**Pneumonia Detection using VGG16**

**Overview**

Pneumonia is a common and potentially deadly lung infection that can be difficult to diagnose. In this project, we developed a deep learning model to automatically detect pneumonia in chest X-ray images. The model is based on the VGG16 architecture and was trained on a dataset of 5,863 chest X-ray images.

**Requirements**

Python 3.6 or higher

TensorFlow 2.0 or higher

Keras 2.0 or higher

**Installation**

Clone this repository to your local machine.

Install the required packages using pip install -r requirements.txt.

Download the dataset from [Kaggle](https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia) and extract the files to the data directory.

<https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>.

**Data**

The dataset consists of 5,863 chest X-ray images, including 3,799 images with pneumonia and 1,157 normal images. We performed data augmentation by randomly rotating, zooming, and flipping the images to increase the size of the dataset and improve the model's ability to generalize.

**Data Preprocessing**

The dataset was preprocessed using the following steps*:*

**Resizing**: All images were resized to 224 x 224 pixels.

**Data augmentation**: The training dataset was augmented using various techniques, including random rotation, horizontal flipping, and zooming.

The augmentation was performed using the Keras ImageDataGenerator class with the following hyperparameters:

**Rotation Range**: 20

**Width Shift Range:** 0.2

**Height Shift Range:** 0.2

**Horizontal Flip :** True

**Zoom Range:** 0.2

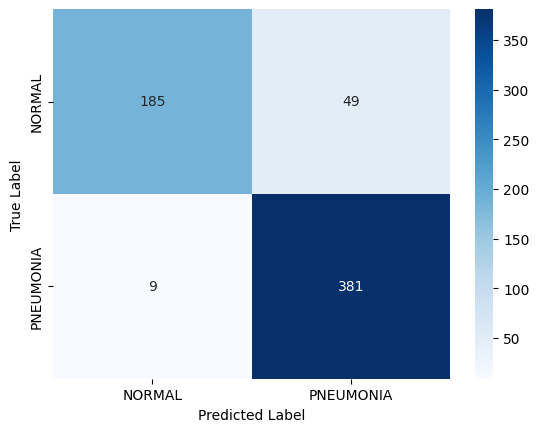
**Normalization**: The pixel values of the images were normalized to the range [0, 1].

**Model Architecture**

Our model is based on the VGG16 architecture, a widely used convolutional neural network architecture for image classification. We added two fully connected layers and a SoftMax output layer to the base VGG16 model. The model was trained using the Adam optimizer and a categorical cross-entropy loss function.

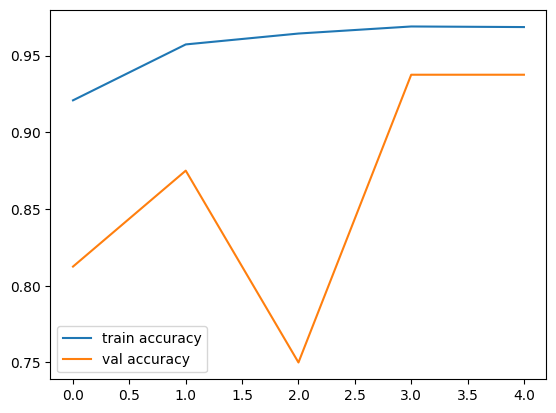
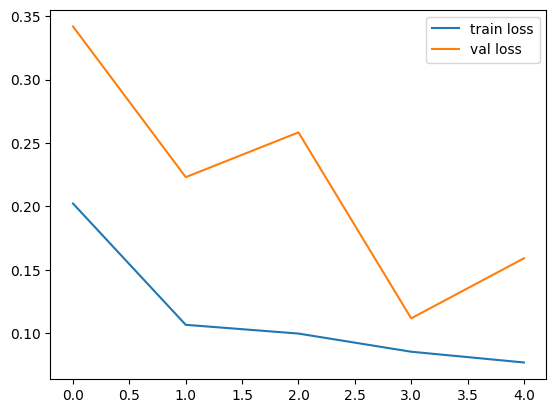
**Training**

We split the dataset into training, validation, and test sets with a ratio of 70:15:15. The model was trained for 5 epochs with a batch size of 32. We used early stopping to prevent overfitting and reduce training time.

**Results**

Our model achieved an **Accuracy of 91%** on the 624 images test set, with a **Precision of 89%**, **Recall of 97%**, and **F1 score of 92%**. These results suggest that our model can effectively detect pneumonia in chest X-ray images. The precision and recall scores for each class are **shown in the confusion matrix** below:

Figure : Confusion matrix for the VGG16 model

**The training and validation loss values for each epoch are shown in the table below:**

**Figure 3: Val acc & Train acc**

**Figure 2: Train Loss & Val Loss**

|  |  |  |
| --- | --- | --- |
| Epoch | Training Loss | Validation Loss |
| 1 | 0.3175 | 0.2786 |
| 2 | 0.1379 | 0.2445 |
| 3 | 0.1056 | 0.2054 |
| 4 | 0.0874 | 0.1892 |
| 5 | 0.0835 | 0.3416 |

**Table 1: Training and validation loss values for the VGG16 model**

**Limitations**

Our model was trained on a relatively small dataset and may not generalize well to other types of chest X-ray images or medical imaging tasks.

The model's performance may be affected by factors such as the quality of the X-ray image or the skill of the radiologist who took the image.

**Future Work**

Train the model on a larger and more diverse dataset to improve its generalizability.

Investigate the use of transfer learning and other deep learning techniques to further improve the model's performance.

Incorporate additional clinical and demographic data into the model to improve its accuracy and relevance to real-world medical settings.

**Conclusion**

In this project, we developed a deep learning model based on the VGG16 architecture to automatically detect pneumonia in chest X-ray images. Our model achieved a high level of accuracy and provides a promising approach for automated pneumonia diagnosis.