# 288434be2b07962957c3d79d75e7b33a.jpg

North South University

CSE299

Section: 3

**Project Title:**

**Covid-19 Detection with image classification**

**Group Name:**

The Enthusiast

**Submitted to:**

Tanjila Farah(TNF)

**Prepared & Submitted by:**

**Name :** Syed Ragib Israq **ID :** 1620873042

**Name :** Taufiqur Rahman Anik **ID :** 1711474042

**ACKNOWLEDGMENT**

First of all, we wish to express our gratitude to the Almighty for giving us the strength to perform our responsibilities and complete the report. The junior project program is essential to bridge the gap between the theoretical knowledge and real life experience as part of Bachelor of Science (BSc) program. This report has been designed to have a practical experience through the theoretical understanding.

We also acknowledge our profound sense of gratitude to all the teachers who have been instrumental for providing us the technical knowledge and moral support to complete the project with full understanding.

It is imperative to show our appreciation for our honourable faculty member **Ms.Tanjila Farah** for her cooperation and relentless help to achieve this milestone. Also, our gratefulness is divine to the North South University, ECE department for providing us a course such as CSE 299 in which we could really work on this project and materialize it the way we have dreamt of. This course helped us to sort out all our mistakes and enhanced our acknowledgment.

We thank our family, faculty members and entourage for their moral support to carve out this project and always offer their support. Without our faculty advisor **Ms. Tanjila Farah**, we would not implement such wonderful project as he helped us to make it better in every possible way.

# Abstract

Currently, the detection of coronavirus disease 2019 (COVID-19) is one of the main challenges in the world, given the rapid spread of the disease. Recent statistics indicate that the number of people diagnosed with COVID-19 is increasing exponentially, with more than 1.6 million confirmed cases; the disease is spreading too many countries across the world. In this study, we analyse the incidence of COVID-19 distribution across the world. We present an artificial-intelligence technique based on a deep convolutional neural network (CNN) to detect COVID19 patients using real-world datasets. Our system examines chest X-ray images to identify such patients. Our findings indicate that such an analysis is valuable in COVID-19 diagnosis as X-rays are conveniently available quickly and at low costs. Empirical findings obtained from 1000 X-ray images of real patients confirmed that our proposed system is useful in detecting COVID-19 and achieves an F-measure range of 95–99%. Additionally, three forecasting methods—the prophet algorithm (PA), autoregressive integrated moving average (ARIMA) model, and long short-term memory neural network (LSTM)—were adopted to predict the numbers of COVID-19 confirmations, recoveries, and deaths over the next 7 days. The prediction results exhibit promising performance and offer an average accuracy of 94.80% and 88.43% in Australia and Jordan, respectively. Our proposed system can significantly help identify the most infected cities, and it has revealed that coastal areas are heavily impacted by the COVID-19 spread as the number of cases is significantly higher in those areas than in non-coastal areas. Keywords: Artificial Intelligence, X-ray, Convolutional Neural Network, Machine Learning, COVID-19.

Table of Contents

[*1*](#_Toc61464965)

[*Abstract 3*](#_Toc61464966)

[*Introduction 5*](#_Toc61464967)

[*Contributions of This Study 7*](#_Toc61464968)

[*I. Related Works 7*](#_Toc61464969)

[*COVID-19 Diagnosis Using Deep Learning: 8*](#_Toc61464970)

[*X-Ray Diagnosis Using Deep Learning: 8*](#_Toc61464971)

[*CT Scan Diagnosis Using Deep Learning: 9*](#_Toc61464972)

[*COVID-19 Infection Prediction Using Machine Learning 10*](#_Toc61464973)

[*II. System Design 10*](#_Toc61464974)

[*Dataset: 10*](#_Toc61464975)

[*Corona Virus statistic: 12*](#_Toc61464976)

[*Experimental Result: 13*](#_Toc61464977)

[*Environment setup 14*](#_Toc61464978)

[*Hardware: 14*](#_Toc61464979)

[*Software: 15*](#_Toc61464980)

[*Timeline 15*](#_Toc61464981)

[*Cost Estimation 15*](#_Toc61464982)

[*Expected Problem and Its Problem Solution 15*](#_Toc61464983)

[*Discussion: 16*](#_Toc61464984)

[*Conclusions 17*](#_Toc61464985)

[*Project Source code: https://github.com/Israq/Covid19-Detector 18*](#_Toc61464986)

[*Reference 19*](#_Toc61464987)

# Introduction

The coronavirus disease (COVID-19) is a global pandemic that was discovered by a Chinese physician in Wuhan, the capital city of Hubei province in mainland China, in December 2019 [1]. Currently, there is no approved human vaccine for combating it. COVID-19 propagation is faster when people are in close proximity. Thus, travel restrictions control the spread of the disease, and frequent hand washing is always recommended to prevent potential viral infections. Meanwhile, fever and cough are the most common infection symptoms. Other symptoms may occur, including chest discomfort, sputum development, and a sore throat. COVID19 may progress to viral pneumonia which has a 5.8% mortality risk. The death rate of COVID-19 is equivalent to 5% of the death rate of the 1918 Spanish flu pandemic. The total number of people infected with COVID-19 worldwide is 5,790,103 as of May 27, 2020 whereas the numbers of reported deaths and recoveries are 357,432 and 2,497,618 respectively. Most of the cases were recorded in the USA, Spain, Italy, France, and Germany, mainland China, UK, and Iran [2]. Saudi Arabia, with 78,541 cases, has the highest number of reported cases among all the Arab countries. Meanwhile, the number of reported cases in Jordan is 720, whereas the numbers of deaths and recoveries are 9 and 586 respectively. The number of reported cases in Australia is 7150, whereas the numbers of deaths and recoveries are 103 and 6579, respectively. Since February 2020, information technology services, such as mobile apps, have been used to curb the potential risk of infection in mainland China. The mobile apps suggest users to self-quarantine and alert the concerned health authorities when someone infected by the virus. They also monitor infected people, and the last persons that they had contact with [3]. Since it was first reported, the disease has spread exponentially across the world and has become an international concern. A research conducted by Jiang et al. [4] revealed that the death rate of COVID-19 is 4.5% across the world. The death rate of patients in the age range of 70–79 years is 8.0%, whereas that of patients above 80 years is14.8%. The authors also confirmed that patients above the age of 50 years with chronic illnesses are at the highest risk and should therefore take special precautions. One of the main threats of COVID-19 is its rapid propagation, with an estimated 1.5–3.5 people getting infected by the disease upon contact with an infected person [5]. This implies that if 10 people are COVID-19 positive, they are more likely to infect 15–35 other people. Therefore, COVID-19 can infect a very large number of people in a few days unless intervention measures are implemented. The standard diagnostic technique is the reverse transcription-polymerase chain reaction (RT-PCR) method [1], a laboratory procedure that interacts with other ribonucleic (RNA) and deoxyribonucleic acids (DNA) to determine the volume of specific ribonucleic acids using fluorescence. RT-PCR tests are performed on clinical research samples of nasal secretions. The samples are collected by inserting a swab into the nostril and gently moving it into the nasopharynx to collect secretions. Although RT-PCR can identify the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strain that causes COVID-19, in some cases, it produced negative test results even though the patients showed progression on follow-up chest computed tomography (CT) scans [6]. In fact, several studies [6-9] have recommend the use of CT scans and X-rays rather than RT-PCR owing to its limited availability in some countries. The detection of COVID-19 symptoms in the lower parts of the lungs has a higher accuracy when using CT scans or X-rays than that when using RT-PCR [7]. In certain cases, CT scans and X-ray tests can be substituted with RT-PCR tests. However, they cannot exclusively address the problem owing to the relatively limited number of radiologists, compared to new residents, and the high volume of re-examinations of infected people who wish to know the progression of their illness. To overcome the challenges of CT scans and X-rays and to assist radiologists, we need to improve the speed of the procedure. This can be achieved by designing advanced diagnostic systems that utilise artificial intelligence (AI) tools. The aim is to reduce the time and effort required to perform CT scans and X-rays of COVID-19-positive patients and evaluate the rate of disease development [7-9]. Radiological imaging is considered an important screening method for COVID-19 diagnosis [10]. Ai et al. [6] demonstrated the consistency of the radiological history of COVID-19-related pneumonia with the clinical nature of the disease. When examined by CT scans, almost all COVID-19 patients have exhibited similar features including groundglass opacities in the early stages and pulmonary consolidation in the latter stages. In fact, the morphology and peripheral lung distribution can be rounded [6]. AI can be used to initially evaluate a COVID-19 patient as an alternative solution to traditional approaches that are time-consuming and labour-intensive. In this paper, we advocate the use of AI to forecast COVID-19 cases and diagnose COVID-19 patients via chest X-ray images.

## **Contributions of This Study**

A. Contributions of This Study The following are the core contributions of this study:

* We propose an automated intelligent system for distinguishing COVID-19 patients from non-patients on the basis of chest X-ray images. Our system instantly reads the structure of a chest X-ray image, leverages hidden patterns to identify COVID-19 patients, and reduces the need for manual pre-processing steps.
* Empirical findings obtained from 1000 chest X-ray images of patients confirmed that our proposed system can detect COVID-19 patients with an accuracy of 95–99%.
* We provide an intelligent prediction system for predicting the number of patients confirmed to have contracted the disease, recovered from the disease, and died from the disease over the next 7 days using three forecasting methods. Our proposed system has been trained and tested on datasets generated from real-world cases and has predicted the numbers of COVID-19 confirmations, recoveries, and deaths in Australia and Jordan with an average accuracy of 94.80% and 88.43%, respectively.
* We highlight the most affected areas and show that coastal areas are heavily impacted by COVID-19 infection and spread as the number of cases in those areas is significantly higher than that in other non-coastal areas. The rest of this paper is organised as follows. Section 2 presents the related works on recent COVID-19 detection and prediction methods for chest X-ray images. Section 3 presents the detailed system design, dataset description, and performance-evaluation metrics. Sections 4 and 5 present the results and discussions, respectively. Section 6 concludes the paper and provides an outlook to future research.

## **I. Related Works**

The analysis and detection of COVID-19 have been extensively investigated in the last few months. The first part of this section addresses issues related to COVID-19 detection based on deep-learning approaches using CT scans and chest X-ray images. The second part reviews the related literatures to assess future estimates of the number of COVID-19 confirmations, recoveries, and deaths. COVID-19 has now become a global pandemic owing to its rapid spread. It is very challenging to detect exposed persons because they do not show disease symptoms immediately. Thus, it is necessary to find a method of estimating the number of potentially infected persons on a regular basis to adopt the appropriate measures. AI can be used to examine a person for COVID-19 as an alternative to traditional time-consuming and expensive methods. Although there are several studies on COVID-19, this study focused on the use of AI in forecasting COVID-19 cases and diagnosing patients for COVID-19 infection through chest X-ray images. Several research areas have implemented AI (e.g. disease diagnoses in healthcare) [11-13]. One of the main advantages of AI is that it can be implemented in a trained model to classify unseen images. In this study, AI was implemented to detect whether a patient is positive for COVID-19 using their chest X-ray image. AI can also be used for forecasting (e.g., how the population will increase over the next 5 years) through existing evidence. Thus, predicting possibilities in the immediate future can help authorities to adopt the necessary measures [14]. Wynants et al. [15] focused on two main concepts. The first concept involved studies related to the diagnosis of COVID-19, and the second involved studies related to the prediction of the number of people who will be infected in the coming days. The study analysis maintained that most of the existing models are poor and biased. The authors suggested that research-based COVID-19 data should be publicly available to encourage the adoption of more specifically designed detection and prediction models

**COVID-19 Diagnosis Using Deep Learning:** The use of machine learning (ML) has been rapidly increasing in various fields including malware detection [16-19], mobile malware detection [20-24], medicine [25-27] and information retrieval [27-31]. In 2012, a modern ML system called deep learning was introduced, which is based on a convolutional neural network (CNN). It won the ImageNet classification competition, the world’s best-known computer-vision competition [32]. Deep-learning algorithms enable computational models composed of multiple processing layers to learn data representation through several abstraction layers. They train a computer model to perform classification tasks directly from pictures, texts, or sounds. According to LeCun et al. [33], deep-learning models feature high accuracies and can improve human output in certain instances.

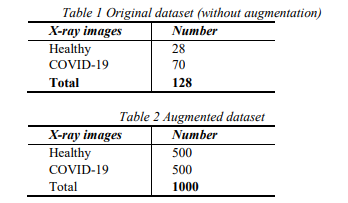
**X-Ray Diagnosis Using Deep Learning:** X-ray machines use light or radio waves as radiation to examine the affected parts of the body because of cancers, lung diseases, bone dislocations, and injuries. Meanwhile, CT scans are used as sophisticated X-ray machines to examine the soft structures of active body parts for better views of the actual soft tissues and organs [34]. The advantages of using X-rays over CT scans are that X-rays are quicker, safer, simpler, and less harmful than CT scans [7, 34]. Narin et al. [7] proposed a CNN-based model to detect COVID-19 patients using 100 chest X-ray images, half of which belong to COVID-19 patients and the other half belong to healthy people. They evaluated three CNN models— ResNet-50, Inception-v3, and Inception-ResNet-v2—using five-fold cross-validation and reported that ResNet-50 had the best detection accuracy (98%). In a similar study conducted by Sethy and Behera [35], the authors extracted features from chest X-ray images using a deep-learning algorithm and classified the images as either infected or healthy using a support vector machine (SVM). The authors employed 11 deep-learning models: AlexNet, VGG16, VGG19, GoogLeNet, ResNet-18, ResNet-50, ResNet-101, Inception-v3, Inception-ResNet-v2, DenseNet201, and XceptionNet. They collected two datasets—the first containing chest X-ray images of 25 infected patients and 25 non-infected patients and the other containing chest X-ray images of 133 infected patients (e.g. MERS, SARS, and ARDS patients) and 133 non-infected patients. They performed separate feature extractions on each dataset using various models and achieved a 95.38% accuracy with ResNet50 and SVM

**CT scan Diagnosis Using Deep Learning:** The CT scan was developed by Godfrey Hounsfield and Allan Cormack in 1972. It utilises an advanced X-ray technology to carefully diagnose delicate internal organs [34]. CT scanning is quick, painless, non-invasive, and precise and can produce three-dimensional (3D) images [34]. CT scans of internal organs, muscles, soft tissues, and blood vessels offer greater clarity than standard X-rays, especially for soft tissues and blood vessels. The main disadvantage of the CT scan is that it is expensive, compared to X-rays [34]. The sensitivity and specificity of RT-PCR for COVID-19 detection have been criticised in several studies [4, 38, 39]. Although RT-PCR is the standard method for this purpose, it generates a relatively large number of false negatives owing to several reasons, including methodological drawbacks, disease stages, and methods of obtaining the specimens, which delay disease diagnosis and control. Therefore, RT-PCR tests are not sufficient for assessing the disease status. Recent results have revealed that nucleic acid testing is not reliable and can only achieve an accuracy of 30–50% [38]. Jiang et al. [4] compared RT-PCR to CT scans and examined 51 patients (29 men and 22 women) with a history of travel to or residency in endemic areas and with severe respiratory and fever symptoms due to unknown causes. The authors obtained a sensitivity of 98% in a non-contrast chest CT scan for the detection of COVID-19, compared to the initial RT-PCR sensitivity of 71%. Owing to the shortage of RT-PCR kits and the growing number of COVID-19 cases, it is important to introduce an automated detection system as an alternative diagnostic method to prevent the spread of COVID-19 [7]

**COVID-19 Infection Prediction Using Machine Learning** Techniques ML is the science of training machines using mathematical models to learn and analyse data. Once ML is implemented in a system, the data are analysed, and interesting patterns are detected. The validation data are then categorised according to the patterns learned during the learning process. As COVID-19 infection has rapidly spread worldwide and international action is required, it is important to develop a strategy to estimate the number of potentially infected people on a regular basis to adopt the appropriate measures. Currently, decision-makers rely on certain decision-making statistics such as imposing lockdowns on infected cities or countries. Therefore, ML can be used to predict the behaviours of new cases to stop the disease from spreading.

## **II. System Design**

Our proposed deep learning-based COVID-19 detection comprises several phases, as illustrated in Figure 1. The phases are summarised in the following five steps:

Step 1: Collect the chest X-ray images for the dataset from COVID-19 patients and healthy persons.

Step 2: Generate almost 1000 chest X-ray images using data augmentation.

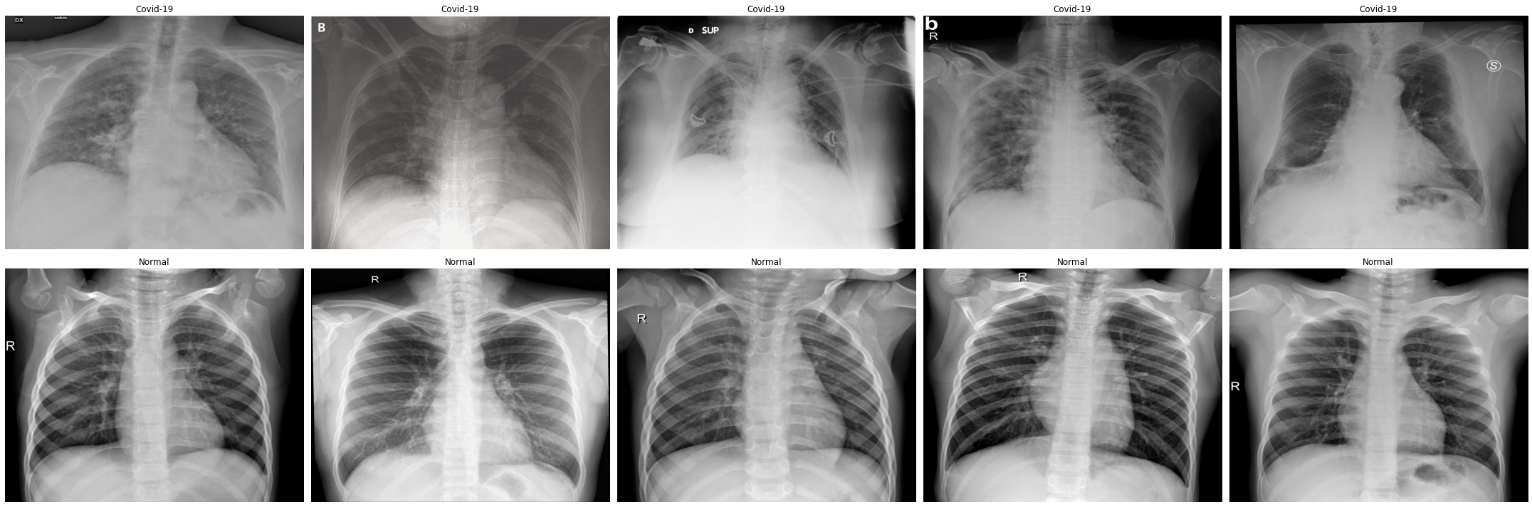
Step 3: Represent the images in a feature space and apply deep learning.

Step 4: Split the dataset into two sets: a training set and a validation set.

Step 5: Evaluate the performance of the detector on the validation dataset.

# **Dataset:**

Two types of datasets were used in the evaluation, the original dataset (without augmentation) and the augmented dataset, which are summarised in Tables 1 and 2, respectively. The dataset contained the following: a) a healthy dataset containing chest X-ray images of healthy persons and b) a COVID-19 dataset containing chest X-ray images of COVID19 patients. The original dataset was obtained from the Kaggle database, and its total number of images is 128.

Figure 1. Chest X-ray images (Normal vs Covid 19)

True-negative (TN), false-positive (FP), and false-negative (FN) scores:

- TP is the proportion of positive COVID-19 chest Xray images that were correctly labelled as positive. - FP is the proportion of negative (healthy) COVID-19 chest X-ray images that were mislabelled as positive.

- TN is the proportion of negative (healthy) chest Xray images that were correctly labelled as healthy. - FN is the proportion of positive COVID-19 chest Xray images that were mislabelled as negative (healthy).

**Accuracy**: This metric measures the percentage of correctly identified cases relative to the entire dataset. The ML algorithm performs better if the accuracy is higher. Accuracy is a significant measure for a test dataset that includes a balanced class. It is computed as follows:

**𝑨𝒄𝒄𝒖𝒓𝒂𝒄𝒚 = (𝑇𝑃 + 𝑇𝑁)/ (𝑇𝑃 + 𝑇𝑁 + 𝐹𝑃 + 𝐹𝑁)**

Precision: This metric is a measure of exactness, which is calculated as the percentage of positive predictions of COVID-19 that were true positives divided by the number of predicted positives. It is computed as follows:

𝑷𝒓𝒆𝒄𝒊𝒔𝒊𝒐𝒏 = 𝑇𝑃/(𝑇𝑃 + 𝐹𝑃) (2)

Recall: This metric is a measure of completeness, which is calculated as the percentage of positives that were correctly identified as true positives divided by the number of actual positives. It is computed as follows:

𝑹𝒆𝒄𝒂𝒍𝒍 = 𝑇𝑃/(𝑇𝑃 + 𝐹𝑁) (3)

F-measure: This is a combination of precision and recall that provides a significant measure for a test dataset that includes an imbalanced class. It is computed as follows:

𝑭 − 𝑴𝒆𝒂𝒔𝒖𝒓𝒆 = 𝟐 ( 𝑃𝑟𝑒𝑐𝑖𝑠𝑖𝑜𝑛 𝑥 𝑅𝑒𝑐𝑎𝑙𝑙/ (𝑃𝑟𝑒𝑐𝑖𝑠𝑖𝑜𝑛 + 𝑅𝑒𝑐𝑎𝑙𝑙) )

**Root Mean Square Error (RMSE):** This metric measures the differences between the actual (𝑥 ) and the predicted ( 𝑥̂𝑖 ) numbers of COVID-19 confirmations, recoveries, and deaths (𝑁). The main advantage of RMSE is that it penalises large prediction errors. RMSE was used to compare the prediction errors of the three prediction algorithms. It is defined as follows:

𝑹𝑴𝑺𝑬 = √ 1 / 𝑁∑ (𝑥𝑖 − 𝑥̂𝑖) 2; 𝑁, 𝑖=1

Correlation coefficient: This metric is often used to evaluate the performance of a prediction algorithm. It is defined as follows:

𝑪𝑪 = (1 – 1 / 𝑁∑|𝑥𝑖 − 𝑥̂𝑖 |; 𝑁, i=1) ∗ 100%

# Corona Virus statistic:

SARS-CoV-2 is a new family of viruses that has never been encountered before. The virus was first discovered in pangolins before its spread to humans [55]. The typical symptoms of COVID-19 include fever, dry cough, fatigue, sputum production, shortness of breath, sore throat, headache, chills, nausea or vomiting, nasal congestion, diarrhoea, haemoptysis, and conjunctival discomfort, although some patients also suffer from general tiredness, runny nose, and loss of taste and/or scent. Figure 2 shows a bar graph of the common COVID-19 symptoms sorted by their percentage of occurrences.

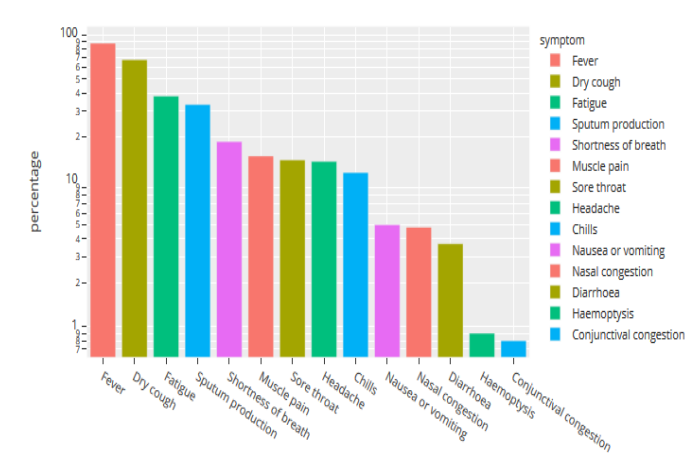


Figure 2. Common COVID-19 symptoms

# Experimental Result:

Firstly, we examined the most infected areas across the world. In Section 4.1, we show that coastal areas are heavily affected by the COVID-19 outbreak as the number of cases in those 𝑨𝒄𝒄𝒖𝒓𝒂𝒄𝒚 = (𝑇𝑃 + 𝑇𝑁)/ (𝑇𝑃 + 𝑇𝑁 + 𝐹𝑃 + 𝐹𝑁) (1) COVID-19 Prediction and Detection Using Deep Learning 173 areas is significantly higher than those in other non-coastal areas.

Figure 3. Areas with the highest number of confirmed cases in Australia

Secondly, we predicted the number of COVID-19 confirmations, recoveries, and deaths in Jordan and Australia over the next 7 days using three well-known time-series forecasting algorithms: PA (prophet algorithm), ARIMA, and LSTM. In Section 4.2, we show how the application of these algorithms allowed us to estimate the forecasting outcomes for certain countries with a detection rate of 99% (Tables 3 and 4) and (Figures. 7 -13). Thirdly, we examined whether chest X-ray images can be used to develop sophisticated classification models for COVID-19 prediction. In Section 4.3, we present the application of a deep-learning algorithm on two datasets. Empirical findings indicated that our proposed system is reliable in detecting COVID-19 and has an F-measure range of 95–99%, as revealed

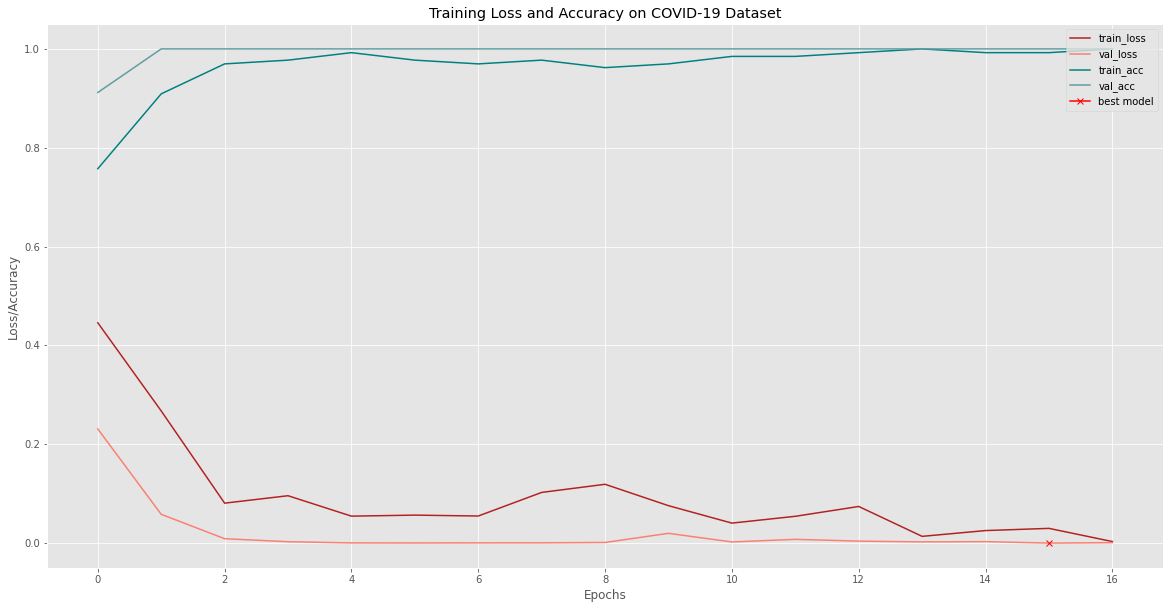


Figure 5: Accuracy Visualization

# Environment setup

## **Hardware:**

For ORB and SVM classiﬁcation, an ordinary high-performance computer is enough,

Like 8 GB ram, i5 (2.3 GHz), and a 256GB solid-state drive (SSD) disk. However,

Training a deep neural network should use GPU to accelerate the process. In this

Report, a Google Cloud GPU is used. A virtual machine instance with four core of

CPU, 8 GB memory, and an NVIDIA GEFORCE is used. Concerning the detail setup

Guide, please refer Google guide and other web pages.

## **Software:**

To test the ORB and SVM classiﬁcation, a python program which was initially used to

Classify plants are ported. It was modiﬁed to use the new dataset and ran it on a

Laptop. An iteration of the test needs about four hours. Because of the CNN-based

Method is computing intensively, so it needs to run on a VM in Google GPU Cloud. To

Test the Capsule network, a python capsule network implementation that aims to detect

Brain tumors was ported to the pneumonia dataset. It also needs to be run on the

GPU VM.

# Timeline

* WEEK (1-3):PROJECT DECISISON
* WEEK 4 :PROJECT PROPOSAL AND LEARNING PROGRESS
* WEEK(5-6):PIPELINE DESIGNING(AUTHENTICATION,PAYMENT METHOD)
* WEEK(7-9):IMPLEMENT TRANSFER LEARNING
* WEEK10 : CREATING WEB APP
* WEEK11 : DEPLOY
* WEEK12 :PROJECT PESENTATION(FINAL)

# Cost Estimation

* TENATIVE

# Expected Problem and Its Problem Solution

During training the model with dataset there may some problem will appear so we’ve to be cautious while training it.Another problem may appear while cutting the final layer of the pertained model so we’ve to determine the final layer. While downloading the boiler plate source code from a known git first the flask boiler plate was not executing first place then the code was modified. Another problem arrived when the transfer learning was executing on google colab notebook there were many bugs arrived regarding to python packages. The package has to be updated to use the latest python libraries. So we’ve installed new keras libraries. Another confusing step is kaggle authentication after few attempts this problem was solved.

# Discussion:

This study provided a forecasting analysis of COVID-19 confirmations, recoveries, and deaths in Australia and Jordan. It further implemented a CNN-based COVID-19 detector to identify COVID-19 infections using X-ray images. Based on the study results, the following conclusions were drawn:

* PA delivered the best performance for COVID-19 prediction over 7 days, compared to LSTM and ARIMA
* The predictions will enable people in both countries to predict their medical needs for tackling the spread of COVID19.
* ARIMA cannot make predictions over the next 1, 2, and 3 days.
* After investigating the number of COVID-19 confirmations, recoveries, and deaths in various countries, we found that coastal areas are significantly impacted by the disease because the numbers of cases in those areas are significantly higher than those in other non-coastal areas. This observation is medically consistent with the propagation capability of viruses in areas with higher humidity rates. Thus, the authors advise healthcare professionals to devote greater attention to coastal regions.
* The use of chest X-ray images is recommended for diagnosing COVID-19 because X-rays are easily obtained at nearby hospitals or clinics fairly quickly and at low costs.
* Our CNN-based COVID-19 detector delivered superior performance in terms of precision, recall, and F-measure.
* The application of ML techniques for COVID-19 diagnosis using our CNN-based COVID-19 detector is recommended.
* It is well known that VGG16 (Wu et al., 2017) outperforms many convolutional networks, such as Google Net and Squeeze Net, and its feature representation capability is beneficial for classification accuracy. Hence, VGG16 is a recommended version of a deep CNN-based algorithm as it makes training easier and quicker. It was implemented in our COVID-19 detector to improve its accuracy in diagnosing COVID-19 in chest X-ray images. Our COVID-19 detector obtained better results when using augmentation. A better training process was achieved as the gap between the training and validation became smaller. Moreover, a more generalized and robust COVID-19 detector was achieved as the F-measure improved from 0.95 to 0.99. Thus, the COVID-19 detector trained on augmented data provides superior performance metrics and is robust for diagnosing COVID-19 in chest X-ray images.

## Conclusions

The rapid spread of COVID-19 across the world and the increasing number of deaths require urgent actions from all sectors. Future prediction of potential infections will enable authorities to tackle the consequences effectively. Furthermore, it is necessary to keep up with the number of infected people by performing regular check-ups, and it is often vital to quarantine infected people and adopt medical measures. Additionally, attention should be given to several other factors to curb the spread of COVID-19, such as the environmental effects and the similarities among the most affected areas, and careful measures should be adopted. In this paper, AI-based techniques were proposed for the prediction and diagnosis of COVID-19:

* Prediction models such as the PA, ARIMA, and LSTM algorithms were used to predict the number of COVID-19 confirmations, recoveries, and deaths over the next 7 days. PA delivered the best performance. It predicted the number of COVID-19 confirmations, recoveries, and deaths in Australia and obtained prediction accuracies of 99.94%, 90.29%, and 94.18%, respectively. It also predicted the number of COVID19 confirmations, recoveries, and deaths in Jordan and obtained prediction accuracies of 99.08%, 79.39%, and 86.82%, respectively. Meanwhile, investigation into more sophisticated forecasting and prediction methods is a subject of a future work.
* A diagnosis model using VGG16 was proposed to detect COVID-19 using chest X-ray images. The model allows the rapid and reliable detection of COVID-19, enabling it to achieve an F-measure of 99% using an augmented dataset. In a future study, we will consider diagnosing COVID-19 in chest CT scan images using the VGG-XX versions and compare their performances using larger datasets. A further contribution of this study is the analysis of the COVID-19 spread and its related statistical data based on its global regional distributions. Thus, two main conclusions were drawn using our AI-based analysis: (i) the most highly infected areas have similar characteristics, and (ii) the spread of the disease in coastal areas is significantly higher than that in other non-coastal areas. Therefore, extra care and attention should be given to coastal cities. In our future work, we will investigate the effects of temperature, humidity, and terrain on the COVID-19 spread in cities and countries.

# **Project Source code: https://github.com/Israq/Covid19-Detector**

## Reference

1. World Health Organization, "Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: interim guidance, 2 March 2020," World Health Organization, World Health Organization2020.
2. Worldometers. (2020, April. 6). Coronavirus Cases. Available: <https://www.worldometers.info/coronavirus/>
3. A. Chen. (2020) China’s coronavirus app could have unintended consequences. MIT Technology Review. Available: [https://www.technologyreview.com/2020/02/13/844 805/coronavirus-china-app-close-contactsurveillance-covid-19-technology/](https://www.technologyreview.com/2020/02/13/844%20805/coronavirus-china-app-close-contactsurveillance-covid-19-technology/)
4. F. Jiang, L. Deng, L. Zhang, Y. Cai, C. W. Cheung, and Z. Xia, "Review of the clinical characteristics of coronavirus disease 2019 (COVID-19)," Journal of General Internal Medicine, pp. 1-5, 2020.
5. Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention," Jama, 2020.
6. T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, et al., "Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," Radiology, p. 200642, 2020.
7. [A. Narin, C. Kaya, and Z. Pamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks," arXiv preprint arXiv:2003.10849, 2020.
8. H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, and M. K. Khan, "Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms," arXiv preprint arXiv:2004.00038, 2020.
9. S. U. K. Bukhari, S. S. K. Bukhari, A. Syed, and S. S. H. SHAH, "The diagnostic evaluation of Convolutional Neural Network (CNN) for the assessment of chest X-ray of patients infected with COVID-19," medRxiv, 2020.
10. H. Shi, X. Han, N. Jiang, Y. Cao, O. Alwalid, J. Gu, et al., "Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study," The Lancet Infectious Diseases, 2020.
11. A. A. Ogunleye and W. Qing-Guo, "XGBoost model for chronic kidney disease diagnosis," IEEE/ACM transactions on computational biology and bioinformatics, 2019.
12. C. Feng, A. Elazab, P. Yang, T. Wang, F. Zhou, H. Hu, et al., "Deep Learning Framework for Alzheimer’s Disease Diagnosis via 3D-CNN and FSBi-LSTM," IEEE Access, vol. 7, pp. 63605-63618, 2019.
13. H. Yin, B. Mukadam, X. Dai, and N. Jha, "DiabDeep: Pervasive Diabetes Diagnosis based on Wearable Medical Sensors and Efficient Neural Networks," IEEE Transactions on Emerging Topics in Computing, 2019.
14. K. Santosh, "AI-Driven Tools for Coronavirus Outbreak: Need of Active Learning and CrossPopulation Train/Test Models on Multitudinal/Multimodal Data," Journal of Medical Systems, vol. 44, pp. 1-5, 2020.
15. L. Wynants, B. Van Calster, M. M. Bonten, G. S. Collins, T. P. Debray, M. De Vos, et al., "Systematic review and critical appraisal of prediction models for diagnosis and prognosis of COVID-19 infection," medRxiv, 2020
16. M. Fu, S.-L. Yi, Y. Zeng, F. Ye, Y. Li, X. Dong, et al., "Deep Learning-Based Recognizing COVID-19 and other Common Infectious Diseases of the Lung by Chest CT Scan Images," medRxiv, 2020.
17. [42] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, et al., "Deep learning system to screen coronavirus disease 2019 pneumonia," arXiv preprint arXiv:2002.09334, 2020.
18. M. Li, Z. Zhang, S. Jiang, Q. Liu, C. Chen, Y. Zhang, et al., "Predicting the epidemic trend of COVID-19 in China and across the world using the machine learning approach," medRxiv, 2020.
19. P. Kumar, H. Kalita, S. Patairiya, Y. D. Sharma, C. Nanda, M. Rani, et al., "Forecasting the dynamics of COVID-19 Pandemic in Top 15 countries in April 2020 through ARIMA Model with Machine Learning Approach," medRxiv, 2020