Polyp Segmentation Using a Hybrid Architecture Based on UNet++ and Swin Transformer

Çağatay Berke Erdaş
Department of Computer Engineering, Faculty of Engineering, Başkent University, Ankara, Turkey

Abstract

The accurate and reliable segmentation of colorectal polyps is of critical importance for the early diagnosis of colorectal cancer, reducing false negative rates during screening, and improving the effectiveness of treatment planning. While deep learning-based methods have made significant advances in this field, the size, color, and morphological diversity of polyps limit the performance of current approaches. This study proposes a novel hybrid approach for polyp segmentation on the widely used Kvasir-SEG dataset, integrating the multi-layered and densely connected architecture of UNet++ with the attention mechanism based on Swin Transformer. The combination of UNet++'s powerful multi-scale feature extraction capability and Swin Transformer's global context capture ability aims to improve the segmentation accuracy of both small and ambiguous polyps. Experimental findings reveal that the proposed hybrid architecture demonstrates superior performance when compared to leading methods in the literature, such as ResUNet, A-DenseUNet, Dilated-UNet-Seg, and EffiSegNet-B4. Notably, the improvements observed in the Dice coefficient and Intersection over Union (IoU) metrics, when evaluated alongside enhancements in boundary delineation, demonstrate that the proposed model possesses high generalization capacity even on small datasets. The results obtained indicate that hybrid CNN-Transformer architectures can be a powerful alternative not only for colorectal polyp segmentation but also for other medical imaging applications. In this respect, the study makes a significant contribution compared to existing approaches in the literature and presents a method that could form the basis for clinical decision support systems to be developed in the future.

Keywords

Polyp Segmentation, UNet++, Swin Transformer, Kvasir-SEG, Medical Image Analysis.