Chẩn đoán bệnh viêm phổi tự động bằng kỹ thuật Transfer Learning

Tác Giả: Lưu Gia Khang

Khoa Công nghệ Thông tin, Trường Đại học Ngoại ngữ - Tin học Thành phố Hồ Chí Minh

Tác giả liên hệ: Email: 18dh110896@st.huflit.edu.vn | Điện thoại: 0936322946

**Tóm tắt**

Viêm phổi là một căn bệnh cực kì nguy hiểm, căn bệnh này có thể gây nguy hiểm đến tính mạng nếu không được chữa trị đúng lúc và do đó, việc chẩn đoán sớm bệnh viêm phổi là rất quan trọng. Bài báo này đề xuất cách tiếp cận transfer learning trên các mô hình học sâu để có thể chẩn đoán bệnh viêm phổi tự động bằng cách sử dụng hình X-quang. Bài báo cung cấp những báo cáo chi tiết trong việc phát hiện chính xác bệnh viêm phổi và sau đó trình bày các phương pháp đã được tác giả áp dụng. Kỹ thuật transfer learning được áp dụng trên hai mạng pretrained Convolutional Neural Network (CNN) là VGG16 và ResNet18. Hai mạng được huấn luyện và đánh giá trên bộ dữ liệu chứa tổng cộng 5910 ảnh X-quang đã được phân loại. Độ chính xác cao nhất đạt được trong việc phân loại hình ảnh đạt được là 99%. Do đó, nghiên cứu được đề xuất có thể hữu ích trong việc chẩn đoán bệnh viêm phổi nhanh hơn.

**Từ khoá:** Computer vision; Deep learning; Image classification; Pneumonia diagnosis.

1. GIỚI THIỆU

Viêm phổi là bệnh nhiễm trùng đường hô hấp dưới gây nhiều biến chứng nghiêm trọng, nếu không kịp thời điều trị có thể dẫn đến tử vong cho người lớn lẫn trẻ nhỏ [[1](#ref1)]. Mỗi năm, trên thế giới có khoảng 450 triệu người mắc bệnh viêm phổi và dẫn đến khoảng 4 triệu ca tử vong [[2](#ref2), [3](#ref3)]. Trẻ em dưới 5 tuổi và người lớn trên 70 tuổi là hai nhóm có nguy cơ mắc bệnh cao nhất. Từ năm 2007 đến 2017, số ca tử vong do viêm phổi ở trẻ em tuy đã giảm 36% nhưng tỉ lệ này lại tăng 34% ở người lớn trên 70 tuổi [[4](#ref4)]. Chụp X-quang ngực là phương pháp phát hiện bệnh tốt nhất nhưng không hoàn hảo. Để có thể đưa ra chẩn đoán chính xác nhất, các bác sĩ cần phải có thêm các thông tin về tình trạng sức khoẻ của bệnh nhân. Các triệu chứng bên ngoài như sốt và tiết dịch cũng phải được tính đến trong quá trình chuẩn đoán. Việc chuẩn đoán bệnh một cách chính xác và nhanh chóng có ý nghĩa rất lớn, điều này có thể đảm bảo các bệnh nhân có thể được tiếp nhận điều trị kịp thời cũng như tiết kiệm được rất nhiều thời gian và tiền bạc.

Nhờ vào lượng dữ liệu dồi dào mà Internet mang lại cùng với sự phát triển của phần cứng máy tính đã giải quyết được các vấn đề về dữ liệu và khả năng tính toán của mạng neuron sâu [[5](#ref5)]. Từ đó, lĩnh vực Deep Learning (DL) ngày càng phát triển và các mạng neuron mới được ra đời. Đối với các thuật toán Machine Learning (ML) truyền thống như SVM hay AdaBoost, ta phải tự trích xuất đặc trưng của dữ liệu đầu vào rồi mới có thể đưa vào thuật toán để phân loại. Tuy nhiên, các mô hình DL có thể tự động trích xuất đặc trưng của dữ liệu đầu vào và tiến hành phân loại dữ liệu [[6](#ref6)]. Vì vậy, mạng neuron vừa bộ trích xuất đặc trưng vừa là bộ phân loại có thể được huấn luyện từ đầu đến cuối.

Trong thời gian gần đây, Convolutional Neural Network (CNN) là lựa chọn tốt nhất để giải quyết các bài toán phân loại hình ảnh. Bằng cách sử dụng các convolutional filter, CNN có thể xử lý hình ảnh tương đương với con người và cấu trúc cực kỳ tối ưu hoá để xử lý hình ảnh dạng 2D và 3D, cũng như khả năng trích xuất các đặc trưng trừu tượng thông qua học tập. So với các lớp Fully Connected (FC) ,các lớp Convolutional của CNN thường có ít tham số hơn nhờ vào tính chất kết nối cục bộ. LeNet [[7](#ref7)], AlexNet [[8](#ref8)], VGGNet [[9](#ref9)], GoogleNet [[10](#ref10)], ResNet [[11](#ref11)] là các kiến trúc mạng CNN nâng cao phổ biến và vượt trội nhất trong từng thời kỳ. Bên cạnh các kiến trúc mạng nơ-ron, các kỹ thuật huấn luyện mô hình cũng được ra đời. Transfer learning là kỹ thuật chuyển giao các đặc trưng mà mạng neuron đã học được từ một bộ dữ liệu rất lớn sang một bộ dữ liệu mới để áp dụng vào bài toán mới [[12](#ref12)]. Huấn luyện toàn bộ mạng neuron từ đầu yêu cầu rất nhiều dữ liệu để có kết quả tốt. Việc có được rất nhiều dữ liệu thường không khả thi trong hầu hết tất cả các trường hợp. Về mặt tính toán, rất tốn kém để đào tạo một mạng neuron sâu trên hàng triệu hình ảnh vì làm như vậy thường đòi hỏi nhiều tuần đào tạo trên nhiều GPU. Đào tạo mạng neuron là một quá trình lặp đi lặp lại. Vì vậy, ngay cả khi ta có khả năng tính toán cần thiết để đào tạo một mạng neuron phức tạp, việc dành hàng tuần thử nghiệm với các siêu tham số khác nhau trong mỗi lần đào tạo cho đến khi đạt được kết quả ưng ý sẽ khiến dự án trở nên rất tốn kém. Ngoài ra, việc sử dụng transfer learning còn giúp tránh overfitting. Việc chuyển giao các đặc trưng được trích xuất từ 1 mạng khác đã xem hàng triệu ảnh sẽ giúp mô hình của ta ít bị overfitting và giúp mô hình của ta tổng quát hoá hơn khi gặp phải các tình huống mới [[13](#ref13)].

Bài nghiên cứu này đề xuất cách tiếp cận transfer learning bằng cách sử dụng hai kiến trúc mạng đã được đào tạo trước khác nhau (VGG16 và ResNet18) và phân tích hiệu suất của chúng để giải quyết vấn đề chẩn đoán bệnh viêm phổi tự động. Đóng góp quan trọng của công trình này là cung cấp phương pháp tiếp cận transfer learning dựa trên CNN bằng cách sử dụng các mô hình được đào tạo trước khác nhau để phát hiện bệnh viêm phổi với độ chính xác cao hơn so với các công trình trước đây. Ngoài ra, bài báo cũng cung cấp các chi tiết phương pháp của công trình, có thể được sử dụng bởi bất kỳ nhóm nghiên cứu nào. Động lực của nghiên cứu này là sử dụng sức mạnh của DL để chẩn đoán bệnh viêm phổi bằng cách phân tích ảnh chụp X-quang với độ chính xác cao.

1. CÔNG TRÌNH LIÊN QUAN

Những cải tiến mới nhất trong mô hình DL và sự sẵn có của bộ dữ liệu khổng lồ đã hỗ trợ các thuật toán đưa ra quyết định có độ chính xác tương đương hay thâm chí vượt trội hơn so với các chuyên gia trong lĩnh vực xử lý ảnh y tế như là chẩn đoán ung thư da [[14](#ref14)], phát hiện rối loạn nhịp tim [[15](#ref15)] và phát hiện bệnh võng mạc tiểu đường [[16](#ref16)]. Chẩn đoán tự động bằng cách phân tích ảnh chụp X-quang phổi ngày càng được quan tâm. Các thuật toán này ngày càng được sử dụng nhiều hơn để tiến hành phát hiện nốt phổi [[17](#ref17)] và chẩn đoán bệnh lao phổi [[18](#ref18)].

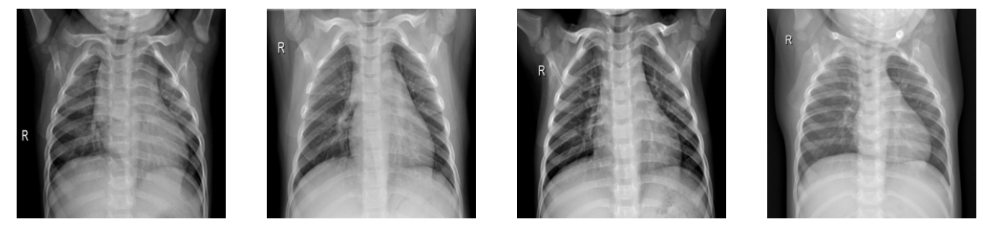
Một vài nhóm nghiên cứu đã đề xuất các cách áp dụng thuật toán DL để pháp hiện bệnh viêm phổi khác nhau và điều cho ra những kết quả đầy hứa hẹn. Tác giả của bài báo [[19](#ref_buil_cnn)] đã đề xuất kiến trúc CNN với sáu lớp để tiết kiệm chi phí tính toán của mô hình và chỉ đạt được độ chính xác cao nhất là 93%. Tác giả của bài báo [[20](#ref_inception)] sử dụng mạng pretrained InceptionV3 đạt được các giá trị Recall và Precision lần lượt là 93.2% và 90.1%. Tác giả của bài báo [[21](#ref_customize_vgg16)] đề xuất cách tinh chỉnh mạng VGG16 và đạt được giá trị Average Under the Curve (AUC) là 99%. Tác giả của bài báo [[22](#ref_densenet)], áp dụng transfer learning với mạng DenseNet201 đạt được các giá trị Recall, Precision và AUC lần lượt là 99%, 97% và 98%.

1. Phương pháp
   1. Tập dữ liệu

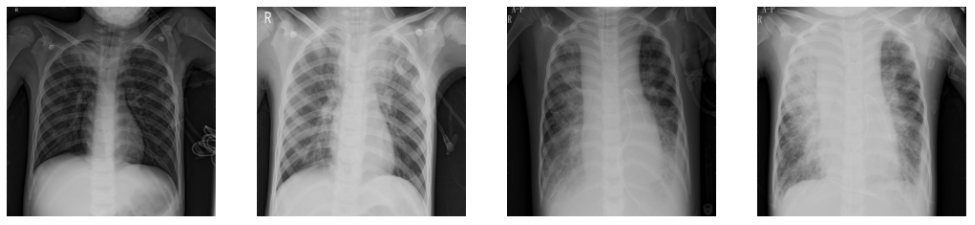
Trong nghiên cứu này, tập dữ liệu hình X-quang ngực được tải xuống từ <https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset>. Tập dữ liệu gốc bao gồm hai thư mục là thư mục train và thư mục test chứa các file ảnh X-quang ngực, cùng với hai tệp csv, trong đó file metadata.csv lưu toàn bộ thông tin của dữ liệu như là tên file ảnh, nhãn, loại tập dữ liệu, .... Tập dữ liệu này có tổng cộng 5910 ảnh, trong đó tập train chứa 5286 ảnh với 3944 ảnh được gán nhãn mắc bệnh viêm phổi và 1342 ảnh được gán nhãn bình thường, tập test chứa 624 ảnh với 390 ảnh được gán nhãn mắc bệnh viêm phổi và 234 ảnh được gán nhãn bình thường. Tất cả các ảnh đều có kích thước khác nhau và 80% tập dữ liệu được thu thập từ các nguồn khác nhau. Bảng [1](#table_dataset) thể hiện số lượng dữ liệu của mỗi lớp để huẩn luyện và kiểm tra mô hình. Hình [1](#fig_normal) là mẫu các hình X-quang được gán nhãn bình thường và Hình [2](#fig_pneumonia) là mẫu các hình X-quang được gán nhãn mắc bệnh viêm phổi của tập dữ liệu.

1. Số lượng dữ liệu của mỗi lớp trên tập train và test.

|  |  |  |  |
| --- | --- | --- | --- |
| Trường hợp | Tập train | Tập test | Tổng |
| Bình thường | 1342 | 234 | 1576 |
| Viêm phổi | 3944 | 390 | 4334 |
| Tổng | 5286 | 624 | 5910 |



1. Các ảnh được gán nhãn bình thường.



1. Các ảnh được gán nhãn bị viêm phổi.
   1. Tiền xử lý dữ liệu và kỹ thuật augmentation

Tất cả ảnh trước khi được đưa vào mô hình đều được resize về một kích thước cố định là 224 x 224 x 3 và giá trị của các pixel được chuẩn hoá trong khoảng từ 0 đến 1.

Data augmentation là kỹ thuật giúp cho mô hình có thể nhìn thấy được nhiều biến thể của dữ liệu bao gồm các phép xoay ảnh, lật ảnh, thay đổi độ sáng, cắt ảnh, ... [[23](#ref_aug)]. Trong quá trình huấn luyện, các ảnh sẽ được biến đổi ngẫu nhiên bởi sự kết hợp giữa các phép biển đổi ảnh từ đó làm giảm sự phụ thuộc của mô hình vào dạng ban đầu của đối tượng trong quá trình học đặc trưng. Kỹ thuật này có thể giúp mạng hạn chế hiện tượng overfitting hay thậm chí có thể cải thiện độ chính xác của mô hình. Trong bài nghiên cứu này, tác giả sử dụng các phép xoay ảnh, thay đổi độ sáng, lật ảnh ngang và zoom ảnh.

Cụ thể, ảnh được xoay trong khoảng 10 độ. Ảnh được làm tối đi trong khoảng 10% và làm sáng trong khoảng 20%. Ảnh còn được phóng to ngẫu nhiên từ 75%. Cuối cùng là phép lật ảnh ngang. Hình [3](#fig_augmentation) là mẫu của các ảnh sau khi áp dụng các phép biến đổi trên một cách ngẫu nhiên trước khi đưa vào mô hình. Các cài đặt thông số áp dụng cho từng phép biến đổi ảnh được sử dụng trong bài nghiên cứu và ý nghĩa của chúng được thể hiện tại Bảng [2](#table_augmentation).



1. Ảnh sau khi áp dụng kỹ thuật augmentation.
2. Chi tiết các thông số áp dụng augmentation.

|  |  |  |
| --- | --- | --- |
| Phép biển đổi ảnh | Thông số | Ý nghĩa |
| Xoay ảnh | 10 | Ảnh được xoay ngẫu nhiên trong phạm vi từ 0 đến 10 độ. |
| Thay đổi độ sáng | (0.8, 1.2) | Ảnh được ngẫu nhiên làm tối đi trong phạm vi tỉ lệ từ 0.8 đến 1 và được ngẫu nhiên làm sáng hơn trong phạm vi tỉ lệ từ 1 đến 1.2. |
| Phóng to ảnh | (0.75, 1) | Phóng to ảnh ngẫu nhiên trong phạm vi tỉ lệ từ 0.75 đến 1. |
| Lật ảnh ngang | True | Cho phép ảnh lật ngang. |

* 1. Pretrained VGG16 và ResNet18

Hai mô hình pretrained VGG16 và ResNet18 đều được huấn luyện trên bộ dữ liệu ImageNet. Bộ phân loại của cả hai mô hình đều được thay đổi hoàn toàn để phù hợp với yêu cầu bài toán. Các siêu tham số dùng để huấn luyện mô hình được tinh chỉnh sau các lần thử nghiệm. Các siêu tham số được đưa ra trong bài nghiên cứu là những siêu tham số cho ra kết quả tốt nhất. Kiến trúc tổng quan sau khi tinh chỉnh trên 2 mạng pretrained VGG16 và ResNet18 được thể hiện ở Hình [4](#fig_architecture).

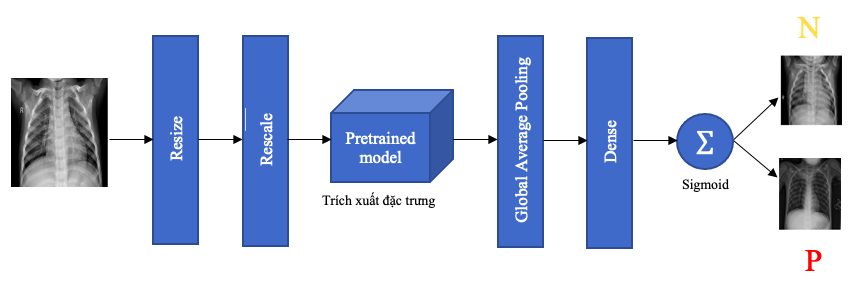
Cụ thể, mô hình mới được tạo nên bởi ba thành phần chính: bộ tiền xử lý dữ liệu, bộ rút trích đặc trưng và bộ phân loại. Bộ tiền xử lý dữ liệu gồm hai lớp resize và rescale. Lớp resize sẽ thay đổi kích thước của ảnh về một kích thước cố định để có thể đưa vào bộ phân lớp và lớp rescale sẽ chuẩn hoá các giá trị pixel của ảnh trong khoảng [0, 255] thành [0, 1]. Với hai lớp này, việc triển khai mô hình áp dụng thực tế sẽ dễ dàng hơn, bởi các ảnh đầu vào không cần phải tiền xử lý trước rồi mới có thể đưa vào mô hình mà được xử lý bởi bộ tiền xử lý dữ liệu.

Các ảnh sau khi được tiền xử lý sẽ được đưa vào bộ rút trích đặc trưng. Bộ rút trích đặc trưng của cả hai mạng pretrained đều được sử dụng lại hoàn toàn, với giá trị khởi tạo của các trọng số là giá trị đã tối ưu trên bộ dữ liệu ImageNet. Cả hai mạng pretrained đều nhận đầu vào là ảnh có kích thước 224 x 224 x 3 và đầu ra ở lớp Convolution cuối cùng là feature map có kích thước 7 x 7 x 512. Toàn bộ các lớp Convolution đều sử dụng hàm kích hoạt ReLU và được huấn luyện lại từ đầu trên bộ dữ liệu mới. Các feature map được đưa ra sau mỗi khổi Convolution của hai mạng VGG16 và ResNet18 được minh hoạ ở hình [5(a)](#fig_featuremap) và [5(b).](#fig_featuremap)

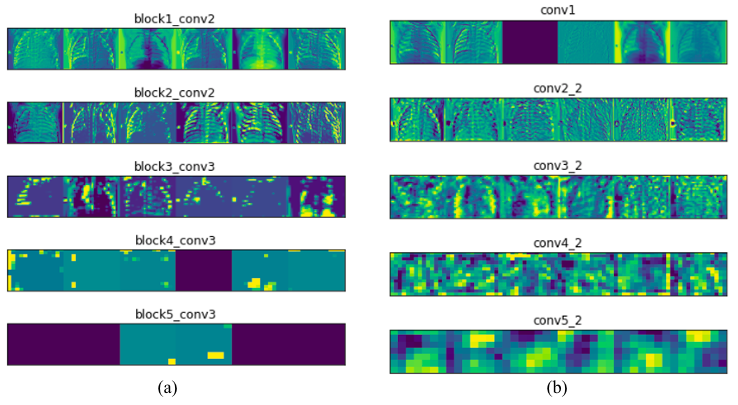
(01)

Bộ phân loại của cả hai mạng có kiến trúc giống nhau, gồm một lớp Global Average Pooling (GAP) và cuối cùng là lớp FC đưa ra kết quả dự đoán. Đầu vào của lớp GAP là feature map có kích thước 7 x 7 x 512 và kết quả đưa ra là vector có kích thước 1 x 1 x 512. Vector này sau đó được đưa vào lớp FC cuối cùng để tính toán và đưa ra kết quả dự đoán là một giá trị xác suất được tính bởi hàm Sigmoid. Nhờ vào cơ chế lấy giá trị trung bình của feature map nên lớp GAP không chứa bất kỳ tham số nào. Vì các mạng trên không chứa bất kỳ lớp FC nào nên mạng được gọi là Fully Convolutional Network (FCN).

(02)



1. Tổng quan kiến trúc áp dụng transfer learning.



1. Feature map sau mỗi khối Convolution của (a) VGG16 và (b) ResNet18.

Để có thể đánh giá hiệu suất của mô hình trong quá trình huấn luyện, mô hình cần thêm một tập validation. Tập validation được lấy ngẫu nhiên 20% từ tập traning. Trong quá trình huấn luyện, hàm loss được sử dụng cho cả hai mạng là hàm Binary-Crossentropy và hàm tối ưu Adam. Vì các trọng số của cả hai mạng đều đã được tối ưu cho tập dữ liệu trước đó, nên khi tái huấn luyện, giá trị learning rate được chọn khá nhỏ cụ thể là 1e-4. Quá trình huấn luyện còn áp dụng kỹ thuật dừng huấn luyện sớm khi giá trị validation loss tốt nhất không giảm ba mươi lần liên tiếp. Quá trình huấn luyện mạng VGG16 dừng ở epoch 67, biểu đồ giá trị loss được thể hiện ở hình [6(a)](#fig_train_vgg) và AUC được thể hiện ở hình [6(b).](#fig_train_vgg) Quá trình huấn luyện mạng ResNet dừng ở epoch 93, biểu đồ loss được thể hiện ở hình [7(a)](#fig_train_resnet) và AUC được thể hiện ở hình [7(b)](#fig_train_resnet).

(03)

Trong đó, vector là vector kết quả dự đoán, vector là vector chứa các nhãn thực của dữ liệu và là tổng số lượng dữ liệu.

Thuật toán tối ưu Adam giúp giải quyết các vấn đề như phụ thuộc vào learning rate và nghiệm của hàm loss có thể rơi vào local minimum. Thuật toán Adam duy trì trung bình bình phương độ dốc (gradient) trước đó ( và cũng đồng thời duy trì trung bình độ dốc trước đó (.

(04)

(05)

(06)

(07)

(08)

Trong đó, , là hai vector moment trong Adam, là đạo hàm tại thời điểm , là learning rate. Giá trị của được lựa chọn phổ biến là 0.9, 0.99 và .

* 1. Metrics đánh giá mô hình

Sau khi hoàn tất quá trình huấn luyện, hiệu suất của hai mạng VGG16 và ResNet18 được đánh giá trên các chỉ số hiệu suất sau đây: precision, recall, F1 score và Area Under the Curve (AUC) hay còn gọi là Average precision (AP).

Đại lượng AUC này là diện tích nằm bên dưới Precision-Recall Curve. Giá trị này là số dương nhỏ hơn hoặc bằng một. Giá trị này càng lớn thì mô hình càng tốt.

Precision thể hiện độ chính xác của mô hình trong khi Recall thể hiện độ nhạy của mô hình trong việc phân loại dữ liệu. Giá trị của precision và recall sẽ thay đổi dựa trên các giá trị threshold khác nhau, thông thường khi precision cao thì recall sẽ thấp và ngược lại. Ứng với mỗi threshold, ta sẽ thu được một cặp (precision, recall). Biểu diễn các điểm (precision, recall) trên đồ thị khi thay đổi threshold từ 0 đến 1 ta sẽ thu được một đường gọi là Precision – Recall Curve hay PR curve.

Giả sử có N threshold để tính precision và recall, với mỗi threshold cho một cặp giá trị (precision, recall) là Pn, Rn, n = 1, 2, ..., N. PR curve được vẽ bằng cách vẽ từng điểm có toạ độ (Rn, Pn) trên trục toạ độ và nối chúng với nhau:

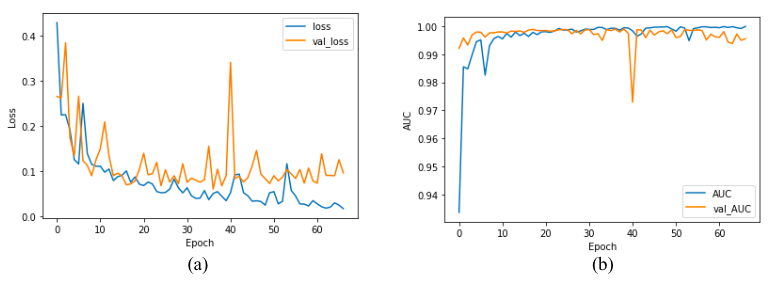
(09)

(10)

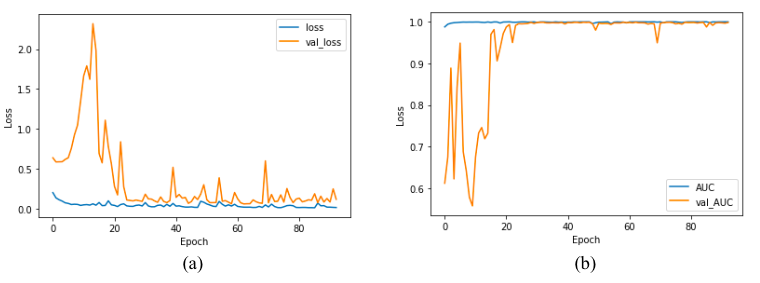
AUC được tính bằng công thức:

(11)

Trong đó, là diện tích hình chữ nhật có chiều rộng và chiều cao



1. Biểu đồ (a) loss và (b) AUC huấn luyện mạng VGG16.



1. Biểu đồ (a) loss và (b) AUC huấn luyện mạng ResNet18.

(12)

Các giá trị true positive (TP), true negative (TN), false positive (FP) và false negative (FN) được thu thập từ confusion matrix. Trong đó, TP được sử dụng để biểu thị số lượng hình ảnh viêm phổi được chẩn đoán là viêm phổi, TN biểu thị số lượng hình ảnh bình thường được chẩn đoán là bình thường, FP biểu thị số lượng hình ảnh bình thường được chẩn đoán là viêm phổi và FN biểu thị số lượng hình ảnh viêm phổi được chuẩn đoán là bình thường.

1. Kết quả

Các giá trị training AUC, training loss, validation AUC và validation loss của hai mạng được thể hiện ở Bảng [3](#table_loss_auc).

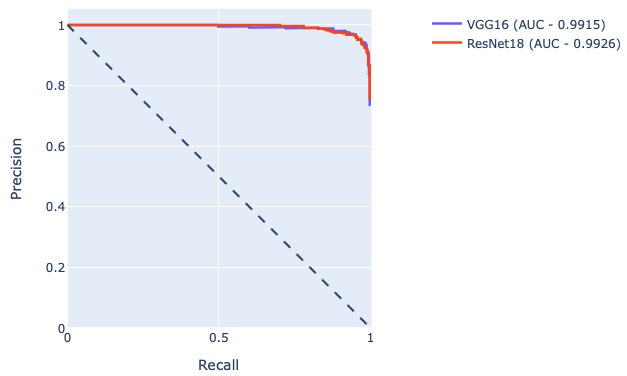
Sau khi hoàn tất quả trình huấn luyện, độ chính xác của mạng được đánh giá trên tập test. Precision – Recall curve của hai mạng được thể hiện ở Hình [8](#fig_prcurve). Giá trị AUC của VGG16 = 99.15% và giá trị AUC của ResNet18 = 99.26%. Với mạng VGG16 giá trị threshold tốt nhất tìm được là 0.7134 và giá trị threshold tốt nhất tìm được của mạng ResNet18 là 0.8163.

Các chỉ số AUC, precision, recall, F1 score dựa trên giá trị threshold tốt nhất được thể hiện ở Bảng [4](#table_metric). Cả hai mô hình đều có hiệu suất rất tốt và tương đương nhau, với các giá trị AUC và F1 score đều gần bằng nhau. Nhìn vào giá trị threshold tốt nhất tìm được ở cả hai mạng, ta có thể thấy rất rõ sự đánh đổi giữa Precision và Recall, cụ thể, lựa chọn threshold thấp hơn thì precision sẽ thấp hơn, recall cao hơn và ngược lại.

Bảng [5](#table_compare) tóm tắt so sánh mức độ hiệu quả của các công trình trong việc chẩn đoán bệnh viêm phổi.

1. Kết quả thu được sau khi huấn luyện ở hai mạng.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mạng | Training AUC | Training loss | Validation AUC | Validation loss |
| VGG16 | 99.93% | 0.05 | 99.90% | 0.06 |
| ResNet18 | 99.99% | 0.0178 | 99.90% | 0.058 |



1. Đường cong PR curve của mạng VGG16 (màu xanh) và PR curve của ResNet18 (màu đỏ).

Hình [9(a)](#fig_confusionmatrix) là confusion matrix của VGG16 và Hình [9(b)](#fig_confusionmatrix) là confusion matrix của ResNet18. Trong số 340 bệnh nhân bị bệnh viêm phổi, VGG16 chẩn đoán đúng 384 trường hợp và 6 trường hợp bị chẩn đoán nhầm thành không có bệnh, trong khi đó, ResNet chẩn đoán đúng 378 trường hợp và 12 trường hợp bị chẩn đoán nhầm thành không có bệnh. Trong số 234 trường hợp bệnh nhân không mắc bệnh, VGG16 chẩn đoán đúng 209 trường hợp và 25 trường hợp bị chẩn đoán nhầm thành có bệnh, ResNet18 chẩn đoán đúng 215 trường hợp và 19 trường hợp bị chẩn đoán nhầm thành có bệnh. Số lượng FP của VGG16 nhiều hơn ResNet18 và số lượng FN của ResNet18 nhiều hơn VGG16 là bởi threshold tốt nhất của ResNet18 lớn hơn threshold tốt nhất của VGG16.

1. Kết luận

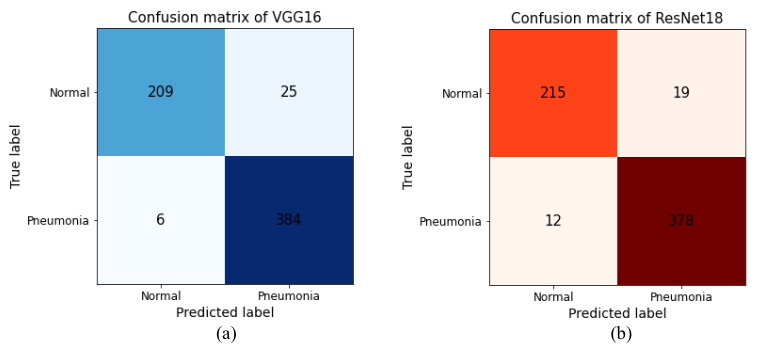
Bài nghiên cứu này trình bày phương pháp áp dụng kỹ thuật transfer learning để có thể chẩn đoán bệnh viêm phổi. Hai mạng VGG16 và ResNet50 đã được đào tạo và thử nghiệm để phân loại bệnh nhân mắc bệnh viêm phổi hay có sức khoẻ bình thường bằng cách sử dụng hình chụp X-quang ngực. Các chỉ số hiệu suất đã cho thấy mạng VGG16 có hiệu suất tốt hơn so với mạng ResNet. Giá trị AUC, precision, recall, f1 score sau khi huấn luyện mạng VGG16 lần lượt là 99.2%, 94.3%, 97.9% và 96.1%. Nghiên cứu này đã cho thấy rằng các mô hình mạng DL và hình ảnh X-quang có thể trích xuất các dấu hiệu sinh học liên quan đến căn bệnh viêm phổi. Có hàng triệu trẻ em chết mỗi năm do căn bệnh có thể gây tử vong này. Can thiệp kịp thời với kế hoạch điều trị thích hợp thông qua chẩn đoán chính xác có thể cứu sống rất nhiều người trên thế giới.

1. Các chỉ số đánh giá hiệu suất của VGG16 và ResNet18.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Mạng | AUC | Precision | Recall | F1 score | Threshold tốt nhất |
| VGG16 | 99.15% | 93.9% | 98.5% | 96.1% | 0.7134 |
| ResNet18 | 99.26% | 95.2% | 96.9% | 96.06% | 0.8163 |

1. So sánh hiệu suất của các công trình chẩn đoán bệnh viêm phổi.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tác giả | Kỹ thuật | Số lượng ảnh | Precision (%) | Recall (%) | AUC (%) | Accuracy (%) |
| Okeke S. et al. [[19](#ref_buil_cnn)] | Xây dựng mạng CNN | 5856 | - | - | - | 93.73% |
| Cohen J.P et al. [[20](#ref_inception)] | Pretrained InceptionV3 | 5232 | 90.1 | 93.2 | - | 92.8% |
| Rajaraman S. et al. [[21](#ref_customize_vgg16)] | Tinh chỉnh VGG16 | 5856 | 97 | 99.5 | 99 | - |
| Rahman T. et al [[22](#ref_densenet)] | Pretrained DenseNet201 | 5247 | 97 | 99 | 98 | - |
| Nghiên cứu này | Pretrained VGG16 | 5910 | 93.9 | 98.5 | 99.15 | - |
| Nghiên cứu này | Pretrained ResNet18 | 5910 | 95.2 | 96.9 | 99.26 | - |



1. Confusion matrix của (a) VGG16 và (b) ResNet18.

TÀI LIỆU THAM KHẢO

1. Vietnam Vaccine JSC, “Bệnh viêm phổi: nguyên nhân, triệu chứng, điều trị và cách phòng ngừa,” VNVC, [https://vnvc.vn/benh-viem-phoi-nguyen-nhan-trieu-chung-dieu-tri-va-cach-phong-ngua](https://vnvc.vn/benh-viem-phoi-nguyen-nhan-trieu-chung-dieu-tri-va-cach-phong-ngua/).
2. Ruuskanen O., Ruuskanen E., Jennings L.C và Murdoch D.R, “Viral pneumonia,” Lancet, tháng 03 năm 2011.
3. Lodha R., Kabra S.K và Pandey R.M, “Antibiotics for community‐acquired pneumonia in children,” Cochrane Database of Systematic Reviews, tháng 06 năm 2013.
4. Harikrishnan S., Jeemon P., Mini G.K., Thankappan K.R. và Sylaja P., “GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017,” Lancet, tháng 11 năm 2018.
5. Benjamin P. và Eliot A., “Hands-On Computer Vision with TensorFlow 2,” Packt Publishing, tr. 24, tháng 05 năm 2019.
6. Mohamed E., “Deep Learning for Vision Systems,” Manning Publications, tr. 32, 2020.
7. Yann LeCun, Léon B., Yoshua B. và Patrick H., “Gradient-Based Learning Applied to Document Recognition,” Proceedings of the IEEE, tháng 11 năm 1998.
8. Alex K., Ilya S. và Geoffrey E.H., “ImageNet Classification with Deep Convolutional Neural Networks,” Communications of the ACM, vol. 60, tháng 06 năm 2017.
9. Karen S. và Andrew Z., “Very Deep Convolutional Networks for Large-Scale Image Recognition,” ICLR, 2014.
10. Christian S., Wei Liu, Yangqing Jia, Pierre S., Scott R., Dragomir A., Dumitru E., Vincent V. và Andrew R., “Going Deeper with Convolutions,” trong Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.
11. Kaiming H., Xiangyu Z., Shaoqing R. và Jian Sun, “Deep Residual Learning for Image Recognition,” 2015.
12. Jason Y., Jeff C., Yoshua B. và Hod L., “How Transferable Are Features in Deep Neural Networks?,” Advances in Neural Information Processing Systems, 2014.
13. Mohamed E., “Deep Learning for Vision Systems,” Manning Publications, tr. 241 - 242, 2020.
14. Andre E., Brett K., Roberto A.N., Justin Ko, Susan M.S., Helen M.B. và Sebastian T., “Dermatologist-level classification of skin cancer with deep neural networks,” Nature, vol. 542, tháng 01 năm 2017.
15. Pranav R., Awni Y.H., Masoumeh H., Codie B. và Andrew Y. Ng, “Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks,” tháng 07 năm 2017.
16. Varun G., Lily P., Marc C., Martin C.S., Derek Wu, Arunachalam N., Subhashini V., Kasumi W., Tom M., Jorge C., Kim R., Rajiv R., Philip C.N., Jessica L.M. và Dale R.W., “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” JAMA, vol. 316, no. 22, tháng 12 năm 2016.
17. Huang P., Park S., Yan R., Lee J., Linda C.C., Cheng T.L., Amira H., Joshua R., Brett T., Chen C., Russel H., David S.E., Malcolm B., Ping Hu, Elliot K.F., Edward G. và Stephen L., “Added Value of Computer-aided CT Image Features for Early Lung Cancer Diagnosis with Small Pulmonary Nodules: A Matched Case Control Study,” Radiology, vol. 286, no. 1, tháng 01 năm 2018.
18. Lakhani P. và Sundaram B., “Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks,” Radiology, vol. 284, no. 2, 2017.
19. Okeke S., Mangal S., Uchenna J. và Do-Un J., “An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare,” Journal of Healthcare Engineering, tháng 03 năm 2019.
20. Cohen J.P., Bertin P., Frappier V, “V. Chester: A Web Delivered Locally Computed Chest X-ray Disease Prediction System,” năm 2019.
21. Rajaraman S., Sema C., Kim I., Geogre T. và Sameer A., “Visualization and Interpretation of Convolutional Neural Network Predictions in Detecting Pneumonia in Pediatric Chest Radiographs,” Appl. Sci., tháng 09 năm 2018.
22. Rahman T., Amith K., Islam K.R., Zaid B.M., Muhammad A.K và Saad K., “Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray,” Appl. Sci., tháng 05 năm 2020.
23. Luis P. và Jason W., “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” tháng 12 năm 2017.