7CS041 - MSC PROJECT DATA SCIENCE

School of Mathematics and Computer
ScienceFaculty of Science and
Engineering

Deep learning assisted framework for chest infection detection

Mian Muhammad Umer Hassan Khan Qureshi

2209340

20/04/2023

1. Abstract

Chest X-ray analysis is an essential tool for diagnosing and monitoring various respiratory diseases, such COVID-19 and pneumonia. However, interpreting chest X-rays can be challenging, even for experienced physicians, and misdiagnosis can have severe consequences for patients. In this project, I will propose a deep learning-powered algorithm for chest disease identification, using a multi-class classification model to distinguish between four different disease categories.I will collect and pre-process a dataset of chest X-rays with annotations indicating the presence or absence of each disease. I will then train a deep convolutional neural network (CNN) on this dataset and optimize its hyperparameters to achieve the best performance. I will evaluate the model's accuracy, precision test set and generate a confusion matrix to understand its strengths and weaknesses in predicting each disease category. Finally, I will deploy the model as mobile application that allows users to upload their own chest X-rays and get a prediction of their disease status. I will build a user interface that is user-friendly and provides clear explanations of the model's outputs and limitations. This project aims to provide an accurate, efficient, and accessible tool for chest disease diagnosis, which can help physicians make more informed decisions and improve patient outcomes.

2. Introduction

Respiratory diseases, COVID-19 and pneumonia are a significant cause of morbidity and mortality worldwide. Early detection and accurate diagnosis of these diseases can greatly improve patient outcomes, but it remains a challenging task for healthcare providers. Chest X-rays are a commonly used diagnostic tool for respiratory diseases due to their accessibility and affordability. However, manual interpretation of chest X-rays is time-consuming and subject to human error, which can lead to misdiagnosis or delayed treatment.

In recent years, machine learning techniques, particularly deep learning algorithms such as Convolutional Neural Networks (CNN), have shown promising results in medical image analysis, including chest X-ray analysis for respiratory disease diagnosis. CNNs are a type of deep neural network that can automatically extract relevant features from medical images, enabling accurate classification of images into different disease categories.

In this context, I propose a multi-class classification model for automated diagnosis of respiratory diseases through chest X-rays using CNN algorithms. The model will be trained using a large dataset of chest X-rays and deep learning techniques to classify chest X-rays into four categories: bronchitis, lung cancer, asthma, and pneumonia.

Here are brief introductions for each respiratory disease:

COVID-19:

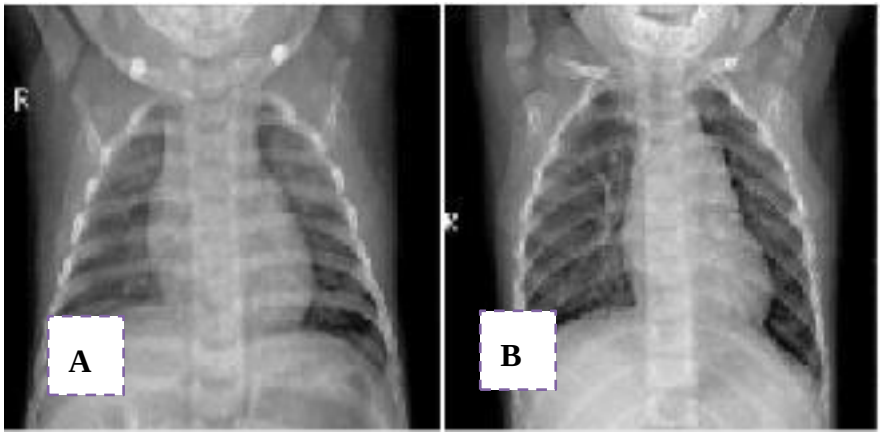
COVID-19 is a contagious respiratory illness caused by the SARS-CoV-2 virus, which was first identified in Wuhan, China, in December 2019. The virus spreads through respiratory droplets when an infected person coughs, sneezes, talks, or breathes, and it can also spread by touching a surface contaminated with the virus and then touching the mouth, nose, or eyes.

Symptoms of COVID-19 can range from mild to severe and include fever, cough, shortness of breath, fatigue, body aches, loss of taste or smell, sore throat, congestion, and diarrhea. In severe cases, COVID-19 can lead to pneumonia, acute respiratory distress syndrome (ARDS), and organ failure, which can be fatal.

Pneumonia:

Pneumonia is an infection of the lungs that can be caused by a variety of pathogens, including bacteria, viruses, or fungi. It can be a serious condition, particularly in young children, the elderly, or those with weakened immune systems. Symptoms may include coughing, fever, chest pain, shortness of breath, and fatigue. Treatment usually involves antibiotics or antiviral medications, as well as supportive care such as rest and hydration.

By developing a deep learning model that can accurately diagnose respiratory diseases through chest X-rays, I aim to provide healthcare providers with a reliable and efficient tool for early detection and treatment of these conditions. The proposed model has the potential to significantly improve patient outcomes and save lives.



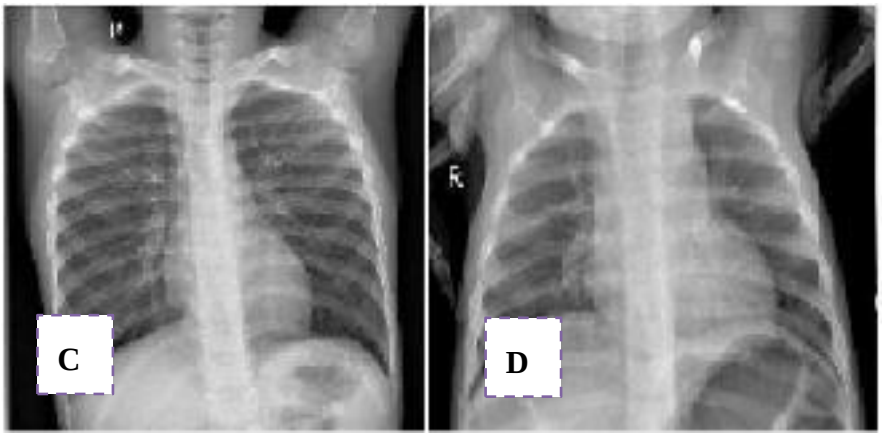


Figure 1:(Without Chest infection images**)**

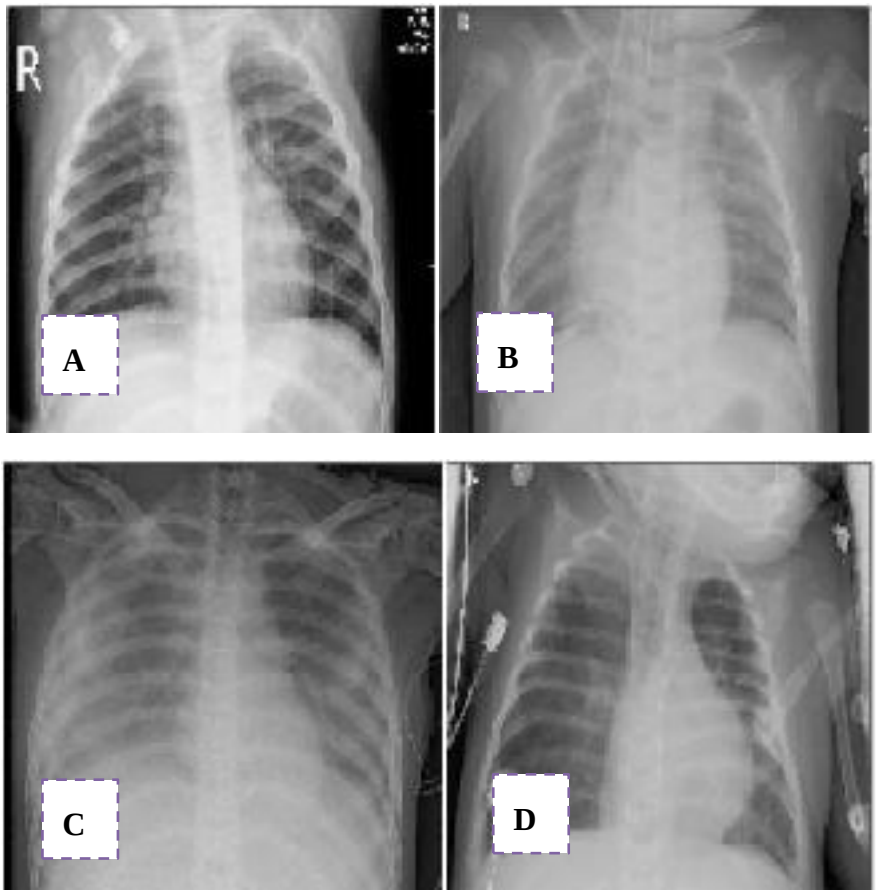


Figure 2: (With Chest infection sample images**)**

The proposed system aims to provide healthcare providers with a reliable and efficient tool for respiratory disease diagnosis, enabling earlier detection and treatment. By leveraging the power of CNN algorithms, the proposed system can automatically extract relevant features from chest X-rays and provide accurate and timely diagnosis, potentially saving lives and improving patient outcomes.

3. Motivation and Scope

The scope of this project is to develop a multi-class classification model for the automated diagnosis of respiratory diseases using X-ray images. The model will be trained to classify a patient's condition into one of four classes, which are asthma, bronchitis, lung cancer, and normal respiratory function. The dataset for this project will consist of X-ray images of patients diagnosed with respiratory diseases, along with demographic and clinical data such as medical history, symptoms, and results from medical tests. The proposed model will use these data to develop a machine learning algorithm that can accurately diagnose respiratory diseases. The model will be developed using deep learning techniques such as convolutional neural networks (CNNs), which have shown promising results in medical image classification tasks. The model will be evaluated using standard metrics such as accuracy, precision.

The motivation for this project is to develop a machine learning-based solution that can aid in the accurate and timely diagnosis of respiratory diseases, ultimately improving patient outcomes and reducing healthcare costs.

4. Related Work

Automated diagnosis of respiratory diseases using machine learning and deep learning techniques has been a topic of research for several years. Various studies have been conducted to develop models for accurate and efficient diagnosis of respiratory diseases.One such study by Rajpurkar et al. (2017) presented a deep learning model based on the Inception architecture for the automated diagnosis of chest. The model showed promising results in the detection of pneumonia, achieving an area under the curve (AUC) score of 0.92.

Overall, these studies demonstrate the potential of machine learning and deep learning techniques in the automated diagnosis of respiratory diseases. Our proposed model builds upon these works and aims to develop a multi-class classification model for the accurate and efficient diagnosis of bronchitis, lung cancer, asthma, and pneumonia using chest X-rays.

5. System Architecture

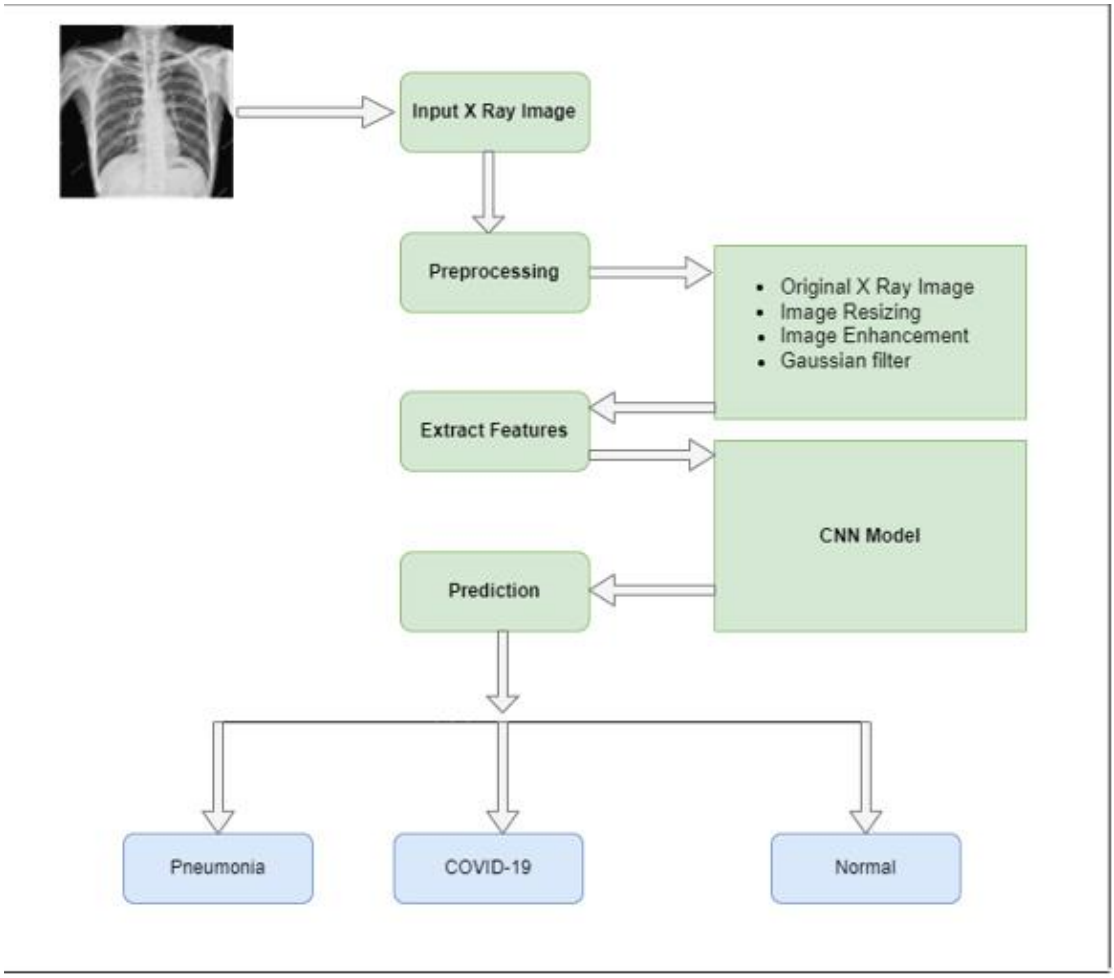


Figure 3 System Architecture

6. Goals and Objectives

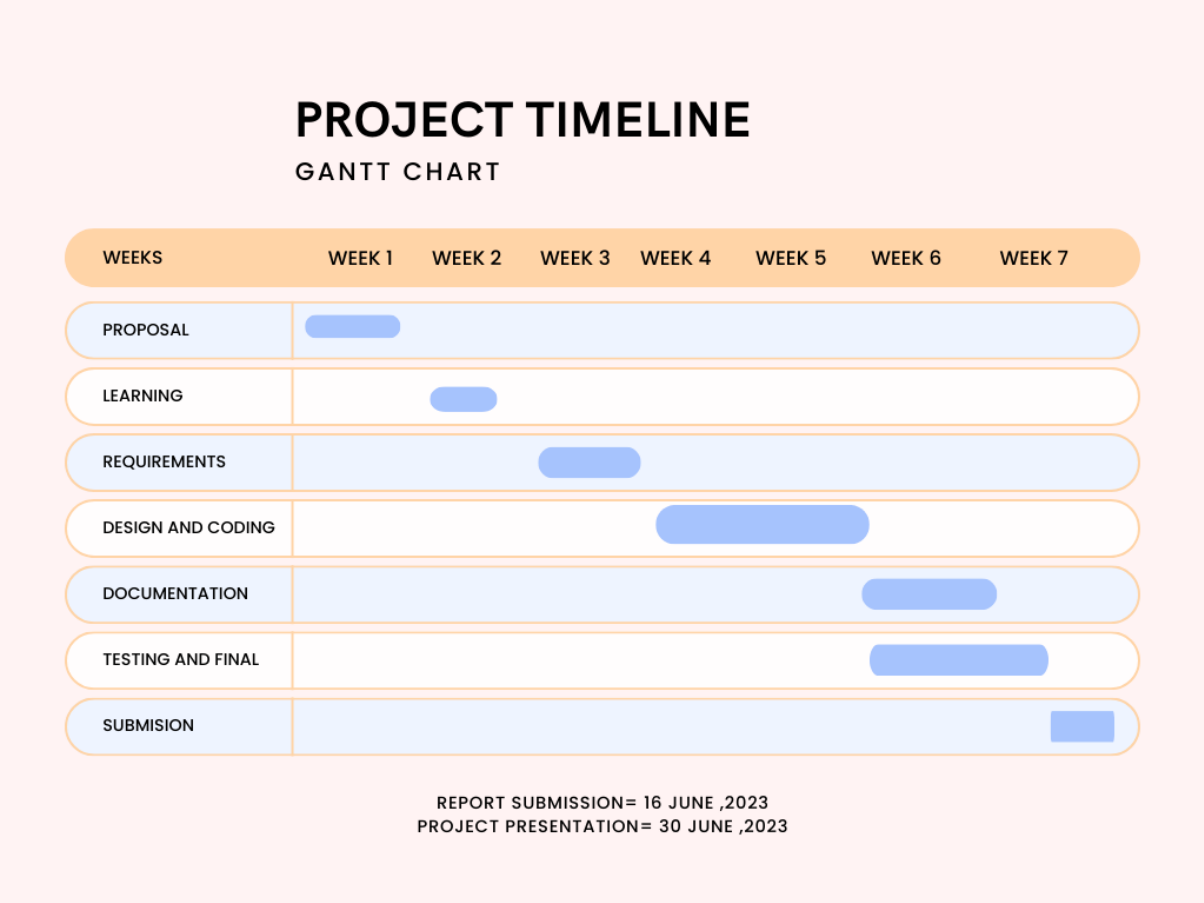
The primary goal of your project is to develop a multi-class classification model for the automated diagnosis of respiratory diseases using X-ray images. Specifically, your model will classify X-ray images into one of four classes: asthma, bronchitis, lung cancer, or normal respiratory function.

The main objective of this project will be to build a system which will helpful for the radiologists and physicians to classify the chest infection accurately and efficiently in X-ray image. I also intending to accomplish following sub-objectives:

- Use in Real time applications.

- Fast Computational time.
- Less prone to human error.
- Robust and flexible to large data amount.
- Compare Accuracy Results with other techniques.
- Cost effective.

Gantt Chart



7. Deliverables and Development Requirements

Following are the requirements to create our project:

- X-RAY images of infected chests by COVID, Pneumonia and Normal.
- X-RAY images of non-infection of chest.

The list of deliverable consists of following:

- Desktop application
- Documentation Development requirements include following software and hardware requirements:
- IDE: PyCharm
- Anaconda
- Dataset based work.

- PyTorch for Libraries
- 8 GB RAM system
- MS WORD
- Programming Language: Python
- Interface: PyQT Designer

Table 1: Tools and Technology

	Tools	Version	Rationale
Tools And Technologies	PyCharm	2019	IDE
	MS Word	2016	Documentation
	MS Power Point	2016	Presentation
	PyQT	Latest	Interface
	Python	Latest	Programming Language

8. Dataset

Data set is available on the following links.

URL: https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia

1. https://github.com/ieee8023/covid-chestxray-dataset
2. https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
3. https://github.com/agchung

Content

Dataset is organized into 2 folders (train, test) and both train and test contain 3 subfolders (COVID19, PNEUMONIA, NORMAL). DataSet contains total 6432 x-ray images and test data have 20% of total images.