Chapter 4

Detection of Covid-19 Infection in CT and X-Ray Images using Transfer Learning Approach

4.1 Introduction

The World Health Organisation (WHO) declared the infection caused by severe acute respiratory syndrome coronavirus (SARS-CoV-2) a global pandemic as the infection spread rapidly throughout the globe. SARS-CoV-2 was renamed COVID-19 by the WHO on Feb 11, 2020.[81].

Deep learning-based approaches [82–84] are gaining considerable attention for computer-based detection of COVID-19 as these approaches are capable of extracting the features on their own and produce more reliable result [85, 86]. Performing a specific task in one domain by using the knowledge acquired from other domain is termed transfer learning [87]. In transfer learning-based approaches, the knowledge is transferred from a source to target for improving the task-specific efficiency of the target [83].

The transfer learning-based methods are categorized into three categories, namely 1) Inductive transfer learning, 2) Transductive transfer learning, and 3) Unsupervised transfer learning. The task of source and target are same in inductive transfer learning while in transductive learning, the task of source and target are different. The unsupervised transfer learning is same as inductive learning, but the target tasks are unsupervised. The machine learning and deep learning approaches which utilized transfer learning showed remarkable improvement in the classification task. These methods reveal the features of the infection, which are not clearly visible in the original images as a better understanding of pathophysiology related to COVID-19 is important to improve the treatment process.

The convolutional neural network (CNN) architectures have shown good efficiency in classifying the infected images [88]. CNN-based binary classifier was used by Asnaoui and Chaw [89] to classify the CT images for detecting the infection. RCNN-based classification technique was proposed by Zreik et al. [90] to classify the coronary artery plaque. Li et al. developed the Cov-Net model to extract

the features of COVID-19 and showed a remarkable improvement in the classification of infected images. Melanoma dermoscopy images were classified by Hagerty et al. [91] with remarkable accuracy using the deep learning-based method. VGG-16-based model for identification of lung infection was proposed by Li et al. [92]. Identification of thoracic diseases was made with the incorporation of DenseNet model [34]. An Inception Net-based model was used by Wang et al. [20] to identify abnormalities of COVID-19. An accuracy of 85.20% was achieved by classifying 1065 CT images. Resnet-based transfer learning technique was proposed by Ayrton, who achieved validation accuracy of about 96%. U-Net ++-based model was used by Chen et al.[93] for covid-1p classification, and the accuracy achieved was 98.85%. The pretrained U-net architecture for segmentation of 3D Lung images was used by Zheng et al. [94] and then they applied the DL method for determination of COVID-19 infection. The accuracy obtained by them was 95.9%. A CT analysis method for detection of COVID-19 infection using artificial intelligence was proposed by Gozes et al. [34]. Recently, DCNN-based structure was presented by Linda et al. [95] for detection of COVID-19 from chest X-ray images. The proposed model classified X-ray and CT images with significant increment in the results as compared with other methods. The results were compared using various evaluation metrics. To avoid the ambiguity caused by some metrics, additional metrics are used to confirm the effectiveness of results. Accuracy and F score sometimes misleads the classification judgment , so Matthews correlation coefficient (MCC) is used for better understanding [96]. The proposed classifier learns the features with the help of transfer learning and produces remarkable results for both CT as well as X-ray images. The main advantage of the proposed technique is the reduction in training time and requirement of fewer data along with remarkable improvement in image evaluation statistics. Finding a faster method to identify COVID-19 infected images is the driving force behind our endeavour. In order to save medical professionals' time, it can aid in the automatic detection of COVID-19 images.

4.2 Methodology

The CT and X-Ray images in the presented work are classified using a transfer learning-based method because both image modalities are utilized to identify COVID-19 symptoms. In this work, the SVM classifier is utilized for classification. The features of the images were extracted using MobileNet V2 architecture. The SVM classifier efficiently classified the infected images after receiving the derived features. The feature extraction process also utilized a number of different networks for comparison. As a comparison tool, the SVM classifier also received the

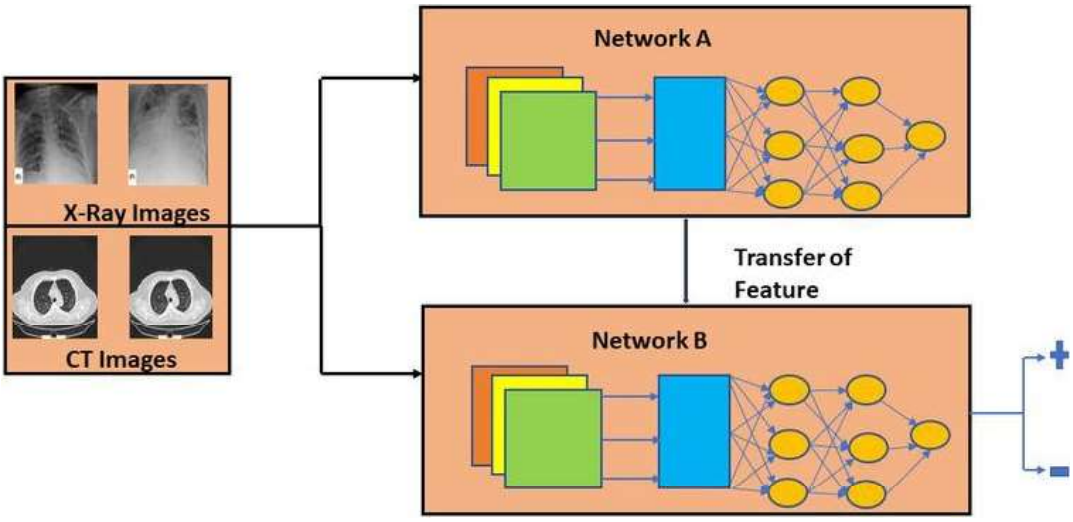

FIGURE 4.1: Graphical Abstract

characteristics that were retrieved by several network designs. Transfer learning was used to expedite the classification of photos that were important from a medical standpoint. The pipeline of the suggested work is depicted in Figure.4.2.

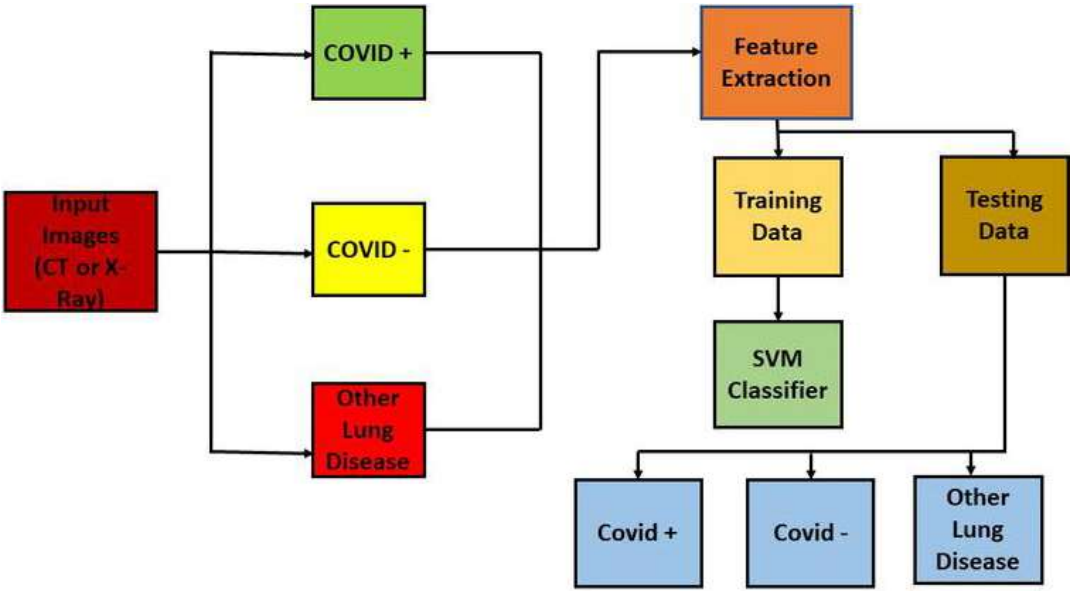

FIGURE 4.2: Pipeline of the proposed work

Transfer learning eliminates the requirement to train the classifier from scratch since it makes use of the knowledge amassed by a network on a large training dataset with labels that are readily available [96]. In the field of machine and deep learning, transfer learning is a type of design

process. Deep learning models typically extract the features of edges in the earliest layers for applications linked to computer vision. The specific features linked to the particular purpose for which the network is designed are extracted by the intermediary layers. Typically, the last layers in a transfer learning approach are taught. This method has the benefit of saving time while training large networks and removing the requirement for big training datasets. Thus, the expertise acquired by one network to do a particular task is utilized to complete a different task. We trained an SVM classifier for the classification of CT and X-ray images infected by COVID-19 using the knowledge gathered by the MobileNet model and other state-of-the-art models for comparison. The entire transfer learning process is depicted in Figure.4.3.

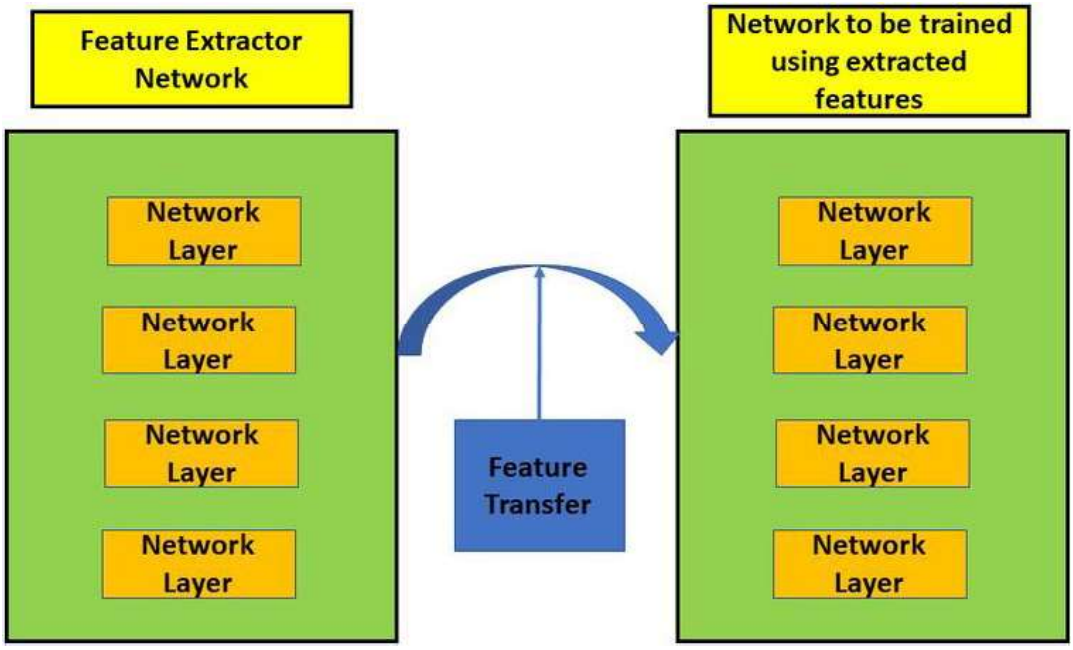

FIGURE 4.3: Feature transfer using transfer learning.

4.2.1 MobileNet-V2

The next-generation architecture for all-purpose computer vision applications is called MobileNet V2 and was created by Google Brain. Low power and low latency models from the MobileNet family are employed to complete a variety of computer vision use cases. The depthwise separable convolution is the fundamental building element for MobileNet V1 . The linear bottlenecks between the layers and the shortcut connections between the bottleneck blocks are included in MobileNet version 2. The in-between inputs and outputs are encoded by the bottlenecks. The middle layers

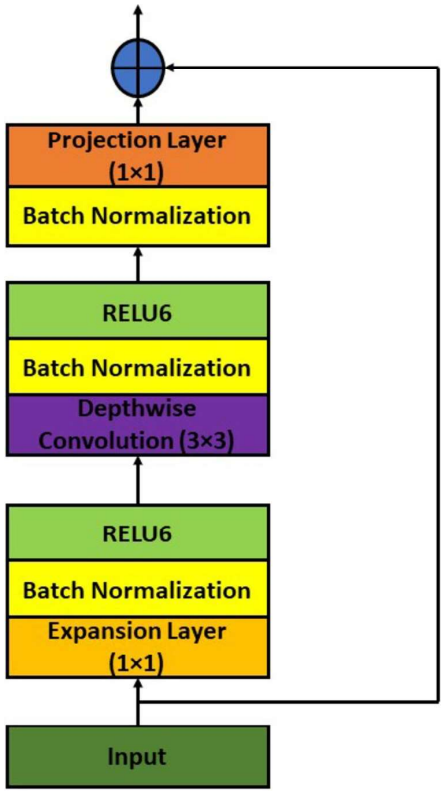

FIGURE 4.4: Depthwise separable convolution block.

are responsible for converting lower-level components like pixels to higher-level images like images. Figure.4.5 depicts the MobileNet V2 bottleneck block that is still in place. Expansion Layer, Depth Wise Convolutional Layer, and Projection Layer are the names of the layers. Depth-wise convolution is incorporated, and each input channel receives a single filter.

A pointwise $1 * 1$ convolution is used to combine the depthwise convolution results. Depthwise separable convolution is the name given to this combination filtering and combining method. Figure.4.4 illustrates how the depthwise separable convolution divides this into two layers, one for filtering and one for combining. Depth-wise convolution with one filter per input channel (input depth) can be written as:

$$\hat{G}_{k,l,m} = \hat{K}_{i,j,m} \cdot \bar{F}_{k+i-1,i+j-1,m} \tag{4.1}$$

where $\hat{K}$ is the depthwise convolutional kernel of size $D_K \cdot D_K \cdot M$ where the m^{th} filter in $\hat{K}$ is applied to the m^{th} channel in $\bar{F}$ to produce the m^{th} channel of the filtered output feature map G. has The convolution cost of depthwise separable convolution is given by the following Equation.4.2

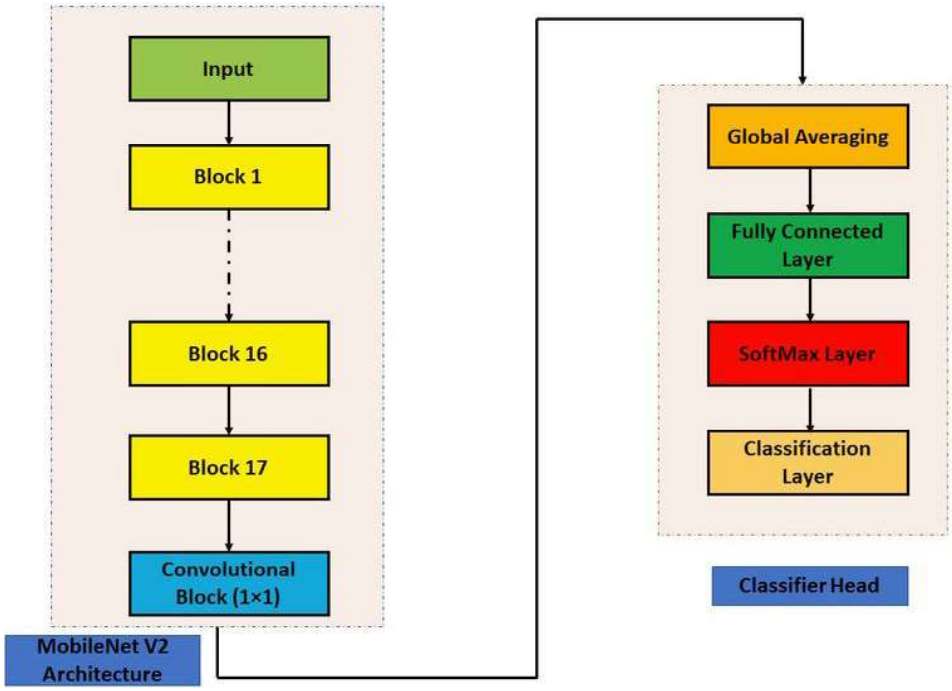

FIGURE 4.5: MobileNet V2 architecture for classification.

$$D_k \cdot D_k \cdot M \cdot D_f \cdot D_f + M \cdot N \cdot D_f \cdot D_f \tag{4.2}$$

Equation.4.2 is the sum of depthwise convolution and 1×1 pointwise convolution where $D_F \cdot D_F$ is the pointwise convolution kernel. The depth-wise separable convolution is represented mathematically altogether in this equation. First part of the equation is depth wise convolution and second part is pointwise convolution. The 1×1 pointwise convolution layer makes the number of channels smaller and is also known as Projection Layer and Bottleneck. The amount of data flowing over the network is decreased. In each layer, Batch Normalization and activation function Relu6 is incorporated. For low-precision applications, RELU6 is more robust than normal RELU. The MobileNet architecture as shown in Figure.4.5, consists of 17 layers of building blocks consecutively which are followed by 1×1 convolutional layer, global average pooling layer, fully connected layer, SoftMax layer, and a classification layer. One feature map is produced by the global average pooling layer for each associated classification category. By calculating the mean of the input's height and width parameters, it conducts down-sampling. The global average pooling layer is positioned before the fully connected layer for image classification tasks in order to minimize the amount of activations

without compromising classification performance. When the activation size is decreased, the network becomes smaller because the downstream fully linked layers will have less weights. The data are finally divided into several classes by the classification layer.

The SVM classifier receives the MobileNetV2 features that were finely extracted. The prime strength of the SVM is its non-probabilistic nature. It separates data across a decision boundary determined by only a small subset of the data. The subset of data that supports the decision boundary is known as support vectors. The other data does not have any effect in determining the position of the boundary. The SVM problem can be understood as a basic optimization problem in which a decision boundary is evaluated which gives the maximum distance between the classes. It first finds the hyperplane that has the maximum margin and this hyperplane is used to predict the class of new data objects. The equation of hyperplane is given as Equation.4.3:

$$w.x = 0 \tag{4.3}$$

The polynomial kernel function was used to train the SVM. The data was split into the ratio of 80:20 for training and testing. The MobileNet V2 extracted the features of COVID-19 from CT as well as X-Ray images. Both types of image features extracted by MobileNet V2 were fed to the SVM classifier for training. The SVM effectively classified the COVID-19-infected images for both modalities, i.e. CT and X-ray. Other five networks were also used to supply the extracted features to the SVM classifier for the identification of images. These networks are Shufflenet, Alexnet, Darknet, Resnet and Inception Net V3. The presented work clearly finds the better method among the compared methods for better identification of COVID-19-infected images. SVM was used as a common platform for all networks. Thus better network can be judged based on the results provided by the SVM. According to the pathological studies of the lung on the autoptic tissues for COVID-19 and viral pneumonia, the widespread thrombosis with microangiopathy and vascular angiogenesis are the main discriminators for texture variations which are efficiently captured by the proposed method [92, 97].

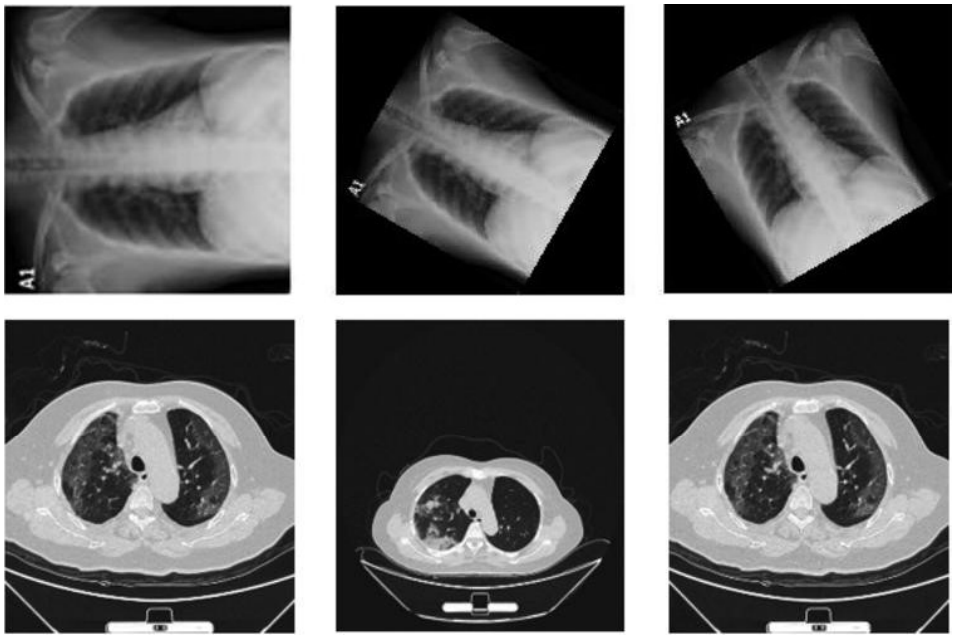

FIGURE 4.6: Sample images of COVID (+) infection. The first row shows X-ray images and the second row shows CT images.

4.3 Dataset Used

4.3.1 X-Ray dataset

In this paper, a different Chest X-Ray and CT dataset was used, as curated by the authors, and is available as open source on Kaggle. It has released 219 COVID-19, 1341 Normal, and 1345 Viral Pneumonia Chest X-ray (CXR) images. In this paper, our goal was to identify the presence of Covid-19, Normal vs Other lung diseases (Viral Pneumonia), and the term 'other lung diseases', which have been used for Viral Pneumonia cases, and It is a multi-class classification problem. In addition to this, 301 images were taken from the GitHub repository of Cohen. To increase the data for training, we augmented the data using data augmentation techniques as the deep learning networks need a large amount of dataset to be trained. The data was rotated at the angles of 300, 600, 900, 1200, 2700, and 3000. In total, we created 3120 COVID-19-infected images, 8046 normal images, and 8070 viral pneumonia-infected images. This augmented data were divided into training and testing.

4.3.2 CT Dataset

The CT dataset is acquired from CT scans for COVID-19 classification database of Kaggle . The database consists of images collected from Union Hospital (HUST-UH) and Liyuan hospital (HUST-LH). The database consists of 5705 non-informative CT (NiCT) images, 4001 positive CT (pCT) images and 9979 negative CT (nCT). Figure.4.6 shows the sample images of COVID-19 infection.

The dataset was resized according to the input requirement of each network. For Mobilenet V2, Shufflenet and Resnet the images were resized to 224×224. For Alexnet, Darknet and Inception V3 the images were resized to 227×227, 256×256, 229×229 respectively.

4.4 Results

This section presents the analysis of the comparative results obtained by the proposed and other existing classifier models. The models used for comparison are Alexnet, Resnet101, Inception V3, Darknet and ShuffleNet. The results obtained are shown in three categories

1. The comparison of accuracy, F scores and MCC.
2. The parameters whose value should be minimum for achieving good performance.
3. other parameters for firm confirmation of results obtained.

The results are categorized into the above-mentioned categories to make them more suitable for analysis. The accuracy, F scores and MCC are evaluated together because the literature review suggests that accuracy and F scores can sometimes mislead the evaluation of classification results. So MCC is used to confirm the reliability of the results as mathematically it takes all four parameters of the confusion matrix in a more balanced manner. The second category is of those parameters which should attain minimum value to confirm the effectiveness of classification. Six more evaluation metrics are used for the firm affirmation of results. Table.4.1 and Table.4.2 show the accuracy, F scores and MCC values obtained by the classifiers for CT and X-Ray images, respectively. Figure.4.7 and Figure.4.8 show the results obtained by other parameters which confirm the robustness of the results obtained by the proposed method for both modalities.

Evaluation metric comparison of CT images				
Networks	**Accuracy**	**F1-Score**	**F2-Score**	**MCC**
MobileNet-V2	0.9942	0.9900	0.9890	0.9852
AlexNet	0.9611	0.9391	0.9643	0.9125
DarkNet	0.9805	0.9716	0.9834	0.9606
InceptionV3	0.9761	0.9635	0.9740	0.9461
ShuffleNet	0.9817	0.9702	0.9777	0.9587
ResNet	0.9756	0.9629	0.9707	0.9448

TABLE 4.1: Evaluation metric comparison of CT images

Evaluation metric comparison of X-Ray images				
Networks	**Accuracy**	**F1-Score**	**F2-Score**	**MCC**
MobileNet-V2	0.9854	0.9770	0.9815	0.9657
AlexNet	0.9556	0.9298	0.9601	0.9001
DarkNet	0.9705	0.9587	0.9540	0.9379
InceptionV3	0.9667	0.9483	0.9683	0.9249
ShuffleNet	0.9702	0.9588	0.9542	0.9380
ResNet	0.9611	0.9391	0.9643	0.9125

TABLE 4.2: Evaluation metric comparison of X-Ray images

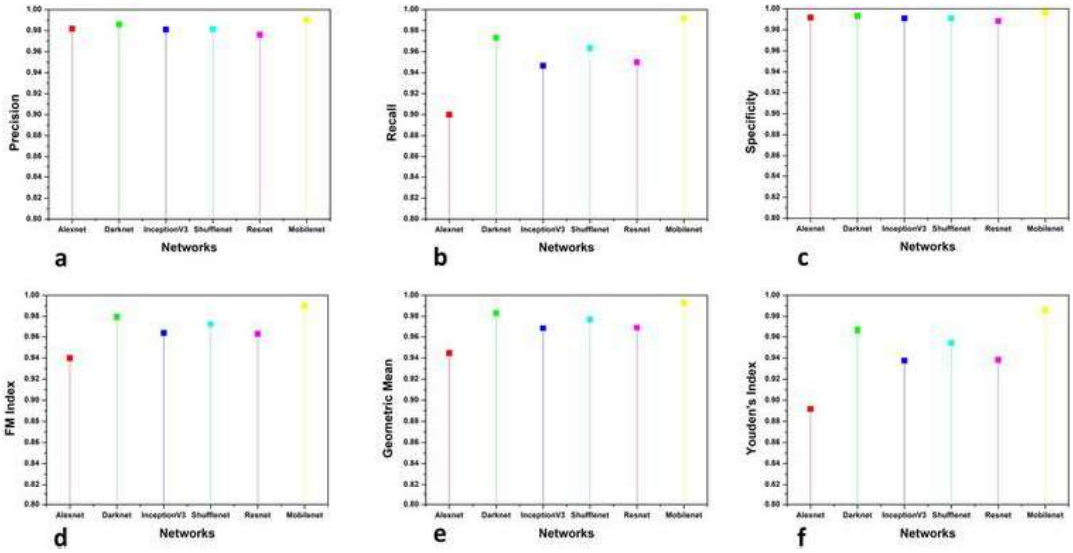

FIGURE 4.7: Evaluation metrics comparison of CT images.

4.5 Discussion

The prime focus of this research is to detect COVID-19 infection through computer-aided analysis. The accurate analysis and detection of COVID-19 infection are quite necessary in the current scenario as the pandemic has placed the whole world in an unprecedented situation by its life-threatening effects. The classifiers segregate the data into COVID (+) and COVID (-) classes

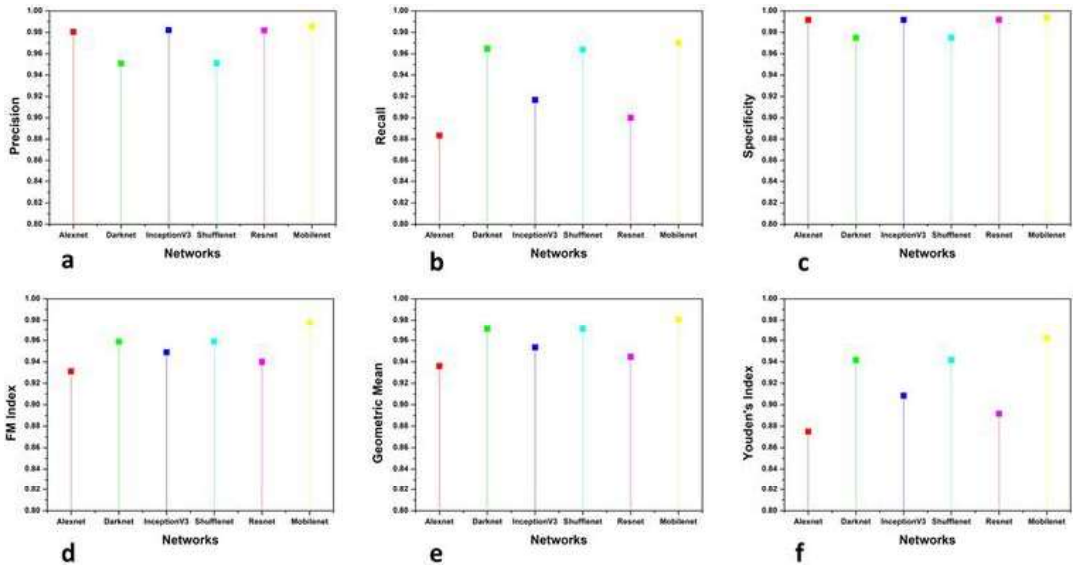

FIGURE 4.8: Evaluation metrics comparison of X-Ray images.

where COVID (+) is the patient infected by COVID-19 and COVID (-) is the patient not having the symptoms of infection. The prediction of infection is made on the basis of the four-parameter values of the confusion matrix, i.e.True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The TP and TN values identify COVID (+) and COVID (-) patients. FP represents the outcomes that are incorrectly recognized as COVID (+) while FN depicts the outcome which is incorrectly classified as COVID (-).

4.5.1 Quantitative Analysis

Accuracy is a measure of accurate identification of the class of a pixel. The evaluation of correctly identified pixels is made using accuracy. Accuracy is computed using the following Equation.4.4:

$$Accuracy(ACC) = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.4}$$

The accuracy for both modalities shows an improvement of more than 1.5%. The distribution of data is not taken into account by the accuracy so the F-score [98] is a better parameter for judging the classification efficiency. It is given by the following Equation.4.5 and 4.6:

$$F1 - Score = \frac{2 * P * R}{P + R} \tag{4.5}$$

and,

$$F0.5 - Score = \frac{1.25 * P * R}{0.25 * P + R} \tag{4.6}$$

where P and R are the Precision and Recall values. The F1-score is defined as the harmonic mean of the precision and recall values with a distance error tolerance. It is a measure of a model's accuracy on a dataset. F0.5 score has the effect of raising the importance of precision and lowering the importance of recall. The F1 score maintains the balance of precision and recall values and has shown an improvement of 1.8% for CT images and 1.9% for X-ray images. The f 0.5 score, which gives more importance to precision as desired for medical analysis is also improved by 1.25 % and 1.4% for CT and X-ray images respectively. Accuracy and F1 measures can sometimes mislead the result interpretation because they fail to consider the ratio of positive and negative elements. The MCC is used to firmly quantitative the analysis of classification as the mathematical properties of MCC handles both dataset imbalance and their invariants effectively. It gives correct predictions for both majorities of the negative cases, and positive cases, independently of their ratios in the overall dataset. The following Equation.4.7 describes MCC mathematically [99]:

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{4.7}$$

MCC has shown a remarkable improvement of respectively 2.6% and 3.8% for CT and X-ray images.This signifies the better classification ability of the proposed model as the MCC has achieved the highest score along with the accuracy and F-scores.

For more analysis and a better understanding of results, we used six additional evaluation metrics for examining the obtained results. Precision, Recall, Specificity, FM index, Geometric mean and Youden's index values are used to compare the results of classification by various methods. Precision which is also known as Positive prediction value (PPV) is the ratio of positive samples which were correctly classified to the total number of samples predicted as positive. It is given by the following Equation.4.8:

$$Precision = \frac{TP}{TP + FP} \tag{4.8}$$

Recall or sensitivity is given as the ratio of positive correctly classified samples to the total number of positive samples. It is also known as true positive rate (TPR) and hit rate. The following Equation.4.9 represents recall:

$$Recall = \frac{TP}{TP + FN} \tag{4.9}$$

The recall is the ability of a model to find all relevant cases within a dataset while the ability of a classification model to identify only the relevant data points is termed precision. A good margin improves the values of precision and recall in both image modalities. Specificity, which is also termed inverse recall or true negative rate (TNR) is expressed as the fraction of correctly classified negative samples to the total count of negative samples. Specificity is expressed mathematically by Equation.4.10:

$$Specificity = \frac{TN}{FP + TN} \tag{4.10}$$

Fowlkes-Mallows Index (FM Index) [38] gives a more accurate representation of unrelated data. The higher value of FM Index gives more significant similarity between classified and ground truth data. It is given by the following Equation.4.11:

$$FM = \sqrt{\left(\frac{TP}{TP + FP}\right) * \left(\frac{TP}{TP + FN}\right)} \tag{4.11}$$

The prime goal of the classification process is to improve sensitivity without sacrificing specificity. This is quite a tough task for imbalanced datasets. Thus Geometric Mean (GM) [100] represents the aggregation of both metrics using the following Equation.4.12:

$$GM = \sqrt{\left(\frac{TP}{TP + FN}\right) * \left(\frac{TN}{TN + FP}\right)} \tag{4.12}$$

The values obtained by the Specificity, FM, and GM have shown considerable improvement shown in the figure. The higher values of sensitivity and specificity indicate that the true values of positive and negative classified samples are greater in the results obtained by the classifier. Thus the classifier is said to have predicted the COVID (+) and COVID (-) samples more accurately than others if these two scores have larger values than other methods in comparison. Fowlkes-Mallows

Index (FM), which gives a more accurate description of unrelated data shows a good degree of improvement in the classification of COVID-19. The last evaluation metric Youden's index or Bookmaker Informedness is an important metric for medical image analysis as it evaluates the discriminative power of the diagnostic test. It evaluates the probability of informed decisions and is more suitable for datasets with imbalanced classes. Youden's index[9] shows an increment of more than 2% for both CT and X-ray images. Mathematically, it is represented by the following Equation.4.13:

$$J = \left(\frac{TP}{TP + FN}\right) + \left(\frac{TN}{TN + FP}\right) - 1 \tag{4.13}$$

where J represents Youden's index. The proposed classifier performs best, as suggested by each evaluation metric. Moreover, it works equally well for both types of medical imaging modality, i.e. CT and X-ray. It takes advantage of transfer learning for training and is more immune to false results predictions.

The proposed method classified the COVID-19-infected images with good accuracy which is supported by several other evaluation metrics [30, 101]. This method is time efficient and accurate as compared with the several methods. Moreover, the computation cost is less as compared to other methods because of depth-wise separable convolution. Lesser number of convolution multiplications tends to save the spatial resolution of the patterns in the images.

4.6 Conclusion

The COVID-19 pandemic has caused a large number of deaths across the globe. An early and accurate diagnosis is necessary for the treatment of this infection. The current need for the diagnosis of COVID-19 is to save the precious time of doctors and to reduce the burden on the present medical system. In conclusion, a transfer learning-based approach is presented in this work for the accurate identification of COVID-19 in CT and X-ray images. This work proposes a computer-aided analysis of accurate detection of COVID-19-infected images. The features finely extracted by the MobileNet V2 model paved the way for this research. The obtained results were evaluated using well-known evaluation metrics, and to the best of our knowledge, no study till now has not compared their method using such a large number of metrics. The advantage of this comparison is that we can be quite sure about the reliability of our results. MCC was used because sometimes accuracy and F

score misleads the prediction of classification as mathematically it is a perfect balance between all four parameters of the confusion matrix. Youden's Index, which is quite a trustworthy statistical measure for medical image analysis, shows outstanding improvement, which proves the medical effectiveness of the results obtained by the proposed model. The CT scan and X-ray facilities are common in most medical organizations. Thus the proposed method can help for the initial identification of The infections caused by viral pneumonia infections and COVID-19 are almost comparable diseases. We believe that with the increase in the dataset, the effectiveness of the classification process can be further improved.