

ĐẠI HỌC BÁCH KHOA HÀ NỘI  
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



BÁO CÁO BÀI TẬP LỚN  
NHẬP MÔN TRÍ TUỆ NHÂN TẠO

Ứng dụng trí tuệ nhân tạo phát hiện một số bệnh hô hấp ở  
người qua ảnh chụp X-quang phổi

Hồ Việt Đức Lương	20200376
Nguyễn Duy Long	20204580
Ngô Trần Anh Thư	20200623
Phạm Xuân Trường	20200655

Giáo viên hướng dẫn:

PGS. TS. Lê Thanh Hương

HÀ NỘI  
Ngày 29 tháng 6 năm 2023

# Mục lục

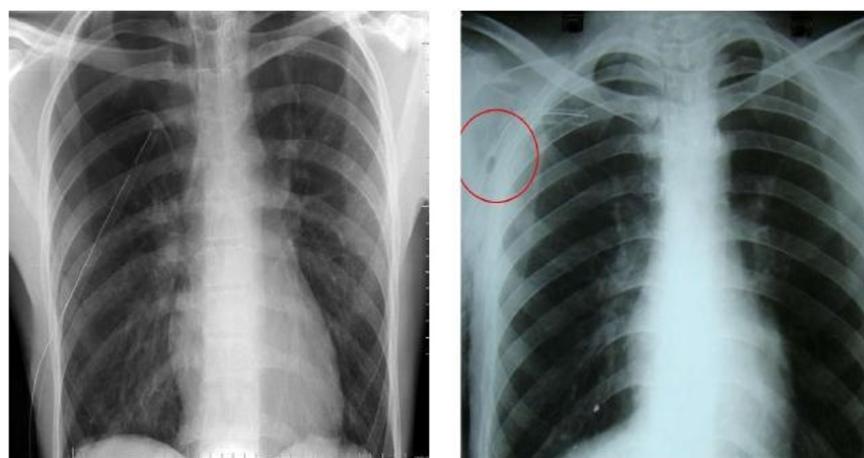
<b>1</b>	<b>Giới thiệu</b>	<b>3</b>
<b>2</b>	<b>Dữ liệu</b>	<b>6</b>
2.1	Dữ liệu sử dụng . . . . .	6
2.2	Tiền xử lý dữ liệu . . . . .	7
2.3	Làm giàu dữ liệu . . . . .	7
2.4	Xử lý mất cân bằng dữ liệu . . . . .	8
2.4.1	Đánh trọng số cho các lớp . . . . .	8
2.4.2	Sử dụng Generative Adversarial Networks - GAN . . . . .	9
<b>3</b>	<b>Các mô hình học máy sử dụng</b>	<b>11</b>
3.1	Mạng MultiLayer Perceptron - MLP . . . . .	11
3.2	Convolutional neuron Network - CNN . . . . .	13
3.3	DenseNet . . . . .	14
<b>4</b>	<b>Thực nghiệm và đánh giá</b>	<b>17</b>
4.1	Cross-Entropy Loss . . . . .	17
4.2	Độ đo sử dụng . . . . .	18
4.3	Thực nghiệm . . . . .	19
4.3.1	Môi trường và tham số . . . . .	19
4.3.2	Kết quả . . . . .	19
<b>5</b>	<b>Phần mềm</b>	<b>21</b>
5.1	Các chức năng của hệ thống . . . . .	21
5.1.1	Tác nhân sử dụng . . . . .	21
5.1.2	Các chức năng . . . . .	21
5.2	Giao diện phần mềm . . . . .	22
<b>6</b>	<b>Tổng kết</b>	<b>25</b>
	<b>Tài liệu tham khảo</b>	<b>28</b>

# Chương 1

## Giới thiệu

Bệnh viêm phổi (Pneumonia) là một trong những bệnh lý phổ biến nhất ở các bệnh nhân mắc bệnh về hô hấp [1], và là nguyên nhân gây tử vong hàng đầu ở trẻ em trên toàn thế giới (chiếm 14% tổng số trẻ em dưới 5 tuổi tử vong, trong đó có đến 20 ~ 45% ca viêm phổi gây ra là do phế cầu khuẩn).

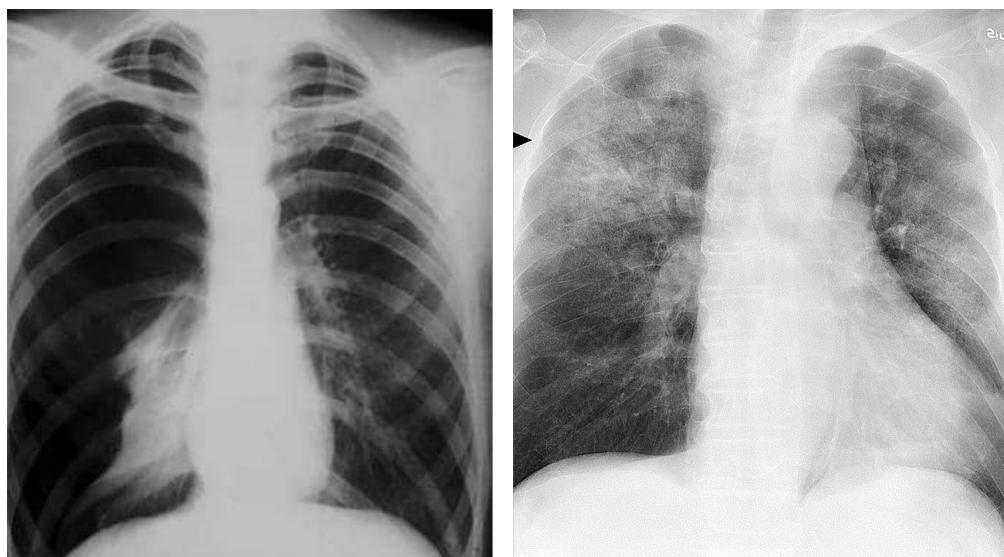
Viêm phổi là tình trạng nhiễm trùng xảy ra ở một hoặc cả hai bên phổi, do vi khuẩn, virus hoặc nấm gây ra. Viêm phổi có thể do nhiều nguyên nhân khác nhau, bao gồm nhiễm trùng virus, vi khuẩn hoặc nấm, hít vào các chất độc hại hoặc thuốc lá, hoặc bị phát triển từ các bệnh lý khác như bệnh tiểu đường hoặc bệnh tim. Việc phát hiện sớm và điều trị kịp thời là rất quan trọng để giảm thiểu các biến chứng của bệnh, nhưng cần phải được thực hiện dưới sự giám sát của các chuyên gia y tế. Nếu không được điều trị đúng cách, bệnh viêm phổi có thể dẫn đến các biến chứng nghiêm trọng và đe dọa tính mạng của bệnh nhân.



Hình 1.1: Hình ảnh X-quang phổi người bị nhiễm bệnh viêm phổi

Covid-19 là một bệnh truyền nhiễm do virus corona mới (SARS-CoV-2) gây ra, được phát hiện đầu tiên tại thành phố Vũ Hán, tỉnh Hồ Bắc, Trung Quốc vào cuối năm 2019 và đã lan rộng ra khắp thế giới trong thời gian ngắn sau đó [2]. Hậu Covid-19 ảnh hưởng đến sức khỏe tâm thần và thể chất của bệnh nhân. Trong hầu hết các trường hợp từ nhẹ đến nặng, bệnh nhân thường gặp các triệu chứng liên quan đến sự hoạt động của phổi như suy hô hấp, xơ phổi, viêm phổi, v.v. Các tổn thương phổi hậu Covid có thể chẩn đoán thông qua các phương pháp như chụp X-quang tim phổi.

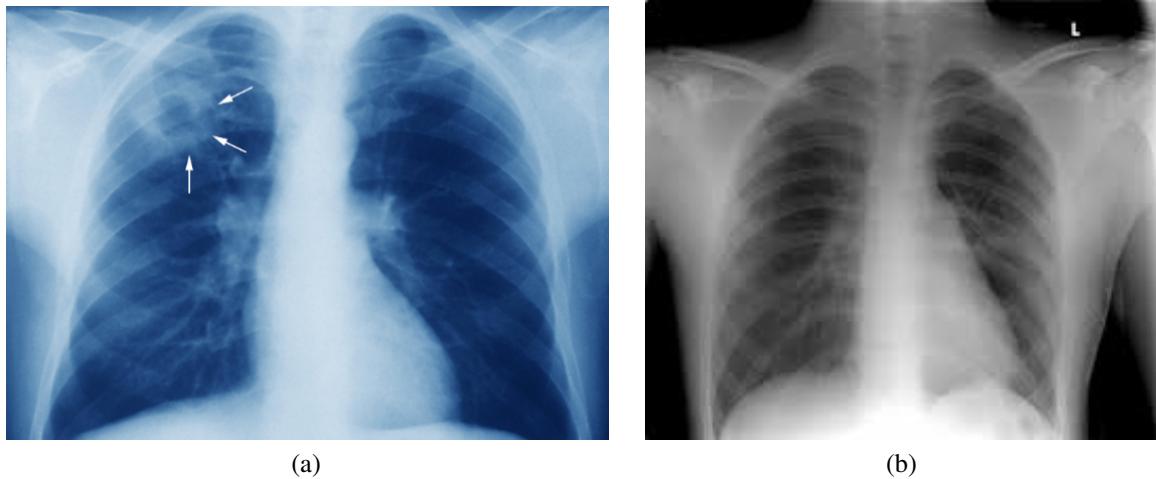
Các loại tổn thương phổi do Covid-19 có thể thấy trên hình ảnh X-quang là: Nốt mờ, đám mờ, kính mờ hình tròn đa ổ; Đông đặc nhu mô phổi vùng ngoại vi đa ổ; Mờ tổ chức kẽ quanh rốn phổi; Dày thành phế quản, dày vách liên tiểu thùy; Xẹp phổi; v.v.



Hình 1.2: Hình ảnh phổi bị tổn thương do tác động của Covid

Bệnh Tuberculosis, còn được gọi là bệnh lao, là một bệnh truyền nhiễm gây ra bởi vi khuẩn Mycobacterium Tuberculosis [3]. Bệnh này thường tác động chủ yếu đến phổi, nhưng cũng có thể ảnh hưởng đến các cơ quan khác trong cơ thể như não, xương, thận, và gan. Bên cạnh đó, Tuberculosis được coi là một trong những bệnh truyền nhiễm tử vong hàng đầu trên thế giới.

Bệnh lao được lây truyền chủ yếu qua hơi hoặc hạt nhỏ chứa vi khuẩn Mycobacterium Tuberculosis từ một người bị bệnh lao hoạt động và bắn vào không khí qua ho hoặc hắt hơi. Các yếu tố gia tăng nguy cơ bị nhiễm khuẩn bao gồm tiếp xúc gần gũi với người mắc bệnh lao, hệ miễn dịch suy yếu, sống trong môi trường không hợp lý, và sử dụng thuốc lá. Triệu chứng phổ biến của bệnh lao bao gồm ho kéo dài, sốt, mệt mỏi, giảm cân, và đau ngực. Tuy nhiên, triệu chứng có thể khác nhau tùy thuộc vào các cơ quan bị ảnh hưởng.



Hình 1.3: Hình ảnh X-quang phổi người bị bệnh lao (a) và tràn khí màng phổi (b)

Tràn khí màng phổi (Pneumothorax) là bệnh trong đó có khí tụ trong khoang màng phổi [4]. Bệnh này xảy ra khi có đường thông từ phổi ra bên ngoài, cho phép khí di chuyển từ phổi vào khoang màng phổi. Điều này làm cho phổi bị xẹp. Các triệu chứng của bệnh tràn khí màng phổi gồm: đau ngực dữ dội, đặc biệt là khi hít thở sâu, khó thở, thở nhanh và nông, ho khan không ra đờm, đau tăng khi ho hoặc hít thở sâu, tim đập nhanh, v.v

Để chẩn đoán bệnh viêm phổi, lao phổi, tràn khí màng phổi hay các ca Covid-19 trở nặng, các bác sĩ thường sử dụng các phương pháp như xét nghiệm da, xét nghiệm nhuỵ hoặc xét nghiệm hình ảnh như chụp X-quang phổi hoặc CT scanner.

Những hình ảnh X-quang trên cho thấy một tiềm năng của việc áp dụng các kỹ thuật học máy để tự động chẩn đoán 1 số bệnh đường hô hấp ở người thông qua ảnh chụp X-quang phổi. Nhóm đề xuất đề tài: “*Ứng dụng trí tuệ nhân tạo phát hiện một số bệnh hô hấp ở người qua ảnh chụp X-quang phổi*”.

**Đề tài:** Ứng dụng trí tuệ nhân tạo phát hiện một số bệnh hô hấp ở người qua ảnh chụp X-quang phổi

**Môn học:** Nhập môn Trí Tuệ Nhân Tạo

**Phát hành:** 30/6/2023

# Chương 2

## Dữ liệu

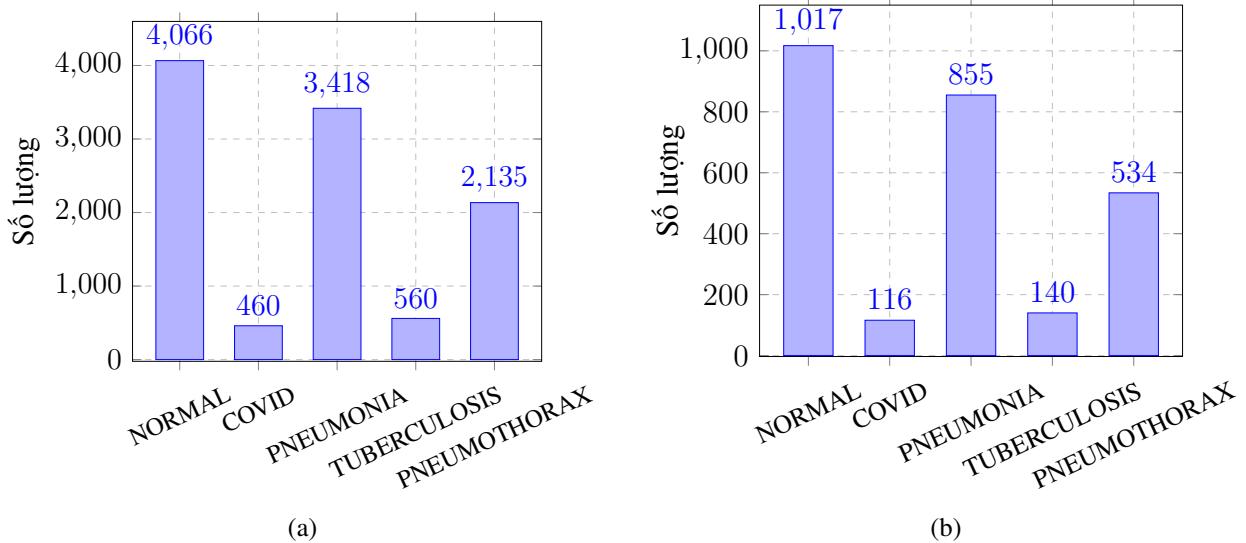
### 2.1 Dữ liệu sử dụng

Trong lĩnh vực học máy, việc thu thập dữ liệu là một bước quan trọng để xây dựng mô hình dự đoán và huấn luyện hệ thống thông minh. Việc thu thập dữ liệu từ các nguồn thường sẵn có là một phần quan trọng trong bài toán chẩn đoán các căn bệnh liên quan đến phổi. Có nhiều nguồn dữ liệu được sử dụng để nắm bắt thông tin về bệnh lý phổi, từ đó cung cấp cơ sở dữ liệu đáng tin cậy cho việc xây dựng mô hình dự đoán và nghiên cứu.

Trong báo cáo này, các bộ dữ liệu ảnh chụp X-quang lớn được thu thập từ nhiều nguồn khác nhau để sử dụng trong quá trình huấn luyện. Các bộ dữ liệu này chứa hàng nghìn hình ảnh phổi từ các bệnh viện, phòng khám và các tổ chức y tế khác. Cụ thể, các nguồn dữ liệu được thu thập từ các nguồn:

- Bệnh nhân viêm phổi: [5] và [6]
- Bệnh nhân COVID-19 trở nặng: [6]
- Bệnh nhân lao phổi: [7]
- Bệnh nhân tràn khí màng phổi: [8]

Tập dữ liệu được chia thành tập huấn luyện và tập kiểm tra (train/test set). Mỗi folder dữ liệu có các subfolder cho mỗi nhãn phân loại của ảnh (gồm NORMAL-bình thường, COVID, PNEUMONIA - viêm phổi, TUBERCOLOSIS - lao phổi, PNEUMOTHORAX - tràn khí màng phổi). Hình 2.1 phân tích về thống kê của số lượng các loại ảnh thu thập được.



Hình 2.1: Số lượng các ảnh của các nhãn  
(a) Trong tập huấn luyện (b) Trong tập test.

## 2.2 Tiền xử lý dữ liệu

Tiền xử lý dữ liệu ảnh X-quang phổi là một bước quan trọng trong quá trình xây dựng mô hình chẩn đoán các bệnh dựa trên hình ảnh X-quang phổi. Dưới đây là một số bước tiền xử lý đã được áp dụng trong báo cáo:

- Điều chỉnh kích thước: Các hình ảnh X-quang phổi có các kích thước khác nhau do được thu thập từ nhiều nguồn. Để đảm bảo sự nhất quán, các hình ảnh được chuẩn hóa về kích thước  $128 \times 128$ .
- Chuẩn hóa hình ảnh: Các mô hình sử dụng ảnh đầu vào gồm 3 chiều (dài, rộng, kênh). Đồng thời, ảnh được chuẩn hóa trong khoảng từ 0 đến 1. Quá trình này giúp giảm kích thước dữ liệu và tăng tốc độ huấn luyện.

## 2.3 Làm giàu dữ liệu

Các kỹ thuật làm giàu dữ liệu được sử dụng:

- Zoom Range (Phóng to/thu nhỏ ảnh): Các hình ảnh có thể được phóng to lên đến 10% hoặc thu nhỏ xuống 10% so với kích thước ban đầu. Kỹ thuật này hữu ích khi các ảnh chụp phổi có vị trí đa dạng, giúp mô hình học được các đặc trưng quan trọng ở các tỷ lệ khác nhau của ảnh.
- Width Shift Range (Dịch ảnh theo chiều rộng): Các hình ảnh có thể được dịch sang trái hoặc sang phải tối đa 10% của chiều rộng ban đầu. Kỹ thuật

này giúp tạo ra những biến thể của ảnh X-quang phổi với vị trí dịch theo chiều rộng khác nhau, đa dạng thêm nguồn dữ liệu.

- Height Shift Range (Dịch ảnh theo chiều cao): Các hình ảnh có thể được dịch lên hoặc xuống tối đa 10% của chiều cao ban đầu. Kỹ thuật này giúp tạo ra những biến thể của ảnh X-quang phổi với vị trí dịch theo chiều cao khác nhau, đa dạng thêm nguồn dữ liệu.
- Bright Range (Thay đổi độ sáng của ảnh): Khi áp dụng phép biến đổi độ sáng lên ảnh, giá trị độ sáng của mỗi pixel trong ảnh sẽ được điều chỉnh theo một giá trị ngẫu nhiên nằm trong khoảng được xác định được biểu diễn dưới dạng một cặp ( $\min_{value}, \max_{value}$ ). Điều này giúp tạo ra sự đa dạng trong độ sáng của ảnh và làm cho mô hình học máy có khả năng nhận diện và phân loại ảnh tốt hơn trong các trường hợp độ sáng khác nhau.
- Shear Range (Nghiêng ảnh): Khi áp dụng phép biến đổi nghiêng lên ảnh, các pixel trong ảnh sẽ được dịch chuyển theo một góc nghiêng ngẫu nhiên nằm trong khoảng được xác định. Điều này giúp tạo ra sự biến đổi hình dạng của ảnh và làm cho mô hình học máy có khả năng nhận diện và phân loại ảnh tốt hơn trong các trường hợp hình dạng khác nhau.

## 2.4 Xử lý mất cân bằng dữ liệu

Mất cân bằng dữ liệu là tình trạng khi số lượng các lớp hoặc nhãn trong tập dữ liệu không đồng đều. Trong dữ liệu ảnh X-quang, mất cân bằng dữ liệu cũng là một vấn đề phổ biến, bởi số lượng bệnh nhân chiếm một số lượng nhỏ trên toàn bộ và điều này có thể ảnh hưởng đáng kể đến hiệu suất và độ chính xác của mô hình.

Trong ảnh y tế, mất cân bằng dữ liệu có thể xảy ra khi số lượng mẫu từ mỗi lớp hoặc bệnh lý khác nhau trong tập dữ liệu không đồng đều. Điều này có thể dẫn đến hiện tượng mô hình bị thiên vị hướng tới các lớp đa số và không đạt được hiệu suất tốt trên các lớp thiểu số. Ví dụ, trong tập dữ liệu ảnh X-quang, ảnh Covid19 chỉ chiếm 5% tổng số ảnh.

Để giải quyết vấn đề mất cân bằng dữ liệu trong bài toán này, một số phương pháp đã được sử dụng như: đánh trọng số cho các lớp và sử dụng mạng sinh ảnh GAN để cân bằng các lớp dữ liệu.

### 2.4.1 Đánh trọng số cho các lớp

Phương pháp đầu tiên được sử dụng là đánh trọng số cho các lớp. Đây là một phương pháp được sử dụng để giải quyết vấn đề mất cân bằng dữ liệu trong bài

toán phân loại. Trong trường hợp mất cân bằng, số lượng mẫu thuộc vào các lớp khác nhau không đồng đều, dẫn đến mô hình có thể thiên vị hướng tới lớp đa số.

Phương pháp này giúp cân bằng ảnh hưởng của các lớp trong quá trình huấn luyện bằng cách gán trọng số khác nhau cho mỗi lớp. Trọng số này được áp dụng vào hàm mất mát trong quá trình tính toán gradient, nhằm tăng cường sự quan trọng của các mẫu thuộc vào lớp thiểu số. Bằng cách làm điều này, mô hình được khuyến khích học tập đặc trưng của lớp thiểu số một cách hiệu quả hơn, từ đó cải thiện khả năng dự đoán cho lớp này.

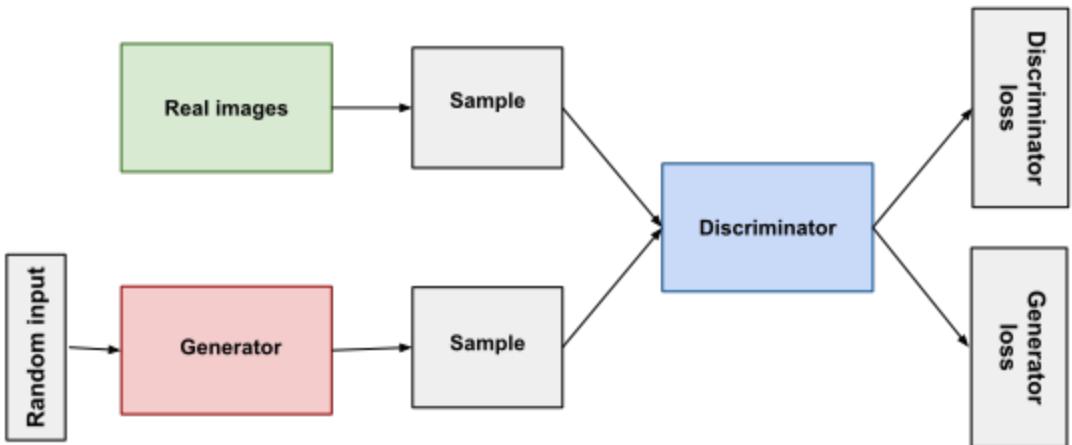
Công thức tính trọng số lớp thường dựa trên tỷ lệ số lượng mẫu của các lớp. Lớp thiểu số sẽ có trọng số cao hơn, trong khi lớp đa số sẽ có trọng số thấp hơn. Việc lựa chọn phương pháp tính trọng số phụ thuộc vào đặc thù của bài toán và mục tiêu cần đạt được.

## 2.4.2 Sử dụng Generative Adversarial Networks - GAN

Ngoài ra, một phương pháp mạnh mẽ khác là sử dụng mạng GAN. Mạng GAN có hai thành phần chính là mô hình sinh (Generator) và mô hình phân biệt (Discriminator). Trên cơ sở dữ liệu ban đầu, mô hình sinh sẽ tạo ra những ảnh X-quang giả. Qua quá trình huấn luyện, mô hình phân biệt phải phân biệt được ảnh X-quang thật và ảnh X-quang giả. Kết quả là, mô hình sinh sẽ học được cách tạo ra những ảnh X-quang mới, tương tự với dữ liệu gốc. Nhờ vậy, ta có thể tạo ra thêm dữ liệu cho lớp thiểu số và cân bằng dữ liệu.

Mạng sinh nhận đầu vào là một vector ngẫu nhiên được gọi là vector tiềm ẩn hoặc vector nhiễu. Qua quá trình lan truyền thuận, mạng sinh biến đổi vector nhiễu thành một đầu ra dữ liệu giả mạo, thường là một hình ảnh, âm thanh, hoặc dữ liệu khác tương ứng với lĩnh vực cụ thể. Trong quá trình huấn luyện, mạng sinh cố gắng cải thiện khả năng tạo ra dữ liệu giả mạo sao cho nó trông giống với dữ liệu thật. Điều này đòi hỏi mạng sinh phải học cách biến đổi vector nhiễu thành các đặc trưng phù hợp và tạo ra dữ liệu có tính chất tương tự với tập dữ liệu huấn luyện. Mạng sinh trong GAN thường có cấu trúc phức tạp, gồm nhiều lớp nơ-ron và sử dụng các kỹ thuật như Transpose Convolution, Pooling, ReLU để tạo ra dữ liệu giả mạo chất lượng cao.

Mạng phân biệt nhận đầu vào là một mẫu dữ liệu, có thể là hình ảnh, âm thanh hoặc dữ liệu khác. Qua quá trình lan truyền thuận, mạng phân biệt đưa ra dự đoán về tính chất của dữ liệu, xác định liệu đó có phải là dữ liệu thật hay dữ liệu giả mạo. Trong quá trình huấn luyện, mạng phân biệt cố gắng học cách phân biệt và đánh giá chính xác sự khác biệt giữa các mẫu dữ liệu thật và giả mạo. Nó cần học những đặc trưng quan trọng để phân loại dữ liệu một cách chính xác và phân biệt dữ liệu giả mạo tạo ra bởi mạng sinh.



Hình 2.2: Mô hình tổng quan hệ thống mạng GAN

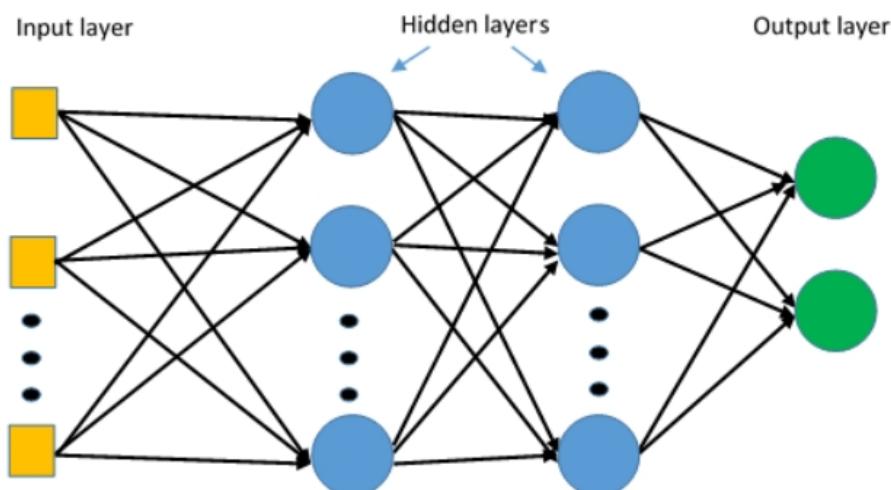
Mạng GAN có ưu điểm là mô hình sinh đơn giản và hiệu quả, có thể dễ dàng sử dụng. Tuy nhiên, việc sử dụng mạng GAN cũng gặp một số hạn chế như mô hình khó đào tạo và khó hội tụ, các ảnh sinh thường giống nhau và dữ liệu ảnh thực để đào tạo chưa đủ nhiều và đa dạng để cải thiện hiệu suất mô hình.

# Chương 3

## Các mô hình học máy sử dụng

### 3.1 Mạng MultiLayer Perceptron - MLP

Mạng MLP là một kiểu mạng neuron nhân tạo (artificial neural network) phổ biến trong lĩnh vực học máy [9]. Mạng MLP được sử dụng rộng rãi trong các bài toán phân loại và hồi quy.

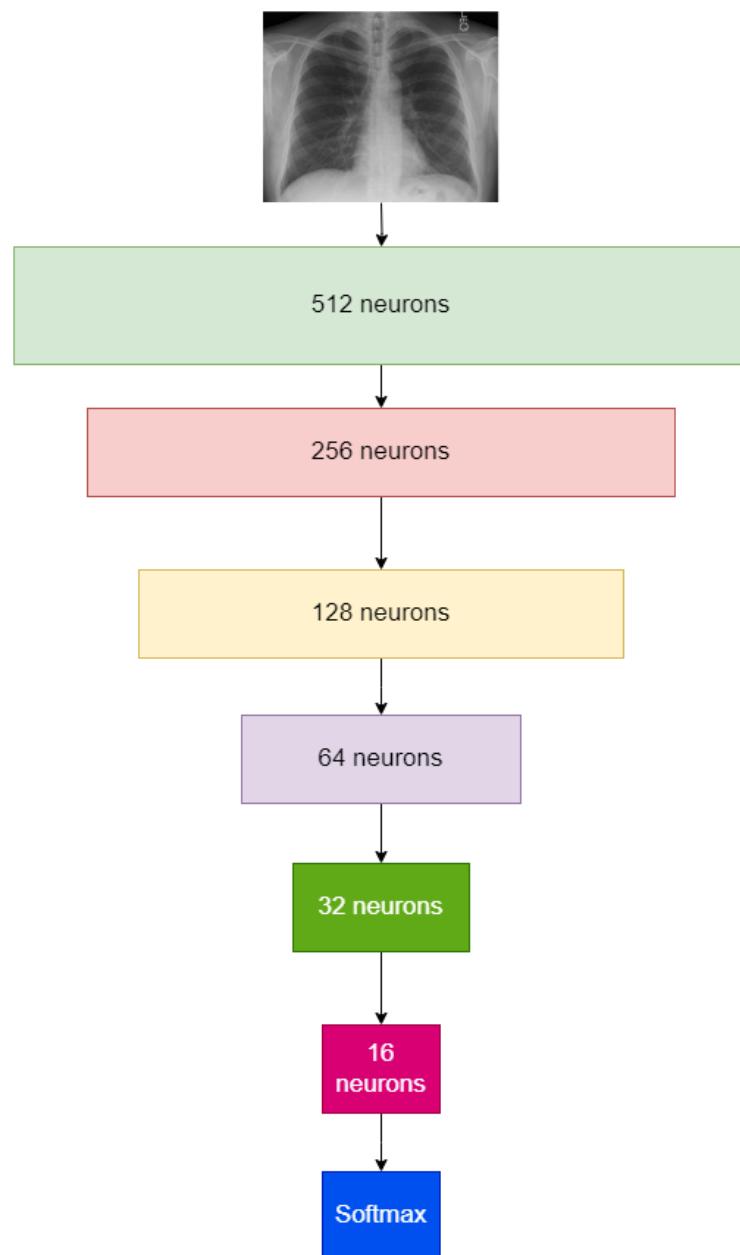


Hình 3.1: Mạng MLP

Mạng MLP bao gồm nhiều lớp neuron ẩn được xếp chồng lên nhau. Mỗi lớp neuron chứa một số neuron kết nối với các neuron trong lớp trước và sau nó. Các neuron trong cùng một lớp là độc lập với nhau và hoạt động song song. Mỗi neuron trong mạng MLP thực hiện hai bước tính toán chính: tổng hợp (sum) và kích hoạt (activation). Trong bước tổng hợp, neuron tính tổng trọng số của các kết nối đầu vào và áp dụng một hàm tổng hợp (ví dụ: tổng có trọng số hoặc tích vô hướng). Sau đó, kết quả được đưa qua một hàm kích hoạt phi tuyến (ví dụ: hàm sigmoid,

hàm ReLU) để tạo ra đầu ra của neuron.

Quá trình huấn luyện mạng MLP thường được thực hiện bằng cách sử dụng thuật toán lan truyền ngược (backpropagation). Thuật toán này dựa trên nguyên lý cập nhật các trọng số của mạng sao cho sai số giữa đầu ra thực tế và đầu ra dự đoán được là nhỏ nhất. Mạng MLP có khả năng học các mô hình phức tạp và khá linh hoạt trong việc xử lý các loại dữ liệu khác nhau. Tuy nhiên, mạng MLP cũng có một số hạn chế, như khả năng bị overfitting nếu không được điều chỉnh tốt và khó khăn trong việc xử lý các mẫu dữ liệu có cấu trúc không rõ ràng.



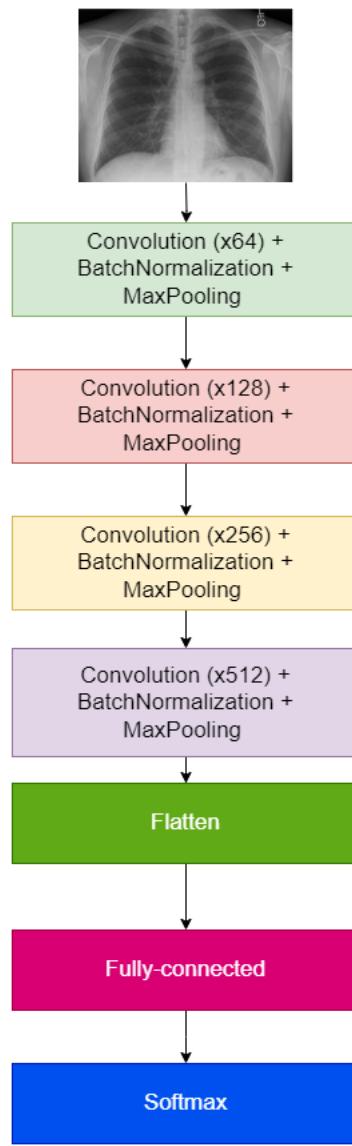
Hình 3.2: Mạng MLP dùng trong bài tập lớn

Trong bài tập lớn này, nhóm sử dụng mạng MLP với 6 hidden layer với số neuron lần lượt là 512, 256, 128, 64, 32, 16 với hàm kích hoạt ReLU, theo sau là

lớp Dropout. Ở layer cuối cùng hàm kích hoạt sẽ là softmax, từ đó đưa ra xác suất thuộc về các lớp của dự đoán.

## 3.2 Convolutional neuron Network - CNN

Mạng CNN là một loại mạng neuron nhân tạo được sử dụng rộng rãi trong lĩnh vực thị giác máy tính và xử lý ảnh [10]. Nó được thiết kế đặc biệt để nhận dạng và trích xuất đặc trưng từ dữ liệu không gian như hình ảnh.



Hình 3.3: Mạng CNN sử dụng trong project

Mạng CNN sử dụng các lớp tích chập (convolutional layers) để tự động học các bộ lọc (filters) nhằm phát hiện các đặc trưng trong ảnh. Các lớp tích chập thực hiện việc trượt các bộ lọc qua ảnh đầu vào và tính toán tổng (convolution) giữa bộ lọc và các vùng tương ứng trong ảnh. Quá trình này giúp mạng học các đặc trưng

cục bộ như cạnh, góc, và các hình dạng đơn giản. Sau đó, các lớp kích hoạt phi tuyến (activation layers) như ReLU được áp dụng để đưa ra đầu ra phi tuyến tính và tăng tính phi tuyến của mạng. Các lớp gộp (pooling layers) được sử dụng để giảm kích thước không gian và số lượng thông tin cần xử lý bằng cách chọn giá trị đại diện từ các vùng chồng lấn trong ảnh. Cuối cùng, các lớp kết nối đầy đủ (fully connected layers) được sử dụng để kết hợp các đặc trưng đã được trích xuất và thực hiện các bước phân loại hoặc dự đoán. Mạng CNN thường kết thúc bằng một lớp softmax để đưa ra xác suất phân loại cho các lớp đầu ra.

Mạng CNN có khả năng học các đặc trưng phức tạp, đồng thời giảm số lượng tham số cần tối ưu so với mạng MLP truyền thống. Điều này làm cho nó phù hợp cho việc xử lý ảnh và nhận dạng đối tượng trong nhiều bài toán. Mạng CNN đã được áp dụng thành công trong nhiều lĩnh vực, bao gồm nhận dạng khuôn mặt, phân loại ảnh, xe tự hành và nhiều ứng dụng khác trong thị giác máy tính. Nó đã đạt được những thành tựu đáng kể và trở thành một công nghệ quan trọng trong lĩnh vực học máy và trí tuệ nhân tạo.

Trong bài tập lớn này nhóm sử dụng 5 Convolution blocks, mỗi block gồm 1 lớp Convolution, 1 lớp Batch normalization và 1 lớp Maxpooling. Cuối cùng, hình ảnh sẽ được đưa qua lớp Flatten và lớp fully connected để đưa ra kết quả dự đoán cho từng lớp bằng hàm softmax.

### 3.3 DenseNet

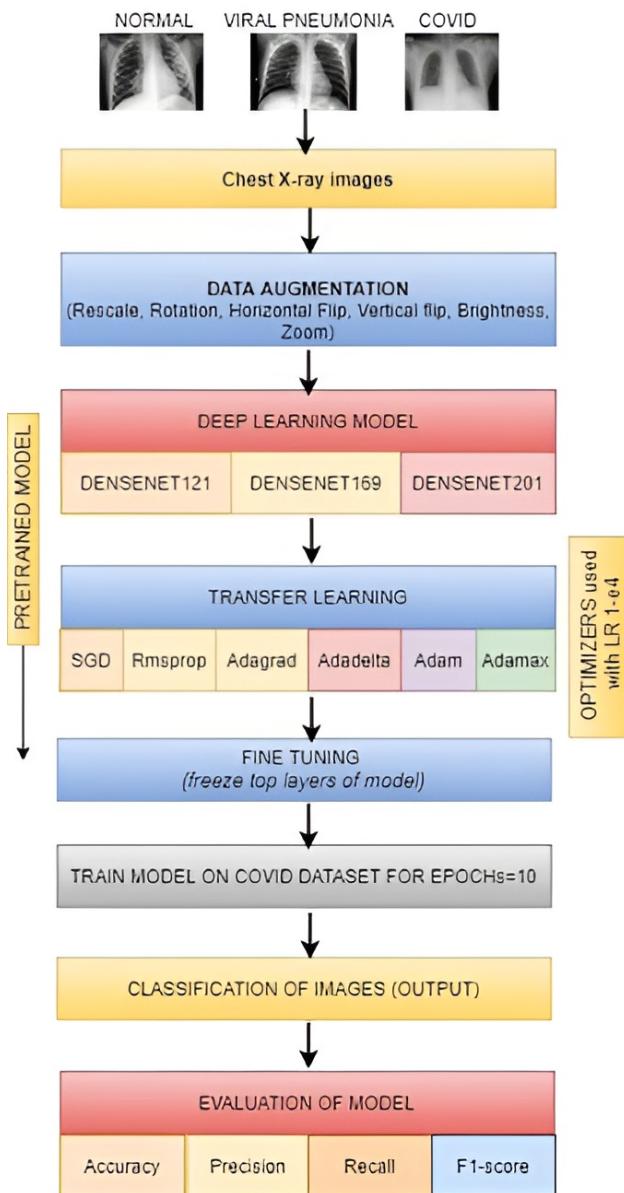
Mạng Densenet (Dense Convolutional Network) là một kiến trúc mạng neuron sâu được giới thiệu bởi Gao Huang et. al. [11]. Mạng Densenet có mục tiêu chính là giải quyết vấn đề mất mát thông tin và sự suy giảm độ chính xác khi mạng mở rộng và sâu hơn.

Trong các mạng neuron truyền thẳng truyền thống như VGG, ResNet, các đặc trưng từ các tầng trước được truyền tiếp qua các tầng sau thông qua việc nối các tầng với nhau. Tuy nhiên, việc truyền thông tin qua các tầng như vậy có thể gây ra mất mát thông tin và làm giảm hiệu suất của mạng. Để giải quyết vấn đề này, mạng Densenet sử dụng một kiến trúc gọi là "kết nối dày đặc" (dense connectivity), trong đó các tầng sau được kết nối với tất cả các tầng trước bằng cách nối đầu ra của các tầng trước vào đầu vào của các tầng sau. Điều này đảm bảo rằng thông tin từ tất cả các tầng trước đó được truyền tiếp và không bị mất đi.

Mạng Densenet được xây dựng dựa trên khối cơ bản gọi là "bottleneck layer", trong đó mỗi khối bao gồm các tầng Convolution - Batch Normalization - ReLU. Tuy nhiên, điểm đặc biệt của Densenet là sự kết hợp của các khối này để tạo thành một mạng sâu. Các khối bottleneck layer được kết nối với nhau thông qua kết nối

dày đặc, tạo ra các đường đi ngắn giữa các tầng và cho phép dữ liệu và gradient được truyền qua mạng một cách hiệu quả.

Mạng Densenet có nhiều lợi ích. Đầu tiên, việc kết nối dày đặc giúp mạng học được các đặc trưng thấp và cao cùng một lúc, từ đó cải thiện khả năng học của mạng. Thứ hai, việc truyền thông tin qua các kết nối ngắn giữa các tầng giúp giảm thiểu sự mất mát thông tin và tăng tính thông tin của mạng. Thứ ba, mạng Densenet có khả năng tái sử dụng đặc trưng tốt hơn, cho phép nó học tốt hơn từ số lượng nhỏ hơn các tham số.



Hình 3.4: Mô hình hệ thống mạng DenseNet

Mạng Densenet đã chứng tỏ hiệu suất tốt trên nhiều tập dữ liệu và bài toán khác nhau, bao gồm phân loại ảnh, nhận diện vật thể, và phân đoạn ảnh. Nó đã trở

thành một trong những kiến trúc quan trọng và được ứng dụng rộng rãi trong lĩnh vực thị giác máy tính. Với tính kết nối mật độ và khả năng học tốt, mạng Densenet đóng góp đáng kể vào sự phát triển của lĩnh vực học sâu và đạt được kết quả ấn tượng trong việc giải quyết các bài toán phức tạp của thị giác máy tính.

# Chương 4

## Thực nghiệm và đánh giá

### 4.1 Cross-Entropy Loss

Cross-entropy loss (hàm mất mát cross entropy) là một hàm mất mát phổ biến được sử dụng trong bài toán phân loại đa lớp [12]. Trong bài toán này, hàm Cross-entropy loss được sử dụng để phân loại cho 5 lớp (1 lớp bình thường và 4 lớp bệnh). Cross-entropy loss được tính bằng cách so sánh giữa xác suất dự đoán của mô hình và nhãn thực tế của mỗi mẫu dữ liệu. Cụ thể, ta sử dụng hàm softmax để chuyển đổi đầu ra của mô hình thành xác suất dự đoán cho từng lớp. Sau đó, ta tính toán mất mát cross entropy bằng cách lấy trung bình của các giá trị cross entropy cho tất cả các mẫu dữ liệu.

Công thức tính toán cross entropy loss cho một tập dữ liệu:

$$L = -\frac{1}{D} \sum_{i=1}^D \sum_{j=1}^5 y_{ij} \log(p_{ij}) \quad (4.1)$$

Trong đó:

- $D$  là số lượng mẫu dữ liệu.
- $y_{ij}$  là nhãn thực tế của mẫu thứ  $i$  của lớp  $j$ .
- $p_{ij}$  là xác suất dự đoán cho mẫu thứ  $i$  của lớp  $j$ .

Cross-entropy loss là một hàm mất mát phù hợp cho bài toán phân loại đa lớp và thường được sử dụng để đánh giá hiệu suất của mô hình trong quá trình huấn luyện và kiểm tra.

## 4.2 Độ đo sử dụng

Để đánh giá kết quả của các mô hình phân loại, báo cáo này sử dụng các độ đo chuẩn *Accuracy A*, *Precision P*, *Recall R* và *F1-score F1*.

Đặt  $N = 5$  là số lớp của các mẫu trong bộ dữ liệu. Với mỗi lớp  $c_i$ ,  $i \in \overline{1, N}$ , gọi:

- $TP_i$  là các mẫu mà mô hình dự đoán đúng là thuộc lớp  $c_i$ .
- $TN_i$  là các mẫu mà mô hình dự đoán đúng là không thuộc lớp  $c_i$ .
- $FP_i$  là các mẫu mà mô hình dự đoán là thuộc lớp  $c_u$  và bị sai.
- $FN_i$  là các mẫu mà mô hình dự đoán là không thuộc  $c_i$  và bị sai.

Khi đó, *Precision P<sub>i</sub>* và *Recall R<sub>i</sub>* của lớp  $c_i$  được tính như sau:

$$\mathbf{P}_i = \frac{TP_i}{TP_i + FP_i}, \quad \mathbf{R}_i = \frac{TP_i}{TP_i + FN_i} \quad (4.2)$$

Đối với tất cả các lớp:

- Độ chính xác *Accuracy A* là tỷ lệ số mẫu phân loại đúng với toàn bộ số mẫu của tập dữ liệu:

$$\mathbf{A} = \frac{\sum_{i=1}^N (TP_i + TN_i)}{\sum_{i=1}^N (TP_i + TN_i + FP_i + FN_i)} \quad (4.3)$$

- *Precision P*, *Recall R* có 2 cách tính vi mô (micro) và vĩ mô (macro):

$$\mathbf{P}_{\text{micro}} = \frac{\sum_{i=1}^N TP_i}{\sum_{i=1}^N (TP_i + FP_i)}, \quad \mathbf{R}_{\text{micro}} = \frac{\sum_{i=1}^N TP_i}{\sum_{i=1}^N (TP_i + FN_i)} \quad (4.4)$$

$$\mathbf{P}_{\text{macro}} = \frac{\sum_{i=1}^N \mathbf{P}_i}{N}, \quad \mathbf{R}_{\text{macro}} = \frac{\sum_{i=1}^N \mathbf{R}_i}{N} \quad (4.5)$$

Macro-average tính toán bằng cách lấy trung bình độ đo của các lớp, do đó độ đo này cho các lớp với trọng số bằng nhau. Micro-average tính toán bằng cách lấy tổng các lớp chia cho tổng phần tử, do đó coi các mẫu với trọng số bằng nhau.

Bộ dữ liệu sử dụng không cân bằng, nên nhóm đánh giá trên độ đo macro-average. Độ đo này cho các lớp có trọng số như nhau, kết quả không bị nghiêng về lớp đa số, giúp việc đánh giá mô hình khách quan hơn.

- *F1-score* F1 là trung bình điều hoà của *Precision* và *Recall*:

$$F1_{macro} = \frac{2P_{macro}R_{macro}}{P_{macro} + R_{macro}} \quad (4.6)$$

## 4.3 Thực nghiệm

### 4.3.1 Môi trường và tham số

Mô hình được cài đặt bằng framework Tensorflow trên máy ảo Kaggle và Colab với GPU Tesla V4 (16GB Ram). Dữ liệu đầu vào của các mô hình là (128, 128, 3). Mô hình được huấn luyện trong 30 epoch với kích thước mini-batch là 128. Các mô hình sử dụng hàm tối ưu Adam với tỷ lệ học ban đầu là 1e-4. Đầu ra của mô hình được đưa qua hàm softmax để tiến hành phân loại (5 lớp). Ngoài ra, mô hình được cài đặt lập lịch để giảm tốc độ học khi metric đánh giá ngừng cải thiện và sử dụng kỹ thuật early stopping khi mà giá trị metric không cải thiện được nữa. Việc lựa chọn siêu tham số huấn luyện mô hình được lấy dựa trên quá trình thực nghiệm.

### 4.3.2 Kết quả

Kết quả thực nghiệm và đánh giá mô hình trên bộ dữ liệu được thể hiện qua Bảng 4.1, Bảng 4.2 dưới đây:

Bảng 4.1: Kết quả so sánh các mô hình khi sử dụng đánh trọng số trên tập test

Mô hình	Accuracy	Precision	Recall	F1-score
MLP + Weight	0.56	0.56	0.57	0.56
CNN + Weight	0.85	0.85	0.85	0.85
<b>DenseNet121 + Weight</b>	<b>0.91</b>	<b>0.92</b>	<b>0.91</b>	<b>0.91</b>

Bảng 4.2: Kết quả so sánh các mô hình khi sử dụng kết hợp với GAN trên tập test

Mô hình	Accuracy	Precision	Recall	F1-score
MLP + GAN	0.57	0.57	0.58	0.58
CNN + GAN	0.87	0.88	0.87	0.87
<b>DenseNet121 + GAN</b>	<b>0.94</b>	<b>0.95</b>	<b>0.94</b>	<b>0.94</b>

Bảng kết quả cho thấy, việc kết hợp mô hình sinh ảnh GAN với DenseNet121 đem lại kết quả vượt trội so với việc sử dụng mô hình CNN và ANN, trong cả hai trường hợp sử dụng đánh trọng số và mô hình GAN. Mô hình DenseNet121

với khả năng trích xuất đặc trưng mạnh mẽ đã cho thấy hiệu quả của mình khi *F1-score* tăng đến 8% so với CNN và 38% so với mô hình ANN đơn giản.

Mô hình DenseNet121 là một mạng neural tích chập (CNN) với kiến trúc đặc biệt, trong đó mỗi tầng kết nối với tất cả các tầng sau nó. Điều này giúp tăng cường quá trình học và trích xuất đặc trưng từ dữ liệu. Mô hình này đã được huấn luyện trên một tập dữ liệu lớn và đạt được hiệu suất cao trong việc nhận dạng và phân loại ảnh. Việc sử dụng mô hình sinh ảnh GAN là một phương pháp tiếp cận độc đáo trong việc tạo ra các ảnh mới từ dữ liệu hiện có. Mạng sinh ảnh cố gắng tạo ra ảnh mới có cấu trúc tương tự như dữ liệu gốc, từ đó cải thiện hiệu suất mô hình.

Kết quả cho thấy rằng việc sử dụng mô hình sinh ảnh GAN kết hợp với DenseNet121 đã cải thiện hiệu suất mô hình tốt hơn. Khi áp dụng phương pháp này, *F1-score* đã được cải thiện đáng kể, đồng thời cải thiện cả độ đo *accuracy*, *precision*, *recall*.

# **Chương 5**

## **Phần mềm**

### **5.1 Các chức năng của hệ thống**

#### **5.1.1 Tác nhân sử dụng**

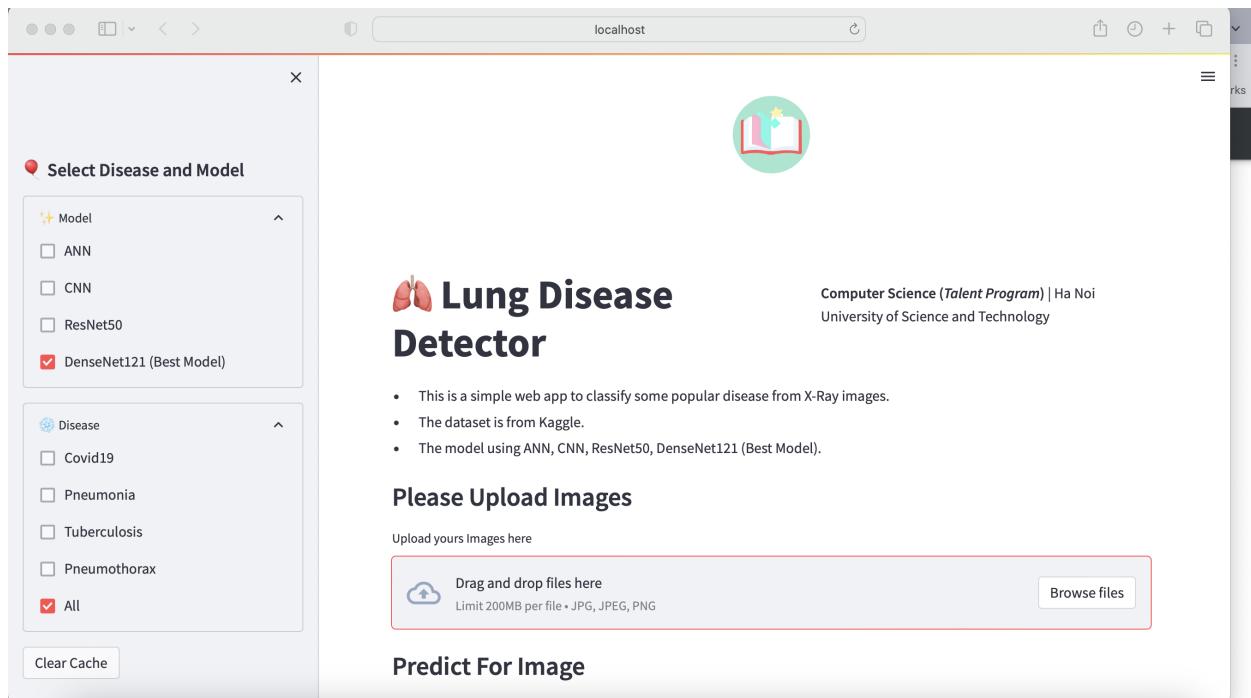
Tác nhân sử dụng là một người dùng bất kỳ có nhu cầu xét nghiệm các bệnh được đề cập, miễn là có ảnh chụp X-quang phổi.

#### **5.1.2 Các chức năng**

Hệ thống cung cấp các chức năng sau:

- Upload một hoặc nhiều hình ảnh chụp X-quang để xét nghiệm.
- Lựa chọn loại bệnh muốn xét nghiệm hoặc toàn bộ.
- Lựa chọn mô hình dự đoán hoặc mặc định là mô hình tốt nhất theo đánh giá của nhóm.
- In ra kết quả dự đoán đối với từng bệnh và từng ảnh chụp.

## 5.2 Giao diện phần mềm



Hình 5.1: Trang chủ

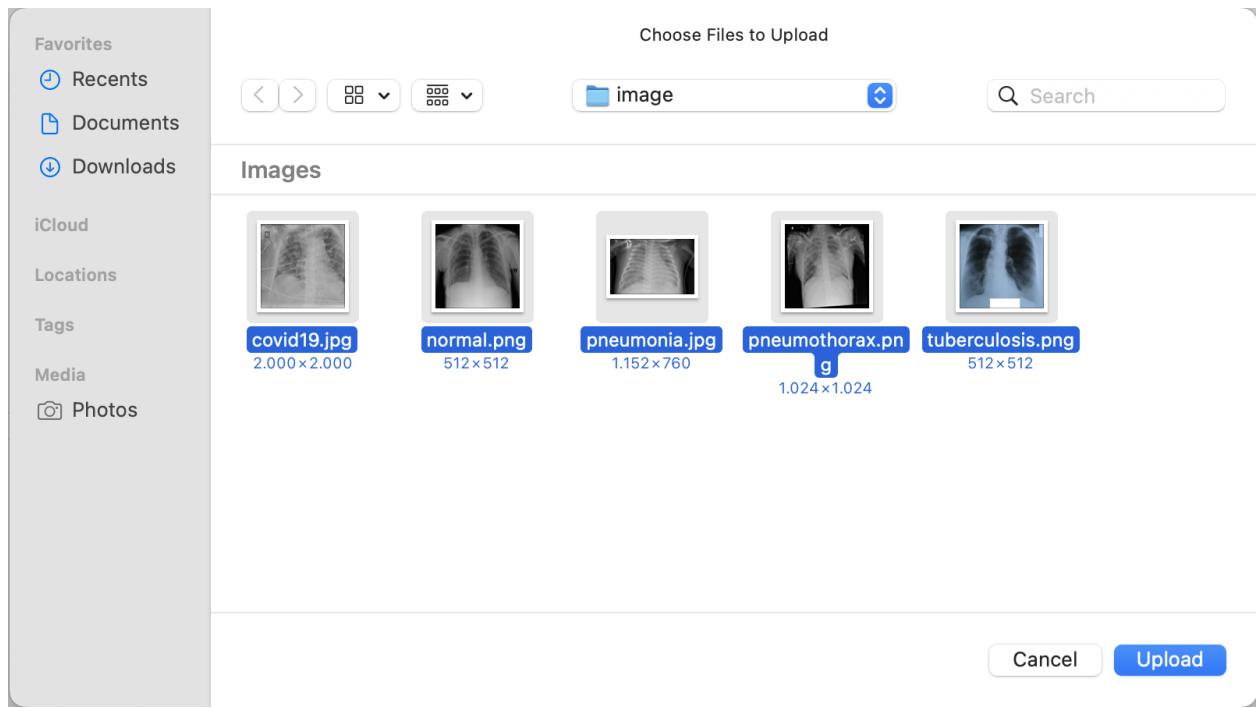
### Please Upload Images

Upload yours Images here

 Drag and drop files here  
Limit 200MB per file • JPG, JPEG, PNG

Browse files

Hình 5.2: Chức năng upload ảnh



Hình 5.3: Các ảnh được chọn để dự đoán

## Predict For Image

Predict For Image

Hình 5.4: Chức năng chẩn đoán cho các ảnh đã được chọn

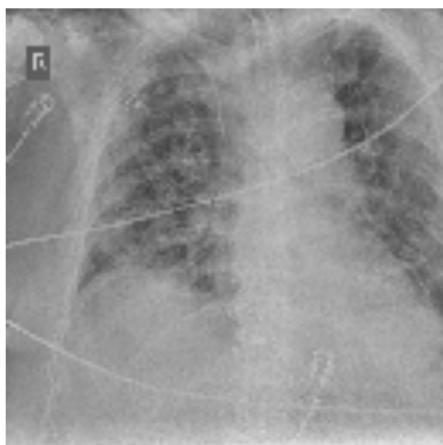
## Predict For Image

Predict For Image

NAME	NORMAL	COVID19	PNEUMONIA	TUBERCULOSIS	PNEUMOTHORAX
covid19	0.00%	100.00%	0.00%	0.00%	0.00%
normal	100.00%	0.00%	0.00%	0.00%	0.00%
pneumonia	0.00%	0.00%	100.00%	0.00%	0.00%
pneumothorax	0.00%	0.00%	0.00%	0.00%	100.00%
tuberculosis	0.00%	0.00%	0.00%	100.00%	0.00%

Hình 5.5: Kết quả chẩn đoán

See more about image



covid19

NORMAL: 0.00%  
COVID19: 100.00%  
PNEUMONIA: 0.00%  
TUBERCULOSIS: 0.00%  
PNEUMOTHORAX: 0.00%



NORMAL: 100.00%  
COVID19: 0.00%  
PNEUMONIA: 0.00%  
TUBERCULOSIS: 0.00%  
PNEUMOTHORAX: 0.00%

Hình 5.6: Xem thêm về chẩn đoán cho các ảnh đã được chọn

# Chương 6

## Tổng kết

Mã nguồn chương trình huấn luyện và thực thi phần mềm có thể được tìm thấy tại <https://github.com/ndlongvn/AI-Project-2022>.

Tổng kết, trí tuệ nhân tạo (AI) có tiềm năng rất lớn trong việc hỗ trợ phát hiện và chẩn đoán các bệnh về phổi. Các mô hình AI có thể phân tích hình ảnh phổi từ X-quang để phát hiện các dấu hiệu bất thường. Chúng cũng có thể phân tích để đưa ra những dự đoán chính xác về khả năng mắc phải một số bệnh phổi nguy hiểm như đột quỵ.

Các mô hình AI dự đoán bệnh sẽ giúp bác sĩ phát hiện sớm, chẩn đoán chính xác và đưa ra phương pháp điều trị tối ưu cho người bệnh. Điều này có ý nghĩa rất quan trọng trong việc nâng cao hiệu quả điều trị, giảm biến chứng và tử vong do các bệnh về phổi.

Đến cuối cùng, các đóng góp của nhóm trong bài tập lớn này có thể được tóm tắt như sau:

- Từ nhu cầu thực tế xem xét và xây dựng bài toán phát hiện một số bệnh hô hấp ở người qua ảnh chụp X-quang phổi.
- Thu thập dữ liệu, bao gồm hơn 10 000 ảnh chụp X-quang phổi, từ nhiều nguồn dữ liệu khác nhau.
- Tiền xử lý dữ liệu, sử dụng nhiều phương pháp làm giàu dữ liệu và xử lý mất cân bằng dữ liệu.
- Đề xuất sử dụng các mô hình học máy huấn luyện và thực nghiệm nhiều lần để chọn được các mô hình tốt nhất với độ chính xác đạt trên 95%.
- Đưa ra tổng kết về mặt thống kê để đánh giá các mô hình học máy sử dụng.

- Xây dựng hệ thống phần mềm có thể được cài đặt trên máy tính người dùng phục vụ thuận tiện sử dụng. Người dùng chỉ cần cài đặt và có thể phát hiện bệnh qua việc upload ảnh chụp X-quang. Hệ thống cho phép đưa ra phán đoán cho đồng thời nhiều ảnh đầu vào và cho phép chọn lựa mô hình hoặc mặc định mô hình tốt nhất. Giao diện phần mềm đảm bảo tính ergonomic.

Tuy nhiên, ứng dụng xây dựng vẫn còn một số hạn chế như thiếu dữ liệu lớn, khó áp dụng cho tất cả các trường hợp phức tạp, khả năng dự đoán còn hạn chế. Do đó, sự kết hợp giữa AI và trí tuệ con người vẫn là phương pháp tối ưu trong chẩn đoán y học hiện nay.

# Tài liệu tham khảo

- [1] C.-C. Kuo, L. A. Jackson, L. A. Campbell, and J. T. Grayston, “Chlamydia pneumoniae (twar),” *Clinical microbiology reviews*, vol. 8, no. 4, pp. 451–461, 1995.
- [2] W. H. Organization *et al.*, “Coronavirus disease (covid-19), 12 october 2020,” 2020.
- [3] A. Natarajan, P. Beena, A. V. Devnikar, and S. Mali, “A systemic review on tuberculosis,” *Indian Journal of Tuberculosis*, vol. 67, no. 3, pp. 295–311, 2020.
- [4] N.-C. Huan, C. Sidhu, and R. Thomas, “Pneumothorax: Classification and etiology,” *Clinics in chest medicine*, vol. 42, no. 4, pp. 711–727, 2021.
- [5] P. Mooney, *Chest x-ray images (pneumonia)*, <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>, Accessed on 2023-04-20, 2018.
- [6] P. Gupta, *Chest x-ray images (covid-19, pneumonia)*, <https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia>, Accessed on 2023-04-20, 2020.
- [7] T. Rahman, *Tuberculosis (tb) chest x-ray dataset*, <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>, Accessed on 2023-04-20, 2021.
- [8] VBookshelf, *Pneumothorax chest x-ray images and masks*, <https://www.kaggle.com/datasets/vbookshelf/pneumothorax-chest-xray-images-and-masks>, Accessed on 2023-04-20, 2020.
- [9] C. Bento, “Multilayer perceptron explained with a real-life example and python code: Sentiment analysis,” *Towards Data Science*, 2021.
- [10] S.-C. Ma, W. Chou, T.-W. Chien, *et al.*, “An app for detecting bullying of nurses using convolutional neural networks and web-based computerized adaptive testing: Development and usability study,” *JMIR mHealth and uHealth*, vol. 8, no. 5, e16747, 2020.
- [11] G. Huang, Z. Liu, and K. Q. Weinberger, “Densely connected convolutional networks,” *CoRR*, vol. abs/1608.06993, 2016. arXiv: 1608 . 06993. [Online]. Available: <http://arxiv.org/abs/1608.06993>.

- [12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.