

BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học

CS2205 - PHƯƠNG PHÁP LUẬN

NGHIÊN CỨU KHOA HỌC

Lớp học

CS2205.APR2023

Giảng viên

PGS.TS. LÊ ĐÌNH DUY

Thời gian

04/2023 - 06/2023

THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo:
<https://youtu.be/6NppQDukRt0>
- Link poster:
<https://github.com/daothuyit/CS2205.APR2023/blob/main/Poster.pdf>
- Link slides:
<https://github.com/daothuyit/CS2205.APR2023/blob/main/Presentation.pdf>

- Họ và Tên: Đào Thị Thu Thủy
- MSHV: 220104015



- Lớp: CS2205.APR2023
- Tự đánh giá (điểm tổng kết môn): 9/10
- Số buổi vắng: 0
- Link Github:
<https://github.com/daothuyit/CS2205.APR2023>

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI

ỨNG DỤNG PHÁT HIỆN TỔN THƯƠNG TRÊN DA QUA HÌNH ẢNH SỬ DỤNG MÔ HÌNH SEGFORMER

TÊN ĐỀ TÀI TIẾNG ANH

APPLICATION OF SKIN LESION DETECTION THROUGH IMAGE USING SEGFORMER MODEL

TÓM TẮT

Đề tài tận dụng tiềm năng của **SegFormer** trong lĩnh vực phân đoạn ảnh (Semantic Segmentation) để phát hiện và phân vùng tổn thương trên da qua hình ảnh. Kết quả dự kiến là tạo ra một mô hình SegFormer tối ưu, cùng với một ứng dụng web tiện ích để tải lên ảnh chụp da và nhận kết quả phân vùng tổn thương. Đề tài này mong đóng góp vào việc phát triển công nghệ ứng dụng trong lĩnh vực y tế, chăm sóc sức khỏe, hỗ trợ chẩn đoán và điều trị các vấn đề về da một cách hiệu quả.

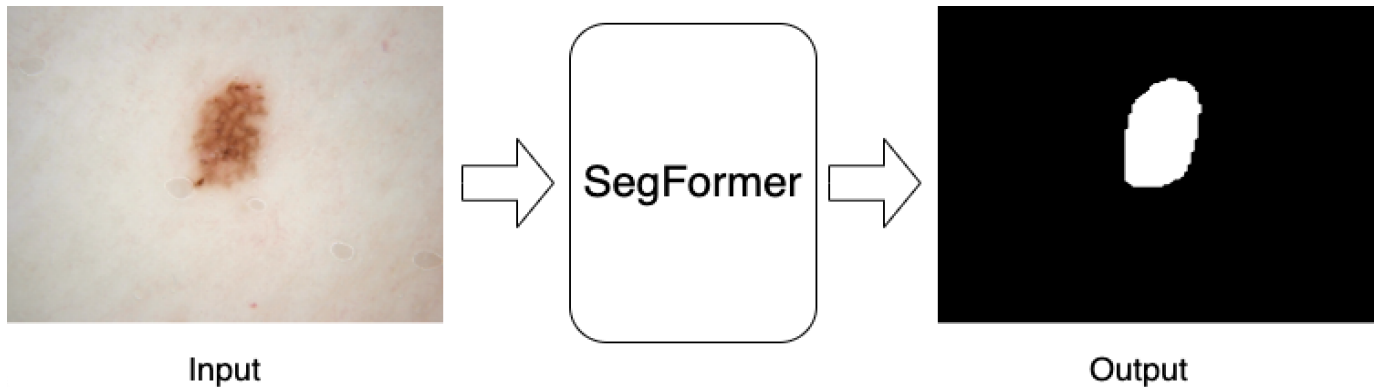
GIỚI THIỆU

Hiện nay, chẩn đoán bệnh da liễu chủ yếu dựa vào sự quan sát trực tiếp của các chuyên gia có thể dẫn đến sai sót và mất thời gian. Trong trường hợp các tổn thương diện rộng, phức tạp hoặc nằm ở các vị trí khó tiếp cận, đánh giá trực quan trở nên khó khăn hơn. Từ đó, dẫn đến sự thiếu chính xác trong chẩn đoán và điều trị các vấn đề về da.

Với sự phát triển của các mô hình học sâu (Deep Learning) và công nghệ thị giác máy tính (Computer vision), các mô hình phân đoạn ngữ nghĩa (Semantic Segmentation) đã cho thấy kết quả tốt trong việc phát hiện và xác định chính xác vùng tổn thương trên da. Trong đề tài này, chúng tôi đề xuất áp dụng mô hình **SegFormer** để xây dựng ứng dụng phát hiện và phân vùng tổn thương trên da qua ảnh. Cụ thể:

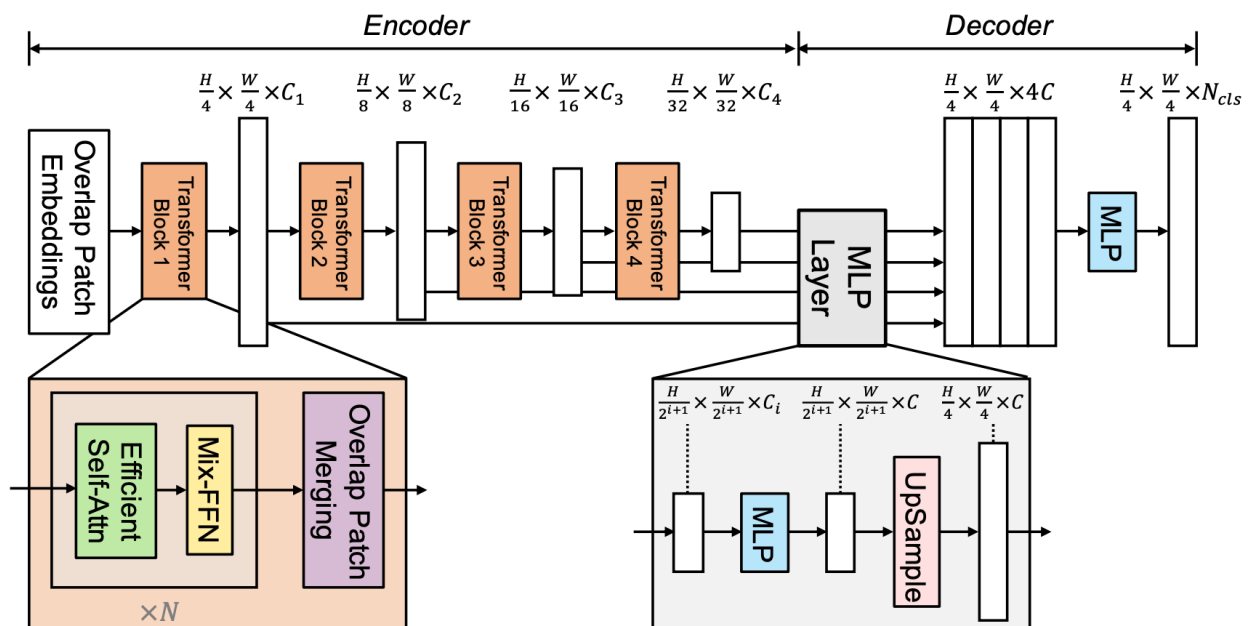
Input: Ảnh chụp da.

Output: Phân vùng vùng da bị tổn thương.



SegFormer:

- Là mô hình Transformer chuyên dụng cho bài toán Phân đoạn ngữ nghĩa (Semantic Segmentation).
- Đạt hiệu năng cao
- Sức mạnh đến chủ yếu từ backbone Mix Vision Transfomer (MiT)



Bộ dữ liệu: ISIC 2018, được Tổ chức Hợp tác Hình ảnh Da Quốc tế (ISIC) xuất bản dưới dạng bộ dữ liệu quy mô lớn về hình ảnh soi da, gồm 2594 hình ảnh đã tiền xử lý để phân đoạn tổn thương da (lesion segmentation)

MỤC TIÊU

- Nghiên cứu kiến trúc và cách hoạt động của mô hình SegFormer, hiểu được tại sao SegFormer là mô hình đơn giản, hiệu quả và đạt độ chính xác cao cho bài toán Phân đoạn ngữ nghĩa (Semantic Segmentation). Khai thác tiềm năng của mô hình SegFormer để nâng cao hiệu suất và độ chính xác trong việc phát hiện tổn thương trên da thông qua hình ảnh.
- Phát triển ứng dụng web để phát hiện và phân vùng tổn thương da qua ảnh từ mô hình SegFormer tối ưu đã tạo.

NỘI DUNG VÀ PHƯƠNG PHÁP

- Tìm hiểu kiến trúc và cách hoạt động của mô hình SegFormer qua bài báo SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers
- Thu thập và tiền xử lý dữ liệu: Bộ dữ liệu ISIC 2018, bao gồm 2594 hình ảnh, là các hình ảnh chứa các trường hợp tổn thương da đã được chẩn đoán. Các hình ảnh được chia thành 3 phần test, train, val, mỗi phần có các hình ảnh soi da và segmentation mask của mỗi hình.
- Xây dựng và huấn luyện mô hình SegFormer. Đánh giá và tối ưu hóa các tham số của mô hình để đạt được hiệu suất tốt nhất.
- Phát triển ứng dụng web cung cấp giao diện thân thiện và dễ sử dụng, giúp người dùng thuận tiện trong việc tải ảnh lên và nhận về thông tin vị trí, phân vùng tổn thương.

KẾT QUẢ MONG ĐỢI

- Từ các kết quả thực nghiệm, tạo được mô hình SegFormer tối ưu, đạt hiệu suất và độ chính xác cao trong việc phát hiện và phân vùng các tổn thương trên da.
- Hoàn thành ứng dụng web để người dùng thể tải lên ảnh chụp da và nhận được kết quả phân vùng tổn thương từ mô hình SegFormer.

TÀI LIỆU THAM KHẢO

[1]. Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, José M. Álvarez, Ping Luo:

SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers.

NeurIPS2021: 12077-12090

[2]. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin:

Attention is All you Need. NIPS 2017: 5998-6008

[3]. Salman H. Khan, Muzammal Naseer, Munawar Hayat, Syed Waqas Zamir, Fahad Shahbaz Khan, Mubarak Shah:

Transformers in Vision: A Survey. CoRR abs/2101.01169 (2021)