNET is the use of skip connection. skip connection as the name suggests, skips some of the layers in the neural network and feeds the output of one layer as the input to the next layers.

If it look closely at this figure you will be able to see the arrows representing the skip connection process. This process helps to fetch and supply the selected features directly from the encoder to the decoder part of the network, which as a result leads to generate a better segmentation mask.[8]

U-Net++is a improved version of U-Net. It was introduced to inhance the performance of U-Net by improving multi-scale feature fusion.It introduces dense convolutional layers in skip connections, allowing a more refined representation of features.

Unlike U-Net, which directly connects encoder and decoder layers, U-Net++ introduces intermediate dense convolutional blocks between them.

It helps in feature enhancement. The added convolutional layers in this architecture refine the feature before merging them, which leads to more better segmentation. U-NET++ architecture has better accuracy then U-NET architecture [9] .

Literature Review

Concurrent improvements in deep learning, especially in Convolutional Neural Networks (CNNs), have crucially improved the performance of object detection and segmentation tasks. Notably, Mask R-CNN [He et al., 2017] has been widely adopted for instance segmentation, allowing precise localization and classification of multiple objects at the pixel level. It extends Faster R-CNN by adding a branch for predicting segmentation masks, proving highly effective for distinguishing between visually similar waste items.

Traditional colonoscopy procedures rely on visual inspection by medical professionals, which may lead to missed polyps, especially when the polyps are small, flat, or hidden within folds of the intestinal wall. As a result, the application of automated methods using artificial intelligence has

gained significant interest in recent years.Among various techniques, deep learning models have shown a remarkable ability to learn visual patterns that are indicative of polyps. One of the widely used architectures in this space is U-Net, which allows pixel-wise segmentation of medical images. [10]
Modified versions, such as ResUNet and U-Net++, aim to improve feature extraction and segmentation precision, making them more suitable for complex shapes and varied polyp appearances.

Object detection models like YOLO and Faster R-CNN have also been explored, particularly for real-time assistance during colonoscopy. These methods are designed to quickly identify and locate polyps in video frames, although they may face limitations in accurately detecting flat or low-contrast polyps.

Author(s) & Year	Methodology / Model	Dataset(s) Used	Main Contributions	Limitations / Challenges
Ronneberger et al. (2015)[6]	U-Net (segmentation network)	General biomedical datasets	Introduced a U-shaped encoder-decoder structure for accurate pixel-wise segmentation.	Not originally tailored for polyp detection; needs domain-specific adaptation.
Akbari et al. (2018)[8]	Deep CNN for detection	CVC-ClinicDB, ETIS-LaribPolypDB	Used convolutional networks to locate and classify polyps with decent accuracy.	Performance declined for small or flat polyps.
Zhang et al. (2019)[8]	ResUNet	Kvasir-SEG, CVC-ClinicDB	Incorporated residual layers into U-Net for improved boundary clarity.	Higher model complexity and resource requirements.

Jha et al. (2021)[7], [8]	DoubleU-Net	Kvasir-SEG, CVC-ClinicDB	Employed dual encoding paths to enhance feature representation for polyp segmentation.	Model is heavy and not suitable for real-time applications.
Brandao et al. (2020)[10], [3]	YOLOv3 (object detection)	Colonoscopy video frames	Delivered real-time detection, aiding live colonoscopic examinations.	Struggles with low-contrast and occluded polyps.

Additionally, recent experiments with attention-based networks and transformer-inspired models aim to make the detection process more context-aware. These systems help the model to focus on regions that are more likely to contain abnormalities, though they are still in the early phases of adoption for this use case. Benchmark datasets such as Kvasir-SEG, CVC-ClinicDB, and ETIS-Larib Polyp DB are frequently used to train and test these models. While they provide a solid starting point, they often suffer from class imbalance and limited diversity, which may hinder the model's ability to generalize across real-world settings.[13]

Despite notable progress, several open problems remain. Many models struggle to adapt to different clinical environments or equipment variations. Additionally, only a few systems have undergone validation in real clinical workflows, making their practical use limited. Therefore, future work should focus on building more efficient, explainable, and robust detection systems that can be trusted in real-time medical diagnostics[14].

Methedology

In this study, we use deep learning techniques for polyp detection through image division. We appoint a widely used model for U-Net architecture, medical image segmentation, so that polyps can be detected from colonoscopy images. The process includes several stages, including dataset preparations, preprocessing, data growth, model design, training and evaluation.

Dataset Description

The CVC-ClinicDB [7]database comprises of frames extracted from colonoscopy videos. The dataset incorporates many examples of polyp frames & corresponding ground truth for them. The actual Truth image consists of a mask corresponding to the region covered by the polyp in the image. The data is accessible in both .png and .tiff formats . a total of 610 images are present in the dataset.

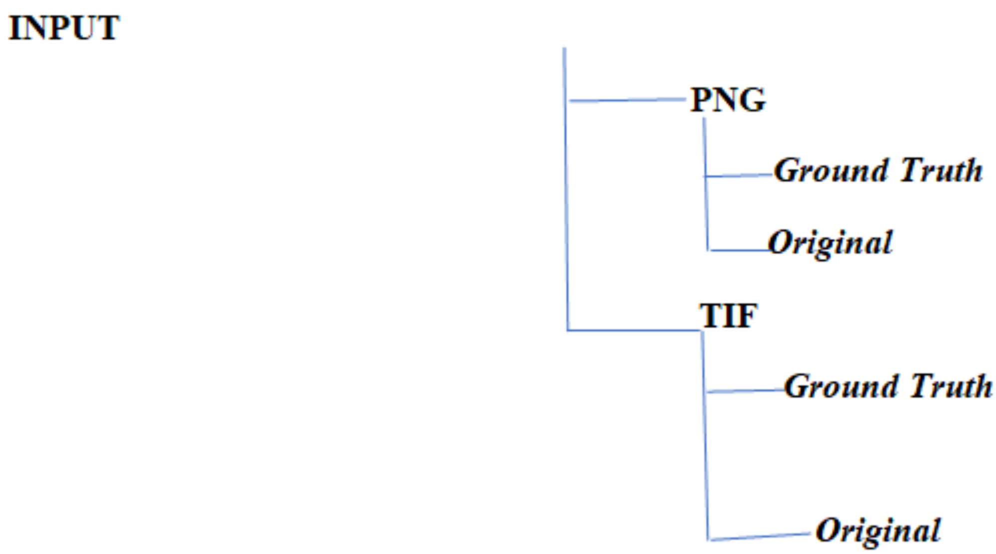

Figure 2 Structure of Dataset

Data Preprocessing

Data preprocessing is an important step in the medical image division. Before training the model, the images go through several steps to ensure consistency and improve performance. These steps include:

Resizing: All images were resized to 256*256 to ensure stability and consistency and optimal use of memory.

Normalization: Normalization is the process which helps the model to learn new features effectively without giving any bias results towards high pixel intensity values and also prevents large gradient updates which can destabilize the training process (pixel values range from 0 -255 - unstable training). Normalization will help easy and fast convergence and better gradient flow , Mainly there are two types of normalization RGB values , either [-1,1] or [0,1] using the transform. Normalize function which takes standard deviation and mean as input.

Color Space Conversion: All images are then converted first to RGB format(standardizing for better compatibility).

image = Image.open(img_path).convert("RGB")

Masks are converted to grey Scale(L mode) , to represent a class and a label.

mask = Image.open(mask_path).convert("L")

Data Augmentation

Because medical datasets are quite small. Data Augmentation helps us to generate new variations of image and improve the generalization. Augmentations enhance model robustness by making it adaptable to real-world variations.

We have used Data Augmentation Techniques Random flipping. In this technique Random horizontal and vertical reflections is created.

Figure 3 sample image conversion from RGB to Grayscale

Data Augmentation

Because medical datasets are quite small. Data Augmentation helps us to generate new variations of image and improve the generalization. Augmentations enhance model robustness by making it adaptable to real-world variations.

We have used Data Augmentation Techniques Random flipping. In this technique Random horizontal and vertical reflections is created.

Model Architecture

We have used UNET AND UNET++ architecture which is a deep learning architecture designed for image segmentation.

U-Net consists of two parts:-

Contracting path(Encoder): It extracts important denatures from the image using cnn layers.

Expanding path(Decoder): It uses transposed convolutions just to restore the size of image. It merges deatures from the Encoder using the process called skip connection to improve the acuuracy of overall image segmentation.

After using we have used U-Net++ for the better segmentation and we have observed that the loss of data is less in U-Net++ than the U-Net.

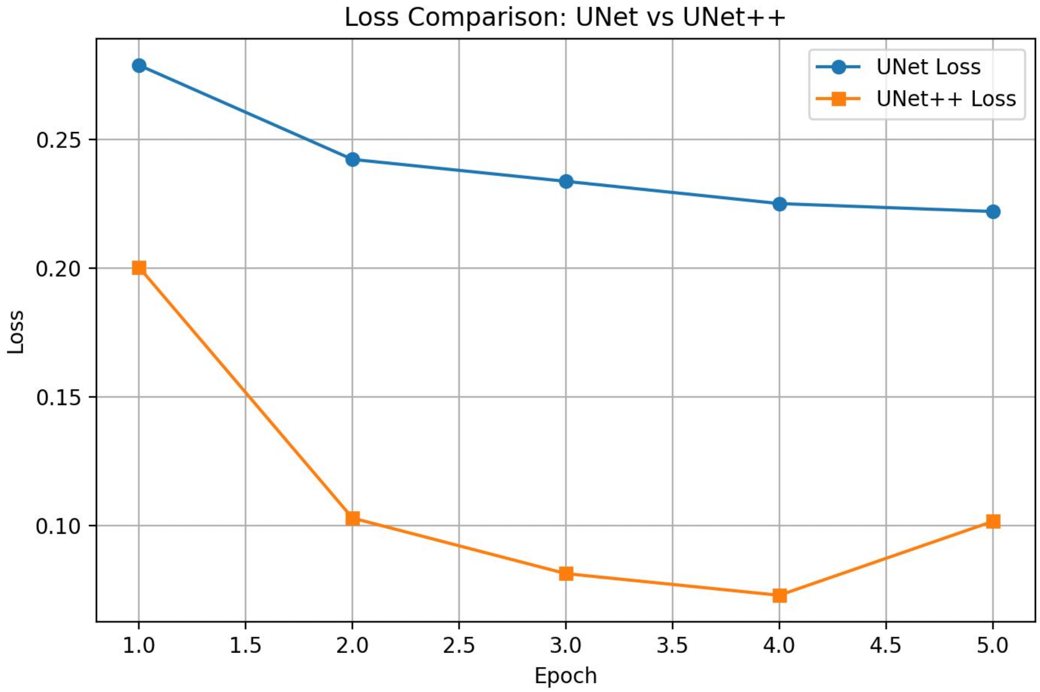

Figure 4 loss comparison between U-NET and U-NET++

As you can see in the above image of graph where the comparison of data loss in both U-Net and U-Net++ is shown. If we see in the graph then we can observe clearly that U-Net++ architecture is the better one.

Model Training

We have use Dice loss through which we can measure the overlap between predicted mask and ground truth masks.We use Adam optimizer for efficient learning . The models learns by reducing the dice loss over multiple training iterations.

Code:-

```
or epoch in range(epochs):
for images, masks in train_loader:
optimizer.zero_grad()
predictions = model(images)
loss = dice_loss(predictions, masks
loss.backward()
optimizer.step()
```

Model Evaluation

We have evaluated the performance of our model using Dice Coefficient and IoU(Intersection over Union) which measures the segmentation accuracy.

Dice Coefficient

Dice Coefficient is used to optimize segmentation model is defined as:-

$$\text{Dice} = 2\ |A \cap B|$$

$$|A|+|B|$$

$|A \cap B|$ is representing the number of overlapping pixels. Dice Loss is defined as:

$$LDice = 1-DSC = 1-2\sum pigi$$
$$\sum pi+\sum gi$$

Intersection over Union(IoU)

IoU is also known as the Jaccerd Index. IoU measures the intersection or overlap between the ground truth and the predicted segmentation.

$$IoU = | A \cap B|$$
$$| A \cup B|$$

Result and Declarations

As a result of , Polyp detection model is performing well in both the architectures U-Net and U-net++, But the accuracy in U-Net++ is better that U-Net comparatively[9][10].

Performance comparison of U-NET and U-NET++

Polyp segmentation models, evaluation of UNET and UNET ++, major performance metrics: accuracy, union (IU) at the intersection, was performed using (IU), F1 score, and recall. These matrix provides the information about the effectiveness of the model in identifying polyps from colonoscopy images.

Accuracy Comparison: Accuracy comparison graph (Fig-E) reflects the performance of both models in various image samples. UNET usually maintains a high accuracy, although UNET ++ displays competitive results in some cases. The accuracy alone does not fully reflect the quality of the model division, especially for small or complex polyp structures .

Intersection over Union(IoU) comparison: The IoU measures overlap between predictive and ground truth masks. IOU comparison graphs (Fig-E) highlight the ups and downs in performance in images. UNET ++ shows a high IU score in some examples, indicating better division purification for specific samples. However, the overall IU values are relatively low, suggesting challenges in dividing complex polyp boundaries .

F1 score and Recall comparison: The F1 score balances and remembers accuracy, making it an important metric for medical image segmentation. The F1 score comparison graph (Fig-E) suggests that UNET + + receives high scores for some samples, indicating better sensitivity in polyp detections. Remember, which measures the ability of the model to correctly detect polyps, reflects similar trends. High recall values for UNET ++ suggest that it is more effective in identifying polyps, but its variability in samples indicates inconsistency in performance.

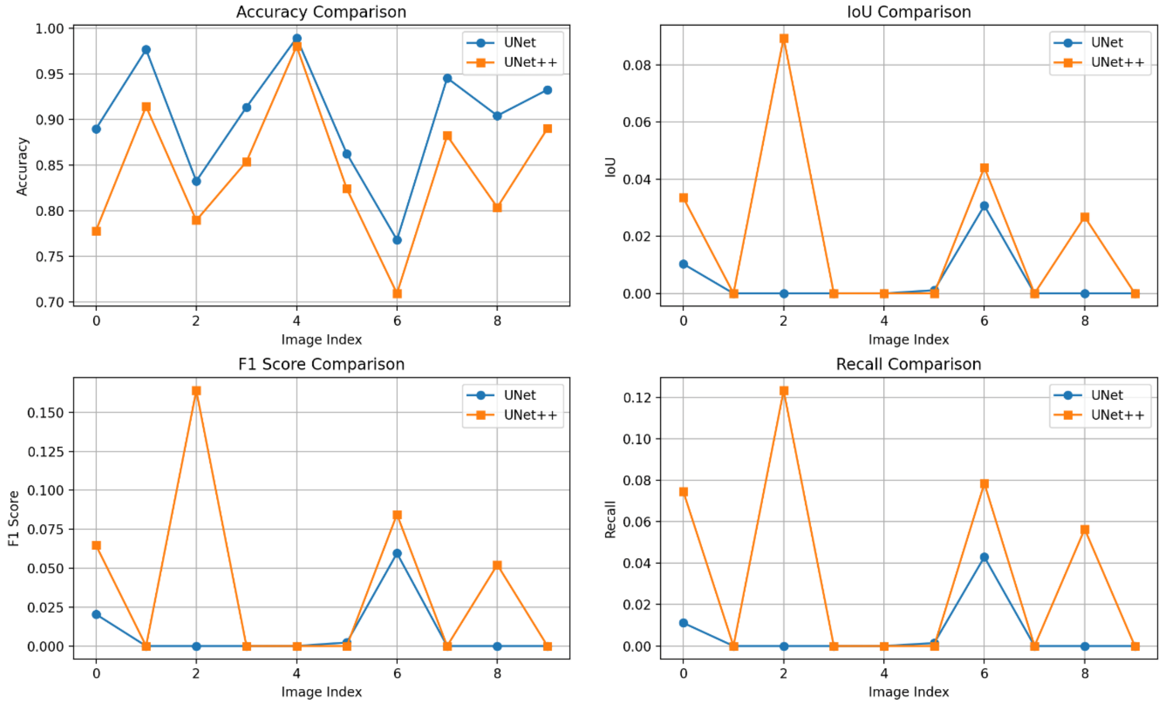

Figure 5 Accuracy, IoU, F1 score and recall comparison

The results suggest that UNET ++ improves the UNET in terms of IoU, recall and F1 scores, making it a better option for polyp segmentation in medical imaging. However, the accuracy remains comparable between the two models, and performance fluctuations suggest that further refinement, such as data growth, better loss functions or post-processing techniques may be required to increase the quality of division.Future reforms can include the integration of meditation mechanisms or transformer-based architecture to improve polyp division accuracy and strength.

Our model serves as a valuable resource for learning purposes, as it has been trained on a relatively small dataset which can provide you the understanding of the polyp detection but it will be better if we perform this on a bigger datasetThe steps and techniques outlined in this paper can be used on the bigger datasets to improve the efficiency, accuracy and for the better segmentation.

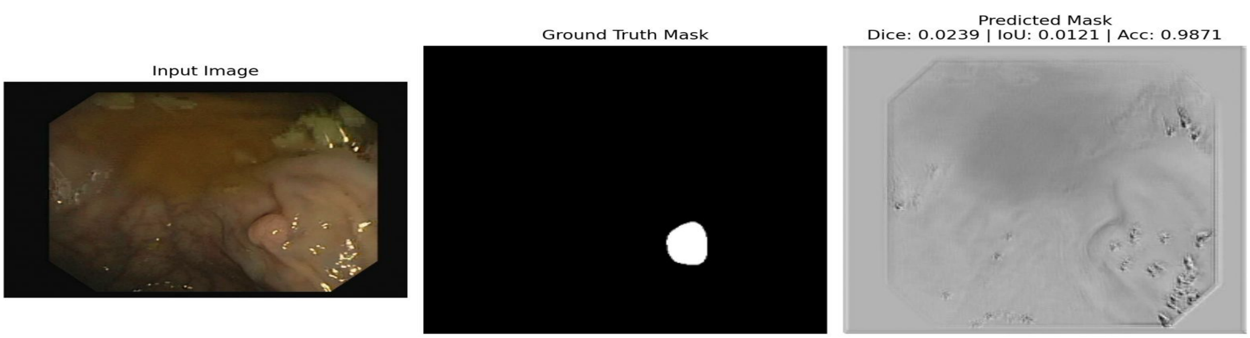

Figure 6 output of polyp detection

Challenges observed during the detection

Despite the promising results of the U-Net model for polyp detection, many challenges faced during the experiment and evaluation stages. One of the primary issues, different samples had variability in polyp shape, size and texture. Polyps can range from small, flat wounds to large, spreading mass, which can lead to effective normalization for the model. Small or sessile polyps often mix in the surrounding mucosa, causing abortion or incomplete division in some cases.

Another major challenge we face during the process of detection is ,Computer barriers created boundaries during training. High-resolution medical images require an increase in significant processing power, training time and memory consumption. Future work can focus on adapting the model architecture to balance accuracy and efficiency for real -time clinical applications.

Conclusion

In This Research paper, we have explored Polyps detection by using the U-Net framework for colonoscopy images, aiming to provide a better way of polyps detection and an early diagnosis of Colorectal cancer. Through our Research and study we came to the conclusion that deep learning-based segmentation, particularly with U-Net and U-Net++, can effectively identify polyps with high accuracy.

We evaluated the effectiveness of UNET and UNET ++ for polyp detection and division in colonoscopy images. Results indicate that UNET ++ improves the UNET in terms of IoU, recall, and F1 scores, which demonstrates its ability to catch fine polyp boundaries and improve partition accuracy. Additionally, the comparison of the disadvantage suggests that UNET ++ changes more efficiently, making it a better option for this medical imaging work.

However, UNET ++, despite showing better performance in many metrics, high -rational and recall values in the IU, highlight the challenges in detecting polyps with different shapes, sizes and textures. This suggests that further adaptation, such as data growth, hybrid deep learning architecture and advanced post-processing technology, can significantly increase the model reliability.

Future work will focus on improving partition performance by integrating the attention mechanism, transformer-based architecture and domain adaptation techniques. Additionally, it would be important to include real -time estimates and deploy these models in clinical settings.

Eventually, accurate polyp division is crucial for early diagnosis and averting of colorectal cancer. The results of this study provides to the advancement of deep learning models in medical image analysis, paving the way for more reliable and automated polyp detection systems.

References

I. 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.

II. Polyp Detection Using U-Net Neural Network Based Algorithm, Bipin Kumar Rai, Devansh Singh, Archita Shukla , Published in International Conference on… 15 March 2024, Medicine, Computer Science.

III. Real-time polyp detection model using convolutional neural networks,Alba Nogueira-Rodríguez, Rubén Domínguez-Carbajales, +11 authors D. Glez-Peña, Published in Neural computing… 21 September 2021, Medicine, Computer Science

IV. Detection of Colorectal Polyps from Colonoscopy Using Machine Learning: A Survey on Modern Techniques, Khaled ELKarazleV. RamanP. ThenC. Chua, Medicine, Computer Science, Italian National Conference on Sensors ,2023

V. Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information, Nima Tajbakhsh, S. Gurudu, Jianming Liang, Published in IEEE Transactions on Medical… 1 February 2016, Computer Science, Medicine

VI. A review on Polyp Detection and Segmentation in Colonoscopy Images using Deep Learning, Krishnendu S , Geetha S , Gopakumar GPublished (First Online): 20-10-2020, ISSN (Online) : 2278-0181, Publisher Name : IJERT

VII. Akbari, M., Mohrekesh, M., Rafiei, S., Haji Maghsoudi, L., Dehnavi, A. M., & Abolghasemi, J. (2018). Polyp segmentation in colonoscopy images using fully convolutional network. In 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME) (pp. 1–5). IEEE (Akbari et al., 2018).

VIII. Zhang, R., Zheng, Y., Poon, C. C. Y., Shen, D., & Lau, J. Y. W. (2019). Polyp detection during colonoscopy using a regression-based convolutional neural network with a tracker. Pattern Recognition, 83, 209–219 (Zhang et al., 2019)

IX. Jha, D., Smedsrud, P. H., Riegler, M. A., Halvorsen, P., de Lange, T., Johansen, D., & Johansen, H. D. (2020). ResUNet++: An advanced architecture for medical image segmentation. In 2020 IEEE International Symposium on Multimedia (ISM) (pp. 225–2255). IEEE (Jha et al., 2020).

X. Brandao, P., Mazomenos, E., & Tsaftaris, S. A. (2018). Towards real-time polyp detection with region-based deep learning. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI) (pp. 115–118). IEEE (Brandao et al., 2018).

XI. A review on Polyp Detection and Segmentation in Colonoscopy Images using Deep Learning, Krishnendu S , Geetha S , Gopakumar G Published (First Online): 20-10-2020, ISSN (Online) : 2278-0181, Publisher Name : IJERT (Krishnendu et al., 2020).

XII. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Ca: Cancer J Clin (2021) 171:209–49. doi: 10.3322/caac.21660 [] [] [] (Sung et al., 2021).

XIII. Bashar MK, Kitasaka T, Suenaga Y, Mekada Y, Mori K. Automatic detection of informative frames from wireless capsule endoscopy images. Med Image Anal (2010) 14:449–70. doi: 10.1016/j.media.2009.12.001 [] [] [] (Bashar et al., 2010).

XIV. Kim T, Lee H, Kim D. (2021). UACANet: Uncertainty augmented context attention for polyp segmentation, in: Proceedings of the 29th ACM International Conference on Multimedia, New York, USA: ACM. pp. 2167–75. doi: 10.1145/3474085.3475375 [] [] (Kim et al., 2021).