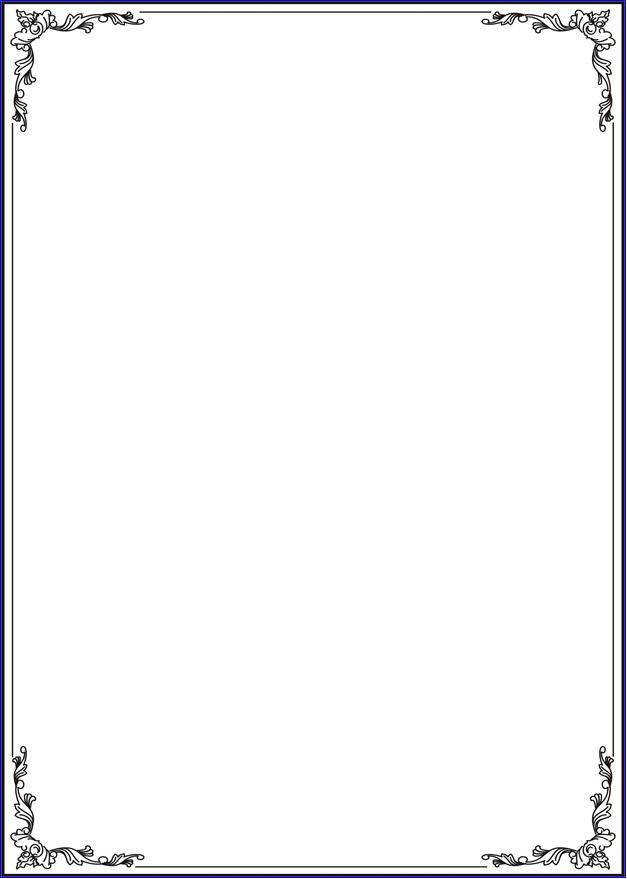
** BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI**

**PHÂN HIỆU TẠI THÀNH PHỐ HỒ CHÍ MINH**



**BÁO CÁO KẾT QUẢ NGHIỆM THU CỦA QUÁ TRÌNH NGHIÊN CỨU KHOA HỌC SINH VIÊN**

**ĐỀ TÀI: NHẬN DẠNG VÀ PHÁT HIỆN CÁC ĐIỂM BẤT THƯỜNG, TỔN THƯƠNG TRÊN NỀN ẢNH X-RAY Ở PHỔI**

**GVHD: Th.S. Phạm Thị Miên**

**Ph.D Nguyễn Đình Hiển**

TP.Hồ Chí Minh , ngày 12 tháng 05 năm 2022

**TABLE OF CONTENT**

[UCHEX-UNET: SOLUTION TO IMPROVE IDENTIFICATION AND DETECTING ABNORMALITIES, LESIONS ON LUNG X-RAYS 3](#_Toc18118)

[LỜI NÓI ĐẦU 4](#_Toc29112)

[ABSTRACT 6](#_Toc19393)

[I. INTRODUCTION 7](#_Toc18229)

[II. RELATED WORK 10](#_Toc10655)

[III. METHODOLOGY 11](#_Toc28603)

[1. Baseline Model: 11](#_Toc21064)

[2. Improve method: 14](#_Toc21807)

[IV. EXPERIMENT AND RESULT 16](#_Toc12608)

[1. Datasets: 16](#_Toc367)

[2. Data pre-processing: 18](#_Toc27567)

[3. Implementation and Metrics: 18](#_Toc17594)

[4. Result: 24](#_Toc14758)

[5. Future work: 26](#_Toc17510)

[V. IMPROVE THE FUTURE WORK 27](#_Toc27276)

[VI. CONCLUSION 28](#_Toc23581)

[VII. REFERENCE 29](#_Toc30157)

**UCHEX-UNET: SOLUTION TO IMPROVE IDENTIFICATION AND DETECTING ABNORMALITIES, LESIONS ON LUNG   
X-RAYS**

***Authors*: Nhan T. HUYNHa, Hien D. NGUYENb,c, Nguyen D.N. TRANc,d,  
 Mien T. PHAMa, Anh M.T. TRANa, Quan P. VOa, Anh D. TRANa**

**a University of Transport and Communications, Campus in Ho Chi Minh City, Vietnam bUniversity Information and Technology c** **Vietnam National University, Ho Chi Minh city, Vietnam**

**d** **Faculty of Mathematics and Computer Science, University of Sciences, Ho Chi Minh  
city, Vietnam**

***Keyword.*** *Deep Learning, Lesion segmentation, U-net  
model, Chest image, Pneumonia detection, ChexNet, Medical image, X-Rays, Diagnosis pneumothorax, Yolov5 for medical image*

**LỜI NÓI ĐẦU**

Sáng tạo và đổi mới là quy tắc mà thế giới này vận hành, vì vậy để tránh không bị xã hội đào thải, chúng ta phải luôn học tập nâng cao kỹ năng, không ngừng cải tiến tư duy trong và ngoài lĩnh vực mà chúng ta đang tìm hiểu như một giải pháp để ta có thể nâng cao năng lực trong bản thân.

Nghiên cứu tìm tòi học hỏi có lẽ là một quá trình lâu dài gắn liền với cuộc đời của mỗi con người chúng ta. Những đại dương sâu thẩm, những căn bệnh siêu hiếm, một vũ trụ bí ẩn và muôn vàn thứ khác xung quanh chúng ta, hàng tỷ tỷ điều chúng ta cần tìm hiểu và sau cùng những thứ ấy luôn đời chúng ta nghiên cứu và khám phá. Nói như vậy để thấy được sự quan trọng và lớn lao mà nghiên cứu khoa học đem đến. Nó đã, đang và sẽ thay đổi và thúc đẩy nền văn minh của nhân loại

The World is flat của Thomas Friedman mà một trong những minh chứng rõ ràng cho việc nghiên cứu khoa học đã tác động trực tiếp vào thế giới quan của con người – Quyển sách như một lời khẳng định rằng chúng ta đã và đang sống trong một kỷ nguyên mới, một kỷ nguyên mà con người có thể xích lại gần nhau hơn, một kỷ nguyên của thông tin và số. Như thế bằng cách nhìn tổng quan trong sự vận động, phát triển, đổi mới và đào thải của thế giới có thể thấy công nghệ thông tin hiện đang là ngành có ảnh hưởng trực tiếp vào trong đời sống cũng như công việc hàng ngày của chúng ta

Như vậy việc nghiên cứu khoa học trong lĩnh vực Công nghệ thông tin là một việc làm hết sức cần thiết và có ý nghĩa to lớn. Với những nghiên cứu đi đôi với ứng dụng tốt sẽ làm nên những thay đổi cho cuộc sống của con người, đẩy mạnh và bắt kịp quá trình thông tin hóa đang diễn ra hiện nay của thế giới

Bài luận này được nhóm chúng tôi là UTC2@iLab viết nhằm tổng kết, đúc rút cũng như thực hành việc nghiên cứu dựa trên lý thuyết được truyền giảng trên lớp và quá trình học hỏi, tìm tòi bán chuyên sâu của chúng tôi ngoài thực tế. Qua đây, chúng tôi xin cảm ơn chân thành đến Th.S Phạm Thị Miên, Y.S Huỳnh Thị Cúc cùng với T.S Nguyễn Đình Hiển đã tận tình dạy bảo, giải đáp những thắc mắc của chúng tôi trong quá trình tìm hiểu và nghiên cứu các vấn đề. Tôi xin dành lời cảm ơn đặc biệt đến Th.S Phạm Thị Miên,Y.S Huỳnh Thị Cúc, Anh Trần Ngọc Đăng Nguyên, Th.S Trần Thị Dung và T.S Nguyễn Đình Hiển đã có những ý kiến quan tâm, đóng góp luận điểm cá nhân và đề xuất ý tưởng cải tiến đến đề tài nghiên cứu của nhóm chúng tôi.

Nội dung bài luận ngoài việc tham khảo tài liệu, nội dung hoàn toàn mang quan điểm của cá nhân, viết theo văn phong cá nhân. Do đó sẽ không thể tránh khỏi những sai sót không muốn. Rất mong quý hội đồng nghiên cứu tích cực nhận xét để cải thiện.

Đại diện nhóm sinh viên thực hiện

Trưởng nhóm

Huỳnh Trọng Nhân

**ABSTRACT**

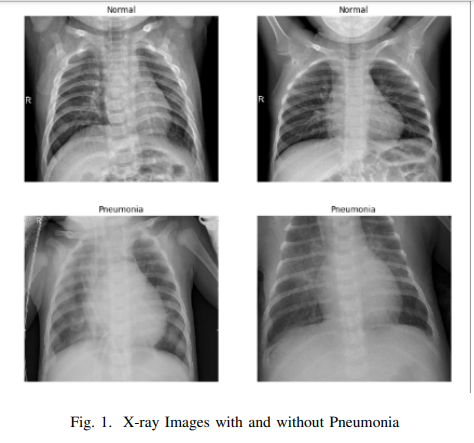
Viêm phổi là một trong những bệnh gây tử vong và có khả năng dẫn đến hậu quả nghiêm trọng trong thời gian ngắn khoảng thời gian, do sự tràn dịch màng bên trong phổi, dẫn đến việc thiếu hô hấp ở phổi. Nếu không có tác động hay sự giải pháp xử lý vào đúng thời điểm, viêm phổi có thể dẫn đến tử vong của các cá nhân. Do đó, chẩn đoán sớm là yếu tố then chốt cùng với sự tiến triển của bệnh[23].

Bài báo cáo này tập trung vào tiến trình sinh học của bệnh viêm phổi và phát hiện nó bằng hình ảnh tia x, tổng quan các nghiên cứu được thực hiện nâng cao mức độ chẩn đoán và trình bày phương pháp và kết quả của quá trình tự động hóa hình ảnh X-quang dựa trên các các thông số nhằm phát hiện bệnh ở giai đoạn rất sớm. Trong nghiên cứu này, chúng tôi đề xuất kiến trúc học sâu của chúng tôi cho nhiệm vụ phân loại, được đào tạo với các hình ảnh đã sửa đổi, thông qua nhiều bước tiền xử lý. Phương pháp phân loại của chúng tôi sử dụng mạng nơ ron phức hợp và kiến trúc mạng dày đặc phân loại các hình ảnh. Phát hiện của chúng tôi mang lại độ chính xác về 84.24%, chỉ xếp sau các mô hình Y.O.L.Ov5 và Detectron.

1. **INTRODUCTION**

Trên toàn cầu, 450 triệu người mắc bệnh viêm phổi trong một năm và 4 triệu người chết vì căn bệnh này. 1 triệu người mỗi năm phải đến bệnh viện chăm sóc và 50 nghìn người chết vì căn bệnh này ở Hoa Kỳ[26]. Sự khác biệt về số lượng giữa tỷ lệ lây nhiễm và tỷ lệ tử vong cho thấy mức độ quan trọng của việc chẩn đoán sớm căn bệnh này. Viêm phổi là một phản ứng viêm hay còn gọi là nhiễm trùng trong túi phổi thường gọi là phế nang. Nó thường do vi khuẩn, vi rút, nấm và các vi khuẩn khác gây ra. Khi vi rút đến phổi, các tế bào bạch cầu hoạt động chống lại vi trùng và tình trạng viêm xảy ra trong các túi khí của phổi. Do đó, các phế nang chứa đầy dịch viêm phổi và dịch này gây ra các triệu chứng như ho, khó thở và sốt. Nếu nhiễm trùng không được can thiệp và xử lý trong thời gian đầu của bệnh, nhiễm trùng viêm phổi có thể lây lan khắp cơ thể và dẫn đến việc tử vong cho cá nhân người bệnh, do không có khả năng trao đổi khí ở phổi. Ngày nay, một trong những kỹ thuật y tế thông thường nhất được sử dụng để chẩn đoán bệnh là chụp X-quang phổi. Khi chùm electron tập trung, được gọi là tia photon, tia X, đi qua các mô của cơ thể, một hình ảnh được tạo ra trên bề mặt nhựa ảnh (ảnh phim)[9,10]. Trong quá trình chẩn đoán, các bác sĩ chuyên môn về X quang ngực tương ứng với các đốm trắng và điểm bất thường trên nền của hình ảnh để xác định vùng nhiễm trùng và các vùng trắng đối với dịch viêm phổi trong phổi[36]. Tuy nhiên, bảng màu hạn chế của hình ảnh X-quang bao gồm các sắc độ đen và trắng, gây ra những hạn chế khi xác định xem có khu vực bị nhiễm trùng trong phổi hay không.

Điều này là do thực tế là cường độ cao của bước sóng trắng xảy ra trên phim ảnh khi chất lỏng trong phổi đủ cao để được coi là một mô đặc và rắn. Nói cách khác, sự chuyển đổi từ một mô chứa đầy không khí (trạng thái bình thường của phổi), được nhìn thấy ở các sắc thái tối hơn, sang một mô dày đặc, đòi hỏi phải có đủ lượng chất lỏng để chuyển bảng màu sang màu sáng hơn. Điều này có nghĩa là để một bác sĩ chuyên khoa X-ray ngực có thể chẩn đoán được thì bệnh viêm phổi phải ở giai đoạn sau[24]. Do đó, việc phát hiện sớm bệnh viêm phổi bị hạn chế do hạn chế màu sắc của hình ảnh X quang. Một hạn chế khác đối với việc chẩn đoán sớm bệnh viêm phổi là việc phát hiện phụ thuộc vào con người. Các bác sĩ X quang chuyên nghiệp cần có đôi mắt được đào tạo đầy đủ để có thể phân biệt giữa sự phân bố màu sắc không đồng nhất của không khí khi đang chảy trong phổi. Điều này có thể được nhìn thấy với các màu sắc khác nhau trên hình ảnh chụp X-quang, nhưng không nó là hoàn toàn không có tác dụng với viêm phổi kết đặc dịch màng phổi[25].



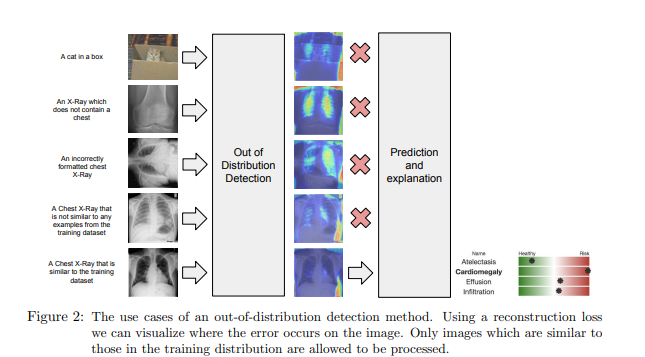
Do đó, việc bác sĩ X quang có thể biết được liệu các đốm trắng trên phim X quang có thực sự tương ứng với chất lỏng hay không. Do biên độ sai số của mắt người, có nhiều trường hợp các bác sĩ X quang không đưa ra được chẩn đoán chính xác. Trong cả hai trường hợp, cho dù đó là chẩn đoán dương tính giả hay âm tính giả, nó đều có những tác động đáng kể đến cơ thể con người. Vì vậy, các phương pháp tính toán trong bước chẩn đoán bệnh là đáng tin cậy về tính nhất quán. Trong ***Fig 1.*** có thể thấy các hình ảnh khác nhau có và không có viêm phổi. Người ta cũng có thể chứng kiến sự không nhạy cảm của hình ảnh người khỏe mạnh so với hình ảnh viêm phổi, điều này cho thấy sự cần thiết của đôi mắt được đào tạo tốt để có thể phân biệt. Đã có những nghiên cứu trước đây được thực hiện liên quan đến việc phát hiện bệnh viêm phổi bằng chụp X-quang ngực thông qua máy học với việc sử dụng bản đồ nhiệt, là hình ảnh hoặc bản đồ biểu thị nhiệt độ thay đổi hoặc bức xạ hồng ngoại được ghi lại trong một khu vực hoặc trong một khoảng thời gian, và phân biệt bệnh lý phổi, là chuyên ngành phụ của bệnh lý ngoại khoa nhằm chẩn đoán và xác định đặc điểm của các bệnh ung thư và không phải ung thư của phổi so với bình thường bằng cách sử dụng máy tính phân tích âm thanh phổi.

1. **RELATED WORK**

Trong số các phương pháp khác được sử dụng bởi các nghiên cứu khác nhau, bài báo cáo này là bài báo cáo đầu tiên trình bày sự tự động hóa của các mô hình khác nhau để chuẩn đoán trên hình ảnh X-quang trong quy mô nghiên cứu học thuật sinh viên mà có thể chẩn đoán viêm phổi ở giai đoạn rất sớm. Trong khi các phương pháp thông thường và phóng xạ đã đề cập có thể có hiệu quả, nghiên cứu của nhóm chúng tôi trình bày một cách tiếp cận học sâu phân loại viêm phổi này. Trong khi nghiên cứu tôi tập trung vào 14 bệnh. Tuy nhiên, do thử nghiệm để chiết xuất 14 loại bệnh khác nhau với một mô hình chỉ có thể đạt được độ chính xác 71,3%. Hơn nữa, LSTM sử dụng nhiều hình ảnh để phân loại một hình ảnh duy nhất, trong khi thử nghiệm và mô hình được đề xuất của chúng tôi chỉ cần mạng nơ-ron được đào tạo trước trọng số để phân loại từng ảnh một. Ngoài ra, của chúng tôi độ chính xác khi thử nghiệm mang lại 78,73%, sử dụng cùng một tập dữ liệu.

Bài báo cáo này trình bày một phương pháp phát hiện sử dụng các bước bổ sung sau thuật toán phân loại. Như một Convolutional thông thường mạng nơ-ron phân loại toàn bộ hình ảnh, các phương pháp phân đoạn như vậy sử dụng phân loại pixelwise, cuối cùng, áp dụng lớp không hợp lệ để phân loại từng pixel một bởi một và cuối cùng tách các đối tượng khác nhau trong một hình ảnh, xương là những thứ phổ biến nhất cho những người đã đề cập nhiệm vụ. Trong một nghiên cứu khác nhằm tiến hành phát hiện sớm không phải cho bệnh viêm phổi mà cho bệnh lồng ngực qua hàng tuần phân loại với mạng nơ-ron tích tụ. Bài báo cáo này phát hiện thành công các mẫu cho bệnh nhân có lồng ngực bệnh hoặc một trong những có thể có bệnh đã đề cập. Chưa, không có hoạt động về phân loại viêm phổi. Trong nghiên cứu này, chúng tôi trình bày một phương pháp mới để phân loại sự tồn tại của bệnh viêm phổi trong hình ảnh chụp X-quang. Chúng tôi đề xuất một hai bước xử lý hình ảnh trước khi đào tạo mô hình học sâu của chúng tôi, trong để làm cho các đặc điểm của hình ảnh X quang rõ ràng hơn và rõ ràng để giảm bớt quá trình phân loại. Chúng tôi, sau đó, thực hiện một mạng nơ-ron phức hợp theo sau bởi một mạng nơ-ron dư mạng cho quá trình phân loại.

Bài báo cáo của chúng tôi trình bày một phương pháp phát hiện các mẫu tổn thương trong khu vực phổi trên nền ảnh X-Rays bằng cách sử dụng kiến trúc Chex\_Net kết hợp phân đoạn tổn thương được xử lý bởi Unet đồng thời được S.O.T.A cùng nhiều kiến khác để tiến hành đo đạc kết quả . Mô hình được đề xuất hiệu quả hơn so với so với kiến trúc Chex\_Net do được sử đào tạo trên kích thước ảnh gốc và được preprocessing rất kỹ [4, 8. 14, 17]. Mô hình này cũng có tiềm năng áp dụng trong các hệ thống nhỏ thực tế.



1. **METHODOLOGY**

Chúng tôi phát triển một mô hình deep learning cho việc phát hiện và phân loại các điểm bất thưởng của các loại bệnh liên quan đến phổi, được mô tả ở ***Figure 3****.* Mô hình này được xây dựng dựa trên kiến trúc ChexNet và S.O.T.A với các kiến trúc phổ biến khác.

1. Baseline Model:

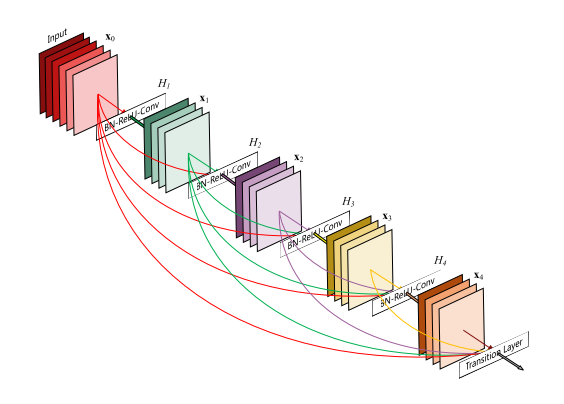
ChexNet[18,19]: là một mạng lưới nơ-ron phức hợp gồm 121-layer Dense Convolutional Network (DenseNet). Đầu vào hình ảnh X-quang ngực và xuất ra xác suất viêm phổi cùng với bản đồ nhiệt khoanh vùng các khu vực của hình ảnh có dấu hiệu viêm phổi nhiều nhất. Chúng tôi đào tạo CheX\_Net trên bộ dữ liệu Vin\_Big\_Data, Chex2pex, NIH được phát hành gần đây, chứa 15.000 hình ảnh X-quang ngực nhìn từ trước được gắn nhãn riêng với tối đa 14 bệnh khác nhau ở lồng ngực, bao gồm cả viêm phổi. Dense\_Nets cải thiện luồng thông tin và độ dốc thông qua mạng, làm cho tối ưu hóa các mạng rất sâu có thể điều chỉnh được. Tôi thay thế lớp được kết nối đầy đủ cuối cùng bằng một lớp có một đầu ra duy nhất, sau đó chúng tôi áp dụng tính phi tuyến tính.

Trong kiến trúc Chex\_Net, mỗi lớp được kết nối với mọi lớp khác, do đó có tên là Mạng kết nối dày đặc. Đối với một mạng có n layers thì sẽ có n\*(n+1) /2 kết nối trực tiếp. Đối với mỗi layer, bản đồ đặc trưng của tất cả các layers trước được sử dụng làm đầu vào và bản đồ đặc trưng của chính nó được sử dụng làm đầu vào cho mỗi layers tiếp theo.

Đây thực sự là nó, đơn giản như điều này nghe có vẻ như vậy, về cơ bản, CheX\_Net sẽ liên kết mọi lớp với mọi lớp khác. Đây là ý tưởng chính mà vô cùng mạnh mẽ. Đầu vào của một lớp bên trong DenseNet là sự ghép nối của các bản đồ đối tượng từ các lớp trước đó.

Từ bài báo[3]:

DenseNets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters.

  
Figure 3. Base Architecture of DenseNet. The origin of ChexNet

Độ hiệu quả và tối ưu của ChexNet đã được kiểm chứng trong thực tế. Cụ thể là một competition được tổ chức giữa 4 bác sĩ chuẩn đoán chuyên khoa hô hấp và chuyên khoa nội với kiến trúc ChexNet-121. Số liệu thu được là F1 score trung bình giữa các bác sĩ và predict của mô hình ChexNet

|  |  |
| --- | --- |
|  | F1 Score (95% CI) |
| Radiologist 1 | 0.383 (0.309, 0.453) |
| Radiologist 2 | 0.356 (0.282, 0.428) |
| Radiologist 3 | 0.365 (0.291, 0.435) |
| Radiologist 4 | 0.442 (0.390, 0.492) |
| Radiologist Avg | 0.387 (0.330, 0.442) |
| CheX\_Net-121 | 0.435 (0.387, 0.481) |

Table 1. We compare radiologists and our model on the F1 metric, which is the harmonic average of the precision and recall of the models. CheXNet achieves an F1 score of 0.435 (95% CI 0.387, 0.481), higher than the radiologist average of 0.387 (95% CI 0.330, 0.442). We use the bootstrap to find that the difference in performance is statistically significant[35].

Chúng tôi thấy rằng mô hình này vượt quá hiệu suất trung bình của bác sĩ X quang trong việc phát hiện bệnh viêm phổi nhiệm vụ. Để so sánh CheXNet với công việc trước đó bằng tập dữ liệu ChestX-ray14, tôi thực hiện các sửa đổi đơn giản đối với CheXNet để phát hiện tất cả 14 bệnh trong tập dữ liệu ChestX-ray14 tăng cường thêm VinBigData, và nhận thấy rằng chúng tôi làm tốt hơn các kết quả được xuất bản tốt nhất trên tất cả 14 bệnh. Tự động phát hiện bệnh từ X-quang ngực ở cấp độ chuyên gia X quang sẽ không chỉ có lợi ích to lớn trong môi trường lâm sàng[5,6,7], nó cũng sẽ là vô giá trong việc cung cấp dịch vụ chăm sóc sức khỏe cho các nhóm dân cư không đủ khả năng tiếp cận với chẩn đoán các chuyên gia hình ảnh.

1. Improve method:

Kiến trúc ChexNet121 lần đầu tiên được đề cập trong bài báo nghiên cứu cả khoa học của Đại Học Standford[9] và được trình bày trong hội nghị Medical Image on International trong giới nghiên cứu[5,7]. Được cải tiến từ kiến trúc DenseNet-121 với hàm những điểm đặc trưng như sau trong Convolutional cuối sẽ thay thế lớp được kết nối đầy đủ cuối cùng bằng một lớp có một đầu ra duy nhất, sau đó chúng tôi áp dụng tính hàm Sigmoid phi tuyến tính.:

Như được mô tả ở Table 1. và từ bài báo gốc[4]. tôi tính điểm F1 cho từng bác sĩ X-quang riêng lẻ và cho CheXNet so với từng nhãn trong số 4 nhãn khác là sự thật cơ bản. Chúng tôi báo cáo trung bình của 4 điểm F1 kết quả cho mỗi bác sĩ X-quang và cho CheXNet, cùng với F1 trung bình của các bác sĩ X quang. Chúng tôi so sánh các nhà nghiên cứu X quang và mô hình của chúng tôi trên chỉ số F1, là giá trị trung bình hài hòa của độ chính xác và thu hồi. CheXNet đạt được điểm F1 là 0,435 (KTC 95% 0,387, 0,481), cao hơn mức trung bình của bác sĩ X quang là 0,387 (KTC 95% 0,330, 0,442). Chúng tôi sử dụng bootstrap để thấy rằng sự khác biệt về hiệu suất là có ý nghĩa thống kê.

Trong mô hình trình bày của chúng tôi hầu như cấu trúc của mô hình cơ sở được giữ nguyên để đo đạt state of the art giữa các mô hình từ đó kiểm chọn ra mô hình để phát triển kiến trúc trong dự án tiếp theo, ở đây chúng tôi chọn 2 kiến trúc: một số đó đã được đề cập chi tiết ở trên còn kiến trúc còn lại là Faster R-CNN:

Tổng quan sơ lược về Faster R-CNN thì đây là một Architechture được cải tiến nhiều nhất với tiền thân ban đầu của nó là R-CNN sau thành Fast R-CNN và giờ là Faster R-CNN. Tương ứng với nhiều cải tiến thì Faster R-CNN là phiên bản cho đến hiện tại được coi là tốt nhất về cả 2 khía cạnh là độ chính xác và thời gian thực thị.

Lý thuyết về ý tưởng thuật toán cơ bản của Faster R-CNN : là sử dụng một mạng CNN mới gọi là Region Proposal Network (RPN) để tìm các region proposal.

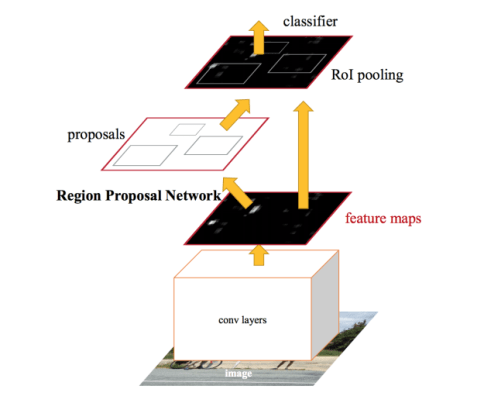


Figure 4. Architechture of Faster R-CNN

Đầu tiên cả bức ảnh được cho qua pre-trained model để lấy feature map. Sau đó feature map được dùng cho Region Proposal Network để lấy được các region proposal. Sau khi lấy được vị trí các region proposal thì thực hiện tương tự Fast R- CNN.

Nhưng mô hình output đầu ra của Faster R-CNN đã được mặc định là một bounding box với kích thước sao cho hệ số IoU 1. Như vậy bài toán là gần như là sự điều chỉnh là sao cho bounding box được predicted phải cận khớp với các Ground-truth bounding box. Hệ số IoU có công thức tổng quát như sau:

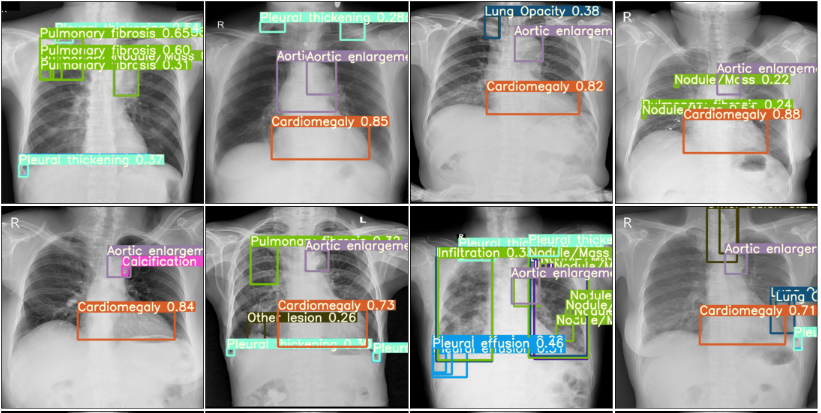


Figure 5. Detection class of pneumonia using by Faster R-CNN architechture

1. **EXPERIMENT AND RESULT**
2. Datasets:

Chúng tôi sử dụng 3 bộ dữ liệu ảnh y khoa ở khu vực lồng ngực[37, 38, 39] để tiến hành phân tích và đào tạo mô hình. Một[37] trong số đó được công khai cho việc nghiên cứu học thuật và đề xuất đánh giá. Còn 2 bộ dữ liệu còn lại được dùng trong 2 cuộc thi học thuật lớn với tổng giá trị giải thưởng lên đến 500.000 USD. Các tập dữ liệu trên nhìn chung đều tương đồng về số phân lớp phân loại trong các tập dữ liệu nên rất thuận tiện cho việc pretrain trên model giữa các bộ dữ liệu để tăng cường thêm trọng số cho mô hình.

CheX-ray14: là một bộ dữ liệu hình ảnh y tế bao gồm 112.120 hình ảnh X-quang nhìn trực diện của 30.805 bệnh nhân duy nhất (được thu thập từ năm 1992 đến 2015) với mười bốn nhãn bệnh phổ biến được khai thác bằng văn bản, được khai thác từ các báo cáo X quang văn bản thông qua kỹ thuật NLP. Nó mở rộng trên ChestX-ray8 bằng cách thêm sáu bệnh khác ở lồng ngực: Phù phổi(Edema), Khí thủng phổi(Emphysema), Xơ nang phế quản(Fibrosis), Dày màng phổi(Pleural thickening) và Thoát vị bẹn(Hernia).

RSNA Pneumonia Detection Datasets: là một cộng đồng quốc tế của các bác sĩ X- quang, nhà vật lý y tế và các chuyên gia y tế khác với hơn 54.000 thành viên đến từ 146 quốc gia trên toàn cầu. Họ nhận thấy tiềm năng ML có thể tự động hóa việc phát hiện ban đầu (sàng lọc hình ảnh) các trường hợp viêm phổi tiềm ẩn để sắp xếp thứ tự ưu tiên và xúc tiến việc xem xét của họ.

VinBigData Chest X-ray Abnormalities Detection: Được thành lập vào tháng 8 năm 2018 và được tài trợ bởi Tập đoàn Vingroup, Viện Dữ liệu lớn của Vingroup (VinBigData) nhằm thúc đẩy nghiên cứu cơ bản và điều tra các công nghệ mới và có tính ứng dụng cao. Viện tập trung vào các lĩnh vực chính của khoa học dữ liệu và trí tuệ nhân tạo: y sinh tính toán, xử lý ngôn ngữ tự nhiên, thị giác máy tính và xử lý hình ảnh y tế. Đội ngũ hình ảnh y tế tại VinBigData thực hiện nghiên cứu thu thập, xử lý, phân tích và tìm hiểu dữ liệu y tế. Họ đang nỗ lực xây dựng các giải pháp hình ảnh y tế quy mô lớn và có độ chính xác cao dựa trên những tiến bộ mới nhất trong trí tuệ nhân tạo để tạo điều kiện cho quy trình làm việc lâm sàng hiệu quả.

Trong cuộc thi này, bạn sẽ tự động khoanh vùng và phân loại 14 loại bất thường ở lồng ngực từ phim chụp X quang phổi. Bạn sẽ làm việc với một tập dữ liệu bao gồm 18.000 lần quét đã được chú thích bởi các bác sĩ X quang có kinh nghiệm. Bạn có thể huấn luyện mô hình của mình với 15.000 hình ảnh được gắn nhãn độc lập và sẽ được đánh giá trên bộ thử nghiệm gồm 3.000 hình ảnh. Những chú thích này được thu thập thông qua nền tảng dựa trên web của VinBigData, VinLab. Thông tin chi tiết về việc xây dựng tập dữ liệu có thể được tìm thấy trong bài báo gần đây của chúng tôi “VinDr-CXR: Một tập dữ liệu mở về chụp X-quang ngực với chú thích của các bác sĩ X quang”.

|  |  |
| --- | --- |
| Datasets | #Cases |
| VinBigData Chest X-ray Abnormalities Detection Datasets | 18.000 DICOM File from 12500 patients |
| CheX-ray 14 Datasets | 112.120 PNG Image File from 30.805 unique patients |
| RSNA Pneumonia Datasets | Total imag in All of stage is nearly 16.000 JPG Image |
| Total | 146.120 Image, it’s include image which was data augumentation |

Table 2. Total number of image on selected datasets

Tiếp theo là việc hiển thị tiệm cận chi tiết thông tin về đối tượng mà ta đang thao tác.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Datasets | Format | Segmentation | Bounding Box | Size |
| VinBigData | DICOM | x | x | 2048 x 2048 |
| RSNA | JPG |  | x | 512 x 512 |
| CheX-ray 14 | PNG |  | x | 512 x 512 |

Table 3. The information of the selected datasets.

1. Data pre-processing:

Với dữ liệu VinBigData đang ở định dạng DICOM cần làm một vài bước chuyển đổi về dạng PNG và resize sao cho cùng kích thước là 512 x 512. Nhưng có một vài bước trước khi xử lý file DICOM đó là phải hiểu phải đọc metadata của dữ liệu để hiểu về tính chất của dữ liệu trước khi tiến hành chuyển đổi.

Còn 2 tập còn lại ta chọn biến đối dữ liệu RSNA theo CheX-ray 14 vì số lượng dữ liệu cùa tập RSNA rất ít nên thuận tiện cho việc giảm tiêu tốn quá nhiều thời gian cho việc tiền xử lý dữ liệu. Hơn nữa CheX-ray 14 là một tập dữ liệu lớn nên để đồng quy dữ liệu này theo RSNA là một việc không cần thiết. Ở đây ta nên áp dụng nguyên lý wisdom of growth vào trong việc chuẩn hóa và xử lý định dạng dữ liệu để tối ưu hóa về mặt thời gian thực thi cho việc tiền xử lý dữ liệu theo nguyên tắc chuẩn hóa này.

Dataloader được sử dụng với sự gia tăng dữ liệu theo trình tự sau: ngẫu nhiên độ sáng; sự tương phản ngẫu nhiên; cây trồng ngẫu nhiên; lật trái và phải ngẫu nhiên; chuẩn hóa pixel. Tỷ lệ đào tạo, xác nhận và kiểm tra lần lượt là 0,65, 0,2 và 0,15

1. Implementation and Metrics:

Kiến trúc mô hình được triển khai và thực hiện theo 2 mô trường khác nhau lệ thuộc vào độ lớn của dữ liệu và kiến trúc phức tạp của mô hình.

Với kiến trúc CheXNet121, nhóm chúng tôi tiến hành xây dựng dựa trên Framework tensorflow, keras, torch và yolov5. Đối với việc áp dụng một lượng lớn dự liệu cho việc đào tạo mô hình nên chúng tôi tiến hành xây dựng kiến trúc này trên một server đặc biệt với cấu hình được mô tả như sau.

CPU: Intel Core i9 9900K 3.6 turbo 5.0 GHz /8 Cores 16 Threads  
 RAM: 32 GB DDR4  
 GPU: 2x GeForce RTX 3060 V2 OC Edition 12GB GDDR6

Còn với Faster R-CNN tôi tiến hành sử dụng YoloV.5 để tiến hành pretrained model dựa các weight của tập ImageNet. Quá trình đào tạo được chạy trong Google Colab với GPU là Tesla 16GB và 12 GB Ram.

Đối với hàm tối ưu của mô hình, tôi sử dụng thuật toán Adam với tỷ lệ learning rate là 1e-3 và mô hình này được đào tạo với 100 epochs và batch-size là 500[27,28].

Vì tính chất của khung ảnh y khoa khung ngực, tức là một khung ngực bình hoàn toàn có thể có nhiều class lồng nhau nên hàm mất mát(Loss function) được sử dụng chính là tổng Binary cross-entropy[29] để có thể đảm bảo rằng việc tính toán mất mát sai lệch giữa các phân lớp trên là chuẩn xác. Tổng quan của hàm Binary cross- entropy[22, 30]:

Nhưng đây là bài toán detect 14 class nên hàm loss được định nghĩa là tổng các loss của các loss trong 14 class.

Accuracy, recall, precision, and F1-Score được sử dụng để đánh giá các mô hình. Đối với một vài bài toán thực tế thì chỉ số đo đạt metrics cần thiết chỉ là accuracy nhưng bài toán y học là một bài toán với vấn đề nhạy cảm. Vì là chuyên môn y học nên việc đánh giá mô hình không chỉ thuộc độ chính xác mà có còn phụ thuộc vào 3 chỉ số khác nhưng đánh giá tổng quát vẫn là F1-score.

Precision được định nghĩa là tỉ lệ số điểm Positive mô hình dự đoán đúng trên tổng số điểm mô hình dự đoán là Positive[15,16]. Recall được định nghĩa là tỉ lệ số điểm Positive mô hình dự đoán đúng trên tổng số điểm thật sự là Positive (hay tổng số điểm được gán nhãn là Positive ban đầu[20,21].

Precision càng cao, tức là số điểm mô hình dự đoán là positive đều là positive càng nhiều. Precision = 1, tức là tất cả số điểm mô hình dự doán là Positive đều đúng, hay không có điểm nào có nhãn là Negative mà mô hình dự đoán nhầm là Positive[13].

Recall càng cao, tức là số điểm là positive bị bỏ sót càng ít. Recall = 1, tức là tất cả số điểm có nhãn là Positive đều được mô hình nhận ra[12].

Tuy nhiên, chỉ có Precision hay chỉ có Recall thì không đánh giá được chất lượng mô hình. Chỉ dùng Precision[11], mô hình chỉ đưa ra dự đoán cho một điểm mà nó chắc chắn nhất. Khi đó Precision = 1, tuy nhiên ta không thể nói là mô hình này tốt.Chỉ dùng Recall, nếu mô hình dự đoán tất cả các điểm đều là positive. Khi đó Recall = 1, tuy nhiên ta cũng không thể nói đây là mô hình tốt[31,32]. Đặc biệt đối với bài toán nhạy cảm như y học.

Khi đó F1-score được sử dụng. F1-score là trung bình điều hòa (harmonic mean) của precision và recall (giả sử hai đại lượng này khác 0). F1-score được tính theo công thức:

Đối Optimize function, Adam optimizer là một sự lựa chọn không thể phù hợp hơn khi viêc của các optimizer là tự động cấu hình lại learning rate theo quá trình học để đạt được hiệu suất tốt hơn.

Để biết hơn về các thức mà Adam optimizer hoạt động cùng qua lại các thuật toán optimizer trước đó để nối các nhược điểm của chúng lại và đối chiếu với Adam.

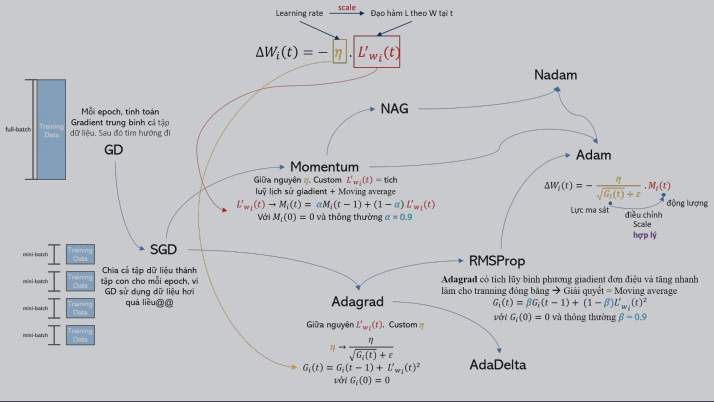


Figure 6. Optimizer connection and inheritant map

Với mục tiêu minimize Objective function(loss function). Một thuật toán tốt là thuật toán tìm ra minimum nhanh và đáng tin cậy (nghĩa là nó không bị mắc kẹt trong local minimal, saddle points, plateau regionsz) mà là global minimum.

Tại sao trong quá trình minimum Loss function thì model càng được cải thiện?

⇒ Vì trong hàm Loss chứa các trọng số, khi minimum hàm loss thì các trọng số thay đổi theo, và nó thay đổi để phù hợp với target đã cho, dẫn đến model của ta càng lúc càng hoàn thiện. Thuật toán sơ khai nhất trong machine learning là gradient descent với việc cập nhật điểm tiếp theo dựa trên độ dốc thay đổi (đạo hàm) của nó:

∆Wi(t) = −ηL0wi (t)

Tuy nhiên nhược điểm của phương pháp này là khi w rơi vào local minimum thì rất khó có thể vượt qua được. Nhiệm vụ của ta giúp w vượt qua được các local minimum.

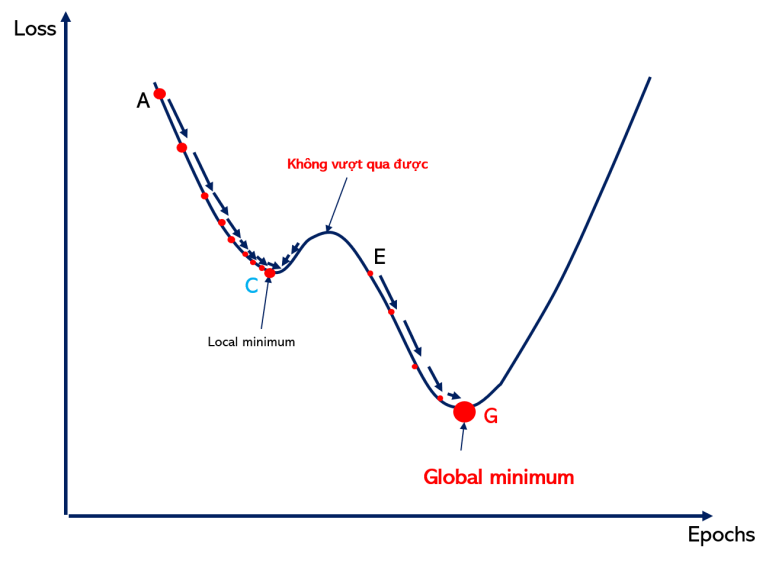


Figure 7. Compare local minium and global minium

Kiến thức này sẽ tập trung vào vấn đề chính bỏ qua 2 yếu tố phụ là momentum và learning rate[1] thì đi thẳng vào các optimizer

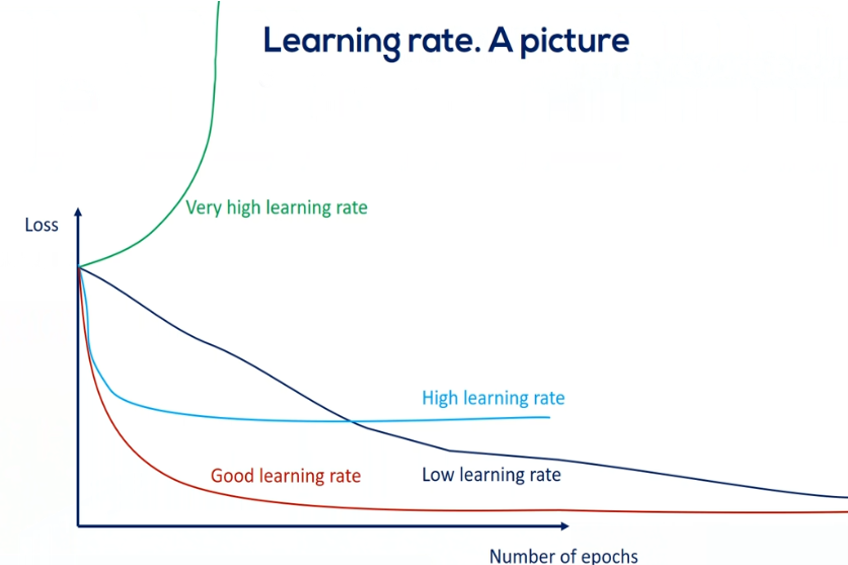


Figure 8. What effect of learning rate in learning curve and ML

Sơ lượt tổng quan về 2 thuật toán AdaGrad và RMSprop.

AdaGrad Không giống như các cách thức trước, learning rate hầu như giống nhau cho quá trình learning, adagrad coi learning rate cũng là một tham số. Nó update tạo các update lớn với các dữ liệu khác biệt nhiều và các update nhỏ cho các dữ liệu ít khác biệt. Adagrad chia learning rate với tổng bình phương của lịch sử biến thiên (đạo hàm)

Trong đó ϵ là hệ số để tránh lỗi chia cho 0, mặc định là 1e−8. G là một diagonal matrix nơi mà mỗi phần tử (i,i) là bình phương của đạo hàm vector tham số tại thời điểm t.

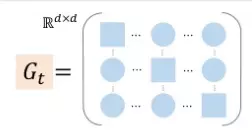


Figure 9. Diagonal Matrix of AdaGrad Optimizer

Một lợi ích dễ thấy của Adagrad là tránh việc điều chỉnh learning rate bằng tay, thường sẽ để default là 0.01 và thuật toán sau đó sẽ tự động điều chỉnh.

Một điểm yếu của Adagrad là tổng bình phương biến thiên sẽ lớn dần theo thời gian cho đến khi nó làm learning rate cực kì nhỏ, làm việc traning trở nên đóng băng.

Còn về phía RMSprop có thề được ngắn gọn là mục đích chính của nó là giải quyết vấn đề của AdaGrad bằng các thay đổi một chút phần hệ số G(t) của AdaGrad bằng moving average:

Lần này tổng bình phương biến thiên không đơn điệu theo thời gian, nó đã được hạn chế về giá trị. Ưu điểm rõ nhất của RMSprop là giải quyết được vấn đề tốc độ học giảm dần của Adagrad. Tuy nhiên thuật toán RMSprop hay AdaGrad có thể cho kết quả nghiệm chỉ là local minimum chứ không đạt được global minimum như Momentum. Vì vậy người ta sẽ kết hợp cả 2 thuật toán Momentum với RMSprop cho ra 1 thuật toán tối ưu Adam.

Adam[2] Giống với Adadelta và RMSprop, nó duy trì trung bình bình phương độ dốc (slope) quá khứ vt và cũng đồng thời duy trì trung bình độ dốc quá khứ mt, giống momentum.

Trong khi momentum giống như một quả cầu lao xuống dốc, thì Adam lại giống như một quả cầu rất nặng và có ma sát (friction), nhờ vậy nó dễ dàng vượt qua local minimum và đạt tới điểm tối ưu nhất (flat minimum)

Ta có công thức Adam:

1. Result:

Table. 4 đã biểu diễn hiệu suất của các mô hình. Cả ba mô hình này đều được đào tạo về bộ dữ liệu.Ngoại trừ mô hình CR-IM-SCRE, cả mô hình cơ sở và mô hình đề xuất của chúng tôi đều được đào tạo với 100 epoch.

Để xác định xem hiệu suất của CheXNet có cao hơn đáng kể về mặt thống kê so với hiệu suất của các mô hình khác không, tôi cũng tính toán sự khác biệt giữa F1 score trung bình của CheXNet và F1 score trung bình. tôi nhận thấy rằng sự khác biệt về F1 scores — 0.051 (95% CI 0.005, 0.084) . do đó kết luận rằng hiệu suất của CheXNet cao hơn và có chính xác trong việc thống kê so với hiệu suất của 2 mô hình còn lại.

|  |  |  |  |
| --- | --- | --- | --- |
| Pathlogoy | Faster R-CNN | CheXNet | Detectron2 |
| Atelectasis (Xẹp phổi) | 0.4705 | 0.8094 | 0.7849 |
| Cardiomegaly (Cơ tim phìn đại) | 0.9106 | 0.9248 | 0.9031 |
| Effusion(Tràn dịch màng phổi) | 0.7346 | 0.8638 | 0.7723 |
| Infiltration(Thâm nhiểm phổi) | 0.7914 | 0.7345 | 0.7431 |
| Mass (Phế cầu khuẩn ở phổi) | 0.9316 | 0.8676 | 0.9129 |
| Nodule(nốt phổi mờ) | 0.7494 | 0.7802 | 0.6863 |
| Pneumonia(Nhiễm trùng phổi) | 0.6919 | 0.7680 | 0.7721 |
| Pneumothorax(Tràn khí màng phổi) | 0.5749 | 0.8887 | 0.6811 |
| Consolidation(Kết đặc ở phổi) | 0.6591 | 0.7901 | 0.7313 |
| Edema(Phù phổi) | 0.7446 | 0.8878 | 0.8155 |
| Emphysema(Khí thủng phổi) | 0.8633 | 0.9371 | 0.8261 |
| Fibrosis(Sơ nan phổi) | 0.8679 | 0.8047 | 0.8356 |
| Pleural Thickening(Dày màng phổi) | 0.7909 | 0.8062 | 0.7531 |
| Hernia(Thoát vị bẹn) | 0.7791 | 0.9164 | 0.8156 |

Table 2. CheXNet outperforms the best published results on all 14 pathologies in the ChestX-ray14 dataset. In detecting Mass, Nodule, Pneumonia, and Emphysema, CheXNet has a margin of >0.05 AUROC over previous state of the art results.

1. Future work:

Đây là phần sẽ đưa ra những failed case trong bài toán trên và đề xuất hướng phát triển cho vấn đề trên.

Đầu tiên có trong việc đào tạo mô hình còn nhiều điểm bất cập như sau:

Vì là ảnh y khoa nên tính chất hướng chuyên môn học thuật về lĩnh vực y khá cao nên yêu cầu tối thiếu cho việc giải quyết bài toán này là cần ít nhất một người thuộc lĩnh vực trên để có thể giải thích tính đúng sai và tường minh của mô hình mặt khác nếu trong trường hợp thiếu sự hỗ trợ từ phía chuyên môn thì việc dự đoán tính đúng sai của mô hình không thể được giải đáp hoặc là bắt buộc người xây dựng kiến trúc mô hình phải nạp một lượng kiến thức chuyên môn từ một lĩnh vực khác hoặc là coi mô hình như một black box và chấp nhận kết quả vô điều kiện.

Fail case đa phần rơi vào trường hợp khi nào vì một số đặc trưng nhất định cùng với tính chất nhất quán của ảnh y khoa mà khung ngoài phạm vị của anh đôi khi lại mang một số đặc trưng sai lệch cụ thể dẫn đến các Feature map[7] của cuối cùng mô hình không nằm gọn trong khung phổi mà có thể bị out range nếu xuất hiện tình trạng trên.

Vấn đề cuối là về phần cứng để xây dựng mô hình. Việc phát triển mô hình và cấu thành các tham số dựa trên bài toán có đặc trưng quá phức tạp đã dẫn đến việc đòi hỏi buộc phải đánh đổi giữa vấn đề tài chính chi tiêu cho việc xây dựng phần cứng và thời gian, tốc độ tính toán của mô hình. Nghĩa là việc ta tối ưu thời gian thực thị thì bắt buộc phải tăng cấu hình phần cứng và được ngược lại. Việc phát triển card đồ họa vào trong lĩnh vực học máy đã vô tình thúc đẩy việc chạy đua phần cứng cho những bài toán chung của xã hội từ đó tạo ra làn ranh giới riêng giữa việc nghiên cứu phát triển và việc học thuật khi bản thân người học muốn truyền tải những kiến thức mà mình tiếp thu được thành sản phẩm thì bắt buộc phải có phần cứng. Từ đó làm nổi lên quan điểm “AI không dành cho người nghèo”.

1. **IMPROVE THE FUTURE WORK**

Đối với những vấn đề trên nhóm chúng tôi phần nào cũng đã đưa giải pháp cơ bản để hoàn thiện những khuyết điểm mà tôi đã đề ra.

Ở vấn đề thứ nhất giải pháp đưa ra cho bài toán này là tạo ra một community giữa những sinh viên ở các nhóm ngành khác nhau từ đó có thể thấy vấn đề của người đôi khi lại là lời giải của người khác. Từ đó việc sự giao tiếp trong cộng đồng sẽ trở nên lớn mạnh và năng lực khi mà tiếp nhận được lời giải từ người khác họ sẽ không đơn thuần dừng lại tại điểm local mà là tiến lên thêm tới global để đạt được giải pháp chung cho mọi trường hợp của bài toán. Hơn nữa tạo ra một community là tạo một giải pháp chung cho mọi vấn đề của mọi lĩnh vực khi trong khi tin có thể đưa ra các giải pháp để hổ trợ y sinh thì toán lại là một là optimization không thể tuyệt vời hơn dành cho tin và lý

Kế hoạch đó hiện tại đang được thực hiện hóa nhưng chỉ ở mức cơ sở khi UTC2@iLabs sẽ có sự liên kết cơn bản với ngành tự động hóa cùng trường để tạo ra sự ràng buộc giữa lý thuyết và thực hành

Vấn đề thứ hai lại liên quan đến cải tiến và kiến trúc mới trong bài báo tương lại khi quan ngại rằng trên nền ảnh X-ray đơn thuần có quá nhiều yếu tố nhiễu dẫn đến việc feature map sau cùng không còn đơn thuần là gói gọn trong khung phổi nữa. Để xử lý bài toán này tôi đề xuất một kiến trúc mới có tên là Uchex-UNet được sáng kiến triển bởi tôi và được phê duyệt ủng hộ bởi Ph.D Nguyễn Đình Hiển, Th.S Trần Ngọc Đăng Nguyên và Ph.D Nguyễn Việt Linh. Mô hình khái quát hóa của kiến trúc này đang được phát triển nên chưa được khái quá hoát. Nhưng ý tưởng chính của nó là khi cho một tấm ảnh X-Ray đi qua mô hình ở lớp đầu ta sẽ cấu hình một kiến trúc UNet[8,14,33,34] để chắc lọc đặc trưng cần quang sát làm bộ phận phổi. Sau khi đã có được sementit segmentation của phổi tao tiến hành đưa vào Kiến CheXNet để xác định và phát hiện phân khu bị tổn thương theo bounding box. Sau cùng ta tiến hành tách những bound box kia và cho qua kiến trúc UNet lần nữa để lấy được sementic segmentation lession trên phổi. Bước cuối ta tiến hành concatenate các lession với ảnh gốc để được phân vùng tổn thương trực quan nhất có thể. Và đây chính là ý tưởng chính để phát triển bài toán trên.

Còn về vấn đề cuối cùng thì có 2 giải pháp để giải quyết. Hiện nay các cloud ML Compiler rất phổ biến ngoài ra các tổ chức này còn cho sử dụng miễn phí card đồ họa để thuận tiện cho việc học máy. Tất nhiên card đồ học miễn phí thì chỉ train để lấy kết quả thôi chứ không thể nào viết article được. Còn giải pháp thức là chuyển từ python sang ngôn ngữ C++[5] sẽ giúp tăng một lượng lớn thời gian thực thi và tính toán từ đó có thể xử lý các bài toán trong thời gian ngắn hơn.

1. **CONCLUSION**

Trong bài báo cáo này, chúng tôi trình bày một công cụ hoàn chỉnh để hỗ trợ chẩn đoán X-quang ngực. Chúng tôi thảo luận về những thách thức của việc phát triển từng khía cạnh của hệ thống và thực hiện các giải pháp thực tế trong việc phát hiện, dự đoán dịch bệnh và giải thích dự đoán. Chúng tôi tin rằng đây là một giải pháp để thu hẹp khoảng cách giữa cộng đồng y tế và Các nhà nghiên cứu DL.

Cũng trong bài báo cáo này nhóm chúng tôi phát triển một thuật toán phát hiện bệnh viêm phổi từ hình ảnh X-quang ngực[6] nhìn trực diện ở cấp độ trực quan nhất khi mà nhìn vào có thể xác định rõ khu vực bị tổn thương. Chúng tôi cũng cho thấy rằng một phần mở rộng đơn giản của thuật toán của chúng tôi để phát hiện nhiều bệnh vượt trội so với trạng thái kỹ thuật trước đây về. chúng tôi hy vọng rằng công nghệ này có thể cải thiện chăm sóc sức khỏe cung cấp và tăng khả năng tiếp cận với chuyên môn hình ảnh y tế ở các nơi trên thế giới, nơi tiếp cận với bác sĩ X quang còn hạn chế.

1. **REFERENCE**

**[1].** Darken, C., Chang, J., & Moody, J. (1992). Learning rate schedules for faster stochastic gradient search. Neural Networks for Signal Processing II Proceedings of the 1992 IEEE Workshop, (September), 1–11. http://doi.org/10.1109/NNSP.1992.253713

**[2].** Diederik P. Kingma\*University of Amsterdam, OpenAI, JimmyLei Ba University of Toronto et al “ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION”. Published as a conference paper at ICLR 2015

**[3].**  Gao Huang; Zhuang Liu; Laurens Van Der Maaten; Kilian Q. Weinberger et al “Densely Connected Convolutional Networks”. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA

**[4].**  Pranav Rajpurkar, Jeremy Irvin, Brandon Yang et al “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning”. Stanford University Department of Computer Science Stanford University Department of Medicine Stanford University Department of Radiology. Correspondence to: Pranav Rajpurkar , Jeremy Irvin . Project website at https://stanfordmlgroup. github.io/projects/chexnet

**[5].**  Elena Limonova, Alexander Sheshkus, and Dmitry Nikolaev, Computational Optimization of Convolutional Neural Networks using Separated Filters Architecture, International Journal of Applied Engineering Research ISSN 0973- 4562 Volume 11, Number 11 (2016) pp 7493

**[6].**  Đoàn Thiện Minh, Trần Văn Lăng, Văn Đình Vỹ Phương, Phan Mạnh Thường et al. “KỸ THUẬT HỌC SÂU ĐỂ GIẢI QUYẾT BÀI TOÁN CHẨN ĐOÁN BỆNH LAO PHỔI”. Kỷ yếu Hội nghị KHCN Quốc gia lần thứ XII về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR); Huế, ngày 07-08/6/2019 DOI: 10.15625/vap.2019.00030

**[7].**  Kaiming He et al. “Delving deep into rectifiers: Surpassing human-level perfor-  
mance on imagenet classification”. In: Proceedings of the IEEE international con-  
ference on computer vision. 2015, pp. 1026–1034.

**[8].**  Hieu Minh Truong and Hieu Trung Huynh. “A Novel Approach For CT-Based  
COVID-19 Classification and Lesion Segmentation Based On Deep Learning”. In:  
The Computer Journal (Mar. 2022). ISSN: 0010-4620. DOI: 10.1093/comjnl/  
bxac015. URL: <https://doi.org/10.1093/comjnl/bxac015>

**[9].** Aras M Ismael and Abdulkadir S ̧ eng ̈ur. “Deep learning approaches for COVID-19 detection based on chest X-ray images”. In: Expert Systems with Applications 164  
(2021), p. 114054.

**[10].** Hieu Nguyen et al. “Design a learning model of mobile vision to detect diabetic  
retinopathy based on the improvement of MobileNetV2”. In: International Journal  
of Digital Enterprise Technology (2021.In press.).

**[11].**  Joseph Paul Cohen, Mohammad Hashir, Rupert Brooks, Hadrien Bertrand et al. “On the limits of cross-domain generalization in automated X-ray prediction”. Proceedings of Machine Learning Research 1–13, 2020. Mila, Université de Montréal

**[12].**  Henry Z. Lo, Joseph Paul Cohen, and Wei Ding. Prediction gradients for feature extraction and analysis from convolutional neural networks. In International Conference on Automatic Face and Gesture Recognition, 2015. doi: 10.1109/FG.2015.7163154.

**[13].**  Tatiana Gabruseva, Dmytro Poplavskiy, Alexandr A. Kalinin · et al. “Deep Learning for Automatic Pneumonia Detection”. 2005.13899v1. University of Michigan Ann Arbor, MI 48109 USA, and Shenzhen Research Institute of Big Data, Shenzhen 518172, Guangdong, China

**[14].** Nguyen N.D.Tran, Hien D.Nguyen, Nhan T.Huynh, Nha P.Tran and Linh V.Nguyen et al. “Segmentation on Chest CT Imaging in COVID- 19 based on the improvement Attention U-net model”. In: Proceedings of the S.O.M.E.T Conference 2022

**[15].** Ke Yan et al. “Holistic and comprehensive annotation of clinically significant findings on diverse CT images: learning from radiology reports and label ontology”.  
In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern  
Recognition (CVPR 2019). 2019, pp. 8523–8532

**[16].** Hien D Nguyen et al. “Ontology-based Integration of Knowledge Base for Build-  
ing an Intelligent Searching Chatbot”. In: Sensors and Materials 33.9 (2021),  
pp. 3101–3123.

**[17].** Mohammad Rahimzadeh, Abolfazl Attar et al. “A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Exception and ResNet50V2”. 2004.08052v2

**[18].** Datasets: ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. {xiaosong.wang,yifan.peng,le.lu,luzh,mohammad.bagheri,rms}@nih.gov

**[19].**  Alexey A. Novikov, Dimitrios Lenis, David Major, Jiˇr´ı Hladuvka, Maria Wimmer, and Katja Buhler et al. Fully Convolutional Architectures for Multi-Class Segmentation in Chest Radiographs

**[20].**  Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong, Bagheri, Mohammadhadi, and Summers, Ronald M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. arXiv preprint arXiv:1705.02315, 2017.

**[21].**  Potchen, EJ, Gard, JW, Lazar, P, Lahaie, P, and Andary, M. Effect of clinical history data on chest film interpretation-direction or distraction. In Investigative Radiology, volume 14, pp. 404–404, 1979

**[22].**  Mollura, Daniel J, Azene, Ezana M, Starikovsky, Anna, Thelwell, Aduke, Iosifescu, Sarah, Kimble, Cary, Polin, Ann, Garra, Brian S, DeStigter, Kristen K, Short, Brad, et al. White paper report of the rad-aid conference on international radiology for developing countries: identifying challenges, opportunities, and strategies for imaging services in the developing world. Journal of the American College of Radiology, 7(7):495–500, 2010

**[23].**  Aydogdu, M, Ozyilmaz, E, Aksoy, Handan, Gursel, G, and Ekim, Numan. Mortality prediction in community-acquired pneumonia requiring mechanical ventilation; values of pneumonia and intensive care unit severity scores. Tuberk Toraks, 58(1):25– 34, 2010.

**[24].**  Cherian, Thomas, Mulholland, E Kim, Carlin, John B, Ostensen, Harald, Amin, Ruhul, Campo, Margaret de, Greenberg, David, Lagos, Rosanna, Lucero, Marilla, Madhi, Shabir A, et al. Standardized interpretation of pediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. Bulletin of the World Health Organization, 83(5):353–359, 2005.

**[25].**  Hopstaken, RM, Whitbread, T, Van Engelshoven, JMA, and Dinant, GJ. Inter-observer variation in the interpretation of chest radiographs for pneumonia in community-acquired lower respiratory tract infections. Clinical radiology, 59(8):743–752, 2004.

**[26].**  Huang, Peng, Park, Seyoun, Yan, Rongkai, Lee, Junghoon, Chu, Linda C, Lin, Cheng T, Hussien, Amira, Rathmell, Joshua, Thomas, Brett, Chen, Chen, et al. Added value of computer-aided ct image features for early lung cancer diagnosis with small pulmonary nodules: A matched case-control study. Radiology, pp. 162725, 2017

**[27].** Ioffe, Sergey and Szegedy, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International Conference on Machine Learning, pp. 448–456, 2015.

**[28].** Esteva, Andre, Kuprel, Brett, Novoa, Roberto A, Ko, Justin, Swetter, Susan M, Blau, Helen M, and Thrun, Sebastian. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639):115–118, 2017.

**[29].** Deng, Jia, Dong, Wei, Socher, Richard, Li, Li-Jia, Li, Kai, and Fei-Fei, Li. Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pp. 248–255. IEEE, 2009.

**[30].** Lakhani, Paras and Sundaram, Baskaran. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, pp. 162326, 2017.

**[31].** Tianqi Chen, Mu Li, Yutian Li, Min Lin, Naiyan Wang, Minjie Wang, Tianjun Xiao, Bing Xu, Chiyuan Zhang, and Zheng Zhang. MXNet: A Flexible and Efficient Machine Learning Library for Heterogeneous Distributed Systems. Neural Information Processing Systems, Workshop on Machine Learning Systems, 2015.

**[32].** O. Oktay, E. Ferrante, K. Kamnitsas, M. Heinrich, W. Bai, J. Caballero, S. Cook, A. de Marvao, T. Dawes, D. ORegan, B. Kainz, B. Glocker, and D. Rueckert, “Anatomically constrained neural networks (acnn): Application to cardiac image enhancement and segmentation,” IEEE Transactions on Medical Imaging, vol. PP, no. 99, pp. 1–1, 2017.

**[33].** S. Singh, D. Hoiem, and D. Forsyth, “Swapout: Learning an ensemble of deep architectures,” in Advances in Neural Information Processing Systems 29, 2016, pp. 28–36.[Online].Available:http://papers.nips.cc/paper/6205-swapout-learning-an-ensemble-of-deep-architectures.pdf

**[34].** L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Semantic image segmentation with deep convolutional nets and fully connected crfs,” arXiv e-prints, arXiv:1412.7062v4 [cs.CV], vol. abs/1412.7062, 2014. [Online]. Available: http://arxiv.org/abs/1412.7062

**[35].** J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.- i. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, “Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists’ detection of pulmonary nodules,” American Journal of Roentgenology, vol. 174, no. 1, pp. 71–74, 2000.

**[36].** Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). 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 (pp. 2097-2106)

**[37].** Datasets ChestX-ray14, the largest publicly available chest X- ray dataset which released by the NIH

**[38].** Datasets VinBigData Chest X-ray Abnormalities Detection Automatically localize and classify thoracic abnormalities from chest radiographs which release by Vingroup Big Data Institute

**[39].** Datasets RSNA Pneumonia Detection Challenge which you can build an algorithm that automatically detects potential pneumonia cases. It was released by Radiological Society of North America Organization