

THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):
<https://www.youtube.com/watch?v=uDZOCa-0d2Q>
- Link slides (dạng .pdf đặt trên Github của nhóm):
<https://github.com/minhtritri/CS2205.CH183/blob/main/Tr%C3%ADC%20V%C3%B5%20Minh%20-%20CS2205.NOV2024.DeCuong.FinalReport.Template.Slide.pdf>
- *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
- *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*
- *Lớp Cao học, mỗi nhóm một thành viên*

- Họ và Tên: Võ Minh Trí
- MSSV: 240101028



- Lớp: CS2205.CH183
- Tự đánh giá (điểm tổng kết môn): 8/10
- Số buổi vắng: 1
- Số câu hỏi QT cá nhân: 4
- Số câu hỏi QT của cả nhóm: 4
- Link Github:
<https://github.com/minhtritri/CS2205.CH183>

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

CẢI THIẾN SCUNET++ TRONG PHÂN ĐOẠN HÌNH ẢNH CT ĐỀ HỖ TRỢ
CHẨN ĐOÁN THUYỀN TẮC PHỔI

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

ENHANCING SCUNET++ FOR PULMONARY EMBOLISM CT IMAGE
SEGMENTATION

TÓM TẮT *(Tối đa 400 từ)*

Thuyên tắc phổi là một bệnh lý nghiêm trọng, hoàn toàn có thể gây nguy hiểm đến tính mạng nếu không được phát hiện và điều trị kịp thời. Chẩn đoán dựa trên hình ảnh cắt lớp vi tính động mạch phổi (Computed Tomography Pulmonary Angiography - CTPA) đòi hỏi các phương pháp phân đoạn chính xác để xác định vùng tổn thương, song vẫn tồn tại nhiều thách thức do sự phức tạp về hình thái và chất lượng hình ảnh. Nghiên cứu này tập trung vào việc cải tiến việc sử dụng mô hình Deep Learning như SCUNet++ vào hành động phân đoạn hình ảnh để tăng hiệu suất phân đoạn hình ảnh CT trong chẩn đoán thuyên tắc phổi.

Với mục đích tối ưu hóa SCUNet++ bằng cách cải thiện kiến trúc mạng, tối ưu hóa quá trình huấn luyện nhằm nâng cao độ chính xác và tốc độ suy luận. Các thực nghiệm sẽ được thực hiện trên các tập dữ liệu được công khai như FUMPE và CAD-PE, với các tiêu chí đánh giá bao gồm hệ số DICE, IoU và khoảng cách Hausdorff. Hiệu suất của mô hình cải tiến sẽ được so sánh với SCUNet++ gốc cũng như các phương pháp phổ biến như UNet và Swin-UNet.

Dự kiến nghiên cứu này sẽ chứng minh rằng mô hình SCUNet++ cải tiến không chỉ nâng cao độ chính xác của phân đoạn hình ảnh mà còn rút ngắn thời gian suy luận, hỗ trợ các bác sĩ chẩn đoán nhanh và hiệu quả hơn. Những đóng góp từ nghiên cứu này có thể mở ra hướng ứng dụng mới cho trí tuệ nhân tạo trong lĩnh vực y khoa, cải thiện chất lượng chăm sóc sức khỏe cho bệnh nhân.

GIỚI THIỆU *(Tối đa 1 trang A4)*

Thuyên tắc phổi (Pulmonary Embolism - PE) là một bệnh lý nguy hiểm, xảy ra khi xuất hiện cục máu đông làm tắc nghẽn động mạch phổi. Việc chẩn đoán PE dựa vào chụp cắt lớp vi tính động mạch phổi (Computed Tomography Pulmonary Angiography - CTPA), giúp xác định chính xác vị trí và kích thước của cục máu đông gây tắc. Tuy nhiên việc phân tích hình ảnh CTPA thủ công không chỉ tốn nhiều thời gian mà còn phụ thuộc vào trình độ của bác sĩ chẩn đoán hình ảnh, dễ dẫn đến sai sót. Hiện nay, các mô hình Deep Learning dựa trên Transformer và CNN, như SCUNet++, đã đạt được những kết quả đáng kể trong việc phân đoạn hình ảnh CT chẩn đoán thuyên tắc phổi, giúp xác định các cục máu đông với độ chính xác cao hơn. Tuy nhiên, mô hình SCUNet++ hiện tại vẫn còn một số hạn chế, đặc biệt trong khả năng trích xuất đặc trưng không gian cục bộ ở encoder và khả năng tái tạo chi tiết ở decoder, dẫn đến hiệu suất phân đoạn chưa tối ưu.

Nghiên cứu này hướng đến việc cải thiện kiến trúc mô hình SCUNet++, tập trung vào Encoder và Decoder nhằm nâng cao chất lượng phân đoạn ảnh CTPA.

Bài toán phân đoạn hình ảnh chụp cắt lớp vi tính động mạch phổi của bệnh nhân được định nghĩa:

- Đầu vào: Một ảnh chụp cắt lớp vi tính động mạch phổi CTPA của bệnh nhân, được biểu diễn dưới dạng ảnh 2D. Ảnh có thể chứa hoặc không chứa vùng thuyên tắc phổi (PE lesions).
- Đầu ra: Một bản đồ phân đoạn (segmentation mask) có cùng kích thước với ảnh đầu vào, trong đó:
 - Vùng chứa cục máu đông (PE) được đánh dấu bằng một nhãn cụ thể (màu đỏ hoặc số 1).
 - Các vùng không bị ảnh hưởng (phổi bình thường, mô xung quanh) được đánh dấu là nền (màu đen hoặc 0).

MỤC TIÊU *(Viết trong vòng 3 mục tiêu)*

- Cải thiện Encoder của SCUNet++ bằng kết hợp CNN trước Swin Transformer,

giúp tăng cường khả năng trích xuất đặc trưng không gian cục bộ và giảm thiểu mất mát thông tin trong quá trình mã hóa.

- Tối ưu hóa Decoder bằng cách sử dụng Attention-based Skip Connection và Feature Fusion Modules, giúp mô hình tái tạo đối tượng chính xác hơn.
- Đánh giá và so sánh hiệu suất của mô hình cải tiến với SCUNet++ gốc và các mô hình phân đoạn phổ biến khác (UNet, Swin-UNet), sử dụng các chỉ số đánh giá như Dice Score, IoU,...

NỘI DUNG VÀ PHƯƠNG PHÁP

Nội dung 1: Thu thập và tiền xử lý dữ liệu

- Mục tiêu: Lựa chọn tập dữ liệu phù hợp với bài toán phân đoạn hình ảnh CTPA trong chẩn đoán thuyên tắc phổi. Tiền xử lý dữ liệu để đảm bảo tính nhất quán.
- Phương pháp: Sử dụng tập dữ liệu công khai như FUMPE và CAD-PE, gồm ảnh CTPA có nhãn về vùng thuyên tắc phổi. Chia tập dữ liệu thành tập huấn luyện (training - 90%) và tập đánh giá (test - 10%).

Nội dung 2: Cải tiến Encoder của SCUNet++

- Mục tiêu:
 - Giúp mô hình nhận diện chính xác hơn các vùng tổn thương nhỏ trong ảnh CTPA.
 - Giảm thiểu mất mát thông tin do self-attention trong Swin Transformer.
- Phương pháp:
 - Thêm CNN trước Swin Transformer Blocks: Sử dụng các lớp 3x3 Convolution trước mỗi Transformer Block để giúp trích xuất đặc trưng cục bộ tốt hơn.
 - Thử nghiệm với các biến thể Encoder khác:
 - ConvNeXt-based Encoder: Tận dụng các tiến bộ của CNN hiện đại để cải thiện hiệu suất.
 - Focal Transformer: Tăng cường khả năng tổng quát hóa bằng cách tập trung vào các vùng quan trọng.

Nội dung 3: Cải tiến Decoder của SCUNet++

- Mục tiêu:
 - Nâng cao khả năng phục hồi biên của cục máu đông.
 - Giữ lại chi tiết nhỏ trong ảnh, tránh mất mát thông tin quan trọng.
 - Tăng cường khả năng tập trung vào vùng tổn thương thông qua Attention.
- Phương pháp:
 - Tích hợp Attention-based Skip Connections: Thay vì kết nối trực tiếp giữa Encoder và Decoder, áp dụng Self-Attention để giúp mô hình tập trung vào các vùng quan trọng hơn.
 - Bổ sung Feature Fusion Modules: Kết hợp thông tin từ nhiều tầng của Encoder để tăng cường chất lượng phân đoạn.

Nội dung 4: Huấn luyện mô hình

- Mục tiêu:
 - Tối ưu hóa mô hình để đạt được độ chính xác phân đoạn cao nhất.
 - Giảm overfitting và đảm bảo mô hình có khả năng tổng quát hóa tốt.
- Phương pháp:
 - Hàm mất mát: Sử dụng Dice Loss và Focal Loss để tối ưu hóa mô hình cho bài toán phân đoạn y khoa.
 - Tối ưu hóa (Optimization): Dùng AdamW Optimizer kết hợp với Cosine Annealing Scheduler để điều chỉnh learning rate hiệu quả hơn.
 - Áp dụng Batch Normalization và Dropout để giảm overfitting.
 - Huấn luyện trên phần cứng: Sử dụng GPU A100 để tăng tốc độ huấn luyện.

Nội dung 5: Đánh giá hiệu suất mô hình

- Mục tiêu:
 - Đánh giá độ chính xác phân đoạn.
 - Kiểm tra tốc độ suy luận và khả năng triển khai thực tế.
 - So sánh mô hình cải tiến với các mô hình hiện có.

- Phương pháp:
 - Đo các chỉ số đánh giá: Dice Score: Đo mức độ trùng khớp giữa vùng dự đoán và vùng thực tế. IoU (Intersection over Union): Đánh giá mức độ giao nhau giữa vùng dự đoán và vùng thực tế. Hausdorff Distance: Kiểm tra độ chính xác của mô hình trong việc nhận diện biên của cục máu đông.
 - Đánh giá tốc độ suy luận: Inference Latency (ms/image): Thời gian mô hình xử lý một ảnh đầu vào. Throughput (ảnh/s): Số lượng ảnh có thể được xử lý mỗi giây.
 - So sánh với SCUNet++ gốc, U-Net và Swin-UNet để kiểm tra hiệu suất của mô hình cải tiến.

KẾT QUẢ MONG ĐỢI

Cải thiện độ chính xác phân đoạn. Mô hình SCUNet++ cải tiến dự kiến sẽ đạt Dice Score $\geq 85\%$ và IoU $\geq 80\%$, giúp xác định vùng thuyên tắc phổi chính xác hơn.

Tối ưu tốc độ suy luận. Kỳ vọng Inference Latency $\leq 200\text{ms/image}$ và Throughput $\geq 5\text{ images/second}$, giúp hỗ trợ chẩn đoán nhanh hơn.

So sánh vượt trội với các mô hình hiện có. Mô hình cải tiến được kỳ vọng sẽ có độ chính xác cao hơn SCUNet++ gốc, Swin-UNet và U-Net.

Khả năng ứng dụng thực tế. Nếu kết quả đạt mong đợi, mô hình có thể triển khai thử nghiệm trên dữ liệu bệnh viện, hỗ trợ bác sĩ trong chẩn đoán thuyên tắc phổi

TÀI LIỆU THAM KHẢO (Định dạng DBLP)

[1]. Yifei Chen, Binfeng Zou, Zhaoxin Guo, Yiyu Huang, Yifan Huang, Feiwei Qin, Qin hai Li, Changmiao Wang:

SCUNet++: Swin-UNet and CNN Bottleneck Hybrid Architecture with Multi-Fusion Dense Skip Connection for Pulmonary Embolism CT Image Segmentation.

WACV 2024: 7759-7767.

[2]. Jianwei Yang, Chunyuan Li, Pengchuan Zhang, Xiyang Dai, Bin Xiao, Lu Yuan, and Jianfeng Gao:

Focal self attention for local-global interactions in vision transformers.

arXiv preprint arXiv:2107.00641, 2021.

[3]. K. O'Shea, R. Nash:

An introduction to convolutional neural networks

arXiv preprint arXiv:1511.08458 (2015)

[4]. Shikhar Mohan, Saumik Bhattacharya, Sayantari Ghosh:

Attention W-Net: Improved Skip Connections for better Representations

arXiv preprint arXiv:2110.08811 (2021)

[5] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie:

A ConvNet for the 2020s

arXiv preprint arXiv:2201.03545 (2022)