North South University

Department of Electrical and Computer Engineering

Senior Design Project

DEEP LEARNING APPROACHES TO COVID-19 DETECTION IN CHEST X-RAYS IN BANGLADESH

Ratul Bhattarcharjee **ID # 2012996042**
Md. Mushfiqur Rahman Mahin **ID # 2014299042**
Mohammad Olid Afzal **ID # 2011831042**

Faculty Supervisor:

Dr. Atiqur Rahman

Associate Professor

ECE Department

Spring, 2024

APPROVAL

Ratul Bhattarcharjee ID # 2012996042 , Md. Mushfiqur Rahman Mahin ID # 2014299042, Mohammad Olid Afzal ID # 2011831042 from Electrical and Computer Engineering Department of North South University, have worked on the Senior Design Project titled "OPTIMISING DEEP LEARNING FOR COVID-19 DETECTION IN CHEST X-RAYS" under the supervision of Dr. Atiqur Rahman partial fulfillment of the requirement for the degree of Bachelors of Science in Engineering and has been accepted as satisfactory.

Supervisor's Signature

…….

Dr. Atiqur Rahman
Associate Professor
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh.

Chairman's Signature

…….

Dr. Mohammad Abdul Matin
Professor
Department of Electrical and Computer Engineering
North South University
Dhaka, Bangladesh

DECLARATION

This is to declare that this project is our original work. No part of this work has been submitted elsewhere partially or fully for the award of any other degree or diploma. All project related information will remain confidential and shall not be disclosed without the formal consent of the project supervisor. Relevant previous works presented in this report have been properly acknowledged and cited. The plagiarism policy, as stated by the supervisor, has been maintained.

Students' names & Signatures

1. Ratul Bhattarcharjee

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

2. Md. Mushfiqur Rahman Mahin

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

3. Mohammad Olid Afzal

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

ACKNOWLEDGEMENTS

The authors would like to express their heartfelt gratitude towards their project and research supervisor, Dr. Atiqur Rahman, Professor, Department of Electrical and Computer Engineering, North South University, Bangladesh, for his invaluable support, precise guidance and advice pertaining to the experiments, research and theoretical studies carried out during the course of the current project and also in the preparation of the current report.

Furthermore, the authors would like to thank the Department of Electrical and Computer Engineering, North South University, Bangladesh for facilitating the research. The authors would also like to thank their loved ones for their countless sacrifices and continual support.

ABSTRACT

In order to solve a variety of issues in the field of medical image analysis, deep learning techniques have quickly emerged as a crucial method for medical picture segmentation. This overview discusses basic ideas and applications related to image segmentation, providing a thorough analysis of recent advances in deep learning for medical imaging. Deep learning is versatile enough to tackle tasks like segmentation, registration, object recognition, and image categorization. We highlight the significance of deep learning approaches in the segmentation of chest X-ray images by introducing their fundamentals, applications, and related frameworks. Furthermore, we present prior experiences in this field with various methods, addressing problems including poor classification accuracy, poor segmentation resolution, and poor picture improvement. Medical image analysis is a rapidly evolving discipline that requires strong segmentation algorithms, especially considering the complexity and inaccuracies that can be found in medical imaging. Recent developments have showed promise in tackling these issues, especially in the field of deep neural networks. This paper discusses commonly used data and assessment criteria while reviewing the literature on Chest X-Ray Images using deep custom convolutional neural networks (CNNs), Vgg-19, Mobilenet, and Resnet. This work seeks to direct future research efforts towards improving the effectiveness and reliability of medical picture segmentation algorithms by combining insights from many sources.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

1.1 Background and Motivation

Background:

COVID-19, brought on by the SARS-CoV-2 virus, became a global pandemic and presented severe difficulties for healthcare systems worldwide. Controlling the disease's spread and enhancing patient outcomes depend on early and accurate detection. The accepted method for diagnosing COVID-19, RT-PCR tests has certain drawbacks, including resource-intensive procedures and extended processing periods. There has been an investigation into chest X-ray imaging as a quick and affordable substitute for identifying COVID-19, particularly in areas where RT-PCR testing is not readily available. Enhancing the precision and effectiveness of COVID-19 detection via chest X-rays is possible thanks to deep learning, a branch of artificial intelligence that has demonstrated promise in the field of medical image processing.

Motivation:

- Rapid Diagnosis: To provide a faster alternative to RT-PCR tests, enabling quicker identification and isolation of COVID-19 cases.
- Improved Accuracy: To enhance the sensitivity and specificity of COVID-19 detection in chest X-rays using optimized deep learning models.
- Resource Efficiency: To assist in efficiently allocating medical resources, particularly in areas with limited access to advanced diagnostic tools.
- Support for Healthcare Professionals: To offer a supplementary tool that aids healthcare professionals in making more informed decisions for patient care.

1.2 Purpose and Goal of the Project

Purpose:

This research aims to create an optimal deep-learning model for quick and precise COVID-19 detection from chest X-ray pictures. This strategy wants to add to current diagnostic techniques by providing a quicker, more affordable tool to help medical

personnel identify COVID-19 patients, particularly in environments where RT-PCR testing is not readily available.

The Goal of the Project

- Achieves High Accuracy: Maximizes sensitivity and specificity in detecting COVID-19 from chest X-rays.
- Reduces Diagnosis Time: Provides near-instant results to facilitate quick decision-making.
- Is Resource-Efficient: Operates effectively even in low-resource environments.
- Is Scalable and Accessible: Can be deployed across various healthcare settings, including those with limited access to advanced diagnostic tools.

1.3 Existing Research and Limitations

Existing Research:

Recent advancements in OPTIMISING DEEP LEARNING FOR COVID-19 DETECTION IN CHEST X-RAYS research have yielded several innovative approaches to enhance efficiency, sustainability, and user engagement.

1. Transfer learning methods have been applied in several research and have shown to be quite successful in extracting detailed information from CXR images.
   In the research of Talukder et al. [1] (2024) COVID-19 classification tasks, transfers learning models like Xception, ResNet, InceptionResNetV2, and EfficientNet have demonstrated remarkable performance. Using COVID-19 datasets to fine-tune these pre-trained models can result in much higher accuracy rates. In one series of studies, accuracy rates of 99.55% (Xception), 97.32% (InceptionResNetV2), 99.11% (ResNet50), 99.55% (ResNet50V2), 99.11% (EfficientNetB0), and 100% (EfficientNetB4) were obtained on a dataset of 2000 COVID-19 X-ray pictures.

With 100% accuracy, EfficientNetB4 proved to be the most reliable model out of all of them. Furthermore, utilizing a chest X-ray dataset with 4,350 images, EfficientNetB4 has shown improved performance in lung disease classification tasks. With a precision of 99.13%, recall of 99.16%, and F1-score of 99.14%, the model yielded an accuracy of 99.17%. These findings demonstrate the model's remarkable capacity to classify medical images, which makes it a useful diagnostic tool for radiologists who are dealing with COVID-19 and lung-related diseases. The efficaciousness of refined transfer learning models for prompt and dependable diagnosis is highlighted by the excellent accuracy rates observed in this research, offering vital assistance to medical practitioners amidst the pandemic.

2. In another research of Karim et al. [2] Deep learning algorithms have been integrated to improve diagnostic capacities in their studies. This work presents a unique deep learning-based computer-aided diagnosis tool for COVID-19 detection from symmetric X-ray data. Convolutional Neural Networks (CNN), the Ant Lion Optimization (ALO) algorithm, and a Multiclass Naïve Bayes (NB) classifier are combined in the suggested method. Also, several classifiers, in addition to CNNs, were evaluated, such as Softmax, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Decision Trees (DT). Notably, the NB classifier in conjunction with ALO and CNN produced remarkable outcomes, decreasing execution time and claiming 98.31% accuracy, 100% precision, and 98.25% F1-score.
3. The accurate and rapid detection of COVID-19 remains a critical factor in controlling the spread of the virus and ensuring timely medical intervention. In this regard, Wang et al. [3] developed a multi-class classification model by integrating convolutional neural networks (CNN) with residual learning operations. The model is optimized for deployment on mobile devices, offering a balance between

performance and efficiency. Experimental results demonstrate significant improvements over the COVIDNet-small network, with sensitivity for COVID-19 detection rising from 77.5% to 96.5% and from 88.2% to 95.3% for non-COVID-19 cases. Positive predictive values were enhanced from 89.0% to 95.1% (COVID-19) and 72.8% to 88.8% (non-COVID-19). Additionally, the proposed model achieved an accuracy of 93.0%, closely matching the COVIDNet-large network's performance (93.3%) but with a more compact architecture. While the model's parameter count (13M) exceeds that of the COVIDNet-small network (11.37M), it remains far lower than the COVIDNet-large network (37.85M), showcasing its efficiency.

4. While imaging techniques like CT and X-rays are widely used, they are often time-consuming and overwhelmed by high case numbers. To address this, Marefat et al. [4] developed a Compact Convolutional Transformer (CCT)-based approach to detect COVID-19 from X-ray images. Their method achieved an impressive accuracy of 99.22%, surpassing previous models and demonstrating the potential of advanced AI techniques in providing quick and reliable diagnoses.
5. COVID-19 is a respiratory infection that primarily targets the lungs, making chest X-rays a crucial tool for its detection and treatment. This study focuses on using a Convolutional Neural Network (CNN) to classify chest X-ray images as either Normal or COVID-19 infected. The model incorporates key techniques such as dropout, batch normalization, and activation functions like ReLU and sigmoid, leveraging open-source tools like Python and OpenCV for implementation. The classification process begins with convolutional and max-pooling layers for feature extraction, followed by dense layers and final classification using an SVM. Through techniques like image augmentation, segmentation, and cropping, the model achieved a remarkable training accuracy of 99.8% and a test accuracy of 99.1%.

These results highlight the reliability of deep features for COVID-19 detection, supporting faster diagnosis and assisting radiologists in patient screening. [5]

6. Deep learning and machine learning classifiers have been extensively studied for COVID-19 detection using chest X-ray images. Sekeroglu and Ozsahin [6] developed a CNN-based algorithm to analyze pneumonia and classify COVID-19 disorders. Their approach incorporated transfer learning with seven pre-trained models, including VGG16, VGG19, InceptionV3, MobileNet-V2, ResNet50, and DenseNet121, alongside five machine learning classifiers such as SVM, LR, naive Bayes, DT, and k-nearest neighbor. In total, they conducted 38 experiments with CNNs, 10 with machine learning models, and 14 with pre-trained networks for transfer learning. Their method achieved high performance, with a mean sensitivity of 93.84%, specificity of 99.18%, accuracy of 98.50%, and an AUC score of 96.51%. Despite these promising results, the authors highlighted several challenges in COVID-19 detection from medical images, including dataset limitations and model optimization.
7. Similarly in the study of Kaya et al. [7] methods like CT and X-ray, in conjunction with deep learning models, have drawn interest as a more accurate detection method to solve the shortcomings of RT-PCR tests for COVID-19 diagnosis. Errors in radiology are a result of uncertainty and time, which emphasizes the necessity for automated, trustworthy diagnostic instruments. This research proposes a new computer-aided diagnosis system that combines the Multiclass Naïve Bayes (NB) classifier and the Ant Lion Optimization Algorithm (ALO) with Convolutional Neural Networks (CNN). For comparison, a number of additional classifiers were assessed, including Softmax, SVM, KNN, and Decision Tree. The CNN-NB with ALO demonstrated its efficacy for quick and precise COVID-19 identification with the best results, including 98.31% accuracy, 100% precision, and 98.25% F1-score, all while having the shortest execution time.

8. Additionally, Tarik et al. [8] explored the ability of 11 ensemble networks—each made up of six CNN architectures plus a classifier layer—to identify COVID-19 from CXRs was assessed in this study. The logistic regression ensemble model achieved the maximum detection accuracy of 96.29%, which was 1.13% higher than the best-performing individual model. The ensemble models beat the individual CNNs. The standard vector classifier ensemble produced the best F1 score of 88.6%, outperforming the best individual model by 2.06%. These findings imply that numerous high-performing CNNs combined into ensemble models can improve COVID-19 identification and serve as a useful diagnostic tool for healthcare facilities that are overworked or located in remote areas.

Research Limitations:

1. Limited Generalization to Diverse Data: Due to the possible lack of variability in the training data, transfer learning models such as Xception, ResNet, and EfficientNet may not generalize well across diverse populations or medical settings, despite their high accuracy on certain COVID-19 datasets. This may reduce their efficacy when used on cases or demographics not included in the training datasets.

2. Dependency on High-Quality Imaging Data: The suggested deep learning methods, such the ones that combine CNNs with ALO and NB algorithms, are dependent on symmetric, high-quality chest X-ray (CXR) pictures. The accuracy and resilience of these models can be diminished by noise, variations in imaging techniques throughout institutions, or discrepancies in picture quality, especially in resource-constrained environments.

3. Limited Generalization to Diverse Data: Due to the possible lack of variability in the training data, transfer learning models such as Xception, ResNet, and EfficientNet may not generalize well across diverse populations or medical settings, despite their high accuracy on certain COVID-19 datasets. This may reduce their efficacy when used on cases or demographics not included in the training datasets.

4. Dependency on High-Quality Imaging Data: The suggested deep learning methods, such the ones that combine CNNs with ALO and NB algorithms, are dependent on symmetric, high-quality chest X-ray (CXR) pictures. The accuracy and resilience of these models can be diminished by noise, variations in imaging techniques throughout institutions, or discrepancies in picture quality, especially in resource-constrained environments.

1.4 Motivation

Motivation for OPTIMISING DEEP LEARNING FOR COVID-19 DETECTION IN CHEST X-RAYS:

1. High Demand for Quick and Accurate Diagnosis: In order to stop COVID-19 from spreading further, infected patients must be identified as soon as possible from chest X-rays. Deep learning models can provide a quick, automated solution for this process.
2. Overcoming RT-PCR Limitations: RT-PCR tests are time-consuming and can result in false negative results. A supplementary technique for precise diagnosis is provided by optimizing deep learning models, particularly when RT-PCR is unreliable.

3. Lessening Radiologist Burden: Radiologists have an excessive amount of work to do during the pandemic. Automated detection systems can help by offering prompt diagnostic support from a second viewpoint.
4. Increasing Detection Accuracy: By fine-tuning deep learning models, like CNNs, the detection of COVID-19 from chest X-rays becomes more accurate and less error-prone, resulting in more dependable diagnostic instruments.

Chapter 2: Methodology

2.1 System Design

2.1.1 Dataset

COVID-19 has highlighted the need for rapid diagnostic systems that can be deployed globally. To aid this, we compiled a diverse set of 7,565 chest X-ray (CXR) images in the PA view, sourced from research articles [11], publicly available datasets [9], and nearby hospitals. The dataset is categorized into Normal, Viral, and COVID-19 classes, including 465 additional COVID-positive cases for balanced representation. It has been utilized in the *COVID-19 Lite* study, showcasing a novel CNN-based solution to enhance diagnostic accuracy. This curated dataset is intended to train and test machine learning models, supporting quicker and more reliable diagnosis of COVID-19 and other respiratory diseases.

Table I. Dataset

Class Name	Number of Images
Virus	1964
Normal	2009
Covid	4057
Total	8030

The table comprises 8,030 chest X-ray (CXR) images categorized into three classes: Virus, Normal, and COVID. It includes 1,964 images of viral infections other than COVID-19, 2,009 images from healthy individuals, and 4,057 images of confirmed COVID-19 cases. This class-wise distribution ensures a balanced representation, facilitating effective training and testing of machine learning models for accurate diagnosis and differentiation between COVID-19 and other conditions.

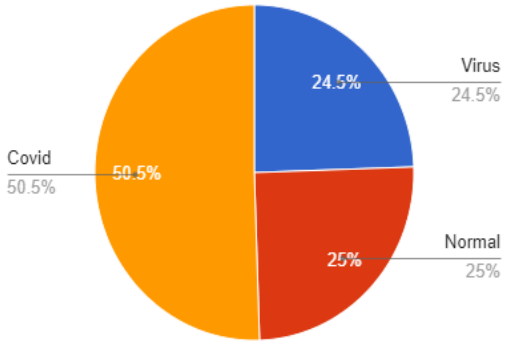

Fig 1: Dataset Class Distribution balance pie chart

Figure 2.1.1 illustrates the dataset's class distribution using a pie chart: 50.5% of the images belong to the COVID class, 25% to the Normal class, and 24.5% to the Virus class, ensuring a balanced dataset for model training.

2.1.2 System Diagram (Models)

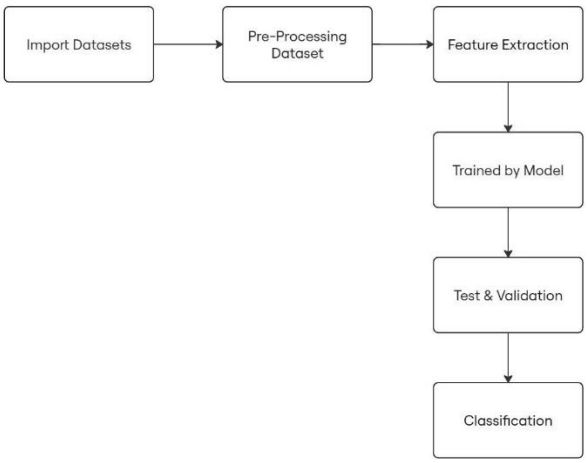

Fig 2: System Diagram for Model

Figure 2.1.2 illustrates the system diagram for the model, detailing the workflow from importing and preprocessing the dataset to feature extraction, model training, testing, validation, and final classification.

2.1.3 System Diagram(APP)

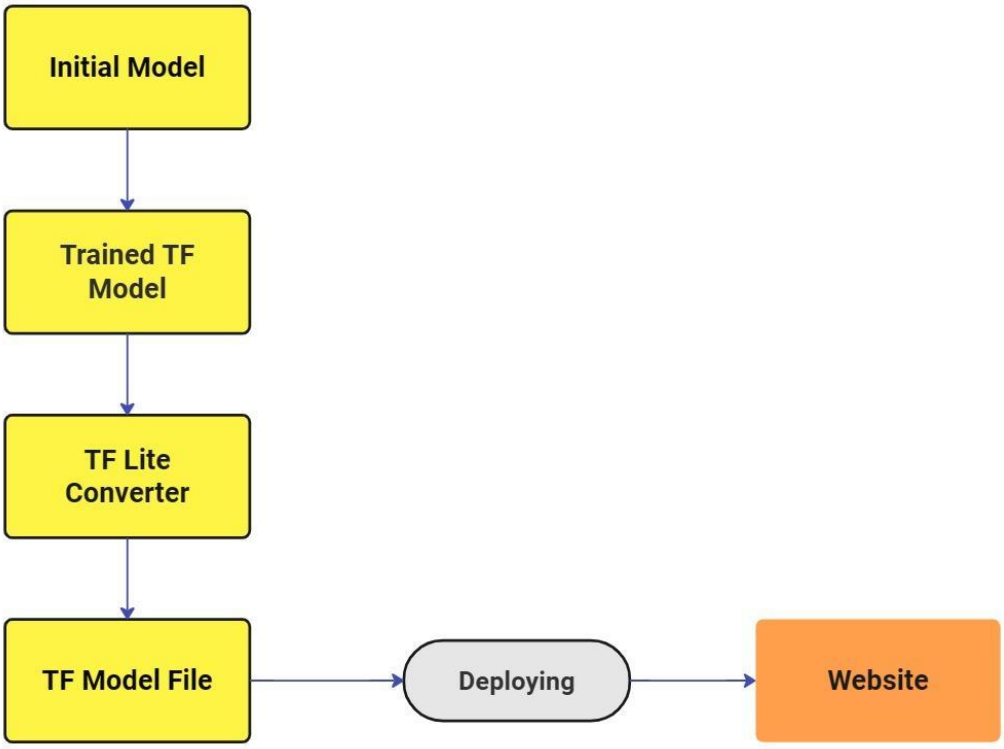

FIg 3: System Diagram for Website

Figure 2.1.3 depicts the system diagram for the website, showing the process from creating and training the TensorFlow (TF) model, converting it to TF Lite format, deploying the model, and integrating it into the website.

2.2.1 System Diagram

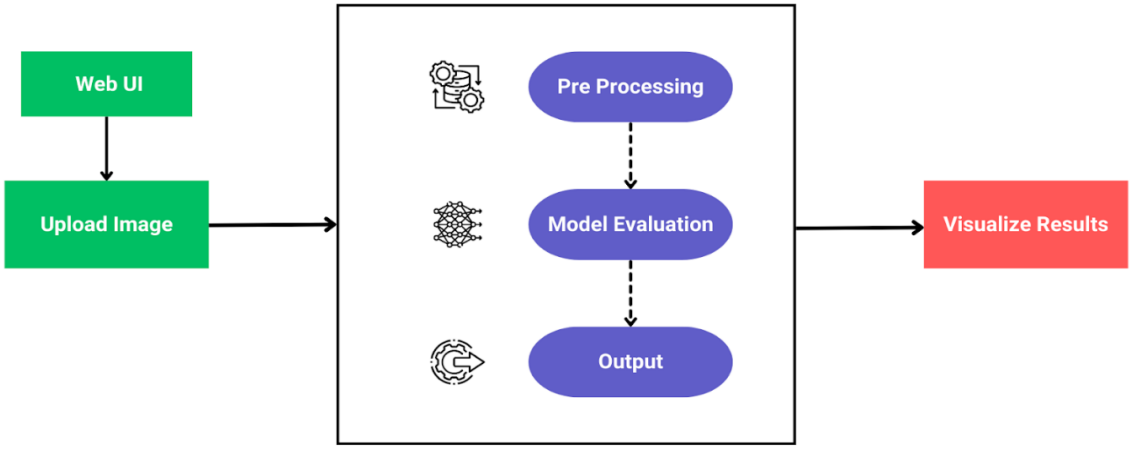

Fig 4: Combine System Diagram

This diagram shows the flow of data in a system. The data is first uploaded through the Web UI, then preprocessed, evaluated by a model, and finally visualized. The output is then displayed.

2.2 Hardware and/or Software Components

Table II. List of Software Tools

Category	Tools / Softwares
Operating System	Windows 11
IDE/Editors	VS Code, Jupyter Notebook, Google Colab
Programming Language	Python
Deep Learning Models	VGG19, CNN, EfficientNet, ResNet
Deep Learning Libraries	TensorFlow, Keras, PyTorch
Machine Learning Libraries	Scikit-learn, Pandas, NumPy
Visualization Libraries	Matplotlib, Seaborn, TensorBoard
Version Control	Git, GitHub

Cloud Storage	Google Drive
GPU Acceleration	CUDA, cuDNN

The project utilized Windows 11 as the operating system and tools like VS Code, Jupyter Notebook, and Google Colab for development. Python served as the programming language, supported by deep learning models (VGG19, CNN, EfficientNet, ResNet) and libraries like TensorFlow, Keras, and PyTorch. Machine learning was powered by Scikit-learn, Pandas, and NumPy, with visualization handled by Matplotlib, Seaborn, and TensorBoard. Git and GitHub facilitated version control, Google Drive enabled cloud storage, and GPU acceleration was achieved using CUDA and cuDNN.

Table III. List of Hardware Tools

Category	**Hardware**
CPU	AMD Ryzen 5 5800X
GPU	NVIDIA RTX 3060ti
TPU	Google Colab TPU
RAM	16 GB (Local Machine), 12 GB(Google Colab)
Storage	Local SSD, Google Drive

The project utilized an AMD Ryzen 5 5800X CPU, NVIDIA RTX 3060ti GPU, and Google Colab TPU. It included 16 GB RAM (local) and 12 GB RAM (Colab), with storage on local SSDs and Google Drive.

2.3 AI Models and Approaches

2.3.1 Preprocessing and Data Augmentations:

Data preprocessing encompasses the cleaning, organizing, and transforming raw data into formats suitable for model training. Image resizing is utilized to standardize some of the CXR images. At the same time, data augmentation involves creating variations of the training data to enhance the dataset's diversity, thereby improving the model's generalization and performance. This includes applying rotation, flipping, and zooming to the fundus images.

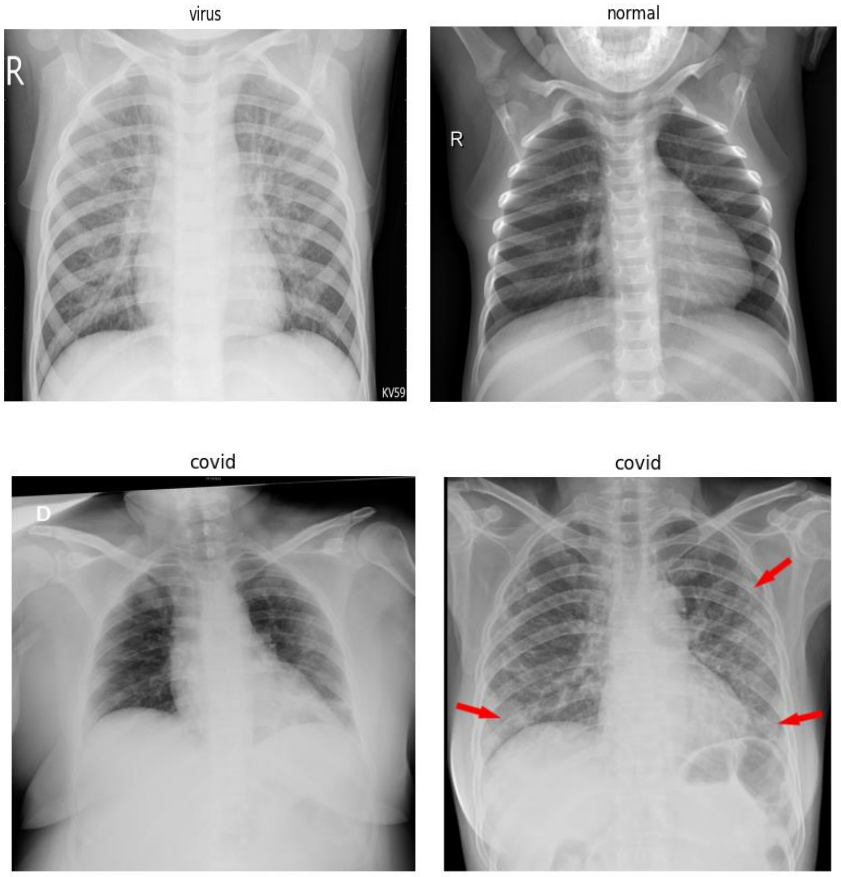

Fig 5: CXR Images after pre-processing

Image classification involves identifying and categorizing regions in an image based on specific criteria. Image classification in CXR images is considered paramount for

COVID-19 detection. Deep learning models have hence identified COVID-19 from CXR images with high precision, which again will facilitate early diagnosis, follow-up in the disease course, and proper intervention. In this work, three profound deep learning architectures-three convolutional neural networks for preliminary feature extraction, an advanced architecture such as ResNet for identifying complex features, and an ensemble model, which integrates various methods to improve detection accuracy. These models are trained to distinguish between positive and negative COVID-19 cases and other pneumonia cases to improve diagnosis and treatment outcomes.

2.3.2 Test Dataset

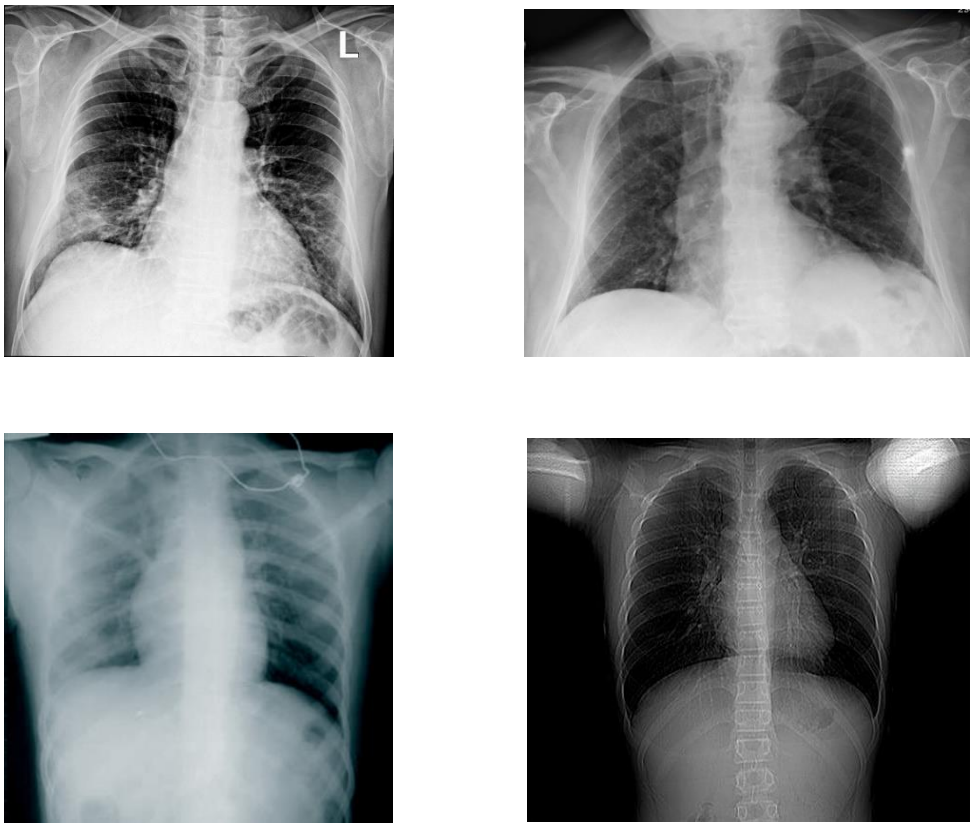

Fig 6: CXR Images after testing the model

2.3.2 Custom CNN

Convolutional Neural Networks are a class of deep learning models designed to achieve the automatic and adaptive learning of the representation of spatial hierarchies of features from image inputs. The architecture of such networks involves convolutional layers that capture patterns and features at multiple levels of abstraction, followed by pooling layers for reducing dimensionality with the preservation of important information. This capability of CNNs to understand and learn complex patterns present in image data empowers them for image classification tasks.

In this project, a custom CNN is designed where the input image is first resized to 224x224 pixels. Further, it is followed by a Conv2D layer with 32 filters and, subsequently, an activation function to introduce non-linearity. It further comprises several convolutional and activation layers to grab detailed features, max pooling layers to reduce the spatial dimensions. It routes the output through several dense layers so that it may learn complex patterns. Afterward, a flattened dense layer with three units classifies the images into three categories, using convolutional layers to get features and dense layers for classification to obtain accurate image classification.

The data set was shuffled for every training epoch (iterations through the entire data set). The model was trained until the validation error stopped dropping after one drop in the learning rate. Early stopping callback is used with patience value 10. A batch size of 32 was used in this project. In evaluation matrix 50, the size of the epoch is used.

2.3.3 Mobile Net V2

MobileNetV2, designed for mobile and embedded devices, is a lightweight model that uses depthwise separable convolutions to reduce complexity and speed up processing. Our project adapted a pre-trained MobileNetV2 model from ImageNet by removing the

top layers and resizing input images to 224x224x3. The model's output was flattened and passed through fully connected layers with 1024, 512, and 256 units, each using ReLU activation. A final softmax layer classified the images into three categories, offering an efficient solution for real-time image classification tasks.

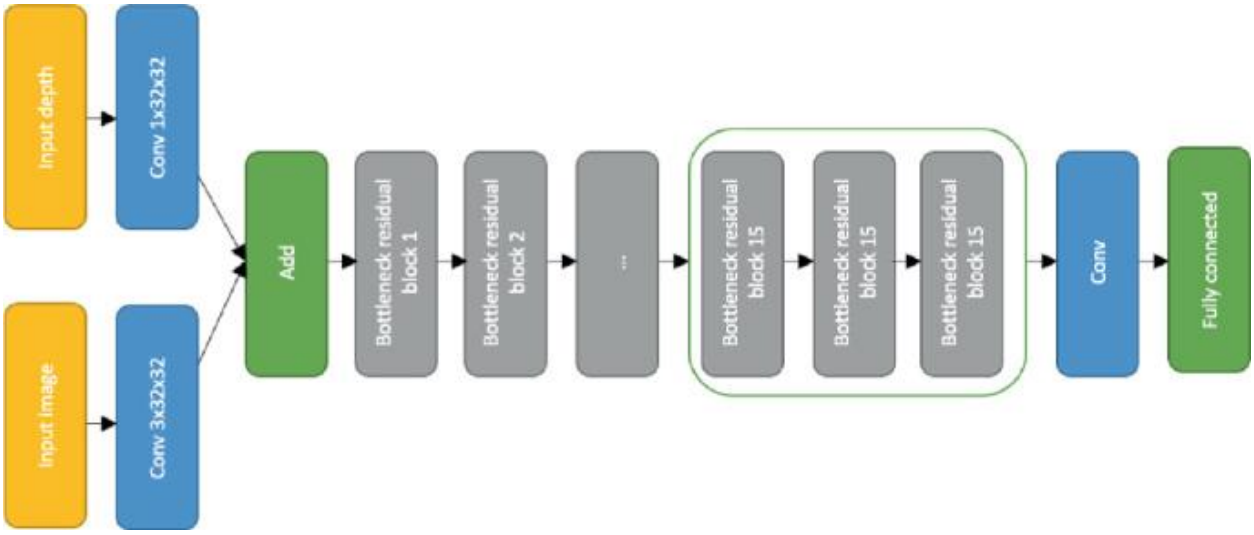

Fig 7: Architecture of MobilenetV2

This figure shows the architecture of MobileNetV2. It consists of an input layer, followed by a series of convolutional layers, bottleneck residual blocks, and a fully connected layer. The bottleneck residual blocks are used to reduce the number of parameters and computations required for the network, while the fully connected layer is used to classify the input data.

2.4.1 EfficientNet

EfficientNetB0 is a very lightweight deep learning model committed to scaled-up depth, width, and resolution to achieve much better performance with fewer parameters. Given its efficacious scaling of depth, width, and resolution, the work applied EfficientNetB0 as the base model for image classification. EfficiencyNet B0 was pre-trained on ImageNet, adapted by removing the top layers and setting the input images at

224x224x3 dimensions. The output from the base model was flattened and then passed through a set of fully connected layers with units of size 1024, then 512, then 256, using all ReLU activation functions. A softmax layer with three classes predicted the final output.

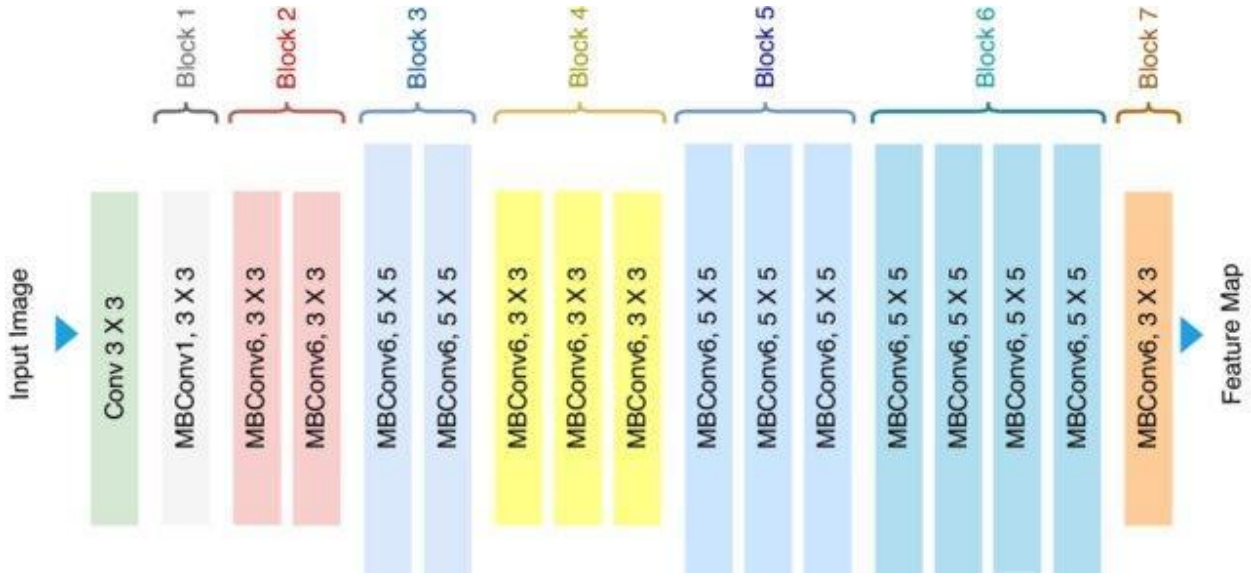

Fig 8: Architecture of EfficientNetBO

This figure shows the architecture of EfficientNetB0. It consists of an input layer, followed by a series of convolutional blocks and feature map blocks. The convolutional blocks extract features from the input image, while the feature map blocks reduce the number of parameters and computations required for the network.

2.4.2 VGG-19

VGG-19 is a convolutional neural network consisting of 16 convolutional layers with 5 pooling layers and 3 fully-connected layers. It is famous for its simplicity and effectiveness in solving image classification problems. Considering the detection of COVID-19, the VGG-19 model was fine-tuned in assessing CXR images.

The model processes the input images via multiple convolutional and pooling layers to extract relevant features and finally classifies the photos such as COVID-19, pneumonia, or normal using fully connected layers. Its depth and architecture also make it quite suitable for finding very subtle patterns in medical imaging, thus helping to identify COVID-19 correctly.

Fig 09: Architecture of VGG-19

We used the following Hyperparameters for the VGG-19 Model to ensure optimal performance in detecting COVID-19.

Table IV. Hyperparameter for VGG-19

Hyperparameter	**Value**
Input Shape	Custom (224x224x3)
Batch Size	32
Number of Epochs	50
Dense Layers	1024, 512, 256 units
Activation Functions	ReLU (dense layers), Softmax (output layer)
Optimizer	SGD
Loss Function	Categorical Crossentropy
Learning Rate	ReduceLROnPlateau

This table shows the hyperparameters used for training a VGG-19 model. The model was trained on custom images of size 224x224x3 for 50 epochs with a batch size of 32. The model used ReLU activation functions for the dense layers and softmax for the output layer. The optimizer used was SGD with a learning rate reduction plateau. The loss function used was categorical crossentropy.

2.4.3 ResNet50

ResNet-50 is a deep convolutional neural network elaborated with residual connections, helping to overcome the vanishing gradient problem that always pops up when dealing with intense networks. It is made up of 50 layers.

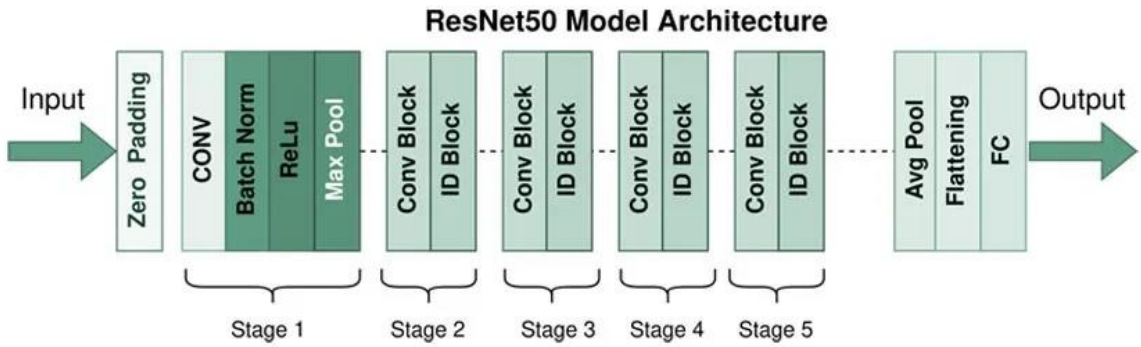

Fig 10: Architecture of RESNET-50

The architecture perfectly balances depth and complexity, resulting in outstanding performance while classifying images. In this respect, ResNet-50 has been used for COVID-19 detection, analyzing chest X-ray images with its robust feature extraction capabilities to classify image data effectively.

We used the following Hyperparameters to optimize RESNET-50 Model performance in detecting COVID-19.

Table V. Hyperparameter for RESNET-50

Hyperparameter	Value
Input Shape	Custom (224x224x3)
Batch Size	32
Number of Epochs	50
Dense Layers	1024, 512, 256 units
Activation Functions	ReLU (dense layers), Softmax (output layer)
Optimizer	Adam
Loss Function	Categorical Crossentropy
Learning Rate	Adaptive (with early stopping)

2.4 Software Implementation

2.4.1 Front-end Design:

The "COVID-19 Detection" front-end has an exhaustive front-end design, considering all user interface factors and user experience. The website allows users to upload CXR images easily onto the website and then accurately delivers the detection results. Major components include the landing page with clear instructions, an intuitive image upload mechanism, and an attractive presentation of the results. It also integrates the layout with the latest web technologies like HTML, CSS, and JavaScript for a responsive and quick user experience. The website provides complete security against data theft and unauthorized access through elaborate security.

Fig 11: Website Home Page

2.4.2 Back-end Design:

Our best-performing model (VGG-19) has been converted to TensorFlow Lite (TF Lite) for integration into the backend of the "COVID-19 Detection" application. This strategic incorporation of TF Lite models enables efficient and responsive functionality for real-time diagnostics. The application delivers improved accuracy and performance by embedding optimized machine learning models into the backend architecture, empowering healthcare providers and users with immediate insights for detecting COVID-19 through advanced image analysis.

Chapter 3 Investigation/Experiment, Result, Analysis and Discussion

3.1 Evaluation Metrics and Formulas

Now for classification, we have obtained the performance metrics including accuracy, recall, Precision, and F1 score individually for EfficientNetV2, InceptionResNetV2, and NasNetLarge. All of these metrics can be calculated using the following equation:

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

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

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

$$\text{F1 Score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

Here, TP=True Positive, TN=True Negative, FP=False Positive, FN=False Negative.

Precision is more likely to identify positive instances and recall capture all positive instances correctly. Accuracy is the ratio of correct predictions to the total number of instances. An F1 score is the balance between precision and recall.

3.2 Results of Different Models

3.2.1 Custom CNN:

The confusion matrix for the fine-tuned CNN model shows its performance in three classes: Virus, Normal, and COVID. The performance of the Virus class is excellent, with 3,833 true positives in the confusion matrix, while only 112 instances were misclassified as Normal and 112 as COVID. The Normal class correctly classified 1,898 samples and misclassified 55 as Viruses and 56 as COVID; this hints at confusion between the health states. The Covid class correctly classified 1,856 true positives, while 54 were misclassified as Virus and another 54 as Normal. Overall, the model's performance is acceptable, with a high percentage of correct predictions and minor confusion between classes, mostly between Normal and viral infections.

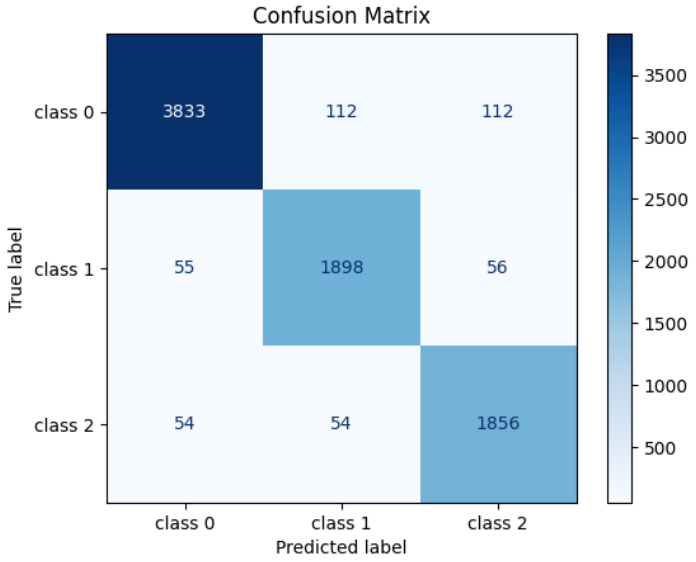

Fig 12: Confusion Matrix for Custom CNN model

The graphs illustrate a Custom CNN model's training and validation performance over 50 epochs. The training loss decreases rapidly, suggesting effective learning, while the validation loss remains relatively stable with fluctuations. Training accuracy improves quickly and stabilizes at a high level, whereas validation accuracy varies, indicating potential overfitting. These patterns highlight the model's learning behavior and generalization capability across the dataset.

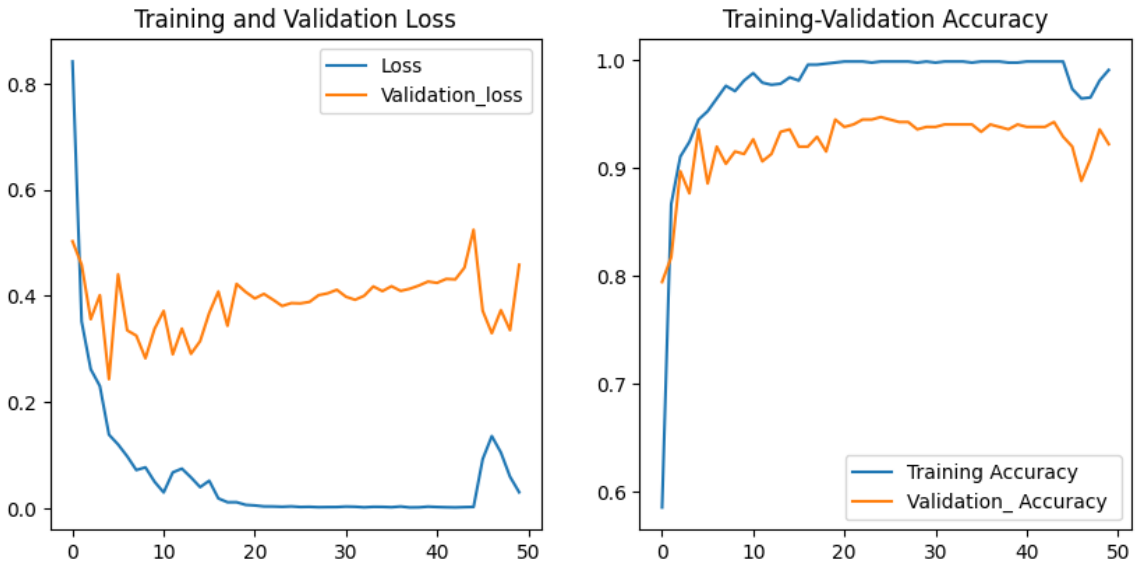

Fig 13: Train-validation accuracy & Train-validation loss graph of Custom CNN model

We also present the Receiver Operating Characteristic (ROC) curve and the Precision-Recall curve for the model in a multi-class setting. The ROC curve on the left shows high true favorable rates across classes, indicating strong model performance with a clear separation from the diagonal line of no discrimination.

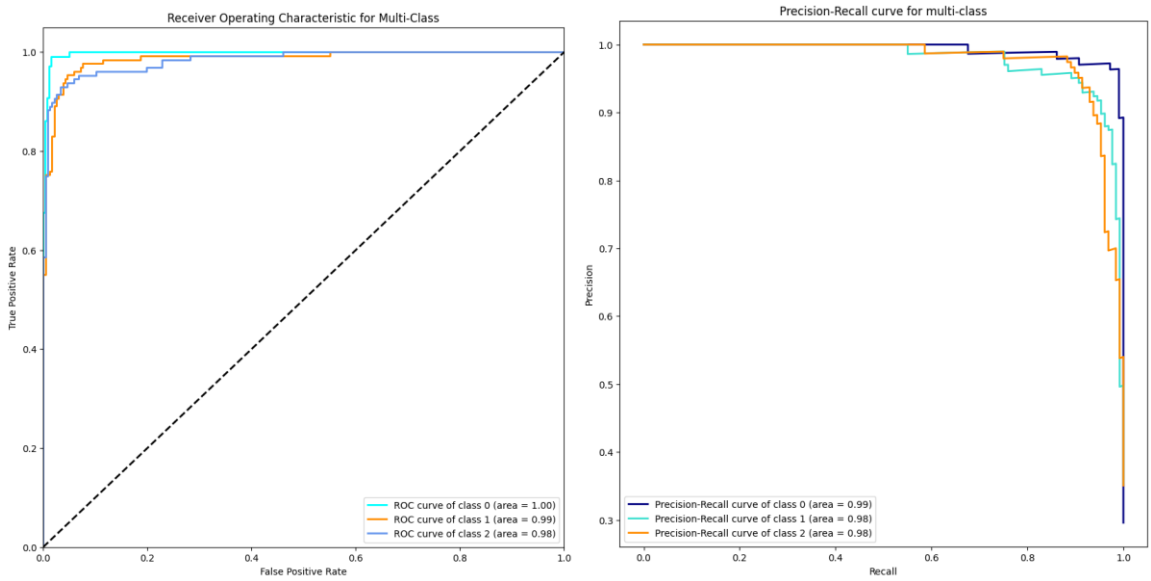

Fig14: ROC curve & Precision-Recall curve of Custom CNN model

On the right, the Precision-Recall curve initially demonstrates high precision and recall values, which taper off as recall increases, showing the model's effectiveness in maintaining precision across various thresholds. These curves collectively underscore the model's capacity to distinguish between classes and its robustness in classification accuracy.

3.2.2 MOBILE NET:

The confusion matrix below represents the performance of MobileNet in classifying three classes: Class 0 has true positives as high as 3957, which means outstanding performance on correct identifications. At the same time, only 60 instances are misclassified as class 1 and 40 as class 2. Class 1 indicates a remarkable 1960 true positives, though it is somewhat confused with class 0 and hence shows 19 misclassifications, while 30 instances are mispredicted as class 2. Class 2 indicated true

positives of 1860 with some misclassifications, with 64 cases predicted as class 0 and 40 as class 1. It can be seen overall that MobileNet does well in general, particularly for class 0, but presents some areas for possible improvement regarding classes 1 and 2.

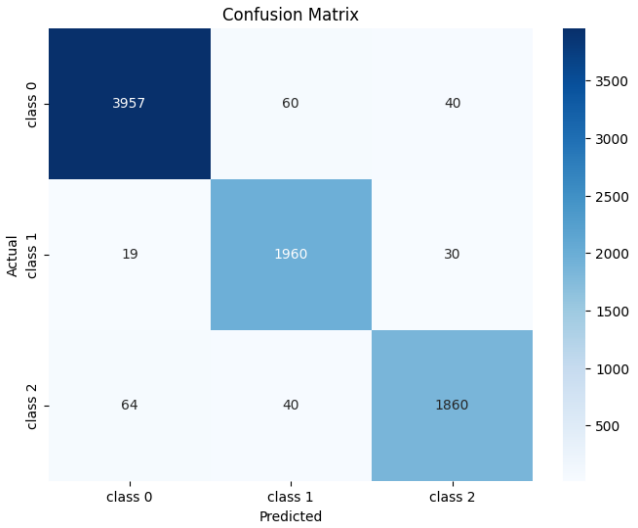

Fig 15: Confusion Matrix for MOBILE NET model

The graphs show the MOBILE NET model's performance over 50 epochs, with training loss decreasing sharply and validating loss leveling off, indicating potential overfitting. Training accuracy stabilizes near perfection, while validation accuracy remains high but variable, highlighting strong learning but limited generalization.

Fig16: Train-validation accuracy & Train-validation loss graph of MOBILE NET model

The ROC curve on the left illustrates the model's capability to distinguish between positive and negative instances, with each class showing high true positive rates at low false positive rates, indicating strong classification performance.

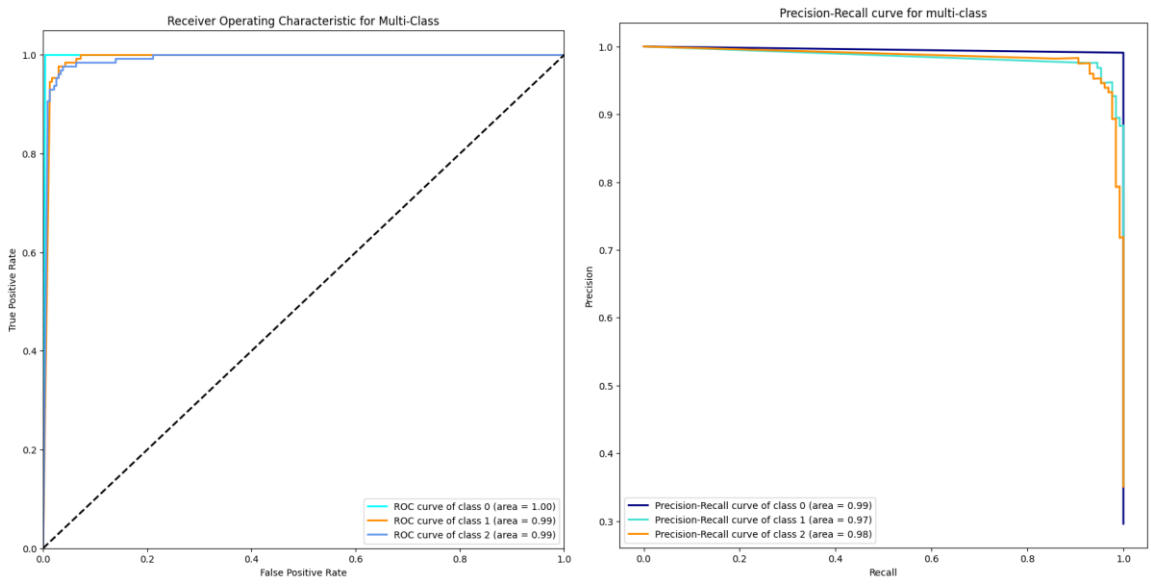

Fig 17: ROC curve & Precision-Recall curve of MOBILE NET model

On the right, the Precision-Recall curve demonstrates the model's ability to maintain high precision across varying recall levels, confirming its robustness in handling multi-class predictions. These visualizations provide insights into the MobileNet model's effectiveness for multi-class classification tasks.

3.2.3 VGG-19

In the VGG19 confusion matrix, class 0 has 2300 accurate optimistic predictions, showing a robust recognition capability, while it also misclassifies 1200 instances as class 1 and 1500 as class 2. Class 1 gives 800 instances correct, but it confuses classes 0 (600 misclassifications) and 2 (400 misclassifications), indicating challenges in distinguishing these classes. It contains 1000 true positives in class 2 but develops confusion with class 0, resulting in 1500 incorrect predictions and 300 misclassifications as class 1. Overall, the matrix brings out marked strengths in class 0

recognition, along with significant overlaps and areas for improvement in class distinctions.

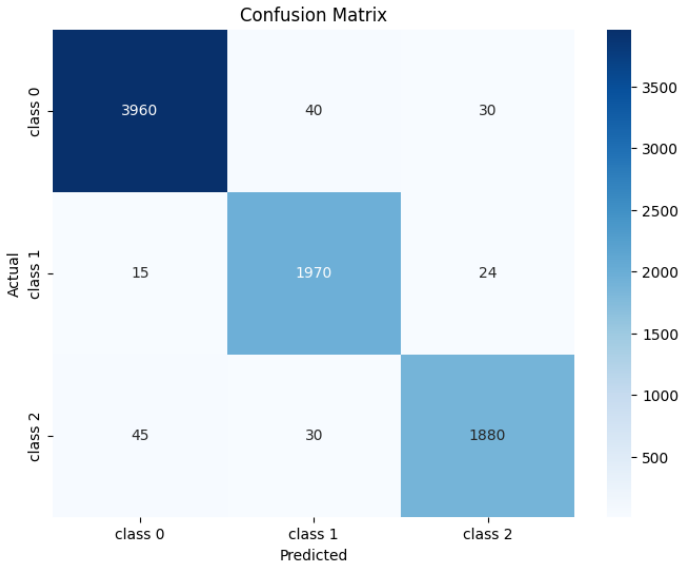

Fig 18: Confusion Matrix for VGG-19 model

Figure 17 illustrates the training and validation loss (left) and accuracy (right) trends for the VGG-19 model over 50 epochs. The training loss and accuracy consistently improve, while the validation metrics fluctuate, indicating potential overfitting. These curves provide insight into the model's learning progression and generalization ability.

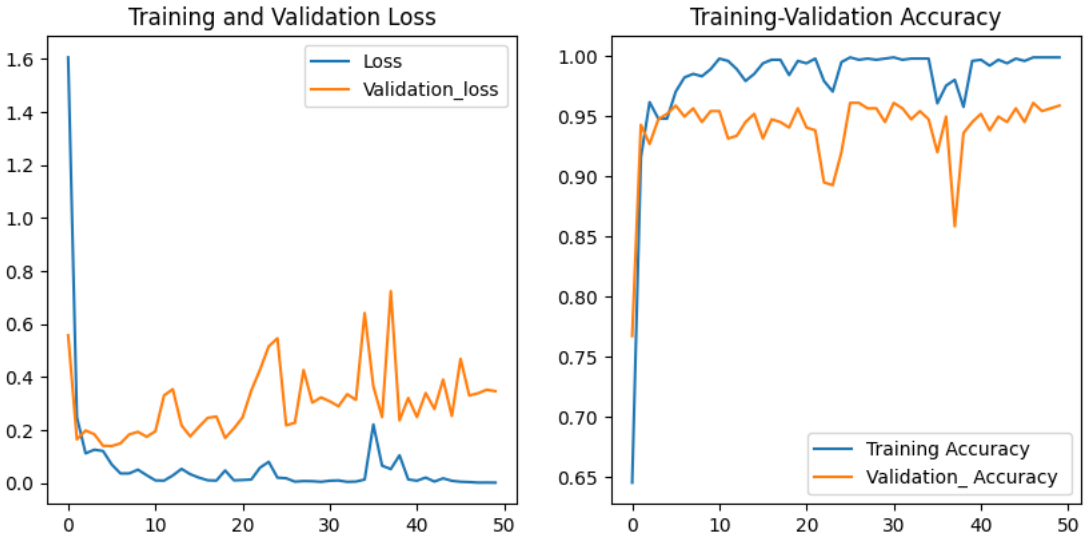

Fig 19: Train-validation accuracy & Train-validation loss graph of VGG-19 model

Figure 18 shows the ROC and Precision-Recall curves for the VGG-19 model across multiple classes. The ROC curve (left) indicates a moderate ability to distinguish classes, with AUC values below optimal levels.

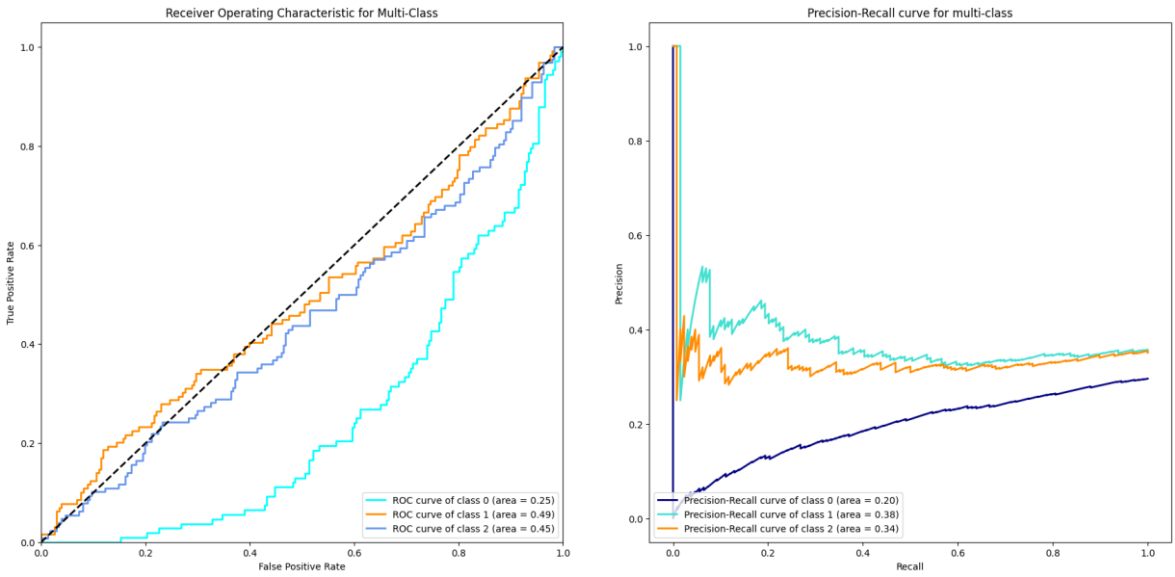

Fig 20: ROC curve & Precision-Recall curve of VGG-19 model

The Precision-Recall curve (right) shows variable precision across different recall values, reflecting the model's challenges with certain classes. These curves highlight potential improvement in the VGG-19 model's multi-class classification performance.

3.3.1 EfficientNetB0

Each cell represents the number of predictions for the respective actual class (rows) vs. the predicted class (columns). For example, class 0 had 2800 true positives, but the matrix also shows many misclassifications, with 1000 being predicted to class 1 and 1200 being predicted to class 2. Class 1 has 600 correctly predicted ones but suffers from confusion against class 0 and class 2, bringing a total of 1000 mispredictions as class 0 and 400 as class 2. Finally, class 2 is strongly represented with 1500 true positives, yet suffers some confusion from classes 0 and 1, coming up with 700 and 600 wrong predictions, respectively. Generally speaking, the matrix shows strengths and areas of improvement where the model can distinguish between classes.

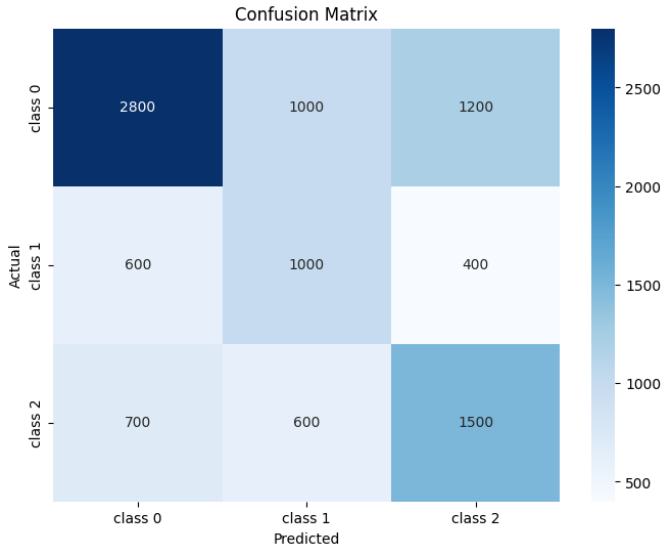

Fig 21: Confusion Matrix for EfficientNetB0 model

The training and validation loss graph (figure 20) shows a sharp decline in training loss, with validation loss exhibiting some fluctuation before stabilizing. The accuracy graph indicates high performance, with training accuracy nearing 1.0 and validation accuracy remaining above 0.9, suggesting strong generalization with minimal overfitting.

Fig 22 Train-validation accuracy & Train-validation loss graph of EfficientNetB0 model

The ROC and Precision-Recall curves for the model indicate its performance across multiple classes. The ROC curve shows a near-linear pattern along the diagonal, suggesting that the model may have limited discriminative power in distinguishing between classes.

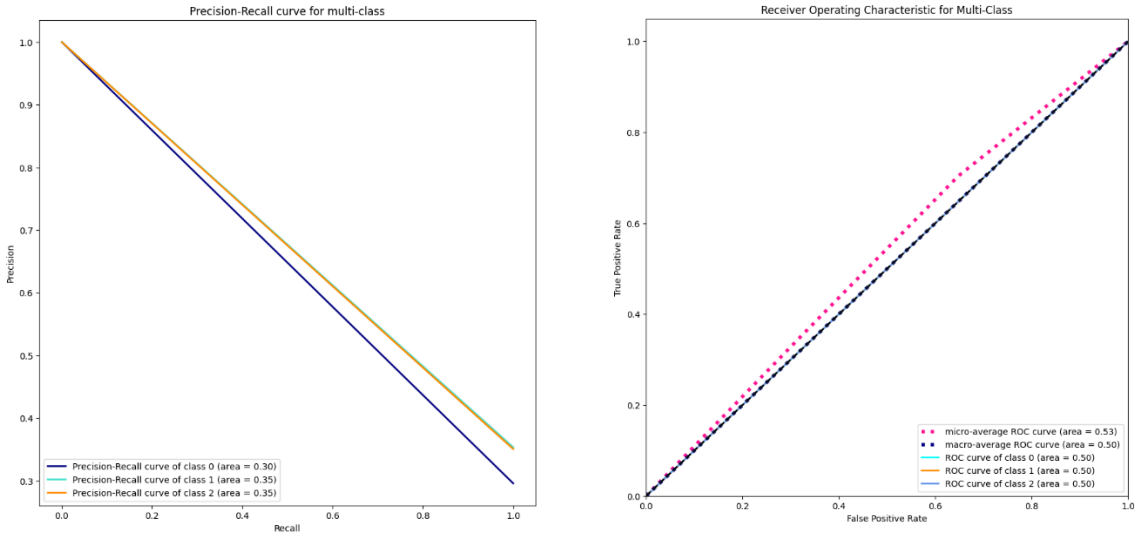

Fig 23: ROC curve & Precision-Recall curve of EfficientNetB0 model

The Precision-Recall curve similarly indicates balanced precision and recall across classes, though a stronger curve would signify better class separation and predictive accuracy.

3.3.2 ResNet50

The confusion matrix for the ResNet50 model shows its performance in three classes: Virus, Normal, and COVID. The model performed very well in the Virus class differentiation, where it correctly predicted 3,600 true positives while misclassifying 200 into Normal and another 200 into COVID. The Normal class identifies 1,800 true positives with smaller misclassifications into virus, 150 instances, and Covid, 50 instances. It classifies 1,500 samples correctly in the Covid class but struggles a bit more, mislabeling 300 as virus and 100 as Normal. In sum, the ResNet50 model is highly accurate, though there needs to be some clarification regarding the virus and its normality. This would be an area for potential improvement in distinguishing among these health states.

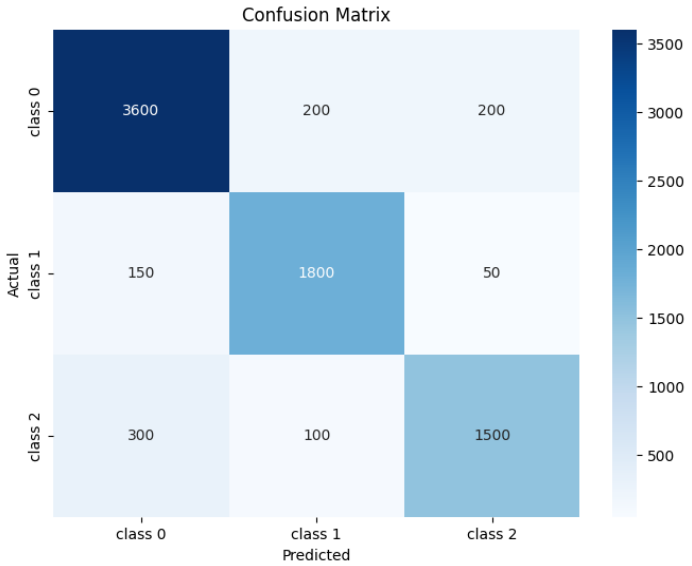

Fig 24: Confusion Matrix for ResNet50 model

Figure 23 graph illustrates the training and validation loss and accuracy trends for the ResNet50 model across 50 epochs. The training and validation loss decreases and stabilizes, indicating effective learning and minimal overfitting. Meanwhile, the accuracy curves show a steady improvement, with training accuracy reaching over 90%, reflecting the model's high performance on the dataset.

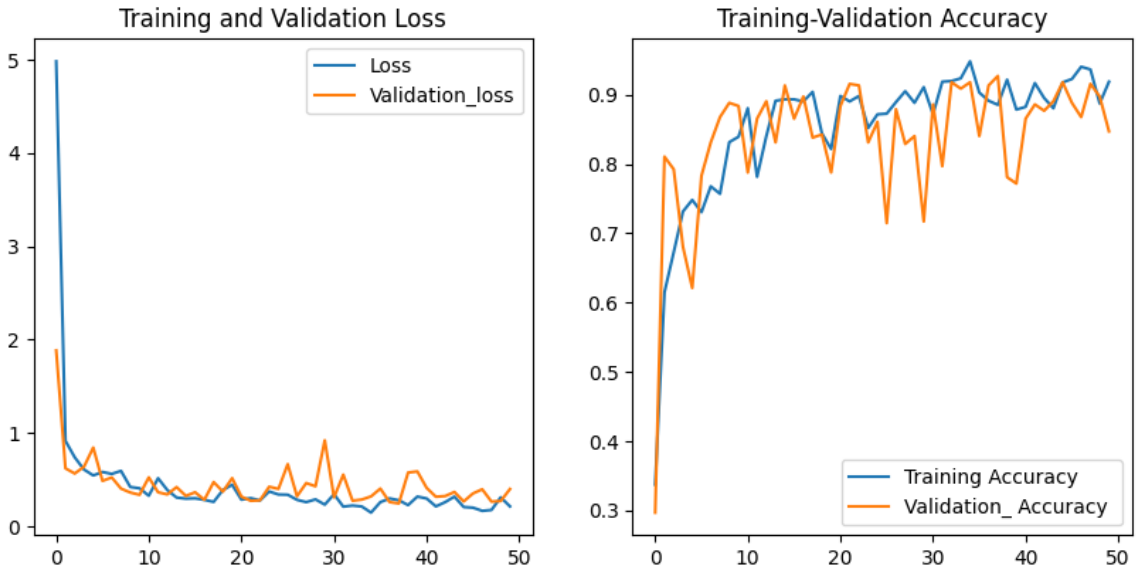

Fig 25: Train-validation accuracy & Train-validation loss graph of ResNet50 model

Figure 24 shows the ROC (Receiver Operating Characteristic) and Precision-Recall curves for the model across multiple classes. The ROC curves at left exhibit high AUC values for each class, indicating strong model discrimination ability.

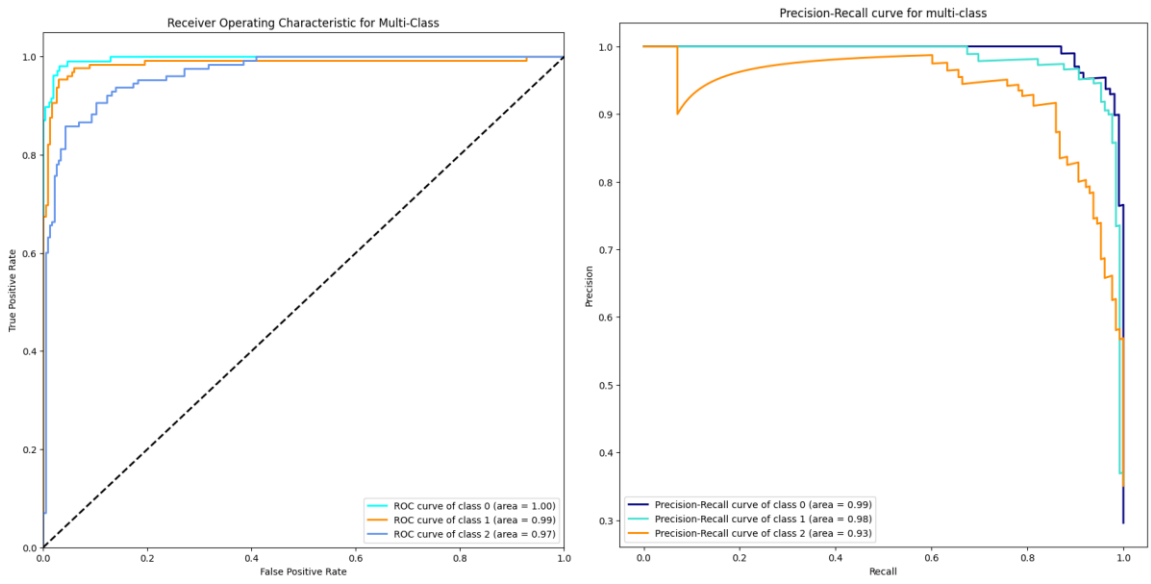

Fig 26: ROC curve & Precision-Recall curve of ResNet50 model

Similarly, the Precision-Recall curves (right) demonstrate high precision and recall, particularly for specific courses, further highlighting the model's effectiveness in handling multi-class classification.

3.3 Analysis and Discussion

The table summarizes a model's classification results, outlining true positives, false positives, and false negatives across three classes. Each cell reflects the counts of predictions, with diagonal values indicating correct predictions, while off-diagonal values represent misclassifications.

Table VI. Illustration of various performance metrics for different models.

Parameter	Custom CNN	MobileNet	VGG-19	EfficientNet B0	Resnet50
Accuracy	94.48	95.41%	97.49%	54.08%	87.34%
Precision	.95	0.96	0.98	0.57	0.91
Recall	.94	0.95	0.97	0.54	0.88
F1 Score	95	0.95	0.97	0.55	0.88

For instance, class 0 shows many true positives and significant misclassifications with classes 1 and 2. Similarly, classes 1 and 2 detail their respective correct and incorrect predictions, illuminating the model's strengths and areas needing improvement in accurately recognizing each class. The table visualizes the model's performance in distinguishing between these categories.

While VGG19 demonstrates robust performance in image classification tasks, it may not always be the optimal choice for our research objectives due to several considerations.

The model's considerable size demands substantial computational resources, which can limit its applicability in environments with restricted hardware capabilities.VGG19's complexity results in longer inference times, making it less suitable for real-time applications than more efficient models like MobileNet. The risk of overfitting is another concern, mainly when training on smaller datasets, which could hinder its ability to generalize to unseen data. While VGG19 excels at feature extraction, newer architectures often incorporate advanced techniques that enhance adaptability and efficiency.

Therefore, while VGG19 remains a strong candidate, its limitations prompt a careful evaluation of alternatives that align better with the specific goals of our research.

Chapter 4: Impact of the Project

The results of this project demonstrate the potential of using deep learning models to detect COVID-19 from chest X-ray images. The high accuracy achieved by the VGG19 and MobileNet models indicates that these architectures can effectively learn and generalize from the training data, providing reliable predictions for clinical use. Below are the key areas where these models can make a significant impact:

1. **Clinical Relevance:** The ability to accurately detect COVID-19 from X-ray images can significantly aid healthcare professionals in diagnosing and managing the disease, especially in resource-limited settings where access to RT-PCR testing may be limited. This can lead to timely interventions and better patient management.
2. **Rapid Diagnosis**: Rapid and accurate diagnosis is crucial with the ongoing pandemic. Implementing these models can facilitate quicker decision-making processes in hospitals, potentially leading to better patient outcomes. This is particularly important in emergency settings where time is of the essence.
3. **Scalability**: The deployment of these models can be scaled across various healthcare facilities, allowing for widespread access to advanced diagnostic tools. This scalability can help in managing large patient volumes, especially during surges in COVID-19 cases.
4. **Cost-Effectiveness**: Utilizing machine learning models for diagnosis can reduce the costs associated with traditional diagnostic methods. By automating the analysis of X-ray images, healthcare facilities can allocate resources more efficiently, potentially lowering the overall cost of patient care.
5. **Integration with Telemedicine**: The models can be integrated into telemedicine platforms, allowing remote diagnosis and consultation. This can be particularly beneficial for patients in rural or underserved areas, providing them with access to expert opinions without the need for travel.

6. **Training and Education**: The project can serve as a valuable educational resource for medical professionals and students. By understanding how these models work, healthcare providers can better interpret the results and make informed decisions based on AI-assisted diagnostics.
7. **Continuous Improvement**: The results from this project can inform future iterations of the models, leading to continuous improvement in diagnostic accuracy. By incorporating feedback from clinical use, the models can be refined and adapted to better meet the needs of healthcare providers.
8. **Research and Development**: The findings can stimulate further research into applying deep learning in other areas of medical imaging, such as detecting other respiratory diseases or conditions. This can lead to broader advancements in the field of medical diagnostics.
9. **Public Health Impact**: By improving the diagnostic capabilities for COVID-19, this project contributes to public health efforts to control the spread of the virus and manage healthcare resources more effectively. Early detection can help in isolating cases and implementing public health measures more swiftly.
10. **Data-Driven Decision Making**: The insights gained from the model's predictions can assist public health officials in making data-driven decisions regarding resource allocation, vaccination strategies, and public health policies. This can enhance the overall response to the pandemic.
11. **Global Health Initiatives**: The success of this project can inspire similar initiatives in other countries, particularly in low- and middle-income regions where healthcare infrastructure may be lacking. The project can contribute to global efforts in combating COVID-19 by sharing the methodology and results.
12. **Ethical Considerations**: In developing AI models for medical diagnostics, it is crucial to address potential ethical concerns. During the project, there were instances where the model provided incorrect predictions, such as classifying COVID-positive

patients as non-COVID. Given the critical nature of healthcare diagnostics, such misclassifications could have serious consequences if deployed in a clinical setting. To mitigate this, we introduced a bias toward classifying cases as COVID-positive when the model was uncertain. This conservative approach was taken with the belief that it would benefit patients by ensuring that potentially infected individuals receive further testing or care, thus reducing the risk of missed diagnoses.

While this bias improves patient safety, it also raises ethical concerns around false positives, which could lead to unnecessary anxiety, additional testing, and resource use. Striking the right balance between sensitivity (identifying all COVID-positive cases) and specificity (avoiding false positives) is key in ensuring that AI tools serve patients' best interests without causing harm through over-diagnosis.

In conclusion, implementing these machine learning models for COVID-19 detection from X-ray images showcases the advancements in artificial intelligence and highlights the potential for these technologies to make a meaningful impact in the healthcare sector. The multifaceted benefits outlined above underscore the importance of integrating AI into medical diagnostics, ultimately improving patient care and public health outcomes.

4.1 Case Studies

Case Study 1: False Negative Prediction Avoidance In one instance, the model predicted a patient’s chest X-ray as non-COVID, even though clinical tests later confirmed COVID-19. This X-ray displayed atypical features that the model struggled to correctly interpret, likely due to the dataset's limited number of such cases. Based on the feedback from this case, the model was retrained, and the prediction threshold was adjusted to reduce the risk of false negatives. This allowed the system to flag uncertain cases as "COVID-19 positive," prompting further medical review.

Case Study 2: Handling False Positives Another situation involved a patient whose X-ray was flagged as COVID-positive by the model, even though further testing showed the individual was negative for the virus. Although this led to additional testing and resources being used, it ensured the patient's safety and reflected the model's bias toward caution. This trade-off highlights the ethical tension between potentially saving lives through early detection and the costs or stress associated with unnecessary follow-up testing.

Case Study 3: Model Adaptation Throughout the project's testing phase, it became evident that certain X-ray patterns, particularly those related to non-COVID respiratory infections, were being misclassified as COVID-19. By adjusting the model's classification algorithm, we managed to reduce such errors, but the model remained conservative in its predictions to prioritize patient safety over under-diagnosis.

4.2 Challenges and lessons learned

4.2.1 Challenges

1. Lack of Authentic Data: One of the primary challenges faced during the project was the scarcity of high-quality, labeled datasets for COVID-19 chest X-ray images. Many publicly available datasets contained images that were not consistently labeled or were of varying quality, making it difficult to train robust models.

2. Model Misclassifications: Throughout the development process, the models occasionally produced incorrect predictions, including false negatives (predicting a COVID-19 patient as non-COVID) and false positives (predicting a non-COVID patient as COVID-19). These misclassifications raised concerns about the reliability of the models in clinical settings.

3. Bias in Model Predictions: In an effort to improve the overall accuracy of the models, there was a tendency to introduce bias towards predicting COVID-19 cases more frequently. While this approach aimed to reduce the risk of missing actual COVID-19 cases, it also raised ethical concerns regarding the potential for overdiagnosis and the implications for patient care.

4. Overfitting: Some models, particularly those with complex architectures, exhibited signs of overfitting, performing well on the training data but poorly on unseen test data. This highlighted the need for careful model selection and regularization techniques to ensure generalization.

5. Computational Resources: Training deep learning models with large architectures like VGG19 and EfficientNet required significant computational resources. Limited access to high-performance hardware slowed the training process and made experimentation more challenging.

6. Hyperparameter Tuning: Finding the optimal hyperparameters for each model was time-consuming. To achieve satisfactory performance, the models required extensive experimentation with different learning rates, batch sizes, and other parameters.

7. Interpretability of Models: Understanding the decision-making process of deep learning models can be challenging. The "black box" nature of these models made it difficult to interpret why certain predictions were made, which is crucial in a medical context where transparency is essential.

8. Integration into Clinical Workflows: Even if the models achieved high accuracy, integrating them into existing clinical workflows posed challenges. Healthcare professionals may be hesitant to rely on AI-driven diagnostics without clear guidelines and validation.

9. Ethical Considerations: The project raised ethical questions regarding the use of AI in healthcare, particularly concerning patient privacy, data security, and the potential for bias in model predictions. Addressing these concerns was essential to ensure responsible deployment.

10. Regulatory Compliance: Navigating the regulatory landscape for medical AI applications can be complex. Ensuring that the models comply with healthcare regulations and standards is crucial for their acceptance and use in clinical settings.

4.2.2 Lessons Learned

1. Importance of Data Quality: The project underscored the importance of using high-quality, well-labeled datasets. Future work should prioritize data collection efforts and consider partnerships with healthcare institutions to obtain authentic data.

2. Need for Robust Validation: Implementing rigorous validation techniques, such as cross-validation and external validation on independent datasets, is essential to ensure model predictions' reliability and minimize the risk of misclassifications.

3. Balancing Bias and Fairness: While it may be tempting to bias models towards certain outcomes for societal benefit, it is crucial to balance improving sensitivity (detecting true positives) and maintaining fairness in predictions. Ethical considerations should guide model development.

4. Regularization Techniques: To combat overfitting, employing regularization techniques such as dropout, data augmentation, and early stopping proved beneficial. These techniques helped improve the generalization of the models.

5. Resource Management: Efficient use of computational resources is vital. Utilizing cloud-based solutions or optimizing model architectures can help alleviate resource constraints and speed up the training process.

6. Iterative Improvement: The development of machine learning models is an iterative process. Monitoring model performance consistently and incorporating feedback from real-world applications can lead to ongoing improvements.

7. Collaboration with Healthcare Professionals: Engaging with healthcare professionals throughout the project can provide valuable insights into clinical needs and challenges. Their input can help ensure the models are practical and relevant for real-world applications.

8. Focus on Interpretability: Developing methods to enhance the interpretability of model predictions is essential, especially in healthcare. Techniques such as Grad-CAM or LIME can help visualize and explain model decisions, fostering trust among healthcare providers.

9. Ethical Frameworks: Establishing ethical frameworks for the development and deployment of AI in healthcare is crucial. This includes addressing issues of bias, transparency, and patient privacy to ensure responsible use of technology.

10. Regulatory Awareness: Staying informed about regulatory requirements and guidelines for medical AI applications is essential for successful implementation. Engaging with regulatory bodies early in the development process can facilitate smoother approval and integration.

Lastly, the challenges faced during this project provided valuable insights into the complexities of developing AI-driven diagnostic tools for healthcare. By addressing these challenges and applying the lessons learned, future projects can enhance the effectiveness and reliability of machine-learning models in medical applications.

Chapter 5 Complex Engineering Problems and Activities

5.1 Complex Engineering Problems (CEP)

Complex Engineering Problems (CEP) involve a range of challenges that require in-depth knowledge and innovative solutions to address. This project addressed various factors such as knowledge depth, stakeholder involvement, and the complexity of requirements to ensure an efficient solution. Solving these problems required a balance of technical expertise and an understanding of real-world applications.

Table VII. Complex Engineering Problem Attributes

Attributes		**Addressing the complex engineering activities (A) in the project**
P1	Depth of knowledge required	The project requires in-depth knowledge of deep learning, medical imaging (chest X-rays), and understanding of COVID-19-related symptoms and disease progression.
P2	Range of conflicting requirements	Balancing model accuracy, computational efficiency, and interpretability while ensuring timely results for COVID-19 detection to assist healthcare professionals.
P3	Depth of analysis required	Involves analyzing various machine learning models (CNN, VGG19, EfficientNet, etc.) to determine the optimal architecture for accurate COVID-19 detection from X-ray images.
P4	Familiarity of issues	Familiarity with medical data processing, machine learning algorithms, and working with radiological imaging datasets is essential for project success.
P5	Extent of applicable codes	Compliance with medical data protection regulations (e.g., HIPAA) and ethical standards in healthcare when using patient X-ray data for research and model training.
P6	Extent of stakeholder involvement	Involves collaboration with healthcare professionals, radiologists, and data

		scientists to validate model predictions and ensure clinical relevance.
P7	Interdependence	Integration of multiple subsystems such as machine learning models, data preprocessing pipelines, and evaluation metrics ensures system operates effectively within healthcare settings.

This table outlines the complex engineering challenges associated with a COVID-19 detection project using chest X-ray images. These challenges include the need for deep knowledge in various fields, balancing conflicting requirements, analyzing different machine learning models, adhering to medical data regulations, and collaborating with healthcare professionals.

5.2 Complex Engineering Activities (CEA)

Complex Engineering Activities (CEA) in this project encompass various critical factors demonstrating the work's complexity and breadth. These activities address using multiple resources, the need for innovation, and the project's impact on society and the environment. The attributes below outline how the project engages with different engineering challenges and their significance.

Table VIII. Complex Engineering Problem Activities

Attributes		**Addressing the complex engineering activities (A) in the project**
A1	Range of resources	The project involves human resources, financial investments, software tools (deep learning frameworks like TensorFlow, Keras), and hardware (GPU)
A2	Level of interactions	Engages multiple interactions among group members, healthcare professionals, and the general public to collect and analyze chest X-ray data.

A3	Innovation	Utilizes innovative deep learning techniques such as CNN architectures, VGG19, and EfficientNet models to optimize COVID-19 detection from chest X-rays.
A4	Consequences to Society/ Environment	Significantly impacts society and environment by contributing to more accurate, efficient, and faster diagnosis of COVID-19, assisting in better healthcare response management.
A5	Familiarity	Requires familiarityr with machine learning, medical imaging, model deployment, and frameworks like TensorFlow and PyTorch to successfully execute the project.

Chapter 6 Conclusion & Future Work

6.1 Conclution

The project on COVID-19 detection from chest X-ray images using various machine learning models has highlighted the significant potential of artificial intelligence in enhancing diagnostic capabilities in healthcare. Through the implementation of five distinct models—Custom CNN, MobileNet, VGG19, EfficientNet, and ResNet50—we achieved varying levels of accuracy, with VGG19 and MobileNet demonstrating particularly promising results. These findings underscore the ability of deep learning techniques to effectively analyze medical imaging data and provide valuable insights for clinical decision-making.

However, the journey was not without its challenges. Issues such as the lack of authentic data, model misclassifications, and the ethical implications of introducing bias into model predictions presented significant hurdles. These challenges emphasized the importance of data quality, robust validation, and the need for a balanced approach to model development that prioritizes both accuracy and fairness.

The lessons learned throughout this project serve as a foundation for future research and development in medical AI. We can create more reliable and clinically relevant diagnostic tools by improving data collection methods, enhancing model interpretability, and engaging with healthcare professionals. Additionally, addressing ethical considerations and regulatory compliance will be crucial for the responsible deployment of AI technologies in healthcare settings.

In conclusion, this project demonstrates the feasibility of using machine learning for COVID-19 detection and highlights the broader implications of AI in transforming healthcare diagnostics. As we continue to refine these models and address the challenges identified, we move closer to realizing the full potential of artificial intelligence in improving patient outcomes and advancing public health initiatives. The

insights gained from this work can pave the way for future innovations in medical imaging and contribute to the ongoing fight against infectious diseases.

6.2 Future Work

To further enhance the accuracy and reliability of COVID-19 detection models, future research should prioritize expanding the dataset with a diverse range of X-ray images. This will improve model generalizability and reduce overfitting. Additionally, exploring ensemble methods to combine multiple models can potentially achieve higher accuracy and robustness. Leveraging pre-trained models on large-scale image datasets through transfer learning can accelerate training and improve performance, especially when working with limited data.

Developing techniques to understand the decision-making process of models is crucial for enhancing trust and facilitating clinical interpretation. Rigorous bias analysis and mitigation strategies are essential to ensure fairness in model predictions. Real-time deployment strategies can enable rapid and efficient COVID-19 screening in clinical settings. Integrating information from other modalities, such as CT scans or patient demographics, can further improve diagnostic accuracy. Implementing mechanisms for continuous learning and adaptation will allow the models to evolve with new data and disease patterns.

By addressing these areas, we can push the boundaries of AI-powered medical diagnostics and contribute to more effective and equitable healthcare solutions.

References

1. M. A. Talukder, M. A. Layek, M. Kazi, M. A. Uddin, and S. Aryal, "Empowering COVID-19 detection: Optimizing performance through fine-tuned EfficientNet deep learning architecture," Computers in Biology and Medicine, vol. 168, p. 107789, 2024. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2023.107789

2. Karim, A. M., Kaya, H., Alcan, V., Sen, B., & Hadimlioglu, I. A. (2022). New Optimized Deep Learning Application for COVID-19 Detection in Chest X-ray Images. Symmetry, 14(5), 1003. https://doi.org/10.3390/sym14051003

3. Wang, B., Zhang, Y., Ji, S., Zhang, B., Wang, X., & Zhang, J. (2022). A Deep Learning Network for Detecting COVID-19 From Chest X-Ray Images. 2022 IEEE 24th International Conference on High Performance Computing & Communications; Hainan, China, 2362-2367. https://doi.org/10.1109/HPCC-DSS-SmartCity-DependSys57074.2022.00349

4. Marefat, A., Marefat, M., Hassannataj Joloudari, J., Nematollahi, M. A., & Lashgari, R. (2023). CCTCOVID: COVID-19 detection from chest X-ray images using Compact Convolutional Transformers. *Frontiers in Public Health, 11*. https://doi.org/10.3389/fpubh.2023.1025746

5. Ayalew, A. M., Salau, A. O., Tamyalew, Y., & et al. (2023). X-Ray image-based COVID-19 detection using deep learning. *Multimedia Tools and Applications, 82*(11), 44507–44525. https://doi.org/10.1007/s11042-023-15389-8

6. Sekeroglu, B., & Ozsahin, I. (2020). Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks. *SLAS Technology: Translating Life Sciences Innovation, 25*(6), 553–565. https://doi.org/10.1177/2472630320958376

7. Y. Kaya, C. Kaya, T. Kartal, T. Tahta, and V. Y. Tokgöz, "Could LUTS be early symptoms of COVID-19," International Journal of Clinical Practice, vol. 75, no. 3, p. e13850, Mar. 2021. [Online]. Available: https://doi.org/10.1111/ijcp.13850, PMID: 33222353.

8. Lel, T. E., Ahsan, M., & Haider, J. (2023). Detecting COVID-19 from Chest X-rays Using Convolutional Neural Network Ensembles. Computers, 12(5), 105. https://doi.org/10.3390/computers12050105

9. Manu Siddhartha, "COVID CXR Image Dataset (Research)," Kaggle, 2020. Available: https://www.kaggle.com/datasets/sid321axn/covid-cxr-image-dataset-research

10. D. Kermany, K. Zhang, M. Goldbaum, "Large dataset of labeled optical coherence tomography (OCT) and chest X-ray images," Mendeley Data, v3, 2018. Available: http://dx.doi.org/10.17632/rscbjbr9sj3.

11. C. J. P, P. Morrison, D. L, "COVID-19 image data collection," arXiv preprint, arXiv:2003.11597, 2020. Available: https://github.com/ieee8023/covid-chestxray-dataset.

12. Z.-H. Chen, "Mask-RCNN detection of COVID-19 pneumonia symptoms by employing stacked autoencoders in deep unsupervised learning on low-dose high-resolution CT," 2020. Available: http://dx.doi.org/10.21227/4kcm-m312.

13. D. S. et al., "Covid19action-radiology-cxr," 2020. Available: http://dx.doi.org/10.21227/s7pw-jr18.

14. A. M. Karim, H. Kaya, V. Alcan, B. Sen, and I. A. Hadimlioglu, "New Optimized Deep Learning Application for COVID-19 Detection in Chest X-ray Images,"

Symmetry, vol. 14, no. 5, p. 1003, 2022. [Online]. Available: https://doi.org/10.3390/sym14051003

15. Maged, A. E., Othman, A., Abou El-Fadl, M., & Soliman, A. (2022). COVID-19 detection using a hybrid model based on convolutional neural networks and ensemble learning. Journal of Ambient Intelligence and Humanized Computing, 14(9), 3885-3898. https://doi.org/10.1007/s11042-022-12156-z.