

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**

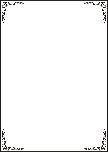
**TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

**======\*\*\*======**

****

**BÁO CÁO BÀI TẬP LỚN**

**HỌC PHẦN TRÍ TUỆ NHÂN TẠO**

ĐỀ TÀI : SỬ DỤNG CONVOLUTIONAL NEURAL NETWORK (CNN) ĐỂ PHÁT HIỆN BỆNH VIÊM PHỔI

| GVHD | Ths. Mai Thanh Hồng |
| --- | --- |
| Nhóm-Lớp | 14-2024IT6097003 |
| Họ tên sinh viên thực hiện |  |
| Lã Văn Cường | 2023602191 |
| Bùi Anh Duy | 2023606682 |
| Nguyễn Thành Hưng | 2023600565 |
| Đỗ Quốc Khánh | 2023605550 |

Hà Nội,2025

**MỤC LỤC**

[LỜI CẢM ƠN 1](#_heading=h.k82ti3bo13c0)

[1. LÝ DO CHỌN ĐỀ TÀI VÀ ỨNG DỤNG TRONG THỰC TẾ 2](#_heading=h.f661uf50yxfp)

[1.1 Lý do chọn đề tài 2](#_heading=h.nmnbw09nghru)

[1.2. Ứng dụng trong thực tế 3](#_heading=h.sr2fcly6sgau)

[2. PHÁT BIỂU BÀI TOÁN  4](#_heading=h.cxrja1cul78u)

[2.1 Đầu vào 4](#_heading=h.aytcsrfsbrbl)

[2.2 Đầu ra 4](#_heading=h.jyqjmcb8rcg7)

[3. PHƯƠNG PHÁP GIẢI QUYẾT BÀI TOÁN  4](#_heading=h.7ftrkmpl47lp)

[3.1 Cơ sở lý thuyết 4](#_heading=h.fqsvy8vf1l3u)

[3.1.1 Khái niệm thuật toán CNN (Convolutional Neural Network)  4](#_heading=h.swq34z50oedx)

[3.1.2 Convolutional là gì? 5](#_heading=h.wbxc7p15ozx4)

[3.1.3 Cấu trúc mạng CNN 5](#_heading=h.o462qg9uh43q)

[3.2 Giải quyết bài toán 10](#_heading=h.f3otax6fdu3a)

[3.2.1 Chương trình 10](#_heading=h.2ohitf6s1t41)

[3.2.2 Khởi tạo mạng CNN 10](#_heading=h.hg8jqf6bwq20)

[3.2.3 Chuẩn bị dữ liệu 11](#_heading=h.qb17ezrdse09)

[4. KẾT QUẢ 11](#_heading=h.7dq16p5e1fl)

[TÀI LIỆU THAM KHẢO 18](#_heading=h.qyoq7id27qeu)

**LỜI CẢM ƠN**

Trước hết, nhóm 14 xin gửi lời cảm ơn chân thành và sâu sắc đến giảng viên bộ môn – Cô Mai Thanh Hồng đã nhiệt tình hướng dẫn, giúp đỡ và giảng dạy trong suốt quá trình thực hiện bài tập lớn này, những kiến thức quý báu và sự động viên của cô đã giúp chúng em hoàn thiện hơn, phát triển hơn trong quá trình nghiên cứu bài tập lớn và đạt được hiệu quả tốt nhất trong suốt quá trình vừa qua.

Đồng thời nhóm 14 chúng em cũng xin gửi lời cảm ơn chân thành tới khoa CNTT trường Đại Học Công Nghiệp Hà Nội và toàn thể các thầy cô giáo trong khoa đã cung cấp những kiến thức và tài liệu, giáo trình giúp chúng em dễ dàng hoàn thiện bài báo cáo đúng thời gian và cùng nhau làm việc hiệu quả nhất.

Với điều kiện thời gian có hạn cũng như kinh nghiệm còn hạn chế nên bài báo cáo sẽ không tránh khỏi những thiếu sót. Nhóm 14 rất mong nhận được sự chỉ bảo, đóng góp ý kiến của cô để nhóm em có điều kiện bổ sung, nâng cao kiến thức của mình, phục vụ tốt hơn công tác thực tế sau này.

Chúng em xin trân thành cảm ơn!

# 1. LÝ DO CHỌN ĐỀ TÀI VÀ ỨNG DỤNG TRONG THỰC TẾ

## 1.1 Lý do chọn đề tài

Viêm phổi là tình trạng viêm cấp tính của phổi do nhiễm trùng. Chẩn đoán ban đầu thường dựa trên chụp X-quang phổi và các dấu hiệu lâm sàng. Nguyên nhân, triệu chứng, điều trị, các biện pháp phòng ngừa và tiên lượng khác nhau tùy thuộc căn nguyên là vi khuẩn, lao, virus, nấm hay ký sinh trùng; cũng như viêm phổi là mắc phải từ cộng đồng, bệnh viện, có thể xảy ra ở bệnh nhân thông khí xâm nhập hay viêm phổi trên một bệnh nhân không có khả năng miễn dịch hay bị suy giảm miễn dịch.

Ước tính có từ 4 đến 5 triệu người ở Hoa Kỳ mắc viêm phổi mỗi năm, trong đó khoảng 55.000 người tử vong. Ở Mỹ, viêm phổi cùng với cúm, là nguyên nhân tử vong xếp thứ 8 và là nguyên nhân tử vong hàng đầu trong các bệnh nhiễm trùng. Viêm phổi là bệnh nhiễm trùng gây tử vong phổ biến nhất ở bệnh viện và ở các nước đang phát triển.

Đặc biệt trong đợt đại dịch Covid 19, những người mắc phải bệnh dù đã điều trị hết bệnh nhưng cơ quan phổi vẫn có thể bị tổn thương, những tổn thương đó là điều kiện thuận lợi khiến cho các bệnh lý lâu năm của người bệnh dễ tiến triển nặng hơn trước, đặc biệt là nhóm bệnh liên quan đến phổi.

Các loại bệnh viêm phổi:

+ Viêm phổi do vi khuẩn: Vi khuẩn là nguyên nhân của hầu hết các trường hợp viêm phổi cộng đồng ở người trưởng thành.

+ Viêm phổi do virus: Virus là nguyên nhân gây ra số trường hợp viêm phổi nhiều thứ hai sau vi khuẩn. Có rất nhiều loại virus gây viêm phổi, chẳng hạn như các loại virus gây ra cảm lạnh cũng như virus cúm.

+Viêm phổi do nấm: Nấm là một tác nhân gây viêm phổi ít phổ biến. Nếu là một người trưởng thành khỏe mạnh thì nguy cơ xuất hiện viêm phổi do nấm là rất thấp, nhưng nếu bị suy giảm miễn dịch vì bất kì lí do nào thì khả năng mắc viêm phổi do nấm sẽ tăng lên.

+Viêm phổi do hóa chất: Viêm phổi do hóa chất là loại viêm phổi đặc thù, rất ít gặp. Nhiều loại hóa chất có thể gây viêm phổi, và chúng có thể ở bất kỳ dạng nào, từ dạng hơi, dạng lỏng cho tới các phân tử rắn. Bên cạnh tổn thương phổi, các hóa chất còn có thể gây hại cho nhiều cơ quan khác.

Vì vậy, với thực trạng bệnh viêm phổi ngày nay diễn biến hết sức phức tạp, nhóm em đã lựa chọn đề tài Sử dụng Convolution Neural Network (CNN) để phát hiện bệnh viêm phổi. Việc sử dụng CNN sẽ phân tích hình ảnh chụp X-Quang phổi và dựa vào đặc tính của bệnh viêm phổi giúp việc phát hiện bệnh nhân có mắc bệnh viêm phổi hay không một cách nhanh chóng hơn với độ chính xác cao. Việc phát hiện bệnh viêm phổi sớm, sẽ giúp bệnh nhân điều trị kịp thời, nhanh chóng và giúp phổi tránh được những tổn thương nặng. Ngoài ra, trong quá trình làm nhóm em sẽ hiểu biết hơn về CNN từ đó nhóm em có thể phát triển và áp dụng CNN vào nhiều bài toán ứng dụng thực tế hơn.

## 1.2. Ứng dụng trong thực tế

Trong Y học hiện nay đã ghi nhận nhiều loại bệnh khác nhau (trên 14 bệnh) liên quan đến phổi và biến chứng thành ung thư phổi, đồng thời đối với mỗi loại bệnh lại có thể chia thành nhiều nguyên nhân khác nhau. Tuy nhiên hầu hết các bệnh nhân mắc bệnh liên quan đến phổi đều có một số triệu chứng tương đồng nhau như những cơn ho, khó thở, đau ngực, nên thông thường rất khó để phân biệt viêm phổi và các bệnh về đường hô hấp do triệu chứng rất giống nhau và điều đó gây nhiều khó khăn cho người bác sĩ chẩn đoán, đặc biệt là những người còn thiếu kinh nghiệm chuẩn đoán. Từ những yếu tố trên, các bác sĩ chẩn đoán bệnh sẽ không thể tránh khỏi những sai sót ngoài ý muốn. Chính vì vậy, cộng đồng vẫn đang tích cực nghiên cứu ra những phương án mới để có thể giảm thiểu mức độ sai sót xuống thấp nhất và có thể giúp cho các bác sĩ thêm phần tự tin về những chuẩn đoán của mình trong bất kì hoàn cảnh nào. Vì vậy, nhóm chúng em quyết định nghiên cứu và ứng dụng xử lý ảnh và AI vào y học, có thể giúp phân biệt và phát hiện bệnh viêm phổi dễ dàng và chính xác hơn.

# 2. PHÁT BIỂU BÀI TOÁN

## 2.1 Đầu vào

Với bài toán nhận biết bệnh nhân có bị bệnh phổi hay không thì đầu vào bài toán sẽ là 1 bức ảnh chụp x-quang vùng phổi của bệnh nhân. Bức ảnh này sẽ được đưa vào bên trong chương trình và chương trình sẽ xử lý dữ liệu. Bài toán sử dụng thuật toán CNN dùng để trích xuất ảnh đặc trưng, qua đó nhóm em sẽ có 2 loại dữ liệu để cho thuật toán training là hình ảnh chụp x-quang bệnh nhân bị viêm phổi và hình ảnh chụp x-quang bệnh nhân phổi bình thường.

Các ảnh được đưa vào chương trình sẽ phải qua một bước tiền xử lý để đưa về phù hợp với mô hình đã huấn luyện

## 2.2 Đầu ra

Đầu ra sẽ là kết quả cho biết tấm phim chụp x-quang của người nào đó có bị mắc bệnh viêm phổi hay không.

# 3. PHƯƠNG PHÁP GIẢI QUYẾT BÀI TOÁN

## 3.1 Cơ sở lý thuyết

### 3.1.1 Khái niệm thuật toán CNN (Convolutional Neural Network)

Thuật toán Convolutional neural network còn gọi là Mạng nơ ron tích chập, thường được viết tắt là CNN. Đây là một trong những mô hình của Deep Learning (Deep learning là tập hợp các thuật toán để cố gắng mô hình dữ liệu trừu tượng hóa ở mức cao bằng cách sử dụng nhiều lớp xử lý với cấu trúc phức tạp hoặc bằng cách khác).

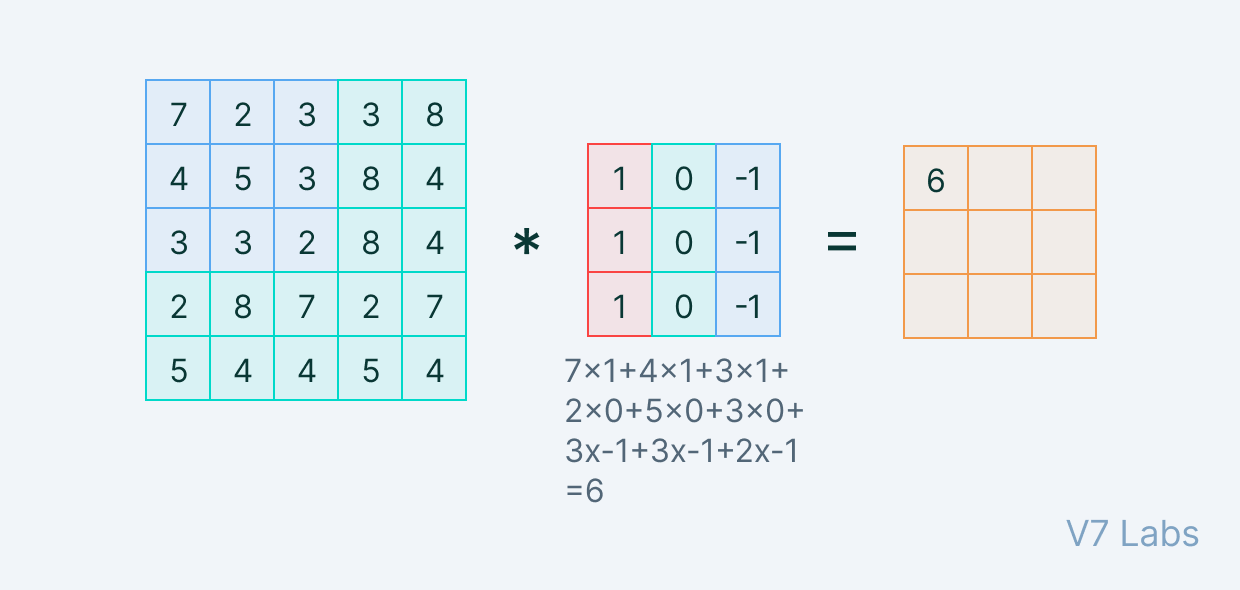
Convolutional Neural Network, đây là một deep neural network architecture. Hiểu đơn giản, nó cũng chính là một dạnh Artificial Neural Network, một Multi-layer Perceptron nhưng mang thêm 1 vài cải tiến, đó là Convolution và Pooling.

Trong đó Artificial neural network (Mạng neuron nhân tạo) là một mô hình toán học hay mô hình tính toán lấy cảm hứng dựa trên cấu trúc của mạng thần kinh. Một mạng neuron bao gồm các nhóm neuron được nối với nhau, trên cơ sở đó thông tin được xử lý

Multi-layer Perceptron ( Mạng neuron truyền thẳng nhiều lớp)

### 3.1.2 Convolutional là gì?

Là một cửa sổ trượt (Sliding Windows) trên một ma trận như mô tả hình dưới:



Các convolutional layer có các parameter(kernel) đã được học để tự điều chỉnh lấy ra những thông tin chính xác nhất mà không cần chọn các feature.

### 3.1.3 Cấu trúc mạng CNN

Mạng CNN là một tập hợp các lớp Convolution chồng lên nhau và sử dụng các hàm nonlinear activation như ReLU và tanh để kích hoạt các trọng số trong các node. Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo.

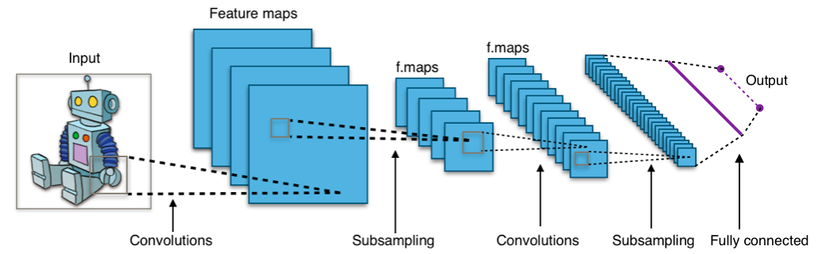
Mỗi một lớp sau khi thông qua các hàm kích hoạt sẽ tạo ra các thông tin trừu tượng hơn cho các lớp tiếp theo. Trong mô hình mạng truyền ngược (feedforward neural network) thì mỗi neural đầu vào (input node) cho mỗi neural đầu ra trong các lớp tiếp theo.

Mô hình này gọi là mạng kết nối đầy đủ (fully connected layer) hay mạng toàn vẹn (affine layer). Còn trong mô hình CNNs thì ngược lại. Các layer liên kết được với nhau thông qua cơ chế convolution.

Layer tiếp theo là kết quả convolution từ layer trước đó, nhờ vậy mà ta có được các kết nối cục bộ. Như vậy mỗi neuron ở lớp kế tiếp sinh ra từ kết quả của filter áp đặt lên một vùng ảnh cục bộ của neuron trước đó.

Mỗi một lớp được sử dụng các filter khác nhau thông thường có hàng trăm hàng nghìn filter như vậy và kết hợp kết quả của chúng lại. Ngoài ra có một số layer khác như pooling/subsampling layer dùng để chắt lọc lại các thông tin hữu ích hơn (loại bỏ các thông tin nhiễu).

Trong quá trình huấn luyện mạng (training) CNN tự động học các giá trị qua các lớp filter dựa vào cách thức mà bạn thực hiện. Ví dụ trong tác vụ phân lớp ảnh, CNNs sẽ cố gắng tìm ra thông số tối ưu cho các filter tương ứng theo thứ tự raw pixel > edges > shapes > facial > high-level features. Layer cuối cùng được dùng để phân lớp ảnh.



Trong mô hình CNN có 2 khía cạnh cần quan tâm là **tính bất biến** (Location Invariance) và **tính kết hợp** (Compositionality). Với cùng một đối tượng, nếu đối tượng này được chiếu theo các gốc độ khác nhau (translation, rotation, scaling) thì độ chính xác của thuật toán sẽ bị ảnh hưởng đáng kể

Pooling layer sẽ cho bạn tính bất biến đối với phép dịch chuyển (translation), phép quay (rotation) và phép co giãn (scaling). Tính kết hợp cục bộ cho ta các cấp độ biểu diễn thông tin từ mức độ thấp đến mức độ cao và trừu tượng hơn thông qua convolution từ các filter.

Mạng CNN sử dụng 3 ý tưởng cơ bản

**các trường tiếp nhận cục bộ** (local receptive field)

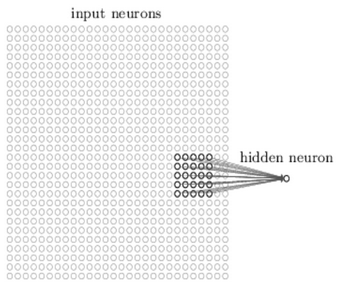
**trọng số chia sẻ** (shared weights)

**tổng hợp** (pooling).

Trường tiếp nhận cục bộ (local receptive field)

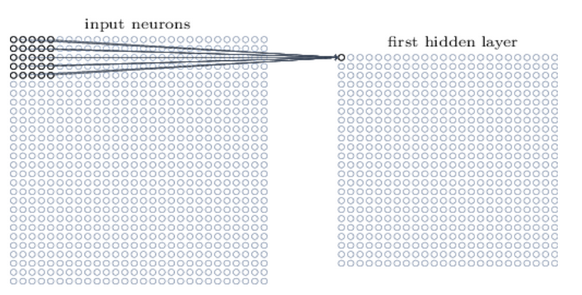
Đầu vào của mạng CNN là một ảnh. Ví dụ như ảnh có kích thước 28×28 thì tương ứng đầu vào là một ma trận có 28×28 và giá trị mỗi điểm ảnh là một ô trong ma trận. Trong mô hình mạng ANN truyền thống thì chúng ta sẽ kết nối các neuron đầu vào vào tầng ảnh.

Tuy nhiên trong CNN chúng ta không làm như vậy mà chúng ta chỉ kết nối trong một vùng nhỏ của các neuron đầu vào như một filter có kích thước 5×5 tương ứng (28- 5 + 1) 24 điểm ảnh đầu vào. Mỗi một kết nối sẽ học một trọng số và mỗi neuron ẩn sẽ học một bias. Mỗi một vùng 5×5 đấy gọi là một trường tiếp nhận cục bộ.

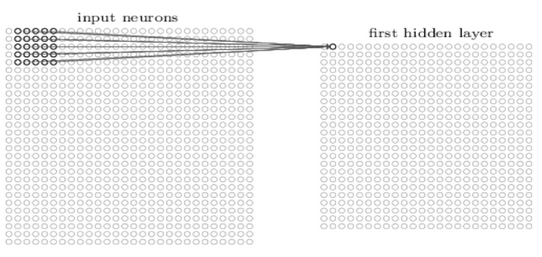


Một cách tổng quan, ta có thể tóm tắt các bước tạo ra 1 hidden layer bằng các cách sau:

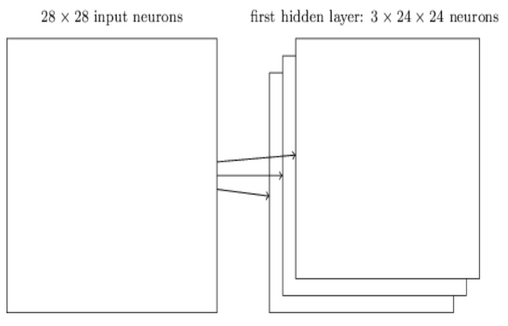
Bước 1 : **Tạo ra neuron ẩn đầu tiên trong lớp ẩn 1**



Bước 2 : **Dịch filter qua bên phải một cột sẽ tạo được neuron ẩn thứ 2.**



với bài toán nhận dạng ảnh người ta thường gọi ma trận lớp đầu vào là feature map, trọng số xác định các đặc trưng là shared weight và độ lệch xác định một feature map là shared bias. Như vậy đơn giản nhất là qua các bước trên chúng ta chỉ có 1 feature map. Tuy nhiên trong nhận dạng ảnh chúng ta cần nhiều hơn một feature map.



Như vậy, local receptive field thích hợp cho việc phân tách dữ liệu ảnh, giúp chọn ra những vùng ảnh có giá trị nhất cho việc đánh giá phân lớp.

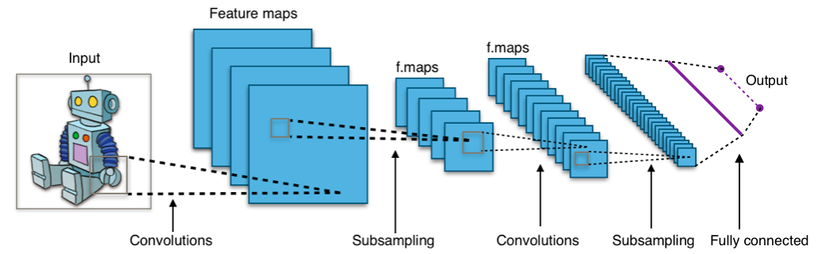
Chia sẻ trọng số (shared weight and bias)

Đầu tiên, các trọng số cho mỗi filter (kernel) phải giống nhau. Tất cả các neuron trong lớp ẩn đầu sẽ phát hiện chính xác feature tương tự chỉ ở các vị trí khác nhau trong hình ảnh đầu vào. Chúng ta gọi việc map từ input layer sang hidden layer là một feature map. Vậy mối quan hệ giữa số lượng Feature map với số lượng tham số là gì?

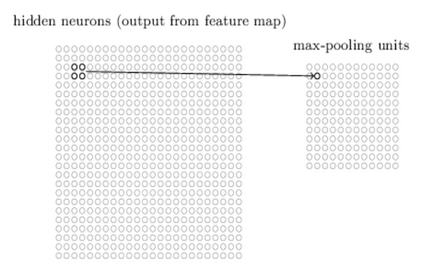
Tóm lại, một convolutional layer bao gồm các feature map khác nhau. Mỗi một feature map giúp detect một vài feature trong bức ảnh. Lợi ích lớn nhất của trọng số chia sẻ là giảm tối đa số lượng tham số trong mạng CNN.

Lớp tổng hợp (pooling layer

Lớp pooling thường được sử dụng ngay sau lớp convolutional để đơn giản hóa thông tin đầu ra để giảm bớt số lượng neuron.

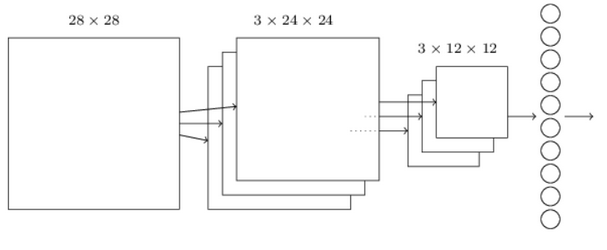


Thủ tục pooling phổ biến là max-pooling, thủ tục này chọn giá trị lớn nhất trong vùng đầu vào 2×2



Như vậy qua lớp Max Pooling thì số lượng neuron giảm đi phân nửa. Trong một mạng CNN có nhiều Feature Map nên mỗi Feature Map chúng ta sẽ cho mỗi Max Pooling khác nhau. Chúng ta có thể thấy rằng Max Pooling là cách hỏi xem trong các đặc trưng này thì đặc trưng nào là đặc trưng nhất. Ngoài Max Pooling còn có L2 Pooling.

Cuối cùng ta đặt tất cả các lớp lại với nhau thành một CNN với đầu ra gồm các neuron với số lượng tùy bài toán.



2 lớp cuối cùng của các kết nối trong mạng là một lớp đầy đủ kết nối (fully connected layer) . Lớp này nối mọi neuron từ lớp max-pooled tới mọi neuron của tầng ra.

Cách chọn tham số cho CNN

Số các convolution layer: càng nhiều các convolution layer thì performance càng được cải thiện. Sau khoảng 3 hoặc 4 layer, các tác động được giảm một cách đáng kể

Filter size: thường filter theo size 5×5 hoặc 3×3

Pooling size: thường là 2×2 hoặc 4×4 cho ảnh đầu vào lớn

Cách cuối cùng là thực hiện nhiều lần việc train test để chọn ra được param tốt nhất.

## 3.2 Giải quyết bài toán

3.2.1 Chương trình

Input: Là một ảnh chụp X-Quang do bác sĩ chup bệnh nhân

Output: Kết quả nhận diện có phải bệnh viêm phổi không

Chương trình gồm chức năng nhận ảnh từ bên ngoài để nhận dạng

### 3.2.2 Khởi tạo mạng CNN

* Input: 12 layers
* 4 : 2D convolution layer
* 4 : MaxPooling2D layer
* 1 : Flatten layer
* 1 : Dropout layer
* 2 : Dense layer
* Output: 2 class
* Pneumonia
* Non-Pneumonia

3.2.3 Chuẩn bị dữ liệu

* 3386 image : chia làm 2 tập train, test
* Training Dataset : 2762 images belonging to 2 classes.
* Testing Dataset : 624 images belonging to 2 classes.

# 4. KẾT QUẢ

Chương trình đã được triển khai bằng ngôn ngữ Python viết trên GoogleColab , sử dụng bộ dataset từ kaggle

[https://drive.google.com/drive/folders/1ARa-tiX9NbtQLv6gT6vLbpcd3l9JGL2s?usp=drive\_link](about:blank)

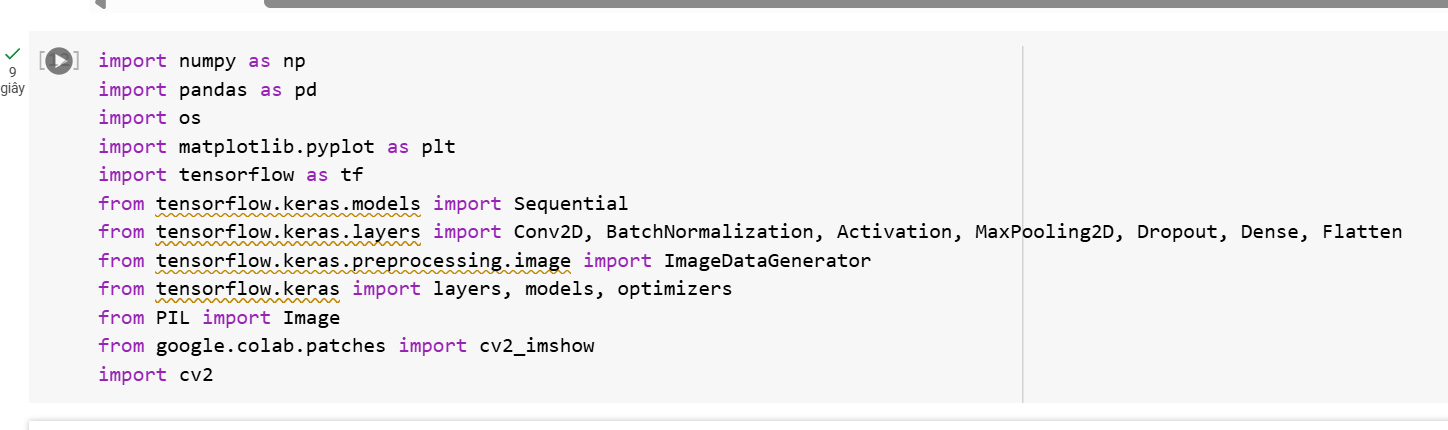
Độ chính xác : 87.17 % trên tập test

Dưới đây là phần chương trình : ở đây chúng em có viết chương trình trên google colab

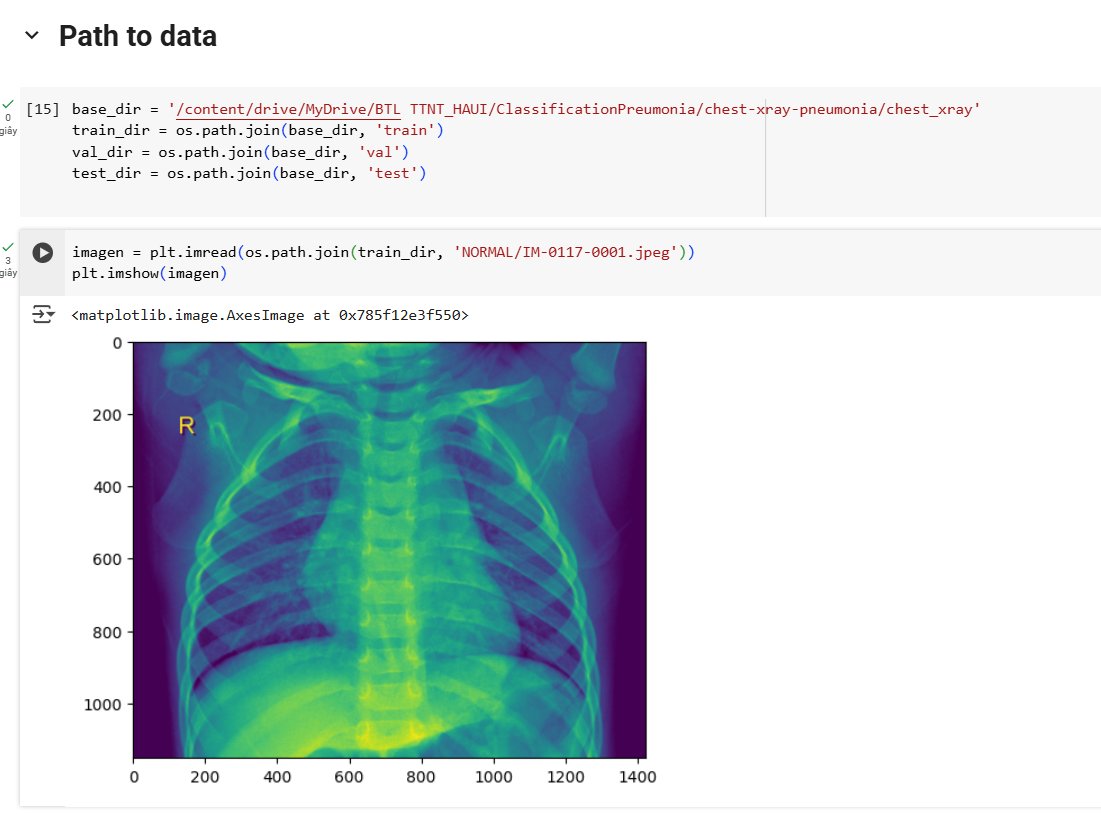
Đầu tiên ta sẽ đi thu thập data : nhóm em đã tìm kiếm được bộ data trên kaggle nên lấy luôn để sử dụng



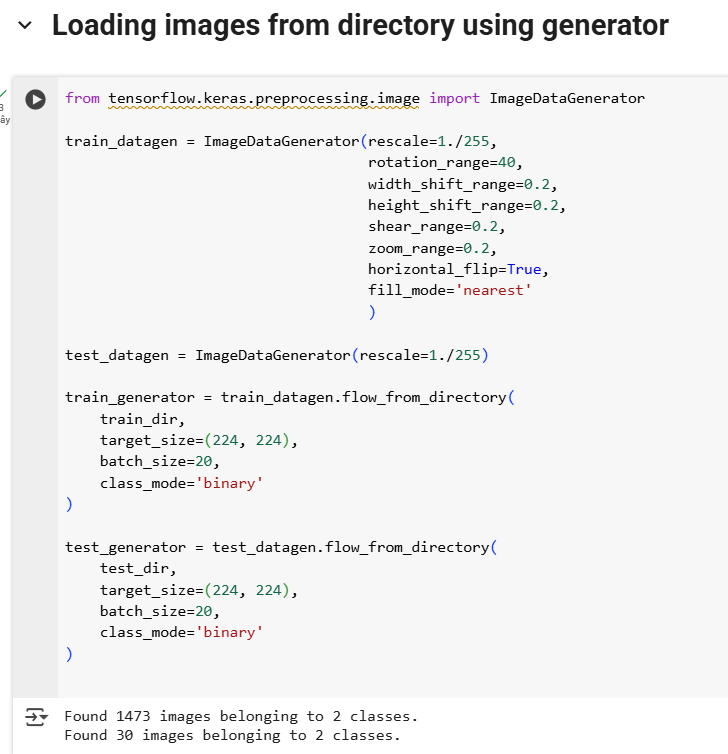
Bước tiếp theo : khai báo các thư viện cần thiết cho chương trình



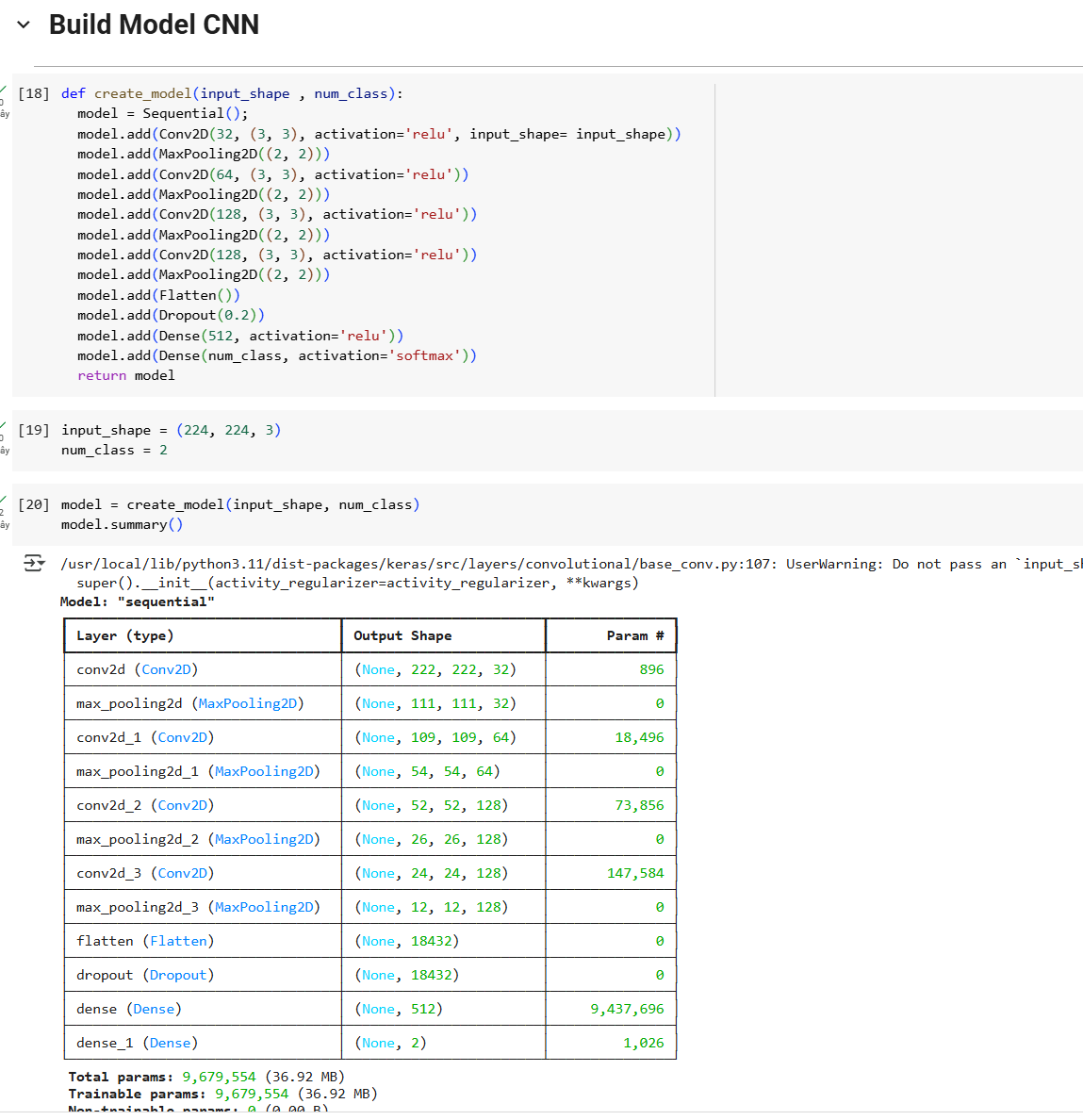
Tiếp đến là : bước lấy đường dẫn tới các mục chứa data



Tiếp theo là : load data và đưa data về cùng 1 kích thước để tiện cho việc huấn luyện mô hình



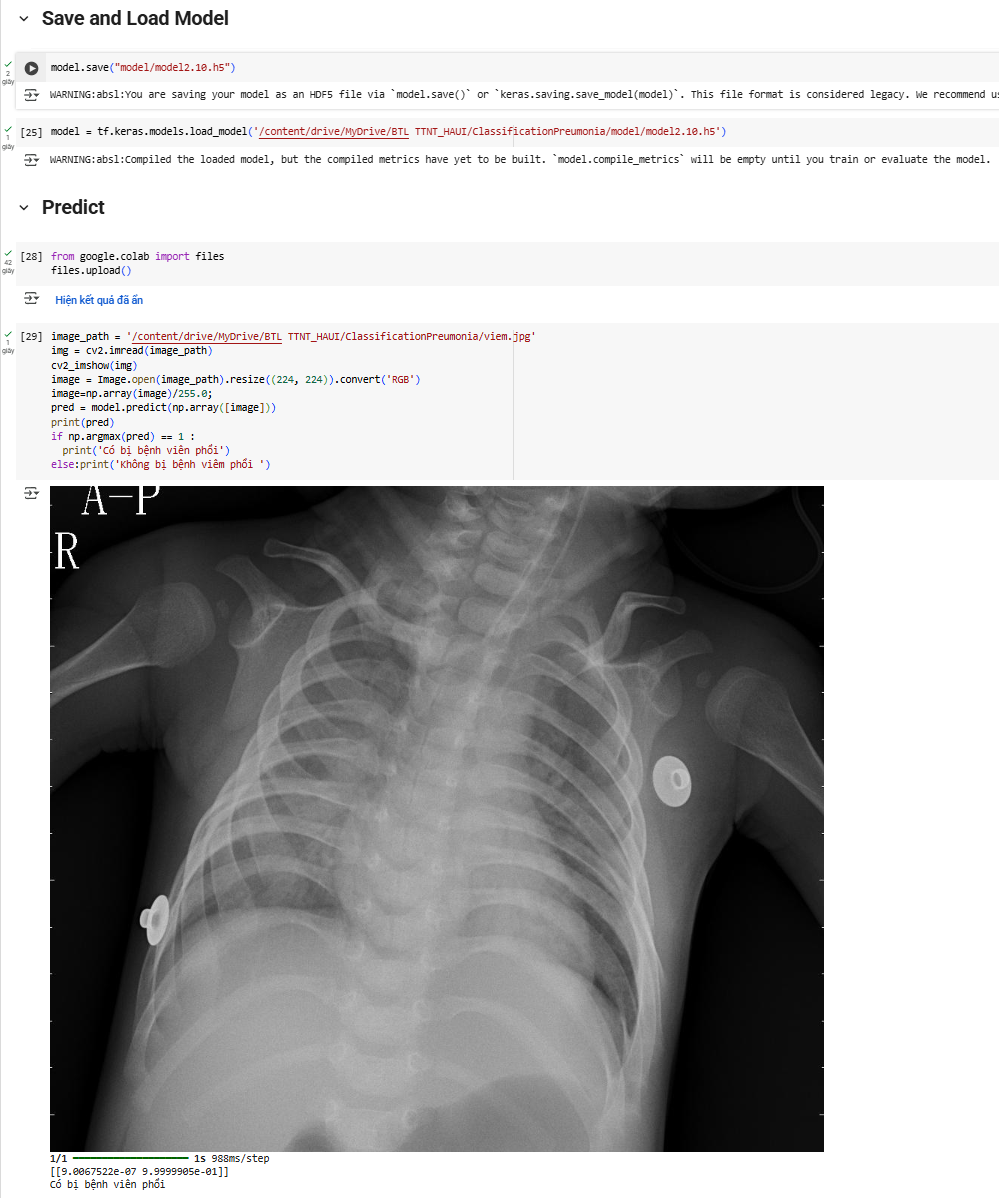
Tiếp đến chúng ta sẽ : xây dựng 1 model CNN cơ bản



Tiếp theo chúng ta sẽ tiến hành train model , ở đây em train với 100 epoch



Và cuối cùng sẽ là bước đánh giá model và test lại model



**TÀI LIỆU THAM KHẢO**

1. Nguyên Phương Nga (ch.b), Trần Hung Cương, Giao trình trí tuê nhân tạo,

Nha xuất bản Thông kê, Ha Nội, 2021.

2. Nguyên Đức Nghĩa, Trí tuê nhân tạo – Cơ sở va ứng dụng, NXB Khoa hoc va

Kỹ thuật, Ha Nội, 2019.

3. Nguyên Tấn Trần Minh Khang. Trí tuê nhân tạo: Cac phương phap va ứng

dụng. NXB Đại hoc Quôc gia TP.HCM, 2016.

4. Mô hình CNN CheXNet: Radiologist-Level Pneumonia Detection on Chest

X-Rays của Stanford.

5. https://www.sciencedirect.com/science/article/pii/S2772442523000436 (truy

cập lúc 11:28 pm 5/6/2025)