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**

......

**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)].

1. Phương pháp
   1. Tập dữ liệu

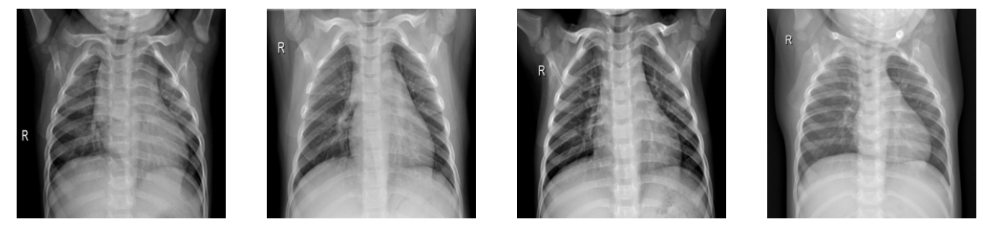
Trong nghiên cứu này, tập dữ liệu hình X-quang ngực được tải xuống từ <https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset>. Tập dữ liệu gốc bao gồm hai thư mục là thư mục train và thư mục test chứa các file ảnh X-quang ngực, cùng với hai tệp csv, trong đó file metadata.csv lưu toàn bộ thông 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](#table01) thể hiện số lượng dữ liệu của mỗi lớp để huẩn luyện và kiểm tra mô hình. Hình [1](#fig01) là mẫu các hình X-quang được gán nhãn bình thường và Hình [2](#fig02) là mẫu các hình X-quang được gán nhãn mắc bệnh viêm phổi của tập dữ liệu.

* 1. Tiền xử lý dữ liệu và kỹ thuật augmentation

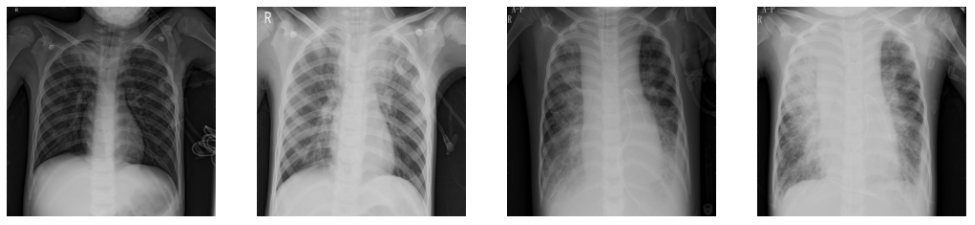
Tất cả ảnh trước khi được đưa vào mô hình đều được resize về một kích thước cố định là 224 x 224 x 3 và giá trị của các pixel được chuẩn hoá trong khoảng từ 0 đến 1.

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

|  |  |  |  |
| --- | --- | --- | --- |
| Nhãn | Tập train | Tập test | Tổng |
| Bình thường | 1342 | 234 | 1576 |
| Viêm phổi | 3944 | 390 | 4334 |
|  |  |  | 5910 |



1. Các ảnh được gán nhãn bình thường.



1. Các ảnh được gán nhãn mắc bệnh viêm phổi.

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, ... [[19](#ref19)]. 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](#aug_img) 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 để áp dụng augmentation được thể hiện ở Bảng [2](#table02).



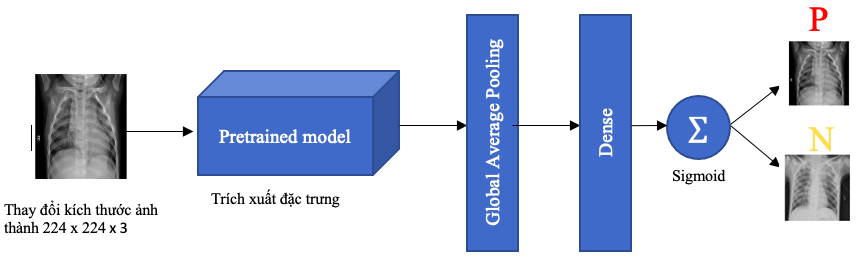
1. Ảnh sau khi áp dụng kỹ thuật augmentation.
2. Chi tiết các thông số áp dụng augmentation.

|  |  |
| --- | --- |
| Phép biển đổi ảnh | Thông số |
| Xoay ảnh | 10 |
| Thay đổi độ sáng | (0.1, 1.2) |
| Phóng to ảnh | (0.75, 1) |
| Lật ảnh ngang | True |

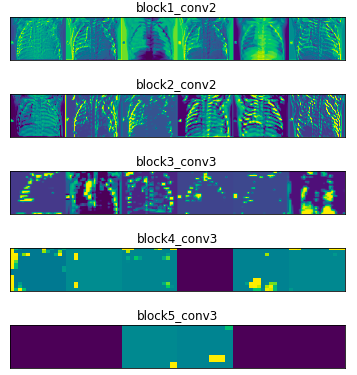
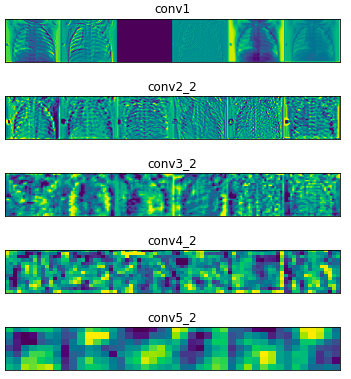
* 1. Pretrained VGG16 và ResNet18

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. Cả hai mạng pretrained đều được tái huấn luyện từ đầu trên bộ dữ liệu mới, với giá trị khởi tạo của các trọng số là giá trị đã được huấn luyện trên bộ dữ liệu ImageNet. Theo sau lớp Convolution cuối cùng là lớp Global Average Pooling (GAP), đầu vào của lớp GAP sẽ là feature map có kích thước H x W x D và kết quả trả ra là vector có kích thước 1 x 1 x D bằng cách lấy trung bình của mỗi feature map. Nhờ vào cơ chế lấy giá trị trung bình của feature map nên lớp GAP không chứa bất kỳ tham số nào. Vector được trả về từ lớp GAP sẽ được đưa vào lớp cuối cùng và trả về kết quả dự đoán. Vì các mạng trên không chứa bất kỳ lớp FC nào nên mạng được gọi là Fully Convolutional Network (FCN). Cả hai mạng đều sử dụng hàm tối ưu Adam [[20](#ref_adam)] với hệ số learning rate ban đầu là 0.0001.

Hình [4](#fig_architecture) là tổng quan kiến trúc của cả 2 mô hình khi áp dụng kỹ thuật transfer learning. Hình [5](#fig_featuremap) các feature map sau khi qua các khối Convolution của hai mạng.



1. Tổng quan kiến trúc áp dụng transfer learning.



1. Hình bên trái là các feature map sau khi đi qua các khổi Convolution của mạng ResNet18 và bên phải là các feature map sau khi đi qua mạng VGG16.
   1. Metrics đánh giá mô hình

Vì tập dữ liệu được sử dụng không được cân bằng nên để có thể đánh giá mô hình một cách chính xác nhất ta phải sử dụng chỉ số Area Under the Curve (AUC) hay còn gọi là Average presion (AP). Đại lượng 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. Hình [6](#fig_auc) là ví dụ về PR curve và AUC.

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:

(1)

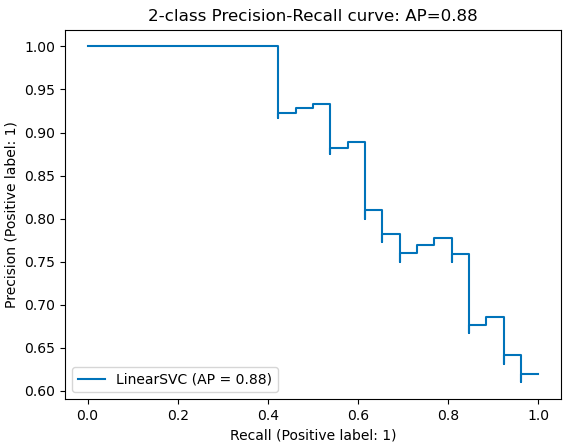
(2)

AUC được xác định bằng:

(3)

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

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. Ví dụ về PR curve và AUC.
2. Kết quả
3. Kết luận

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. Luis P. và Jason W., “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” tháng 12 năm 2017.
20. Diederik P.K. và Jimmy L.B., “ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION,” ICLR, 2015.