

RAPUMobileNet: A LIGHTWEIGHT HYBRID MODEL FOR REAL-TIME POLYP SEGMENTATION

Nguyễn Kiều Vinh - 230101030

Tóm tắt

- Lớp: CS2205.CH190
- Link Github của nhóm: <https://github.com/kieuvinhhh/CS2205.CH190>
- Link YouTube video:
- Thành viên: Nguyễn Kiều Vinh
-



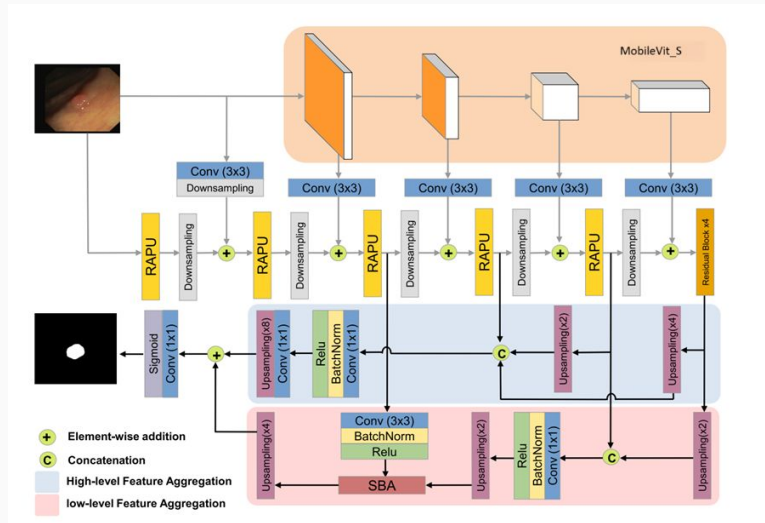
Giới thiệu

- Ung thư đại trực tràng là một trong những nguyên nhân gây tử vong hàng đầu trên toàn thế giới, và polyp được xem là những tổn thương tiền ung thư có thể phát hiện và loại bỏ thông qua nội soi. Tuy nhiên, việc phát hiện và phân đoạn chính xác polyp từ ảnh nội soi vẫn là một thách thức lớn do sự đa dạng về hình dạng, kích thước, kết cấu và mức độ tương phản thấp với nền xung quanh.
- Những tiến bộ trong học sâu, đặc biệt là các mạng CNN, đã đạt được nhiều kết quả tích cực trong phân đoạn ảnh y tế. Tuy nhiên, hầu hết các mô hình CNN hiện tại chủ yếu dựa trên đặc trưng cục bộ và thường bỏ sót mối quan hệ không gian toàn cục giữa các vùng ảnh – yếu tố quan trọng trong việc xác định chính xác ranh giới polyp.
- Nghiên cứu này đề xuất một mô hình hybrid mới kết hợp giữa kiến trúc MobileVit và mạng CNN lightweight nhằm khai thác cả thông tin toàn cục và chi tiết cục bộ. Mô hình được thiết kế theo dạng encoder-decoder, sử dụng MobileVit để trích xuất đặc trưng toàn cục và CNN để tái tạo bản đồ phân đoạn với độ chính xác cao. Nghiên cứu này cũng tích hợp cơ chế Channel Attention giúp tăng cường khả năng tập trung vào các vùng quan trọng, đặc biệt hữu ích trong các trường hợp polyp có biên mờ hoặc nhỏ.

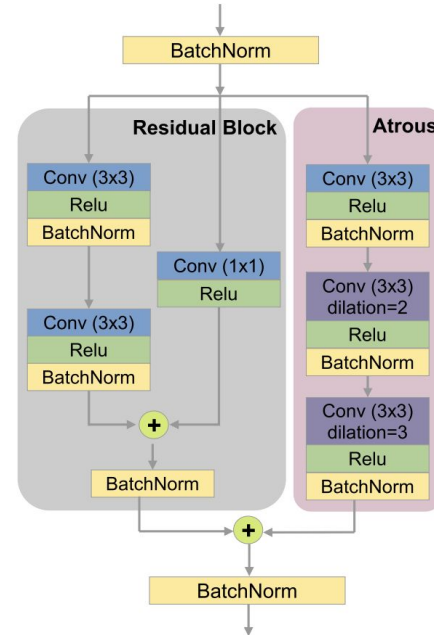
Mục tiêu

- Có thể giảm FLOPs để triển khai được trên các thiết bị di động với độ trễ thấp
- Giữ local-global context nhưng tối ưu hóa kernel
- Giữ được độ chính xác cao

Nội dung và Phương pháp



Hình 1: Tổng quan kiến trúc RAPUNet



Hình 2: Kiến trúc của khối RAPU

Kết quả dự kiến

Triển khai được trên các thiết bị di động với độ trễ thấp

Giữ được độ chính xác cao

Tài liệu tham khảo

- [1]. Hyunnam Lee, Juhan Yoo: MetaFormer and CNN Hybrid Model for Polyp Image Segmentation. IEEE 2024-08-24
- [2]. Sachin Mehta, Mohammad Rastegari: MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer. ICLR 2022