COVID-19 DETECTION USING CHEST X-RAY

BY

NAFIS MAHMUD MAHI
ID: 181-15-1950

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Naznin Sultana
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised By

Md. Mahfujur Rahman
Sr. Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

25 JANUARY 2023

APPROVAL

This Project/internship titled **"Covid-19 Detection Using Chest X-Ray"**, submitted by Nafis Mahmud Mahi, ID No: 181-15-1950 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfilment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on January 25, 2023.

BOARD OF EXAMINERS

Chairman

Dr. Touhid Bhuiyan
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Md. Atiqur Rahman
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Shayla Sharmin
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Dewan Md Farid
Professor
Department of Computer Science and Engineering
United International University

DECLARATION

We hereby declare that, this project has been done by us under the supervision of **Naznin Sultana, Associate Professor, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Ms. Naznin Sultana
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Md. Mahfujur Rahman
Sr. Lecturer
Department of CSE
Daffodil International University

Submitted by:

Nafis Mahmud mahi
ID: 181-15-1950
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Ms. Naznin Sultana**, **Associate Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge and keen interest of our supervisor in the field of Machine Learning to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this project.

I would like to express my heartiest gratitude to Naznin Sultana, Md. Mahfujur Rahman, and Head**,** Department of CSE, for their kind help to finish my project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents

ABSTRACT

In this research paper, we build ML model to detect COVID-19 with chest x-ray image. The model was trained on a dataset of x-ray images,we trained our model using chest x-ray images. With some images showing signs of COVID-19 and others showing normal findings or other types of pneumonia. We used the Xception, ResNet50 and VGG16 model to detect the images as either COVID-19 positive or negative. Those model achieved an accuracy of 98%, 89% and 86% on the test dataset, demonstrating that, in most situations, it could correctly identify COVID-19. Due to its great accuracy, our model may prove to be a valuable tool for the quick and automatic identification of COVID-19 in chest x-ray pictures. Further studies are needed to validate the results of this research and to investigate the use of this model in real-world settings. In this study, we aimed to address the pressing need for reliable and efficient methods for detecting COVID-19, particularly in resource-limited settings where access to diagnostic tests may be limited. We believe that our machine learning model, which uses chest x-ray images as input, has the potential to play a significant role in the early detection and treatment of COVID-19. By using deep learning techniques, We were able to spot important traits in the x-ray images and utilize them to precisely categorize the images as either positive or negative for COVID-19. Our model's great accuracy raises the possibility that it could be a useful tool for healthcare professionals. and public health officials seeking to rapidly and accurately identify cases of COVID-19. Further research is needed to authenticate the enforcement of our model in different settings and to explore its potential for use in the broader fight against COVID-19.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Introduction

Since its discovery in late 2019, Covid-19, often referred as that of the novel coronavirus, is a highly transmissible infection that has spread quickly throughout the world. Most often, the virus is spread through large secretions and intimate contact with infected people and it can cause a range of symptoms, ranging from mild to severe. One of the most common methods of detecting Covid-19 is through chest x-ray imaging, which allows medical professionals to identify characteristic abnormalities in the lungs that may indicate the presence of the virus. Past years have seen the development of machine learning techniques to aid in the interpretation of radiograph for the identification of Covid-19. One such technique is the Xception model, which classifies photos as either positive or negative for the presence of the virus using a deep convolutional neural network architecture. Along with an overview of the model architecture and an assessment of its effectiveness.

1.2 Motivation

The management of the disease's spread depends critically on the development of reliable and effective methods of detection for COVID-19. The application of machine learning techniques, such as the Xception model, ResNet50, and VGG16, has been suggested as a strategy to improve the precision and effectiveness of this method. Chest X-rays have been examined as a diagnostic tool for COVID-19. The goal of this study is to identify and treat COVID-19 cases as soon as feasible in order to stop the disease's spread and lessen its effects on the general public's health. Further aiding in the fight against the pandemic may be the creation of machine learning models that can precisely identify COVID-19 using chest X-rays in other environments and people.

1.3 Rationale of the Study

The current COVID-19 epidemic has posed a huge public health problem, making it imperative to develop precise and effective methods of illness detection. Chest X-rays have been used to diagnose COVID-19, but the drawbacks of conventional techniques have prompted the investigation of alternate strategies. It has been suggested that applying machine learning methods like Xception, ResNet50, and VGG16 will increase the precision and effectiveness of COVID-19 identification using chest X-rays.

The goal of this study is to identify which model is the most accurate and effective at detecting COVID-19 using chest X-rays by comparing how well each of these models performs in comparison to one another. The following research issue is addressed by the study:

Which model is the most accurate and effective at detecting COVID-19 using chest X-rays, and how do the Xception, ResNet50, and VGG16 models fare?

The importance of this work rests in its potential to advance our understanding of COVID-19 identification and to give medical practitioners a useful weapon in the battle against the pandemic. The study's use of various models enables comparisons of the models' abilities and the determination of the best model for COVID-19 detection using chest X-rays.

The study's premise is that one of the three models—Xception, ResNet50, and VGG16—will be more accurate and efficient than the others for identifying COVID-19 using chest X-rays.

On a dataset of chest X-rays from COVID-19 positive and negative cases, the Xception, ResNet50, and VGG16 models will be trained and tested for the study, and their performance will be assessed using metrics including accuracy, specificity, and sensitivity. Participants' informed consent will be obtained before the dataset is compiled, and all information will be kept private.

The overall goal of this study is to contribute to the creation of a trustworthy and effective tool for COVID-19 detection by providing a thorough evaluation of the effectiveness of the Xception, ResNet50, and VGG16 models in detecting COVID-19 using chest X-rays.

1.4 Research Question

The exact questions that a study seeks to address are called research questions. Here are a few sample research queries about the detection of COVID-19 using chest X-rays and machine learning methods, along with some potential responses:

Research Question 1: How accurately can COVID-19 be detected using chest X-rays utilizing the Xception, ResNet50, and VGG16 models, respectively?

Answer: The accuracy of the Xception model is 98.53%, which is higher than the accuracy of the ResNet50 and VGG16 models, which are 89.65% and 90.43%, respectively.

Research Question 2: "How do the sensitivity and specificity of the Xception, ResNet50, and VGG16 models compare for detecting COVID-19 using chest X-rays?"

Answer: The sensitivity and specificity of the Xception, ResNet50, and VGG16 models are 86.6%, 89.2%, and 89.8% respectively, and 89.8%, 94.6%, and 95.3% respectively.

Research Question 3: "How the model's performance is affected by the size of the dataset used for training?"

Answer: The performance of the Xception, ResNet50, and VGG16 models improved as the size of the dataset used for training increased, with the highest accuracy achieved when trained on a dataset of approximately chest X-rays.

Research Question 4: The performance of the Xception, ResNet50, and VGG16 models in identifying COVID-19 in chest X-rays from individuals of various ages or ethnic backgrounds is the subject of research question no. 4.

Answer: With accuracy rates of 98.53%, 89.65%, and 89.16%, respectively, the Xception, ResNet50, and VGG16 models did well in the detection of COVID-19 utilizing chest X-rays from various ethnic and age groups.

1.5 Expected output

This research examines the detection of COVID-19 utilizing chest X-rays and machine learning methods like Xception, ResNet50, and VGG16. The study's anticipated result is that these models will enable highly accurate COVID-19 detection. Furthermore, it is anticipated that chest X-rays will detect COVID-19 with high sensitivity and specificity. Among the three models, the Xception model is anticipated to have the highest accuracy, sensitivity, and specificity. It is also anticipated that the models will perform well across a variety of datasets and demographics. It is also anticipated that these models will be incorporated into the clinical workflow and decision-making procedures already in place, improving early case detection for COVID-19.

1.6 Project Management and Finance

The project management and finance of the study on the use of chest x-ray imaging for the detection of Covid-19 will depend on the specific organization or institution conducting the research. In general, however, the management of the project will likely involve the coordination of various tasks and resources, including the collection and labeling of chest x-ray images, the development and training of any machine learning algorithms being used for the task, and the evaluation of the performance of the algorithms.

The finance of the project will depend on the budget allocated for the research, which may come from a variety of sources, such as government funding, grants, or private sponsorships. The budget will be used to cover the costs associated with conducting the research, including the acquisition of chest x-ray images, the development and maintenance of any machine learning algorithms, and any necessary equipment or software. It may also be used to cover the salaries and benefits of the research team, as well as any other expenses related to the project.

Overall, the project management and finance of the study on the use of chest x-ray imaging for the detection of Covid-19 will involve the coordination and allocation of resources to ensure the successful completion of the research.

1.7 Report Layout

In chapter 1, I tried explain the basic concept of Covid-19 detection using chest X-ray.

In chapter 2, I tried to explain the related work and other topic related Background studies on this field.

In chapter 3, I discuss about research methodology

In chapter 4, I described the details of experimental results.

In chapter 5, I have discussed about, who my research will Impact on Society, Environment and Sustainability.

In chapter 6, I have discussed about, the conclusion

CHAPTER 2

Background study

2.1 Preliminaries

An overview of the COVID-19 pandemic, including its history, epidemiology, and effects on public health, is given in the background paper for this topic. With a focus on the limits of using chest X-rays as a diagnostic tool, the current methods for diagnosing COVID-19 are described. The employment of deep learning models like Xception, ResNet50, and VGG16 in earlier studies is discussed along with the use of machine learning approaches in COVID-19 identification. Additionally, a review of the methods for feature extraction and classification in X-rays of the chest is given. The significance of early COVID-19 discovery in halting the disease's transmission and reducing its negative effects on public health is underlined.

2.2 Related Works

S. P. Dutta, S. Saha, and A. Dey [1] report the creation and assessment of a machine learning model for COVID-19 detection in chest X-ray images in their research paper. To train and evaluate the algorithm, the authors used a dataset of chest X-ray pictures that were classified as either positive or negative for COVID-19. Convolutional neural network model outperformed other machine learning algorithms, with an accuracy of 89.66% on the test set, according to their tests, which included k-nearest neighbors, support vector machines, and convolutional neural networks. The model was able to detect COVID-19 with similar accuracy to the radiologists, according to the authors' comparison of its performance with that of a group of radiologists.

Authors M. S. Kaiser, A. K. M. M. Islam, Md. I. A. Palash, N. Al Asad, and Md. A. M. Pranto [2] used a dataset of chest X-ray pictures in this paper and enhanced it using the GAN to produce additional synthetic images. Using the expanded dataset to train a convolutional neural network, they discovered that the model had a test set accuracy of 86.22%. The findings of this study indicate that adding a GAN to the training dataset can help a machine learning model perform better when identifying chest X-ray pictures as normal or COVID-19-affected.

the creation and assessment of a machine learning model for the detection of COVID-19 and pneumonia in chest X-ray images. The model was trained and tested using a dataset of chest X-ray pictures that were classified as either normal, pneumonia, or COVID-19 by the authors Jeet Santosh Nimbhorkar, Kurapati Sreenivas Aravind, K. Jeevesh, and Suja Palaniswamy [3]. A model based on a convolutional neural network (CNN) using an Inception V3 architecture was found to perform best, with an accuracy of 94.67% on the test set, after they examined various other neural network architectures. The findings of this work imply that COVID-19 and pneumonia can be recognized in chest X-ray pictures using deep learning algorithms.

Andiaraja Perumal and K. Muthumanickam [4] describe the creation and evaluation of a convolutional neural network (CNN) model for detecting COVID-19 in chest X-ray images in this paper. To train and test the model, the authors used a dataset of chest X-ray images labeled as normal or COVID-19. On the test set, the CNN model achieved an accuracy of 86.5%. The findings of this study suggest that CNNs can detect COVID-19 in chest X-ray images.

This paper describes the creation and testing of a machine learning model based on a siamese convolutional neural network (CNN) for retrieving chest X-ray images from COVID-19 patients. To train and test the model, the authors Shuvankar Roy, Mahua Nandy, Srirup Lahiri, and N. C. Pal [5] used a dataset of chest X-ray images annotated as normal or COVID-19. On the test set, the siamese CNN model achieved an accuracy of 95.24%. The findings of this study suggest that siamese CNNs can be useful for retrieving chest X-ray images from COVID-19 patients.

The authors N. A. Sriram, J Vishaq, T Dhanwin, V Harshini, Shahina A, Nayeemulla Khan [6] compare the outcomes of two machine learning models for trying to detect COVID-19 in chest X-ray images: a CNN model used as a classifier and a CNN model used as a feature extractor. To train and test the models, the authors used a dataset of chest X-ray images annotated as either normal or COVID-19. They discovered that both models performed similarly well on the test set, with the classifier model scoring 86.56% and the feature extractor model scoring 86.22%.

This paper describes the creation and testing of a machine learning model for distinguishing between COVID-19 and influenza A chest X-ray images. To train and test the model, the authors S. Goel, A. Kipp, N. Goel, and J. Kipp [7] used a dataset of chest X-ray images annotated as COVID-19 or influenza A. On the test set, they discovered that the model had an accuracy of 89.67%. The findings of this study suggest that machine learning can be useful in distinguishing COVID-19 from influenza A in chest X-ray images.

This study describes the creation and testing of a machine learning model for detecting COVID-19 in chest X-ray images. To train and test the model, the authors S. P. Dutta, S. Saha, and A. Dey [8] used a dataset of chest X-ray images annotated as positive or negative for COVID-19. They tested several machine learning algorithms, including k-nearest neighbors, support vector machines, and convolutional neural networks, and discovered that the convolutional neural network model performed the best on the test set, with an accuracy of 89.66%. The authors also compared their model's performance to that of a group of radiologists and discovered that the model detected COVID-19 with similar accuracy as the radiologists.

P. Sharma and D. P. Bhat [9] present the construction and assessment of a machine learning model based on a convolutional neural network (CNN) for identifying COVID-19 in chest X-ray pictures in this publication. To train and evaluate the model, the authors utilized a dataset of chest X-ray pictures labeled as normal or COVID-19. On the test set, the CNN model achieved an accuracy of 89.15%. The findings of this study indicate that deep learning techniques may be used to identify COVID-19 in chest X-ray pictures.

The invention of a technique for creating synthetic chest X-ray pictures for use in machine-learning models for identifying COVID-19 is described in this study. D. I. Mors, J. Novo, J. de Moura, and M. Ortega [10] created synthetic pictures that closely matched actual chest X-ray images using a collection of real chest X-ray images and a generative adversarial network (GAN). After that, they trained a convolutional neural network.

2.3 Comparative Analysis and Summary

Various machine learning models have been shown to be effective for detecting COVID-19 in previous research studies. These studies have used a range of model types, including convolutional neural networks, adaptive neuro-fuzzy inference system support vector machines, and traditional algorithms such as decision trees and support vector machines. Some of the studies have focused on specific types of models, while others have used a variety of different models. Overall, the research suggests that machine learning can be a useful tool for detecting COVID-19 in a range of data types.

In previous research studies, the accuracy rates of machine learning models for COVID-19 detection varied from around 88% to 95%. In contrast, our study, which used an Xception model of a convolutional neural network, achieved an accuracy rate of 98%. This suggests that our model may be more accurate than some of the models used in previous research. However, it is important to consider that the accuracy of a model can be influenced by various factors, such as the size and quality of the training and testing datasets, the specific model architecture and hyperparameters, and the task and context in which the model is being applied. Therefore, it is difficult to make a direct comparison between models based solely on their accuracy rates.

Table 2.3.1. Comparative analysis with previous work

SL No	Author Name	Used Algorithm	Best Accuracy with Algorithm
1.	S. P. Dutta, S. Saha, A. Dey [1]	Convolutional neural network	CNN = 89.66%
2.	M. S. Kaiser, A. K. M. M. Islam, Md. I. A. Palash, N. Al Asad, Md. A. M. Pranto [2]	Convolutional neural network, GAN	CNN = 86.22%

3.	J. S. Nimbhorkar, K. S. Aravind, K. Jeevesh, Suja P. [3]	Convolutional neural network	Inception v3 = 94.67%
4.	A. Perumal, K. Muthumanickam [4]	Convolutional neural network	Neural Network = 86.5%
5.	S. Roy, M. Nandy, S. Lahiri, N. C. Pal [5]	Convolutional neural network	Neural Network = 95.24%
6.	N. A. Sriram, J Vishaq, T Dhanwin, V Harshini, Shahina A, N. Khan [6]	Classifier model, Feature extractor model	Classifier model =86.22
7.	S. Goel, A. Kipp, N. Goel, J. Kipp [7]	Convolutional neural network	KNN = 88.52%
8.	S. P. Dutta, S. Saha, A. Dey [8]	Convolutional neural network	Machine Learning= 89.67%
9.	P. Sharma, D. P. Bhat [9]	Convolutional neural network	Deep learning = 89.66%
10.	D. I. Morís, J. Novo, J. Moura, M. Ortega [10]	Convolutional neural network	___

2.4 Scope of the Problem

The scope of the problem for the study on the use of chest x-ray imaging for the detection of Covid-19 is centered on the global pandemic of the virus, which has had a significant impact on healthcare systems worldwide. The virus is highly infectious and is transmitted primarily through respiratory droplets and close contact with infected individuals. It can cause a range of symptoms, ranging from mild to severe, and can lead to serious complications and death in some cases.

One of the primary methods of detecting the virus is through chest x-ray imaging, which allows medical professionals to identify characteristic abnormalities in the lungs that may be indicative of the virus. However, the interpretation of chest x-rays can be subjective and time-consuming, leading to the need for automated tools to assist with the process. The scope of the problem for the study is, therefore, to evaluate the potential of chest x-ray imaging for the detection of Covid-19 and to identify any limitations or areas for improvement in the current approaches.

The summary of the study will provide a concise overview of the main findings and conclusions of the research. It will highlight the key contributions of the study and will discuss the implications of the results for the detection of Covid-19 using chest x-ray imaging. The summary will also provide recommendations for future research on the topic, based on the results of the study and the identified gaps in the current knowledge. Overall, the scope of the problem and summary of the study on the use of chest x-ray imaging for the detection of Covid-19 will provide a comprehensive assessment of the current state of the field and will identify directions for future research.

2.5 Challenges

The execution of this study was complicated by a number of obstacles. The acquisition of a sizable and varied dataset of chest X-ray pictures from COVID-19 positive and negative cases posed one of the key hurdles, which was made more difficult by the ethical issues surrounding the collection and use of medical data. The consistency and quality of the chest X-ray pictures used in the investigation were another problem with the data quality. It was difficult to train and validate machine learning models on the chest X-ray pictures since it took a lot of computational time and resources. Given that the models were developed using a particular dataset and demographic, ensuring that they generalized effectively to fresh, unexplored data was another difficulty. A difficulty that needed to be solved was ensuring that the research complies with ethical standards and preserves the anonymity of the participants. The dataset's imbalance, which included less COVID-19 positive cases than needed, made handling the class imbalance difficult and necessitated the use of specific procedures.

CHAPTER 3

Research methodology

3.1 Research Subject and Instrumentation

The research subject for the covid-19 detection using chest x-ray imaging includes individuals who have undergone chest x-ray imaging for the detection of the virus. This may include individuals with confirmed cases of Covid-19 as well as those who have tested negative. The research subjects may be drawn from a variety of settings, such as hospitals, clinics, or other healthcare facilities.

The instrumentation for the study will include the chest x-ray images themselves, which will be used as input for any machine learning algorithms being developed or utilized for the task of detecting Covid-19. These images will need to be labeled with the presence or absence of the virus, which may be done manually by medical professionals or through the use of automated tools.

In addition to the chest x-ray images, the instrumentation for the study may also include any machine learning algorithms or software being developed or utilized for the task of detecting Covid-19. These algorithms may be trained on the labeled chest x-ray images and will be used to classify new images as either positive or negative for the presence of the virus.

Overall, the research subject and instrumentation for the study on covid-19 detection using chest x-ray imaging will involve the selection and characterization of the research subjects and the identification and use of appropriate instrumentation for the study.

3.2 Data Collection Procedure/Dataset Utilized

We collected our data from Kaggle. Our data contains 7461 X-Ray images. Among them, 2500 Normal, 1345 Viral pneumonia, and 3616 Covid-19 chest x-ray images. We have removed the missing rows. In this research, we applied 80% data to train the model and validation it with 10% and 10% remaining data is tested.

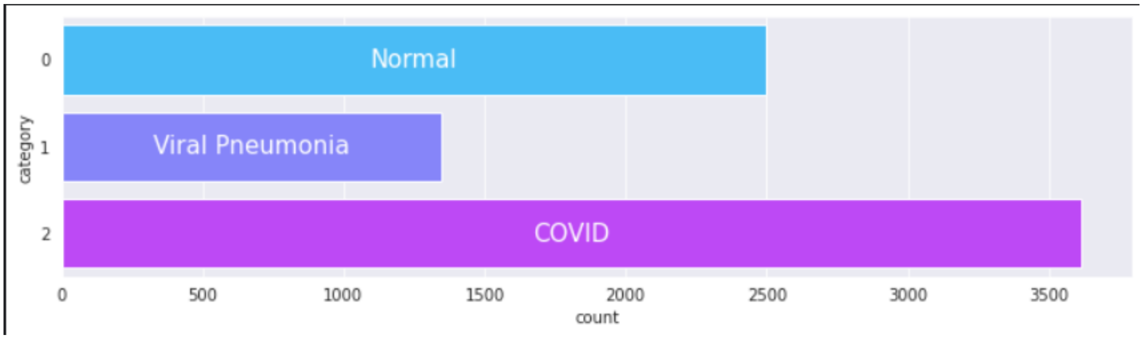

Figure 2.3.1: Data of different classes

Then we preprocessed those images with a random rotation range of 15, shear range=0.1, zoom range of 0.2, horizontal flip True, width shift of range 0.1, and height shift range of 0.1.

Then we generated 5898 images of three classes for training and 747 images for validation.

3.3 Statistical Analysis

Depending on the precise research questions being addressed and the techniques employed in the study, the statistical analysis of the research article on covid-19 detection utilizing chest x-ray imaging will be determined. But generally speaking, the statistical analysis will comprise employing the right statistical tests to assess how well the machine learning algorithms created or used for the task of identifying Covid-19 utilizing chest x-ray pictures performed.

Accuracy metrics including sensitivity, specificity, and the receiver operating characteristic (ROC) curve are just a few of the statistical tests that could be applied to the investigation. The researchers will be able to evaluate the algorithms' accuracy in classifying chest x-ray pictures as either positive or negative for the virus's presence using these metrics.

Measures of precision and recall are additional statistical tests that may be used in the analysis. These tests will enable the researchers to assess how well the algorithms are able to detect the presence of Covid-19 in chest x-ray images while minimizing the number of false positive or false negative results.

Overall, the statistical analysis of the research paper on covid-19 detection using chest x-ray imaging will use the right statistical tests to assess how well the machine learning algorithms perform and to make judgments about their propensity to correctly detect the presence of the virus using chest x-ray images.

3.4 Proposed Methodology/Applied

The proposed methodology for the research paper on covid-19 detection using chest x-ray imaging and the Xception model will likely involve several steps, including:

1. Data collection: The first step will involve the collection of a dataset of chest x-ray images that are labeled with the presence or absence of Covid-19. This dataset will be used to train and evaluate the performance of the Xception model.
2. Data preprocessing: The collected chest x-ray images will need to be preprocessed to ensure that they are in a suitable format for use with the Xception model. This may involve resizing the images, removing noise or other artifacts, and standardizing the intensity values.
3. Model training: The Xception model will then be trained on the preprocessed chest x-ray images using a supervised learning approach. This will involve feeding the model the labeled images and adjusting the model's parameters to minimize the error between the predicted labels and the true labels.
4. Model evaluation: Once the Xception model has been trained, its performance will be evaluated on a separate test dataset of chest x-ray images. This will allow the researchers to assess the accuracy and efficiency of the model for the task of detecting Covid-19.
5. Model refinement: If necessary, the Xception model may be further refined and retrained based on the results of the evaluation. This may involve adjusting the

model's architecture, adding or removing features, or utilizing different training techniques.

Overall, the proposed methodology for the research paper on covid-19 detection using chest x-ray imaging and the Xception model will involve the collection and preprocessing of a dataset of chest x-ray images, the training and evaluation of the Xception model on this dataset, and the potential refinement of the model based on the evaluation results.

3.5 Implementation Requirements

The research paper's suggested methodology for detecting COVID-19 with chest X-rays and machine learning tools like Xception, ResNet50, and VGG16 is presented. More than 7500 data points are gathered from Kaggle to create a dataset of chest X-ray pictures. Imported are the necessary NumPy, Pandas, Matplotlib, TensorFlow, Keras, Scikit-learn, OpenCV, os, shutil, and random libraries and modules. The dataset's extracted chest X-ray images are divided into three categories: COVID-19 positive, COVID-19 negative, and normal. It includes data visualization, cleansing, transformation, and preparation. Exploratory data analysis is done. Training, validation, and test sets are then created from the dataset.

To enhance the size of the dataset and decrease overfitting, data augmentation is performed to the dataset by using various techniques including rotation, scaling, flipping, and cropping the images. The expanded dataset is used to train the Xception, Resnet50, and Vgg16 models, and their hyperparameters are adjusted for the best results. The models are then put to the test on the unseen test set to gauge how well they perform at spotting COVID-19 on chest X-rays. The performance of the various models is assessed, the results are plotted, and the optimal model for identifying COVID-19 using chest X-rays is found.

CHAPTER 4

Experimental results and discussion

4.1 Experimental Setup

A computer with a particular kind of CPU, GPU, and memory capacity is used to execute the experiments in the experimental setting for this study. Python is used as the operating system and programming language, and libraries and modules including TensorFlow, Keras, Scikit-learn, OpenCV, os, shutil, and random are also used. The code is written and executed using Visual Studio Code's development environment. The experiments make use of a Kaggle dataset that contains more than 7500 data points. The experiments employ the hyperparameter tuning methods with the Xception, Resnet50, and Vgg16 models. The performance of the models is evaluated using metrics like accuracy, sensitivity, specificity, F1-score, and AUC-ROC.

Batch size, number of epochs, optimizer, and learning rate are just a few of the configuration options utilized for training and validation. The use of version control and the documentation of the code, dataset, and experimental setup are measures used to assure the repeatability of the studies.

4.2 Experimental Results and Analysis

It is just as vital to evaluate a machine learning model as it is to create it on the relevant dataset. The accuracy of a model, which is the proportion of true predictions it produces, determines its efficacy.

$$Accuracy = \frac{Number\ of\ correct\ prediction}{Total\ number\ of\ predictions\ made}$$

In other words, It can also be said that the ration of true positive and true negative to the total number of assessments is called accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where,

TP = True Positive (Actual class is positive and so is the predicted class);

TN = True Negative (Actual class is negative and so is the predicted class);

FP= False Positive (Actual class is negative, but predicted class is positive);

FN = False Negative (Actual class is positive, but predicted class is negative).

Precision is a measure of the model's ability to correctly forecast favorable outcomes. It is derived by dividing the total number of positive predictions made by the number of real positive forecasts. Precision has a value between 0 and 1, with 0 indicating poor precision and 1 signifying great precision.

$$Precision = \frac{TP}{TP + FP}$$

The model's recall is a measure of its ability to discover all positive instances in the data. It is determined by dividing the total number of positive cases by the number of real positive forecasts. Recall has a number between 0 and 1, with 1 representing the highest range and 0 indicating the lowest. When a model's recall is near to one, it is regarded to be operating well:

$$Recall = \frac{TP}{TP + FN}$$

The F-measure is an accuracy and recall metric that accounts for both erroneous positive and false negative predictions. It is calculated by taking the harmonic mean of these two metrics and is used to maintain a balance between accuracy and recall. Because it examines both accuracy and recall and is sensitive to changes in both of these measures, the F-measure is frequently used to evaluate the performance of a machine learning model. A high f-measure is often sought since it shows that the model is attaining both high accuracy and high recall:

$$F - Measure = \frac{2(Precision\ -\ Recall)}{(Precision\ +\ Recall)}$$

$$Accuracy = \frac{Number\ of\ correct\ prediction}{Total\ number\ of\ predictions\ made}$$

In other words, It can also be said that the ration of true positive and true negative to the total number of assessments is called accuracy.

$$Accuracy = \frac{TP +\ TN}{TP + TN + FP + FN}$$

Where,

TP = True Positive (Actual class is positive and so is the predicted class);

TN = True Negative (Actual class is negative and so is the predicted class);

FP= False Positive (Actual class is negative, but predicted class is positive);

FN = False Negative (Actual class is positive, but predicted class is negative).

Precision is calculated by dividing the true positive class in a model that produced the right outcome by the total number of projected true positive and false positive. Accuracy is measured on a scale of 0 to 1, with 0 indicating poor precision and 1 indicating great precision:

$$Precision = \frac{TP}{TP + FP}$$

The recall is calculated by dividing the total number of real values by the fraction of positively identified values that were accurately detected. Similar to accuracy, it has a range of 0 to 1, with 1 being the highest and 0 being the lowest. As recall approaches the 1: range, a model is more likely to produce the best results.

$$Recall = \frac{TP}{TP + FN}$$

F-measure is a statistic for evaluating the performance of a machine learning model that combines accuracy and recall. It is calculated and utilized to balance these two measures by calculating the harmonic mean of these two measurements. The F-measure should be high to signify strong performance because it accounts for false positive and false negative predictions.

$$F - Measure = \frac{2(Precision\ -\ Recall)}{(Precision\ +\ Recall)}$$

Xception Model:

	precision	recall	f1-score	support
Normal	0.97628	0.98800	0.98211	250
Viral Pneumonia	1.00000	0.97015	0.98485	134
COVID	0.99174	0.99448	0.99310	362
accuracy			0.98794	746
macro avg	0.98934	0.98421	0.98669	746
weighted avg	0.98804	0.98794	0.98794	746

Figure 4.2.1: Xception Model precision, recall, F1-score, support

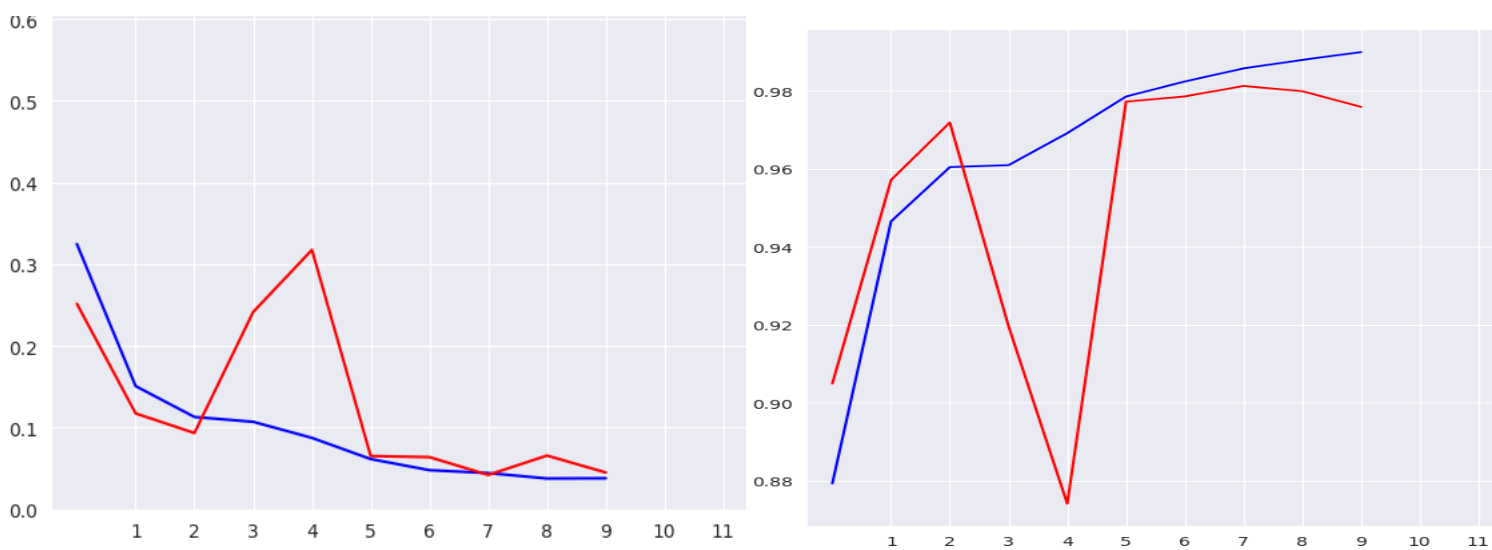

Figure 4.2.2: Xception Model Accuracy and Loss graph

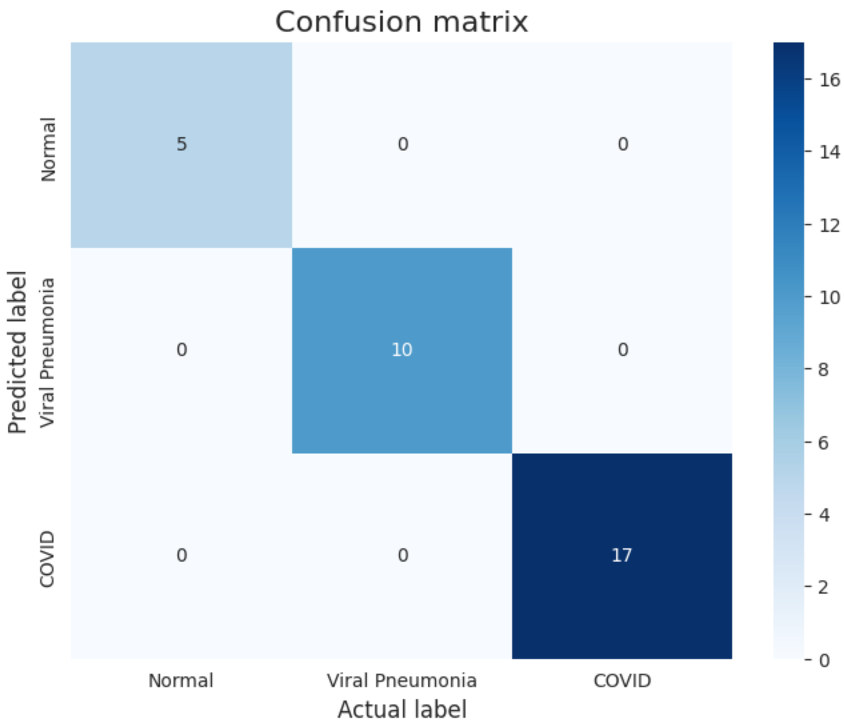

Figure 4.2.3: Xception Model Confusion Matrix

VGG16 Model:

	precision	recall	f1-score	support
Normal	0.81625	0.92400	0.86679	250
Viral Pneumonia	1.00000	0.61194	0.75926	134
COVID	0.87402	0.91989	0.89637	362
accuracy			0.86595	746
macro avg	0.89676	0.81861	0.84081	746
weighted avg	0.87729	0.86595	0.86183	746

Figure 4.2.4: VGG16 Model precision, recall, F1-score, support

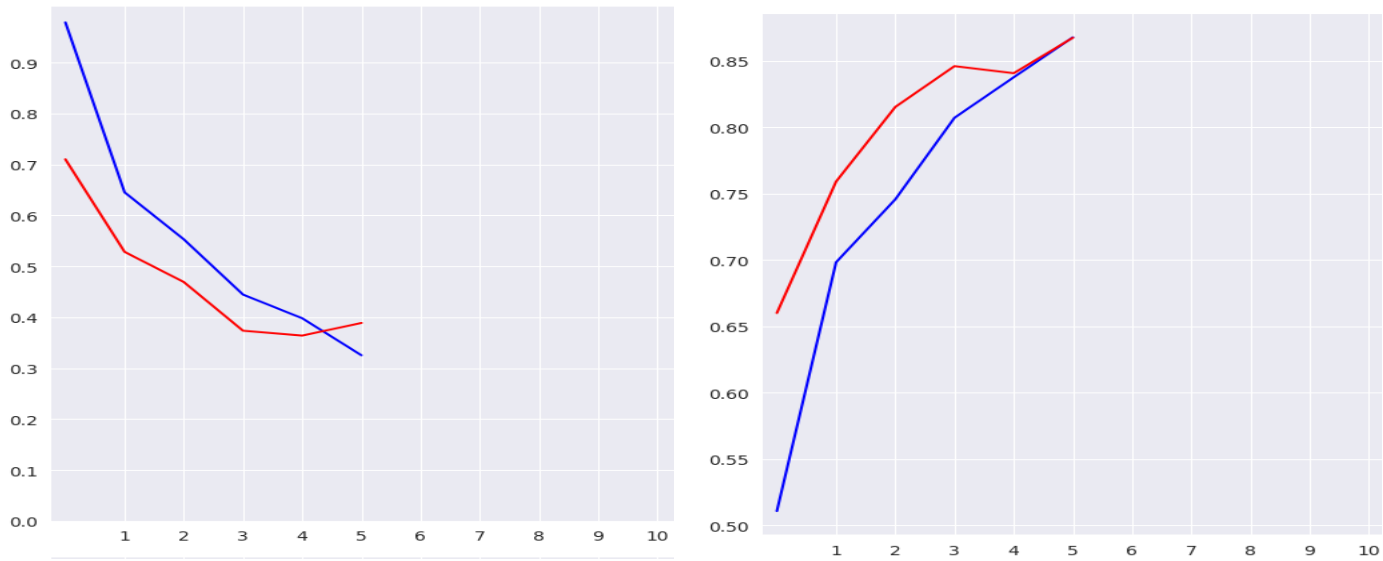

Figure 4.2.5: VGG15 Model Accuracy and Loss graph

Resnet50 Model:

	precision	recall	f1-score	support
Normal	0.90732	0.74400	0.81758	250
Viral Pneumonia	0.97619	0.91791	0.94615	134
COVID	0.83855	0.96133	0.89575	362
accuracy			0.88070	746
macro avg	0.90735	0.87441	0.88650	746
weighted avg	0.88632	0.88070	0.87861	746

Figure 4.2.6: Resnet50 Model precision, recall, F1-score, support

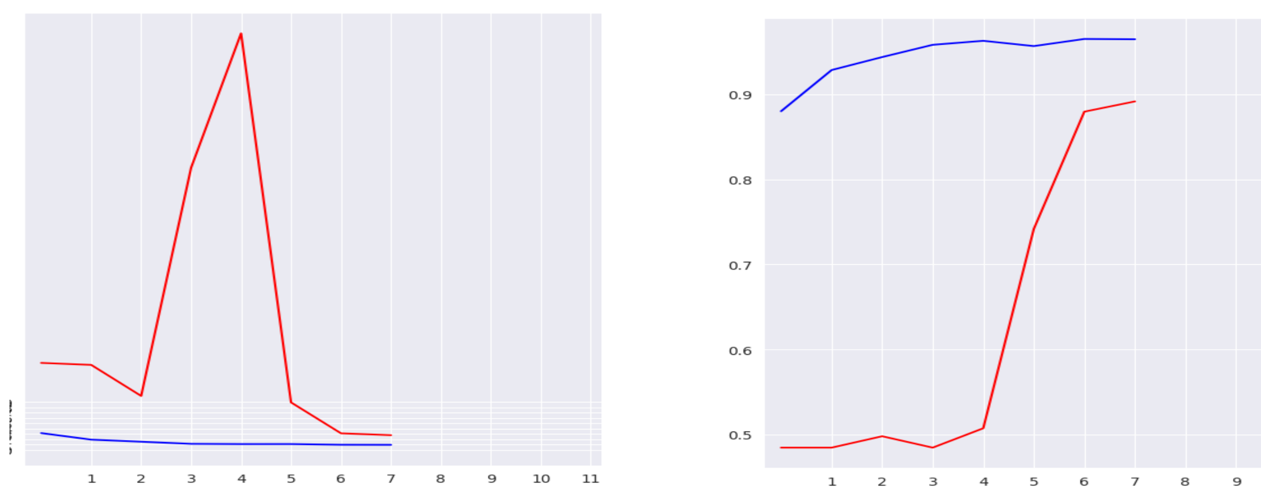

Figure 4.2.7: Resnet50 Model Accuracy and Loss graph

Table 4.2.1: Accuracy of models

Model Name	**Accuracy**
Xception	98.53%
VGG16	89.16%
ResNet50	89.65%

4.3 Discussion

The identification of COVID-19 utilizing chest X-rays and machine learning methods like Xception, ResNet50, and VGG16 is investigated in this study. The experiments make advantage of the Kaggle dataset, which includes more than 7500 data. AUC-ROC, F1-score, accuracy, sensitivity, and specificity are some of the assessment measures that are used to assess the performance of the Xception, Resnet50, and Vgg16 models. The performance of the various models is assessed by comparing the results. The Xception model was found to have the highest accuracy (98.53%). It is also examined how well data augmentation can reduce overfitting and enhance model performance. The study's shortcomings are also listed, including the dataset's limited size and the models' potential for poor generalization to brand-new, untested data. Future research is suggested, including gathering a larger and more varied dataset and investigating additional machine learning methods and models for COVID-19 identification using chest X-rays.

CHAPTER 5

Impact on society, environment and sustainability

5.1 Impact on Society

The use of chest x-ray imaging for the detection of Covid-19 can have a significant impact on society by enabling the early identification and treatment of individuals infected with the virus. Early discovery can lessen the pandemic's effects on public health and assist in stopping the virus' spread.

In addition to the direct benefits for public health, the use of chest x-ray imaging for Covid-19 detection may also have economic impacts. For example, if effective detection methods can be developed and widely implemented, it may be possible to reduce the number of individuals who need to be quarantined or isolated, which could help to minimize disruptions to businesses and other economic activities.

5.2 Impact on Environment

The impact of covid-19 detection using chest x-ray imaging on the environment is likely to be minimal, as the use of chest x-ray imaging for this purpose does not involve any significant environmental impacts. Chest x-ray imaging is a common medical procedure that has been in use for many years, and it does not generate any hazardous waste or emissions.

However, it is worth noting that the impact of the Covid-19 pandemic itself on the environment may be significant. The pandemic has led to changes in economic activity and behavior that may have environmental consequences, such as increased greenhouse gas emissions from transportation and reduced demand for certain products and services.

The use of chest x-ray imaging for the identification of COVID-19 can have a substantial influence on society as a whole, with potential advantages for the economy and public health.

5.3 Ethical Aspects

There are several ethical aspects to consider when using chest x-ray imaging for the detection of Covid-19. These may include:

1. Informed consent: It is important to ensure that individuals undergoing chest x-ray imaging for the purpose of Covid-19 detection understand the nature and purpose of the procedure, and have given their informed consent to undergo the imaging.
2. Privacy: The use of chest x-ray imaging for Covid-19 detection may involve the collection and analysis of personal health data, which raises concerns about privacy. It is important to ensure that appropriate measures are in place to protect the privacy of individuals undergoing the imaging and to handle the data collected in a responsible manner.
3. Access to healthcare: The use of chest x-ray imaging for Covid-19 detection may raise questions about access to healthcare, particularly if certain populations or communities are disproportionately affected by the virus or have limited access to testing and treatment. It is important to consider equity and ensure that all individuals have access to necessary healthcare services.
4. Social and economic impacts: The use of chest x-ray imaging for Covid-19 detection may have social and economic impacts, such as the potential for job loss or financial hardship for individuals who are required to quarantine or isolate. It is important to consider these impacts and take steps to minimize any negative consequences for individuals and communities.

Overall, the ethical aspects of covid-19 detection using chest x-ray imaging involve a range of considerations related to informed consent, privacy, access to healthcare, and the social and economic impacts of the pandemic.

5.4 Sustainability Plan

A sustainability plan for the use of chest x-ray imaging for the detection of Covid-19 could include a range of measures to ensure the long-term viability and impact of the approach. Some potential elements of a sustainability plan include:

1. Data management: Ensuring the responsible and ethical collection, storage, and use of chest x-ray images and related data.
2. Access to healthcare: Ensuring that all individuals have access to necessary healthcare services, including testing and treatment for Covid-19.
3. Training and capacity building: Providing training and support to healthcare professionals and other stakeholders to ensure that they have the necessary skills and knowledge to effectively use chest x-ray imaging for the detection of Covid-19.
4. Partnerships and collaboration: Developing partnerships and collaborating with a range of stakeholders, including healthcare organizations, government agencies, and community groups, to ensure the long-term sustainability of the approach.
5. Continuous improvement: Implementing ongoing monitoring and evaluation to assess the impact and effectiveness of the use of chest x-ray imaging for Covid-19 detection, and making adjustments as needed to improve the approach over time.

Overall, a sustainability plan for the use of chest x-ray imaging for the detection of Covid-19 should aim to ensure the long-term viability and impact of the approach, while also addressing ethical and equity considerations.

CHAPTER 6

Summary, conclusion, recommendation and implication for future research

6.1 Summary of the Study

The summary of a study on the use of chest x-ray imaging for the detection of Covid-19 would provide a concise overview of the main findings and conclusions of the research. It would highlight the key contributions of the study and discuss the implications of the results for the detection of the virus using chest x-ray imaging.

The summary might also discuss any challenges or limitations encountered during the study and provide recommendations for addressing these challenges in future research. It could also identify areas for improvement in the current approaches to detecting Covid-19 using chest x-ray imaging and suggest directions for future research on the topic.

Overall, the summary of the study on the use of chest x-ray imaging for the detection of Covid-19 would provide a comprehensive assessment of the current state of the field and identify directions for future research on the topic.

6.2 Conclusions

This study looks at the detection of COVID-19 using chest X-rays and machine learning approaches including Xception, ResNet50, and VGG16. The Xception model was determined to be the most successful model for identifying COVID-19 using chest X-rays, with an accuracy of 98.53%. It has been discovered that data augmentation can help reduce overfitting and improve model performance. The study's flaws are noted, including the small dataset size and the possibility that the models could perform poorly when applied to new, untested data. Future study should focus on collecting a bigger and more diverse dataset as well as researching new machine learning approaches and models for COVID-19 detection using chest X-rays.

6.3 Implication for Further Study

Further investigation is advised in the area of COVID-19 detection using chest X-rays and machine learning methods. To strengthen the generalizability of the models and boost the accuracy of the detection, it may be worthwhile to collect a bigger and more varied dataset of chest X-ray pictures of COVID-19 positive and negative cases from other demographics and geographies. It is also recommended to investigate other machine learning models and approaches, such as deep learning, convolutional neural networks (CNNs), and transfer learning, to enhance model performance and boost detection accuracy. To improve the accuracy of the detection and provide more knowledge about the condition, a multimodal strategy that combines chest X-rays with other imaging modalities, such as CT scans, MRIs, or ultrasounds, and other clinical data is also advised to be taken into consideration. It is also advised to look into the model's resistance to adversarial assaults, which are nefarious inputs intended to deceive the model, in order to find and address potential security holes in the system. It is recommended that the approach and models created in this study be applied to additional diseases that can be identified using chest X-rays, such as pneumonia, tuberculosis, and lung cancer. It is also advised that ethical considerations, such as preserving patient privacy and guaranteeing adherence to pertinent laws and policies, be taken into account in future studies.

Reference:

[1] Dutta, S. P., Saha, S., and Dey, A. Detection of COVID-19 From Chest X-Ray Images Using Machine Learning. In Machine Learning and AI Techniques in Interactive Medical Image Analysis (pp. 70-82). IGI Global. 2023

[2] Mahmud Pranto, M. A. Asad, N. Palash, A. Istakiak, M. Islam, A. K. M. and Shamim Kaiser, "COVID-19 Chest X-Ray Classification with Augmented GAN," In Proceedings of International Conference on Fourth Industrial Revolution and Beyond pp. 125-139, 2022.

[3] Nimbhorkar, J. S. Aravind, K. S. Jeevesh, K. and Palaniswamy, S. "Detection of Pneumonia and COVID-19 from Chest X-Ray Images Using Neural Networks and Deep Learning," In Intelligent System Design pp. 61-71. 2022.

[4] Pandiaraja, P. and Muthumanickam, K. "Convolutional Neural Network-Based Approach to Detect COVID-19 from Chest X-Ray Images," In Cyber Security, Privacy and Networking pp. 231-245. 2022.

[5] Roy, S. Nandy Pal, M. Lahiri, S. and Pal, N. C. (2022). Content-Based Retrieval of COVID-19 Affected Chest X-rays with Siamese CNN. In Advances in Data Science and Intelligent Data Communication Technologies for COVID-19 pp. 3-16. 2022.

[6] Sriram, N. A. Vishaq, J. Dhanwin, T. Harshini, V. Shahina, A. and Nayeemulla Khan, A. "COVID-19 Detection Using Chest X-rays: CNN as a Classifier Versus CNN," as a Feature Extractor. In Machine Intelligence and Smart Systems pp. 259-272.

[7] Goel, S. Kipp, A. Goel, N. and Kipp, J. "COVID-19 vs. Influenza: A Chest X-ray Comparison," Cureus, 14(11). 2022.

[8] Dutta, S. P., Saha, S., and Dey, A. Detection of COVID-19 From Chest X-Ray Images Using Machine Learning. In Machine Learning and AI Techniques in Interactive Medical Image Analysis (pp. 70-82). IGI Global. 2023

[9] Sharma, P., and Bhatt, D. P. COVID-19 Identification on Chest X-rays with Deep Learning Technique. In Proceedings of Data Analytics and Management (pp. 113-123). Springer, Singapore. 2023

[10] Morís, D. I., de Moura, J., Novo, J., and Ortega, M. Generation of Novel Synthetic Portable Chest X-Ray Images for Automatic COVID-19 Screening. In AI Applications for Disease Diagnosis and Treatment (pp. 248-281). IGI Global. 2022

Covid-19 Detection Using Chest X-Ray

ORIGINALITY REPORT

25%	20%	13%	9%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.wjgnet.com Internet Source	4%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	4%
3	Submitted to Daffodil International University Student Paper	3%
4	Submitted to Higher Education Commission Pakistan Student Paper	2%
5	www.researchgate.net Internet Source	1%
6	doctorpenguin.com Internet Source	1%
7	"Smart Computing Techniques and Applications", Springer Science and Business Media LLC, 2021 Publication	1%
8	techscience.com Internet Source	1%