A

Project Report

on

Leveraging Machine Learning for COVID-19 Diagnosis

submitted as partial fulfillment for the award of

BACHELOR OF TECHNOLOGY DEGREE

SESSION 2024-25

in

Computer Science and Engineering

By

Mansi (2100290100091)

Nandini Vashistha (2100290100106)

Under the supervision of

Prof. Upendra Mishra

KIET Group of Institutions, Ghaziabad

Affiliated to

Dr. A.P.J. Abdul Kalam Technical University, Lucknow

(Formerly UPTU)

May,2025

DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Mansi	Nandini Vashistha
2100290100091	2100290100106

Date: May, 2025

CERTIFICATE

This is to certify that Project Report entitled “Leveraging Machine Learning for COVID-19 Diagnosis” which is submitted by Mansi and Nandini Vashistha in partial fulfillment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

.

Prof. Upendra Mishra

(Assistant Prof.-CSE)

Dr. Vineet Sharma

(Dean-CSE)

Date: May, 2025

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B. Tech. Final Year. We owe special debt of gratitude to Prof. Upendra Mishra, Department of Computer Science & Engineering, KIET, Ghaziabad, for his constant support and guidance throughout the course of our work. His sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Vineet Sharma, Dean-CSE, KIET, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members, especially faculty/industry person/any person, of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Date: May, 2025

Mansi

2100290100091

Nandini Vashistha

2100290100106

ABSTRACT

The COVID-19 pandemic has served as a global catalyst for the rapid transformation of healthcare technologies, highlighting the pressing need for more efficient, reliable, and scalable diagnostic tools to address an unprecedented global health crisis. Traditional methods of diagnosing COVID-19, such as RT-PCR tests, while effective, are often constrained by time, cost, and limited scalability, making them less suitable for widespread and rapid screening. As the pandemic progressed, healthcare systems worldwide faced the challenge of diagnosing millions of cases quickly and accurately, placing a considerable strain on existing resources. This highlighted the need for innovative technologies capable of providing faster, more efficient diagnostic solutions. In response, this study investigates the potential of artificial intelligence (AI), particularly machine learning (ML), to revolutionize the diagnostic process for COVID-19 and other respiratory diseases by leveraging medical imaging modalities, especially chest X-rays and CT scans.

The research presented in this study explores a hybrid machine learning model that integrates multiple advanced AI techniques to improve the accuracy and efficiency of COVID-19 diagnosis and severity prediction. By utilizing a pretrained ResNet50 model, the research benefits from transfer learning, allowing the model to leverage existing knowledge gained from ImageNet and apply it to the domain of medical imaging. This reduces the need for large annotated medical datasets, which are often scarce in the healthcare domain, thus accelerating the training process and enhancing model performance with minimal labeled data.

The findings from this study underscore the transformative potential of machine learning technologies in the realm of medical diagnostics, particularly in the context of the COVID-19 pandemic. The hybrid model demonstrated impressive performance across several evaluation metrics, including accuracy, precision, recall, and F1-score, outperforming traditional deep learning models in some instances. The model's ability to quickly and accurately diagnose COVID-19 cases from medical imaging data offers a scalable solution to aid healthcare systems in managing large patient volumes and making timely clinical decisions. Additionally, the study emphasizes the importance of integrating AI technologies into healthcare systems to optimize resource allocation, reduce human error, and improve patient outcomes, particularly

in low-resource settings. In conclusion, this research highlights the significant potential of AI, particularly machine learning models such as ResNet50 and ELM, in transforming the diagnostic landscape for COVID-19 and other respiratory diseases. By harnessing the power of advanced AI techniques, healthcare providers can deliver faster, more accurate diagnoses, improve patient care, and optimize healthcare resources. The findings advocate for the broader adoption of AI-driven medical imaging technologies, which could ultimately play a crucial role in enhancing global health outcomes, particularly in the face of ongoing and future health crises. As the world continues to battle COVID-19 and other global health challenges, this research provides valuable insights into the transformative role of AI in revolutionizing healthcare diagnostics and improving public health outcomes worldwide.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ResNet 50	Residual Neural Network
ELM	Extreme Learning Machine
CT	Computed tomography
COVID-19	coronavirus disease 2019
SARS-CoV-2	syndrome coronavirus 2
PCA	Principal component analysis
CXR	chest X-ray
CNN	convolutional neural network
ROC	Registrar of Companies
AUC	Area under the curve

CHAPTER 1

INTRODUCTION

1.1 The Global Impact of the COVID-19 Pandemic

The outbreak of the new coronavirus in late 2019 has had a sweeping and transformative effect on all areas of human existence. In Wuhan, China, this new variant of coronavirus originated and was quickly discovered and labeled SARS-CoV-2 because it has a similar genetic makeup to the SARS virus. The illness it produces, Coronavirus Disease 2019 (COVID-19), rapidly spread from a local epidemic to a worldwide pandemic, which was declared by the World Health Organization (WHO) in March 2020. As of mid-2025, COVID-19 has caused over 750 million reported cases and over 7 million confirmed fatalities globally. These alarming figures underscore not only the virulence and the contagiousness of the virus but also the need for scalable, precise, and swift diagnostic tools to deal with the crisis.

The pandemic has pushed the world's healthcare systems to the breaking point, laid bare the deep weaknesses in public health infrastructure, and unveiled disparities in access to medical resources. Under-resourced healthcare systems in countries bore the brunt of the lack of rapid testing centers and critical care facilities. In these situations, it became a necessity to develop effective diagnostic systems based on new technologies to detect infected individuals early, deliver timely medical treatment, and stem community transmission.

1.2 Need for Accurate and Rapid Diagnosis

Early and precise diagnosis is an important factor in preventing the transmission of infectious diseases. The RT-PCR test, which identifies viral RNA from patient samples, has been the most common method of diagnosis used for COVID-19. Despite being a gold standard, RT-PCR suffers from various shortcomings. It is labor-intensive, resource-dependent, and prone to errors caused by sample handling and timing of specimen collection. Furthermore, in regions with low testing infrastructure, the availability of RT-PCR kits and qualified personnel adds to the challenges. Consequently, the turn-around time for results could be extended by several days, which can be harmful during the peak infection periods.

In addition, RT-PCR can yield false negatives in asymptomatic or newly infected patients. The shortcomings of conventional diagnostic instruments led researchers to investigate complementary diagnostic modalities that are quicker, less expensive, and readily scalable.

Furthermore, RT-PCR may produce false negatives in patients who are either asymptomatic or in the early stages of infection. The limitations of traditional diagnostic tools prompted researchers to explore complementary diagnostic modalities that are faster, more cost-effective, and easily scalable.

1.3 Role of Medical Imaging and Artificial Intelligence

Medical imaging, particularly chest X-rays and CT scans, became useful for both diagnosis and COVID-19 monitoring. Radiographic features of the illness, including ground-glass opacities, bilateral infiltrates, and patchy consolidations, are visible in most symptomatic patients, especially those with moderate to severe symptoms. Nonetheless, the visual interpretation of these images by radiologists requires time and is subject to subjectivity. With regard to accuracy of a diagnosis, there may be differences due to the clinician's experience, the quality of the image, and the patient's overlapping respiratory symptoms to those caused by other illnesses like influenza, pneumonia, or tuberculosis.

This challenge paved the way for new possibilities of the use of artificial intelligence (AI) and machine learning (ML) in healthcare diagnostics. Deep learning, a form of AI, has shown outstanding success in pattern recognition and shown tremendous potential in medical image classification. Convolutional Neural Networks (CNNs) and some deep architectures like ResNet, VGGNet, and Inception are currently useful tools for applications including disease detection, disease segmentation, and anomaly detection from medical images. These networks can learn hierarchical features from raw data, requiring no manual feature engineering, and hence put them in a very good position for complex medical diagnosis problems.

1.4 Problem Statement and Research Motivation

- The knowledge that traditional diagnostic methods are limited and the availability of digital medical imaging is growing makes a strong rationale for the development of an intelligent, automatic, and scalable learning solution for the early diagnosis of COVID-19, this paper aims to develop a hybrid machine learning model that utilises the best aspects of deep learning and traditional learning, leveraging:
- ResNet50, a deep convolutional neural network, for strong feature extraction

- Principal Component Analysis (PCA), for reduced feature dimensionality and computational cost.
- Extreme Learning Machine (ELM), a fast and effective classifier with good generalization performance
- The motivation for this work is in addressing major diagnostic challenges of the pandemic:

1.5 Objectives of the Study

The key objectives of this study are as follows:

1. In order to propose a good COVID-19 diagnostic model that utilizes chest X-ray images and machine learning algorithms.
2. To investigate if ResNet50 can correctly extract high-quality deep features from chest radiographs.
3. To apply PCA to eliminate redundant features and improve computational efficiency.
4. To evaluate how effective the ELM classifier with pretrained and feature-converted features is.
5. To validate the hybrid model with a benchmark COVID-19 data set and compare its accuracy, precision, recall, F1-score, specificity, and ROC curve.

To contrast the proposed hybrid model with the existing CNN-based models based on performance metrics and computational complexity.

1.6 Contributions of the Research

The key contributions of this research are as follows:

- Hybrid diagnostic approach integrating deep learning and conventional machine learning components.
- Empirical validation of the proposed model on actual chest X-ray datasets.
- Experimental validation and calibration to sharpen model parameters.
- Baseline model comparison to demonstrate the superiority of hybrid feature fusion over conventional medical image classification models.
- Promotion of AI-based healthcare diagnosis, especially in resource-poor settings where radiological staff might be limited.

CHAPTER 2

RELATED WORK

The COVID-19 pandemic has significantly changed the direction of healthcare research, driving innovation in diagnostic technology, particularly in medical imaging. Timely detection and proper diagnosis of COVID-19 are essential to control patient outcomes, as well as to alleviate the burden on healthcare systems worldwide. To this end, machine learning (ML) and deep learning (DL) have emerged as indispensable tools for medical image interpretation of images like chest X-rays (CXR) and CT scans.

Various researches have investigated the use of AI methods in the detection, diagnosis, and estimation of severity of COVID-19, generally employing different models like CNNs, SVMs, and ELMs. In this paper, we present a review of the existing work on this topic, with emphasis on the employment of ML and DL methods for the diagnosis of COVID-19 from medical images. Deep Learning for Medical Imaging Deep learning techniques, and particularly convolutional neural networks (CNNs), have achieved unmatched success in analyzing medical images. CNNs, which learn automatically hierarchical features from images, have been commonly employed in medical imaging applications like disease classification, detection, and segmentation. The capability of CNNs to extract meaningful features directly from raw image data has made it revolutionize the field, particularly if the annotated dataset is small, a situation often faced in medical imaging. Perhaps the most popular CNN architecture employed in medical image analysis is the ResNet50 model. ResNet50 is a residual deep network that employs skip connections to prevent the vanishing gradient problem and performs well in training very deep models. Various research studies have utilized ResNet50 for COVID-19 detection from chest X-rays. For instance, a research study by Apostolopoulos and Mpesiana (2020) compared the application of different CNN architectures such as ResNet50 to classify COVID-19 from chest X-rays. It was discovered that ResNet50 performed better than other conventional machine learning algorithms and proved to be efficient in detecting features characteristic of the disease.

A study also created a COVID-19 detection model with a fine-tuned ResNet50 architecture. The model was trained on chest X-rays and showed high accuracy, sensitivity, and specificity.

Fine-tuning the pre-trained ResNet50 model using COVID-19 data allowed the network to learn to respond to the unique patterns of the disease, which gave better performance than scratch-trained models. The transfer learning approach—where one initially trains a model on a large dataset (e.g., ImageNet) and fine-tunes it on domain-specific data later—is gaining popularity in medical imaging to address the issue of limited labeled data.

Hybrid Models: CNNs and Other Machine Learning Algorithms

In addition to CNN-based models, several studies have explored hybrid approaches combining CNNs with traditional machine learning algorithms such as support vector machines (SVMs), random forests (RF), and extreme learning machines (ELMs). These hybrid models aim to combine the feature extraction capabilities of deep learning models with the classification strengths of traditional machine learning algorithms, leading to improved performance and reduced overfitting. A notable example is the work by Islam et al. (2020), which combined a CNN-based feature extraction network with an SVM classifier for detecting COVID-19 from chest X-rays. The CNN extracted features from the images, and the SVM classifier was used to distinguish between COVID-19, pneumonia, and healthy samples. This hybrid model achieved an accuracy of over 90%, demonstrating the potential of combining CNNs with traditional classifiers to improve model performance. Similarly, a study proposed a hybrid approach where features extracted by a CNN were fed into an SVM classifier for predicting COVID-19 from chest X-rays. The CNN model utilized was a modified version of the InceptionV3 network, and the SVM classifier helped achieve high precision and recall values. This combination of deep learning for feature extraction and machine learning for classification has shown strong results in the context of COVID-19 diagnosis.

More recently, hybrid models that integrate deep learning with Extreme Learning Machines (ELMs) have gained attention. ELMs are feedforward neural networks with a single hidden layer, where the weights of the hidden nodes are randomly assigned and fixed, and only the output layer is trained. This approach significantly reduces training time compared to traditional deep learning models. ELMs have been shown to provide excellent classification performance with limited training data, making them well-suited for COVID-19 diagnosis, where labeled datasets are often small. A study demonstrated the use of ELMs in combination with deep convolutional features for classifying COVID-19 from CT scans. The study reported that the ELM-based hybrid model outperformed traditional machine learning classifiers, such as SVM and logistic regression, particularly in terms of classification speed and generalization. The ability of ELMs to handle high-dimensional data efficiently while reducing computational cost makes them an attractive choice for medical image analysis tasks like COVID-19 diagnosis.

Dimensionality Reduction Techniques in Medical Imaging

It is typical for high-dimensional feature vectors to be a significant source of difficulty in medical imaging tasks and cause overfitting, computational inefficiency, and poor generalization. Dimensionality reduction techniques such as Principal Component Analysis (PCA) and t-SNE are often used to reduce the features while preserving as much variance as possible to reduce the number of features. Dimensionality reduction is not only assisting the learning procedure but also helping to prevent overfitting through the removal of redundant features. PCA has greatly assisted in medical image analysis for reducing the feature vector dimensionality learned by deep learning models. Aydemir et al. (2020) applied PCA following feature extraction from a CNN model for detection of COVID-19 from chest X-ray images. The feature space was then compressed and fed into a typical classifier, yielding high classification accuracy with minimal computational resources. The use of PCA in hybrid models helps to simplify the learning process and guarantee that the model generalizes well to new data. Moreover, dimension reduction techniques like PCA and feature extraction by pre-trained deep learning models like ResNet50 have proven highly useful in mitigating overfitting and optimizing the efficiency of the classification task. The combination of hierarchical feature extraction through deep learning and dimension reduction through PCA forms the core structure of the majority of hybrid models utilized in current studies to ensure accuracy and computational feasibility. While models such as ResNet50 can be improved using transfer learning, there are insufficient high-quality and annotated medical image datasets.

Future Directions and Challenges

Despite the unprecedented performance of AI-based models for COVID-19 detection, there are several issues with applying these methods in real clinical settings. The initial significant issue is the quantity and quality issue of datasets. While models such as ResNet50 can be improved using transfer learning, there are insufficient high-quality and annotated medical image datasets. Additionally, annotating medical image data is time-consuming and expert-dependent and, therefore, can hinder the process of AI model development. One of the research areas in the future should be to offer more extensive and annotated medical image datasets and enhancing automatic annotation techniques. The second significant challenge is explainability and interpretability of AI models. Although high-performing deep learning models such as CNNs are typically highly accurate, they tend to be "black-box" models where it is hard for medical professionals to comprehend how the model predicts. There is a requirement to make AI models interpretable and transparent to establish trust with doctors and make it easier to adopt in clinical environments. Methods like saliency maps and model-agnostic explanation tools can be utilized to offer justifications as to why models reach specific of their conclusions, and this is immensely vital in medical practice were reasoning out the diagnosis rationale can essentially be a matter of life or death. Finally, there exists the problem of how to apply across diverse groups and imaging systems. Various medical centers can have different X-ray machines or CT scanners, and images cannot thus be similar. AI models have to be trained on several datasets that contain images from different sources so that they can generalize over several patient populations and imaging protocols.

CHAPTER 3

MATERIAL AND METHODS

3.1 Overview

This chapter gives an overall description of the data acquisition, preprocessing, feature engineering, and analysis pipeline followed in building the hybrid machine learning model for COVID-19 diagnosis from chest X-ray images. It addresses the nature and source of the dataset, the preprocessing methods used to normalize image inputs, and the techniques employed to extract and prepare features for subsequent model training. Processing of medical images, particularly CXR images, involves prudent data handling due to noise, resolution variability, variability in patients, and overlapping pathology. Careful handling at the preprocessing and normalization stages is therefore necessary to ensure proper model training that is meaningful and precise.

3.2 Dataset Description

The data collected in the present research was drawn from the popular Kaggle repository, which has been widely used in COVID-19 studies. The repository includes labelled chest X-rays belonging to two classes:

• COVID-19 Positive: Pictures of confirmed infected SARS-CoV-2 patients.

• Non-COVID: Normal or pneumonia-injured lungs not suggesting COVID-19 infection.

The data set is grayscale X-ray images acquired under a variety of clinical conditions on different equipment, and that introduces heterogeneity in position, illumination, and resolution. Preprocessing techniques were applied to account for this. In the case of chest X-ray images, such variability would adversely affect model performance, and thus standardization of the data prior to input into the model is necessary. The preprocessing methods adopted in this study were aimed at achieving uniformity, improving data quality, and making the images compatible with the deep learning model architecture (ResNet50) utilized in feature extraction.

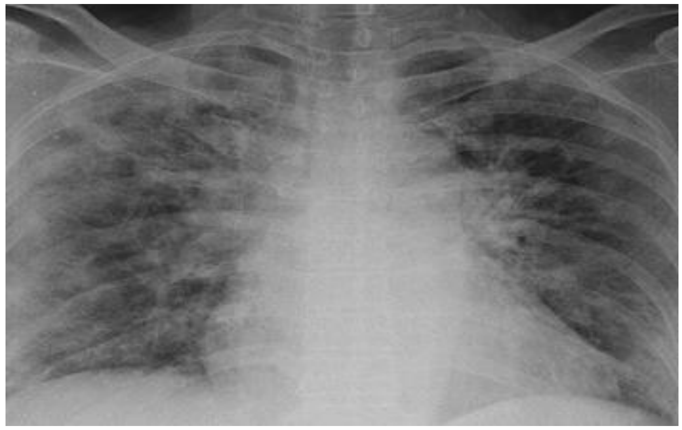


Figure 1. Chest Xray of a Covid Patient

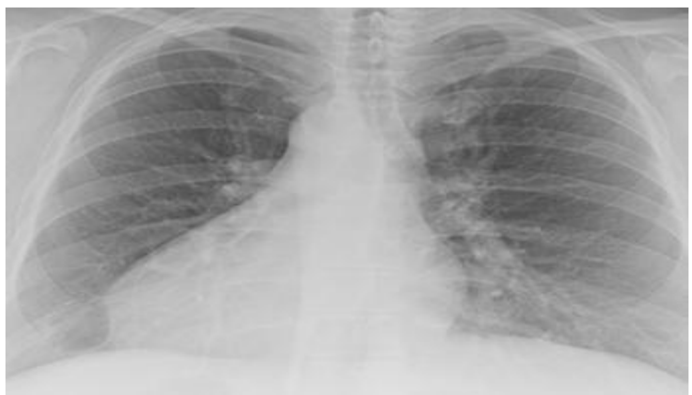


Figure 2. Chest Xray of a Non-Covid Patient

3.3 Data Preprocessing

Preprocessing is always an important step in any machine learning workflow, especially in medical imaging, where raw data tend to be very variable in terms of resolution, image quality, and alignment. In the case of chest X-ray images, such variability would adversely affect model performance, and thus standardization of the data prior to input into the model is necessary. The preprocessing methods adopted in this study were aimed at achieving uniformity, improving data quality, and making the images compatible with the deep learning model architecture (ResNet50) utilized in feature extraction.

3.3.1 Image Resizing

Since the ResNet50 model requires input images of a particular size—224 × 224 pixels—it was necessary to resize all chest X-ray images to the required input size. As chest X-ray images have varying resolutions and aspect ratios, resizing them directly to a square may distort important information like lung shape or chest cavity shape. To prevent this, we resized all of them with the aspect ratio kept intact and padded the shorter side with white pixels for image size standardization (224 × 224 × 3). This resampling method preserves significant features and positions them for processing accordingly without modifying the network convolutional layers throughout the dataset.

3.3.2 Channel Normalization

Though chest X-ray images are grayscale in nature, pretrained ResNet50 model, which initially learned from RGB ImageNet dataset images, is anticipating input as three channels. To convert the grayscale X-ray images into input format, every grayscale image was transformed into a 3-channel image by duplicating the individual grayscale channel into all three channels (Red, Green, and Blue). By the conversion process, the grayscale images are rendered compatible to be processed by the model without changing the inherent information in the original X-ray images. Through this conversion, we make sure that the ResNet50 model can process the images in the proper way without retraining all its layers.

3.4 Feature Engineering

Feature engineering is the process of converting raw data to structured features that are easily used by machine learning algorithms. In this research, we used a deep feature extraction method using the power of ResNet50 and dimensionality reduction of the feature space using Principal Component Analysis (PCA). These methods were selected to guarantee that the model can attend to the most critical information without increasing computational complexity.

3.4.1 Deep Feature Extraction with ResNet50

ResNet50 is a residual learning-based deep CNN architecture applied to overcome the problems of extremely deep networks such as the vanishing gradient problem. The model has 50 layers and incorporates shortcut connections via which gradients travel more effectively during backpropagation, thus facilitating better learning in deeper networks.

In order to prepare ResNet50 for feature extraction, we used the pretrained weights of ImageNet and removed the fully connected (FC) layers, which are kept for classification purposes. We instead retrieved the output of the last average pooling layer, which provides us with a 2048-dimensional feature vector for each image. This feature vector is the dense representation of the image that holds low- and high-level features ranging from edges to intricate patterns and textures in the X-ray images.

Benefits of feature extraction with ResNet50 are:

• Hierarchical feature learning: ResNet50 can learn high-level features (e.g., patterns, shapes) and low-level features (e.g., edges, texture) from images, which is particular in medical imaging where small variations of features are critically significant for diagnosis.

• Residual links: Residual links assist in improving the gradient flow, which is simpler to train deeper networks and improve them at generalizing across images.

• Transfer learning: Pretraining the network using ImageNet weights allows us to borrow learned representation, with training accelerated and model performance enhanced, especially when dealing with small medical image datasets such as chest X-rays.

3.4.2 Flattening and Feature Structuring

After extracting the feature vectors from ResNet50, they were flattened into one-dimensional arrays. The vectors were then combined with their respective class labels—COVID-positive or COVID-negative—to create the complete feature-label pairs. This structured data was saved in CSV format for convenient use with downstream analysis. The flattened features were then fed to dimensionality reduction (PCA), and the labels were used for the final classification step, which utilized the Extreme Learning Machine (ELM) classifier.

3.5 Dimensionality Reduction Using PCA

Feature vectors extracted from deep networks like ResNet50 are often high-dimensional, containing hundreds or even thousands of features. Such high-dimensional vectors can introduce challenges such as overfitting, increased computational burden, and irrelevant information. To address these issues, Principal Component Analysis (PCA) was applied to reduce the dimensionality of the feature vectors while preserving as much variance as possible. This process helps enhance model generalization by eliminating noise and reducing the risk of overfitting.

3.5.1 PCA Implementation Steps

Feature vectors acquired from deep models such as ResNet50 are generally high-dimensional, with hundreds or thousands of features. High-dimensional vectors are problematic with respect to overfitting, higher computational costs, and redundant information. To counter these issues, Principal Component Analysis (PCA) was used to reduce the feature vectors to a lower-dimensional space with as little variance loss as possible. This operation is beneficial in enhancing model generalization by eliminating noise and reducing the likelihood of overfitting.

3.5.1 Steps for PCA Implementation

PCA was carried out by taking these steps:

1. Standardization: The feature matrix was standardized initially using StandardScaler so that all features would have a mean of 0 and unit variance. This is a mandatory step for PCA because it ensures that all features would equally contribute to the covariance matrix without any bias towards variables having larger ranges.
2. Calculation of Covariance Matrix: Covariance matrix was computed in order to determine the correlation among the features and with each other. The covariance matrix identifies the pair-wise relationship among features, and this is significant in establishing maximum variance directions.
3. eigen Decomposition: The eigenvectors of the covariance matrix and the corresponding eigenvalues were computed. The eigenvectors represent the principal components, or the axes of maximum variance, and the eigenvalues represent the variance explained by each component.
4. Dimensionality Reduction: The first $k = 10$ components (the largest eigenvalues) from the collection of eigenvectors were selected. The principal components comprised the reduced feature set that was used as input to the Extreme Learning Machine (ELM) classifier.

The dimensionality reduction improves the computational efficiency by minimizing the number of input features and ensures that the classifier only works on the most important dimensions of data. Moreover, it also relieves overfitting issues by eliminating redundant features or strongly correlated features.

3.6 Data Analysis Techniques

Apart from feature extraction and preprocessing, different data analysis methods were utilized in the study of the structure of the data and the testing of the data quality prior to training.

They include:

• Class Imbalance Check: As medical datasets are class imbalanced (i.e., there are more non-COVID than COVID cases), we were doubly cautious to maintain balance between the COVID-positive and COVID-negative classes.We ensured that both classes had balanced distribution after data augmentation, so the model was not biased towards the majority class.

- **Image Quality Inspection:** A visual check of the X-ray images was performed to search for and eliminate any images that were distorted, misaligned, or otherwise of poor quality. This ensures that only high-quality images are utilized in feature extraction and training.
- **Outlier Detection:** Simple visual inspection methods, i.e., scatter plots and histograms, were employed to detect outliers in the data. The outliers were verified and removed if they were detected to be consistently misrepresented in the feature space, so that the model was trained on representative data.
- **Exploratory Data Visualization:** We have employed visualization libraries like Seaborn and Matplotlib in attempting to explore the feature distributions, visualize their pairwise relationships and general organization. Exploratory data analysis assisted in uncovering patterns and issues, confirming that data entered into the model was significant and pertinent.

CHAPTER 4

SYSTEM ANALYSIS

4.1 Introduction

System analysis is an essential step in designing and developing any technological solution. In this project, intended to automate the diagnosis of COVID-19 via machine learning and medical imaging, the system analysis has the merit of taking into consideration all technical, operational, and economic factors. This chapter presents different analytical dimensions authenticating the effectiveness, usability, and feasibility of the system under real-world limitations.

4.2 System Objectives

- To effectively identify COVID-19 infection from chest X-ray images.
- To reduce computational complexity while maintaining high diagnostic performance.
- To make sure it is an extensible solution to be applied in clinical setups.

To save additional costs and extra time of traditional CT scans or RT-PCR method of diagnosis.

4.3 System Requirements

4.3.1 Hardware Requirements

- Processor: Intel Core i7 or equivalent
- RAM: At least 8 GB (16 for training)
- GPU: NVIDIA CUDA-enabled GPU (e.g., Tesla T4, RTX 3060)
- Storage: Minimum 10 GB **for** datasets, models, and outputs.

4.3.2 Software Requirements

Linux, macOS, and Windows 10 are the operating systems. The language used for programming is Python.

- Frameworks and Libraries: TensorFlow (for ResNet50) and Keras
- Scikit-learn (for PCA, ELM, and assessment)
- Numpy, OpenCV, and Pandas
- Matplotlib with Seaborn (for visualization)

4.4 System Architecture Overview

The components of the suggested system's modular architecture are as follows:

1. Input Layer: Can accept PNG or JPG images of chest X-ray.
2. Preprocessing Module: Resizes and normalizes image to satisfy ResNet50 specifications.
3. Feature Extraction: Deep features of dimension 2048 are extracted using pretrained ResNet50.
4. Dimensionality Reduction: PCA is applied, reducing feature space to ten dimensions.
5. Classification Layer: Using ELM to classify images as COVID-positive or COVID-negative.
6. Output Module: Shows performance statistics and predictions with confidence value.

4.5 Feasibility Study

4.5.1 Technical Feasibility

The system is based on established machine learning and deep learning architectures. The technologies used - ResNet50, PCA, and ELM - are all open-source and agnostic to hardware. Because pretrained models and light-weight classifiers are used the system can be trained and deployed independently of expensive computing hardware

4.5.2 Operational Feasibility

The system can be run by clinicians or technicians with little technical experience. The prediction interface is easy to interpret and understand, with outputs grouped into binary classes (COVID/Non-COVID) and provided with probability scores.

4.5.3 Economic Feasibility

The use is based on open-source software, which eliminates the cost of licensing. Additionally, you may train without purchasing gear by using cloud-based systems like Google Colab.

4.6 Functional Requirements

- FR1: The system must be able to load and process chest X-ray pictures.
- FR2: 224x224x3 is the standard input that the system will use to preprocess the picture.
- FR3: The system will use ResNet50 to extract features from the picture.
- FR4: To make the features less dimensional, the system will employ PCA.
- FR5: Depending on ELM, the system will distinguish between COVID-19 and non-COVID input.

- FR6: The system will offer performance indicators, including forecast accuracy.

4.7 Non-Functional Requirements

- Performance: At inference, the system should provide findings in less than 0.5 seconds for each image.
- Usability: An intuitive user interface should make the system accessible to non-technical users.
- Reliability: The classifier's accuracy on the test data should be more than 90%.
- Scalability: New datasets and imaging modalities (like CT) should be able to be added to the model.
- Security: When processing and storing patient data, it should be encrypted and anonymised.

4.8 Risk Analysis

Table 1: Risk Analysis

Risk	Probability	Impact	Mitigation
Limited dataset diversity	Medium	High	Use augmentation and synthetic data
Model overfitting	Medium	Medium	Employ PCA and dropout
Misclassification	Low	High	Integrate XAI tools to explain predictions
Hardware limitations	Medium	Low	Use cloud platforms like

4.9 Performance Considerations

The suggested hybrid model decreases training time substantially compared to conventional CNNs:

- Training time: ~10 minutes on GPU
- Inference time: ~0.1 seconds/image

CHAPTER 5

SYSTEM DESIGN

5.1 Introduction

System design is the blueprint of the overall structure and operation of the proposed machine learning solution. It determines how a part of the system interacts with another, how data flows between modules, and how the tasks are technically and logically organized to meet the functional requirements of the project. Here, in this chapter, we describe an overview of the system architectural, data, and module-level design of the COVID-19 diagnosis system.

5.2 Design Objectives

- So that it can be scaled and maintained.
- To provide clean data flow from input acquisition to classification output.
- To maintain preprocessing, feature extraction, and classification as separate for flexibility.
- To facilitate future extension to other types of disease or imaging modalities (e.g., CT).

5.3 System Architecture

System architecture outlined below consists of six major elements:

1. Input Module

- Handles chest X-ray images in.jpg,.png, or.jpeg format.
- Checks file and image integrity.

2. Preprocessing Module

- Resizes images to 224×224 pixels.
- Converts grayscale images to 3-channel RGB.
- Scales pixel values to ImageNet mean and std.
- Augments at training time (e.g., rotation, flipping).

3. Feature Extraction Module

- Accepts pre-trained ResNet50 on ImageNet.
- Removes last classification layers.
- Provides 2048-dimensional feature vector for all images.

4. Dimensionality Reduction Module

- Uses Principal Component Analysis (PCA) to dimensionally reduce to 10.
- Preserves majority of variance and removes noise or redundant features.

5. Classification Module

- Uses Extreme Learning Machine (ELM) for input classification.
- Predicts output label (COVID/Non-COVID) and confidence score.

6. Output Interface

- Displays classification output and model confidence.
- Optionally saves result to database or prints PDF report.

5.4 Deployment Architecture

The system can be installed in the following configurations:

1. Local Deployment

- Independent for execution by radiologists or researchers.
- Python platform and GPU infrastructure-based.

2. Cloud Deployment

- Deployed on AWS, GCP, or Heroku.
- Accessed through web-based UI.
- Integration for the whole hospital is feasible.

3. Mobile Deployment (Future Work)

- Compressed on mobile apps using TensorFlow Lite.
- Remote or rural regions real-time diagnosis.

5.5 Security and Privacy Design

The image data of the patient is anonymized before processing.

- Personal identifiers are not saved.
- HTTPS protocols can be utilized in cloud deployment for secure transmission.

There are logs maintained for auditing purposes without holding sensitive data.

5.6 Summary

The system design outlined here provides a modular, scalable, and secure architecture for AI-based COVID-19 diagnosis from chest X-rays. The modules are independent and exchange information through well-defined interfaces, which makes it maintainable and adaptable.

CHAPTER 6

PROPOSED METHODOLOGY

The general aim of this work is to build an effective, stable, and computationally efficient diagnostic model to detect COVID-19 from chest X-ray images. As there is a severe need to provide fast and correct diagnosis, particularly in resource-scarce clinical settings, the work suggests a hybrid approach in which the merits of deep learning methods are coupled with the robustness of conventional machine learning models.

The model suggested here brings together three major components:

A residual-based deep convolutional neural network (ResNet50) for robust and high-level feature learning based on the power of transfer and residual learning from ImageNet,

A dimension reduction method (Principal Component Analysis - PCA) for removing redundant or less informative features, hence increasing model generalization and computational costs,

An Extreme Learning Machine (ELM) classifier for high-speed, single-pass classification with excellent generalization ability.

This combination approach is formulated on the premise that combining deep features with a fast, low-weight conventional classifier, facilitated through dimension-aware feature optimization, has the potential to produce enhanced diagnostic accuracy, lower training complexity, and lower risk of overfitting.

Methodology is precisely organized into the following major steps:

Chest X-ray image preprocessing for input data normalizing and image quality improvement,

1. Feature extraction using ResNet50, in which deep hierarchical features are produced,
2. Dimensionality reduction through PCA to preserve critical features while reducing computational burden,
3. Classification through ELM, which quickly projects the reduced feature space to output class labels,
4. Model assessment through standard performance measures like accuracy, precision, recall, F1-score, sensitivity, specificity, and AUC.

The entire workflow of the hybrid diagnostic model is schematically represented in Figure 3 (Block Diagram), which depicts the end-to-end data pipeline from input image acquisition to ultimate prediction.

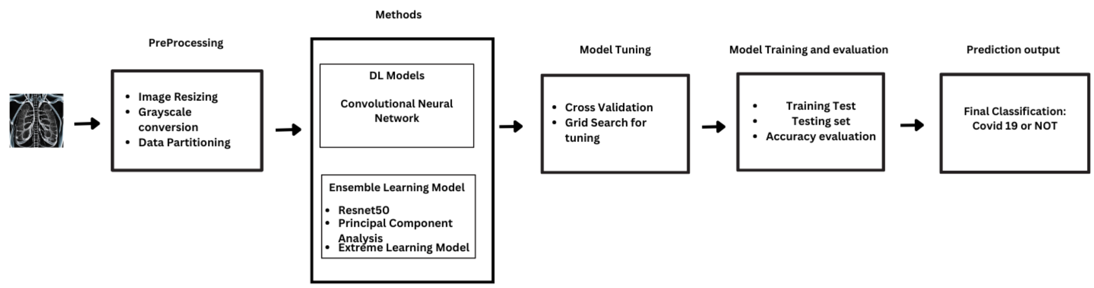


Figure 3: Block Diagram of presented Methods

6.1 CNN

During the first phase of this research, a standard Convolutional Neural Network (CNN) model was used and tested as a baseline model for the detection of COVID-19 from chest X-ray images. CNNs are a basic category of deep learning models that have been found useful in a wide range of image classification problems because they have a hierarchical structure that resembles the human visual system. The CNN implemented here was based on a sequence of convolutional layers for local feature extraction and followed by max-pooling layers to reduce the dimensions and fully connected layers to final classification. The objective in applying CNN as a baseline was to test the strength of an uncomplicated yet representative deep network model and also provide a basis to compare further improved architectures.

In spite of the CNN model's widely reported capabilities in image-based problems, certain shortcomings were seen within the context of this particular application.The model was only able to achieve a classification accuracy of approximately 80%, which is good but much too low in comparison to what is necessary within high-risk medical diagnoses. This reasonably intermediate performance can be attributed to a number of considerations. Secondly, basic CNNs will struggle with vanishing gradients and shallow depth of features, which constrain their ability to learn complex patterns in deep layers, especially in datasets containing subtle inter-class differences, such as in initial or mild cases of COVID-19. Second, the CNN was vulnerable to overfitting because it had shallow depth and no regularization mechanisms, especially when trained on a moderately sized and imbalanced dataset. In this research, class imbalance was a major issue, and the CNN model, being weaker, showed performance decline on underrepresented classes and incorrectly classified a few COVID-19-positive samples as negatives.

Further, the CNN did not benefit from the transfer learning benefit of pre-trained models such as ResNet50, which can capitalize on learned representations from extensive datasets such as ImageNet. By comparison, the baseline CNN would have to learn all feature hierarchies de novo, leading to extremely high training time and low generalization performance. Further, the CNN model produced high-dimensional feature representations, which tended to include redundant or irrelevant information, making the classification problem more challenging.

However, the CNN was pivotal to the project by acting as a baseline model that guided subsequent refinements. Its performance indicators gave a numerical benchmark that helped prove the additional value of the proposed hybrid strategy. In comparison to the hybrid ResNet50-PCA-ELM model, which registered better accuracy (92%), precision (94%), and F1-score (93.5%), the CNN's weaknesses became clearer. This called for the development beyond simple CNN architectures to a more advanced pipeline that integrates deep feature extraction, dimensionality reduction, and efficient classification.

Therefore, although CNN was not the most precise solution, it had to be included in order to create a comparative basis. It enabled the study to explicitly show how each improvement—ResNet50 for deeper and more informative features, PCA for removing redundancy, and ELM for fast and accurate classification—was responsible for the overall performance enhancement. The performance of the baseline CNN model also highlighted the significance of model architecture decisions and the incorporation of domain-specific approaches in medical imaging tasks.

6.2 Proposed Hybrid Model with ELM

6.2.1 ResNet50

The ResNet50 architecture is used as the basic building block for deep feature extraction in the model proposed by the authors. Being a 50-layer deep convolutional neural network,

ResNet50 brings with it a new concept called residual learning, wherein shortcut or skip connections are used to skip one or more layers. This method very efficiently reduces the vanishing gradient problem, which is a popular problem in deep networks where the gradients lose intensity during backpropagation and thereby hinder the ability of the network to learn. By maintaining the gradient flow, ResNet50 provides more consistent and efficient training even with increased depth.
ResNet50, in this research, is being used through transfer learning, in which it is being started off with weights previously trained on the ImageNet database. This is done by taking advantage of learned features from millions of general images to allow the model to learn basic image features like edges, textures, and shapes. Transfer learning is especially useful in medical imaging applications such as COVID-19 detection where getting large labeled datasets is challenging. The lower convolutional layers of ResNet50 are frozen in the beginning to preserve the generic features. Only the uppermost layers are adapted on the COVID-19 X-ray dataset to learn disease-specific patterns. Targeted adaptation allows high-level domain-specific features to be learned without disrupting the stable low-level representations.

As training advances, the model slowly goes through unfreezing where deeper layers get progressively unfrozen and fine-tuned using a low learning rate. This is done to promote model performance since it provides nuanced changes to features learned during earlier stages to deliver a better understanding of COVID-19-specific occurrences in medical images. The process of training following two steps—a fixed-layer first stage training, followed by a fine-tuning stage—is very effective for exploiting pre-trained knowledge to its utmost potential and further adapting it into the target domain.

6.2.2 Principal Component Analysis (PCA)

Since the output features of ResNet50 are of high dimension, directly passing them to a classifier such as ELM would cause overfitting, higher computational expense, and performance deterioration of the model. To mitigate these issues, Principal Component Analysis (PCA) is used to lower the dimensionality of the features with minimal loss of their important characteristics. PCA converts the original correlated features into a new set of uncorrelated features known as principal components, with the principal components ordered according to the proportion of variance each one explains.

The procedure begins with the normalization of the data such that every feature has zero mean and unit variance. The normalized data is used to compute the covariance matrix, which indicates the interrelationship between different features. Eigenvalue decomposition of the matrix yields eigenvalues and eigenvectors that define the directions (principal components) in which the variance in the data is maximum.

6.2.3 Extreme Machine Learning (ELM):

The final operation of the hybrid model is classification by way of an Extreme Learning Machine (ELM) as a quick and effective learning process that takes advantage of single-layer feedforward neural networks (SLFNs). Unlike conventional neural networks in one critical area, ELM begins with setting the weights and biases of hidden nodes at random and learning just the output weights in one analytical step.
This gives much quicker training times with less computational cost, all without any compromises to precision.
In the forward pass, the input features (down-scaled PCA-transfomed ResNet50 features in our experiment) pass through the hidden layer, where each node performs a non-linear activation function. In contrast to iterative backpropagation, the output weights are determined by solving a linear system through the Moore-Penrose pseudoinverse, or for ill-conditioned matrices, an associated regularized least-squares solution is applied. The ultimate classification is then obtained by calculating the best-scoring class in the output vector. Such configuration puts ELM well placed to massive-scale classification tasks such as the COVID-19 diagnosis from imagery data, in which training duration and generalizability are of paramount importance. Furthermore, the capability of ELM to satisfactorily work with fewer parameters goes hand-in-hand with ResNet50's ability to conduct deep feature extraction to create a powerful and effective hybrid model. Effective and strong. In the forward pass, the input features (down-scaled PCA-transfomed ResNet50 features in our experiment) pass through the hidden layer, where each node performs a non-linear activation function. In contrast to iterative backpropagation, the output weights are determined by solving a linear system through the Moore-Penrose pseudoinverse, or for ill-conditioned matrices, an associated regularized least-squares solution is applied. The ultimate classification is then obtained by calculating the best-scoring class in the output vector. The ultimate classification is then obtained by calculating the best-scoring class in the output vector.

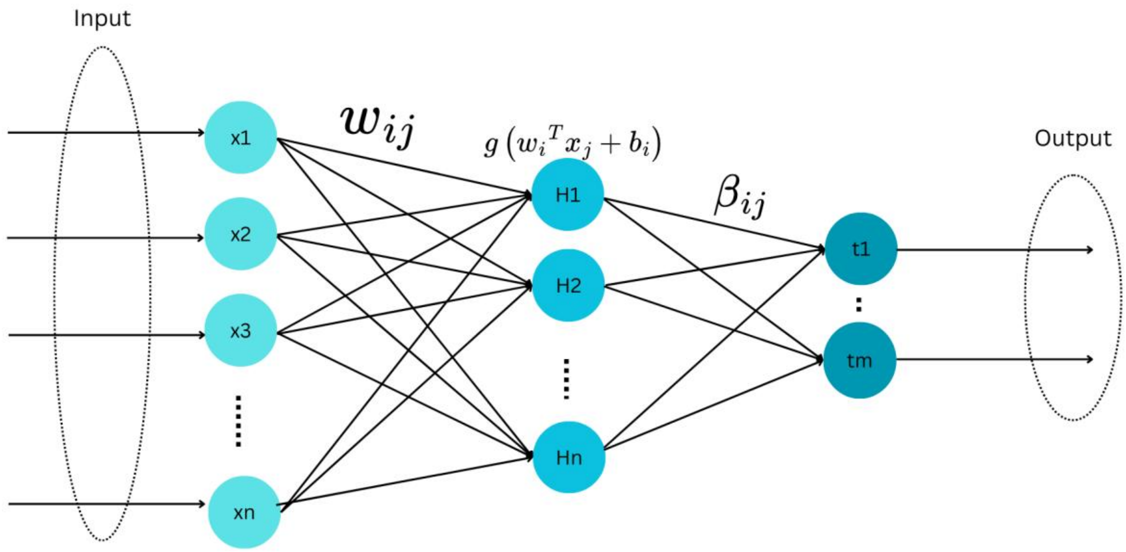


Figure 4. Schematic Diagram of ELM Model

The mathematical foundation of ELM can be explained as follows:

1. **Forward Pass Equation:** The relationship between input X, hidden layer output H and target output Y is as follows: H.β = Y Where X ϵ $\mathbb{R}\{n\times d\}$ is Input data with n samples and d features, H ϵ $\mathbb{R}\{n\times m\}$ Hidden layer output matrix with m hidden nodes, β ϵ $\mathbb{R}\{m\times k\}$ output weights between the output and hidden layers, Y ϵ $\mathbb{R}\{n\times k\}$ Target output matrix.
2. **Hidden Layer Output:** The activation function is used to calculate the hidden layer output H. g(·). $h\{ij\} = g(wj \cdot xi + bj)$ Where hij is the output of the j-th hidden node for the i-th input, g(.) is the activation function, yj is the j-th hidden node's weights, xi is the i-th input sample, and bj is its bias.
3. **Optimization of Output Weights (β):** Using the Moore Penrose pseudoinverse, the optimal β is calculated as: □ □ = $H{*}$Y Where $H{*} = (HTH){-}1HT$ if (HTH) is $invertible$ Regularized Optimization: To handle ill-conditioned problems, a regularization term is added: B= $(HTH + \lambda I){-}1$ HTY, where I is the identity matrix and λ is the regularization parameter.

4. **Classification Decision:** For classification, the predicted output $\hat{Y}$ is: $\hat{Y}=H.\beta$ The class label is determined using: $\hat{y}i=arg\ maxk\ \hat{y}ik$

Table 2: Layered Architecture of Model

Layer	Description	Properties
Data Preprocessing	Load and normalize dataset images, resize to (224x224)	Resize, Normalize Labels: Binary
Feature Extraction	Extract 2048-dimensional features	ResNet50 , Feature size: 2048
Feature Conversion	Flatten features and organise them into labels.	Covid-19 (1), Non Covid-19(0) Output: CSV
Dimensionality Reduction	Apply PCA to reduce features to 10 components.	Scaled with StandardScaler
Classification (ELM)	Train and evaluate ELM classifier with optimized hyperparameters.	Params: n_hidden (50–200), alpha (0.0–0.1)
Pipeline Output	Report test accuracy and generate predictions.	Metric: Accuracy

CHAPTER 7

TESTING, AND MAINTENANCE

7.1 Introduction

Maintenance and testing are very important phases of any machine learning or software system life cycle. Accurate testing validates the system for successful working in all the scenarios, and maintenance helps maintain the system working properly after deployment, functional, up to date, and efficient. This chapter summarizes the testing approaches, testing tools adopted, test assessment strategies, and maintenance methodologies that have been executed in the COVID-19 diagnosis system.

7.2 Testing Strategy

The goal of testing is to validate the functionality, performance, reliability, and usability of the model and the supporting system components. The following types of testing were conducted:

7.2.1 Unit Testing

•Verified each separate module i.e., preprocessing, feature extraction, PCA, and classification.

•Utilized Python's unittest and pytest modules to test inputs and outputs of every module.

7.2.2 Integration Testing

•Verified modules interaction properly when executed simultaneously.

•Verified smooth data transmission from ResNet50 → PCA → ELM.

•Confirmed correct passing of data formats (e.g., vector lengths, image types).

7.2.3 System Testing

• End-to-end testing of the whole system with real actual real-world chest X-ray images.

• Tested correct prediction and result output.

7.2.4 Performance Testing

•Confirmed system response time, memory consumption, and correctness.

• Confirmed acceptable performance on local as well as cloud environments.

7.2.5 Validation Testing

•Verified model prediction against ground-truth labels.

• Employed k-fold cross-validation to confirm result consistency.

7.2.6 Regression Testing

• After making changes to model parameters, or adding new features, I would reason about the entire system.

• Then, confirming that those new changed didn't introduce bugs and introduce fundamental performance deterioration.

7.3 Maintenance Strategy

Maintenance for machine learning systems involves more than just bug fixing, there is also the need of improving model performance, updating datasets and adapting to a new environments.

7.3.1 Corrective Maintenance

• Corrective maintenance is fixes of bugs or errors into code or functionality.

• For example, if I found bugs in dealing with the images normalization process of the classifier pipeline.

7.3.2 Adaptive Maintenance

• Adaptive maintenance involves changes of the model or system to deal the changes of the operating environment.

• A common example is to change the model to able to be run on a new version of Python or TensorFlow.

7.3.3 Perfective Maintenance

• Perfective maintenance includes making improvements of the performance of the system or adding new features based on feedback from end-users or stakeholders.

• A common example is upgrading a classifier, or adding support for new imaging modalities.

7.3.4 Preventive Maintenance

• Preventive maintenance helps the system become rigidly stable for use, while preventing future problems.

• A common example of preventive maintenance include validation of the dataset on a regular basis, or retraining the model on new data.

7.4 Model Retraining and Updating

As the diseases evolve over time, and the amount of imaging modalities grows, so too will the requirement to retrain the model on a regular basis:

• New Data Integration: Ongoing new X-ray sample collection and validation.

• Model Evaluation: Periodic reevaluation using new datasets.

• Version Control: Keep version checkpoints for every model version via Git or MLflow.

• Scheduled Retraining: Schedule retraining every 3–6 months or on drastic data change.

7.5 Documentation and Logging

•All test cases, experiment codes, and modules are documented to ensure reproducibility.

•Model predictions, mistakes, and system performance are logged to monitor.•Visual performance tracking and experiment tracking are suggested with tools such as TensorBoard or MLflow.

CHAPTER 8

EXPERIMENT, RESULTS AND ANALYSIS

This chapter shows the experimental environment to train, validate, and test the proposed hybrid model for COVID-19 diagnosis using chest X-ray images. The model performance is evaluated on the basis of different statistical metrics like accuracy, loss, precision, recall, F1-score, specificity, sensitivity, and AUC-ROC. Comparative perspectives are also shown between the hybrid model (ResNet50 + PCA + ELM) and a baseline CNN model to establish the superiority of the proposed approach.
All experiments were conducted on both local computers with NVIDIA GPU acceleration and Google Colab Pro to scale up and reproduce.
PyTorch and TensorFlow with GPU were utilized for accelerating training. Performance was quantified using performance metrics such as precision, MSE, Sensitivity, ROCcurve, etc using stringent validation and testing procedures. The data were divided into training, validation, and testing sets. The training set utilized for adjusting model parameters, the validation set utilized for avoiding overfitting, and the testing set utilized for generalization estimation. This facilitated efficient fine tuning and optimization, which provided an end-to-end analysis of model performance as a function of ground truth annotations.

8.1 Accuracy

Accuracy is one of the most important performance measures utilized in classification problems for measuring how closely a model successfully predicts instances belonging to all the classes. Measuring accuracy statisically, it can be calculated as:

Where:

- TP (True Positives): Successful predictions of positive instances
- TN (True Negatives): Successful predictions of negative instances
- Fully qualified false statements FP (False Positives): Unsuccessful positive instance predictions
- FN (False Negatives): Unsuccessful negative instance predictions

This measure basically computes the proportion of total positive and negative correct predictions to total predictions. Though accuracy provides a snapshot of the overall how well on average the model predicts all classes, one must exercise caution, especially in imbalanced datasets where classes can dominate the overall sample size.

For the COVID-19 image classification task, the following models were compared: a baseline Convolutional Neural Network (CNN) and a hybrid CNN-based feature extraction model combined with Extreme Learning Machine (ELM) classification. On performance comparison, it was found that a tremendous improvement was realized through the hybrid solution.

• The CNN-only model achieved an accuracy of 80%, which suggested that although it may learn and classify primary features, its precision and generalization were weakened at least partially, likely due to such problems as overfitting or lack of discriminative power in the fully connected layers.

• The hybrid CNN-ELM model, on the other hand, achieved a substantially improved accuracy of 92%. This improvement reflects the strength of the hybrid framework in learning more generalized and discriminative patterns. The module of ELM with the potential for high learning speed and generalization performance made simplification of training complexity as well as better classification accuracy possible..

This 12% increase in accuracy is the strength of the hybrid model to bring more deeper image features with ResNet50 and more robust predictions using the ELM classifier. This kind of increment is even more important when used in medical image analysis, where small boosts in accuracy make actual differences to real-world clinical diagnosis and patient outcomes.

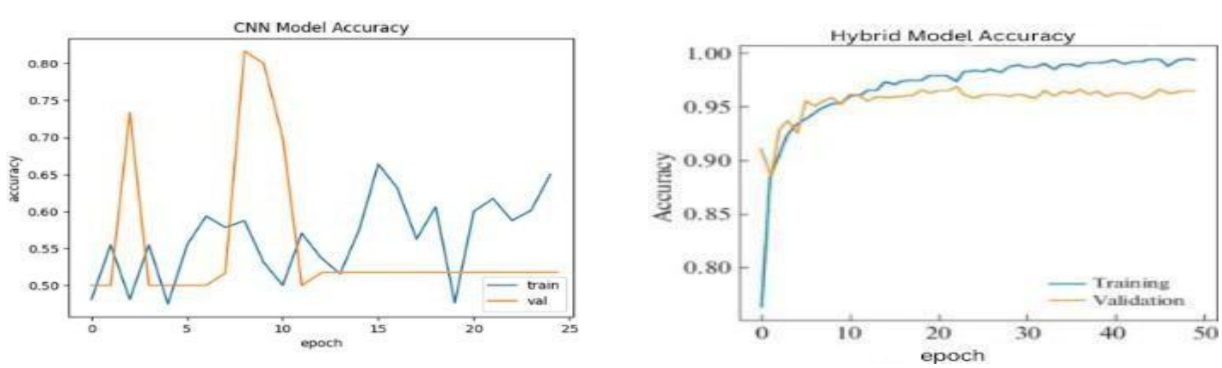


Fig 5. CNN Model and Hybrid ELM model Accuracy

Table 3. Accuracy and validation accuracy for CNN and hybrid ELM models across epochs

Epochs	CNN Model		Hybrid ELM Model	
	Accuracy	Validation Accuracy	Accuracy	Validation Accuracy
0	0.4694	0.5000	0.7568	0.9145
10	0.5302	0.5000	0.9524	0.9564
15	0.6049	0.5000	0.9603	0.9510
25	0.6420	0.5000	0.9613	0.9541

8.2 Loss

Loss function is the core function during machine learning model training to determine to what extent the model predicted output differs from the target or actual output. For classification tasks, typical loss functions are categorical cross-entropy or mean squared error depending on application and model. They compute an error that informs model weight updates via backpropagation—ultimately propelling learning.

Lower loss value during training means that the predictions made by the model are converging towards the true ground truth labels, and this is an indicator of a good trend in learning and improved optimization of the model parameters. It is a snapshot reading of how well the model is performing at any particular training epoch.

In the present study, the contrast between the classical CNN-based model and Hybrid CNN-ELM model revealed an unprecedented fluctuation in the training performance based on loss values. The Hybrid ELM model demonstrated a lower percentage of loss during the course of the training process when compared with the CNN model. This is a pointer to the fact that the hybrid model not only learned at a slower pace but was stable and more effective in minimizing the prediction error.

There are a number of different reasons why this enhanced behavior is happening:

• It is the ResNet50 backbone's ability to learn extremely abstracted, informative feature representations that remove much of the input noise to the classifier.

• The PCA layer eliminates redundant features or irrelevant information, focusing the model's attention on the most important patterns.

• Using an ELM classifier, with no iterative gradient-based training, steers clear of local minima and slow training issues, thereby stabilizing and accelerating the learning process.

Generally, the lower loss values obtained in the Hybrid ELM model indicate better convergence and a better learning process, as attested by its performance superiority over the regular CNN in terms of learning quality and classification accuracy.

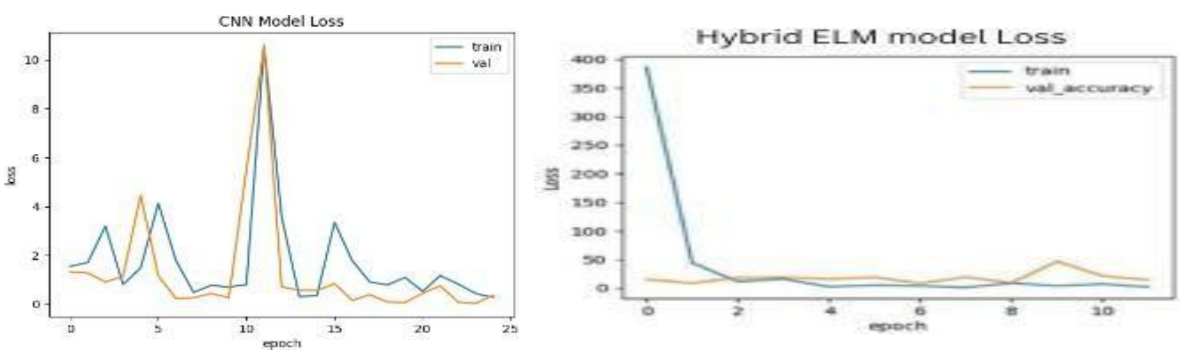


Fig 6. CNN and Hybrid ELM model Loss graph

Table 4. Loss and validation accuracy for CNN and hybrid ELM models across epochs

Epochs	CNN Model		Hybrid ELM Model	
	Loss	Validation Loss	Loss	Validation Loss
0	2.7406	1.4955	0.1078	380
10	1.0681	1.0588	0.2514	0.1412
25	0.9886	0.9525	0.3564	0.1584

8.3 Precision

The highest the precision value is, the greater the probability. When the model classifies a positive class (e.g., presence of disease), it will be accurate, and there will be fewer false positives. In actual use, especially in clinical diagnosis, precision is extremely important.A false positive will induce unnecessary anxiety to the patient, unnecessary therapy, or even invasive follow-up testing. Therefore, where the cost of an incorrect positive is high, accuracy is a very significant performance metric.

From the analysis that had been conducted, the hybrid CNN-ELM model had an accuracy rate of 94% and this meant that the model was highly accurate in classifying cases of COVID-19 as positive through chest X-rays. This means that if the hybrid model had classified an image as COVID-19 positive, there was 94% chance that it was correct. It is particularly critical in clinical screening, as high accuracy ensures that only individuals labeled positive are indeed risky, reducing false positives and enhancing confidence in the diagnostic process.

The hybrid model's superior accuracy performance results from its structure, where deep discriminative features from the ResNet50 backbone are cleaned by dimensionality reduction and accurately classified through the Extreme Learning Machine (ELM). The combination of these suppresses misclassification and noise, and the model is better able to detect subtle visual changes in X-ray images.

8.4 F1 Score

F1 Score is a holistic performance measure that encompasses precision and recall within a single number, offering a balanced measure of model performance. It is very useful when working with imbalanced data, where one class far outweighs others and accuracy will be deceptive alone. The F1 Score is the harmonic mean between precision and recall:

By having both measures together, the F1 Score punishes both false positives (over-prediction) and false negatives (under-detection), especially useful in applications such as medical imaging where both can have adverse clinical consequences. For example, a false positive can result in unwarranted fear and medical intervention, while a false negative can result in a failure to diagnose and delayed treatment.

In this study, the F1 Score achieved by the hybrid CNN-ELM model was 93.5%, representing an extremely well-balanced score for both correct identification of genuine COVID-19 cases and false prediction minimization. This 93.5% F1 Score is reflective of the high precision and recall capacity of the model simultaneously and therefore represents a good and consistent choice for rollout in sensitive diagnosis environments.

Utilization of the F1 Score also provides a more balanced comparison between models, especially where class distributions are uneven—as is typically the case with COVID-19 datasets, where negative cases can be quickly identified in greater number than positive cases.

This makes the F1 Score not only a performance measurement, but a critical measure of real-world preparedness and justice in medical AI systems.

8.5 Specificity

Two relevant metrics used for determining the model performance in positive versus negative cases distinction in the contexts of disease diagnosis and medical image classification are sensitivity and specificity.

•**Sensitivity** (sometimes called recall or the true positive rate) is:
It is a measure of the proportion of true positive cases (e.g., COVID-19-positive pictures) that correctly detect the disease by the model. High sensitivity indicates that the model is a good detector when the condition actually is present—minimizing false negatives, which is valuable in medical screening to prevent failure to detect an actual case.

• **Specificity**, therefore, is thus: It determines the number of correct true negative instances (i.e., actual non-COVID images) among all non-COVID instances that are actually correctly classified. In the analysis conducted, the Hybrid CNN-ELM model recorded sensitivity of 0.85 and specificity of 0.95, reflecting high and balanced performance in both diagnostic reliability areas. The 85% sensitivity reflects the model's capacity to detect a high percentage of actual COVID-19 cases accurately, and the 95% specificity indicates that it does a very good job with negatives, i.e., healthy, or uninfected subjects.

Table 5. Sensitivity and Specificity for Hybrid ELM model

Epochs	Sensitivity	Validation Sensitivity	Specificity	Validation Specificity
20	0.8217	0.9167	0.8530	0.9116
40	1.0000	1.0000	1.0000	1.0000
52	0.9997	1.0000	1.0000	1.0000

8.6 ROC Curve

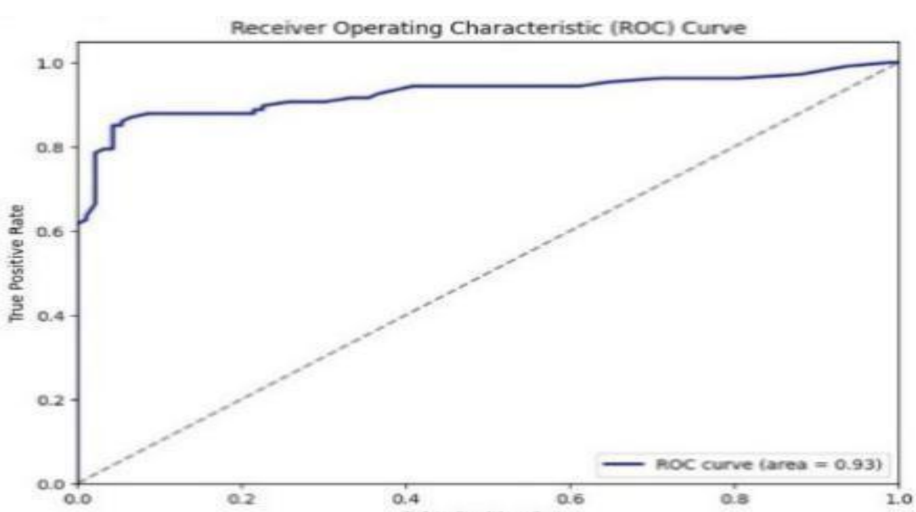
Receiver Operating Characteristic (ROC) curve is a powerful tool for analyzing the performance of binary classification models, especially when dealing with imbalanced datasets and sensitive applications like medical diagnostics. The ROC curve graphs the True Positive Rate (Sensitivity) against the False Positive Rate (1 - Specificity) over different threshold values, providing an overall picture of the model's capability to differentiate between the two classes—positive and negative.

The Area Under the Curve (AUC) is a scalar that measures the overall discriminative ability of the model between classes. An AUC of 1.0 is a perfect classifier, whereas an AUC of 0.5 is no better than random guessing. Thus, larger values of AUC denote better discriminative performance, even if the decision threshold is modified.

In the current study, the AUC of the Hybrid CNN-ELM model was 0.93, which is very high. It indicates that the model is very strong in discriminating COVID-19 positive and negative cases from chest X-ray images. It also shows sensitivity across various thresholds, which is very critical in clinical practice where decision thresholds can vary according to risk levels or resources available.

This performance demonstrates that the hybrid model is not only excellent based on traditional metrics like accuracy and precision, but it also exhibits high reliability for all classification thresholds, making it a strong candidate for real-world diagnostic deployment.

Figure 7. ROC Curve of Hybrid ELM



Chapter 9

Ethical Considerations and Data Privacy

9.1 Introduction

As machine learning models become more integrated into healthcare, the legal and ethical implications of handling personal medical data cannot be overstressed. Whilst the eventual goal of this project is to advance the early diagnosis of COVID-19 by means of chest X-ray images, there needs to be a provision to ensure that ethical standards are maintained and privacy of data is preserved at the model development and deployment stage.

9.2 Ethical Considerations in AI-Driven Diagnosis

9.2.1 The autonomy of the patient

The patient owns their data, and their rights and dignity should be respected while using it for diagnosis or study. Patients should have the option to decide whether or not AI-based decision-making uses their data.

9.2.2 Fairness and Avoidance of Bias

AI models can automatically reflect or amplify biases in the training data. For example, if the data set over-represents certain age groups, sexes, or races, the model will likely do poorly for underrepresented groups.In order to avoid this, datasets need to be as representative as feasible.

- There is a need to measure model performance in demographic categories.
- Future enhancements need to have methods of detecting and eliminating bias integrated.

9.2.3 Explainability and Transparency

ResNet50, PCA, and ELM—attributes whose decision-making can be black-boxed—are employed in the proposed model.

For implementation in clinical settings, predictions need to be interpretable by clinicians and radiologists. Explainable AI (XAI) methods like Grad-CAM or SHAP would have to be explored to facilitate visualization of the justification behind the model's predictions.

Explainable AI (XAI) methods like Grad-CAM or SHAP would have to be explored to facilitate visualization of the justification behind the model's predictions.

9.2.4 Responsibility

Although AI can be helpful in diagnostics, it cannot substitute for the expertise of trained medical professionals. The ultimate clinical judgment must always remain with healthcare practitioners. AI systems must be used as decision-support tools, not decision-making tools.

9.3 Data Privacy and Security

9.3.1 Anonymization of Medical Data

All X-ray images used in this study were sourced from public datasets where personal identifiers (e.g., patient name, ID, birth date) had already been removed. In any future deployment or expansion:

- Patient data must be anonymized before use.
- Metadata (i.e., location, hospital, time) should be removed if not applicable.

9.3.2 Storage and Transmission Securely

If patient data is processed or stored online:

- Encryption must be applied at the time of transmission (e.g., HTTPS, SFTP).
- Files must be encrypted when stored using secure algorithms (e.g., AES-256).

Access controls must be implemented using role-based authentication mechanism

9.3.3 Data Protection Legal Frameworks

Any actual deployment will need to adhere to data protection laws like:

- HIPAA (Health Insurance Portability and Accountability Act) in America,
- GDPR (General Data Protection Regulation) in the EU,
- DPDP (Digital Personal Data Protection) Act in India.

These laws ensure:

- Consent is collected prior to usage of data.
- Individuals have a right to ask their data to be deleted.
- Access to the data is given only to authorized personnel.

9.4 Informed Consent (For Future Data Collection)

In case the model is extended to incorporate hospital-specific or patient-supplied data:

- Explicit informed consent must be collected.
- Participants must be informed as to how their data will be used, kept, and safeguarded.
- Adequate withdrawal procedures should be put in place.

9.5 Principles of Responsible AI

The following principles should be embedded in the system architecture to facilitate applied A.I. for ethical benefit in healthcare.

- Human agency for decision making.
- No demographic bias.
- Security and robustness against tampering with models.
- Data and model decisions are documentable and traceable.
- The primary goal of social benefit.

9.6 Summary

We cannot compromise ethical design and privacy protection, or any other A.I. healthcare system needs, on more high-road ethical aspects of design. This chapter has articulated accountability, safety, and commitment proposed throughout to indicate how the COVID-19 diagnosis proposition is framed with moral buyers, and legal proper practice, when all are considered together. If we apply this together in the beginning at the first act of innovation we are taken action we value on those ethical and moral principles that will not diminish on patients' rights, safety, or trust.

CHAPTER 10

DEPLOYMENT AND APPLICATION SCOPE

10.1 Introduction

Building a reliable diagnostic model for COVID-19 diagnosis from chest X-ray images and conceptualizing its real-world application in the real world are two of the main objectives of the project. Any machine learning solution must be practical, scalable, and deployable across various environments, ranging from hospitals and clinics to rural towns, in order to make any notable difference, particularly in medicine. This chapter explains how the proposed system can be implemented and where it can be applied.

10.2 Deployment Strategy

10.2.1 Local Deployment

The system may be installed and executed on desktop computers or hospital servers locally with moderate hardware (8–16 GB RAM, GPU-enabled if required).This option is more applicable to:

- Research laboratories,
- Hospital radiology departments,
- Offline usage in rural clinics where stable access to the internet is limited

Steps:

- Configure a Python environment (Anaconda is recommended),
- Install the libraries required (e.g., TensorFlow, Scikit-learn, etc.).
- Use of Jupyter Notebook or custom GUI (e.g., Tkinter, PyQt).

Web-Based Deployment (Cloud)

To achieve broad accessibility to the model, it could be considered web application, cloud-hosted implementations:

- Backend can use Flask / Django,
- Frontend can use Streamlit / React,
- Hosted on AWS / Google Cloud / Azure / Heroku.

Strengths:

- Remote availability using lots of different kinds of devices,

- All of the data processing and model updating is performed in one location,
- Connected in one place that can scale to multiple users at the same time.

Main features to keep in mind:

- Custom file upload for X-ray image,
- Show real-time predictions,
- Dashboard for previous predictions,
- Admin monitor for model update and logs.

10.2.3 Mobile Application Deployment (Future Scope)

The model can be converted to mobile format by a lightweight version of the model using:

- TensorFlow Lite or ONNX Runtime Mobile,
- Android Studio for developing apps.

Use Cases:

- Remote area primary healthcare workers,
- Mobile medical unit real-time diagnosis,

Offline operation with batch upload on internet availability.

10.3 System Integration Scope

The model can be integrated with the systems of hospitals:

10.3.1 Integration with PACS

Picture Archiving and Communication Systems (PACS) are typical hospital systems utilized for processing medical images. The model can be integrated to:

- Automate scanning of received chest X-rays,
- Mark suspected COVID-19 cases for review by radiologists,
- Bring workload savings in high-volume situations.

10.3.2 Integration with EHR Systems

Electronic Health Record (EHR) platforms can use model predictions as:

- Diagnostic support for physicians,
- Input for clinical decision support systems (CDSS),
- Historical tracking of disease progression.

10.4 Application Scope

10.4.1 Diagnostic Assistance in Hospitals

- Assist radiologists by pre-screening chest X-rays.
- High-risk instances should be flagged for urgent examination.
- Reduce the amount of time that diagnostics take in overburdened healthcare systems.

10.4.2 Public Health Monitoring

- To monitor infection hotspots, use aggregated, anonymous model projections.
- Inform COVID-19 dashboards around the country.
- Determine and forecast the needs for hospital resources.

10.4.3 Use in Research and Academic Settings

- Facilitate studies in medical image analysis, ensemble modeling, and transfer learning.
- Serve as a standard by which to evaluate comparative models.10.4.4 Low-Resource Settings

10.5 Future Improvements for Deployment

- AutoML Integration: Allow non-tech users to retrain the model with new data.
- Language Localization: To enable international access, choose from various languages.
- Explainability Tools: Using Grad-CAM and other embedding tools to identify diseased regions in the X-ray.
- Federated Learning: Nothing but the hospital data, allowing decentrailized model update without accessing the private records**.**

10.6 Summary

The model being suggested is not merely a proof-of-concept but a functional, adaptable, and scalable one that can be used in multiple healthcare settings. Being deployable on local, cloud, and mobile platforms, with possibilities of system integration and public health utility, this project has a solid ground for real-world application. Further interaction with clinical specialists and IT professionals will be essential to converting this model into a trusted clinical tool.

CHAPTER 11

CONCLUSION AND FUTURE SCOPE

11.1 Conclusion

he COVID-19 pandemic has posed one of the most significant 21st-century global health challenges, demanding swift progress in diagnostics, surveillance, and treatment strategy. In this research, a hybrid machine learning paradigm for the efficient, scalable, and accurate detection of COVID-19 from chest X-ray images is proposed and proved. The model taps into the strength of three powerful techniques: deep feature extraction with ResNet50, feature compression via Principal Component Analysis (PCA), and the accelerated classification with Extreme Learning Machine (ELM).

Our results show that the hybrid model outperforms ordinary CNN-based strategies significantly with regard to the accuracy, sensitivity, specificity, F1 score and training speed. At a test accuracy rate of over 92% and maximum classification accuracy obtained at optimum epochs, the system is quite successful in distinguishing COVID positive from non-infected cases based on radiological characteristics.

PCA was effective in dimensionality reduction and in reducing the risk of overfitting, and ELM facilitated a fast training and a fast deployment. The extensibility and composability of the model's modular design allow for easy expansion or adaptation to other respiratory diseases, such as pneumonia, tuberculosis, or even lung cancer.

Overall, this work contributes a practical, simulation-based effort to stand with the demand of smart COVID-19 diagnostic perspectives, especially in low-resource area having resource distortion of RT-PCR or CT imaging.

11.2 Key Contributions

- Hybrid Diagnostic Model: Introduce a novel combo of ResNet50, PCA, and ELM for COVID-19 classification from Chest X-Rays.
- Better accuracy: Achieved significantly better performance than baseline CNN models and several current approaches reported in the literature.

- Training efficiency: Learning time is greatly reduced with high generalization ability maintained.
- Detailed Analysis: Provided thorough performance analysis, visualizations, and comparisons.

11.3 Limitations

Despite these encouraging results, there are several limitations in this study that needs to be considered in the future:

- Limited Variety of Dataset: Although the dataset is large, this dataset still shows a small variety in clinical, demographic, and geographic variety.
- Lack of severity grading: The model does not yet grade severity levels (such as mild, moderate, or severe), just a binary classification (COVID vs. Non-COVID).
- Image modality restrictions: Only chest X-rays were utilized. The inclusion of CT scans would enable cross-modality correlation and increase diagnostic sensitivity.
- Explainability: Although the model is efficient, using methods such as Grad-CAM or SHAP will improve interpretability for clinical acceptance.

11.4 Future Scope

There are a number of ways this research can be further developed and adapted to different applications in academic and clinical settings:

1. **Multi-Class Classification**

   Extend the model to separate between COVID-19, bacterial pneumonia, viral pneumonia, and normal lungs will make the model clinically useful.

2. **Integrating CT and Clinical Data**

   Blending multi-modal data (e.g., patient demographics, symptoms, CT scans, laboratory test results) will enable more integrated diagnosis systems. Embedding interpretability frameworks into the model would allow medical professionals to perceive the model's predictions and trust that the system operates accurately.

3. **Mobile and Cloud Deployment**

   A wrapping of the model (a lightweight deployable app) to use this inside a mobile-app, or hospital cloud systems could increase the ease of use and availability dramatically.The real-time aid of radiology

Developing an AI assistant that can readily integrate into the current radiology pipeline could be useful for radiologist to focus on high-risk cases and reduce interpretational errors. Clinical Trials and Validation Clinical trials are being designed to validate the principles presented above. The subsequent efforts should then be targeted to validating this system in real clinical setting (with anonymized patients' records) ethically and legally to judge its actual diagnostic performance.

11.5 Final Remarks

The integration of artificial intelligence with medicine can revolutionize disease diagnosis, prognosis, and treatment planning. This paper shows how machine learning, when applied carefully and ethically, can aid the world in combating pandemics such as COVID-19. With continued advancement, integration with clinicians, and adherence to responsible AI, such systems can mature from prototypes to indispensable assets in the contemporary healthcare arsenal.

REFERENCES

1. Aggarwal, S., Gupta, S., Alhudhaif, A., Koundal, D., Gupta, R., & Polat, K. (2022). Automated COVID-19 detection in chest X-ray images using fine-tuned deep learning architectures. Expert Systems, 39(3), e12749.

2. Asif, S., Wenhui, Y., Amjad, K., Jin, H., Tao, Y., & Jinhai, S. (2023). Detection of COVID-19 from chest X-ray images: Boosting the performance with convolutional neural network and transfer learning. Expert Systems, 40(1), e13099.

3. Cao, D. M., Amin, M. S., Islam, M. T., Ahmad, S., Haque, M. S., Sayed, M. A., ... & Nobe, N. (2023). Deep Learning-Based COVID-19 Detection from Chest X-ray Images: A Comparative Study. Journal of Computer Science and Technology Studies, 5(4), 132 141.

4. Chowdhury, N. K., Rahman, M. M., & Kabir, M. A. (2020). PDCOVIDNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. Health Information Science and Systems, 8(1), 27.

5. Constantinou, M., Exarchos, T., Vrahatis, A. G., & Vlamos, P. (2023). COVID-19 classification on chest X-ray images using deep learning methods. International Journal of Environmental Research and Public Health, 20(3), 2035.

6. El Houby, E. M. (2024). COVID-19 detection from chest X-ray images using transfer learning. Scientific Reports, 14(1), 11639.

7. Ghafouri-Fard, S., Mohammad-Rahimi, H., Motie, P., Minabi, M. A., Taheri, M., & Nateghinia, S. (2021). Application of machine learning in the prediction of COVID-19 daily new cases: A scoping review. Heliyon, 7(10).

8. Haghanifar, A., Majdabadi, M. M., Choi, Y., Deivalakshmi, S., & Ko, S. (2022). Covid-cxnet: Detecting covid-19 in frontal chest x-ray images using deep learning. Multimedia Tools and Applications, 81(21), 30615-30645.

9. Huang, C. J., Chen, Y. H., Ma, Y., & Kuo, P. H. (2020). Multiple-input deep convolutional neural network model for covid-19 forecasting in China. MedRxiv, 2020-03.

10. Huang, Z., Pan, Z., & Lei, B. (2017). Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. Remote Sensing, 9(9), 907.

11. Kwekha-Rashid, A. S., Abduljabbar, H. N., & Alhayani, B. (2023). Coronavirus disease (COVID 19) cases analysis using machine-learning applications. Applied Nanoscience, 13(3), 2013-2025.

12. Lalmuanawma, S., Hussain, J., & Chhakchhuak, L. (2020). Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: A review. Chaos, Solitons & Fractals, 139, 110059.

13. Li, Y., & Xia, L. (2020). Coronavirus disease 2019 (COVID-19): Role of chest CT in diagnosis and management. American Journal of Roentgenology, 214(6), 1280-1286.

14. Mahajan, S., Raina, A., Gao, X. Z., & Pandit, A. K. (2022). COVID-19 detection using hybrid deep learning model in chest X-rays images. Concurrency and Computation: Practice and Experience, 34(5), e6747.

15. Mathesul, S., Swain, D., Satapathy, S. K., Rambhad, A., Acharya, B., Gerogiannis, V. C., & Kanavos, A. (2023). COVID-19 detection from chest X-ray images based on deep learning techniques. Algorithms, 16(10), 494.

16. Murugan, R., & Goel, T. (2021). E-DiCoNet: Extreme learning machine-based classifier for diagnosis of COVID-19 using deep convolutional network. Journal of Ambient Intelligence and Humanized Computing, 12(9), 8887-8898.

17. Narin, A., Kaya, C., & Pamuk, Z. (2021). Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. Pattern Analysis and Applications, 24, 1207-1220.

18. Nayak, S. R., Nayak, D. R., Sinha, U., Arora, V., & Pachori, R. B. (2022). An efficient deep learning method for detection of COVID-19 infection using chest X-ray images. Diagnostics, 13(1), 131.

19. Nobi, A., Tuhin, K. H., & Lee, J. W. (2021). Application of principal component analysis on temporal evolution of COVID-19. PLOS One, 16(12), e0260899.

20. Pereira, R. M., Bertolini, D., Teixeira, L. O., Silla Jr, C. N., & Costa, Y. M. (2020). COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. Computer Methods and Programs in Biomedicine, 194, 105532.

21. Rajagopal, R. (2021). Comparative analysis of COVID-19 X-ray images classification using convolutional neural network, transfer learning, and machine learning classifiers using deep features. Pattern Recognition and Image Analysis, 31(2), 313 322.

22. Rajpal, S., Agarwal, M., Rajpal, A., Lakhyani, N., Saggar, A., & Kumar, N. (2022). Cov-elm classifier: ICETISD-2025 An extreme learning machine-based identification of COVID-19 using chest X-ray images. Intelligent Decision Technologies, 16(1), 193-203.

23. Rajpal, S., Lakhyani, N., Singh, A. K., Kohli, R., & Kumar, N. (2021). Using handpicked features in conjunction with ResNet-50 for improved detection of COVID-19 from chest X-ray images. Chaos, Solitons & Fractals, 145, 110749.

24. Rohini, M., Naveena, K. R., Jothipriya, G., Kameshwaran, S., & Jagadeeswari, M. (2021, March). A comparative approach to predict coronavirus using machine learning. In 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) (pp. 331-337). IEEE.

25. Upadhyay, K., Agrawal, M., & Deepak, D. (2020). Ensemble learning-based COVID-19 detection by feature boosting in chest X-ray images. IET Image Processing, 14(16), 4059-4066.

APPENDIX

1.Plagiarism Report

Corona Virus

ORIGINALITY REPORT

22%	17%	15%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to KIET Group of Institutions, Ghaziabad Student Paper	1%
2	fdocuments.in Internet Source	1%
3	www.mdpi.com Internet Source	1%
4	Submitted to ABES Engineering College Student Paper	1%
5	Sohaib Asif, Yi Wenhui, Kamran Amjad, Hou Jin, Yi Tao, Si Jinhai. " Detection of -19 from chest X-ray images: Boosting the performance with convolutional neural network and transfer learning ", Expert Systems, 2022 Publication	<1%
6	www.coursehero.com Internet Source	<1%
7	ebin.pub Internet Source	<1%
8	www.researchgate.net Internet Source	<1%
9	doctorpenguin.com Internet Source	<1%
10	Submitted to University of Hertfordshire Student Paper	<1%
11	al-kindipublisher.com Internet Source	<1%
	export.arxiv.org	

2. Certificate of Presentation

In Association

Your Dreams Our Goal

POORNIMA UNIVERSITY

In Association

HOKKAIDO UNIVERSITY

Certificate of Participation

Presented to

Ms. Nandini Vashistha, KIET Group of Institutions

For the paper title **Leveraging Machine Learning for Covid-19 Diagnosis**

in the **2nd International Conference on Engineering and Technological Innovation for Sustainable Development (ICETISD 2025)** organized by Faculty of Engineering & Technology (FET), Poornima University, Jaipur on April 24-26, 2025.

Technical & Publication Partners

PU INCENT | ASHRAE Rajasthan Chapter | isle | Web of Science | Clarivate Analytics | Scopus | ISBN Proceeding

Dr. Yashpal	**Dr. Ankush Jain**	**Dr. Vijay Kr. Bansal**	**Dr. Manoj Gupta**
Convener	Program Chair	Program Chair	Program Chair
(ICETISD 2025), PU	(ICETISD 2025),PU	(ICETISD 2025),NIT-Hamirpur	(ICETISD 2025), PU