*Article*

# Automatic Classification Approach for Detecting COVID-19 using Deep Convolutional Neural Networks

**Sajib Sarker [1,†], Ling Tan [1,†], Ma Wen Jie [1], Rong Shan Shan [1], Md. Arfan Ali [2,†], Muhammad Bilal [2,\*], Zhongfeng Qiu [2], Sanjit Kumar Mondal [3] and Pravash Tiwari [4]**

[1]  School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China; sajibsarker@nuist.edu.cn

[2]  School of Marine Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China, md.arfanali@nuist.edu.cn; zhongfeng.qiu@nuist.edu.cn

[3]  Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disaster, Institute for Disaster Risk Management, School of Geographical Science, Nanjing University of Information Science and Technology, 219 Ningliulu, Nanjing 210044, China; 20185111002@nuist.edu.cn

[4]  House no. 71, Galkopakha, Kathmandu, Nepal; pravashtiwarigeop@hotmail.com

\*  Correspondence: muhammad.bilal@connect.polyu.hk

†  These authors contributed equally to this work.

**Abstract:** The COVID-19 pandemic situation has created even more difficulties in the quick identification and screening of the COVID-19 patients for the medical specialists. Therefore, a significant study is necessary for detecting COVID-19 cases using an automated diagnosis method, which can aid in controlling the spreading of the virus. In this paper, the study suggests a Deep Convolutional Neural Network-based multi-classification approach (COV-MCNet) using eight different pre-trained architectures such as VGG16, VGG19, ResNet50V2, DenseNet201, InceptionV3, MobileNet, InceptionResNetV2, Xception which are trained and tested on the X-ray images of COVID-19, Normal, Viral Pneumonia, and Bacterial Pneumonia. The results from 3-class (Normal vs. COVID-19 vs. Viral Pneumonia) showed that only the ResNet50V2 model provides the highest classification performance (accuracy: 95.83%, precision: 96.12%, recall: 96.11%, F1-score: 96.11%, specificity: 97.84%) compared to rest of the models. The results from 4-class (Normal vs. COVID-19 vs. Viral Pneumonia vs. Bacterial Pneumonia) demonstrated that the pre-trained model DenseNet201 provides the highest classification performance (accuracy: 92.54%, precision: 93.05%, recall: 92.81%, F1-score: 92.83%, specificity: 97.47%). Notably, the ResNet50V2 (3-class) and DenseNet201 (4-class) models in the proposed COV-MCNet framework showed higher accuracy compared to the rest six models. This indicates that the designed system can produce promising results to detect the COVID-19 cases on the availability of more data. The proposed multi-classification network (COV-MCNet) significantly speeds up the existing radiology-based method, which will be helpful to the medical community and clinical specialists for early diagnosis of the COVID-19 cases during this pandemic.

**Keywords:** COVID-19; chest X-ray images; deep convolutional neural network; COV-MCNet; deep learning

## 1. Introduction

The COVID-19 pandemic has triggered large-scale emergencies across the globe that constitutes a public health crisis as well as humanitarian and development crises worldwide. The 2019 ''Novel Corona Virus (COVID-19)'' as the disease name coined by the World Health Organization (WHO) is a disease associated with infectious respiratory illness. COVID-19 was first identified in Wuhan city,

China, at the end of December 2019. Since then COVID-19 cases have spread gradually across the globe, affecting almost 216 countries, infecting 30,905,162 and killing 958,703 people (until 21 September 2020; Worldometer, 2020). The COVID-19 was declared a pandemic by the WHO on March 11, 2020 [1], taking into account the exponential increase in the number of daily cases and severe impact on people's lives, health systems, and economies all over the world. Coronavirus (COVID-19) has some other species: Severe Acute Respiratory Syndrome (SARS) and Respiratory Syndrome of the Middle East (MERS).[2]. Since COVID-19 is a newly discovered from the large family of viruses, therefore, the global community of medical experts is actively working on strategies to control and find a suitable remedy for this virus [3]. The COVID-19 transmission is higher than other viruses and has common symptoms targeting the respiratory systems of the infected, like the other species of coronaviruses. Figure 1 shows the rising COVID-19 cases worldwide, as of September 21, 2020.
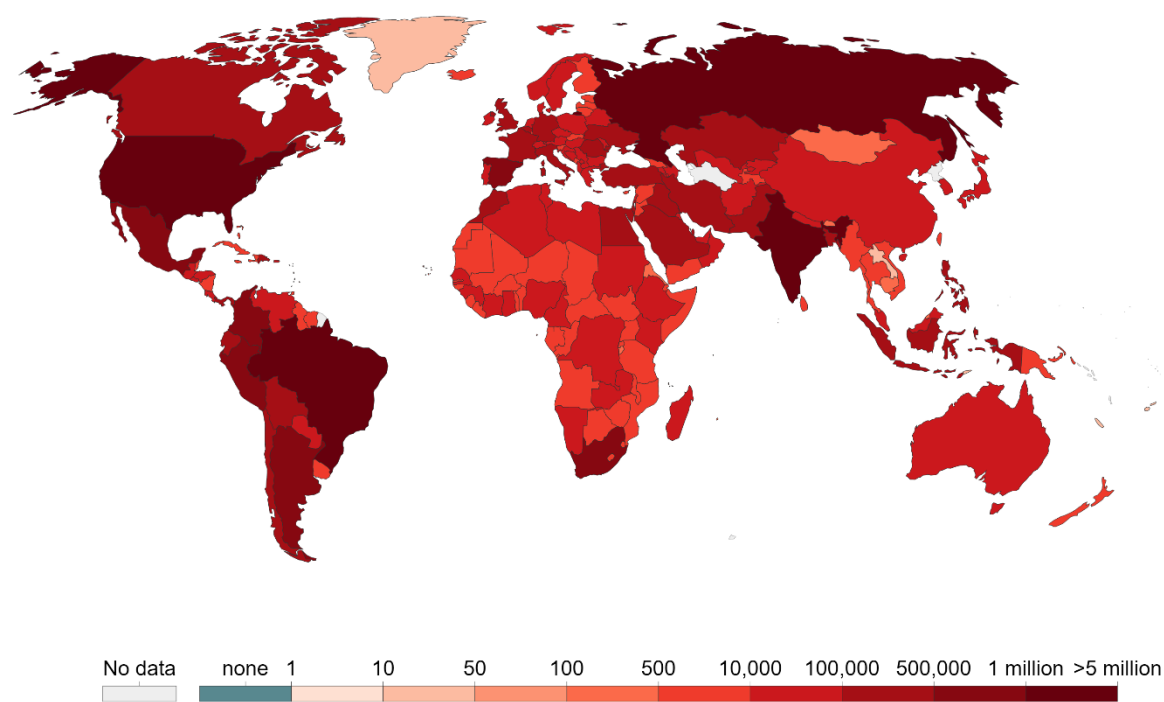


**Figure 1.** Cumulative COVID-19 cases worldwide, from 31 December 2019 to 21 September 2020 (Source: European CDC).

The primary symptoms of COVID-19 patients are similar to that of seasonal flu (e.g. fever, dry cough, and shortness of breath). However, the lethality of COVID-19 is far more serious and infectious and then the normal flu; concerning severe cases, the infection may lead to pneumonia, kidney disorder, multi-organ failure, difficulty in breathing, and death [4]. However, it is important to mention that this newly identified virus has no history of human contact, which implies the high vulnerability of getting infected from it. Also, one of the major challenges to control and stop the spreading of the virus is the provision of an effective testing system. Therefore, experts are now even more urged to find a cost-effective testing system to cope up with the virus.

Identifying and isolating patients infected with COVID-19 is an important step in managing this global pandemic. To control the spread, health care professionals need an appropriate screening system to determine who has come into contact with the infected people. This method is called as touch tracing. However, the real-time Reverse Transcription Polymerase Chain Reaction (RT-PCR) is one of the recent widely used procedures for detecting COVID-19 [5]. Depending on the RT-PCR test, it may take a few hours to 2 days to find out the results of the test. In this period of the crowning time of the COVID-19 outbreak, many countries are facing a shortage of ventilators and RT-PCR test kits. Among several alternatives to the RT-PCR screening method for diagnosing COVID-19, healthcare experts and researchers are exploring the possibility to use chest radiography imaging tests. Many

researchers reported that chest radiological imaging such as computed tomography (CT) and X-ray might be expedient in the primary diagnosis of this disease [6]. While patients with COVID-19, an abnormality can show on either a chest X-ray or CT scan, but the absence of an abnormality on either a chest X-ray or CT scan does not necessarily exclude COVID-19. Therefore, most experts and medical societies believe that chest radiology imaging tests can be an effective tool in detecting COVID-19.

The diagnosis methods PCR tests and CT scans both are comparatively expensive [7] and sometimes are required for critical patients to perform more selective tests. X-ray imaging is relatively cost-effective and usually used for lung infection detection or segmentation and has proven convenient for COVID-19 detection as well [8]. Several studies have been proposed for the detection of COVID-19 using chest X-ray and CT images [9-14]. Since COVID-19 attacks the epithelial cells that affect our lung area, medical specialists use X-ray images to diagnose pneumonia, lung irritation, boils, and/or other lung diseases. And almost all hospitals have X-ray imaging machines, but sometimes for rural areas, it might be difficult to use X-ray imaging. Without the enthusiastic evaluation kits, it might be conceivable to use X-rays to monitor for COVID-19. Again, a downside is that the X-ray examination needs a professional radiologist and takes a considerable amount of time, which is valuable since there is still a significant upsurge in infected cases globally. Thus, it is essential to develop an automated method of study to save valuable time for medical specialists. Previous studies have been conducted exploring along this line, one such study is the work by Fan et al. [15], based on a novel COVID-19 Lung Infection Segmentation Deep Network (Inf-Net) for identifying infected regions using chest CT images. Along with that, Waheed et al. [16] developed an Auxiliary Classifier Generative Adversarial Network (ACGAN) based model called CovidGAN to generate synthetic chest X-ray (CXR) images. Hasan et al. [17] study extract features from CT images using deep learning and a Q-deformed entropy algorithm to classify COVID-19, pneumonia, and normal cases after that features are classified using a long short-term memory (LSTM) neural network classifier. They achieved 99.68% accuracy. Wang et al. [18] presented Artificial Intelligence's based deep learning methods to extract COVID-19's specific graphical features. Panwar et al. [19] proposed nCOVnet a DCNN based method for detecting the COVID-19 by examining the X-rays of patients. They achieved the training accuracy is up to 97% using a different number of images. Makris et al. [20] proposed a pre-trained based model with transfer learning for classifying COVID-19. Their best two pre-trained models achieved an accuracy of 95%. Narin et al. [21] implemented the same experiment with three different CNN models (ResNet50, InceptionV3, and InceptionResNetV2) among these models ResNet50 pre-trained model obtained 98% for 2-class classification. So far the results connecting deep learning to COVID-19 diagnosis have been very promising.

In this study, we implement a deep convolutional neural network (CNN)-based automated classification system for identifying COVID-19 infected cases from chest radiology images. The proposed network is called COV-MCNet combining with different pre-trained models that classify three types of pneumonia; COVID-19, viral pneumonia, and bacterial pneumonia. The proposed system was implemented for 3-class (Normal, COVID-19, and Viral Pneumonia) and 4-class (Normal, COVID-19, Viral Pneumonia, and Bacterial Pneumonia) classification using eight pre-trained models (VGG16, VGG19, ResNet50V2, InceptionV3, InceptionResNetV2, DenseNet201, MobileNet, and Xception). The proposed network produced promising results, even though using a small dataset (300 Normal, 240 COVID-19, 300 Viral Pneumonia, and 300 Bacterial Pneumonia).

## 2. Materials and Methods

### 2.1. Dataset

This study has used a total of 1140 images (240 COVID-19, 300 Normal, 300 Viral Pneumonia, and 300 Bacterial Pneumonia) to develop the multi-classification network (COV-MCNet). The COVID-19 X-ray images are sourced from the GitHub repository [22] and the rest three dataset (normal, viral pneumonia, and bacterial pneumonia) were obtained from the Kaggle repository [23]. Therefore, these datasets have been used for feature extraction based on different deep learning architectures. Details of the used dataset as shown in Table 1. Since this study focused primarily on

the detection of COVID-19 infected cases, therefore, the MERS, SARS, and ARDS virus images were not considered. The two datasets are examined separately in the COV-MCNet proposed models. Figure 2 shows several chest X-ray images of normal, COVID-19, viral pneumonia, and bacterial pneumonia patients.
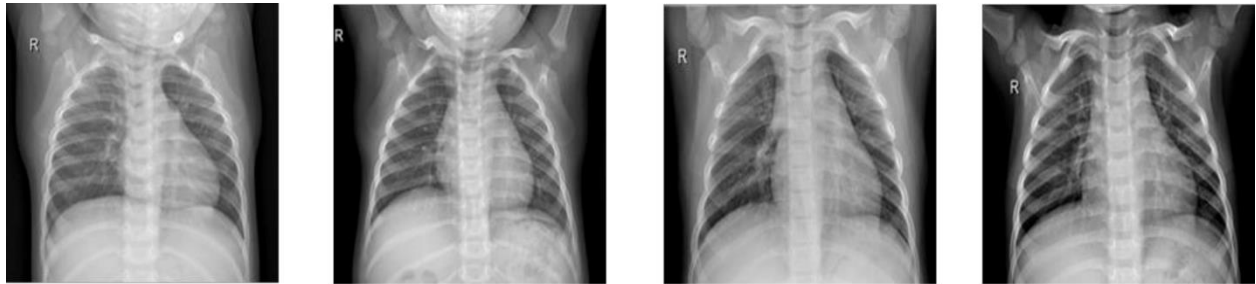
(a) Chest X-ray images of normal patients

(b) Chest X-ray images of COVID-19 patients

(c) Chest X-ray images of viral pneumonia patients

(d) Chest X-ray images of bacterial pneumonia patients

**Figure 2.** Samples of prepared chest X-ray images dataset. a) Normal, b) COVID-19, c) Viral Pneumonia, and d) Bacterial Pneumonia.

**Table 1.** Details of Dataset used in the present study.

| Classification Task | Classes | Number of datasets |
|---|---|---|
| 3-Class Classification | Normal, COVID-19, Viral Pneumonia | 300 Normal, 240 COVID-19, 300 Viral Pneumonia |
| 4-Class Classification | Normal, COVID-19, Viral Pneumonia, Bacterial Pneumonia | 300 Normal, 240 COVID-19, 300 Viral Pneumonia, 300 Bacterial Pneumonia |

*2.2. Proposed COV-MCNet*

Deep learning methods are widely used in a variety of studies such as image classification, segmentation, and skin disease detection of medical statistics [24, 25]. The study proposed a state-of-the-art deep learning image classifier, namely COV-MCNet (Multi-classification network) based on a deep convolutional neural network (CNN). The COV-MCNet uses eight different pre-trained models which are assembled into 3- and 4- classes to classify COVID-19, normal, viral pneumonia, and bacterial pneumonia cases. The entire methodology is divided into three steps: input and pre-processing steps, pre-trained models, and finally training and classification process. ImageNet is an image database with over 14 million images belonging to over 20 thousand categories created for image recognition competitions [26]. The VGG16 and VGG19 [27] model is an improved version of the convolutional neural network (CNN). These models have small convolution filters (3x3) to get a deeper and more complex network. These two models differ in the depth of convolution, pooling, and fully connected layers. The ResNet50V2 [28] is the upgrade version of ResNet50. The ResNet50V2 model has Deep Residual Networks, which is eight times deeper compared to the VGG nets. This architecture is based on skip connection, which allows us to take activation from one layer and feed it to the future layer. InceptionV3 [29] aims to utilize the additional computation as competently as likely by appropriately factorized convolutions and aggressive regularization. The model 48 layers deep along with pooling and fully connected layers. Inception-ResNet-v2 [30] is the mutual architecture of the Inception with residual connections. This architecture is 164 layers deep. As a result, the network has erudite rich feature demonstrations for an extensive range of images. DenseNet201 (Densely Connected Convolutional Networks) [31] has 201 layers on the ImageNet dataset and it has some compelling advantages: they improve the vanishing-gradient difficulty, fortify feature propagation, boost feature reuse, and significantly reduce the number of parameters. MobileNet [32] is an effective model for mobile and entrenched vision applications. This model uses depthwise separable convolutions based on a rationalized architecture to build light weight deep neural networks. Xception [33] a 71 layers deep convolutional neural network architecture enthused by Inception, where Inception modules have been substituted with depthwise distinguishable convolutions. The network trained on more than a million images from the ImageNet database. A schematic representation of the proposed network is shown in Figure 3.
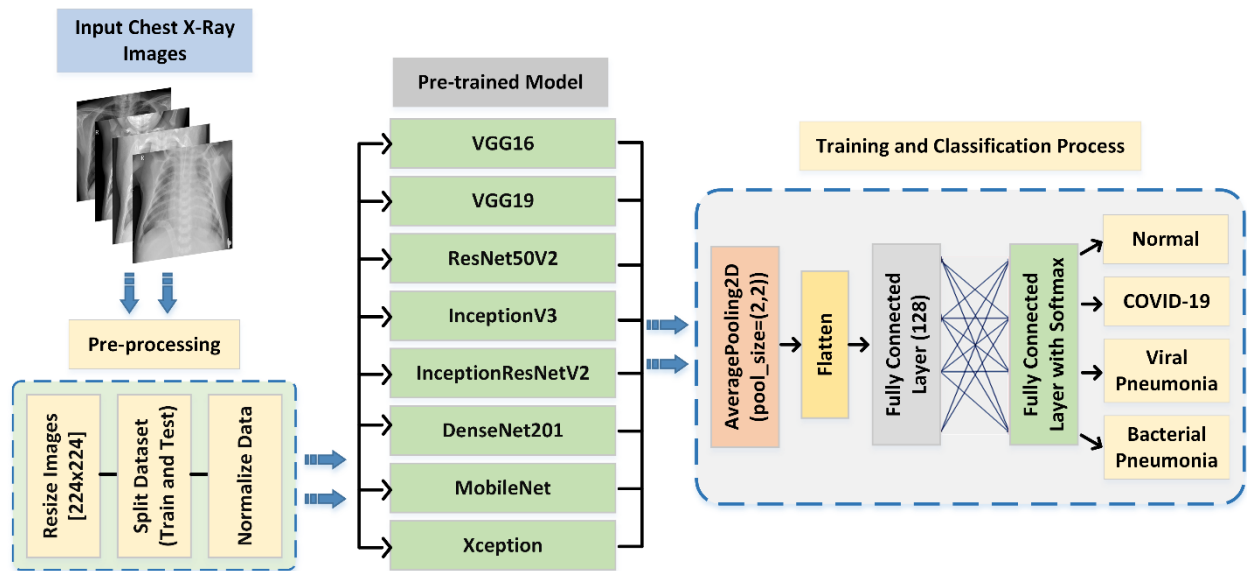
**Figure 3.** Graphical workflow of proposed COV-MCNet models for the detection of Normal, COVID-19, Viral Pneumonia and Bacterial Pneumonia patients.

### 2.2.1. Input and Pre-processing steps

Since the properties of the image (width and height) vary for chest X-ray images of normal, COVID-19, viral pneumonia, and bacterial pneumonia, therefore, the study has used a fixed size of 224 x 224 pixels. Following that, 80% of the data are used as the training dataset and 20% of them are used to evaluate the trained model. Finally, to obtain the decimal values (0 to 1), we normalized the data by dividing 255.

### 2.2.2. Pre-trained Models

Pre-trained models are trained on a large benchmark dataset as a starting point to solve different problems. In this study, eight different pre-trained models (e.g. VGG16, VGG19, ResNet50V2, InceptionV3, InceptionResNetV2, DenseNet201, MobileNet, and Xception) have been used for multi-classification (3-class and 4-class). All the models have different convolution and pooling layers which extract the features from images and classifier categorize the images from extracted features.

### 2.2.3. Training and classification process

In the final step, we fine-tuned the pre-trained models with deep learning image classifiers for detecting COVID-19, normal, viral pneumonia, and bacterial pneumonia cases. In the training and classification process, AveragePooling2D have used for all the models to calculate the average for each patch of the feature map with pool size (2, 2). Afterward, we flattened the activations to create a vectorized feature map and connected two fully connected layers; one layer contained 128 nodes, and the other consisted of 3 and 4 for 3-class classification and 4-class classification, respectively. Subsequently, the activations from the second fully connected layer were fed into a softmax layer, which provided the probability for each of normal, COVID-19, viral pneumonia, and bacterial pneumonia.

### 2.3. Experimental Setup

Python programming language was used for the experiments to training the proposed COV-MCNet framework and Jupyter Notebook as an editor for executing the codes. The background running environment is built-up using deep learning framework TensorFlow (1.14) and Keras package [34]. All experiments were carried out on CPU Intel Core i7 9700K - (32 GB/2 TB HDD/128 GB SSD/Windows 10 Home/4 GB Graphics) and equipped with GPU NVIDIA GeForce RTX 2080Ti. The COV-MCNet framework was trained with random initialization weights using the SGD

(Stochastic Gradient Descent) optimizer. The batch size and learning rate are experimentally set to 10, 0.0001, and the number of epochs is set to 20 to avoid overfitting for all experiments.

*2.4. Performance Metrics*

To test the classification performance of pre-trained models in the COV-MCNet, the following metrics have been implemented in this study to show the classified or misclassified cases. The performance metrics are calculated based on True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values.

2.4.1. Accuracy

It measures the ratio of correctly classified cases with respect to the whole dataset. If the accuracy is higher, that means the models perform better. The accuracy is a portion of the predicted or classified value to its actual value. It represented as follows:

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

2.4.2. Precision

It measures the percentage of correctly classified as positive out of all positive cases. It is defined as follows:

$$\text{Pre} = \frac{TP}{TP + FP} \tag{2}$$

2.4.3. Recall

The recall is computed as the ratio of positives that were correctly predicted as true positives divided by the number of actual positives. It is calculated as follows:

$$\text{Rec} = \frac{TP}{TP + FN} \tag{3}$$

2.4.4. F1-score

F1 Score is calculated based on the scores of precision and recall. It provides the classification capability of the model. F1 score measures the test's accuracy. If the F1 score presents the best value, that means perfect precision and recall. It is calculated as follows:

$$F1 - S = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right) \tag{4}$$

2.4.5. Specificity

It is also called True Positive Rate (TPR) which measures the ratio of actual negatives that are correctly labeled. It is represented as follows:

$$\text{Spe} = \frac{TN}{TN + FP} \tag{5}$$

TP is the proportion of positive cases that are correctly classified as positive; FP is the proportion of negative cases that are misclassified as positive; TN is the proportion of negative that is correctly classified as negative and FN is the proportion of positive that is misclassified as negative by the proposed model.

**3. Results and discussion**

3.1.3-Class Classification training, and validation accuracy and loss

Figure 4 shows the training and validation accuracy with their loss values for the 3-class classification based on the eight pre-trained models (ResNet50V2, VGG16, DenseNet201, MobileNet,

VGG19, Xception, InceptionResNetV2, and InceptionV3). The running time intervals for all pre-trained models were setup into the $20^{th}$ epoch to avoid overfitting of the models. The ResNet50V2 model (Figure 4a) showed the highest validation accuracy (95.83%) compared to the VGG16 (94.64%) (Figure 4b), DenseNet201 (94.05%) (Figure 4c), MobileNet (93.45%) (Figure 4d), VGG19 (93.45%) (Figure 4e), Xception (92.26%) (Figure 4f), InceptionResNetV2 (91.67%) (Figure 4g), and InceptionV3 (88.69%) (Figure 4h).As a result, the ResNet50V2 model provides an advantage over the other pre-trained models used in COV-MCNet both training and validation stages. The evaluation outputs of the best performance model (ResNet50V2) for 3-class classification are shown in Figure S1.  Moreover, it can be seen from Figure 4 that a greater variation in loss values at the beginning of the training for all the eight pre-trained models, which may be because of using the less number of COVID-19 datasets as compared to the other three datasets (normal, viral pneumonia). Several previous studies were conducted to detect COVID-19 positive (infected) and negative (normal) and pneumonia cases based on chest X-ray images. For example, Ozturk et al. [35] proposed DarkNet architecture, a convolutional neural network-based model to detect and classify COVID-19 cases from X-ray images. DarkNet achieved 87.02% accuracy for 3-class classification. Asif and Wenhui [36] proposed a transfer learning-based deep CNN model Inception V3 architecture to classify COVID-19 pneumonia and the study reported 96% accuracy. Ioannis et al. [37] proposed a deep transfer learning-based classification task. They achieved 93.48% for the three-class performance. In comparison to these studies, the ResNet50V2 model in our proposed network (COV-MCNet) showed high accuracy than Ozturk et al. [35] and Ioannis et al. [37] and comparable accuracy with Asif and Wenhui [36].
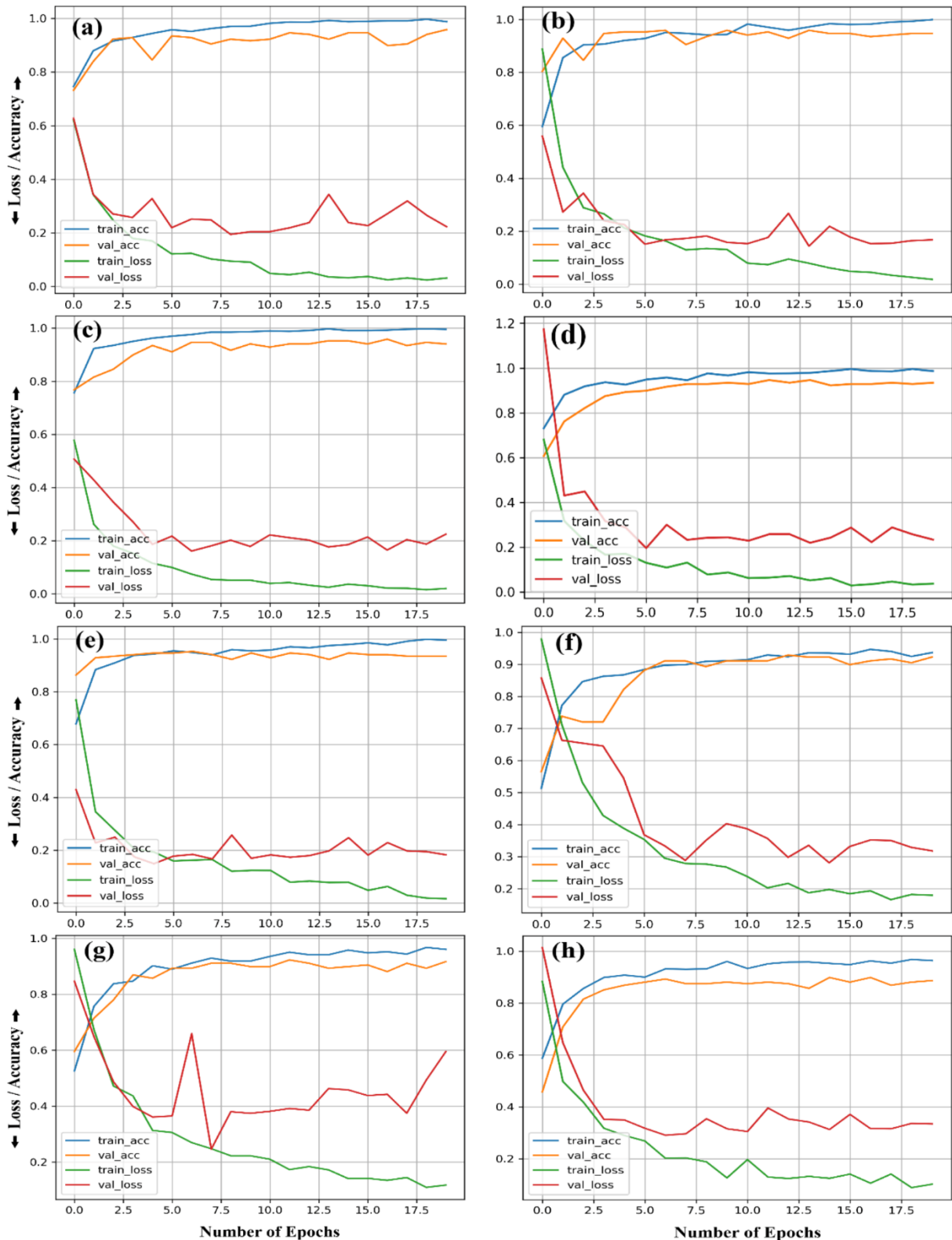
**Figure 4.** The training accuracy (train_acc), loss (train_loss), and validation accuracy (val_acc), loss (val_loss) curves of all pre-trained models in the COV-MCNet for 3 classes: a) ResNet50V2, b) VGG16, c) DenseNet201, d) MobileNet, e) VGG19, f) Xception, g) InceptionResNetV2, and h) InceptionV3.

3.2.3-class classification Confusion Matrix and ROC curve

The confusion matrix (CM) and the receiver operating characteristic curve (ROC) for the 3-class classification problem shown in Figures 5 and 6, respectively. Rows of the confusion matrix correspond to an actual class while columns represent to the predicted class and the color intensity

specifies the probability of each element in a row. It can be seen from the confusion matrix of 3-class classification (Figure 5) that the pre-trained model classified COVID-19 cases better than other classes of normal and viral pneumonia. Besides, the roc curve (Figure 6) plots the TPR against FPR which measures the classification performance on the various threshold. It is important to mention from the ROC curve of the ResNet50V2 that the average area under the curve (AUC) value of 1.00 defines the COVID-19 (i.e. Class 1), AUC ~ 0.98 indicates normal (Class 0), AUC ~ 0.97 reveals viral pneumonia (Class 2).
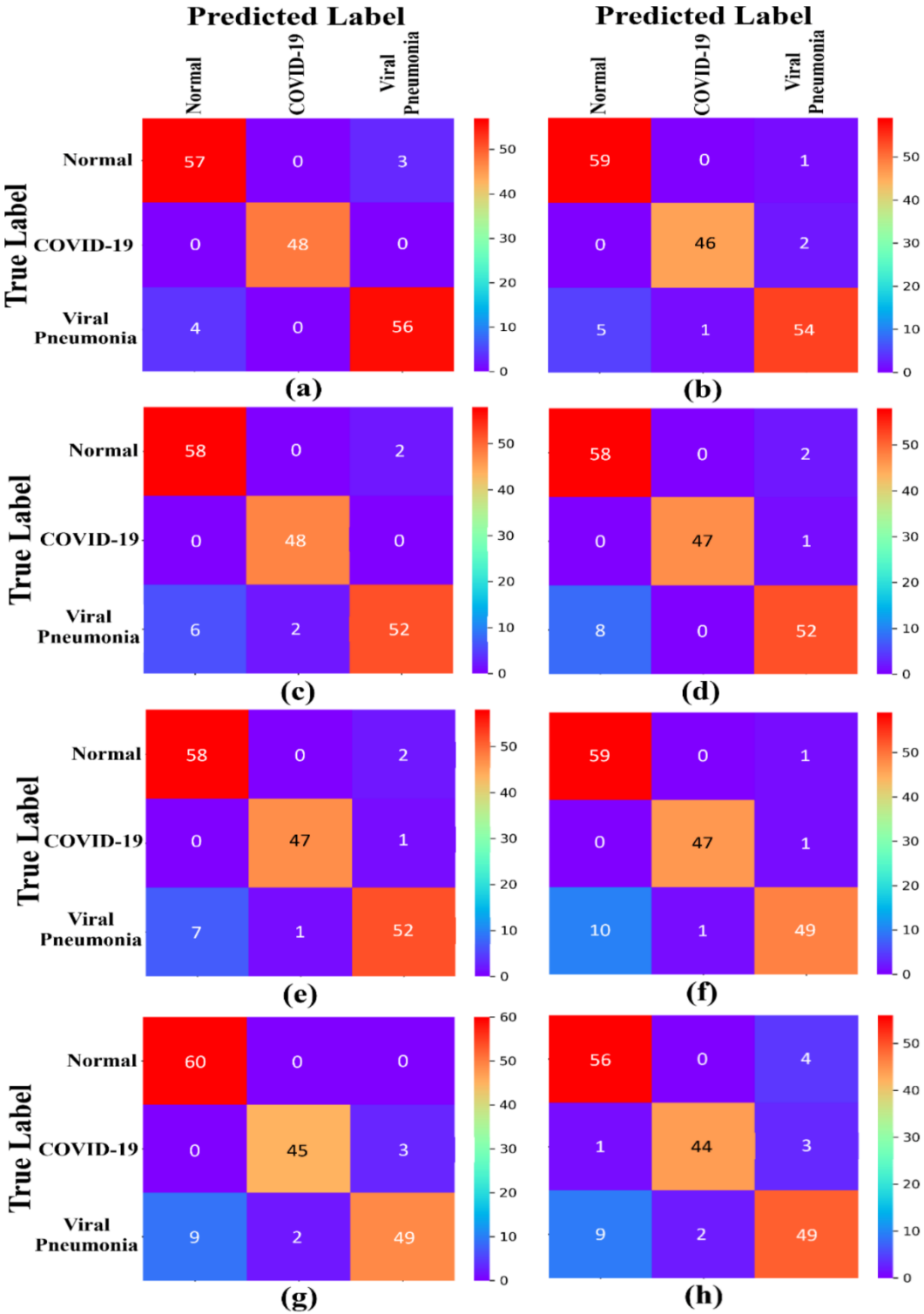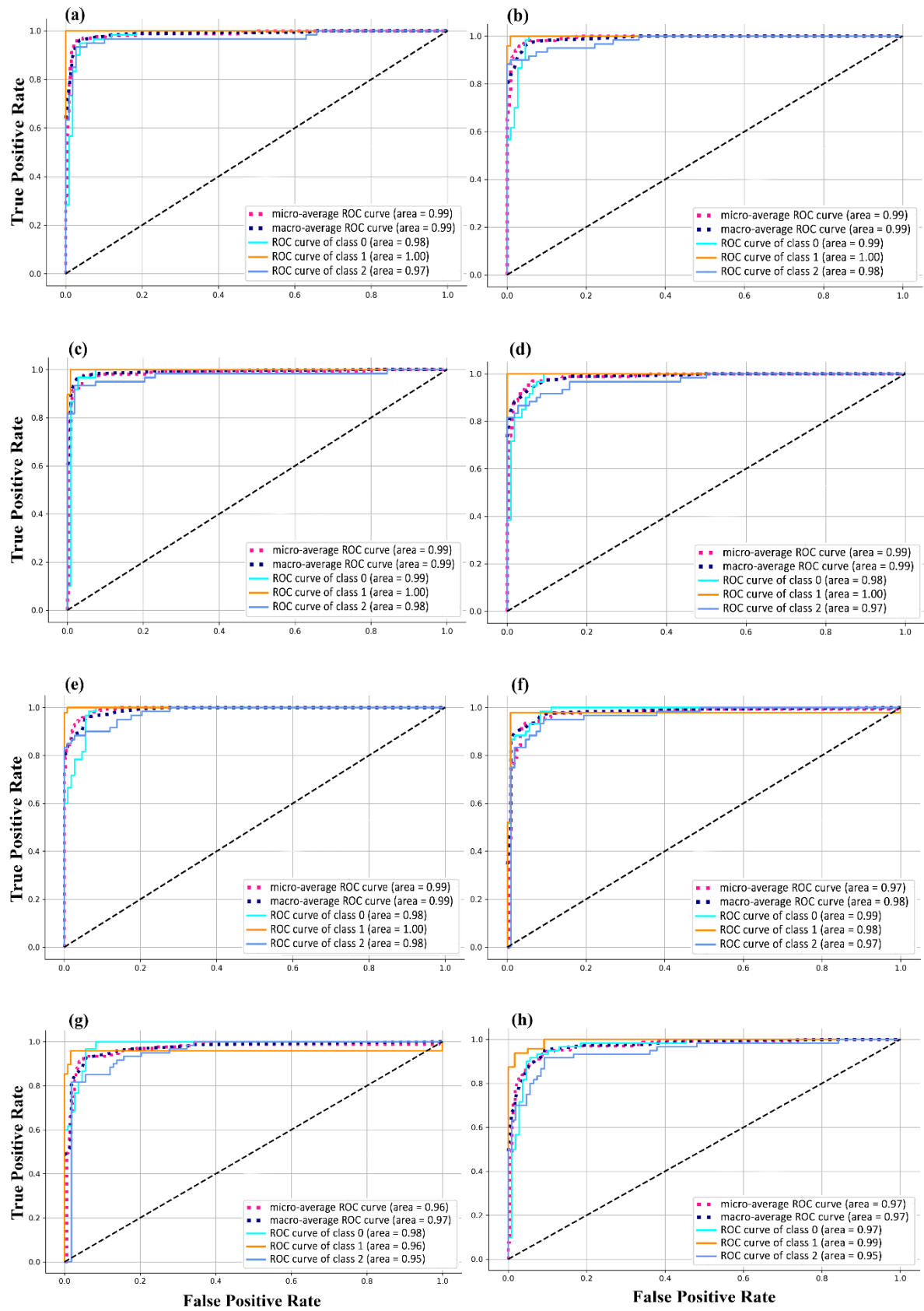
**Figure 5.** The confusion matrix results of 3-class classification obtained using pre-trained models in the COV-MCNet: a) ResNet50V2, b) VGG16, c) DenseNet201, d) MobileNet, e) VGG19, f) Xception, g)

InceptionResNetV2, and h) InceptionV3. Here, the diagonal red, light orange, and dark orange are true positive (TP) and the light and dark blue are the miss classifications of our model.

**Figure 6.** ROC Curve of pre-trained models for 3-class classification in the COV-MCNet: (Class 0: Normal, Class 1: COVID-19, Class 2: Viral Pneumonia). a) ResNet50V2, b) VGG16, c) DenseNet201, d) MobileNet, e) VGG19, f) Xception, g) InceptionResNetV2, and h) InceptionV3.

### 3.3.3-class comparative performance metrics of pre-trained models in the COV-MCNet

Table 2 demonstrates the performance metrics of the eight pre-trained models used in the proposed network for 3-class classification. It can be noticed from the performance metrics of pre-trained models used in COV-MCNet (Table 2) that the best classification performance was recorded by the ResNet50V2 model for each class such as Normal (Precision: 93.44%, Recall: 95%, F1-Score: 94.21%, Specificity: 96.30%), COVID-19 (Precision: 100%, Recall: 100%, F1-Score: 100%, Specificity: 100%), Viral Pneumonia (Precision: 94.92%, Recall: 93.33%, F1-Score: 94.12%, Specificity: 97.22%). As the ResNet50V2 model has deeper residual networks compared to the other models, which also use skip connections to skip training from a few layers and connects directly to the output hence, these results recommended that the ResNet50V2 model is robust and superior than the other tested models.

**Table 2.** The performance metrics of eight pre-trained models for 3-class classification. Where Precision (Pre), Recall (Rec), F1-score (F1-S), Specificity (Spe) values.

| Models | Classes | TP | TN | FP | FN | Pre | Rec | F1-S | Spe |
|---|---|---|---|---|---|---|---|---|---|
| ResNet50V2 | Normal | 57 | 104 | 4 | 3 | 93.44 | 95 | 94.21 | 96.30 |
| | COVID-19 | 48 | 120 | 0 | 0 | 100 | 100 | 100 | 100 |
| | Viral Pneumonia | 56 | 105 | 3 | 4 | 94.92 | 93.33 | 94.12 | 97.22 |
| VGG16 | Normal | 59 | 103 | 5 | 1 | 92.19 | 98.33 | 95.16 | 95.37 |
| | COVID-19 | 46 | 119 | 1 | 2 | 97.87 | 95.83 | 96.84 | 99.17 |
| | Viral Pneumonia | 54 | 105 | 3 | 6 | 94.74 | 90 | 92.31 | 97.22 |
| DenseNet201 | Normal | 58 | 102 | 6 | 2 | 90.63 | 96.67 | 93.55 | 94.44 |
| | COVID-19 | 48 | 118 | 2 | 0 | 96 | 100 | 97.96 | 98.33 |
| | Viral Pneumonia | 52 | 106 | 2 | 8 | 96.30 | 86.67 | 91.23 | 98.15 |
| MobileNet | Normal | 58 | 100 | 8 | 2 | 87.88 | 96.67 | 92.07 | 92.59 |
| | COVID-19 | 47 | 120 | 0 | 1 | 100 | 97.92 | 98.95 | 100 |
| | Viral Pneumonia | 52 | 105 | 3 | 8 | 94.55 | 86.67 | 90.44 | 97.22 |
| VGG19 | Normal | 58 | 101 | 7 | 2 | 89.23 | 96.67 | 92.80 | 93.52 |
| | COVID-19 | 47 | 119 | 1 | 1 | 97.92 | 97.92 | 97.92 | 99.17 |
| | Viral Pneumonia | 52 | 105 | 3 | 8 | 94.55 | 86.67 | 90.44 | 97.22 |
| Xception | Normal | 59 | 98 | 10 | 1 | 85.51 | 98.33 | 91.47 | 90.74 |
| | COVID-19 | 47 | 119 | 1 | 1 | 97.92 | 97.92 | 97.92 | 99.17 |
| | Viral Pneumonia | 49 | 106 | 2 | 11 | 96.08 | 81.67 | 88.29 | 98.15 |
| InceptionResNetV2 | Normal | 60 | 99 | 9 | 0 | 86.96 | 100 | 93.03 | 91.67 |
| | COVID-19 | 45 | 118 | 2 | 3 | 95.74 | 93.75 | 94.73 | 98.33 |
| | Viral Pneumonia | 49 | 105 | 3 | 11 | 94.23 | 81.67 | 87.50 | 97.22 |
| InceptionV3 | Normal | 59 | 98 | 10 | 1 | 85.51 | 98.33 | 91.47 | 90.74 |
| | COVID-19 | 48 | 119 | 1 | 0 | 97.96 | 100 | 98.97 | 99.17 |
| | Viral Pneumonia | 49 | 107 | 1 | 11 | 98 | 81.67 | 89.09 | 99.07 |

The results (Table 3) show that the ResNet50V2 pre-trained model in the proposed network (COV-MCNet) showed an accuracy of 95.83% in detecting COVID-19 for 3-class and the achieved precision, recall, F1-score, and specificity values were 96.12%, 96.11%, 96.11%, and 97.84%, respectively.

**Table 3.** Accuracy (Acc), Precision (Pre), Recall (Rec), F1-Score (F1-S), and Specificity (Spe) results for 3-class classification (Normal vs. COVID-19 vs. Viral Pneumonia) of proposed COV-MCNet model.

| Classification Models | Performance Metrics (%) | | | | |
|---|---|---|---|---|---|
| | Acc | Pre | Rec | F1-S | Spe |
| **ResNet50V2** | **95.83** | **96.12** | **96.11** | **96.11** | **97.84** |
| VGG16 | 94.64 | 94.93 | 94.72 | 94.77 | 97.25 |
| DenseNet201 | 94.05 | 94.31 | 94.45 | 94.25 | 96.97 |
| MobileNet | 93.45 | 94.14 | 93.75 | 93.82 | 96.60 |
| VGG19 | 93.45 | 93.90 | 93.75 | 93.72 | 96.64 |
| Xception | 92.26 | 93.17 | 92.64 | 92.56 | 96.02 |
| InceptionResNetV2 | 91.67 | 92.31 | 91.81 | 91.75 | 95.74 |
| InceptionV3 | 88.69 | 93.82 | 93.33 | 93.18 | 96.33 |

3.4.4-class classification training and validation accuracy and loss

Figure 7 shows the training and validation accuracy with their loss values for the 4-class classification based on the eight pre-trained models (DenseNet201, VGG16, MobileNet, ResNet50V2, InceptionV3, Xception, InceptionResNetV2, and VGG19). The training time for all pre-trained models has been conducted up to the 20th epoch to avoid overfitting. The DenseNet201 model (Figure 7a) showed the highest validation accuracy (92.54%) compared to the VGG16 (90.35%) (Figure 7b), MobileNet (89.04%) (Figure 7c), ResNet50V2 (87.72%) (Figure 7d), InceptionV3 (86.84%) (Figure 7e), Xception (85.09%) (Figure 7f), InceptionResNetV2 (84.65%) (Figure 7g), and VGG19 (83.77%) (Figure 7h). The evaluation outputs of the best performance model (DenseNet201) for 4-class classification are shown in Figure S2. Moreover, loss values exhibited a greater variation at the beginning of the training for all the eight pre-trained models, which may be due to using the less number of COVID-19 datasets as compared to the other three datasets (normal, viral pneumonia, and bacterial pneumonia) (Figure 7). To the best of our knowledge, there are only two studies about 4-class classification were found based on CoroNet Xception and COVID-Net. For example, Khan et al. [38] detected COVID-19 cases based on the CoroNet Xception pre-trained model and reported an accuracy of 89.6%. In contrast, Wang and Wong [39] proposed a deep neural network-based model, namely COVID-Net, they achieved 92.4% accuracy. In comparison to these studies, the DenseNet201 model in our proposed network (COV-MCNet) showed high accuracy than Khan et al. [38] and comparable accuracy with Wang and Wong [39].
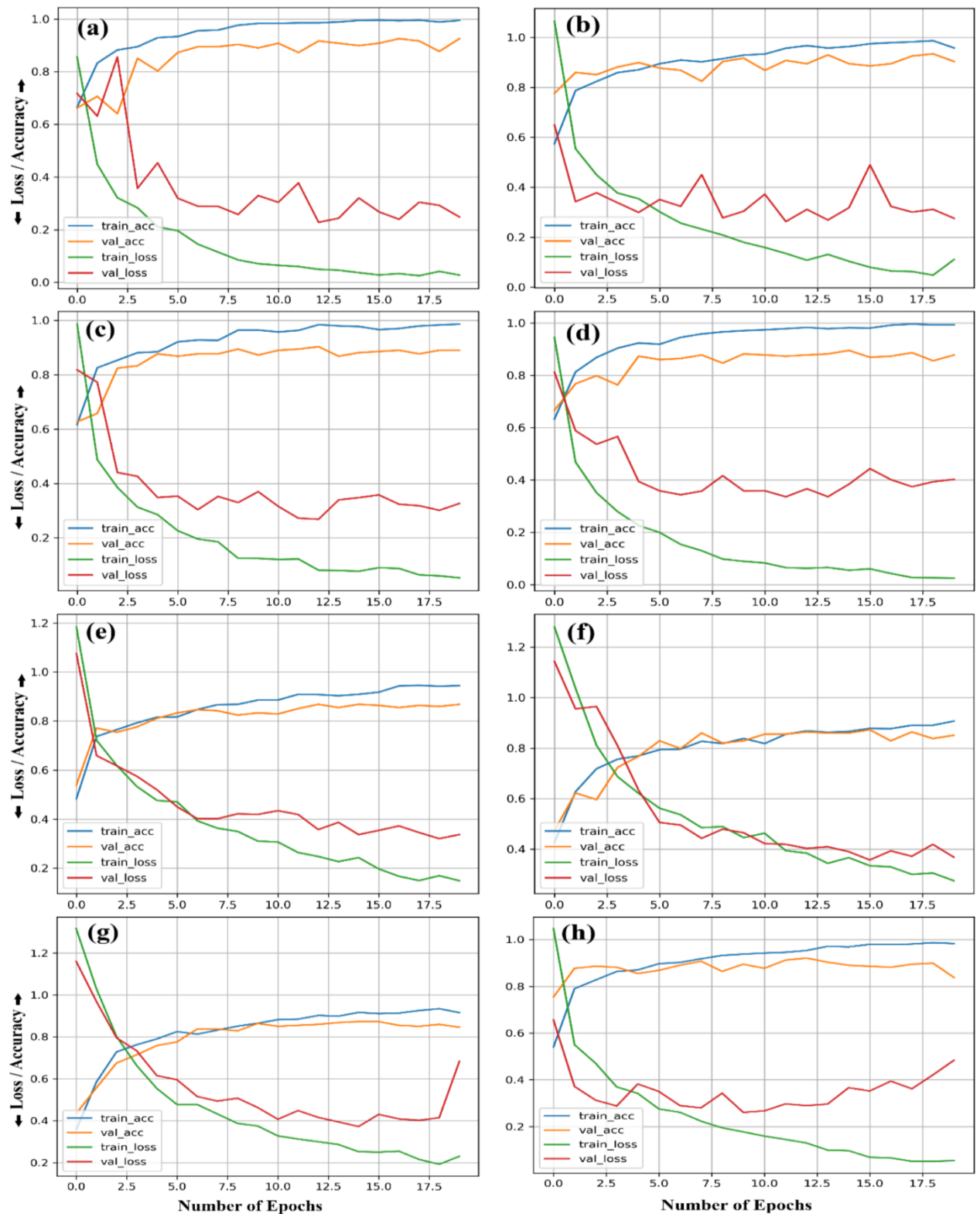
**Figure 7.** The training accuracy (train_acc), loss (train_loss) and validation accuracy (val_acc), loss (val_loss) curves of all pre-trained models in the COV-MCNet for 4 classes: a) DenseNet201, b) VGG16, c) MobileNet, d) ResNet50V2, e) InceptionV3, f) Xception, g) InceptionResNetV2, and h) VGG19.

*3.5. The 4-class classification Confusion Matrix and ROC curve*

Figure 8 and 9 shows the confusion matrix (CM) and the receiver operating characteristic curve (ROC) for the 4-class classification problem, respectively. Rows of the confusion matrix correspond to an actual class while columns represent to the predicted class and the color intensity specifies the

probability of each element in a row. The results (Figure 8) show that the pre-trained models classified COVID-19 cases better than other classes of normal, viral pneumonia, and bacterial pneumonia. Besides, the roc curve (Figure 9) plots the TPR against FPR which measures the classification performance on the various threshold. In Figure 9a, AUC ~ 1.00 represents COVID-19 (i.e. Class 1), AUC ~ 0.99 represents normal (Class 0), AUC ~ 0.97 represents viral pneumonia (Class 2) 0.97, and AUC ~ 0.98 represents bacterial pneumonia (Class 3).
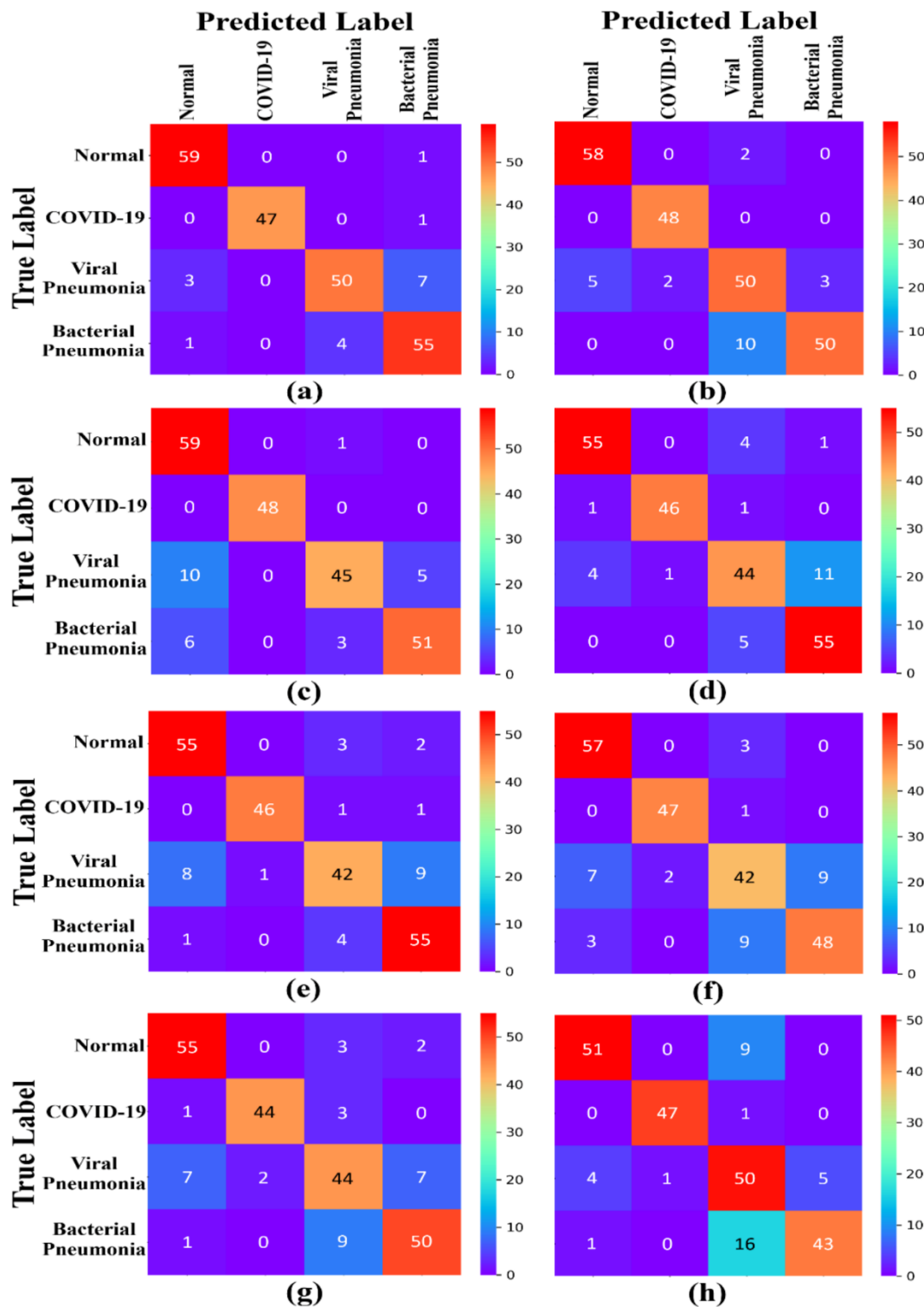


**Figure 8.** The confusion matrix results of 4-class classification obtained using pre-trained models in the COV-MCNet: a) DenseNet201, b) VGG16, c) MobileNet, d) ResNet50V2, e) InceptionV3, f) Xception, g) InceptionResNetV2, and h) VGG19. Here, the diagonal red, light orange, and dark orange are true positive (TP), and the light and dark blue, cyan is the miss classifications of our model.
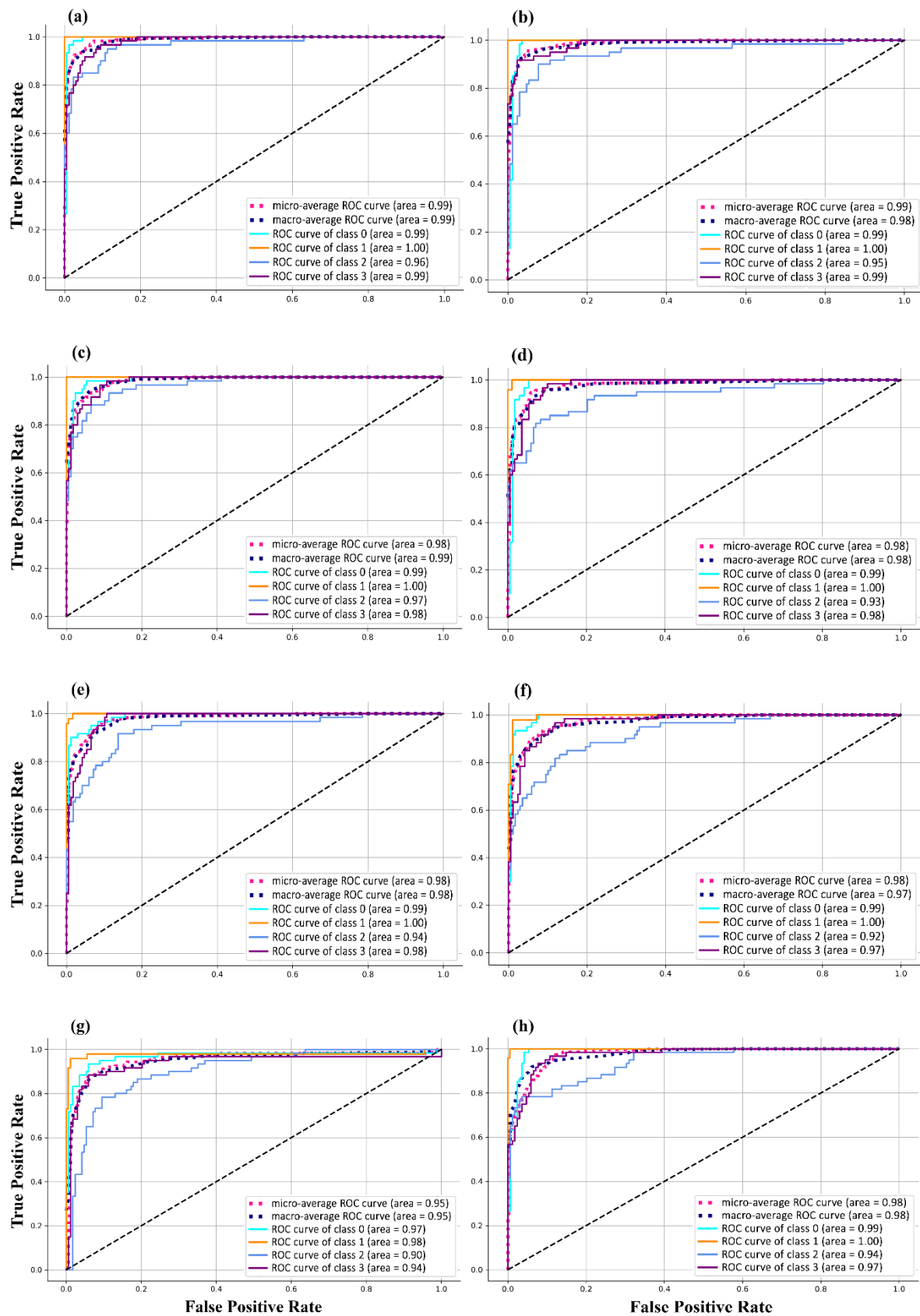
**Figure 9.** ROC Curve of pre-trained models for 4-class classification in the COV-MCNet: (Class 0: Normal, Class 1: COVID-19, Class 2: Viral Pneumonia, Class 3: Bacterial Pneumonia). a) DenseNet201, b) VGG16, c) MobileNet, d) ResNet50V2, e) InceptionV3, f) Xception, g) InceptionResNetV2, and h) VGG19.

3.6.4-class comparative performance metrics of pre-trained models in the COV-MCNet

Table 4 demonstrates the performance metrics of the eight pre-trained models used in the proposed network for 4-class classification. It can be noticed that the DenseNet201 model showed the best classification performance for each class such as Normal (Precision: 93.65%, Recall: 98.33%, F1-Score: 95.93%, Specificity: 97.62%), COVID-19 (Precision: 100%, Recall: 97.92%, F1-Score: 98.95%, Specificity: 100%), Viral Pneumonia (Precision: 92.59%, Recall: 83.33%, F1-Score: 87.72%, Specificity: 97.62%), Bacterial Pneumonia (Precision: 85.94%, Recall: 91.67%, F1-Score: 88.71%, Specificity: 94.64%). As the DenseNet201 model classifier uses features of all complexity levels which inclines to provide further smooth decision boundaries. Also, it has comparatively more layers (i.e., 201 layers) than the rest models as well as improves the vanishing-gradient difficulty, fortify feature propagation, and boost feature reuse, which significantly reduces the number of parameters. Therefore, these results recommended that the DenseNet201 model is robust and superior to the other tested models in terms of precision, recall, F1-score, and specificity.

**Table 4.** The performance metrics of eight pre-trained models for 4-class are presented based on Precision (Pre), Recall (Rec), F1-score (F1-S), Specificity (Spe) values.

| Models | Classes | TP | TN | FP | FN | Pre | Rec | F1-S | Spe |
|---|---|---|---|---|---|---|---|---|---|
| DenseNet201 | Normal | 59 | 164 | 4 | 1 | 93.65 | 98.33 | 95.93 | 97.62 |
| | COVID-19 | 47 | 180 | 0 | 1 | 100 | 97.92 | 98.95 | 100 |
| | Viral Pneumonia | 50 | 164 | 4 | 10 | 92.59 | 83.33 | 87.72 | 97.62 |
| | Bacterial Pneumonia | 55 | 159 | 9 | 5 | 85.94 | 91.67 | 88.71 | 94.64 |
| VGG16 | Normal | 58 | 163 | 5 | 2 | 92.06 | 96.67 | 94.31 | 97.02 |
| | COVID-19 | 48 | 178 | 2 | 0 | 96 | 100 | 97.96 | 98.89 |
| | Viral Pneumonia | 50 | 156 | 12 | 10 | 80.65 | 83.33 | 81.97 | 92.86 |
| | Bacterial Pneumonia | 50 | 165 | 3 | 10 | 94.34 | 83.33 | 88.49 | 98.21 |
| MobileNet | Normal | 59 | 152 | 16 | 1 | 78.67 | 98.33 | 87.41 | 90.48 |
| | COVID-19 | 48 | 180 | 0 | 0 | 100 | 100 | 100 | 100 |
| | Viral Pneumonia | 45 | 164 | 4 | 15 | 91.84 | 75 | 82.57 | 97.62 |
| | Bacterial Pneumonia | 51 | 163 | 5 | 9 | 91.07 | 85 | 87.93 | 97.02 |
| ResNet50V2 | Normal | 55 | 163 | 5 | 5 | 91.67 | 91.67 | 91.67 | 97.02 |
| | COVID-19 | 46 | 179 | 1 | 2 | 97.87 | 95.83 | 96.84 | 99.44 |
| | Viral Pneumonia | 44 | 158 | 10 | 16 | 81.48 | 73.33 | 77.19 | 94.05 |
| | Bacterial Pneumonia | 55 | 156 | 12 | 5 | 82.09 | 91.67 | 86.62 | 92.86 |
| InceptionV3 | Normal | 55 | 159 | 9 | 5 | 85.94 | 91.67 | 88.71 | 94.64 |
| | COVID-19 | 46 | 179 | 1 | 2 | 97.87 | 95.83 | 96.84 | 99.44 |
| | Viral Pneumonia | 42 | 160 | 8 | 18 | 84 | 70 | 76.36 | 95.24 |
| | Bacterial Pneumonia | 55 | 156 | 12 | 5 | 82.09 | 91.67 | 86.62 | 92.86 |
| Xception | Normal | 57 | 158 | 10 | 3 | 85.07 | 95 | 89.76 | 94.05 |
| | COVID-19 | 47 | 178 | 2 | 1 | 95.92 | 97.92 | 96.91 | 98.89 |
| | Viral Pneumonia | 42 | 155 | 13 | 18 | 76.36 | 70 | 73.04 | 92.26 |
| | Bacterial Pneumonia | 48 | 159 | 9 | 12 | 84.21 | 80 | 82.05 | 94.64 |
| InceptionResNetV2 | Normal | 55 | 159 | 9 | 5 | 85.94 | 91.67 | 88.71 | 94.64 |
| | COVID-19 | 44 | 178 | 2 | 4 | 95.92 | 91.67 | 93.75 | 98.89 |
| | Viral Pneumonia | 44 | 153 | 15 | 16 | 74.58 | 73.33 | 73.95 | 91.07 |
| | Bacterial Pneumonia | 50 | 159 | 9 | 10 | 84.75 | 83.33 | 84.03 | 94.64 |
| VGG19 | Normal | 51 | 163 | 5 | 9 | 91.07 | 85 | 87.93 | 97.02 |
| | COVID-19 | 47 | 179 | 1 | 1 | 97.92 | 97.92 | 97.92 | 99.44 |
| | Viral Pneumonia | 50 | 158 | 10 | 10 | 83.33 | 83.33 | 83.33 | 94.05 |
| | Bacterial Pneumonia | 43 | 163 | 5 | 17 | 89.58 | 71.67 | 79.63 | 97.02 |

It is observed from Table 5, the DenseNet201 pre-trained model in the proposed study (COV-MCNet) showed better results in detecting COVID-19 for 4-class with accuracy, precision, recall, F1-score, and specificity are 92.54%, 93.05%, 92.81%, 92.83%, and 97.47%, respectively.

**Table 5.** Accuracy (Acc), Precision (Pre), Recall (Rec), F1-Score (F1-S), and Specificity (Spe) results of all the pre-trained models used in COV-MCNet for 4-class classification (Normal vs. COVID-19 vs. Viral Pneumonia vs. Bacterial Pneumonia).

| Classification Models | Performance Metrics (%) | | | | |
|---|---|---|---|---|---|
| | Acc | Pre | Rec | F1-S | Spe |
| **DenseNet201** | **92.54** | **93.05** | **92.81** | **92.83** | **97.47** |
| VGG16 | 90.35 | 90.76 | 90.83 | 90.68 | 96.75 |
| MobileNet | 89.04 | 90.40 | 89.58 | 89.48 | 96.28 |
| ResNet50V2 | 87.72 | 88.28 | 88.13 | 88.08 | 95.84 |
| InceptionV3 | 86.84 | 87.48 | 87.29 | 87.13 | 95.55 |
| Xception | 85.09 | 85.39 | 85.73 | 85.44 | 94.96 |
| InceptionResNetV2 | 84.65 | 85.30 | 85 | 85.11 | 94.81 |
| VGG19 | 83.77 | 90.48 | 84.48 | 87.20 | 96.88 |

## 4. Summary and Conclusion

In addition to these studies in the literature, the main advantages of our study summarized as follows: A deep convolutional neural network-based framework COV-MCNet was designed with eight different pre-trained models (VGG16, VGG19, ResNet50V2, InceptionV3, InceptionResNetV2, DenseNet201, MobileNet, and Xception) for the detection of COVID-19. The proposed study has used a total of 1140 images (240 COVID-19, 300 Normal, 300 Viral Pneumonia, and 300 Bacterial Pneumonia) to develop the multi-classification network. COV-MCNet classified 3-class (Normal, COVID-19 and Viral Pneumonia) and 4-class (Normal, COVID-19, Viral Pneumonia, and Bacterial Pneumonia) cases without any feature extraction or selection techniques. The proposed multi-classification approach (COV-MCNet) can assist medical specialists for diagnosis X-ray image related diseases. Our proposed model achieved an accuracy of 95.83% and 92.54% for 3- and 4-classes, respectively. Moreover, the number of data we have used in this study is less nonetheless the proposed study obtained superior performance for both classification tasks (3- and 4-class) as compared to the other studies.

The primary limitation of the study is the shortage of COVID-19 image data used for the training of deep learning models. In the future, we intend to improve the proposed model by collecting more radiology data. Besides, we will test the designed multi-classification network (COV-MCNet) with CT images for COVID-19 detection and compare the achieved results with the proposed model which was trained using chest X-ray images.

As the COVID-19 cases are still increasing daily, quick identification of COVID-19 patients is can be one of the effective steps towards preventing the spread of the virus into the non-affected community. Thus and so, this study has proposed a multi-classification approach, namely COV-MCNet based on eight different pre-trained models (VGG16, VGG19, ResNet50V2, InceptionV3, InceptionResNetV2, DenseNet201, MobileNet, and Xception) to detect COVID-19 patients automatically. The suggested models could successfully detect the COVID-19 infected cases based on the 3-class and 4-class classification. The 3-class classification demonstrated the ResNet50V2 best classification model of COVID-19 infected cases with an accuracy of 95.83%, while 4-class classification revealed the DenseNet201 model with an accuracy of 92.54%. The study achieved promising results in comparison to similar studies with small datasets, which can be beneficial for medical specialists to make decisions and gain deeper knowledge about COVID-19 cases. The classification performance of the method can still be improved by increasing the number of training datasets. Also, the study still needs scientific testing but with higher performance, it can pave a way towards a modern and efficient diagnosis of the COVID-19. In the future, we aim to collect more

radiology images of COVID-19 from local hospitals to make more superior results using the suggested model.

## References

1. World Health Organization (WHO). Coronavirus disease (COVID-19) pandemic (2020). https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (Accessed 25 August 2020)
2. G. Gabutti, E. d'Anchera, F. Sandri, M. Savio, and A. Stefanati, "Coronavirus: Update Related to the Current Outbreak of COVID-19," *Infectious Diseases and Therapy*. 2020, doi: 10.1007/s40121-020-00295-5.
3. P. S. Masters, "The Molecular Biology of Coronaviruses," in *Advances in Virus Research*, 2006, pp. 193–292.
4. E. Mahase, "Coronavirus covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate," *BMJ*, vol. 368, p. m641, Feb. 2020.
5. Y. M. Arabi *et al.*, "Middle East Respiratory Syndrome," *N. Engl. J. Med.*, vol. 376, no. 6, pp. 584–594, Feb. 2017, doi: 10.1056/NEJMsr1408795.
6. W. Guan *et al.*, "Clinical characteristics of coronavirus disease 2019 in China," *N. Engl. J. Med.*, 2020, doi: 10.1056/NEJMoa2002032.
7. E. Livingston, A. Desai, and M. Berkwits, "Sourcing Personal Protective Equipment During the COVID-19 Pandemic," *JAMA*, vol. 323, no. 19, p. 1912, May 2020, doi: 10.1001/jama.2020.5317.
8. H. Y. F. Wong *et al.*, "Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19," *Radiology*, vol. 296, no. 2, pp. E72–E78, Aug. 2020, doi: 10.1148/radiol.2020201160.
9. X. Xu *et al.*, "A Deep Learning System to Screen Novel Coronavirus Disease 2019 Pneumonia," *Engineering*, Jun. 2020, doi: 10.1016/j.eng.2020.04.010.
10. C. Butt, J. Gill, D. Chun, and B. A. Babu, "Deep learning system to screen coronavirus disease 2019 pneumonia," *Appl. Intell.*, Apr. 2020, doi: 10.1007/s10489-020-01714-3.
11. L. Li *et al.*, "Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, Aug. 2020, doi: 10.1148/radiol.2020200905.
12. P. K. Sethy, S. K. Behera, P. K. Ratha, and P. Biswas, "Detection of Coronavirus Disease (COVID-19) Based on Deep Features," *Int. J. Math. Eng. Manag. Sci.*, 2020, doi: 10.20944/preprints202003.0300.v1.
13. kishore Medhi, K. Medhi, and I. Hussain, "Automatic Detection of COVID-19 Infection from Chest X-ray using Deep Learning," *medrxiv*, 2020, doi: 10.1101/2020.05.10.20097063.
14. S. Kumar, S. Mishra, and S. K. Singh, "Deep Transfer Learning-based COVID-19 prediction using Chest X-rays," *medRxiv*, 2020, doi: 10.1101/2020.05.12.20099937.
15. D.-P. Fan *et al.*, "Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images," *IEEE Trans. Med. Imaging*, pp. 1–1, 2020, doi: 10.1109/TMI.2020.2996645.
16. A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, "CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection," *IEEE Access*, vol. 8, pp. 91916–91923, 2020, doi: 10.1109/ACCESS.2020.2994762.
17. A. M. Hasan, M. M. AL-Jawad, H. A. Jalab, H. Shaiba, R. W. Ibrahim, and A. R. AL-Shamasneh, "Classification of Covid-19 Coronavirus, Pneumonia and Healthy Lungs in CT Scans Using Q-Deformed Entropy and Deep Learning Features," *Entropy*, vol. 22, no. 5, p. 517, May 2020, doi: 10.3390/e22050517.
18. S. Wang *et al.*, "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *medRxiv*, 2020, doi: 10.1101/2020.02.14.20023028.

19. H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, and V. Singh, "Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet," *Chaos, Solitons & Fractals*, vol. 138, p. 109944, Sep. 2020, doi: 10.1016/j.chaos.2020.109944.

20. A. Makris, I. Kontopoulos, and K. Tserpes, "COVID-19 detection from chest X-Ray images using Deep Learning and Convolutional Neural Networks." doi: 10.1101/2020.05.22.20110817.

21. Narin, A., Kaya, C., & Pamuk, Z. (2020). Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. *arXiv preprint arXiv:2003.10849*.

22. Cohen J.P., Morrison P., & Dao L. (2020). COVID-19 image data collection, arXiv:2003.11597, https://github.com/ieee8023/covid-chestxray-dataset (accessed on 25 August 2020).

23. Paul Mooney. (2018). Chest X-Ray Images (Pneumonia) https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia (accessed on 25 August 2020).

24. U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," *Multimed. Tools Appl.*, vol. 77, no. 8, pp. 9909–9924, Apr. 2018, doi: 10.1007/s11042-018-5714-1.

25. P. Ribalta Lorenzo *et al.*, "Segmenting brain tumors from FLAIR MRI using fully convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 176, pp. 135–148, Jul. 2019, doi: 10.1016/j.cmpb.2019.05.006.

26. O. Russakovsky, J. Deng, H. Su, and J. Krause, "Imagenet large scale visual recognition challenge," *International journal of*, 2015, [Online]. Available: https://link.springer.com/article/10.1007/s11263-015-0816-y.

27. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015.

28. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

29. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.

30. C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," *31st AAAI Conf. Artif. Intell. AAAI 2017*, Feb. 2016, [Online]. Available: http://arxiv.org/abs/1602.07261.

31. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.

32. H. A. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2009, doi: 10.1016/S1507-1367(10)60022-3.

33. F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 1800–1807, doi: 10.1109/CVPR.2017.195.

34. E. Bisong, "TensorFlow 2.0 and Keras," *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. pp. 347–399, 2019, doi: 10.1007/978-1-4842-4470-8_30.

35. T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, p. 103792, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103792.

36. S. Asif and Y. Wenhui, "Automatic Detection of COVID-19 Using X-ray Images with Deep Convolutional Neural Networks and Machine Learning," 2020. doi: 10.1101/2020.05.01.20088211.

37. I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/s13246-020-00865-4.

38. A. I. Khan, J. L. Shah, and M. M. Bhat, "CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images," *Comput. Methods Programs Biomed.*, vol. 196, p. 105581, Nov. 2020, doi: 10.1016/j.cmpb.2020.105581.

39. Wang, L., & Wong, A. (2020). COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest Radiography Images. *arXiv preprint arXiv:2003.09871*.