

THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo (tối đa 5 phút):
<https://youtu.be/9jFVCQE9L2Y>
- Link slides (dạng .pdf đặt trên Github của nhóm):
<https://github.com/tricaodev/CS2205.CH201/blob/main/Tr%C3%ADC%20Cao%200%C4%90%E1%BB%A9c%20-%20CS2205.SEP2025.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: Cao Đức Trí
- MSSV: 250101069



- Lớp: CS2205.CH201
- Tự đánh giá (điểm tổng kết môn): 8/10
- Số buổi vắng: 1
- Số câu hỏi QT cá nhân: 3
- Số câu hỏi QT của cả nhóm: 15
- Link Github:
<https://github.com/tricaodev/CS2205.CH201>

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

ỨNG DỤNG THỊ GIÁC MÁY TÍNH ĐỂ TỰ ĐỘNG PHÁT HIỆN BẤT THƯỜNG TRÊN PHIM X-QUANG NGỰC

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

USING COMPUTER VISION TECHNIQUES TO AUTOMATICALLY DETECT ABNORMALITIES IN CHEST X-RAYS

TÓM TẮT *(Tối đa 400 từ)*

X-quang ngực (Chest X-ray – CXR) là xét nghiệm hình ảnh phổ biến, chi phí thấp và được dùng rộng rãi để sàng lọc nhiều bệnh lý (viêm phổi, tràn dịch màng phổi, xẹp phổi, tim to, tràn khí màng phổi...). Tuy nhiên, khối lượng CXR lớn, chất lượng ảnh không đồng đều và phụ thuộc kinh nghiệm bác sĩ có thể gây chậm trễ trong phát hiện bất thường hoặc bỏ sót ca nặng. Đề tài này nghiên cứu ứng dụng kỹ thuật thị giác máy tính để tự động phát hiện bất thường trên CXR nhằm hỗ trợ sàng lọc và ưu tiên đọc phim (triage).

Hướng tiếp cận đề xuất xây dựng mô hình học sâu cho bài toán phân loại đa nhãn (multi-label classification) trên CXR: đầu vào là ảnh X-quang, đầu ra là xác suất xuất hiện của từng loại bất thường (hoặc nhãn “abnormal/normal” ở mức tổng quát). Quy trình gồm: tiền xử lý ảnh (chuẩn hoá kích thước/cường độ), tăng cường dữ liệu, huấn luyện mạng CNN hiện đại (DenseNet/EfficientNet hoặc Vision Transformer) với hàm mất mát phù hợp mất cân bằng lớp, và đánh giá bằng AUROC, sensitivity/specificity trên tập kiểm thử. Để tăng tính tin cậy khi ứng dụng, nghiên cứu bổ sung mô-đun giải thích (explainability) bằng heatmap (Grad-CAM) nhằm chỉ ra vùng ảnh đóng góp vào dự đoán, đồng thời phân tích lỗi thường gặp.

Kết quả kỳ vọng là một mô hình đạt hiệu năng cạnh tranh so với baseline, kèm trực quan hoá vùng nghi ngờ giúp người dùng kiểm tra nhanh. Sản phẩm demo có thể nhận ảnh CXR và trả về danh sách bất thường có khả năng cao + mức độ tự tin + heatmap, từ đó hỗ trợ bác sĩ/nhân viên y tế trong quy trình sàng lọc ban đầu.

GIỚI THIỆU *(Tối đa 1 trang A4)*

X-quang ngực là phương tiện chẩn đoán hình ảnh tuyến đầu trong nhiều tình huống lâm sàng nhờ tốc độ nhanh và chi phí thấp. Trong thực tế bệnh viện, số lượng phim CXR mỗi ngày rất lớn, trong khi thời gian đọc phim có hạn. Ngoài ra, ảnh X-quang có thể nhiễu, khác biệt thiết bị chụp, tư thế bệnh nhân, hoặc chất lượng phơi sáng, khiến việc đọc phim khó khăn hơn. Những yếu tố này làm tăng nguy cơ bỏ sót tổn thương sớm hoặc chậm phát hiện các tình trạng cấp cứu như tràn khí màng phổi.

Sự phát triển của học sâu và thị giác máy tính mở ra khả năng tự động phân tích CXR. Nếu xây dựng được mô hình phát hiện bất thường đủ chính xác, hệ thống có thể đóng vai trò “trợ lý sàng lọc”: gợi ý nhãn bất thường, ưu tiên các ca nghiêm trọng để bác sĩ đọc trước, hoặc hỗ trợ đào tạo/kiểm soát chất lượng. Tuy nhiên, bài toán CXR có nhiều thách thức: nhãn thường là đa nhãn (một ảnh có thể có nhiều bất thường), dữ liệu mất cân bằng (một số bệnh hiếm), nhãn có thể nhiễu (gán nhãn từ báo cáo), và yêu cầu giải thích cao khi ứng dụng y tế.

Do đó, đề tài tập trung vào: (1) lựa chọn mô hình phù hợp cho phân loại đa nhãn CXR; (2) thiết kế quy trình huấn luyện/đánh giá chặt chẽ; (3) bổ sung giải thích bằng heatmap để tăng khả năng kiểm chứng; và (4) xây dựng demo minh họa quy trình hỗ trợ sàng lọc. Kết quả nghiên cứu kỳ vọng đóng góp một pipeline có thể tái sử dụng cho các bài toán tương tự trong phân tích ảnh y khoa.

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

1. **Xây dựng mô hình thị giác máy tính** phát hiện bất thường trên X-quang ngực theo bài toán phân loại nhị phân (normal/abnormal) và/hoặc phân loại đa nhãn theo nhóm bệnh phổ biến.
2. **Thiết kế và đánh giá thực nghiệm** với quy trình tiền xử lý, xử lý mất cân bằng dữ liệu, và thước đo phù hợp (AUROC, sensitivity, specificity, F1).
3. **Tăng tính tin cậy và khả năng sử dụng** bằng mô-đun giải thích (Grad-CAM/heatmap) và phân tích lỗi, hướng tới demo hỗ trợ sàng lọc.

NỘI DUNG VÀ PHƯƠNG PHÁP

1) Dữ liệu và nhãn

- **Nguồn dữ liệu (tuỳ điều kiện nhóm):** ChestX-ray14 / CheXpert / MIMIC-CXR (hoặc dữ liệu nội bộ nếu được phép).
- **Nhãn:**
 - Cách A: nhị phân *Normal vs Abnormal* (dễ triển khai, phù hợp demo).
 - Cách B: đa nhãn theo các bất thường phổ biến (ví dụ: Pneumonia, Effusion, Atelectasis, Cardiomegaly, Pneumothorax...).
- **Chia tập:** train/val/test theo bệnh nhân để tránh rò rỉ (leakage).

2) Tiền xử lý & tăng cường dữ liệu

- Chuẩn hoá kích thước (ví dụ 224×224 hoặc 512×512), chuẩn hoá cường độ pixel.
- Augmentation: xoay nhẹ, dịch chuyển, thay đổi độ sáng/contrast (hạn chế biến đổi làm sai cấu trúc y khoa).
- Xử lý mất cân bằng: class weights, focal loss, hoặc oversampling.

3) Mô hình đề xuất (baseline → nâng cấp)

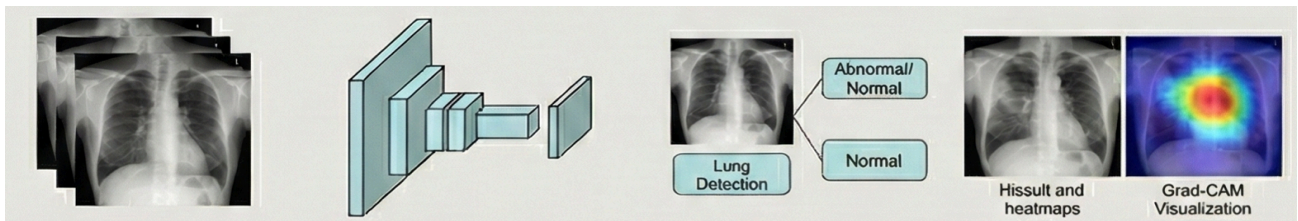
- **Baseline:** CNN đơn giản hoặc ResNet-18/34 huấn luyện từ pretrained ImageNet.
- **Mô hình chính:** DenseNet-121 / EfficientNet-B0/B2 (mạnh cho ảnh y khoa), hoặc Vision Transformer (nếu đủ tài nguyên).
- **Thiết lập bài toán:**
 - Multi-label: sigmoid từng lớp + Binary Cross-Entropy / Focal Loss.
 - Binary: sigmoid + BCE.

4) Đánh giá

- Chỉ số: **AUROC** theo từng nhãn và trung bình; **Sensitivity/Specificity** ở ngưỡng chọn; F1-score (nếu phù hợp).
- Phân tích lỗi: nhầm do chất lượng ảnh, bệnh hiếm, nhãn nhiễu, vùng tổn thương nhỏ.
- **Giải thích:** Grad-CAM tạo heatmap cho nhãn dự đoán để xem mô hình “nhìn” vào vùng hợp lý hay không.

5) Demo sản phẩm

- Input: 1 ảnh CXR.
- Output:
 - Danh sách nhãn bất thường có xác suất cao
 - Nhãn tổng quát normal/abnormal
 - Heatmap vùng nghi ngờ (Grad-CAM)
- Mục tiêu demo: hỗ trợ sàng lọc/ưu tiên đọc phim, không thay thế chẩn đoán bác sĩ.



KẾT QUẢ MONG ĐỢI

- Một pipeline end-to-end cho phát hiện bất thường CXR (tiền xử lý → mô hình → đánh giá).
- Mô hình đạt hiệu năng tốt hơn baseline (AUROC cải thiện rõ rệt) và ổn định trên tập kiểm thử.
- Heatmap giải thích trực quan giúp người dùng kiểm chứng.
- Demo chạy được: nhập ảnh → xuất kết quả dự đoán + heatmap.

TÀI LIỆU THAM KHẢO (Định dạng DBLP)

- [1]. Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, Andrew Y. Ng:
CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.
CoRR abs/1711.05225, 2017.
- [2]. Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, Ronald M. Summers:
ChestX-ray8: Hospital-Scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases.
CVPR 2017: 3462-3471.

[3]. Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silvana Ciurea-Ilcus, Chris Chute, Henrik Marklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, Jayne Seekins, David A. Mong, Safwan S. Halabi, Jesse K. Sandberg, Ricky Jones, David B. Larson, Curtis P. Langlotz, Bhavik N. Patel, Matthew P. Lungren, Andrew Y. Ng:

CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison.

AAAI 2019: 590-597.

[4]. Alistair E. W. Johnson, Tom J. Pollard, Seth Berkowitz, Nathaniel R. Greenbaum, Matthew P. Lungren, Chih-Ying Deng, Roger G. Mark, Steven Horng:

MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports.

Scientific Data 6: 317, 2019. *(Nếu cần đúng DBLP thuần hội nghị/journal CS, có thể thay bằng bản arXiv/CoRR của nhóm tác giả.)*

[5]. Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger: Densely Connected Convolutional Networks.

CVPR 2017: 2261-2269.

[6]. Mingxing Tan, Quoc V. Le:

EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.

ICML 2019: 6105-6114.

[7]. Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra:

Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization.

ICCV 2017: 618-626.