

ẢNH HƯỞNG CỦA CÁC KỸ THUẬT TIỀN XỬ LÝ TRONG BÀI TOÁN PHÁT HIỆN BỆNH TRÊN ẢNH X-QUANG NGỰC

Đỗ Phạm Phúc Tính^{1,2}, Nguyễn Trần Gia Thé^{1,2}, Cao Đình Duy Ngọc^{1,2}, Ngô
Huỳnh Trưởng^{1,2}, and Nguyễn Tất Bảo Thiện^{1,3}

¹ Trường Đại học Công Nghệ Thông Tin, ĐHQG TP.HCM

² {20522020, 20521940, 20521661, 20522085}@gm.uit.edu.vn

³ thienntb@uit.edu.vn

Tóm tắt nội dung Trong sự phát triển mạnh mẽ của trí tuệ nhân tạo, việc ứng dụng các kỹ thuật học máy, học sâu đóng vai trò quan trọng trong chẩn đoán bệnh trên hình ảnh X-Quang, điều này giúp phát hiện sớm và điều trị kịp thời cho các bệnh nhân. Bên cạnh các thuật toán, việc tiền xử lý hình ảnh trước khi huấn luyện cũng ảnh hưởng đáng kể đến hiệu suất. Do đó, ở bài toán này, chúng tôi tiến hành thực nghiệm để kiểm tra mức độ hiệu quả của một số phương pháp tiền xử lý hình ảnh cơ bản như tăng cường độ tương phản (histogram equalization và CLAHE), khử nhiễu (bilateral filter), biến đổi phân ngưỡng (thresholding transformations), tạo hiệu ứng viền (unsharp masking) trên bộ dữ liệu VinDr CXR với ba mô hình: YOLOv5, YOLOv8 và Faster R-CNN. VinDr CXR là bộ dữ liệu X-quang ngực của Việt Nam do VinDr lab nghiên cứu và công bố vào cuối năm 2020. Đây là một trong số ít các bộ dữ liệu về X-quang có đánh dấu các vị trí bất thường và được thực hiện bởi nhiều bác sĩ chuyên khoa có kinh nghiệm, vì vậy đây là bộ dữ liệu đáng tin cậy để tiến hành thực nghiệm. Kết quả thực nghiệm cho thấy các phương pháp tăng cường ảnh cơ bản không mang lại hiệu quả cho mô hình.

Từ khóa: Phát hiện đối tượng, VinDr CXR, Tiền xử lý, YOLO v5, YOLO v8, Fast R-CNN, X-quang

1 GIỚI THIỆU

Sự phát triển mạnh mẽ của thị giác máy tính kéo theo những yêu cầu về quá trình xử lý hình ảnh để phù hợp với các tác vụ. Việc áp dụng các kỹ thuật tiền xử lý giúp cải thiện chất lượng và độ tin cậy của dữ liệu hình ảnh, tạo ra các đặc trưng hữu ích và chuẩn bị cho các bước tiếp theo trong quá trình xử lý hình ảnh.

Trong lĩnh vực hình ảnh X-quang, tiền xử lý đóng một vai trò đặc biệt quan trọng. Hình ảnh X-quang chứa thông tin quan trọng về sự xuất hiện của các bất thường trong cơ thể, nhưng đồng thời cũng có nhược điểm như độ tương phản

thấp, nhiễu và định dạng không đồng nhất. Do đó, việc tiền xử lý hình ảnh trở thành một yếu tố quan trọng. Tuy vậy, việc tìm ra kỹ thuật tiền xử lý nào phù hợp với bài toán cũng là một thách thức cho các nhà phát triển. Vì vậy ở đề tài này, chúng tôi sẽ thực nghiệm hiệu quả của một số kỹ thuật tiền xử lý trên ảnh X-quang.

Chúng tôi sử dụng bộ dữ liệu VinDr CXR [15] để thực nghiệm và đánh giá các phương pháp tiền xử lý hình ảnh trong việc chẩn đoán bệnh dựa trên ảnh X-quang. Bộ dữ liệu VinDr CXR gồm 100.000 ảnh X-quang được thu thập từ hai bệnh viện lớn tại Việt Nam. Các tác giả đã công bố một phần bộ dữ liệu với kích thước 18.000 ảnh gồm 6 nhãn toàn cục và 22 nhãn cục bộ. Đặc biệt, tất cả các hình ảnh đã được gán nhãn bằng tay bởi 17 bác sĩ chuyên khoa có kinh nghiệm, đảm bảo tính chính xác và uy tín của dữ liệu. Chúng tôi chỉ sử dụng 14 nhãn (những nhãn có số lượng đủ lớn để mô hình có thể học một cách hiệu quả) của bộ dữ liệu và chỉ sử dụng những hình ảnh có đề cập đến sự bất thường.

Trong quá trình thực nghiệm, chúng tôi áp dụng các mô hình tiên tiến như YOLOv5, YOLOv8 và Faster R-CNN để phân loại và phát hiện bất thường trên ảnh. Các phương pháp tiền xử lý mà chúng tôi tiếp cận trong bài toán này bao gồm tăng cường độ tương phản (histogram equalization), khử nhiễu (bilateral filter), biến đổi phân ngưỡng (thresholding transformations), tạo hiệu ứng viền (unsharp masking) để phân tích, đánh giá và đưa ra kết luận về sự ảnh hưởng của các kỹ thuật tiền xử lý dữ liệu.

Qua nghiên cứu này, chúng tôi hy vọng rằng sẽ mang lại cái nhìn tổng quát về sự ảnh hưởng của các phương pháp tiền xử lý vào quá trình chẩn đoán bệnh dựa trên ảnh X-quang. Bên cạnh đó, việc so sánh và đánh giá các phương pháp tiền xử lý khác nhau cũng sẽ giúp định hình những hướng phát triển tiềm năng trong lĩnh vực này.

Các phần còn lại trong bài báo như sau: Phần 2 trình bày tình hình các nghiên cứu liên quan trong những năm gần đây. Phần 3 giới thiệu về bộ dữ liệu VinDr CXR. Phần 4 mô tả các phương pháp thay đổi chất lượng ảnh được chúng tôi sử dụng trong bài nghiên cứu này. Phần 5 trình bày về kiến trúc của các mô hình được sử dụng. Phần 6 đưa ra những thông tin về các thí nghiệm trong bài nghiên cứu. Phần 7 là kết quả đạt được và phân tích lỗi. Phần 8 đưa những kết luận chúng tôi đạt được cùng với đề ra hướng phát triển.

2 CÔNG TRÌNH LIÊN QUAN

Phát hiện các bệnh trong hình ảnh X-quang là một tác vụ đang được quan tâm trong những năm gần đây trong lĩnh vực y tế kết hợp trí tuệ nhân tạo. Để phát triển và phục vụ tác vụ này, nhiều công trình nghiên cứu với nhiều phương pháp, bộ dữ liệu mới đã được công bố. Trong bài báo [1], các tác giả đã tiếp cận một vấn đề phân loại đa nhãn trong hình ảnh X-quang, với hai nhiệm vụ quan trọng: xác định sự hiện diện của một bệnh và phân loại bệnh đó là COVID-19 hay viêm phổi. Trong nghiên cứu này, ba bộ dữ liệu X-quang đã được sử dụng. Bộ dữ liệu thứ nhất bao gồm các hình ảnh thu thập từ bộ dữ liệu được giới thiệu trong bài báo "COVID-19 image data collection", và có sẵn tại [2]. Bộ dữ

liệu thứ hai được công bố trong công trình "Automated Detection of COVID-19 Cases Using Deep Neural Networks with X-ray Images" [16]. Bộ dữ liệu thứ ba là ChestX-ray8 [23].

Ngoài công trình trên, các tác giả của bài báo [17] đã sử dụng hai bộ dữ liệu khác để phục vụ phân loại và phát hiện bệnh. Đó là CheXpert [8], một bộ dữ liệu gồm các hình ảnh X-quang ngực lấy từ các bệnh viện với hơn 174.000 ảnh bất thường, 17.000 hình ảnh bình thường ở định dạng JPG, và RSNA CXR [19], một bộ dữ liệu cung cấp bởi Hiệp hội X-quang Hoa Kỳ (RSNA) chứa hơn 17.800 ảnh bất thường, 18.500 ảnh bình thường ở định dạng DICOM. Việc sử dụng các bộ dữ liệu phong phú và đa dạng này giúp tăng tính đa dạng và độ chính xác của các mô hình phân loại và phát hiện bệnh.

Bên cạnh đó, trong bài báo [14], tác giả đã thực hiện tác vụ phát hiện bệnh trên hình ảnh X-quang bằng bộ dữ liệu SNA Pneumonia Detection Challenge⁴ và một bộ dữ liệu về hình ảnh X-quang ngực trẻ em là pediatric chest X-ray. Bộ dữ liệu RSNA Pneumonia Detection Challenge là một bộ dữ liệu lớn chứa số lượng lớn hình ảnh X-quang. Điều này giúp tăng tính đa dạng và độ phức tạp của các trường hợp bệnh trong quá trình huấn luyện và đánh giá mô hình. Bài báo [4] đã thử nghiệm một số phương pháp tiền xử lý hình ảnh cơ bản đến kết quả của bài toán phân loại các nhãn normal, COVID-19, Pneumonia và đánh giá, so sánh độ hiệu quả giữa các kỹ thuật.

Ngoài ra, trong cuộc thi VinBigData Chest X-ray Abnormalities Detection Competition⁵, một cuộc thi quan trọng trong lĩnh vực này, bộ dữ liệu VinDr-CXR [15] đã được sử dụng cho các nhiệm vụ phân loại và phát hiện các bất thường trong hình ảnh X-quang ngực. Bộ dữ liệu này bao gồm các hình ảnh X-quang từ nhiều bệnh viện và chứa 14 loại bệnh khác nhau. Cuộc thi VinBigData đã tạo ra một sự cạnh tranh và khám phá trong lĩnh vực phát hiện bệnh trên hình ảnh X-quang.

Bên cạnh những nghiên cứu trên cuộc thi trên, bộ dữ liệu VinDr CXR cũng đã được sử dụng để thực hiện thử nghiệm phát hiện các loại bệnh bởi nhóm VNOpenAI⁶. Điều này đóng góp vào việc nghiên cứu và phát triển các phương pháp phân loại và phát hiện bệnh sử dụng hình ảnh X-quang đối với tình trạng sức khỏe của người dân Việt Nam.

3 DATASET

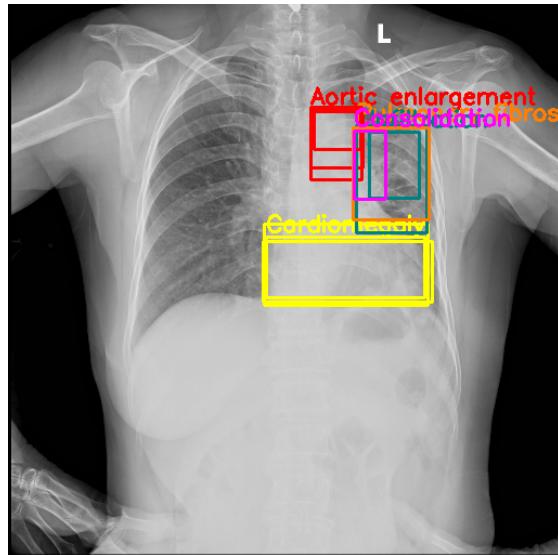
Bộ dữ liệu VinDR-CXR [15] đã được giới thiệu vào cuối năm 2020 [15]. Bộ dữ liệu này bao gồm hơn 100.000 hình ảnh định dạng DICOM được thu thập từ năm 2018 đến năm 2020 tại Bệnh viện 108 và Bệnh viện Đại học Y Hà Nội. 17 bác sĩ chuyên khoa đã gắn nhãn cho các hình ảnh với ít nhất 8 năm kinh nghiệm. Các tác giả đã xây dựng bộ dữ liệu này cho VinBigData Chest X-ray

⁴ <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge>

⁵ <https://www.kaggle.com/c/vinbigdata-chest-xray-abnormalities-detection/>

⁶ <https://vnopenai.github.io/ai-doctor/vision/lung-abnormalities-detection-xray/intro/>

Abnormalities Detection Competition⁷ trên nền tảng Kaggle. Bộ dữ liệu cho cuộc thi bao gồm 18.000 hình ảnh với 15 nhãn, bao gồm một tập huấn luyện gồm 15.000 hình ảnh và một tập kiểm tra gồm 3.000 hình ảnh. Ba bác sĩ đã gán nhãn cho các hình ảnh trong tập huấn luyện, trong khi năm bác sĩ đã gán nhãn cho tập kiểm tra. Hình minh họa dưới đây là một ví dụ trong bộ dữ liệu.



Hình 1: Một ví dụ trong bộ dữ liệu.

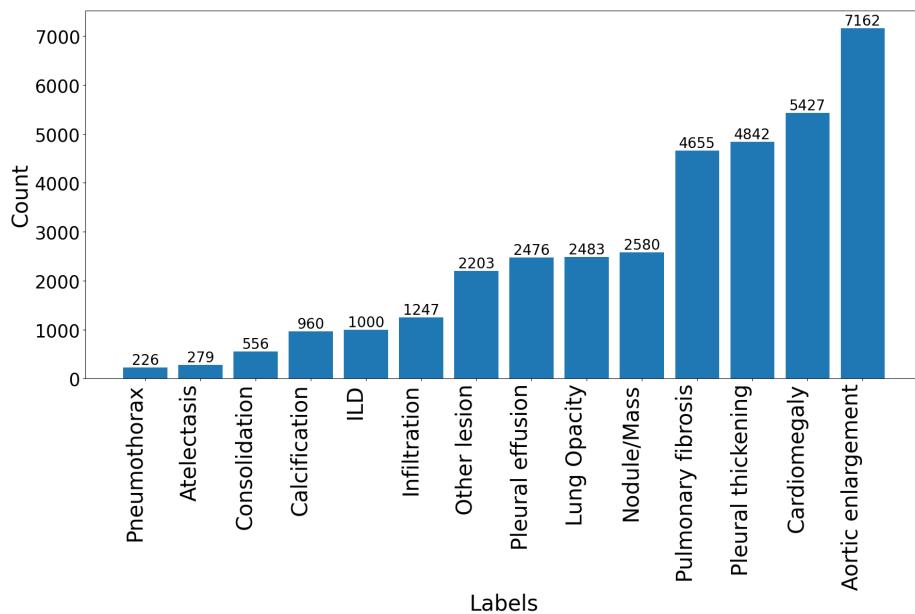
Chúng tôi chỉ thu thập hình ảnh và nhãn từ tập huấn luyện, bởi vì tác giả không công khai nhãn cho bộ kiểm thử. Đối với nhiệm vụ phát hiện, chúng tôi yêu cầu có các tọa độ điểm giới hạn cho các nhãn. Tuy nhiên, nhãn "No finding" không có các tọa độ điểm này, khiến việc phát hiện trở nên không thể. Do đó, chúng tôi đã loại bỏ các hình ảnh có nhãn "No finding". Ngoài ra, chúng tôi đã điều chỉnh kích thước hình ảnh và chuyển đổi định dạng của chúng. Sau quá trình này, chúng tôi thu được 4.394 hình ảnh có kích thước 512x512 pixel, định dạng PNG, 14 nhãn (các nhãn được chúng tôi mô tả ở bảng 4. Những hình ảnh này được sử dụng để thử nghiệm hiệu suất các kỹ thuật tăng cường chất lượng huấn thông qua các mô hình. Số lượng hình ảnh trong các tập huấn luyện, phát triển và kiểm tra lần lượt là 3.515, 439 và 440(tỉ lệ 8:1:1).

Hình 2 và bảng 1 cho thấy phân phối các nhãn trong toàn bộ bộ dữ liệu, của 3 tập không đồng đều. Một số nhãn, như "Aortic enlargement" and "Cardiomegaly" xuất hiện nhiều hơn rất nhiều lần so với các nhãn khác như "Pneumothorax", "Atelectasis" và "Consolidation". Điều này cho thấy sự mất cân đối trong bộ dữ liệu, khi một số nhãn được đại diện quá nhiều so với những nhãn khác. Sự mất cân đối này có thể ảnh hưởng đến hiệu suất của mô hình. Các tập dữ liệu có cùng phân phối giúp cho việc đánh giá hiệu suất chính xác

⁷ <https://www.kaggle.com/c/vinbigdata-chest-xray-abnormalities-detection/>

Bảng 1: Thống kê số lượng bbox của mỗi nhãn trên từng tập dữ liệu.

| Nhãn | Toàn bộ | Huấn luyện | Phát triển | Kiểm tra |
|---------------------------|---------|------------|------------|----------|
| Aortic enlargement | 7162 | 5657 | 748 | 757 |
| Cardiomegaly | 5427 | 4294 | 565 | 568 |
| Pleural thickening | 4842 | 3938 | 472 | 453 |
| Pulmonary fibrosis | 4655 | 3770 | 451 | 413 |
| Nodule/Mass | 2580 | 2109 | 219 | 259 |
| Lung Opacity | 2483 | 2044 | 213 | 233 |
| Pleural effusion | 2476 | 2031 | 212 | 219 |
| Other lesion | 2203 | 1779 | 212 | 212 |
| Calcification | 1247 | 1021 | 121 | 131 |
| Infiltration | 1000 | 803 | 108 | 105 |
| ILD | 960 | 741 | 88 | 89 |
| Consolidation | 556 | 454 | 53 | 49 |
| Atelectasis | 279 | 233 | 27 | 24 |
| Pneumothorax | 226 | 183 | 22 | 16 |



Hình 2: Thống kê số nhãn của bộ dữ liệu sau khi bỏ nhãn No finding.

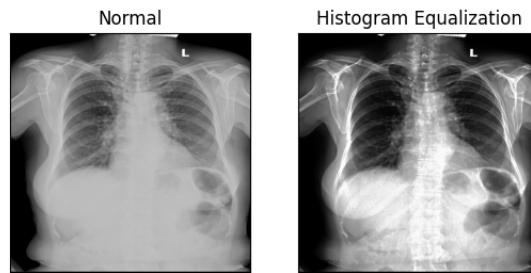
hơn. Số lượng ảnh ít hơn số lượng bounding box của mỗi nhãn là do trên một ảnh có thể có nhiều vị trí của bounding box được gán cùng 1 nhãn.

4 CÁC PHƯƠNG PHÁP TIỀN XỬ LÝ

Trong phần này, chúng tôi trình bày các phương pháp tiền xử lý hình ảnh được áp dụng trong bài toán.

4.1 Histogram equalization

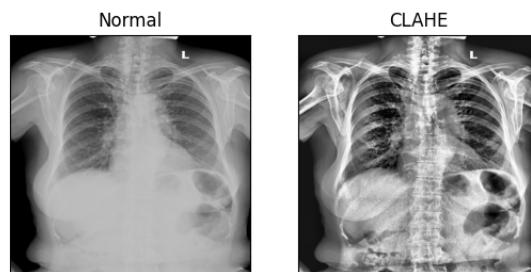
Phương pháp Histogram equalization được sử dụng để cân bằng histogram của hình ảnh. Việc cân bằng histogram giúp phân phối giá trị pixel trên toàn bộ phạm vi độ sáng, tăng độ tương phản và làm nổi bật các chi tiết quan trọng trong hình ảnh y khoa. Phương pháp này đặc biệt hữu ích đối với các hình ảnh có độ tương phản thấp hoặc bị méo do sự khác biệt về ánh sáng và điều kiện chụp.



Hình 3: Histogram equalization

4.2 CLAHE

CLAHE (Contrast Limited Adaptive Histogram Equalization) được sử dụng để cải thiện độ tương phản của hình ảnh. CLAHE tạo ra một ảnh đầu ra có độ tương phản tốt hơn bằng cách cân bằng histogram của các vùng nhỏ trong hình ảnh. Để hạn chế hiện tượng tăng nhiễu, CLAHE áp dụng một giới hạn đối với biến đổi histogram. Điều này đảm bảo rằng sự cân bằng tần số chỉ áp dụng trong một phạm vi hạn chế, từ đó giữ lại các chi tiết quan trọng trong hình ảnh, đây là điểm cải tiến so với phương pháp Histogram equalization ở phần trên.

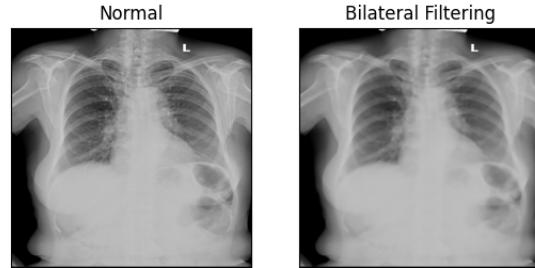


Hình 4: CLAHE

4.3 Bilateral Filtering

Phương pháp Bilateral Filtering được sử dụng để làm mờ hình ảnh và giảm nhiễu. Bilateral Filtering bảo toàn các cạnh và đặc trưng quan trọng trong hình

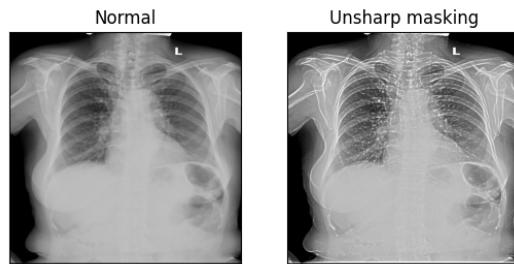
ảnh trong khi loại bỏ nhiễu thông qua việc áp dụng một bộ lọc dựa trên không gian và độ tương tự.



Hình 5: Bilateral Filtering

4.4 Unsharp masking

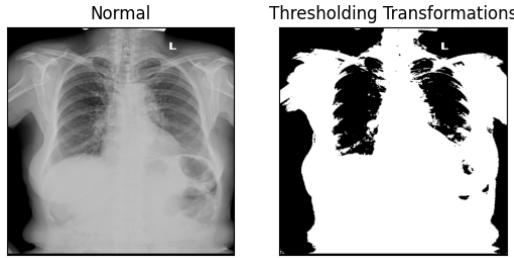
Unsharp masking là một phương pháp tiền xử lý được sử dụng để làm nổi bật các đường viền và chi tiết trong hình ảnh. Phương pháp này bao gồm làm mờ ảnh ban đầu bằng một bộ lọc Gaussian, sau đó trừ ảnh đã làm mờ từ ảnh gốc để tạo ra một ảnh nền mờ.



Hình 6: Unsharp masking

4.5 Thresholding Transformations

Phương pháp Thresholding Transformations [13] được sử dụng để chuyển đổi hình ảnh thành hình ảnh nhị phân dựa trên ngưỡng. Điều này giúp tách các đối tượng quan trọng và nền trong hình ảnh. Chúng tôi áp dụng các phương pháp như Thresholding đơn giản, Adaptive Thresholding và Otsu Thresholding để chọn ngưỡng phù hợp và tạo ra một hình ảnh nhị phân phù hợp cho việc phân tích và trích xuất thông tin.



Hình 7: Thresholding Transformations

5 MÔ HÌNH

5.1 YOLOv5

YOLOv5 được phát hành vào năm 2020 bởi Ultralytics [9], một mô hình thị giác máy tính dùng để phát hiện đối tượng. YOLOv5 đã thực hiện một số thay đổi kiến trúc như hình 8, đáng chú ý nhất là việc chia mô hình thành ba thành phần, bao gồm backbone, neck và head.

Dầu tiên, **Backbone** của YOLOv5 là Darknet53, một kiến trúc mạng được thiết kế để trích xuất đặc trưng. Nó sử dụng cửa sổ filter nhỏ và kết nối dư để tăng cường khả năng. Wang et al. [22] đã đề xuất khái niệm (CSPNet) Cross Stage Partial connections, cho phép kiến trúc đạt được luồng gradient phong phú hơn trong khi giảm độ phức tạp tính toán.

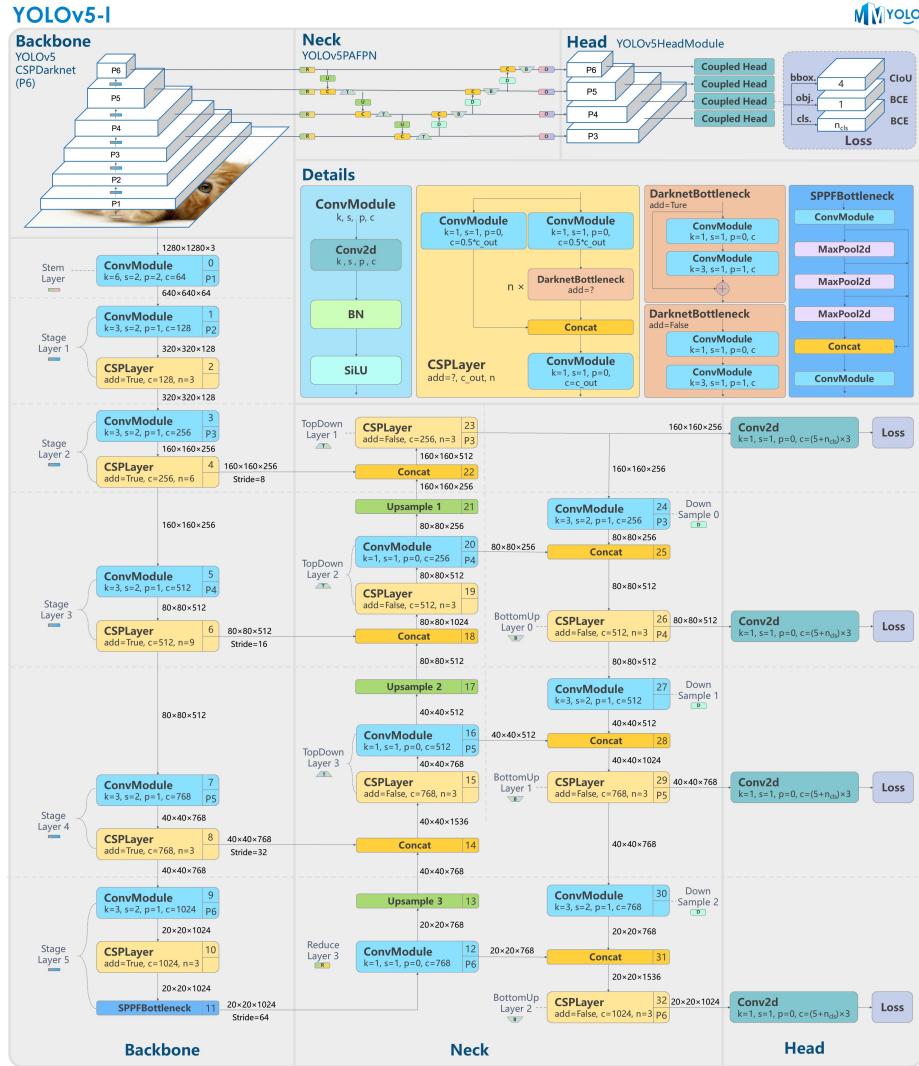
Neck của YOLOv5, mô tả bởi Teven et al. [20], chịu trách nhiệm kết nối giữa backbone và head. Nhiệm vụ của neck là tập hợp và tinh chỉnh các đặc trưng được trích xuất bởi backbone, tập trung vào cải thiện thông tin không gian và ngữ nghĩa qua các tỷ lệ khác nhau. Kiến trúc bao gồm mô-đun Spatial Pyramid Pooling (SPP) [7] và CSP-Path Aggregation Network [22].

Head của YOLOv5 bao gồm ba nhánh dự đoán khác nhau về các đặc trưng và thông tin về các đối tượng trong ảnh đầu vào. Cụ thể, các nhánh này dự đoán bounding boxes, xác suất các lớp và confidence scores cho các đối tượng.

YOLOv5 cung cấp năm phiên bản: YOLOv5n (nano), YOLOv5s (small), YOLOv5m (medium), YOLOv5l (large), and YOLOv5x (extra large), trong đó chiều rộng và độ sâu của các mô-đun tích chập thay đổi để phù hợp với các ứng dụng cụ thể và yêu cầu phần cứng. Trong báo cáo này, chúng tôi sử dụng phiên bản YOLOv5l để tiến hành huấn luyện.

5.2 YOLOv8

YOLOv8 là phiên bản mới nhất của mô hình phát hiện đối tượng YOLO, được phát hành vào tháng 1 năm 2023 bởi Ultralytics [10]. Phiên bản mới này có cấu trúc tương tự như các phiên bản trước đó như hình 9 và giới thiệu nhiều cải tiến so với các phiên bản trước của YOLO.



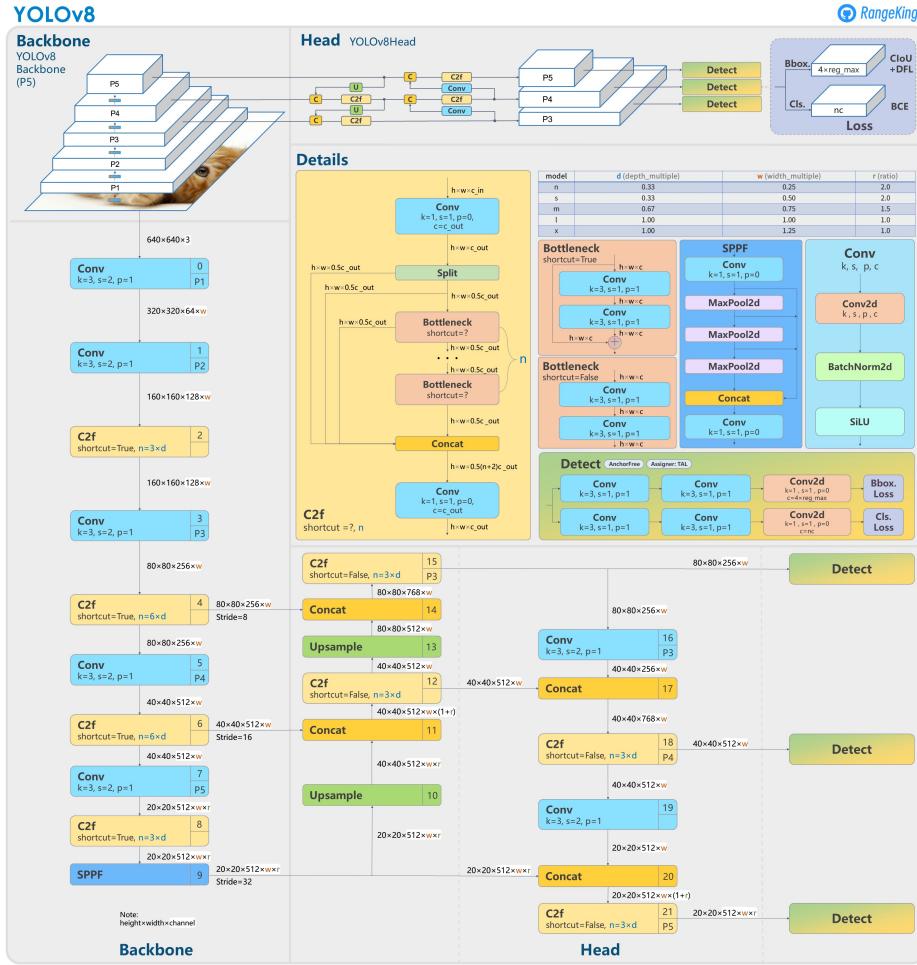
Hình 8: Kiến trúc của YOLOv5.

Nguồn: <https://github.com/open-mmlab/mmyolo/tree/main/configs/yolov5>

YOLOv8 có cùng backbone với YOLOv5 nhưng có một số thay đổi trên CSPLayer, hiện được gọi là C2f module [21]. C2f module kết hợp các đặc trưng cấp cao với thông tin ngữ cảnh để cải thiện độ chính xác phát hiện đối tượng.

YOLOv8 sử dụng mô hình không gắn kết (anchor-free) với kiến trúc Head tách rời để xử lý độc lập các nhiệm vụ về objectness, classification, và regression. Thiết kế này cải thiện độ chính xác tổng thể.

Các hàm mất mát trong YOLOv8 bao gồm ClIoU [24] (Complete Intersection over Union) và DFL [11] (Distribution Focal Loss) cho việc mất mát bounding box và mất mát phân loại nhị phân. Các hàm mất mát này đã chứng minh hiệu



Hình 9: Kiến trúc của YOLOv8.

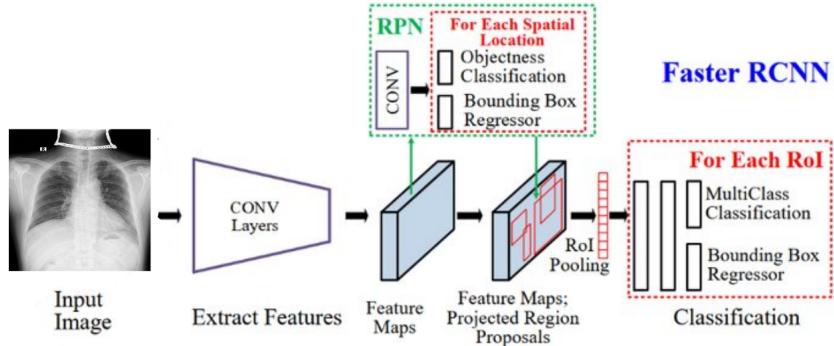
Nguồn ảnh: <https://github.com/ultralytics/ultralytics/issues/189>

suất cải thiện trong phát hiện đối tượng, đặc biệt là khi phát hiện và xác định vị trí các đối tượng nhỏ hơn.

YOLOv8 cung cấp năm phiên bản: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra large), trong đó chiều rộng và độ sâu của các mô-đun tích chập thay đổi để phù hợp với các ứng dụng cụ thể và yêu cầu phần cứng. Trong báo cáo này, chúng tôi sử dụng phiên bản YOLOv8l để tiến hành huấn luyện.

5.3 Faster Regional Convolution Neural Networks

Faster R-CNN [18] là một mở rộng của phương pháp R-CNN [6] và Fast R-CNN [5] trong việc phát hiện đối tượng. Phương pháp R-CNN ban đầu yêu



Hình 10: Phương pháp Faster R-CNN trong bài báo

cầu thực hiện truyền thuận của Mạng Neural Tích chập cho mỗi khoảng 2000 đề xuất vùng trong một hình ảnh. Fast R-CNN đã cải thiện điều này bằng cách chia sẻ tính toán đặc trưng tích chập giữa các đề xuất khác nhau, giúp tăng tốc quá trình phát hiện. Tuy nhiên, Fast R-CNN vẫn phụ thuộc vào phương pháp đề xuất vùng, chẳng hạn như Selective Search hoặc Edge Box để tạo ra các đề xuất. Để khắc phục hạn chế này, Ren và đồng nghiệp [18] đã giới thiệu Mạng Đề Xuất Vùng (Region Proposal Network).

Faster R-CNN kết hợp các ưu điểm của RPN và Fast R-CNN thành một khung công việc thống nhất. RPN khởi tạo các hộp tham chiếu vuông với tỷ lệ khía cạnh và tỷ lệ quy mô đa dạng tại mỗi vị trí trên bản đồ đặc trưng tích chập. Sau đó, các hộp này được ánh xạ thành các vector đặc trưng, được đưa vào các lớp kết nối đầy đủ để phân loại danh mục đối tượng và điều chỉnh hộp. Bằng cách chia sẻ đặc trưng tích chập với Fast R-CNN, RPN cho phép tính toán đề xuất vùng hiệu quả cao. Nó được huấn luyện end-to-end để tạo ra các đề xuất vùng chất lượng cao cho hình ảnh đầu vào có kích thước tùy ý.

6 Thiết kế thí nghiệm

6.1 Xử lý hình ảnh

Chúng tôi tiến hành thí nghiệm với 6 bộ dữ liệu: bộ dữ liệu chứa các hình ảnh gốc, không có bất kỳ sửa đổi nào, các bộ dữ liệu tương ứng các phương pháp thay đổi chất lượng đã đề cập ở section 4. Kết quả của bộ dữ liệu gốc như một điểm sàn để so sánh với các kỹ thuật tiền xử lý khác. Trong thí nghiệm này, ở phương pháp CLAHE, chúng tôi sử dụng ngưỡng giới hạn độ tương phản (cliplimit) là 20. Đối với phương pháp khử nhiễu Bilateral Filtering, hai tham số sigma (điều chỉnh độ mờ của Gaussian) và strength (độ mạnh của hiệu ứng tăng cường cạnh) lần lượt là 1 và 1.5. Giá trị ngưỡng (thresh) bằng 127 được sử dụng ở phương pháp Thresholding Transformers.

Bảng 2: Các phương pháp thay đổi chất lượng ảnh.

| Data | Phương pháp xử lí |
|-----------|------------------------------|
| Normal | Không xử lí |
| Process 1 | CLAHE |
| Process 2 | Bilateral Filtering |
| Process 3 | Unsharp masking |
| Process 4 | Thresholding Transformations |
| Process 5 | Histogram equalization |

6.2 Độ đo

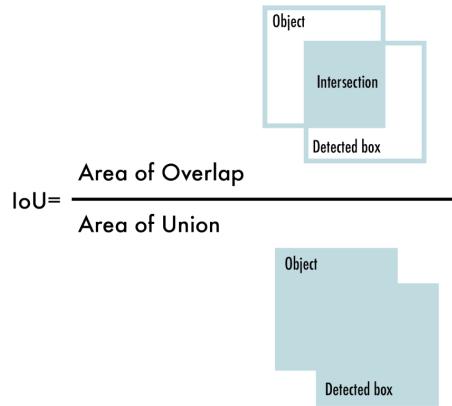
Mean Average Precision (mAP), là phương pháp đo thông thường được sử dụng để đánh giá hiệu suất của các mô hình phát hiện đối tượng. Trong YOLOv1 và YOLOv2, tập dữ liệu được sử dụng để huấn luyện và đánh giá là PASCAL VOC 2007 và VOC 2012 [3]. Tuy nhiên, từ YOLOv3 trở đi, đặc biệt là YOLOv5 và YOLOv8 trong báo cáo này, tập dữ liệu được sử dụng là Microsoft COCO (Common Objects in Context) [12]. Tiêu chuẩn hiện tại sử dụng COCO AP để đánh giá chi tiết hơn về mức độ hoạt động của một mô hình ở các ngưỡng IoU khác nhau. Độ chính xác trung bình (mAP) dựa trên các số liệu Precision và Recall, Handling multiple object categories và Intersection over Union(IoU).

Precision và Recall: độ đo mAP tích hợp đường cong Precision và Recall, biểu đồ Precision so với Recall cho các ngưỡng tin cậy khác nhau. Độ đo này cung cấp một đánh giá cân bằng về Precision và Recall bằng cách xem xét diện tích dưới đường cong Precision và Recall.

Handling multiple object categories: Độ đo mAP giải quyết vấn đề này bằng cách tính toán độ chính xác trung bình (AP) cho từng loại đối tượng riêng biệt và sau đó lấy trung bình của các AP này cho tất cả các loại đối tượng. Phương pháp này đảm bảo rằng hiệu suất của mô hình được đánh giá cho từng loại đối tượng một cách độc lập, cung cấp một đánh giá toàn diện hơn về hiệu suất tổng thể của mô hình.

Intersection over Union (IoU): chỉ số IoU cho biết mức độ chồng chéo giữa các tọa độ bounding box được dự đoán và bounding box thực tế (ảnh). Giá trị IoU cao cho thấy các tọa độ bounding box dự đoán gần giống với các tọa độ bounding box thực tế.

Trong báo cáo này, chúng tôi lựa chọn **sử dụng cả mAP50 và mAP50-95** giúp đánh giá hiệu suất của các mô hình, các thuật toán object detection một cách toàn diện. mAP50 cho biết độ chính xác đối tượng với một ngưỡng trung bình, trong khi mAP50-95 cung cấp thông tin về chất lượng phát hiện đối tượng trên nhiều mức độ khác nhau của IoU.



Hình 11: Intersection over Union (IoU). Nguồn: [20]

7 KẾT QUẢ

7.1 Kết quả các mô hình

Bảng 3: Kết quả mô hình

| Model | Data | Scores | | | |
|--------------|-----------|-----------|--------|--------|----------|
| | | Precision | Recall | mAP50 | mAP50-95 |
| YOLOv5 | Normal | 0.403 | 0.333 | 0.322 | 0.15 |
| | Process 1 | 0.386 | 0.307 | 0.294 | 0.136 |
| | Process 2 | 0.422 | 0.317 | 0.314 | 0.147 |
| | Process 3 | 0.433 | 0.321 | 0.322 | 0.152 |
| | Process 4 | 0.35 | 0.198 | 0.188 | 0.0859 |
| | Process 5 | 0.444 | 0.317 | 0.323 | 0.147 |
| YOLOv8 | Normal | 0.397 | 0.347 | 0.338 | 0.169 |
| | Process 1 | 0.374 | 0.301 | 0.291 | 0.148 |
| | Process 2 | 0.37 | 0.323 | 0.306 | 0.156 |
| | Process 3 | 0.397 | 0.326 | 0.325 | 0.166 |
| | Process 4 | 0.245 | 0.165 | 0.0988 | 0.0365 |
| | Process 5 | 0.284 | 0.324 | 0.265 | 0.098 |
| Faster R-CNN | Normal | 0.166 | 0.204 | 0.131 | 0.098 |
| | Process 1 | 0.125 | 0.147 | 0.101 | 0.077 |
| | Process 2 | 0.071 | 0.078 | 0.064 | 0.047 |
| | Process 3 | 0.073 | 0.09 | 0.069 | 0.052 |
| | Process 4 | 0.027 | 0.031 | 0.006 | 0.01 |
| | Process 5 | 0.135 | 0.15 | 0.096 | 0.073 |

Nhìn vào bảng 3 có thể thấy rằng với mô hình Yolov8 huấn luyện trên bộ dữ liệu gốc (Normal) đạt kết quả cao nhất trong cả 3 mô hình ở thang đo mAP 50 và mAP50-90 với số điểm lần lượt là 0.338 và 0.169. Mô hình Yolov5 đạt kết

quả cao nhất ở cách xử lí Unsharp masking (Process 3) nhưng không quá chênh lệch với kết quả trên bộ dữ liệu gốc (Normal). Mô hình cuối là Faster R-CNN thì kết quả cao nhất ở bộ dữ liệu gốc (Normal). Nhưng điểm số ở mô hình này khá thấp so với các mô hình Yolo.

7.2 Thảo luận

Dữ liệu ảnh x-quang thường có đặc điểm đối tượng nhỏ và mật độ đối tượng cao có thể gây khó khăn cho mô hình dự đoán. Với phương pháp dự đoán trên lối ô, chia nhỏ ảnh, các mô hình YOLO có thể dự đoán được những chi tiết nhỏ trong ảnh x-quang. Trong khi đó, mô hình Faster R-CNN sử dụng mạng CNN (Convolutional Neural Network) thiên về tổng quát hình ảnh nên mô hình này có kết quả không tốt bằng các mô hình YOLO.

Để phân tích sâu hơn, chúng tôi chọn mô hình YOLO V5 để phân tích vì điểm số ở hầu hết các process đều cao hơn hai mô hình còn lại.

Process 1 + Process 5

Hai phương pháp CLAHE (Process 1) và Histogram equalization (Process 5) là hai phương pháp cải thiện độ tương phản của hình ảnh, vì vậy chúng tôi chọn bộ data có kết quả cao hơn để tiếp tục phân tích. Ở đây có một số nhãn có điểm mAP50 và mAP50-95 khá thấp như Calcification và Other lesion ở thang đo mAP50 (dưới 0.2) và Pulmonary fibrosis, Lung Opacity, ILD, Pleural thickening, Other lesion, Calcification ở thang đo mAP50-95 (dưới 0.01).

| PROCESS 5 V5 | | | | | |
|--------------------|-----------|-------|-------|-------|----------|
| Class | Instances | P | R | mAP50 | mAP50-95 |
| Aortic enlargement | 757 | 0.744 | 0.435 | 0.673 | 0.384 |
| Cardiomegaly | 568 | 0.833 | 0.407 | 0.524 | 0.348 |
| Pleural effusion | 219 | 0.515 | 0.402 | 0.475 | 0.174 |
| Consolidation | 49 | 0.437 | 0.327 | 0.347 | 0.127 |
| All | 3528 | 0.444 | 0.317 | 0.323 | 0.147 |
| Atelectasis | 24 | 0.745 | 0.25 | 0.32 | 0.132 |
| Infiltration | 105 | 0.419 | 0.333 | 0.309 | 0.123 |
| Node/Mass | 259 | 0.401 | 0.293 | 0.284 | 0.143 |
| Pulmonary fibrosis | 413 | 0.302 | 0.31 | 0.259 | 0.0874 |
| Lung Opacity | 233 | 0.335 | 0.318 | 0.256 | 0.098 |
| Pneumothorax | 16 | 0.268 | 0.312 | 0.254 | 0.145 |
| ILD | 89 | 0.331 | 0.266 | 0.237 | 0.095 |
| Pleural thickening | 453 | 0.329 | 0.289 | 0.233 | 0.0704 |
| Other lesion | 212 | 0.313 | 0.222 | 0.185 | 0.0715 |
| Calcification | 131 | 0.239 | 0.275 | 0.172 | 0.0664 |

Hình 12: Kết quả của các nhãn trên bộ process 5 với YOLO V5.

Việc cải thiện độ tương phản thông qua việc cân bằng histogram có thể làm nổi bật các chi tiết của ảnh nhưng nhược điểm là làm mất đi hoặc gây nhiễu các pixel mang giá trị cao trong việc chuẩn đoán bệnh. Ở bộ dữ liệu này, ảnh gốc có thể đã có độ tương phản tốt ban đầu, hoặc chỉ số tương phản chưa phù hợp nên việc áp dụng tăng tương phản có thể làm mất hoặc nhiễu đi những chi tiết nhỏ, đặc biệt trong hai lớp Calcification, Other Lesion. Nhưng việc bù trừ cho nhau những ưu và nhược điểm thì phương pháp này không có độ hiệu quả cũng

như chênh lệch đáng kể so với bộ dữ liệu gốc.

Process 2

Ở phương pháp này, hai nhãn có điểm thấp nhất theo thang điểm mAP50 là ILD và Other lesion. Với Bilateral Filtering, phương pháp này làm cho ảnh mờ đi để khử nhiễu. Với ILD, hình ảnh có sự tăng đục của phổi, tăng mờ của các vùng phổi, và hiện diện của các vết sẹo hoặc cấu trúc dày đặc bất thường trong mô nội mỏ phổi. Các hộp chữ nhật của nhãn ILD thường lớn, bao xung quanh phổi, vì vậy khi gia tăng độ làm mờ, hình ảnh của bệnh cũng bị mờ đi trong khi ảnh gốc chưa qua nhiều xử lý nên việc khử nhiễu trên ảnh gốc cũng không mang lại nhiều ý nghĩa. Những lý do trên khiến cho nhãn ILD có kết quả thấp trong phương pháp tiền xử lý này.

| PROCESS 2 VS | | | | | |
|--------------------|-----------|-------|-------|-------|----------|
| Class | Instances | P | R | mAP50 | mAP50-95 |
| Aortic enlargement | 757 | 0.729 | 0.461 | 0.686 | 0.391 |
| Cardiomegaly | 568 | 0.809 | 0.407 | 0.551 | 0.358 |
| Pleural effusion | 219 | 0.427 | 0.411 | 0.456 | 0.18 |
| Consolidation | 49 | 0.437 | 0.306 | 0.354 | 0.146 |
| All | 3528 | 0.422 | 0.317 | 0.314 | 0.147 |
| Infiltration | 105 | 0.392 | 0.313 | 0.289 | 0.107 |
| Atelectasis | 24 | 0.485 | 0.25 | 0.286 | 0.135 |
| Pneumothorax | 16 | 0.367 | 0.312 | 0.272 | 0.159 |
| Nodule/Mass | 259 | 0.474 | 0.251 | 0.262 | 0.121 |
| Lung Opacity | 233 | 0.359 | 0.339 | 0.254 | 0.0992 |
| Pulmonary fibrosis | 413 | 0.267 | 0.337 | 0.247 | 0.0784 |
| Pleural thickening | 453 | 0.267 | 0.306 | 0.222 | 0.07 |
| Calcification | 131 | 0.3 | 0.313 | 0.201 | 0.0759 |
| ILD | 89 | 0.351 | 0.268 | 0.183 | 0.0787 |
| Other lesion | 212 | 0.237 | 0.168 | 0.139 | 0.0539 |

Hình 13: Kết quả của các nhãn trên bộ process 2 với YOLO V5.

Process 3

Dây là phương pháp mang lại hiệu quả lớn nhất trong số năm phương pháp được thực nghiệm. Điều này có thể lý giải rằng Unsharp masking chỉ áp dụng hiệu chỉnh trên các thành phần cường độ cao, như biên cạnh và chi tiết. Các vùng trơn trong ảnh sẽ không bị ảnh hưởng nhiều, do đó, Unsharp masking giúp cải thiện chất lượng ảnh mà không làm mất thông tin quan trọng.

| PROCESS 3 VS | | | | | |
|--------------------|-----------|-------|-------|-------|----------|
| Class | Instances | P | R | mAP50 | mAP50-95 |
| Aortic enlargement | 757 | 0.773 | 0.445 | 0.677 | 0.387 |
| Cardiomegaly | 568 | 0.75 | 0.414 | 0.539 | 0.361 |
| Pleural effusion | 219 | 0.424 | 0.406 | 0.431 | 0.163 |
| Pneumothorax | 16 | 0.418 | 0.375 | 0.37 | 0.193 |
| Consolidation | 49 | 0.492 | 0.336 | 0.338 | 0.155 |
| All | 3528 | 0.433 | 0.321 | 0.322 | 0.152 |
| ILD | 89 | 0.478 | 0.247 | 0.279 | 0.115 |
| Infiltration | 105 | 0.537 | 0.221 | 0.278 | 0.11 |
| Pulmonary fibrosis | 413 | 0.311 | 0.337 | 0.275 | 0.091 |
| Lung Opacity | 233 | 0.386 | 0.352 | 0.269 | 0.103 |
| Atelectasis | 24 | 0.357 | 0.292 | 0.259 | 0.141 |
| Nodule/Mass | 259 | 0.37 | 0.305 | 0.254 | 0.12 |
| Pleural thickening | 453 | 0.312 | 0.3 | 0.223 | 0.0699 |
| Calcification | 131 | 0.265 | 0.328 | 0.193 | 0.0758 |
| Other lesion | 212 | 0.187 | 0.142 | 0.118 | 0.0464 |

Hình 14: Kết quả của các nhãn trên bộ process 3 với YOLO V5.

So với các phương pháp tiền xử lý khác như CLAHE (Contrast Limited Adaptive Histogram Equalization) hay Histogram equalization, Unsharp masking có thể giữ lại các đặc trưng và thông tin quan trọng trong ảnh gốc một cách tốt hơn, đồng thời tạo ra ảnh với độ tương phản cao hơn và chi tiết rõ ràng hơn.

Process 4

Process 4 cho kết quả rất thấp trên các mô hình. Đầu tiên là việc chọn ngưỡng phù hợp để phân chia các pixel có thể chưa phù hợp. Tiếp đến là việc Thresholding Transformation chuyển đổi ảnh xám sang ảnh nhị phân, gây ra mất mát thông tin về các chi tiết nhỏ cũng như đường biên của các đối tượng trong ảnh cũng gây khó khăn cho các mô hình trong quá trình huấn luyện và dự đoán. Qua đó, có thể thấy được việc chỉ thuần áp dụng phương pháp này là không hiệu quả mà đòi hỏi phải kết hợp một số phương pháp khác như histogram equalization hay unsharp masking để có thể giữ lại được nhiều chi tiết cho ảnh hơn.

| PROCESS 4 VS | | | | | |
|--------------------|-----------|-------|--------|--------|----------|
| Class | Instances | P | R | mAP50 | mAP50-95 |
| Aortic enlargement | 757 | 0.695 | 0.451 | 0.655 | 0.345 |
| Cardiomegaly | 568 | 0.675 | 0.426 | 0.555 | 0.309 |
| Pleural effusion | 219 | 0.257 | 0.324 | 0.24 | 0.084 |
| Pneumothorax | 16 | 1 | 0.177 | 0.203 | 0.091 |
| All | 3528 | 0.35 | 0.198 | 0.188 | 0.0859 |
| Atelectasis | 24 | 0.658 | 0.167 | 0.184 | 0.0771 |
| Consolidation | 49 | 0.221 | 0.127 | 0.157 | 0.0763 |
| Lung Opacity | 233 | 0.211 | 0.18 | 0.121 | 0.0401 |
| ILD | 89 | 0.174 | 0.247 | 0.114 | 0.0385 |
| Infiltration | 105 | 0.197 | 0.19 | 0.1 | 0.0408 |
| Pleural thickening | 453 | 0.165 | 0.135 | 0.0857 | 0.0239 |
| Pulmonary fibrosis | 413 | 0.165 | 0.145 | 0.0848 | 0.0254 |
| Nodule/Mass | 259 | 0.222 | 0.0927 | 0.0648 | 0.0261 |
| Other lesion | 212 | 0.139 | 0.0755 | 0.0476 | 0.0175 |
| Calcification | 131 | 0.115 | 0.0399 | 0.0226 | 0.00805 |

Hình 15: Kết quả của các nhãn trên bộ process 4 với YOLO V5.

Nhìn chung, sự ảnh hưởng của các phương pháp còn phụ thuộc vào đặc điểm, đặc trưng của từng loại nhãn có tương thích với phương pháp tiền xử lý hay không. Bởi vì điểm mAP50 hay mAP50-95 là điểm trung bình tất cả các nhãn (không phân biệt số lượng có trong mỗi nhãn) nên có những nhãn với số lượng ít hoặc không phù hợp với phương pháp dẫn đến độ chính xác tổng thể của bộ dữ liệu không vượt trội so với những hình ảnh không xử lý.

8 KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Trong nghiên cứu này, chúng tôi đã tiến hành thử nghiệm trên ba mô hình là YOLov5, YOLov8 và Faster R-CNN, cùng với năm phương pháp tiền xử lý gồm Histogram equalizer, CLAHE, Bilateral Filtering, Unsharp masking và Thresholding Transformations. Mục tiêu của chúng tôi là đánh giá tính hiệu quả của các phương pháp tiền xử lý này trên bộ dữ liệu VinDr CXR.

Kết quả thu được từ các thử nghiệm của chúng tôi cho thấy không có phương pháp tiền xử lý nào hoàn toàn hiệu quả trên bộ dữ liệu VinDr CXR. Mặc dù một số phương pháp cho thấy có những cải thiện ở một số khía cạnh, nhưng về

tổng thể, hiệu suất của các mô hình không được cải thiện bởi các phương pháp tiền xử lý đã sử dụng. Điều này cho thấy rằng các phương pháp đã chọn có thể không phù hợp cho bộ dữ liệu và nhiệm vụ cụ thể này.

Trong tương lai, chúng tôi cần cải tiến độ tin cậy của việc kết luận tính hiệu quả và phân tích các phương pháp tiền xử lý cải tiến hơn. Điều này có thể đạt được thông qua việc thử nghiệm trên nhiều tham số khác nhau và sử dụng các phương pháp đánh giá hiệu suất chính xác và toàn diện hơn. Đồng thời, cần tiếp tục cập nhật các nghiên cứu về công nghệ mới và xu hướng trong lĩnh vực tiền xử lý ảnh, để áp dụng những phương pháp tiền xử lý tiên tiến nhất và có hiệu quả nhất cho công việc phân tích ảnh X-quang. Bằng cách giải quyết những khía cạnh này, chúng ta có thể tiến xa hơn trong lĩnh vực phân tích X-quang nói chung và X-quang ngực nói riêng. Dòng góp vào việc phát triển các công cụ chẩn đoán chính xác và hiệu quả hơn cho hệ thống chăm sóc sức khỏe.

Tài liệu

1. Brunese, L., Mercaldo, F., Reginelli, A., Santone, A.: Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays. Computer Methods and Programs in Biomedicine **196**, 105608 (2020). <https://doi.org/https://doi.org/10.1016/j.cmpb.2020.105608>, <https://www.sciencedirect.com/science/article/pii/S0169260720314413>
2. Cohen, J.P., Morrison, P., Dao, L.: Covid-19 image data collection. arXiv preprint arXiv:2003.11597 (2020)
3. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International journal of computer vision **88**, 303–338 (2010)
4. Gielczyk, A., Marciniak, A., Tarczewska, M., Lutowski, Z.: Pre-processing methods in chest x-ray image classification. PLOS ONE **17**, 1–11 (04 2022). <https://doi.org/10.1371/journal.pone.0265949>, <https://doi.org/10.1371/journal.pone.0265949>
5. Girshick, R.: Fast r-cnn. arxiv 2015. arXiv preprint arXiv:1504.08083 (2015)
6. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 580–587 (2014)
7. He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE transactions on pattern analysis and machine intelligence **37**(9), 1904–1916 (2015)
8. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haghgoo, B., Ball, R., Shpanskaya, K., et al.: Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In: Proceedings of the AAAI conference on artificial intelligence. vol. 33, pp. 590–597 (2019)
9. Jocher, G.: YOLOv5 by Ultralytics. <https://doi.org/10.5281/zenodo.3908559>, <https://github.com/ultralytics/yolov5>
10. Jocher, G., Chaurasia, A., Qiu, J.: YOLO by Ultralytics, <https://github.com/ultralytics/ultralytics>
11. Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J., Yang, J.: Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection. Advances in Neural Information Processing Systems **33**, 21002–21012 (2020)

12. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part V 13. pp. 740–755. Springer (2014)
13. Maini, R., Aggarwal, H.: A comprehensive review of image enhancement techniques (2010)
14. Mao, Y., Xue, F.F., Wang, R., Zhang, J., Zheng, W.S., Liu, H.: Abnormality detection in chest x-ray images using uncertainty prediction autoencoders. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part VI 23. pp. 529–538. Springer (2020)
15. Nguyen, H.Q., Lam, K., Le, L.T., Pham, H.H., Tran, D.Q., Nguyen, D.B., Le, D.D., Pham, C.M., Tong, H.T.T., Dinh, D.H., Do, C.D., Doan, L.T., Nguyen, C.N., Nguyen, B.T., Nguyen, Q.V., Hoang, A.D., Phan, H.N., Nguyen, A.T., Ho, P.H., Ngo, D.T., Nguyen, N.T., Nguyen, N.T., Dao, M., Vu, V.: Vindr-cxr: An open dataset of chest x-rays with radiologist’s annotations (2020)
16. Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Rajendra Acharya, U.: Automated detection of covid-19 cases using deep neural networks with x-ray images. Computers in Biology and Medicine **121**, 103792 (2020). <https://doi.org/https://doi.org/10.1016/j.combiomed.2020.103792>, <https://www.sciencedirect.com/science/article/pii/S0010482520301621>
17. Rajaraman, S., Kim, I., Antani, S.K.: Detection and visualization of abnormality in chest radiographs using modality-specific convolutional neural network ensembles. PeerJ (2020)
18. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems **28** (2015)
19. Shih, G., Wu, C.C., Halabi, S.S., Kohli, M.D., Prevedello, L.M., Cook, T.S., Sharma, A., Amorosa, J.K., Arteaga, V., Galperin-Aizenberg, M., et al.: Augmenting the national institutes of health chest radiograph dataset with expert annotations of possible pneumonia. Radiology: Artificial Intelligence (2019)
20. Terven, J., Cordova-Esparza, D.: A comprehensive review of yolo: From yolov1 to yolov8 and beyond. arXiv preprint arXiv:2304.00501 (2023)
21. Wang, C.Y., Bochkovskiy, A., Liao, H.Y.M.: Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 7464–7475 (2023)
22. Wang, C.Y., Liao, H.Y.M., Wu, Y.H., Chen, P.Y., Hsieh, J.W., Yeh, I.H.: CspNet: A new backbone that can enhance learning capability of cnn. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. pp. 390–391 (2020)
23. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (July 2017)
24. Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., Ren, D.: Distance-iou loss: Faster and better learning for bounding box regression. In: Proceedings of the AAAI conference on artificial intelligence. vol. 34, pp. 12993–13000 (2020)

A Mô tả các nhãm

Ở mục này chúng tôi đưa ra các tên nhãm bằng tiếng Việt và mô tả về định nghĩa của các nhãm có trong bộ dữ liệu

Bảng 4: Mô tả các nhãm.

| Nhãm Tiếng Anh | Nhãm Tiếng Việt | Mô Tả |
|---------------------------|---------------------|---|
| Aortic enlargement | Phình động mạch | Một khuyết tật được đặc trưng bởi sự rộng hơn bình thường của động mạch phổi chính |
| Cardiomegaly | Phì đại tim | Kích thước của tim to hơn bình thường ở người lớn và tỷ lệ tim phổi lớn hơn 0,5. |
| Pleural thickening | Dày màng phổi | Bất kỳ dạng dày màng nào liên quan đến màng phổi vách hoặc màng phổi nội tạng. |
| Pulmonary fibrosis | Seo phổi | Sự dư thừa của mô seo trong phổi. |
| Nodule/Mass | Khối u/ hạch | Bất kỳ khối u chiếm không gian nào đơn lẻ hoặc nhiều lẻ. |
| Lung opacity | Sự mờ phổi | Bất kỳ sự mờ tập trung hoặc tổng thể không bình thường nào trong các lĩnh vực phổi (thể mở rộng bao gồm nhưng không giới hạn đến sự kết dính, khối, khối u, vôi hóa, sự dày thêm liên kết, v.v.). |
| Pleural effusion | Tràn dịch màng phổi | Sự tích tụ dịch lỏng trong không gian màng phổi. |
| Other lesion | Bệnh lý khác | Các bệnh lý khác không có trong danh sách các phát hiện hoặc bất thường được đề cập ở trên. |
| Calcification | Vôi hoá | Tích tụ muối canxi trong phổi. |
| Infiltration | Sự thâm nhập | Chất bất thường tích tụ dồn trong tế bào hoặc mô cơ thể hoặc bất kỳ chất hoặc loại tế bào nào xuất hiện trong hoặc lan truyền như qua những khoảng trống (không gian liên kết và / hoặc bìa phổi) của phổi, là bất thường đối với phổi hoặc tích tụ nhiều hơn bình thường trong đó. |

Bảng 4 – Tiếp tục từ trang trước

| Nhân Tiếng Anh | Nhân Tiếng Việt | Mô Tả |
|--|--------------------|---|
| Interstitial lung disease (ILD) | Bệnh phổi nội mô | Sự tác động đến mô hỗ trợ của phổi gây ra sự xuyên thấu tinh thể mịn hoặc thô hoặc các khối nhỏ. |
| Consolidation | Phổi đông đặc | Bất kỳ quá trình bệnh lý nào làm cho bọng phế quản được lắp đầy với chất lỏng, mủ, máu, tế bào (bao gồm tế bào khối u) hoặc các chất khác dẫn đến vùng thấu quang phân khu, phân tán hoặc đa trung tâm không rõ ràng. |
| Atelectasis | Phổi rỗng | Sụp một phần của phổi do giảm lượng không khí trong bọng phế quản dẫn đến mất thể tích và tăng mật độ. |
| Pneumothorax | Tràn khí màng phổi | Sự hiện diện của khí (không khí) trong không gian màng phổi. |