1 Methodology

1.1 Dataset

In this study we used dataset from kaggle[1]. The dataset comprises subfolders for each image category (Pneumonia/Normal) and is arranged into three folders (train, test, and val). There are two categories (Pneumonia/Normal) and 5,863 X-ray images (JPEG). Anterior-posterior chest X-ray images were chosen from retrospective cohorts of pediatric patients from Guangzhou Women and Children's Medical Center, Guangzhou, aged one to five. Every chest X-ray image was taken as a standard clinical procedure for the patients. we used this dataset because it it a large dataset of chest X-ray images with labels indicating whether each image contains pneumonia or not.

Table 1: Dataset Distribution Normal/Pneumonia

Subset	**Normal**	**Pneumonia**
Training	1341	3875
Validation	8	Row 8
Testing 1	234	390
Total	1583	4272

In the figure.1 , The normal chest X-ray (left panel) depicts clear lungs without any areas of abnormal pacification in the image. Bacterial pneumonia (middle) typically exhibits a focal lobar consolidation, in this case in the right upper lobe (white arrows), whereas viral pneumonia (right) manifests with a more diffuse "interstitial" pattern in both lungs.

Figure 1: Chest X-ray comparisons—Normal (left), bacterial pneumonia (middle), and viral pneumonia (right)

1.2 Data Pre-Processing

Resize all image to a consistent size and normalize pixel value. In order to feed input images into neural networks that require fixed input dimensions, resizing makes sure that each image has a consistent size. Normalization improves an image, enlarging its brightness to fill the entire dynamic range to reduce the distribution of noise.

split the dataset into training, validation, and test set then train the model using the training set and validate it on the validation set. Training a model to identify patterns and features in chest X-ray images that are suggestive of the presence or absence of pneumonia is the first step in using convolutional neural networks (CNNs) to detect pneumonia.

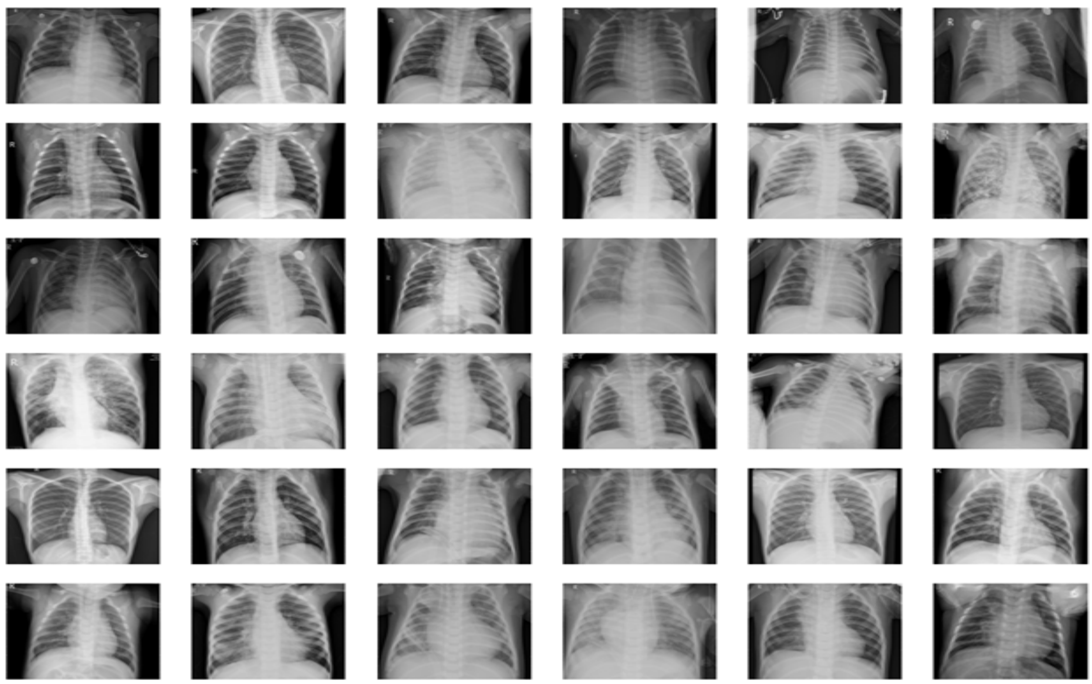

Figure 2: Example of Dataset After Pre-Pocessing

1.3 Model Architechture

The input data is fed into the model for feature extraction following the pre-processing stage. Features from the images can be extracted in this step and fed into the classification for use in the following prediction processes.Two deep learning models, RestNet50 and VGG16, have been used in this work. These models include input, convolution, pooling, dense, and output, and they are responsible for a more thorough and effective feature extraction procedure.The ReLu and sigmoid activation functions, different optimizers algorithm , and 0.0001 (learning rate) are all used.

Following the extraction of features, the training dataset—which comprises roughly 70 percent of the total data and 30 percent of the test data set—is used to train the entire model. Figure 3 illustrates the suggested framework

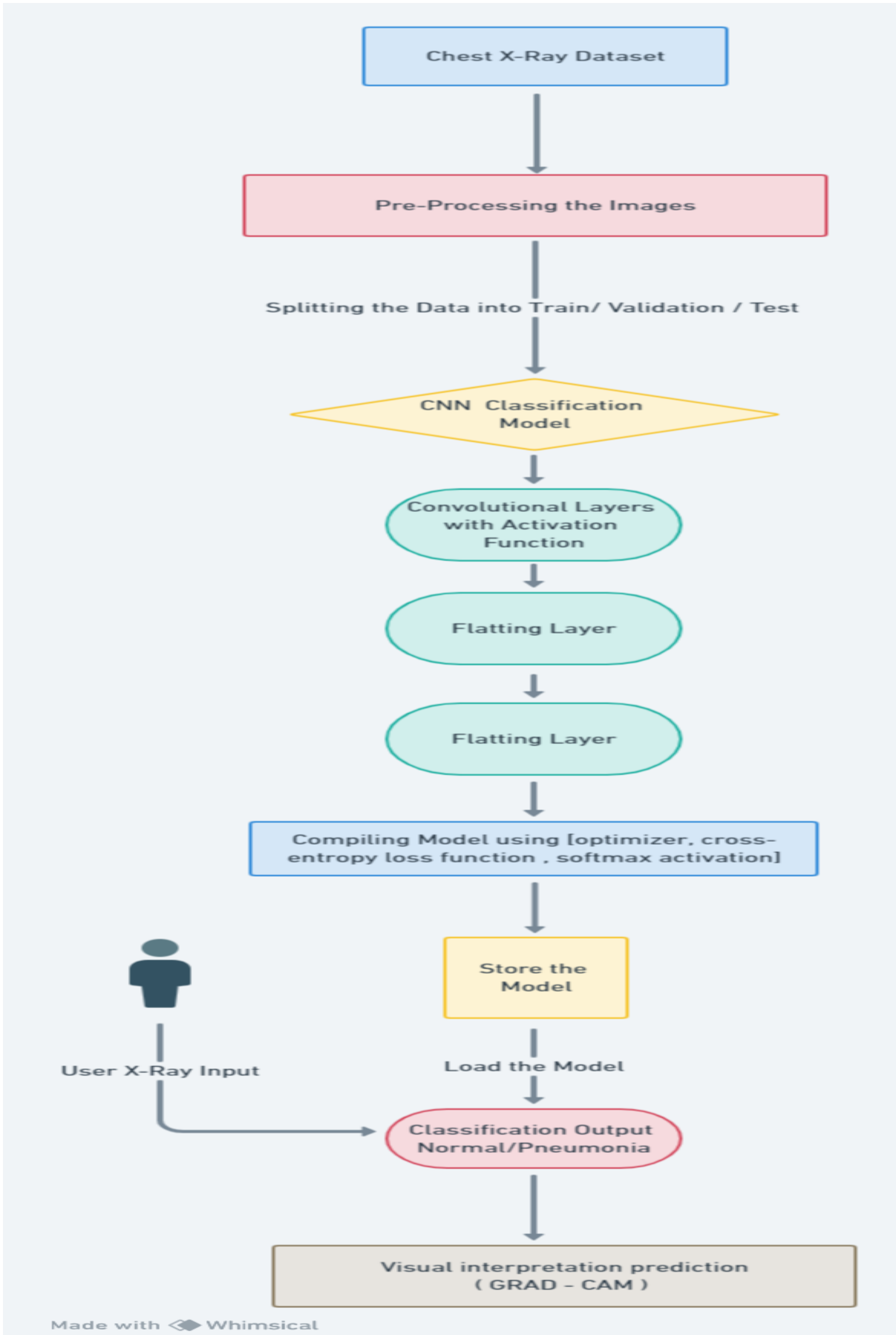

Figure 3: Proposed Framework Model for Pneumonia Prediction

1.4 Evaluation Metrics

As indicated in Table 1, the estimated results of the suggested model are based on a number of measures, including accuracy, precision, F1-score, and recall. Four concepts, "false positive," "true positive," "false negative," and "true negative," need to be clearly defined in order to use these measurements. The term "false positive (FP)" describes samples that are expected to belong to positive classes but actually belong to negative classes. Samples that are both positive and members of the positive class are referred to as "true positives (TP)". Samples classified as "false negative (FN)" are those that are expected to belong to the negative classes but really belong to the positive classes. Samples that are successfully predicted and correspond to the negative classes are referred to as "true negatives (TN)".

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

1.5 Prototype results

RestNet50: A number of evaluation metrics were calculated with the lowest possible error value, and the suggested model performed well in terms of pneumonia prediction. With an overall accuracy of 97 percent, precision of 98 percent, and recall of 96 percent, the model has projected a far superior performance; the F1-score for this dataset, as shown in Table 2, is 97 percent.

Model	Evaluation Metrics			
	Accuracy	Precision	Recall	F1 Score
RestNet50	0.97	0.98	0.96	0.97

Table 2: Model Evaluation Results

For this dataset, the outcomes are also shown in a ROC graph, as shown in Fig. 4, which is helpful in displaying the degree of separatability. ROC highlights the correlation between recall and precision. It is used to graphically represent the evaluation of binary classification. Other ways produce a single value to represent performance. The ROC curve is a well-known curve for identifying a good classifier; the closer the ROC is to the upper left of the graph, the better the model.

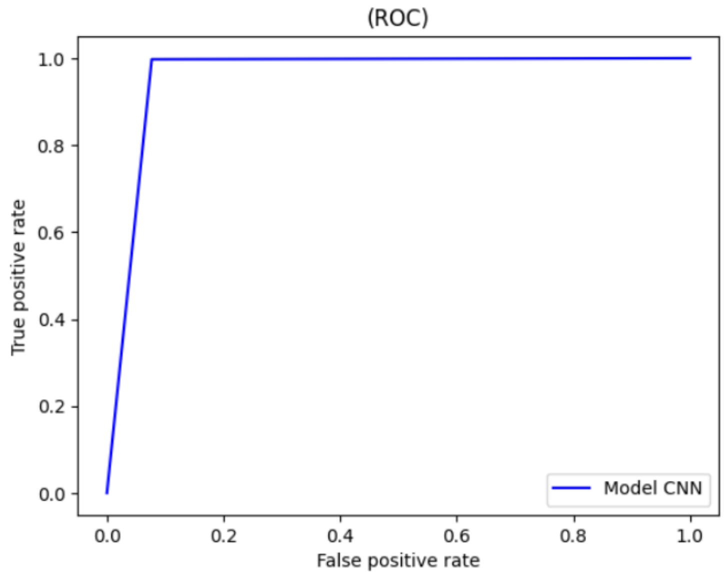

Figure 4: ROC Curve for RestNest50

Figure .5 We evaluate our models by computing accuracy and accuracy validation. It measures the ratio of correctly predicted instances to the total instances. Accuracy is a common metric for classification problems. In addition to Monitor the training and validation loss during the training process. The model is learning when the training loss decreases, but to make sure the model is not overfitting, it is crucial to keep an eye on the validation loss.

1.6 Model Interpretation

Using a technique called Grad-CAM (Gradient-weighted Class Activation Mapping), one can see which areas of an image are crucial for the prediction of a deep learning model. In this study we Implement Grad-CAM techniques to visualize the regions of the image that are important for the model's prediction and localize disease specific region in chest x-ray images aiding in the models interoperability.

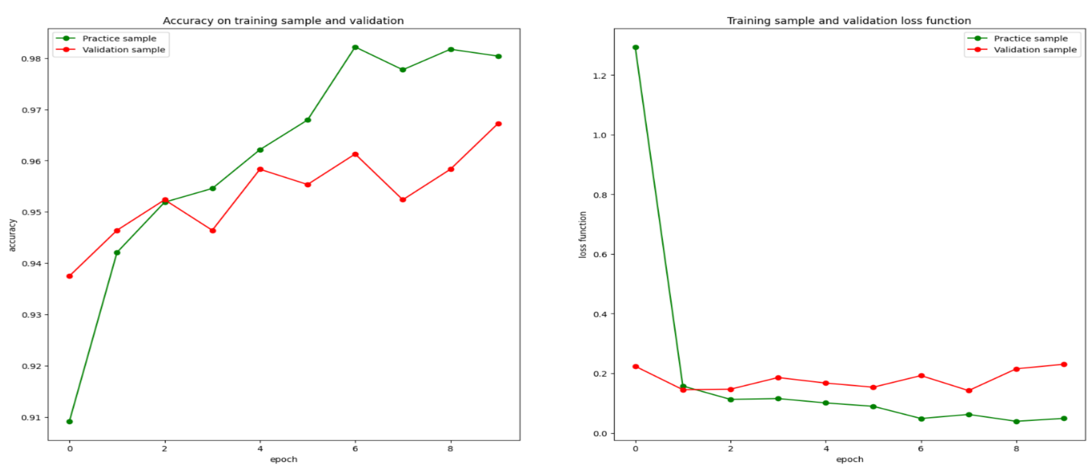

Figure 5: Accuracy and Loss Function on Training and Validation