

NHẬN DẠNG VIÊM PHỔI QUA PHIM CHỤP X-QUANG SỬ DỤNG KỸ THUẬT DEEP LEARNING PNEUMONIA RECOGNITION THROUGH X-RAY IMAGING USING DEEP LEARNING TECHNIQUES

NGUYỄN QUAN KHÁNH^{1,a}, ĐẶNG THỊ XUÂN TIÊN², TRẦN VĂN THUẬN³,
ThS. TỪ VĂN TRỌNG⁴, TS. TRẦN THỊ TỐ QUYÊN⁵, TS. NGUYỄN VŨ LÂM⁶

¹Trường Cao đẳng Kiên Giang

²Trường Tiểu học Tân Hạnh B Vĩnh Long

³Trường Cao đẳng nghề Vĩnh Long

⁴Trường Cao đẳng Kiên Giang

⁵Trường Đại học Cần Thơ

⁶Trường Cao đẳng Kiên Giang

^aTác giả liên hệ: nqkhanh@kgc.edu.vn

Nhận bài (Received): 03/03/2024; Phản biện (Reviewed): 28/03/2024; Chấp nhận (Accepted): 08/04/2024

TÓM TẮT

Viêm phổi là một căn bệnh phổ biến và nguy hiểm, có thể gây tử vong và tác động lớn đến sức khỏe cộng đồng. X-quang ngực là một phương pháp chẩn đoán không xâm lấn và phổ biến để xác định viêm phổi. Phát hiện và chẩn đoán viêm phổi một cách chính xác là rất quan trọng để đảm bảo điều trị kịp thời và ngăn chặn sự lây lan của căn bệnh. Trong bối cảnh này, giải thuật học sâu, đặc biệt là mạng nơ-ron tích chập (CNN), đã được áp dụng để nhận dạng viêm phổi qua hình ảnh X-quang. Trong bài báo này, chúng tôi trình bày kết quả của việc sử dụng hai mô hình Deep Learning phổ biến là VGG16 và ResNet 50V2 để nhận dạng viêm phổi trên tập dữ liệu X-quang. Tập dữ liệu được sử dụng bao gồm 5.856 hình ảnh X-quang, chia thành ba thư mục: train (5216 hình), test (624 hình), và val (16 hình). Trong đó, có 2 loại chính là Viêm phổi (4.273 hình) và Bình thường (1.583 hình). Kết quả thử nghiệm cho thấy hiệu suất ấn tượng của hai mô hình này trong việc nhận dạng viêm phổi, với VGG16 đạt 89.1% và ResNet 50V2 đạt 90.7%. Nghiên cứu này cung cấp một cái nhìn quan trọng về việc phát triển các mô hình hiệu quả và hiệu suất cao vào việc cải thiện quá trình chẩn đoán và điều trị cho bệnh nhân viêm phổi trong lĩnh vực y học.

Từ khóa: Viêm phổi, X-quang, Deep Learning, CNN, VGG16, ResNet 50V2, Nhận dạng hình ảnh.

ABSTRACT

Pneumonia is a common and dangerous disease that can lead to death and have significant impacts on public health. Chest X-ray is a non-invasive and widely used diagnostic method for identifying pneumonia. Accurate detection and diagnosis of pneumonia are crucial to ensure timely treatment and prevent the spread of the disease. In this context, deep learning algorithms, especially convolutional neural networks (CNNs),

have been applied to identify pneumonia through X-ray images. In this paper, we present the results of using two popular deep learning models, VGG16 and ResNet 50V2, to identify pneumonia in X-ray datasets. The dataset used consists of 5,856 X-ray images, divided into three folders: train (5,216 images), test (624 images), and val (16 images). There are two main categories: Pneumonia (4,273 images) and Normal (1,583 images). The experimental results show impressive performance of these two models in pneumonia identification, with VGG16 achieving 89.1% accuracy and ResNet 50V2 achieving 90.7%. This study provides important insights into the development of effective and high-performance models to improve the diagnosis and treatment process for pneumonia patients in the medical field.

Keywords: Pneumonia, X-ray, Deep Learning, CNN, VGG16, ResNet 50V2, Image Recognition..

1. MỞ ĐẦU

1.1. Giới thiệu

Việc sử dụng giải thuật học sâu để phát hiện và chẩn đoán viêm phổi qua phim chụp X-quang mang lại nhiều lợi ích quan trọng. Đầu tiên, nó giúp tăng tốc quá trình chẩn đoán và giảm thời gian đợi của bệnh nhân. Thứ hai, việc sử dụng mô hình tự động có khả năng giảm thiểu sai sót của con người và đảm bảo tính nhất quán trong việc chẩn đoán. Cuối cùng, nó mở ra cánh cửa cho việc tăng cường khả năng tiếp cận chẩn đoán y tế ở các khu vực có nguồn lực y tế hạn chế.

Chẩn đoán viêm phổi một cách kịp thời và chính xác là vô cùng quan trọng để đảm bảo can thiệp kịp thời và ngăn chặn sự lây lan của căn bệnh. Những tiến bộ gần đây trong hình ảnh y tế và học máy đã tạo điều kiện cho việc phát triển các thuật toán tinh vi có khả năng phân biệt những mẫu tinh tế biểu hiện viêm phổi trong hình ảnh X-quang. Bằng cách tận dụng các bộ dữ liệu lớn bao gồm hàng ngàn hình ảnh X-quang, nghiên cứu này cố gắng đánh giá hiệu quả của các mô hình VGG16 và ResNet 50V2 trong việc phân loại chính xác các trường hợp viêm phổi. Bộ dữ liệu đã được tổ chức và phân loại một cách cẩn

thận thành hai nhóm chính là viêm phổi và bình thường, đó là nền tảng nghiên cứu của chúng tôi.

Thông qua thử nghiệm nghiêm túc và phân tích, nghiên cứu này mong muốn làm sáng tỏ về hiệu quả của các phương pháp học sâu và nhằm mục đích đóng góp vào lĩnh vực chẩn đoán y học bằng việc đánh giá hiệu suất của chúng trong việc nhận diện viêm phổi.

1.2. Những nghiên cứu liên quan

Enes AYAN và cộng sự. [1] Chẩn đoán bệnh viêm phổi từ hình ảnh X-quang ngực bằng phương pháp học sâu. Sử dụng hai mô hình mạng lưới thần kinh tích chập nổi tiếng Xception và VGG16 để chẩn đoán bệnh viêm phổi với phương pháp học chuyển giao và tinh chỉnh trong giai đoạn huấn luyện; trong đó VGG16 deep learning model độ chính xác là: 84.5%.

Xianghong Gu và cộng sự. [2] Sử dụng Deep Convolutional neural network (CNN); phương pháp đề xuất được đánh giá trên bộ dữ liệu của Trung tâm Y tế Phụ nữ và Trẻ em Quảng Châu, Trung Quốc, với tổng số 4.513 bệnh nhân nhi, từ 1 đến 9 tuổi, trong giai đoạn từ 2003 đến 2017. Độ chính xác mà tác giả đạt được là: 80.4% .

Archit Khatri và cộng sự. [3] Sử dụng

tiếp cận Earth Mover's Distance (EMD) với số hình ảnh sử dụng: 144 hình ảnh để xác định chính xác phổi bị viêm phổi bị nhiễm trùng với phổi không bị nhiễm trùng bình thường trong đó tiền xử lý các hình ảnh để chỉ có hình ảnh của phổi, sau đó thực hiện một số tỷ lệ, xoay và chuẩn hóa cường độ để có một bộ ảnh X-quang phổi có kích thước/hình dạng đồng nhất, sau đó tính toán EMD và so sánh kết quả với độ chính xác là: 83.3%.

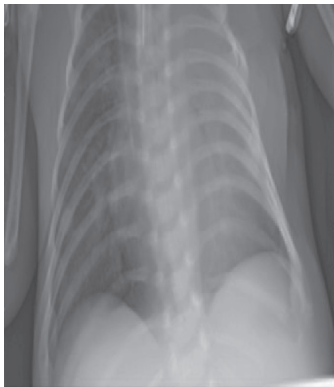
1.3. Đặc điểm của bệnh viêm phổi

Viêm phổi là một bệnh nhiễm trùng gây viêm ở túi khí của một hoặc cả hai phổi. Túi khí có thể bị nhiễm dịch hoặc mủ, dẫn đến triệu chứng như ho kèm đờm hoặc mủ, sốt, cảm giác ớn lạnh và khó thở. Nguyên

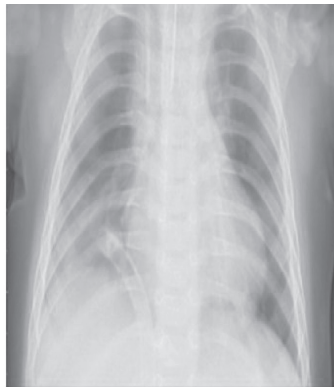
nhân của viêm phổi có thể là do nhiều loại sinh vật như vi khuẩn, vi rút và nấm.

Bệnh viêm phổi có thể có mức độ nghiêm trọng từ nhẹ đến đe dọa tính mạng. Người mắc bệnh nghiêm trọng nhất thường là trẻ sơ sinh và trẻ nhỏ, người cao tuổi trên 65 tuổi, cũng như những người có vấn đề về sức khỏe hoặc hệ miễn dịch suy yếu. [4, 5]

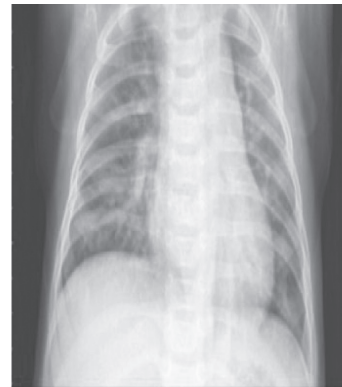
Thường thì, bác sĩ sẽ đặt chẩn đoán dựa trên các triệu chứng lâm sàng và kết quả của các xét nghiệm hình ảnh như X-quang ngực. Ngoài ra, chẩn đoán cũng có thể dựa trên các triệu chứng lâm sàng và kết quả của các xét nghiệm phân tích dịch phế quản hoặc mẫu nhày cổ.



Viêm phổi do Vi khuẩn



Viêm phổi do Vi khuẩn



Viêm phổi do Virus

Hình 1: Hình ảnh một số bệnh viêm phổi do vi khuẩn hoặc virus

1.4. Vai trò của học sâu trong nhận dạng bệnh viêm phổi

Học sâu, đặc biệt là mạng nơ-ron tích chập (CNN), đóng vai trò quan trọng trong nhận dạng bệnh viêm phổi bằng cách cung cấp các công cụ mạnh mẽ để phân tích và hiểu hình ảnh y khoa. Việc áp dụng mô hình học sâu CNN cho hình ảnh X-quang ngực có thể cải thiện đáng kể độ chính xác trong việc nhận dạng các dấu hiệu của bệnh viêm phổi. [6]

Các nhà nghiên cứu đã chỉ ra rằng việc sử dụng mô hình CNN để nhận dạng bệnh viêm phổi trên hình ảnh CT scanner đã đạt được hiệu quả cao, vượt qua khả năng nhận dạng của các phương pháp truyền thống. [7]

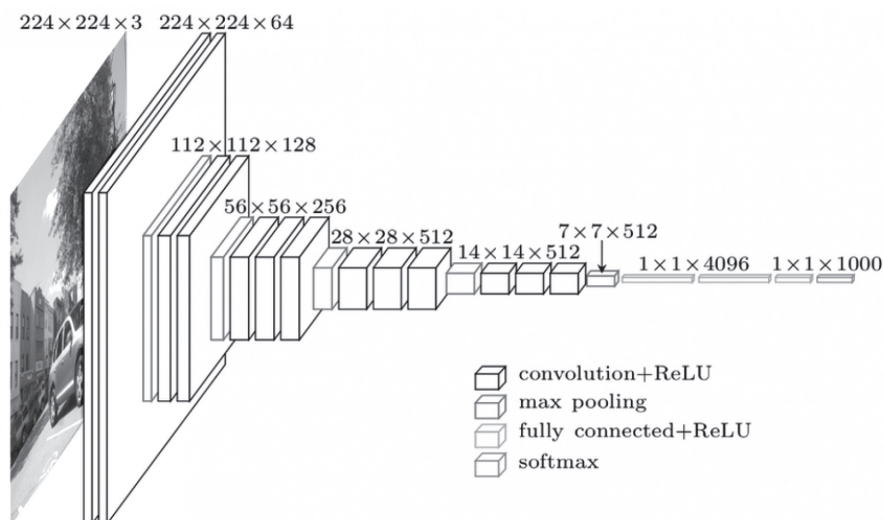
1.5. Các mạng nơ-ron trích xuất đặc trưng

1.5.1 VGG16

Kiến trúc của VGG16 [8] (Very Deep

Convolutional Networks for Large-Scale Image Recognition) bao gồm một chuỗi các lớp convolutional (tích chập), lớp

pooling và các lớp fully connected (hoàn toàn kết nối). Dưới đây là chi tiết kiến trúc của VGG16:



Hình 2: VGG16 với 16 lớp, bao gồm các lớp tích chập, lớp pooling, và các lớp fully connected (Nguồn: [11])

Input Layer: Kích thước: 224x224 pixel (đối với hình ảnh RGB). Số kênh (channels): 3 (một cho mỗi màu đỏ, lục, và xanh).

Convolutional Layers: Gồm 13 lớp convolutional, mỗi lớp được kích thước 3x3. Số lượng bộ lọc (filters) tăng dần từ 64 ở lớp đầu tiên lên đến 512 ở lớp cuối cùng. Hàm kích hoạt: Rectified Linear Unit (ReLU).

Pooling Layers: Sử dụng lớp pooling kích thước 2x2 với bước (stride) là 2 để giảm kích thước không gian của đặc trưng.

Fully Connected Layers: Bao gồm 3 lớp fully connected với 4096 neuron ở mỗi lớp. Hàm kích hoạt: ReLU (trong các lớp ẩn). Lớp đầu ra có 1000 neuron (tương ứng với số lượng lớp trong tập dữ liệu ImageNet). Lớp đầu ra sử dụng hàm kích hoạt Softmax để tạo xác suất cho việc phân loại.

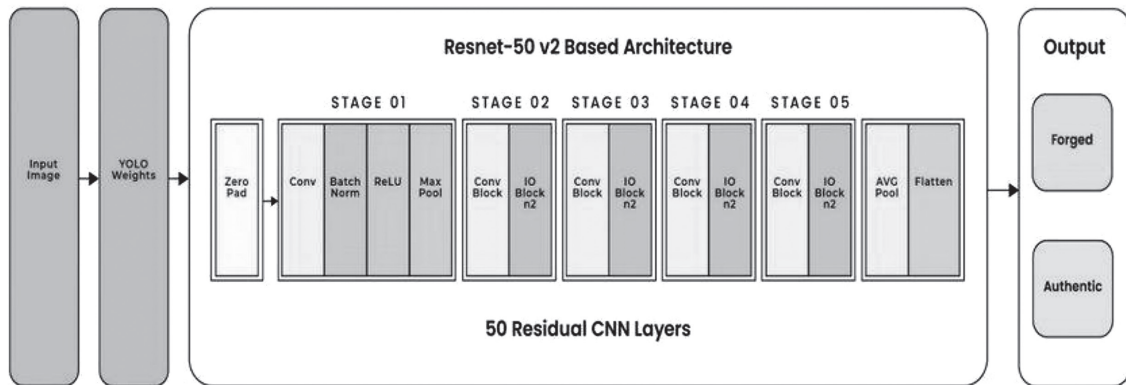
Dropout: Một kỹ thuật regularization

được áp dụng để giảm nguy cơ overfitting. Một tỷ lệ dropout 50% được sử dụng trước các lớp fully connected.

1.5.2 ResNet 50V2

ResNet 50V2 sử dụng khái niệm của các khối residue để xây dựng một mạng nơ-ron sâu mà vẫn có khả năng huấn luyện tốt và tránh được hiện tượng biến mất đạo hàm (vanishing gradient). Các khối residue cho phép thông tin từ các lớp trước đó được truyền tiếp qua mạng mà không bị giảm mất trong quá trình lan truyền ngược. Điều này giúp mạng nơ-ron có thể học được các đặc trưng phức tạp từ dữ liệu.[9]

Kiến trúc chính: ResNet 50V2 có một kiến trúc sâu với tổng cộng 50 lớp, bao gồm các lớp convolutional, batch normalization và các lớp fully connected. Kiến trúc này sử dụng các khối cơ bản có cấu trúc residual (còn được gọi là identity blocks) để tránh vấn đề biến mất đạo hàm trong quá trình lan truyền ngược.



Hình 3: ResNet 50V2 kiến trúc sâu với tổng cộng 50 lớp (Nguồn: [12])

Identity blocks: Trong ResNet 50V2, các identity blocks cho phép thông tin truyền qua mà không bị biến dạng nhiều. Điều này giúp cho việc huấn luyện mạng nơ-ron sâu trở nên dễ dàng hơn và giảm thiểu hiện tượng biến mất thông tin.

Cải tiến Batch Normalization: ResNet 50V2 sử dụng Batch Normalization không chỉ sau mỗi lớp convolutional mà còn sau mỗi hàm kích hoạt (ReLU), giúp cải thiện tốc độ và độ chính xác của quá trình huấn luyện.

Khối bottleneck: Một đặc điểm quan trọng của ResNet 50V2 là việc sử dụng các khối bottleneck trong các lớp sâu của mạng, giúp giảm độ phức tạp tính toán mà vẫn duy trì hiệu suất của mô hình.

Regularization: ResNet 50V2 thường sử dụng các kỹ thuật regularization như dropout hoặc weight decay để giảm nguy cơ overfitting trong quá trình huấn luyện.

Kích thước đầu vào: ResNet 50V2 có thể chấp nhận đầu vào có kích thước đa dạng, nhưng thường được sử dụng với các hình ảnh có kích thước 224x224 pixel.

1.6 Các độ đo đánh giá mô hình

Positive Rate (Tỷ lệ dự đoán tích cực): Tỷ lệ giữa số lần mô hình dự đoán tích cực đúng (True Positive - TP) và tổng số dự

đoán tích cực (TP + False Positive - FP). Nó cho biết khả năng của mô hình trong việc dự đoán chính xác các trường hợp tích cực.

$$Positive\ Rate = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)} \quad (1)$$

Recall (Độ bao phủ): Tỷ lệ phần trăm các trường hợp tích cực (TP) mà mô hình đã phát hiện được, tính bằng tỉ lệ giữa TP và số lượng thực sự của các trường hợp tích cực (TP + False Negative - FN). Recall đo lường khả năng của mô hình trong việc bắt kịp các trường hợp tích cực.

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \quad (2)$$

F1 Score: Là một trung bình điểm số hài hòa giữa precision và recall. Trong đó, Precision là tỷ lệ giữa số lần dự đoán tích cực đúng (TP) và tổng số lần dự đoán tích cực (TP + False Positive - FP).

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

Accuracy (Độ chính xác): Tỷ lệ giữa số lần dự đoán đúng và tổng số dự đoán, bao gồm cả các dự đoán tích cực (TP) và âm (TN), được tính bằng công thức:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

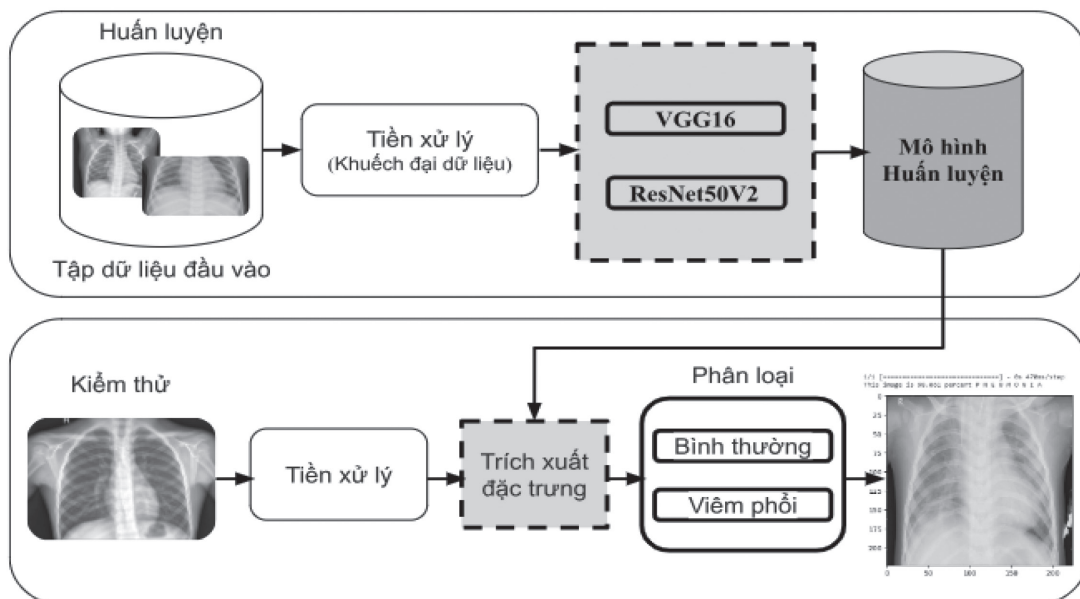
Độ chính xác đo lường tỷ lệ các dự đoán chính xác của mô hình trên toàn bộ dữ liệu. [10]

2. KẾT QUẢ NGHIÊN CỨU

2.1. Phương pháp đề xuất

Trong nghiên cứu này, chúng tôi đề

xuất mô hình tổng quát như Hình 4. Mô hình tổng quát gồm có hai giai đoạn: Giai đoạn huấn luyện và kiểm thử.



Hình 4: Mô hình đề xuất nhận dạng viêm phổi qua phim chụp x-quang

2.1.1 Huấn luyện mô hình

Để huấn luyện mô hình, chúng tôi đã áp dụng một loạt các phép biến đổi hình ảnh để tăng cường dữ liệu đầu vào; rotation: Ở mỗi lần huấn luyện, ảnh được xoay ngẫu nhiên với góc độ từ -30 đến $+30$ độ để tạo ra sự đa dạng trong dữ liệu đầu vào. Horizontal Flip và Vertical Flip: Ảnh có thể được lật ngang và lật dọc ngẫu nhiên để tạo ra thêm các biến thể của ảnh. Rescale: Giá trị của mỗi pixel trong ảnh đã được chuẩn hóa bằng cách chia cho 255. Điều này giúp tăng tốc độ quá trình huấn luyện và cải thiện hiệu suất của mô hình. ZCA Whitening: Kỹ thuật ZCA Whitening được áp dụng để giảm thiểu hiện tượng overfitting và giúp mô hình học được các đặc trưng chung từ dữ liệu. Chúng tôi đề xuất phương pháp VGG16 và ResNet 50V2; loại bỏ lớp fully connected ở đỉnh của mạng, chỉ giữ lại các lớp convolution và pooling, sử dụng trọng số được huấn luyện trước trên tập dữ liệu

ImageNet, kích thước đầu vào của ảnh là 224×224 pixel với 3 kênh màu (RGB), sử dụng max pooling để trích xuất đặc trưng từ các lớp convolution cuối cùng của mạng; số lớp đầu ra của mạng là 2, tương ứng với 2 lớp là “viêm phổi” và “bình thường”.

2.1.2 Kiểm thử

Ở giai đoạn kiểm thử, dữ liệu đầu vào là tập dữ liệu hình ảnh được mô tả trong phần 2.2.1 sau đó đưa qua các mô hình mạng được đề xuất từ giai đoạn thứ nhất để dự đoán.

2.2 Kết quả thực nghiệm

2.2.1 Các kịch bản áp dụng

Ở quá trình tiền xử lý, ảnh đầu vào sẽ được chuẩn hóa về kích thước thành 224×224 , sau đó đưa qua các mô hình mạng được đề xuất từ giai đoạn thứ nhất để dự đoán.

Chúng tôi đề xuất 4 kịch bản áp dụng với tham số huấn luyện như sau:

Bảng 1: Các kịch bản được đề xuất và các tham số huấn luyện

| Kịch bản | Kiến trúc mạng | Epochs | Steps per epoch | Batch Size | Input Shape | Learning rate | Class Mode | Pooling |
|----------|----------------|--------|-----------------|------------|-------------|---------------|------------|---------|
| 1 | VGG16 | 1000 | 1 | 16 | 224, 224, 3 | 0.001 | binary | max |
| 2 | VGG16 | 1000 | 1 | 32 | 224, 224, 3 | 0.001 | binary | max |
| 3 | ResNet 50V2 | 1000 | 1 | 16 | 224, 224, 3 | 0.001 | binary | max |
| 4 | ResNet 50V2 | 1000 | 1 | 32 | 224, 224, 3 | 0.001 | binary | max |

2.2.2. Môi trường cài đặt và tập dữ liệu thực nghiệm

a. *Môi trường cài đặt:* Hệ thống được cài đặt bằng ngôn ngữ Python và chạy trên cùng một môi trường Python 3 Google Compute Engine (T4 GPU, RAM 50GB, Disk 225GB). Thư viện hỗ trợ huấn luyện mô hình mạng sử dụng là Tensorflow và Keras.

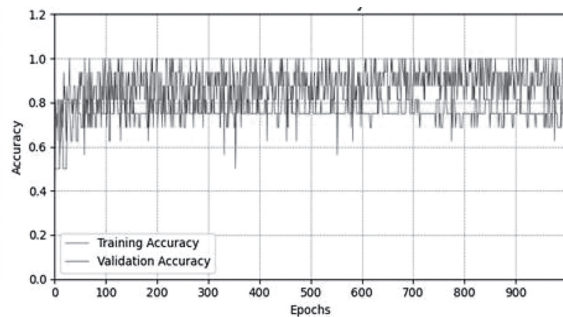
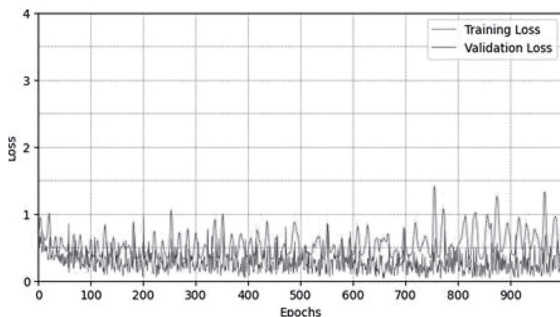
b. *Tập dữ liệu thực nghiệm:* Tập dữ liệu được sử dụng bao gồm 5.856 hình ảnh X-quang, chia thành ba thư mục: train (5216 hình), test (624 hình), và val (16 hình). Trong

đó, có 2 loại chính là Viêm phổi (4.273 hình) và Bình thường (1.583 hình). Hình ảnh X-quang ngực (trước-sau) được chọn từ đoàn hệ hồi cứu của bệnh nhân nhi từ một đến năm tuổi từ Trung tâm Y tế Phụ nữ và Trẻ em Quảng Châu. Phiên bản 2 được phát hành ngày 6 tháng 1 năm 2018.

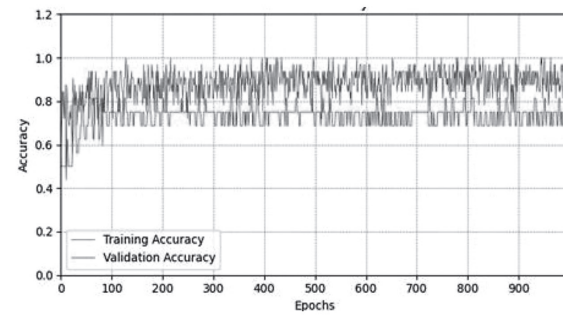
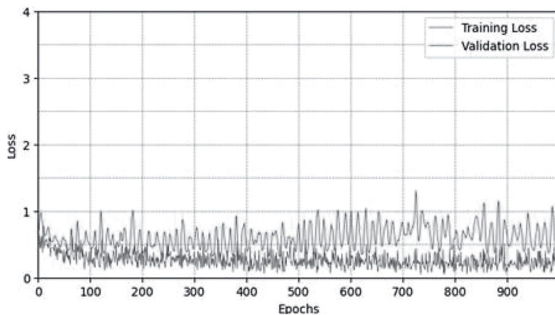
2.2.3. Các kết quả thực nghiệm

a. *Độ đo Loss và Accuracy*

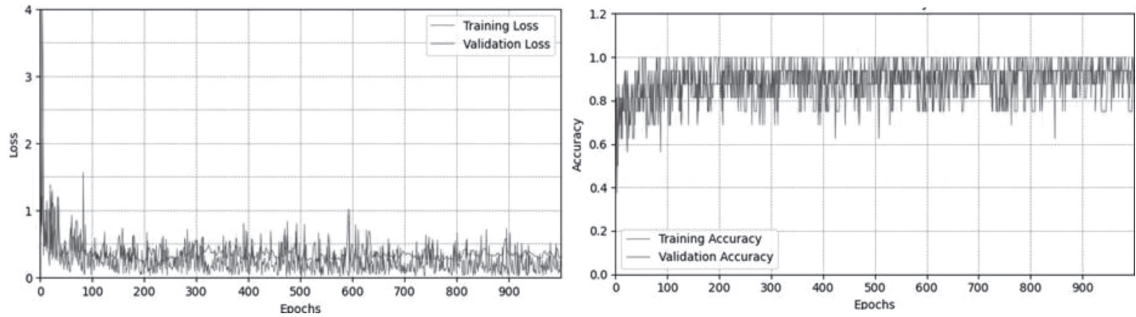
Sau quá trình huấn luyện Hình 5 thể hiện giá trị Loss và Accuracy của các kịch bản. Độ chính xác các kịch bản lần lượt là: 81.41%, 89.1%, 90.7%, 89.9%.



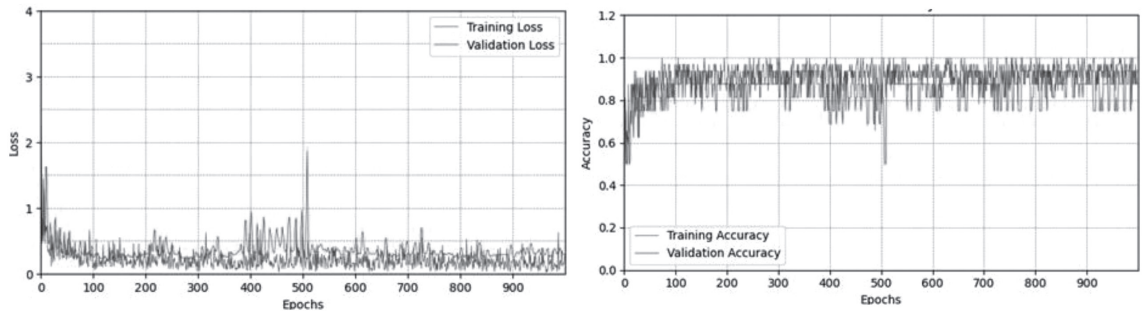
Hình 5 a. Biểu đồ độ đo Accuracy và Loss của VGG16, Batch Size 16 (Độ chính xác: 81.41%)



Hình 5 b. Biểu đồ độ đo Accuracy và Loss của VGG16, Batch Size 32 (Độ chính xác: 89.1%)



Hình 5 c. Biểu đồ độ đo Accuracy và Loss của ResNet 50V2, Batch Size 16 (Độ chính xác: 90.7%)

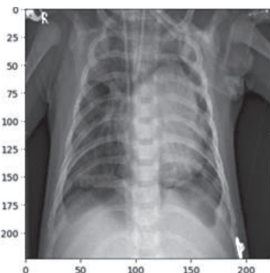


Hình 5 d. Biểu đồ độ đo Accuracy và Loss của ResNet 50V2, Batch Size 32 (Độ chính xác: 89.9%)

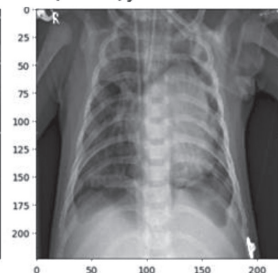
b. Một số kết quả thực nghiệm

Hình 6 trình bày một số kết quả thực nghiệm lần lượt: VGG16 (Batch 16),

Nhãn thực tế: VIÊM PHỔI
 Nhãn dự đoán: VIÊM PHỔI
 Độ tin cậy: 82.41%

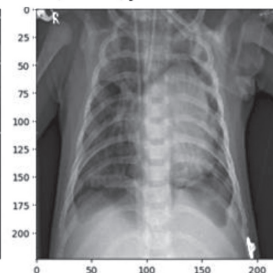


Nhãn thực tế: VIÊM PHỔI
 Nhãn dự đoán: VIÊM PHỔI
 Độ tin cậy: 93.47%

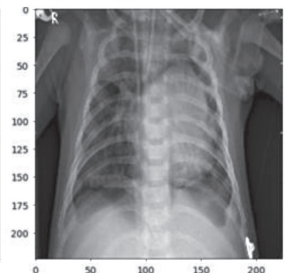


VGG16 (Batch 32), ResNet 50V2 (Batch 16), ResNet 50V2 (Batch 32) của quá trình nhận dạng:

Nhãn thực tế: VIÊM PHỔI
 Nhãn dự đoán: VIÊM PHỔI
 Độ tin cậy: 97.24%



Nhãn thực tế: VIÊM PHỔI
 Nhãn dự đoán: VIÊM PHỔI
 Độ tin cậy: 96.34%



Hình 6: Một số kết quả thực nghiệm

3. KẾT LUẬN

Trong nghiên cứu này kết quả đánh giá thử nghiệm; độ chính xác: ResNet 50 V2 đạt độ chính xác cao hơn VGG16 trong cả hai trường hợp Batch Size. Batch Size 16 cho độ chính xác thấp hơn Batch Size 32 trong cả hai kiến trúc. Độ chính xác cao nhất đạt được là 90.71% với ResNet 50 V2 và Batch Size 16. Mất mát: ResNet50V2

có mất mát thấp hơn VGG16 trong cả hai trường hợp Batch Size. Mất mát thấp nhất đạt được là 25.25% với ResNet50V2 và Batch Size 32. Biểu đồ: VGG16 có sự dao động lớn hơn trong độ chính xác và mất mát so với ResNet 50V2. Biểu đồ cho thấy sự hội tụ tốt cho cả hai kiến trúc mạng nơ-ron.

Sự khác biệt giữa các mô hình: ResNet

50 V2 được biết đến với cấu trúc “skip connection” hay “residual connections”, giúp giảm vấn đề biến mất gradient khi huấn luyện các mạng sâu. Điều này không những giúp mạng nhanh chóng hội tụ mà còn cải thiện đáng kể độ chính xác của mô hình. Trong khi đó, VGG16, một kiến trúc mạng nơ-ron sâu khác, dù có khả năng học sâu và rộng qua từng lớp, nhưng lại dễ gặp phải vấn đề về việc biến mất gradient, do đó làm hạn chế khả năng học của mạng.

Ảnh hưởng của batch size: Ngoài ra, việc thử nghiệm với hai Batch Size khác nhau (16 và 32) cũng đã mở ra những hiểu biết mới về ảnh hưởng của Batch Size đến quá trình huấn luyện. Mặc dù Batch Size 16 mang lại độ chính xác cao nhất cho ResNet 50 V2, sự chênh lệch giữa Batch Size 16 và 32 không quá lớn, điều này cho thấy sự linh hoạt và ổn định của ResNet 50 V2 dưới các điều kiện huấn luyện khác nhau. Điều này cũng phản ánh rằng, dù Batch Size lớn hơn

cung cấp nhiều thông tin hơn trong mỗi lần cập nhật trọng số, nhưng việc lựa chọn Batch Size phù hợp cũng cần cân nhắc đến khả năng tính toán và bộ nhớ sẵn có.

Qua biểu đồ độ chính xác và mất mát, rõ ràng là VGG16 có sự dao động lớn hơn so với ResNet 50 V2, điều này không chỉ cho thấy sự khác biệt về cấu trúc giữa hai mô hình mà còn phản ánh sự ổn định trong quá trình học của ResNet 50 V2. ResNet 50 V2 không chỉ đạt được sự hội tụ nhanh chóng mà còn duy trì được sự ổn định qua từng epoch, điều này là rất quan trọng trong việc phát triển một mô hình đáng tin cậy và hiệu quả.

Như vậy: ResNet 50 V2 cho hiệu suất tốt hơn VGG16 trong cả hai trường hợp Batch Size. Batch Size 32 cho hiệu suất tốt hơn Batch Size 16. Mô hình ResNet 50V2 với Batch Size 32 đạt được độ chính xác cao nhất (90.71%) và mất mát thấp nhất (25.25%).

TÀI LIỆU THAM KHẢO

- [1] Ayan, E.; Unver, H.M. (2019) *Diagnosis of Pneumonia from Chest X-ray Images Using Deep Learning*. In Proceedings of the 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 24–26 April 2019; pp. 1–5.
- [2] Gu, X.; Pan, L.; Liang, H.; Yang, R. (2018) *Classification of Bacterial and Viral Childhood Pneumonia Using Deep Learning in Chest Radiography*. In Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia—MUM '04; Association for Computing Machinery (ACM), Guiyang China, 16–18 March 2018; pp. 88–93.
- [3] Khatri, A.; Jain, R.; Vashista, H.; Mittal, N.; Ranjan, P.; Janardhanan, R. (2020) *Pneumonia Identification in Chest X-ray Images Using EMD*. In Internet of Things—Applications and Future; Springer Science and Business Media LLC: Singapore, 2020; pp. 87–98.
- [4] E. Donnall Thomas, M.D, (1990) *the winner of the 1990 Nobel Prize in Medicine, achieved what every expert said was impossible*. <https://www.mayoclinic.org/diseases-conditions/pneumonia/symptoms-causes/syc-20354204>
- [5] Goldman L, et al., eds. (2016) *Overview of pneumonia*. In: *Goldman-Cecil Medicine*.

- 25th ed. Philadelphia, Pa.: Saunders Elsevier; 2016. <http://www.clinicalkey.com>. Accessed April 18, 2016.
- [6] Wang, G., Liu, X., Shen, J. et al. (2021) *A deep-learning pipeline for the diagnosis and discrimination of viral, non-viral and COVID-19 pneumonia from chest X-ray images*. Nat Biomed Eng 5, 509–521 . <https://doi.org/10.1038/s41551-021-00704-1>.
- [7] Stephen M Humphries , Devlin Thieke , David Baraghoshi , Matthew J. Strand , Jeffrey J Swigris , Kum Ju Chae , Hye Jeon Hwang , Andrea S. Oh , Kevin R. Flaherty , Ayodeji Adegunsoye , Renea Jablonski (2023) *Deep Learning Classification of Usual Interstitial Pneumonia Predicts Outcomes*. , <https://doi.org/10.1164/rccm.202307-1191OC>.
- [8] Karen Simonyan, Andrew Zisserman (2014) *Very Deep Convolutional Networks for Large-Scale Image Recognition*, arXiv: 1409. 1556
- [9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun (2016) *Identity Mappings in Deep Residual Networks*. arXiv:1603.05027v3
- [10] Meysam Vakili, Mohammad Ghamsari and Masoumeh Rezaei (2020). *Performance Analysis and Comparison Of Machine and Deep Learning Algorithms for IoT Data Classification*. arXiv:2001.09636 or arXiv:2001.09636v1.
- [11] Manolis Loukidakis, Jos'e Cano, Michael O'Boyle, (2018) "*Accelerating Deep Neural Networks on Low Power Heterogeneous Architectures*", Institute for Computing Systems Architecture School of Informatics, University of Edinburgh, UK, January 2018
- [12] Emad Ul Haq Qazi *, Tanveer Zia and Abdulrazaq Almorjan, (2021) *Deep Learning-Based Digital Image Forgery Detection System*, <https://doi.org/10.3390/app12062851>