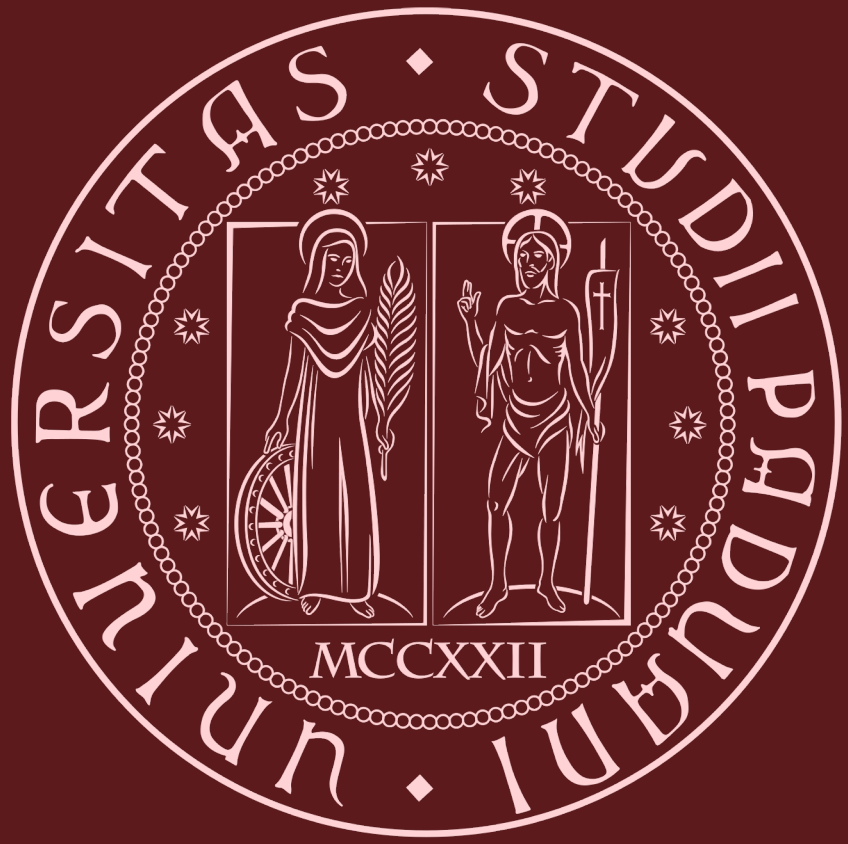
TRANSFER LEARNING APPROACH FOR DETECTING COVID-19 AND PNEUMONIA FROM CHEST X-RAY IMAGES

Fatıma Rabia YAPICIOĞLU – 2049536

HUMAN DATA ANALYTICS COURSE

UNIVERSITY OF PADOVA



CONTENTS

INTRODUCTION

Could we have saved many lives with a correct diagnosis in the early stage?

- The COVID-19 disease outbreak result in many people have severe respiratory problems and it's now been two years since the World Health Organization declared a Covid-19 pandemic.
- The appearance of numerous mutations with increasing extremely infectious levels has hampered the possibility of finally eradicating this pandemic.
- The RT-PCR testing has a high false-negative rate, as well as a long time to obtain results and kits that are inconvenient in some areas.

LITERATURE

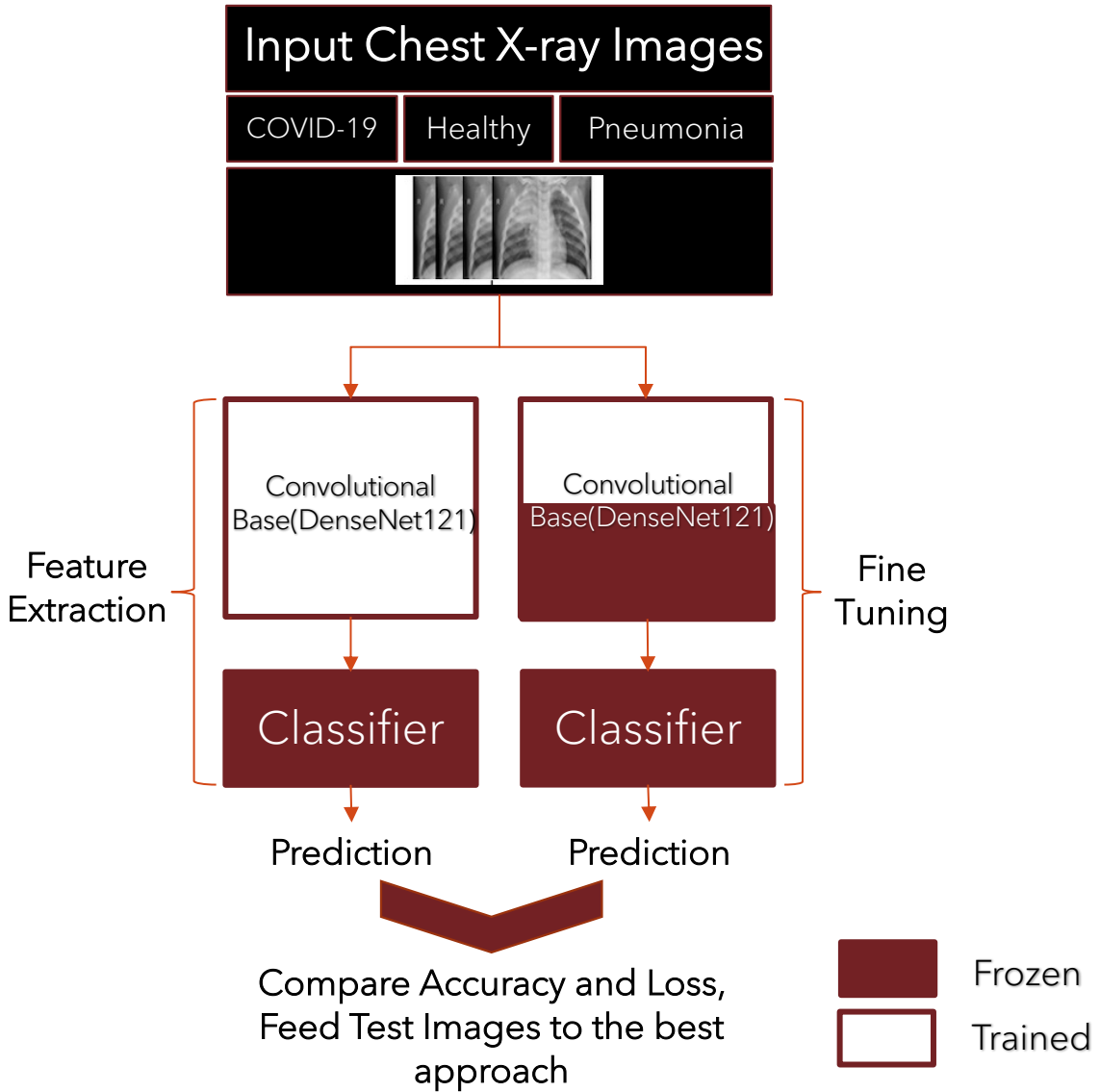
- Since the beginning of pandemics researchers spending great effort on the development and evolution of systems which are helpful in forecasting, detection, organizing of pandemics.

1	Chakraborty S, Paul S, Hasan KMA. **A Transfer Learning-Based Approach with Deep CNN for COVID-19- and Pneumonia-Affected Chest X-ray Image Classification**. SN Comput Sci. 2022;3(1):17. doi: 10.1007/s42979-021-00881-5. Epub 2021 Oct 26. PMID: 34723208; PMCID: PMC8547126.	VGG-19 Architecture, classification between COVID-19, Healthy, and Pneumonia. Larger data with MongoDb.
2	J L G, Abraham B, M S S, Nair MS. A computer-aided diagnosis system for the classification of COVID-19 and non-COVID-19 pneumonia on chest X-ray images by integrating CNN with sparse autoencoder and feed forward neural network. Comput Biol Med. 2022 Feb;141:105134. doi: 10.1016/j.compbiomed.2021.105134. Epub 2021 Dec 14. PMID: 34971978; PMCID: PMC8668604.	COVID-19 and Non-COVID-19 Pneumonia, Sparse Autoencoder and CNN architecture for multiclass classification.
3	Daniel Arias-Garzón, Jesús Alejandro Alzate-Grisales, Simon Orozco-Arias, Harold Brayan Arteaga-Arteaga, Mario Alejandro Bravo-Ortiz, Alejandro Mora-Rubio, Jose Manuel Saborit-Torres, Joaquim Ángel Montell Serrano, Maria de la Iglesia Vayá, Oscar Cardona-Morales, Reinel Tabares-Soto, COVID-19 detection in X-ray images using convolutional neural networks, Machine Learning with Applications,Volume 6,2021,100138,ISSN 2666-8270,	VGG19 and U-Net to process these images and classify them as positive or negative for COVID-19.
4	Kaur P, Harnal S, Tiwari R, Alharithi FS, Almulihi AH, Noya ID, Goyal N. A Hybrid Convolutional Neural Network Model for Diagnosis of COVID-19 Using Chest X-ray Images. *International Journal of Environmental Research and Public Health*. 2021; 18(22):12191. https://doi.org/10.3390/ijerph182212191	Extracts deep learning (DL) features by applying the InceptionV4 architecture and Multiclass SVM classifier to classify and detect COVID-19 infection into four different classes. Bacterial or Viral Pneumonia as two additional classes.

OBJECTIVE

- This study will provide a comparison between two ways of customizing a pre-trained model using transfer learning approach by fine-tuning and feature extraction, thus will show an effective way of using pre-trained models for classifying between COVID-19, Pneumonia, and Healthy images.
- Test data will be feed into the best neural network training approach and then the evaluation metrics will be produced according to it.

- Thus, the project will discover an efficient and quick way to train a pre-trained model which is able to make multiclass classification.



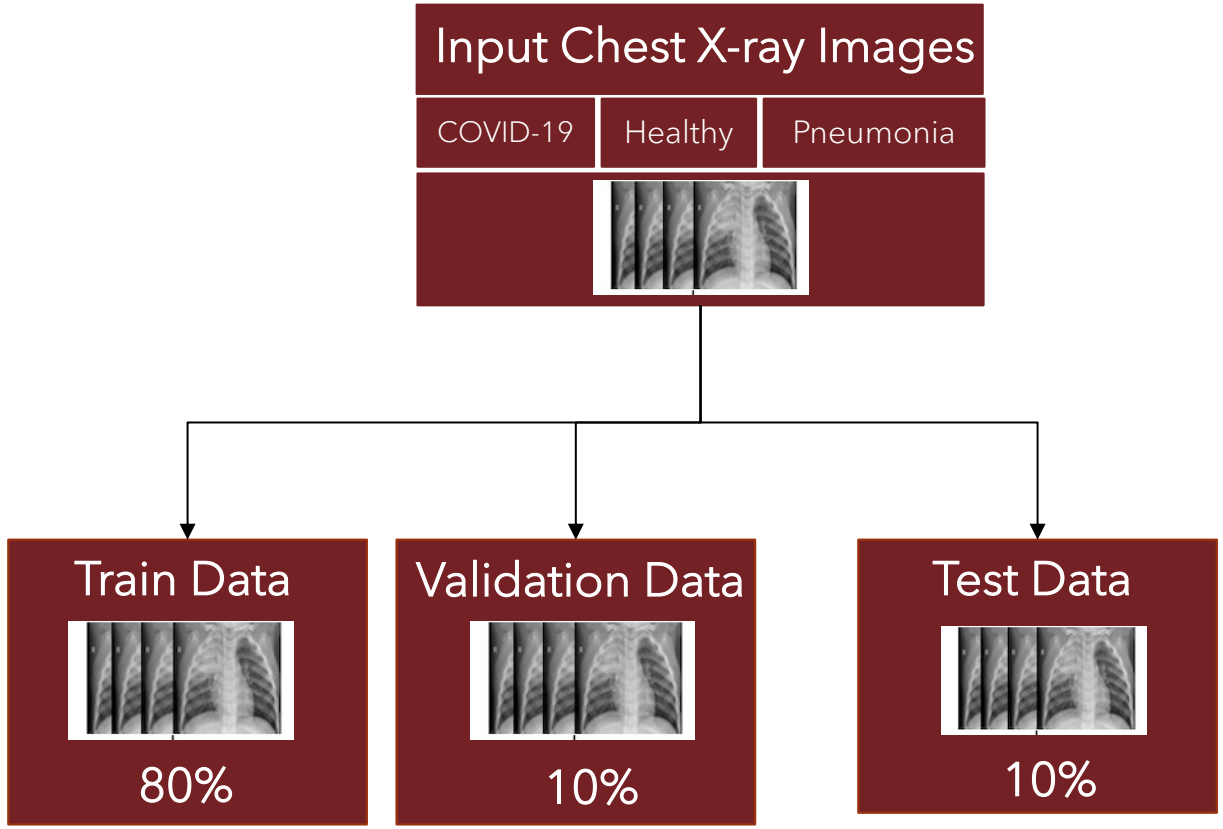
DATASET DESCRIPTION

COVID19, Pneumonia and Normal Chest X-ray PA Dataset

The dataset is organized into 3 folders (covid, pneumonia , normal) and metadata.csv which contain chest X-ray posteroanterior (PA) images. X-ray samples of COVID-19 were retrieved from different sources for the unavailability of a large specific dataset. The dataset consist of 1525 chest x-ray images per each class which sums up to 4575 in total. The dataset satisfies the great expectation of neural networks from the perspective of keeping balance in the dataset among different classes.

Institution: Khulna University of Engineering and Technology

Source: Asraf, Amanullah; Islam, Zabirul (2021), aˆ CœCOVID19, Pneumonia and Normal Chest X-ray PA Datasetaˆ C, Mendeley Data, V1, doi: 10.17632/jctsfj2sfn.1



MATERIALS AND METHODS

- Materials for storing data and coding can be listed as:

  - As a dataset, COVID19, **Pneumonia and Normal Chest X-ray PA Datas**et by the Khulna University of Engineering and Technology has been utilized. Download and store it in the **Google Drive.**
  - For model building, testing and evaluation (coding), the **Google Colab** environment has been used. The libraries; Tensorflow, Numpy, Pandas, Matplotlib, Seaborn, Keras Image Data Generator, Keras Callbacks Model Check Point, Keras Applications Densenet121, OpenCV.
  - In order to build **Demo** software desktop application **PyQT5** Tools have been used.

MATERIALS AND METHODS

Dataset Preparation
Fine-Tuning
Feature Extraction
Trained
Frozen

Start

Input Chest X-ray Images
Pneumonia
Healthy
COVID-19

(224 x 224 x 3)
Data Augmentation

(32x 224 x 224 x 3)
keras.applications preprocess input

The input pixels values are scaled between 0 and 1

Validation Data 10%
Training Data 80%
Test Data 10%

(32 x 224 x 224 x 3)

Base Model DenseNet121
100
(32 x 7x 7x 1024)
GlobalAveragePooling2D (32 x 1024)
Prediction Layer With 3 Nodes (32 x 3)
Softmax
$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$
Covid
Healthy
Pneumonia

Base Model DenseNet121
(32 x 7x 7x 1024)
GlobalAveragePooling2D (32 x 1024)
Prediction Layer With 3 Nodes (32 x 3)
Softmax
$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$
Covid
Healthy
Pneumonia

Probabilities for each class

Finish

DEMO App

Precision, Recall, F1-Score

	precision	recall	f1-score	support
0	0.81	0.93	0.87	14
1	0.86	0.67	0.75	9
2	1.00	1.00	1.00	9
accuracy			0.88	32
macro avg	0.89	0.87	0.87	32
weighted avg	0.88	0.88	0.87	32

Grad-CAM

Confusion Matrix

Training and Validation Accuracy & Loss

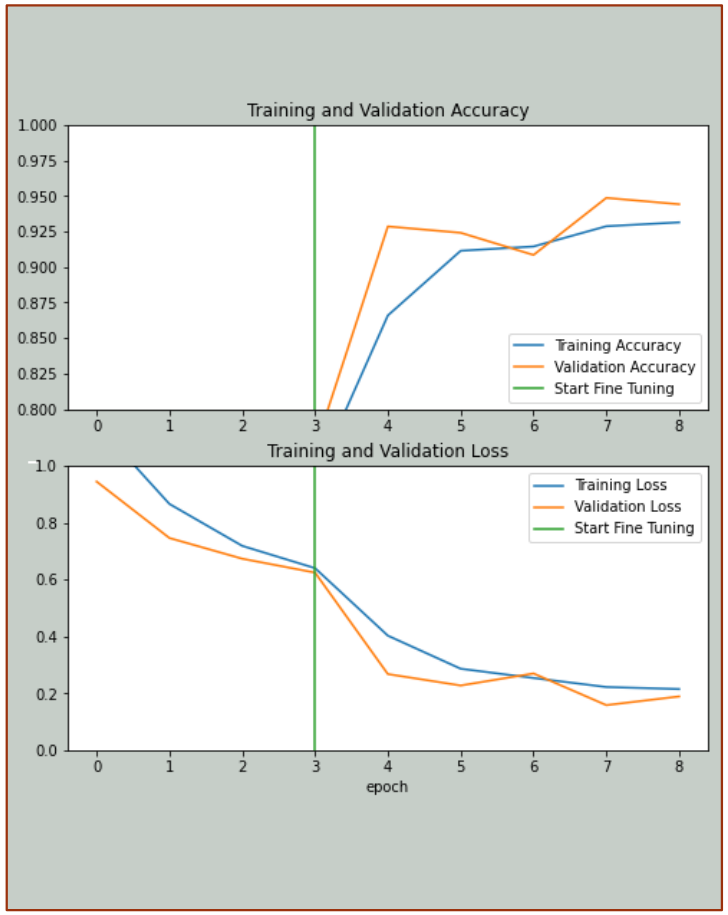
Single Super Test Method

Predicted Diagnosis is: covid
covid % 54.676456
normal % 23.009481
pneumonia % 22.314072

RESULTS

1- Training & Validation Accuracy & Loss

During the implementation of Feature Extraction after 4 epochs, the model achieved 77.23% test accuracy. Then additional 4 epochs has been applied to train the model after fine-tuning approach adjustments, the model validation accuracy reached up to 94.42%.



2 – Test Accuracy

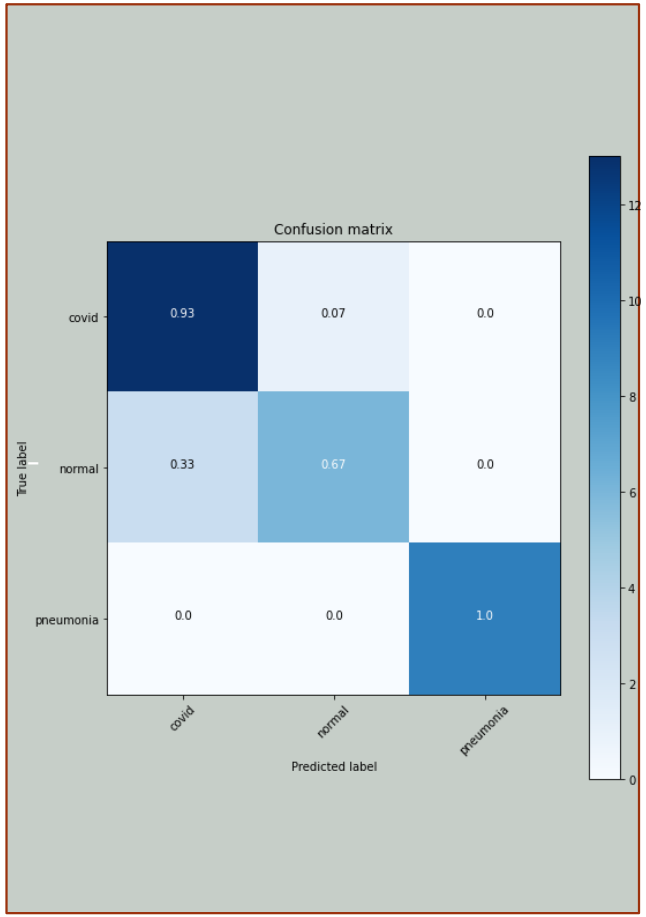
Test dataset partitioned from the whole dataset as 10% consisting 480 images. Meaning that unseen 480 images for the model to test. The model that we created achieved Accuracy rate of 94.58%.

RESULTS

3- Confusion Matrix

Accuracy is a misleading metric!
Accuracy, the proportion of correct classifications among all classifications, is very simple and very "intuitive" measure, yet it may be a poor measure for imbalanced data.

We should have a look at the Precision, Recall, F1-Score and Confusion Matrix results.



Precision	What proportion of positive identifications was actually correct?
Recall	What proportion of actual positives was identified correctly?

4 – Precision, Recall, and F1-Score

Recall would be a better metric because we don't want to accidentally discharge an infected person and let them mix with the healthy population thereby spreading the contagious virus.
The model has difficulties identifying the actual healthy images, the proportion of correctly identified actual healthy chest x-rays (Recall) is smaller when compared to others.

Class	Precision	Recall	F1-Score
COVID-19	0.81	0.93	0.87
Normal	0.86	0.67	0.75
Pneumonia	1.00	1.00	1.00
Average	0.89	0.87	0.87

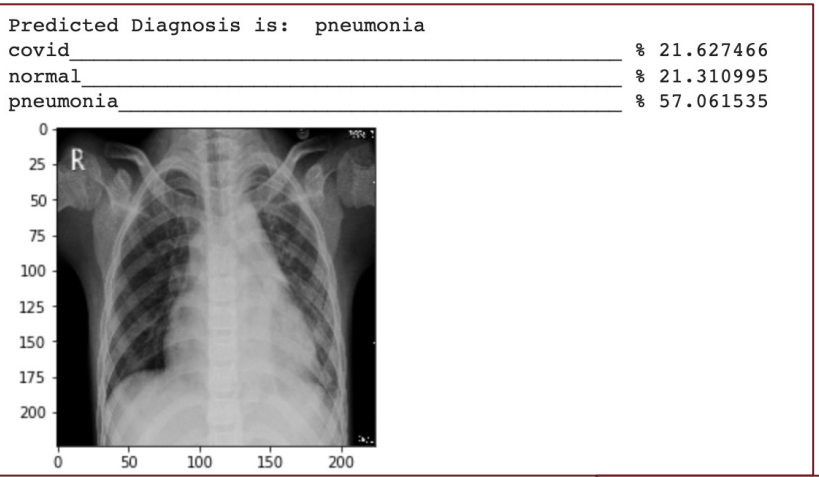
TABLE 1: Class, Precision, Recall, F1-Score,

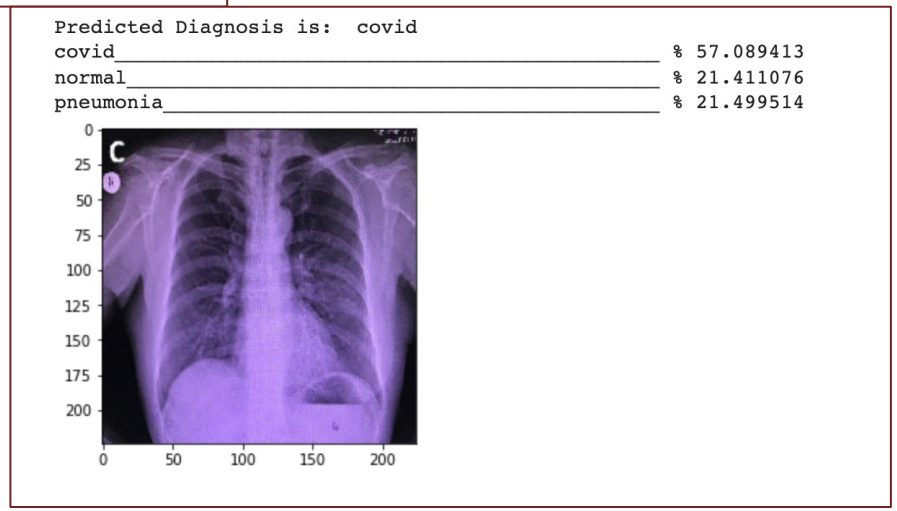
RESULTS

5-Probability Measures
Single Super Test Method

Since we utilized the softmax activation function after the final dense layer, it is possible to get logits, and produce the probabilities belong to each class during the decision mechanism of our model.

This method can take a path to any image and produces probabilities for each class belong to it.

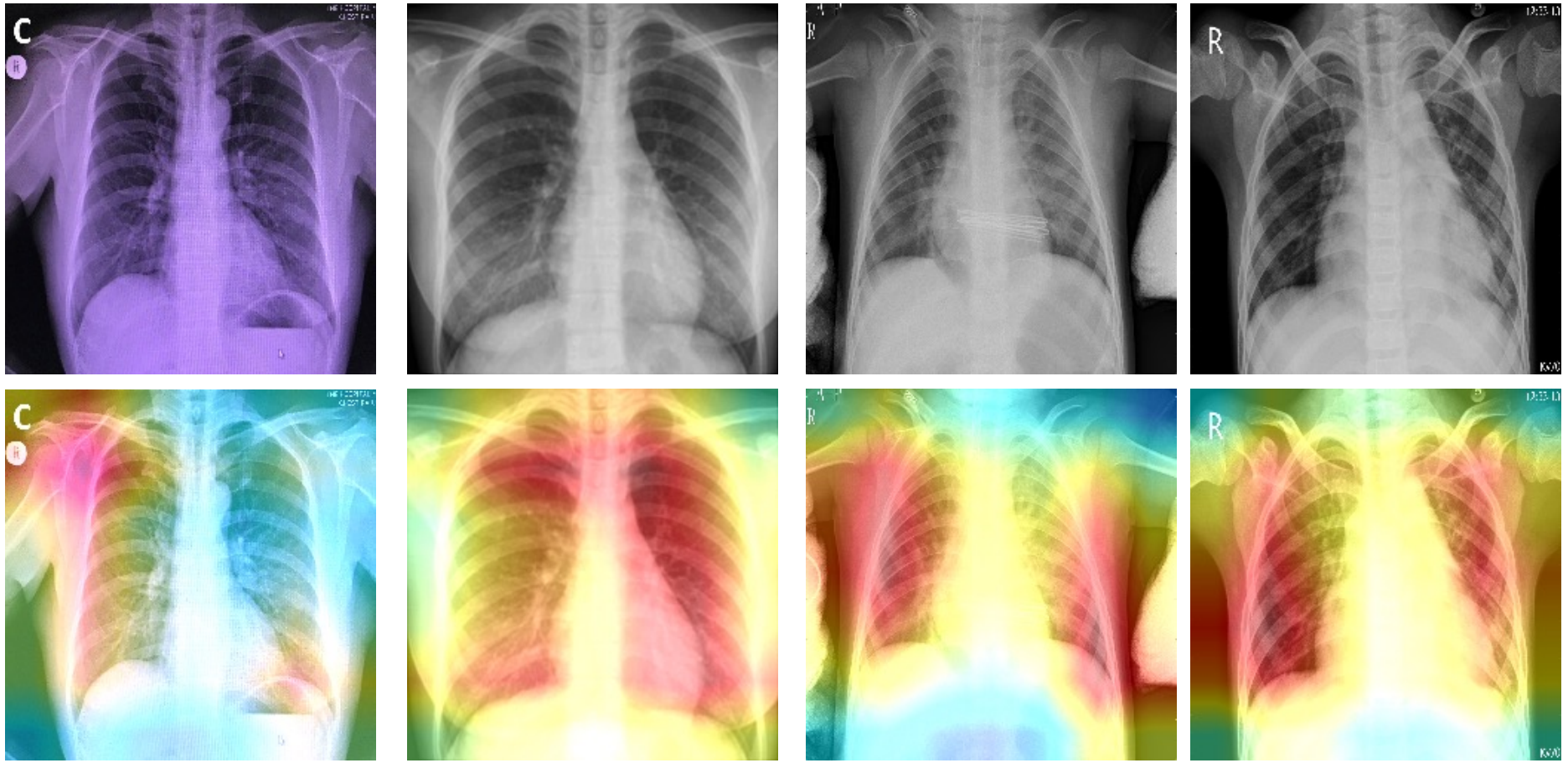




GRAD-CAM VISUALIZATION

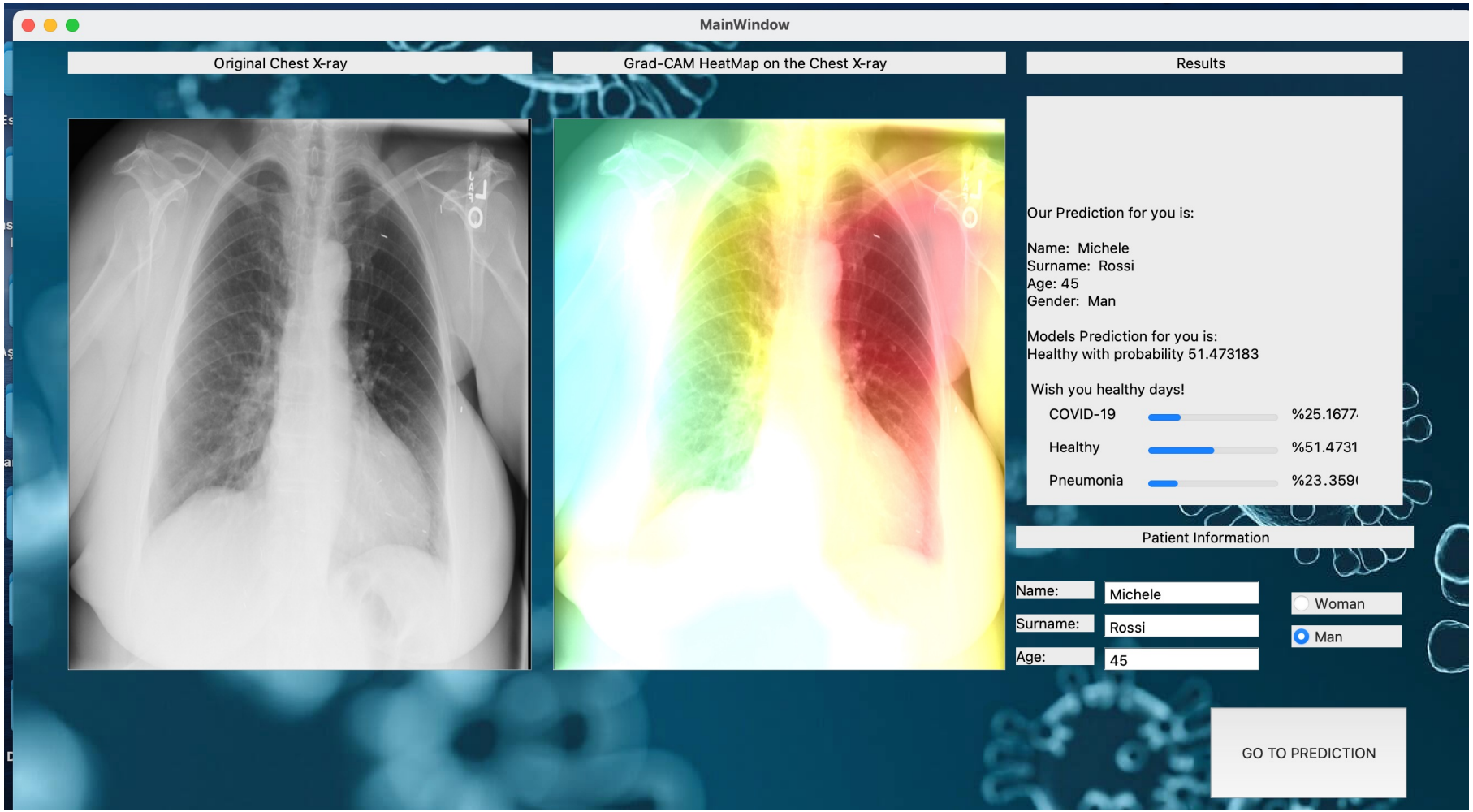
The Grad-CAM technique has been used to visualize the regions of input that are "essential" for predictions from these models or visual explanations using COVID-19 and Pneumonia chest X-ray.

COVID-19. COVID-19 PNEUMONIA PNEUMONIA

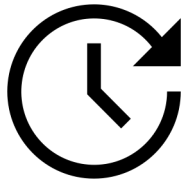


DEMO APPLICATION

The demo application has been created by using PyQT5 and Python.
It is a desktop application which is able to make predictions with probabilities between COVID-19, Healthy, Pneumonia.



CONCLUSIONS & FUTURE WORK



- Fine-tuning approach performance on test data is better than using existing representations on the pre-trained network in Feature Extraction approach.
- Training dataset can be improved with more data samples in order to get more accurate results.
- Other base models can be used such as Xception, Inception, and then the performances can be compared.
- Base models can be combined with different network architectures such as autoencoders.

REFERENCES

- 1. Asraf, Amanullah; Islam, Zabirul (2021), aˆ CœCOVID19, Pneumonia and Normal Chest X-ray PA Datasetaˆ C, Mendeley Data, V1, doi: 10.17632/jctsfj2sfn.1
- 2. Chakraborty S, Paul S, Hasan KMA. A Transfer Learning-Based Approach with Deep CNN for COVID-19- and Pneumonia-Affected Chest X-ray Image Classification.
- SN Comput Sci. 2022;3(1):17. doi: 10.1007/s42979-021- 00881-5. Epub 2021 Oct 26. PMID: 34723208; PMCID: PMC8547126.
- 3. Mohammad Rahimzadeh, Abolfazl Attar, A modified deep convolutional neural network for detecting COVID- 19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2, Informatics in Medicine Unlocked, Volume 19,2020, 100360, ISSN 2352- 9148.https://doi.org/10.1016/j.imu.2020.100360.
- 4. Ali Abbasian Ardakani, Alireza Rajabzadeh Kanafi, U. Rajendra Acharya, Nazanin Khadem, Afshin Mohammadi, Application of deep learning technique to manage COVID- 19 in routine clinical practice using CT images: Results of 10 convolutional neural networks, Computers in Biology and Medicine, Volume 121, 2020, 103795, ISSN 0010-4825.
- 5. Ferhat Ucar, Deniz Korkmaz, COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus dis- ease 2019 (COVID-19) from X-ray images, Medical Hy- potheses, Volume 140, 2020, 109761, ISSN 0306-9877, https://doi.org/10.1016/j.mehy.2020.109761.
- 6. Bejoy Abraham, Madhu S. Nair, Computer-aided detec- tion of COVID-19 from X-ray images using multi-CNN and Bayesnet classifier, Biocybernetics and Biomedical Engineer- ing, Volume 40, Issue 4, 2020, Pages 1436-1445, ISSN 0208- 5216, https://doi.org/10.1016/j.bbe.2020.08.005.
- 7. Liqa A. Rousan, Eyhab Elobeid, Musaab Karrar, Yousef Khader, Chest X-ray findings and temporal lung changes in patients with covid-19 pneumonia, BMC Pulm. Med. 20 (1) (2020) 1aˆ C"9. https://bmcpulmmed.biomedcentral.com/articles/10.1186/s12890- 020-01286-5.
- 8. Manop Phankokkruad. 2020. COVID-19 Pneumonia Detection in Chest X-ray Images Using Transfer Learn- ing of Convolutional Neural Networks. In Proceedings of the 3rd International Conference on Data Science and In- formation Technology (DSIT 2020). Association for Com- puting Machinery, New York, NY, USA, 147aˆ C"152. DOI:https://doi.org/10.1145/3414274.3414496.
- 9. Gu J, Yang L, Li T, Liu Y, Zhang J, Ning K, Su D. Temporal relationship between serial RT-PCR results and serial chest CT imaging, and serial CT changes in coron- avirus 2019 (COVID-19) pneumonia: a descriptive study of 155 case in China. Eur Radiol. 2021 Mar;31(3):1175-1184. doi: 10.1007/s00330-020-07268-9. Epub 2020 Sep 15. PMID: 32930834; PMCID: PMC7491042.

THANK YOU

QUESTIONS ARE WELCOME.