


THÔNG TIN CHUNG CỦA BÁO CÁO

- Link YouTube video của báo cáo (tối đa 5 phút):
(<https://www.youtube.com/watch?v=AWq7uw-36Ng>)
- Link slides (dạng .pdf đặt trên Github):
(<https://github.com/mynameuit/CS2205.APR2023/TenDeTai.pdf>)
- Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in

<ul style="list-style-type: none">• Họ và Tên: Lê Võ Đình Kha• MSSV: 18520872 	<ul style="list-style-type: none">• Lớp: CS2205.CH1702• Tự đánh giá (điểm tổng kết môn): 9/10• Số buổi vắng: 0• Link Github:
---	---

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

ỨNG DỤNG HỖ TRỢ CHUẨN ĐOÁN SỰ BẤT THƯỜNG TỪ ẢNH X - QUANG CỘT SỐNG

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

APPLICATION FOR DETECTING SPINAL ABNORMALITIES FROM SPINE X-RAY FILMS

TÓM TẮT *(Tối đa 400 từ)*

Trí tuệ nhân tạo (AI) đang ngày càng được áp dụng rộng rãi trong cả nghiên cứu cơ bản và ứng dụng, sự phát triển của AI trong lĩnh vực nghiên cứu cơ bản đã mở ra nhiều cánh cửa cho nhiều lĩnh vực khác nhau, trong đó lĩnh vực y tế được đánh giá có tiềm năng hứa hẹn.

Đề tài tập trung vào xây dựng một ứng dụng có khả năng phát hiện các bất thường trên ảnh X-quang cột sống. Với đề tài này, ba mô hình phát hiện đối tượng, bao gồm Faster RCNN, RetinaNet và FCOS, được lựa chọn triển khai để đánh giá khả năng và giới hạn của từng phương pháp. Mục tiêu là tìm hiểu sâu về các mô hình này, so sánh đánh giá hiệu suất của chúng trên tập dữ liệu được sử dụng và xác định phương pháp phù hợp nhất để chẩn đoán các bất thường trên ảnh X-quang cột sống.

Từ phương pháp được lựa chọn, mô hình sinh ra có thể được ứng dụng lâm sàng trong thực tế để hỗ trợ các bác sĩ trong quá trình chẩn đoán, đưa ra các liệu pháp chữa trị kịp thời và hiệu quả cho bệnh nhân điều trị.

GIỚI THIỆU

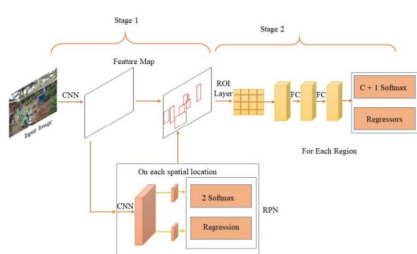
Sự bùng nổ của deep learning đã mang lại một cuộc cách mạng trong việc áp dụng vào bài toán xử lý ảnh, đặc biệt là với mô hình CNN. Khác với các phương pháp truyền thống, CNN có thể không cần tạo đặc trưng thủ công và tự động học các đặc trưng từ dữ liệu. Điều này giúp tiết kiệm thời gian xử lý ảnh và mang lại kết quả tốt hơn.

Trong lĩnh vực y tế, việc áp dụng các mô hình học sâu cũng trở thành xu hướng để giải quyết các bài toán khác nhau trong đó có các bài toán xử lý ảnh y khoa. Trong đề tài này sẽ tập trung vào xây dựng một ứng dụng phát hiện bất thường trên ảnh X-quang cột sống. Mục tiêu là tự động phát hiện các bất thường và hỗ trợ các bác sĩ trong quá trình chẩn đoán, đưa ra kết quả chính xác.

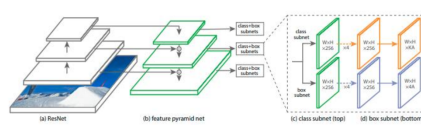
Mô hình phát hiện đối tượng trong thị giác máy tính là các hệ thống được thiết kế để nhận diện và xác định vị trí của các đối tượng trong ảnh. Có hai loại chính của mô hình phát hiện đối tượng là

mô hình phát hiện đối tượng một giai đoạn (one-stage object detection model) và mô hình phát hiện đối tượng hai giai đoạn (two-stage object detection model).

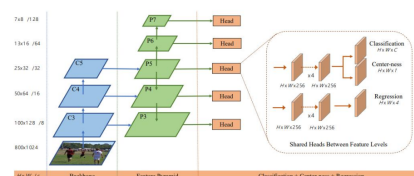
Các mô hình như Faster RCNN, RetinaNet và FCOS thường được sử dụng để phát hiện đối tượng bất thường trên ảnh X-quang. Faster RCNN và Retina đã được áp dụng trên bộ dữ liệu được lựa chọn [9] với mAP@0.5 31.38 cho Faster RCNN và 28.09 cho RetinaNet. Dựa trên việc khảo sát các nghiên cứu đã có, đề tài sẽ áp dụng hai mô hình một giai đoạn là RetinaNet, FCOS với mạng trích xuất đặc trưng là ResNet50 và mô hình hai giai đoạn áp dụng là Faster RCNN với các mạng trích xuất đặc trưng là VGG16, Resnet và Resnet FPN. Từ đó sẽ đánh giá so sánh để tìm ra mô hình phù hợp áp dụng phương pháp phù hợp nhất, xây dựng ứng dụng để cho việc giải quyết bài toán nêu trên.



Kiến trúc mô hình Faster RCNN



Kiến trúc của RetinaNet



Kiến trúc của mô hình FCOS

Input: Một ảnh X-quang cột sống.

Output: Các bất thường nằm thuộc 7 loại bất thường bao gồm gai xương, trượt đốt sống, hẹp khe khớp, xẹp đốt sống, hẹp lỗ liên hợp, kết hợp xương, các bất thường khác và vị trí của các bất thường.

MỤC TIÊU

Xác định vị trí bất thường và loại thương tổn: Phát triển một mô hình học sâu (chọn một trong ba mô hình Faster RCNN, RetinaNet và FCOS), huấn luyện mô hình có thể xác định những bất thường hoặc các loại thương tổn từ ảnh X-quang cột sống.

Hỗ trợ quyết định chẩn đoán cho bác sĩ: Chọn mô hình đạt hiệu suất tốt trong khi huấn luyện, xây dựng ứng dụng có áp dụng mô hình, cho phép người dùng tải lên một file ảnh X-quang cột sống và ứng dụng sẽ trả về một hình ảnh kết quả có hiển thị các vị trí bất thường và loại thương tổn đã được phân loại từ file ảnh ban đầu.

Tăng hiệu quả trong quá trình điều trị: Tạo ra công cụ hỗ trợ chẩn đoán và đánh giá, giúp tăng hiệu quả trong quá trình điều trị cho bệnh nhân. Khi chẩn đoán bất thường một cách chính xác và nhanh

chóng, các bác sĩ có thể đưa ra quyết định điều trị kịp thời và tiết kiệm thời gian.

NỘI DUNG VÀ PHƯƠNG PHÁP

Nội dung: Bài toán nhận diện và phát hiện đối tượng trong ảnh.

Phương pháp: Đề xuất tính các chỉ số để đánh giá hiệu suất và độ chính xác của các phương pháp phát hiện đối tượng. Phương pháp thực hiện để xây dựng mô hình phát hiện đối tượng là học chuyển tiếp và tinh chỉnh mạng, cụ thể là fine-tuning.

- Intersection over Union: Chỉ số IoU được tính bằng cách lấy diện tích của phần giao giữa kết quả phát hiện và đối tượng thực tế, chia cho tổng diện tích của phần hợp giữa cả hai.
- Average Precision: Là một chỉ số phổ biến trong đánh giá các bài toán phát hiện đối tượng. Nó tính toán diện tích dưới đường cong Precision-Recall.
- 11-point interpolation: Là một phương pháp đánh giá AP bằng cách tính toán AP tại 11 điểm độc lập trên đường cong Precision-Recall.
- 101-point interpolation: Là một phương pháp tương tự như 11-point interpolation, nhưng tính toán AP tại 101 điểm trên đường cong Precision-Recall.
- Mean Average Precision (mAP): Là giá trị trung bình của các AP được tính toán cho tất cả các lớp đối tượng.
- **Học chuyển tiếp và tinh chỉnh mạng:** Transfer learning là một phương pháp trong học máy và mạng nơ-ron sử dụng mô hình đã được huấn luyện trên một tác vụ để giải quyết một tác vụ khác có liên quan. Các lợi ích của transfer learning bao gồm: Học từ kiến thức trước đó; Tận dụng dữ liệu nhỏ; Tăng hiệu suất; Pretraining (Tiền huấn luyện); Fine-tuning (Tinh chỉnh).
- **Fine-tuning:** Fine-tuning là quá trình điều chỉnh lại các trọng số của một mô hình đã được huấn luyện trước đó để nâng cao hiệu suất trên một tác vụ mới hoặc tập dữ liệu mới.

Nội dung: Xây dựng bộ dữ liệu để áp dụng huấn luyện các mô hình.

Phương pháp: Áp dụng trên bộ dữ liệu VinDr-SpineXR thuộc Công ty Cổ phần VinBigData (thuộc Tập đoàn Vingroup).

- Tập dữ liệu gốc bao gồm 10,469 hình ảnh X-quang cột sống được thu thập từ hai bệnh viện là Bệnh viện Đại học Y Hà Nội và Bệnh viện Trung ương Quân đội 108. Bộ dữ liệu ban đầu được chia thành hai tập training và testing theo tỉ lệ 80% và 20%. Trong đề tài lần này, đề xuất tập training sẽ được tách ra một phần dùng làm tập validation sao cho tỉ lệ các tập

training, validation, testing lần lượt là 70%, 10% và 20%.

Nội dung: Môi trường huấn luyện và huấn luyện các mô hình.

Phương pháp: Huấn luyện trên môi trường của Kaggle và huấn luyện các mô hình Faster RCNN, RetinaNet và FCOS để đưa ra kết quả so sánh. Dựa vào kết quả so sánh, chọn mô hình cho kết quả tốt nhất áp dụng mô hình vào chương trình ứng dụng.

- **Huấn luyện các mô hình:** Bộ dữ liệu huấn luyện được chia thành ba tập train, val, test với các tỉ lệ 70%, 10% và 20%. Cạnh ngắn của ảnh được resize với kích thước ngẫu nhiên từ 640 đến 800 pixel. Các mô hình được huấn luyện với 20 epoch (ở mạng ResNet101 và ResNet152 số epoch là 25) với learning rate ban đầu là $1e-2$ giảm đi 10 lần sau 10 epoch. Hàm tối ưu sử dụng là SGD. Mometum là 0.9, weight decay là 0.0001.

Nội dung: Chọn mô hình phù hợp và xây dựng các chức năng cơ bản cho ứng dụng.

Phương pháp: Xây dựng các chức năng cho ứng dụng: Chức năng xem kết quả chẩn đoán; Xem danh sách dự án được phân công, danh sách các hình ảnh cần chẩn đoán và nhãn cần chẩn đoán; Quản lý dự án.

- **Môi trường ứng dụng:** Chương trình có thể xây dựng bằng ngôn ngữ Python với framework áp dụng là Django. Một số thư viện được áp dụng trong quá trình xây dựng bao gồm: Pytorch, Torchvision, cv2, PIL,.... Chương trình được triển khai với Ngrok để có thể đưa được ứng dụng lên Internet.

KẾT QUẢ MONG ĐỢI

- Kết quả so sánh khả quan giữa các mô hình phát hiện đối tượng (Faster RCNN, RetinaNet và FCOS). Hy vọng mô hình Faster RCNN với mạng backbone ResNet152 FPN cho kết quả tốt nhất trong các mô hình thử nghiệm.
- Triển khai được chương trình ứng dụng cơ bản trong lĩnh vực chẩn đoán ảnh y khoa (ảnh X-quang cột sống).
- Có kết quả khả thi để tiếp tục phát triển hướng nghiên cứu, đưa ra các phương pháp nâng cao hiệu quả, các biện pháp tăng cường dữ liệu, thay đổi các tham số đầu vào cho mô hình để cải thiện độ chính xác, thu thập dữ liệu mới huấn luyện cho mô hình hoặc kết hợp thêm các bộ dữ liệu khác để có khả năng nhận diện giai đoạn bệnh hiện tại và dự đoán về các trạng thái giai đoạn tiếp theo của bệnh.

TÀI LIỆU THAM KHẢO (Định dạng DBLP)

- [1] S. K. B. Sangeetha, V. Muthukumaran, K. Deeba, H. Rajadurai, V. Maheshwari, and G. T. Dalu, "Multiconvolutional Transfer Learning for 3D Brain Tumor Magnetic Resonance Images," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–9, Aug. 2022, doi: 10.1155/2022/8722476.
- [2] E. Karypidis, S. G. Mouslech, K. Skoulariki, and A. Gazis, "Comparison Analysis of Traditional Machine Learning and Deep Learning Techniques for Data and Image Classification," *WSEAS TRANSACTIONS ON MATHEMATICS*, vol. 21, pp. 122– 130, Mar. 2022, doi: 10.37394/23206.2022.21.19.
- [3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2012. Accessed: Jul. 09, 2023. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html
- [4] S. P. Singh, L. Wang, S. Gupta, H. Goli, P. Padmanabhan, and B. Gulyás, "3D Deep Learning on Medical Images: A Review," *Sensors*, vol. 20, no. 18, p. 5097, Sep. 2020, doi: 10.3390/s20185097.
- [5] "Bone fracture detection through the two-stage system of Crack-Sensitive Convolutional Neural Network," *Informatics in Medicine Unlocked*, vol. 22, p. 100452, Jan. 2021, doi: 10.1016/j.imu.2020.100452.
- [6] E. Yahalomi, M. Chernofsky, and M. Werman, "Detection of Distal Radius Fractures Trained by a Small Set of X-Ray Images and Faster R-CNN," vol. 997, pp. 971–981, 2019, doi: 10.1007/978-3-030-22871-2_69.
- [7] X. Zhang et al., "A New Window Loss Function for Bone Fracture Detection and Localization in X-ray Images with Point-based Annotation." *arXiv*, Jan. 04, 2021. Accessed: Jul. 09, 2023. [Online]. Available: <http://arxiv.org/abs/2012.04066>
- [8] F. Hardalaç et al., "Fracture Detection in Wrist X-ray Images Using Deep LearningBased Object Detection Models," *Sensors*, vol. 22, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/s22031285.
- [9] H. T. Nguyen et al., "VinDr-SpineXR: A deep learning framework for spinal lesions detection and classification from radiographs." *arXiv*, Jun. 24, 2021. Accessed: Jul. 09, 2023. [Online]. Available: <http://arxiv.org/abs/2106.12930>
- [10] "Ground Truth Annotated Femoral X-Ray Image Dataset and Object Detection Based Method

for Fracture Types Classification | IEEE Journals & Magazine | IEEE Xplore.”
<https://ieeexplore.ieee.org/document/9214464?denied=> (accessed Jul. 09, 2023).