

Tumor and Polyp Segmentation in Medical Images Using Deep Learning

S.Navya sri, Dr. Om Prakash Samantray(Associate professor)

CSE(DS),Anil Neerukonda Institute Of Technology And Sciences,Andhra Pradesh
, Viskhapatanam, India

Email: navyasrisabbi@gmail.com

Abstract— Medical image segmentation is one of critical parts in disease detection in early stage for example brain tumors and gastrointestinal polyps. The task of segmentation helps doctors in early diagnosis and then treatment. However, image segmentation is challenging task and time consuming. Even a single image requires a long time for accurate segmentation and also may produce inaccurate results for different human interpreters.

We proposed a novel system, which is based on deep learning, for the tumor and polyp segmentation from the medical images. In the proposed system, a U-Net-based model is utilized for segmenting the tumors and polyps from the preprocessed input images. The system generates the tumor and polyp masks that precisely detect the abnormally grown portions from the brain MRI and the colonoscopy images. Through a number of experimental results, the efficiency of the proposed system for the tumor segmentation has been evaluated and the quantitative assessment metrics, such as the Dice Score (0.9193), the IoU (0.8506), the Precision (0.9023), and the Recall (0.9369) have been calculated and presented for the polyp segmentation Dice Score (0.9716), the IoU (0.9447), the Precision (0.9447), and the Recall (0.9594). The obtained results and the corresponding output of the tumor segmentation are graphically represented and the efficiency of the model in detecting the tumors has been illustrated.

Keywords: Medical Image Segmentation, Deep Learning, U-Net, Brain Tumor Segmentation, Polyp Segmentation, Dice Score, IoU, Precision, Recall.

I. Introduction

Medical image analysis plays an important role in the modern healthcare systems for the early detection and treatment of diseases. Accurate segmentation of abnormal structures in images is crucial for the effective diagnosis

and treatment of the diseases. There are a variety of medical abnormalities that require precise segmentation. In this work, we have focused on the tumors of the brain and the gastrointestinal polyps, which can be life threatening if not detected in time and can transform into colorectal cancer if left untreated after their polyp formation.

Tumors are abnormal growths of cells. These can occur in the brain and affect the normal function of the brain. They can be life threatening and must be detected as early as possible. Polyps in the gastrointestinal tract are small benign growths which can develop into cancer, most commonly into colorectal cancer, and must be detected and treated in time.

Manual segmentation methods are typically time-consuming, and even more so as the amount of images is increasing exponentially. In order to improve efficiency, automatic methods have been developed and are increasingly used in medical applications. By using Artificial Intelligence (AI) in combination with Deep Learning (DL) techniques, high-quality segmentation results can be achieved automatically within a short time frame.

The field of Deep Learning, and in particular of Convolutional Neural Networks (CNNs), has been rapidly growing in recent years, especially for what concerns image processing and computer vision applications. A huge number of architectures have been proposed in the state of the art, each of them with its strengths and weaknesses. In the specific field of biomedical image segmentation, one of the most commonly used architectures is U-Net. This network is based on an encoder-decoder structure and it is able to perform segmentation by exploiting the features of the input images. U-Net is able to extract both the spatial information and the contextual information and, thanks to

the skip connections between the encoder and the decoder, it is able to perform a fine segmentation of the images. Thanks to these characteristics, U-Net has been widely used for biomedical image segmentation and it has been implemented for the segmentation of many different types of images.

In this work, we have implemented a deep learning-based system to automatically segment all kinds of tumors and polyps from the corresponding medical images (MRI and colonoscopy images) by using U-Net architecture. The input images are processed in order to detect abnormal areas and to produce the corresponding segmentation masks. The performance of the proposed system is evaluated by using well known segmentation measures including Dice Score, IoU, Precision and Recall.

The proposed U-Net for automatic segmentation of tumors and polyps in medical images showed a high level of performance. The produced segmentations were analyzed using the established metrics for segmentation such as the Dice Score, IoU, precision, and recall. The results show a good performance for the U-Net with a Dice Score of 0.9193, an IoU of 0.8506, a precision of 0.9023, and a recall of 0.9369 for the segmentation of tumors and for polyps Dice Score 0.9716, the IoU 0.9447, the precision 0.9447, and the recall 0.9594. This model is able to automatically produce a mask that is able to highlight the abnormal structures present in images, and thus to support the medical image analysis. The system could also support the doctors in the decision process and support the automated clinical analysis.

II. Literature Review

The segmentation of medical images has become an important task in the field of medicine. It helps to detect abnormalities such as tumors and polyps early, which can lead to an effective cure if treated on time. A number of approaches have been put forth by researchers to improve the performance of a model that is used for segmenting images. These models are generally based on deep learning approach that help improve the feature extraction of the images.

In 2020, the automatic model for polyp segmentation by using U-Net with ResNet50 as the backbone of the encoder was proposed in [1]. This model, incorporated the encoder-decoder structure of the typical U-Net with a pre-trained ResNet50 as the powerful encoder for feature extraction, achieved very accurate results in polyp segmentation on the public Kvasir-SEG dataset.

Focus U-Net, proposed in 2021, utilizes a dual attention mechanism that enhances spatial and channel feature learning for the improvement of the accuracy of segmentation. The Focus U-Net achieves good results on the segmentation of polyps in the CVC-ClinicDB and Kvasir-SEG databases.

In 2022, UPolySeg [3] by introducing some changes in the pre-processing step and also by tuning the parameters of the U-Net architecture in order to enhance the performance for the polyp segmentation, achieved a reliable performance on the Kvasir-SEG dataset for polyp segmentation.

UNet++ [4] extend a basic U-Net for the introduction of skip connection. By this extension, UNet++ is able to improve feature learning and to reduce the information loss between the encoder and the decoder of the U-Net architecture. For tumor segmentation, a high accuracy could be achieved by using the BraTS dataset.

A Transformer-based model for the medical segmentation was introduced in 2023 in [5]. This approach attempts to enhance the global features by learning from a CNN and subsequently from the Transformer in the following sequence. The study assessed the approach using the datasets for the tumor and polyp segmentation, i.e. the BraTS and Kvasir-SEG datasets.

In 2023, an approach for the attention-based segmentation of tumors in medical images was proposed in [6]. In this approach, a modified U-Net, so-called Attention U-Net, was used, where so-called attention-gates were integrated in order to highlight relevant parts of the image, while unnecessary information is suppressed in the process. The approach was tested on the datasets LiTS and BraTS using the corresponding metrics in order to check the segmentation accuracy.

Dilated U-Net-Seg is a novel architecture introduced in 2024 that employs dilated convolutions for better feature extraction at multiple scales, especially to aid in the detection of small polyps in colonoscopies. The model showed better segmentation performance using the CVC-ClinicDB and Kvasir datasets.

Another study from 2024 proposed a ConvNeXt and Transformer-based segmentation model. It combined encoder-decoder architecture with attention mechanisms to improve feature extraction and segmentation

performance. This model achieved promising results on the Kvasir-SEG dataset.

In 2025, PolypSeg-GradCAM combined the U-Net architecture with Grad-CAM to enhance explainability in segmentation results. This method helped pinpoint important regions responsible for predictions and improved model transparency using the Kvasir-SEG dataset.

A deep learning-based survey on polyp segmentation published in 2025 reviewed around 146 segmentation models and compared their performance across several datasets. The study pointed out the strengths, weaknesses, and future areas for improvement of deep learning methods in medical image segmentation.

From the review of previous studies, it is clear that deep learning approaches provide better segmentation performance than traditional methods. Among various models, the U-Net architecture remains one of the most commonly used and effective methods due to its simple design and reliable segmentation abilities. Therefore, this project adopts the U-Net model for tumor and polyp segmentation in medical images.

III. METHODOLOGY

A. System Architecture

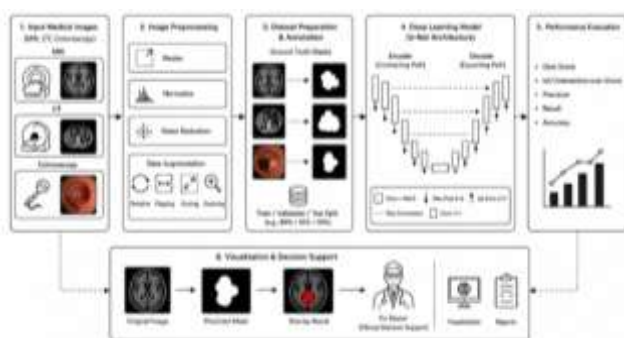


Figure 1. Proposed System Architecture for Tumor and Polyp Segmentation Using U-Net

Figure 1 shows the proposed system architecture for automatic tumor and polyp segmentation with the U-Net deep learning model. The system includes image preprocessing, dataset preparation, model training, segmentation prediction, and performance evaluation.

First, medical images, such as MRI and colonoscopy images, undergo preprocessing. This includes resizing, normalization, and noise reduction. Data augmentation

methods, like rotating, flipping, and scaling, are used to improved model performance and reduce overfitting.

The processed image and ground truth masks are provide to U-Net model for the training. The encoder extracts key features from the input images, while the decoder reconstructs accurate segmentation masks using skip connections. After training, the model predicts tumor and polyp regions from new medical images.

Finally, the performance of the proposed system is assessed using metrics like Dice Score, IoU, Precision, Recall, and Accuracy. The generated outputs consist of predicted masks and overlay visualization results for medical image analysis.

B. Dataset Collection

The proposed model uses publicly available medical imaging datasets:

- BraTS Dataset for brain tumor segmentation
- Kvasir-SEG Dataset for polyp segmentation

These datasets contain medical images and corresponding ground truth masks. They are used for supervised deep learning training.

C. Image Preprocessing

Medical image preprocessing is done to improve image quality and segmentation accuracy.

The preprocessing steps include:

- Image resizing to fixed dimensions
- Image normalization
- Noise reduction
- Contrast enhancement

All images and masks are converted into suitable formats before training the model.

D. Data Augmentation

To increase dataset diversity and reduce overfitting, data augmentation techniques are used during training. The augmentation methods include:

- Rotation
- Horizontal flipping
- Vertical flipping
- Zoom transformation

These techniques improve the robustness and generalization ability of the segmentation model.

E. Proposed U-Net Architecture

This segmentation framework uses the U-Net architecture, consisting of encoder and decoder networks connected via skip connections.

1. Encoder Network

Medical images contain important features. The encoder uses:

- Convolution layers
- The ReLU activation function
- Max pooling

To extract the important features and reduce the image complexity.

2. Decoder Network

The decoder uses:

- Up-sampling layers
- Feature concatenation
- Convolution operations to reconstruct

3. Skip Connections

In the U-Net architecture, skip connections transfer feature sets from encoder layers to decoder layers, conserve fine details, and add accuracy to segmentation.

The U-Net model is used to identify the tumor and polyps with segmentation masks.

F. Model Training

Model Training

The model is developed using any of the deep learning frameworks, TensorFlow, or PyTorch. The dataset is split into:

- 80% Training Data
- 10% Validation Data
- 10% Testing Data

Training Parameters

Parameter	Value
Optimizer	Adam
Learning Rate	0.0001
Batch Size	4

Epochs	40
Loss Function	Dice Loss + Binary Cross Entropy
Input Image Size	256×256

The segmentation accuracy and overlap performance improve using this combined loss function

G. Segmentation Prediction

Once the model has been trained, it predicts the segmentation masks for previously unobserved medical photographs. These masks successfully identify the regions of tumors and polyps in MRI and colonoscopy images.

The segmentation results show a notable advancement in the detection of abnormal regions and in the localization of boundaries.

H. Performance Evaluation

To assess the performance of the proposed system, conventional segmentation metrics of the Dice Score, IoU, Precision, and Recall were used.

•Dice Score

$$\text{Dice} = \frac{2|P \cap G|}{|P| + |G|}$$

•Intersection over Union (IoU)

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|}$$

•Precision

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

•Recall

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

IV. Experimental Results and Discussion

S.No	Metric	Tumor Values	Polyp values
1	Dice score	0.9193	0.9716
2	IoU	0.8506	0.9447
3	Precision	0.9023	0.9841
4	Recall	0.9369	0.9594

Table 2. Performance Evaluation metrics of tumor and polyp

According to the experimental results using the proposed U-Net model, the segmentation resulted in the tumor being identified with a Dice Score of 0.9193, an IoU score of 0.8506, with Precision and Recall scores of 0.9023 and 0.9369 and for polyp Dice Score of 0.9716, an IoU score of 0.9447, with Precision and Recall scores of 0.9841 and 0.9594 respectively. The segmentation masks successfully highlighted the tumor in the abnormal region and can be useful in the automated analysis of medical images. The performance evaluation results of the U-Net segmentation model are shown in Table 1.

1. Result and Discussion

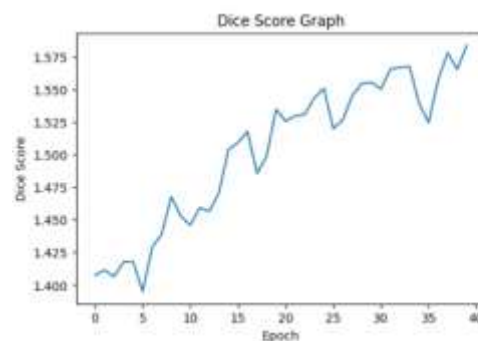
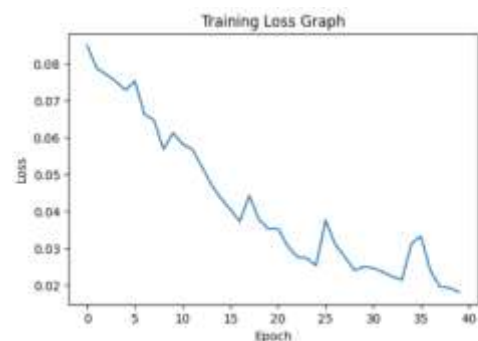
The proposed model of U-Net is an improved architecture deep learning model for automatic image segmentation of tumors and polyps in medical images. For the assessment of the proposed model, the most common assessment metrics were used; Dice Score, IoU, Precision and Recall.

The results of the experiments demonstrated that the proposed framework, based on deep learning, achieved a high performance in a detection task of road surface in complex scenarios, characterized by a Dice Score of 0.9193, an IoU of 0.8506, a Precision of 0.9023 and a Recall of 0.9369 for polyp Dice Score of 0.9716, an IoU score of 0.9447, a Precision of 0.9841 and a Recall of 0.9594.

```

Epoch [1/40] Loss: 0.0949, Dice Score: 1.4075
Epoch [2/40] Loss: 0.0703, Dice Score: 1.4114
Epoch [3/40] Loss: 0.0721, Dice Score: 1.4004
Epoch [4/40] Loss: 0.0753, Dice Score: 1.4126
Epoch [5/40] Loss: 0.0729, Dice Score: 1.4177
Epoch [6/40] Loss: 0.0751, Dice Score: 1.3961
Epoch [7/40] Loss: 0.0661, Dice Score: 1.4298
Epoch [8/40] Loss: 0.0640, Dice Score: 1.4385
Epoch [9/40] Loss: 0.0569, Dice Score: 1.4677
Epoch [10/40] Loss: 0.0613, Dice Score: 1.4530
Epoch [11/40] Loss: 0.0581, Dice Score: 1.4439
Epoch [12/40] Loss: 0.0509, Dice Score: 1.4594
Epoch [13/40] Loss: 0.0522, Dice Score: 1.4563
Epoch [14/40] Loss: 0.0474, Dice Score: 1.4706
Epoch [15/40] Loss: 0.0436, Dice Score: 1.5036
Epoch [16/40] Loss: 0.0407, Dice Score: 1.5091
Epoch [17/40] Loss: 0.0373, Dice Score: 1.5177
Epoch [18/40] Loss: 0.0449, Dice Score: 1.4856
Epoch [19/40] Loss: 0.0376, Dice Score: 1.4987
Epoch [20/40] Loss: 0.0351, Dice Score: 1.5347
Epoch [21/40] Loss: 0.0354, Dice Score: 1.5255
Epoch [22/40] Loss: 0.0307, Dice Score: 1.5298
Epoch [23/40] Loss: 0.0278, Dice Score: 1.5308
Epoch [24/40] Loss: 0.0274, Dice Score: 1.5434
Epoch [25/40] Loss: 0.0254, Dice Score: 1.5508
Epoch [26/40] Loss: 0.0306, Dice Score: 1.5199
Epoch [27/40] Loss: 0.0311, Dice Score: 1.5267
Epoch [28/40] Loss: 0.0278, Dice Score: 1.5454
Epoch [29/40] Loss: 0.0241, Dice Score: 1.5545
Epoch [30/40] Loss: 0.0291, Dice Score: 1.5534
Epoch [31/40] Loss: 0.0247, Dice Score: 1.5580
Epoch [32/40] Loss: 0.0236, Dice Score: 1.5657
Epoch [33/40] Loss: 0.0224, Dice Score: 1.5689
Epoch [34/40] Loss: 0.0215, Dice Score: 1.5674
Epoch [35/40] Loss: 0.0313, Dice Score: 1.5264
Epoch [36/40] Loss: 0.0333, Dice Score: 1.5244
Epoch [37/40] Loss: 0.0239, Dice Score: 1.5576
Epoch [38/40] Loss: 0.0196, Dice Score: 1.5781
Epoch [39/40] Loss: 0.0193, Dice Score: 1.5654
Epoch [40/40] Loss: 0.0182, Dice Score: 1.5830
Training Completed!

```



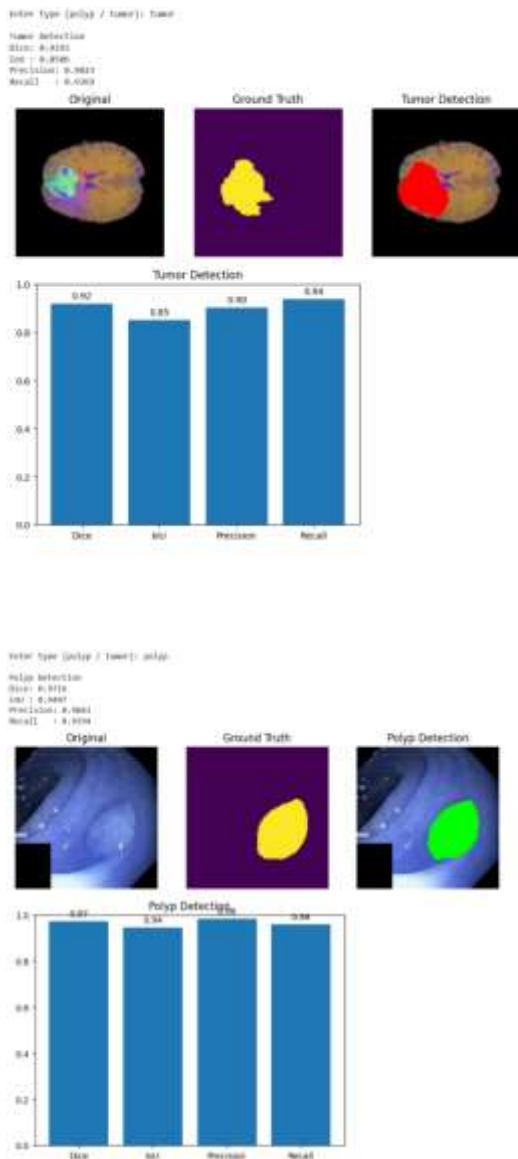


Figure 2. Output Results Showing Original image, Ground Truth Mask and Predicted Segmentation

V. CONCLUSION

We built a tumor and polyp segmentation tool integrating a U-Net inspired architecture and deep learning. The primary objective was to lessen the burden of manual segmentation in medical image review.

From the experiments we conducted, we demonstrated that our model accurately completed segmentation and defined features for accurate clinical judgement and diagnosis.

We demonstrated that a modified U-Net architecture learns the ability to perform segmentation tasks in medical imaging. We have demonstrated that this approach has the potential to be a strong tool in medical imaging and segmentation, particularly with the optimization of large image segmentation in a clinical setting.

VI. References

- [1] Y. Zhang, H. Li, and M. Sun, "Explainable Deep Learning for Polyp Segmentation using Grad-CAM," 2025.
- [2] T. Chen, S. Liu, and X. Wang, "Dilated U-Net for Small Object Medical Segmentation," *Computers in Biology and Medicine*, 2024.
- [3] J. Ma, Y. He, F. Li, et al., "Medical Image Segmentation using Deep Learning: A Survey," *Artificial Intelligence Review*, 2023.
- [4] P. Bilic, P. F. Christ, E. Vorontsov, et al., "The Liver Tumor Segmentation Benchmark (LiTS)," *Medical Image Analysis*, 2023.
- [5] X. Huang, L. Wang, and J. Zhang, "Hybrid CNN-Transformer Network for Tumor Segmentation," *Biomedical Signal Processing and Control*, 2023.
- [6] A. Hatamizadeh, Y. Tang, D. Nath, et al., "UNETR: Transformers for 3D Medical Image Segmentation," *WACV*, 2022.
- [7] S. Minaee, Y. Boykov, F. Porikli, et al., "Image Segmentation using Deep Learning: A Survey," *IEEE TPAMI*, 2022.
- [8] D. Jha, S. Ali, K. Emanuelsen, et al., "A Comparative Study on Colorectal Polyp Segmentation using Deep Learning," *IEEE Access*, 2021.
- [9] S. Ali, D. Jha, N. Ghatwary, et al., "An Objective Comparison of Deep Learning-based Polyp Segmentation Methods," *Scientific Reports*, 2021.
- [10] H. Cao, Y. Wang, J. Chen, et al., "Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation," 2021.
- [11] D. Jha, P. H. Smedsrud, M. A. Riegler, et al., "Kvasir-SEG: A Segmented Polyp Dataset," *MMM*, 2020.
- [12] R. Azad, M. Asadi-Aghbolaghi, M. Fathy et al., "Bi-directional ConvLSTM U-Net with Densely Connected Convolutions," *ICC*