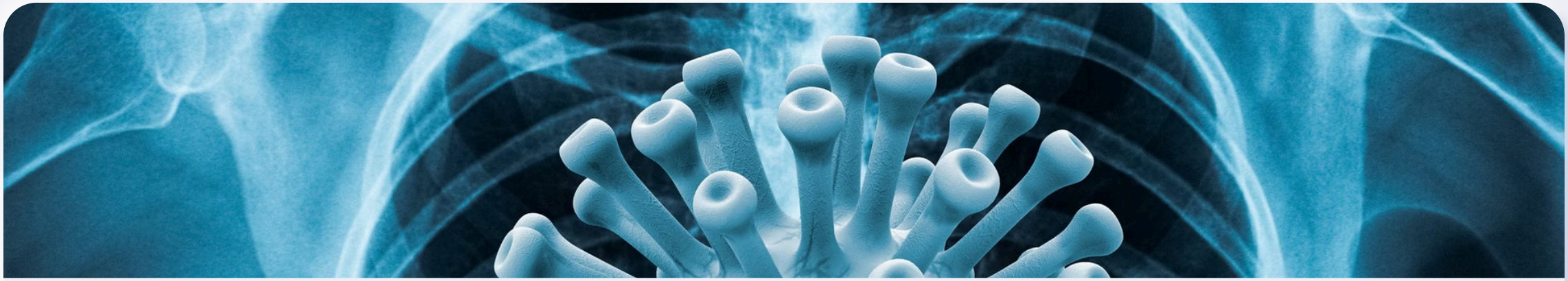
# Covid-19 Detection using Chest X-rays


**Group:**

- M Bharadwaj (521158)
- K Shashi Kumar (521154)
- Y Syamsai (521261)



# Introduction

Our model leverages the power of deep learning to analyze chest X-ray images and accurately identify Covid-19 cases. Through extensive training on a large dataset, the model has learned to detect specific patterns and abnormalities associated with the virus.

In this project, we used a Convolutional Neural Network (CNN), a class of deep learning models most commonly applied to analyzing visual imagery. CNNs are particularly effective for image classification tasks due to their ability to capture spatial hierarchies.

Let's embark on this enlightening journey together and uncover the potential of chest X-ray-based Covid-19 detection!

# Understanding CNN Architecture

## Convolutional Neural Networks

CNNs are deep learning models specifically designed for image recognition and processing.

## Feature Extraction

CNN architecture excels at automatically extracting features from input images.

## Neural Network Layers

CNNs consist of convolutional, pooling, and fully connected layers for learning representations.

## Training and Optimization

Training involves updating the model's parameters using data to optimize performance.

# Importance of Chest X-Rays in Covid-19 diagnosis

- **Early Detection:** Chest X-rays help in early detection of Covid-19 symptoms.
- **Severity Assessment:** They aid in assessing the severity of lung involvement.
- **Monitoring Progress:** X-rays are crucial for monitoring the progress of the disease and treatment.

# Pipeline

1. Data Collection
2. Data Preprocessing
3. Model Building
4. Model Training
5. Model Evaluation
6. Model Deployment

# Data Source

The dataset for this project was obtained from Kaggle.


www.kaggle.com

**Chest Xray for covid-19 detection**

Chest Xray Dataset for detecting covid-19

## Description of the Dataset

The dataset comprises two types of chest X-ray images:

1. Chest X-Ray images from patients infected by COVID-19.
2. Chest X-Ray images from normal cases, i.e., patients not infected by COVID-19.

All the files in the dataset are in either of JPEG/JPG/PNG format.

## Data Split

The dataset was split into three subsets: training, validation, and prediction.

- Training Set: The training set was used to train the model. It contains two subfolders, ‘Covid’ and ‘Normal’, each containing 144 images. This balanced dataset helps in training the model to accurately classify both classes.
- Validation Set: The validation set was used to tune the model parameters and prevent overfitting. It also contains two subfolders, ‘Covid’ and ‘Normal’, each containing 30 images.
- Prediction Set: The prediction set was used to evaluate the final performance of the model. It contains 23 images.

This method of splitting the data ensures that the model is trained and validated on different subsets of the data, which helps in assessing the model’s ability to generalize to unseen data.

# Preprocessing Chest X-Ray Images for CNN

## Data Augmentation

Involves generating new training samples by applying random transformations such as rotation, scaling, and flipping to the existing images. This helps to increase the diversity and robustness of the training dataset, improving the model's ability to generalize and handle variations in real-world X-ray images.

## Loading in Batches

In addition to data augmentation, we load the images in batches using data generators. This approach allows us to efficiently manage memory and process large datasets. The data generators enable us to load and preprocess a batch of images at a time, feeding them into the CNN model during the training process.

## Image Resizing and Normalization

Furthermore, as part of the preprocessing pipeline, we resize the images to a consistent size and perform normalization to ensure standardized input for the CNN architecture.

# Gaussian Filtering

## 1 Read Image in Grayscale

- We start by reading the chest X-ray images in grayscale.
- Grayscale simplifies the image, reducing it from 3 color channels (RGB) to one channel.

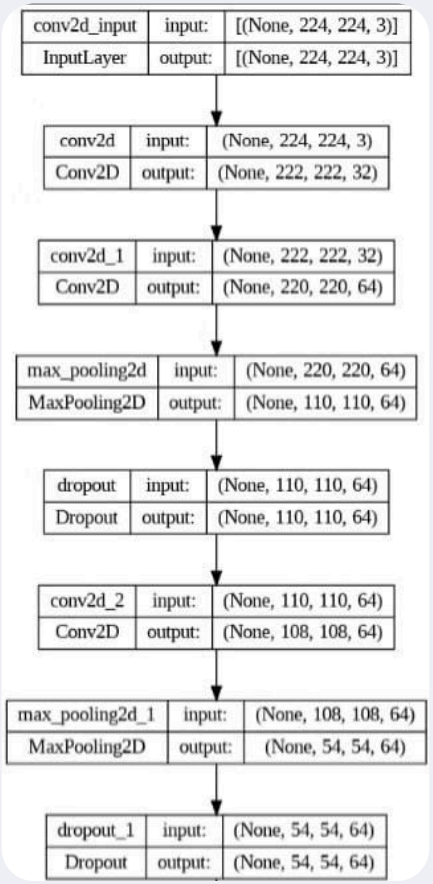
## 2 Add Extra Channels Using Gaussian Filters

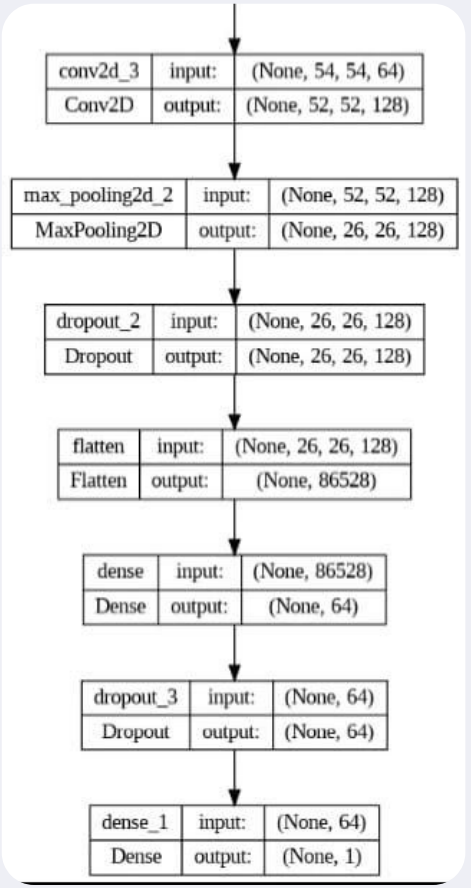
- We then add two extra channels to the grayscale image.
- These channels are created by applying Gaussian filters to the original image.
- Gaussian filters help to highlight the details in the image by blurring.
- By using filters with different standard deviations (sigma), we can capture features at various scales.

## 3 Result: Enhanced Image Data

- The result is an image with 3 channels - the original grayscale image and two filtered images.
- This enriched image data can help the model to better learn and capture the features for detecting COVID-19 in chest X-ray images.

# Model Architecture



# Model Training

To improve the training process and monitor the model's performance, we utilized two important callbacks: early stopping and model checkpoint.

## 1 Early Stopping Callback

The early stopping callback allows us to stop training the model if it's not improving anymore. It monitors a specified metric (e.g., validation loss) and stops training if the metric doesn't improve for a certain number of epochs.

## 2 Model Checkpoint Callback

The model checkpoint callback saves the best model weights during training. It monitors a specified metric (e.g., validation loss) and saves the model weights whenever the metric improves. This ensures that we have the best model at the end of training.

# Model Training and Evaluation

## Loss Function

Binary Cross-Entropy

- Suitable for binary classification problems.
- It quantifies the difference between the predicted probabilities and the actual class.

## Optimizer

Adam

- An adaptive learning rate optimization algorithm.
- It’s efficient and requires little memory.

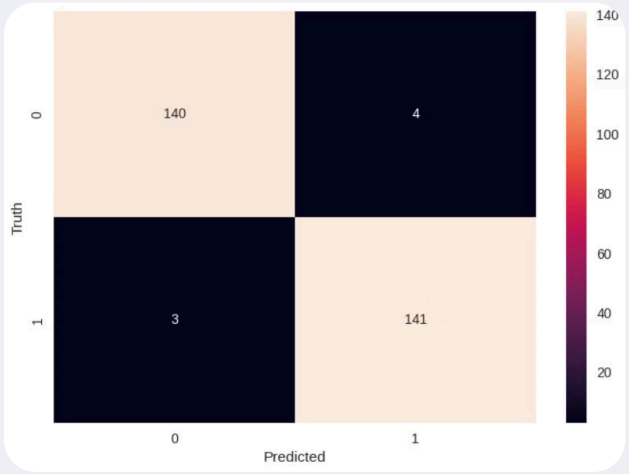
## Metrics

Training and Validation Accuracy

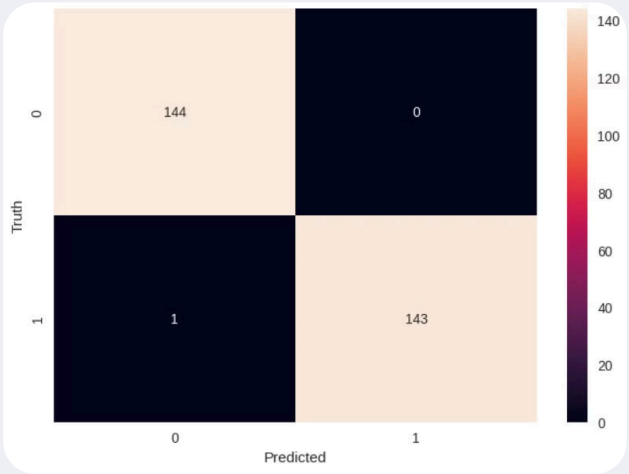
- Accuracy is the ratio of correct predictions to total predictions.
- Training accuracy indicates how well the model learned from the training data.
- Validation accuracy shows how well the model generalizes to new, unseen data.

# Results

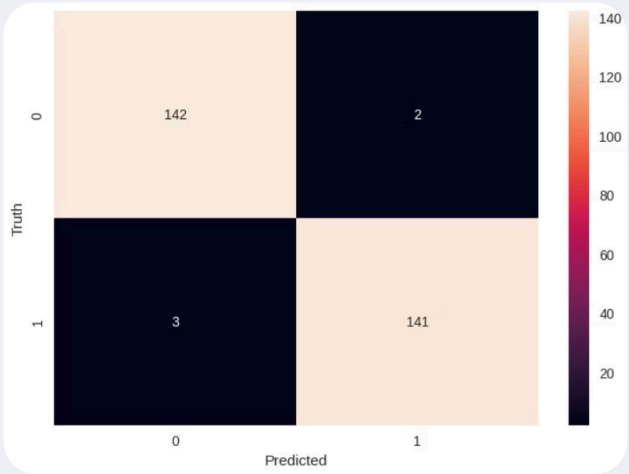
| Model | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| **Custom Model** | 97.57 | 97.90 | 97.22 | 97.56 |
| **VGG16** | 99.65 | 99.31 | 100 | 99.65 |
| **InceptionV3** | 98.26 | 97.93 | 98.61 | 98.27 |




Custom Model




VGG16




InceptionV3

# Challenges and Limitations of Using Chest X-rays for Covid-19 Detection

### Accuracy

The challenge lies in ensuring the accuracy and reliability of Covid-19 detection using chest x-rays.

### False Negatives

There is a risk of false negatives, where chest x-rays may not consistently detect Covid-19 infections.

### Human Interpretation

Understanding the CNN's decision-making process for COVID-19 detection remains a challenge in medical contexts.

# Future Developments and Research Opportunities

1

## Expanded Dataset

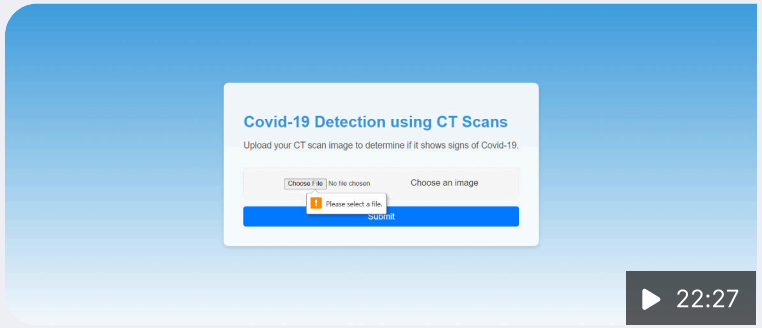
Collecting a larger dataset from diverse demographics for better representation.

2

## Clinical Deployment

Preparing the model for integration with healthcare systems for real-time diagnosis.

# Project YouTube Link

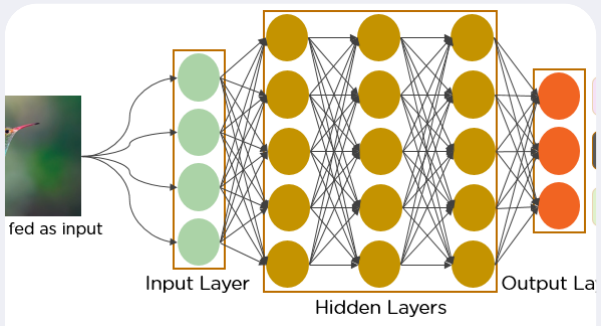



YouTube

**Covid-19 Detection using Chest X-rays Deep Learning project**

#ml #machinelearning #deeplearning #cnn #mlproject #mlprojects

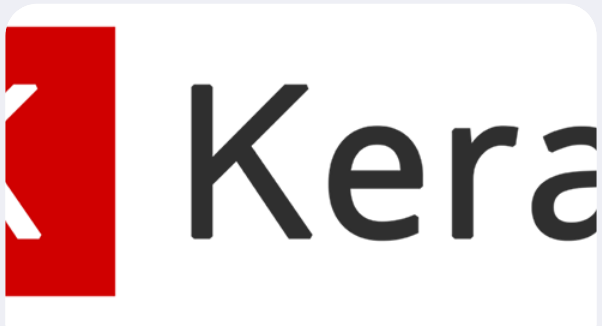
# References




Analytics Vidhya

**Introduction to Convolutional N...**

Learn about Convolutional Neural Networks (CNNs) for understanding...




keras.io

**Keras documentation: Convolut...**

Keras documentation

Datagen

**Image Classification Using CNN...**

Image classification is the task of assigning a label to an input image,...