Human Data Analytics Course Final Project

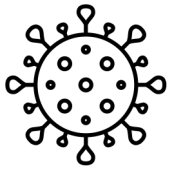
DETECTING COVID-19 AND PNEUMONIA FROM CHEST X-RAY IMAGES

Prof. Michele Rossi

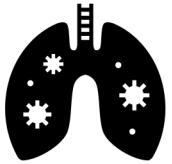
Physics of data

Asal Rangrasiasl - 2046832

22/02/2023

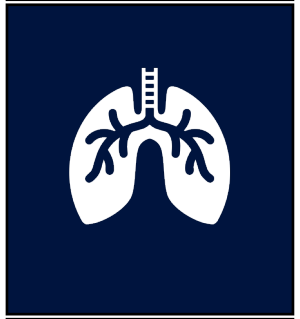


CONTENTS



INTRODUCTION

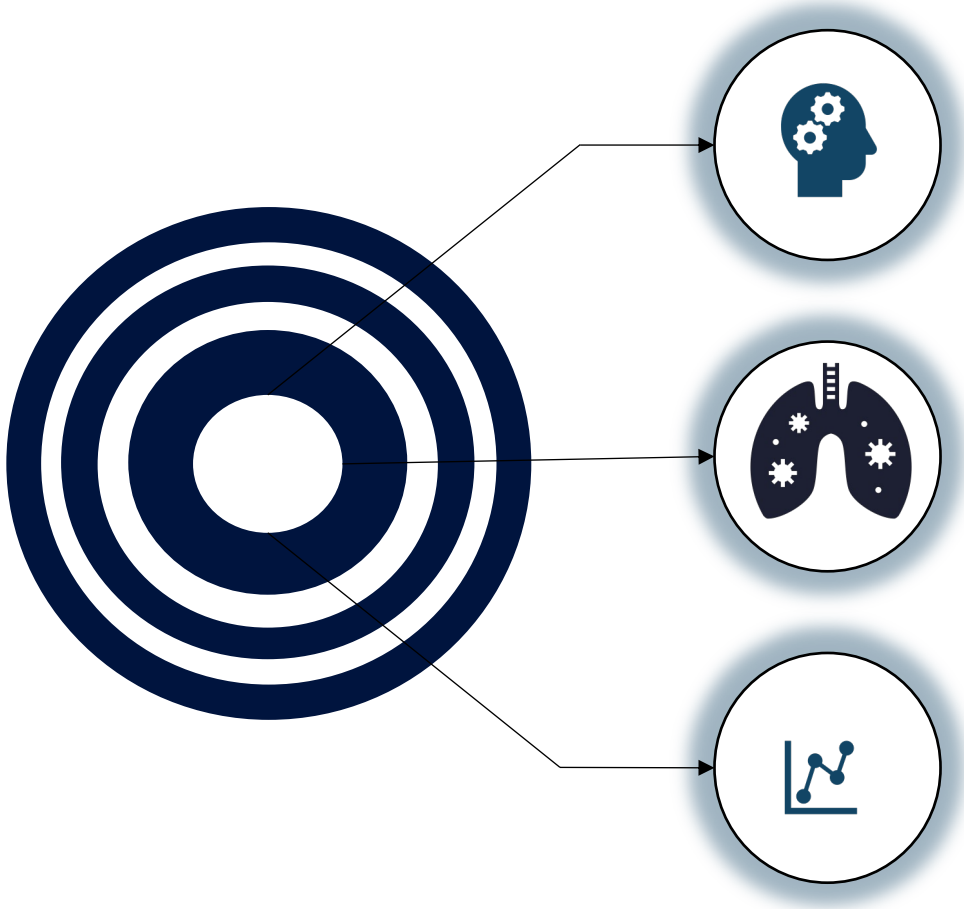
As the need for medical treatments grows exponentially, there has been an unprecedented demand for the quick diagnosis of illnesses.



This study proposes a computer-aided diagnosis system for automatic disease detection using Chest X- ray images of three classes COVID-19, Pneumonia, and Healthy.

The prediction can be done through deep learning algorithms, such as convolutional neural networks (CNNs). The algorithms are trained on a large dataset of X-ray images and can accurately classify the images into different disease categories

AIM OF THE PROJECT



Proposed Method :
A CNN model would take an X-ray chest image as input and then classify the image into different categories, such as ”normal” or ”covid” and ”pneumonia”.

Feature Extraction :
CNN uses filters to scan the input image and extract local features, such as edges and shapes

Evaluate:
Training and Validation accuracy of CNN for classification, Precision , Recall , F1-score and confusion matrix

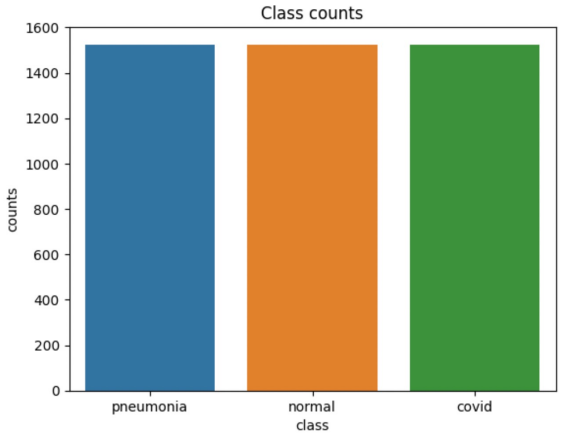
MATERIAL AND METHODS

Understanding The Context	Retrieving The Data	Dataset Preparation	Building Models and Training	Evaluation of Models and Reporting
study about different methods for classification of x-chest	Chest X-ray PA DatasetPaper published by Amanullah Asraf, Zabirul Islam	Adding the labels for each group of dataset	We built two models to compare and evaluate for the best: **CNN** and **pre-trained DenseNet**.	As a final result, we produced and compared accuracy and loss plots and confusion matrix.

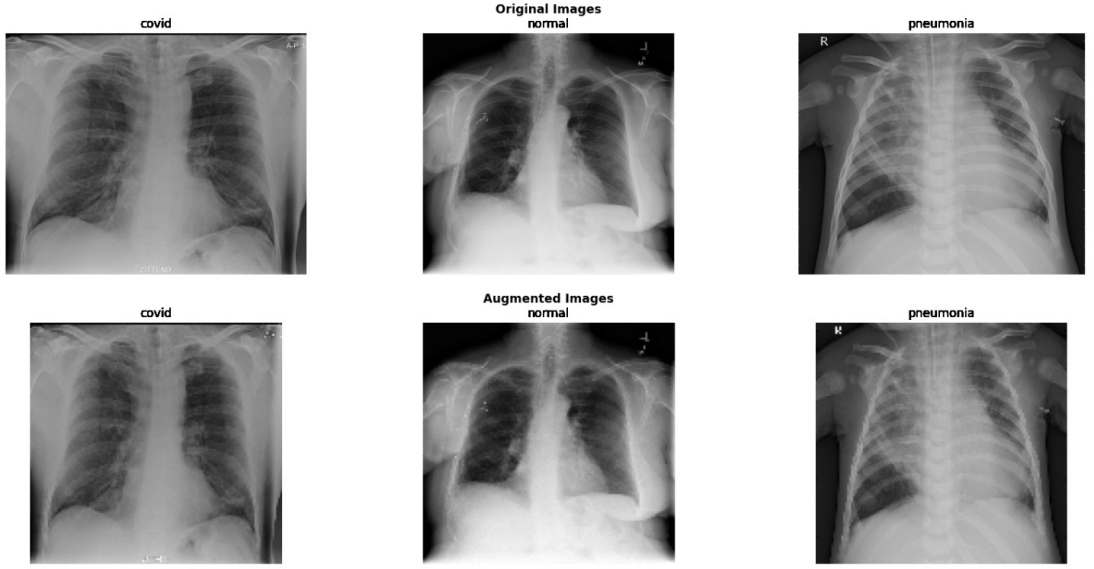
DATASET DESCRIPTION

Chest X-ray PA Dataset: The dataset consist of 1525 chest x-ray images per each class which sums up to 4575 in total

- Data augmentation
- Splitting: 70% training, 30% testing.



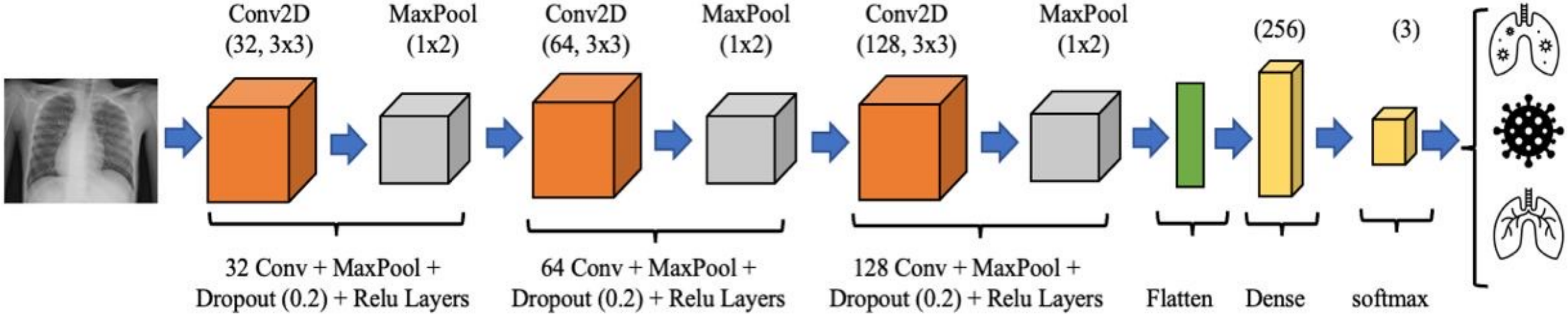
Class 0	• COVID-19
Class 1	• NORMAL
Class 2	• PNEUMONIA



METHODOLOGY: CNN

Convolutional Neural Network (CNN)

- Commonly used for image classification tasks.
- A CNN consists of multiple layers, including convolutional layers, activation layers, and pooling layers, that extract and simplify the features in an image, reducing the dimensionality of the data and making it easier to classify.



METHODOLOGY: CNN

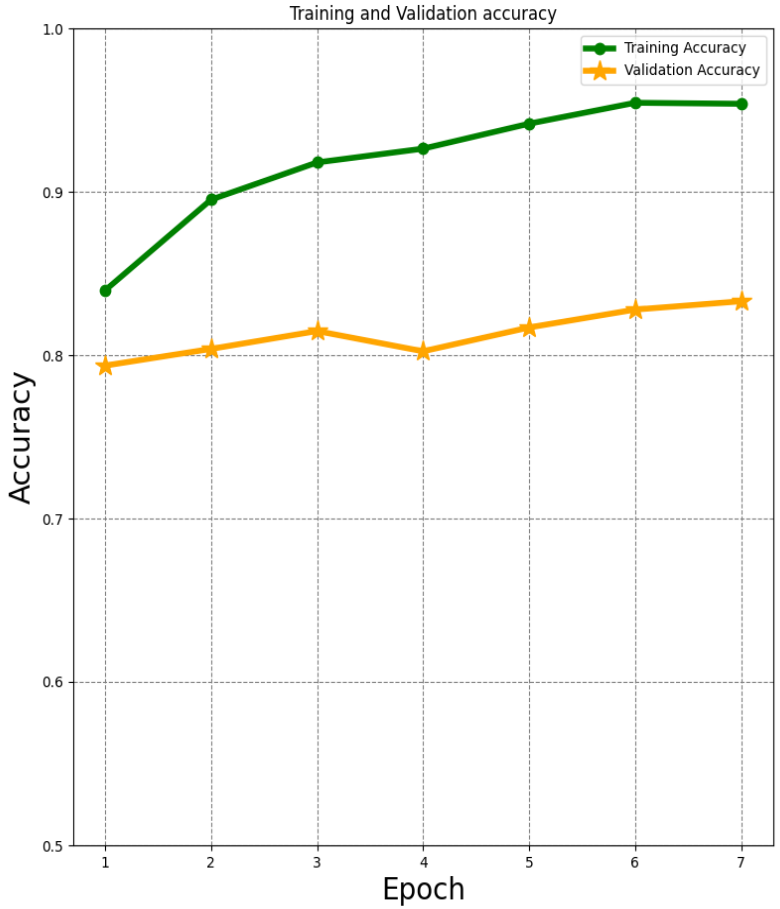
A CNN model would take an X-ray chest image as input and then classify the image into different categories, such as ”normal” or ”covid” and ”pneumonia”.

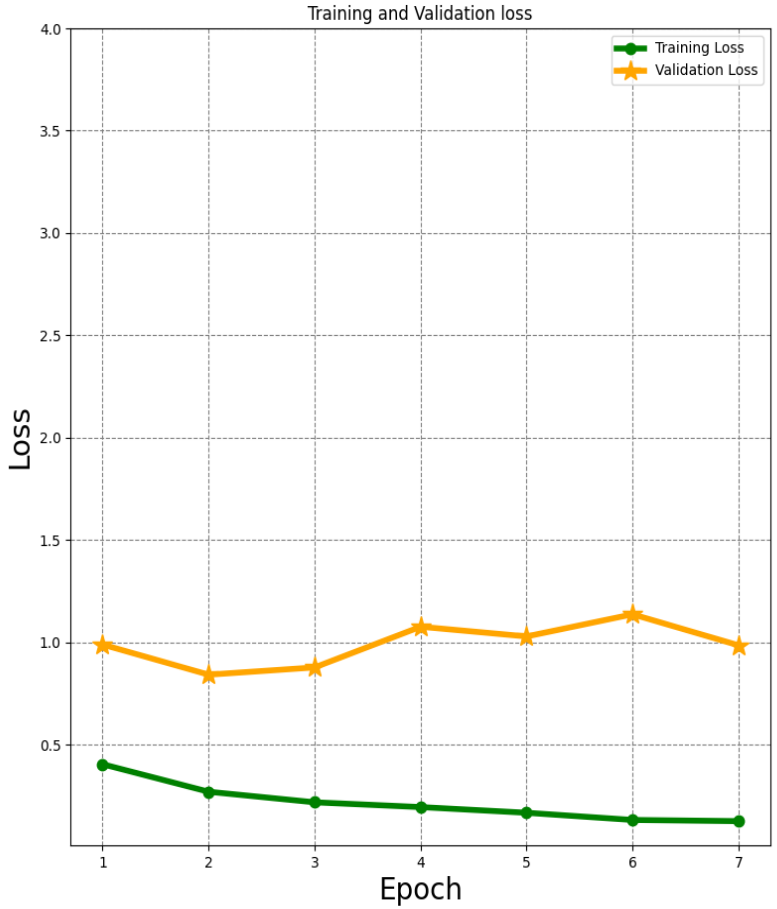
- **Kernel**
  - Size=(3,3)
- **Dropout**
  - Rate = 0.2 , 0.1
- **Batch size**
  - 32
- **epoch**
  - 20
- **Optimizer**
  - Adam
  - Learning rate = 0.0001
- **Loss**
  - Categorical cross entropy

layer	Shape	Model Setting
Input	128 × 128	
Conv2D-1 MaxPooling2D-1 Dropout-1	32 × 3 × 3 1 × 2	activation=Relu Kernel-size=(3,3) dropout-rate=0.2
Conv2D-2 MaxPooling2D-2 Dropout-2	64 × 3 × 3 1 × 2	activation=Relu Kernel-size=(3,3) dropout-rate=0.2
Conv2D-3 MaxPooling2D-3 Dropout-3	128 × 3 × 3 1 × 2	activation=Relu Kernel-size=(3,3) dropout-rate=0.2
Flatten Dense-1 Dropout-3	256	activation=Relu dropout-rate=0.1
Output	3	activation=Softmax

RESULTS

Training & validation accuracy : model achieved 90.01% and 80.38% train and test accuracy, respectively





RESULTS

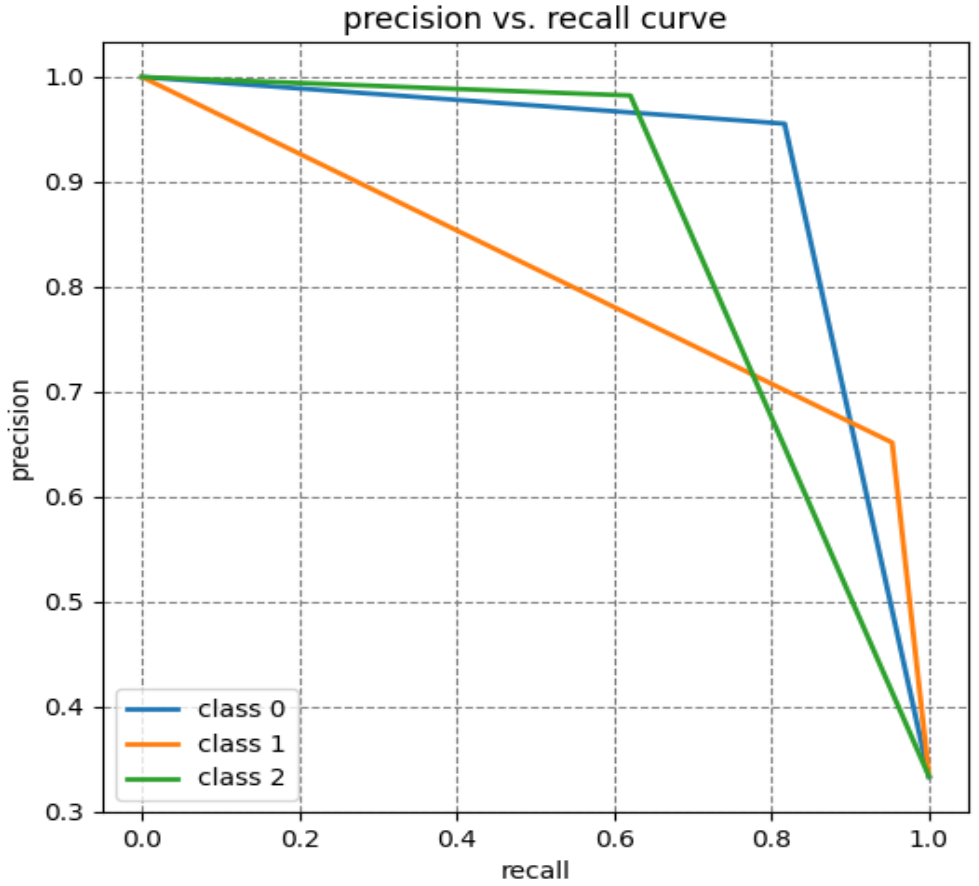
Precision (1), Recall (2), and F1-score (3) are used to evaluate the classification model performance

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives}$$

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

	precision	recall	f1-score
0	0.96	0.82	0.88
1	0.65	0.95	0.77
2	0.98	0.62	0.76



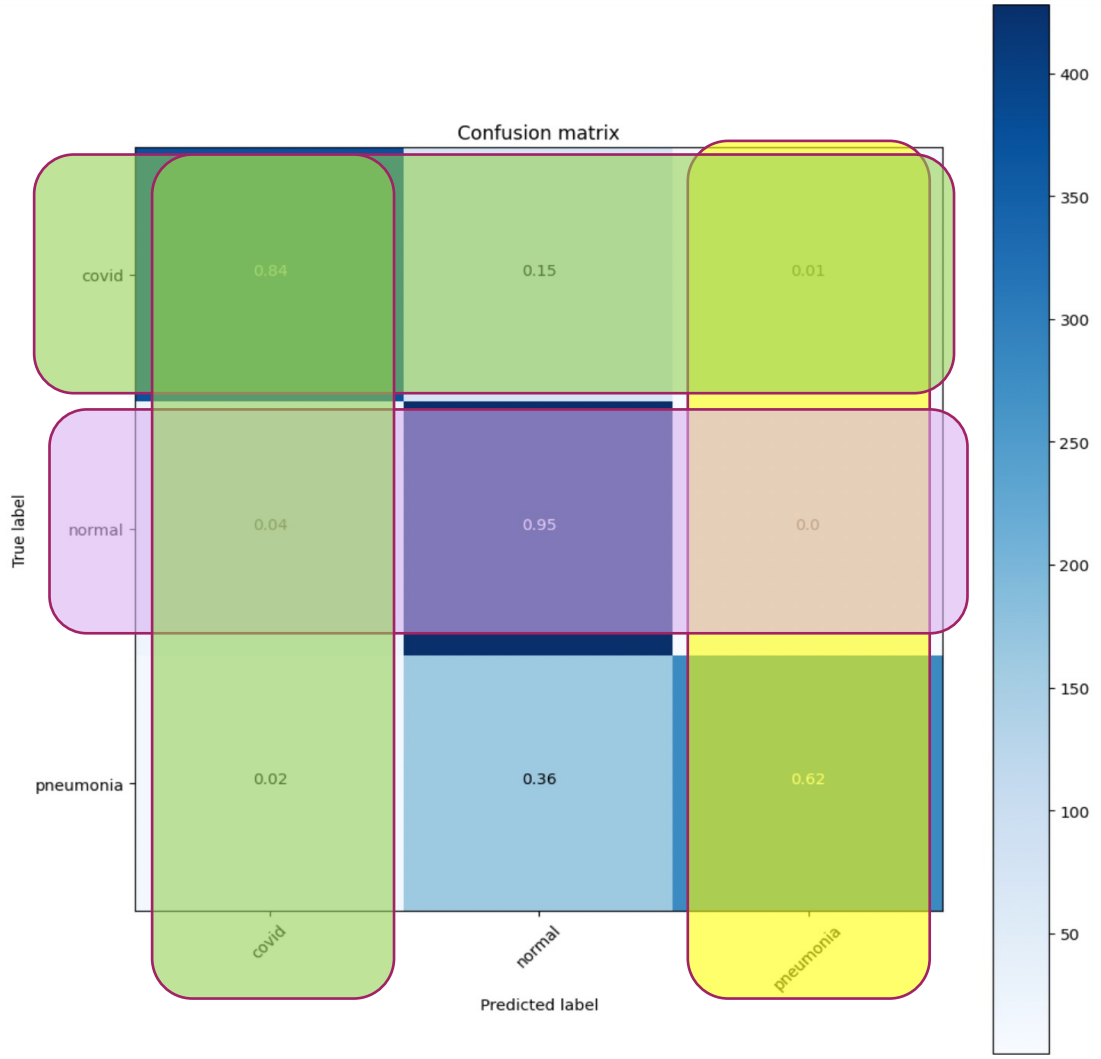
Class 0	• COVID-19
Class 1	• NORMAL
Class 2	• PNEUMONIA

RESULTS

A confusion matrix for the test dataset

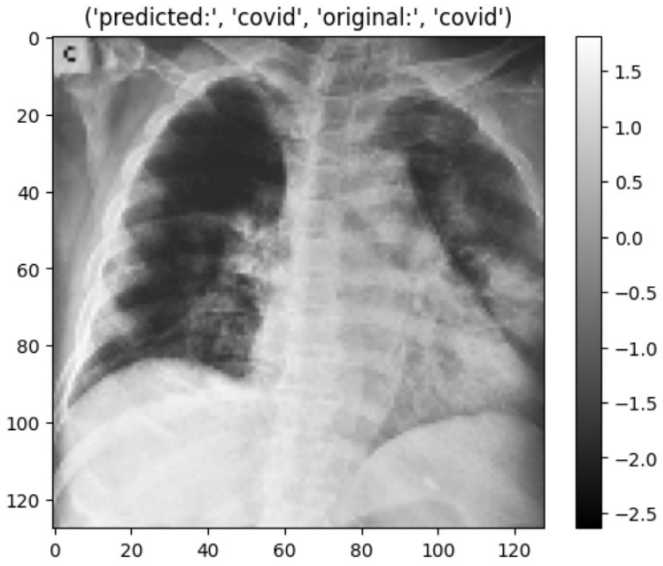
Best class per different Performance metrics:

- Precision:
  - pneumonia
- Recall:
  - normal
- F1-score:
  - Covid

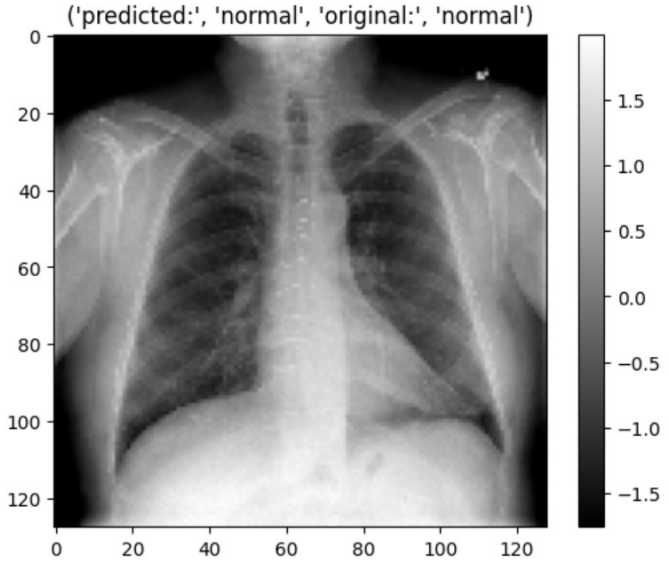


Model Performance of CNN

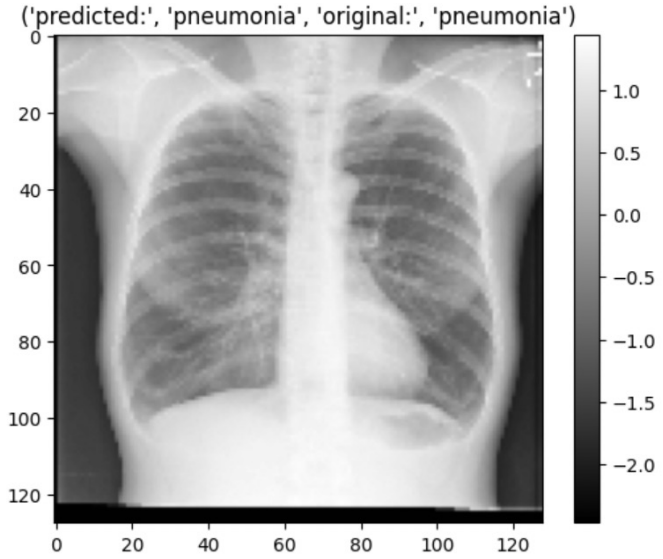
Demo : one sample per each class of images from a batch of test dataset to show the prediction.



Predicted : Covid
Original : Covid

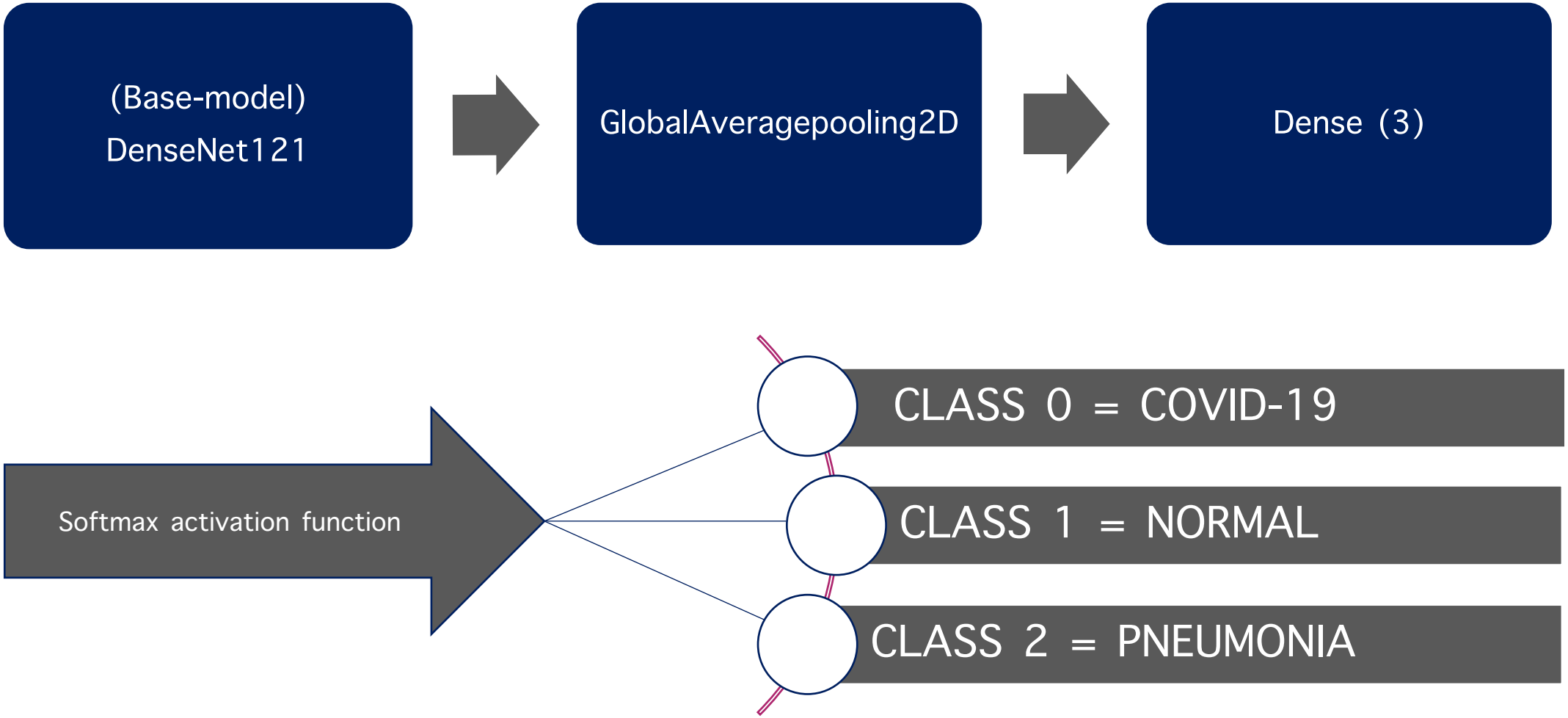


Predicted : Normal
Original : Normal



Predicted : Pneumonia
Original : Pneumonia

Pre-trained DenseNet



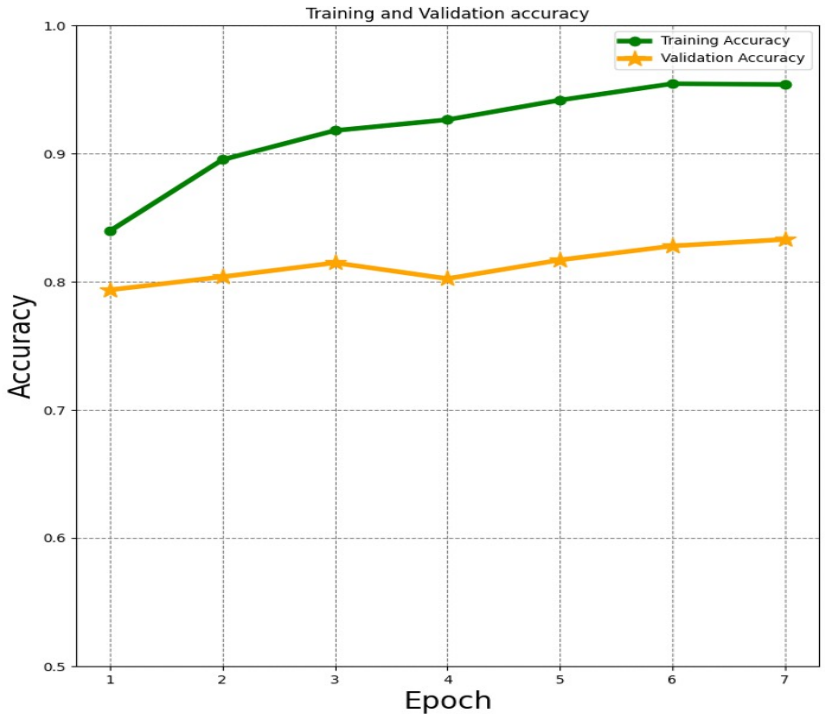
COMPARISION:

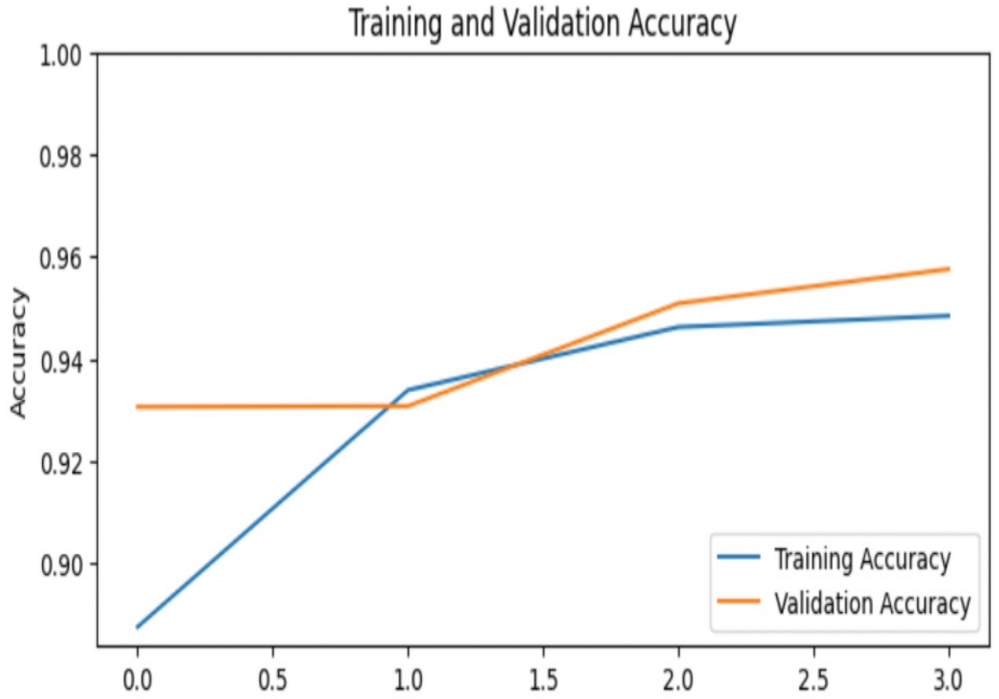
CNN

Test accuracy : 80.37 %

Convolutional Base (DenseNet121)

Test accuracy : 95.76 %

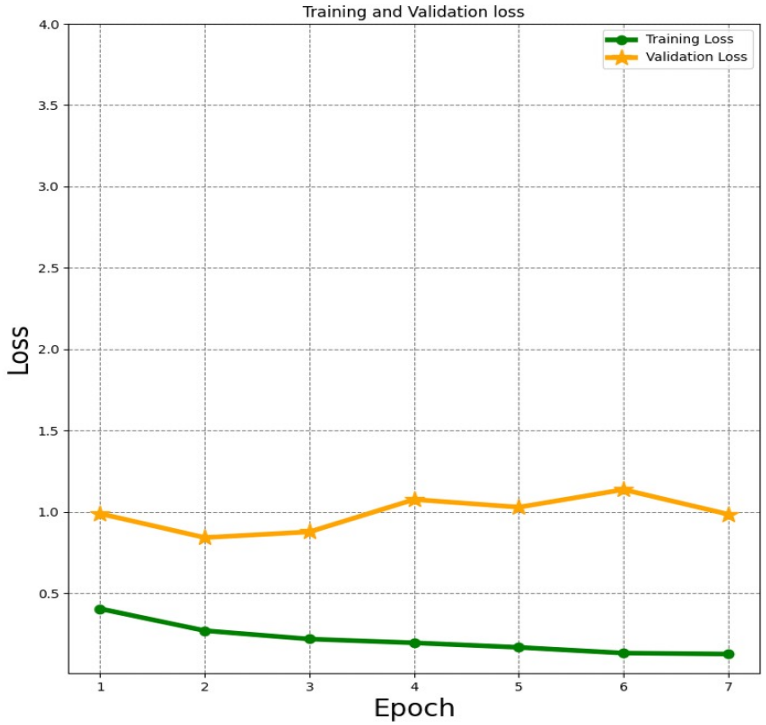




COMPARISION:

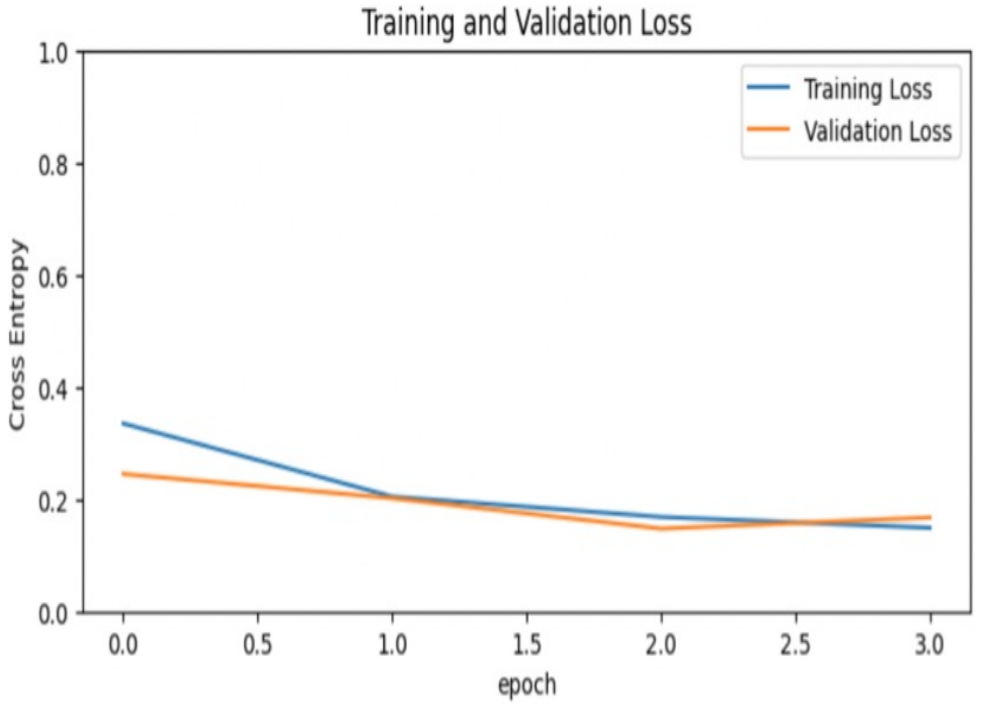
CNN

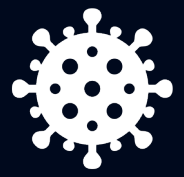
Training and validation loss



Convolutional Base
(DenseNet121)

Training and validation loss





THANK YOU

QUESTIONS ARE WELCOME