**Methodology**

**Dataset:** The chest X-ray image dataset for the COVID-19 classification was downloaded from GitHub. For the pneumonia diseases, the dataset of X-ray images was collected from Kaggle (a huge source of public datasets). Both datasets were merged to form a collective dataset of COVID-19 and pneumonia chest X-ray images. The collected dataset contains the 468 COVID-19 images and 3875 pneumonia images. Pneumonia images were further divided into two classes (Bacterial pneumonia and Viral pneumonia). Lastly, the dataset was base on four classes and the images count of each class is available in Table 1.

Table 1: Number of total images in each class of prepared dataset.

|  |  |  |
| --- | --- | --- |
| Class Label | Class Name | Total images |
| 0 | Normal | 1341 |
| 1 | Bacterial Pneumonia | # |
| 2 | Viral Pneumonia | # |
| 3 | COVID-19 | 468 |

**Preprocessing:** All the X-ray images were into gray-scale images. As the dataset of chest X-ray images was composed of large size images, so the most significant step of preprocessing was the resizing of the images. All the images were resized to 224 x 224 x 1 to deal with reasonable computational time during the training of deep learning model. Deep learning models required the sufficient amount of data for the reliable training and excellent result of the model. The augmentation technique was applied to increase the dataset to a reasonable amount. The vertical flip, horizontal flip, rotation, translation and blurring techniques were used for the augmentation of data. Lastly, we converted the augmented grayscale image to 3 channel images for fine tuning.

**Model:** For the Classification of COVID-19 and pneumonia, the pretrained model of PyTorch (cheXNet) was used for training. CheXNet model is based on 121 convolutional layer (Dense Network of DenseNet) and trained on the 14 dataset of Chest X-rays images. DenseNet uses the pretrained weights of ImageNet and replace the fully connected layer with the single output layer. It also replaces the SoftMax function with the nonlinearity Sigmoid function. The hyperparameters of the CheXNet were: optimizer=Adam, batch size=16, and learning rate=0.001 for training. The count of training and testing images is available in table 2.

Table 2: Count of X-ray images in training and testing set after augmentation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class Label | Class Name | Total images | After Augmentation | Training | Testing |
| 0 | Normal | 1341 | # | # | 233 |
| 1 | Bacterial Pneumonia | # | # | # | 242 |
| 2 | Viral Pneumonia | # | # | # | 148 |
| 3 | COVID-19 | 468 | # | # | 30 |

The CheXNet pretrained model was selected as base model for training. The dataset was split into 80% and 20% into training and testing set respectively. For the training of the model on 4 classes rather than 14 classes the output layer of the model was modified. The split of 0.2% of training set was used for validation during the training of the model. The Adam optimizer with batch size 8 was used for the training of the model. Later, the model was fine-tuned by changing the learning rates and optimization algorithm. The model was tested on the best fine-tuned model by calculating the evaluation measures using Eq 1-4. In Eq 1-4, the TP, TN, FP and FN are abbreviated from true positive, true negative, false positive and false negative respectively. As the false predicted covid-19 X-ray images were more costly for our system, so the main goal of the model was to achieved the high score of recall.

**Results**

In this study, we presented the fully automated deep learning model for the Classification of COVID-19 virus among the other viral viruses. The chest X-ray images of four classes including COvid-19, Bacterial Pneumonia, Viral Pneumonia and Normal images were collected from Kaggle and GitHub (see methodology).

Firstly, the CheXNet pretrained model was trained on three classes (COVID-19, Pneumonia and Normal). We merged the Viral Pneumonia and Bacterial Pneumonia into single parent class and label as Pneumonia. The 80% split of dataset was used for training data and 20% split for testing data. The Adam optimizer with different learning rates was used to trained the model. By comparing the evaluation scores of the model with different learning rates, model performed best with Adam optimizer on 0.001 learning rate. CheXNet showed the approximately 98% accuracy with the significant score of other evaluation measures (Table 3). The model was evaluated on testing data by plotting the confusion matrix (Figure 1). The confusion matrix showed that all the COVID-19 samples are correctly classified. But for the Pneumonia images, model predict the 2 samples of Pneumonia as Normal Chest images. The False positive and false Negative rate of the model confirmed that the model is fair enough to deploy in real world environment. The ROC curve of the trained model is also presented in Figure 2.

Table 3: Evaluation scores of the CheXNet model on three classes.

|  |  |  |  |
| --- | --- | --- | --- |
| **Class** | **AUROC** | **Sensitivity** | **PPV** |
| Normal Lung | 0.9788 | 0.761 | 0.989 |
| Bacterial Pneumonia | 0.9798 | 0.961 | 0.881 |
| Viral Pneumonia | 0.9370 | 0.872 | 0.721 |
| COVID-19 | 0.9994 | 1.000 | 0.938 |

|  |  |
| --- | --- |
| Figure 1: Confusion matrix of CheXNet calculating using testing data. | Figure 2: ROC curve of CheXNet model on three classes using testing data |

Secondly, the dataset with their original classes (c=4) was used to train the CheXNet model. Same configuration settings were followed for the training of the model. The confusion matrix of the model was calculated by the testing data (Figure 3). The accuracy, precision, recall and F1-Score were calculated by confusion matrix in Table 4. The confusion matrix showed that the model still correctly classified the all samples of COVID-19. While the model incorrectly classified the 18 Viral sample as Bacterial and 7 Bacterial samples as Viral. The model also wrongly classified the 18 and 43 Normal Samples as Bacterial and Viral Pneumonia samples. The significant false negative score of the model showed the robustness of trained model and made the model capable to made predictions in real world environment after clinical trials. The ROC curves of CheXNet models on four classes is also presented in Fig 4.

Table 4: Evaluation scores of the CheXNet model on four classes.

|  |  |  |  |
| --- | --- | --- | --- |
| **Class** | **AUROC** | **Sensitivity** | **PPV** |
| Normal Lung | 0.9788 | 0.761 | 0.989 |
| Bacterial Pneumonia | 0.9798 | 0.961 | 0.881 |
| Viral Pneumonia | 0.9370 | 0.872 | 0.721 |
| COVID-19 | 0.9994 | 1.000 | 0.938 |

|  |  |
| --- | --- |
| Figure 3: Confusion matrix of CheXNet calculating using testing data. | Figure 4: ROC curve of CheXNet model on four classes using testing data. |

**Discussion**

In this study the transfer learning technique was used for the classification of Pneumonia and COVID-19 disease with Normal samples using Chest X-ray images. Firstly, the CheXNet model with their pretrained weights was used for the classification of Chest X-ray of three classes (Pneumonia, COVID-19, Normal). The newly trained CheXNet model showed the significant scores for the selected evaluation measures in Table 3. But the model was unable to classify the type of Pneumonia and also classify the 59 Normal samples as pneumonia samples. It showed that the pneumonia is the most deceptive class for the model and model hardly distinguished between Normal and Pneumonia samples. Later, the Pneumonia class divided into two major types of Pneumonia (Bacterial Pneumonia and Viral Pneumonia). Once again, the CheXNet model with their pretrained weights was used for the classification of Pneumonia and COVID-19 disease on four classes (Bacterial Pneumonia, Viral Pneumonia, COVID-19, Normal). After the training of the CheXNet model, the evaluation scores showed slightly decline on testing set. The fall of evaluation score was due to the miss classification of Normal samples. As the model was confused between normal and pneumonia samples in first model and predict some wrong classification for normal class, now the Pneumonia class in further divided into Bacterial and Viral classes. Now the model is more hardly distinguished between Pneumonia classes and Normal samples. That’s why the model wrongly classified to Normal samples into Pneumonia classes and Pneumonias classes into Normal samples. However, for the classification of COVID-19 chest X-ray images each model is robust enough to made prediction on real world data. For the significant result of Pneumonia image classification, there is need to perform some advance image processing steps. These steps will emphasize the affected area of Pneumonia images and distinguished them from the Normal images. Resultantly, the same models will be robust enough to made predictions for Pneumonia images.

**Conclusion**

In this study, we used the Chest X-ray image of Pneumonia and COVID-19 disease from Kaggle and GitHub. For the training of the model, the count of image was increase up to appropriate amount by augmentation. The resizing and scaling technique were used for the training of model in time period. The CheXNet pretrained model was used with pretrained weights. By the transfer learning technique, CheXNet model was trained on three and four classes for the classification of Pneumonia and COVID-19. The CheXNet model showed the 98% and 96% accuracy on the three and four classes respectively. For the COVID-19 class, model showed the accurate results. Although, the Pneumonia class was deceptive for the model and model predict some wrong classification between Pneumonia class and Normal class. However, the misclassification of the Pneumonia and Normal class can be handle by performing some image processing steps in future study.